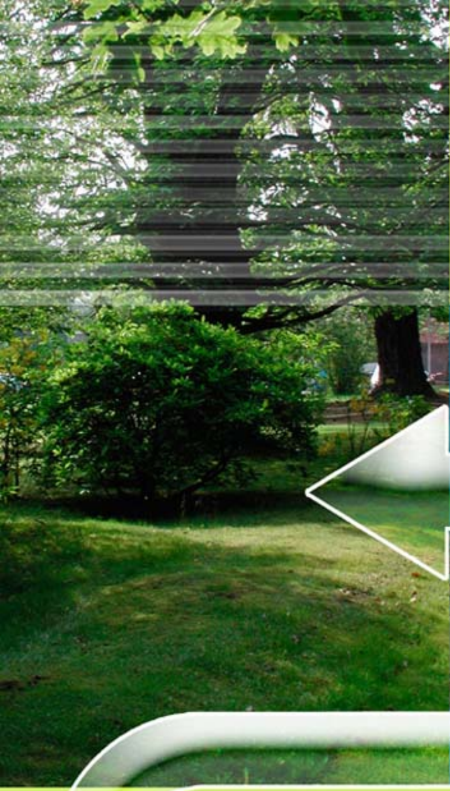


Editors:
C. A. Brebbia,
M. E. Conti &
E. Tiezzi

Management of

Natural Resources, Sustainable Development and Ecological Hazards

WIT PRESS

Management of Natural Resources, Sustainable Development and Ecological Hazards

WIT*PRESS*

WIT Press publishes leading books in Science and Technology.
Visit our website for the current list of titles.
www.witpress.com

WIT*eLibrary*

Home of the Transactions of the Wessex Institute.
Papers presented at Ravage of the Planet 2006 are archived in the WIT elibrary
in volume 99 of WIT Transactions on Ecology and the Environment (ISSN 1743-3541).
The WIT electronic-library provides the international scientific community with
immediate and permanent access to individual papers presented at WIT conferences.
http://library.witpress.com.

First International Conference on
the Management of Natural Resources,
Sustainable Development and Ecological Hazards

The Ravage of the Planet

Honorary President

A. Perez Esquivel
Nobel Prize for Peace

Conference Chairmen

C. A. Brebbia
Wessex Institute of Technology, UK

M. E. Conti
University of Rome 'La Sapienza', Italy

E. Tiezzi
University of Siena, Italy

International Scientific Advisory Committee

D. Almorza Gomar
R. Barber
C. Borrego
F. Botre
L. Campanella
G. Cecchetti
C. Cecchi
A. H-D. Cheng
M. da Conceicao Cunha
W. P. De Wilde
S. E. Jørgensen
D. Kaliampakos
K. L. Katsifarakis
J. Kretzschmar
J. W. S. Longhurst
Ü. Mander
T. Martonen
F. C. B. Mascarenhas
M. Maugeri
G. Passerini
F. Patania
V. Popov
M. Portapila
P. F. Ricci
C. Rossi
G. Sciutto
B. Sunden
W. Timmermans
C. Trozzi
E. Uspuras
A. van Timmeren

Organised by
Wessex Institute of Technology, UK
University of Siena, Italy

Sponsored by
WIT Transactions on Ecology and the Environment

WIT Transactions on Ecology and the Environment

Transactions Editor

Carlos Brebbia
Wessex Institute of Technology
Ashurst Lodge, Ashurst
Southampton SO40 7AA, UK
Email: carlos@wessex.ac.uk

Editorial Board

This page intentionally left blank

Management of Natural Resources, Sustainable Development and Ecological Hazards

Editors

C. A. Brebbia
Wessex Institute of Technology, UK

M. E. Conti
University of Rome 'La Sapienza', Italy

E. Tiezzi
University of Siena, Italy

WIT*PRESS* **Southampton, Boston**

C. A. Brebbia
Wessex Institute of Technology, UK

M. E. Conti
University of Rome 'La Sapienza', Italy

E. Tiezzi
University of Siena, Italy

Published by

WIT Press
Ashurst Lodge, Ashurst, Southampton, SO40 7AA, UK
Tel: 44 (0) 238 029 3223; Fax: 44 (0) 238 029 2853
E-Mail: witpress@witpress.com
http://www.witpress.com

For USA, Canada and Mexico

Computational Mechanics Inc
25 Bridge Street, Billerica, MA 01821, USA
Tel: 978 667 5841; Fax: 978 667 7582
E-Mail: infousa@witpress.com
http://www.witpress.com

British Library Cataloguing-in-Publication Data

A Catalogue record for this book is available
from the British Library

ISBN: 978-1-84564-048-4
ISSN: 1746-448X (print)
ISSN: 1743-3541 (online)

The texts of the papers in this volume were set individually by the authors or under their supervision. Only minor corrections to the text may have been carried out by the publisher.

Printed in Great Britain by Athenaeum Press Ltd., Gateshead.

Preface

This book contains the papers presented at the International Conference on the Management of Natural Resources, Sustainable Development and Ecological Hazards, which was held in Patagonia, Argentina. The Conference name was shortened to 'The Ravage of the Planet' to emphasize the urgency of the problems under discussion.

The state of our planet continues to deteriorate at an alarming rate. We have arrived at a situation where we need to determine urgent solutions before we reach a point of irreversible damage.

Our civilisation has fallen into a self-destructive process whereby natural resources are consumed at an ever-increasing rate. This process has now spread across the planet in search of further sources of energy and materials. The aggressiveness of this quest is such that the future of our planet hangs in the balance. The problem is compounded by the pernicious effect of the resulting pollution.

With these concerns in mind, the Co-Editors of this book decided to call an international conference organised by the Wessex Institute of Technology with the sponsorship of their respective institutions. The concerns were discussed with Dr. Adolfo Perez-Esquivel, Nobel Peace Prize winner, whose courageous defence of human rights extends to the need to stop the violence against our environment. A peaceful and just society is only possible if it is able to achieve a degree of harmony with its environment. The destruction of natural resources and the pollution of the environment will result in increasingly violent confrontations and an unstable world. These arrangements led Perez-Esquivel to support the meeting by accepting an invitation to become its Honorary President.

Much has been written about reaching sustainability in different contexts, but the concept itself needs to be defined within the framework of many different disciplines in order to arrive at optimal solutions. Therefore this conference is essentially a trans-disciplinary meeting in order to find appropriate sustainable solutions, i.e. those involving collaboration across a wide range of disciplines.

Patagonia was selected as the location for the Meeting as it is one of the few unspoilt parts of the planet, now under threat from over-development and exploitation of its many resources. Bringing the region to the attention of the international scientific and learned community may help to highlight new ways in which regions

such as Patagonia can be saved for future generations. The success of the Conference, which is demonstrated by the number and quality of the papers in this book, is a clear demonstration of the importance that our colleagues attach to the environmental and ecological problems threatening our planet. The organisers were pleasantly surprised by the response.

The editors are indebted to many people for the success of the Conference and the quality of the book. Firstly, our gratitude goes to all participants who travelled great distances to Bariloche in order to attend the conference and present the papers contained in this book. Many others helped to make the Conference a reality. Our greatest debt is to Dr. Perez-Esquivel who, by agreeing to be Honorary President, lends credibility to a meeting, which, in the early stages, was difficult to conceive and organise.

We hope that the Conference has fulfilled its main objective, which was to take stock of the present situation and help to formulate constructive proposals and policies for a way forward.

The Editors
Bariloche, Argentina
December 2006

Contents

Section 7: Water resources

Section 7.1: Management of complex systems under extreme conditions (organised by L. Fagherrazzi)

Section 8: Air

Section 9: Soil

Section 10: Energy

Section 1
The re-encounter

This page intentionally left blank

Holistic sustainability: local culture and global business—a unique opportunity

M. B. Neace
Mercer University, USA

Abstract

This paper discusses the role of global enterprises and local culture as primary inputs to attain long-term sustainable development. Several models are discussed showing that today's linear business systems are overly anthropocentric, disconnective, and monolithic in their approach to sustainable development. A business philosophy inclusive of local socio-culture stakeholders respecting their diversity and creativeness as vital constituents to the attainment of sustainable development is proposed. The discussion parallels the UN Global Compact. Broadening business planning to include local stakeholders, incorporating the precautionary principle, and knowledge sharing in a holistic manner with local stakeholders can lead to sustainable communities. The models presented demonstrate this holistic business stratagem.

1 Introduction

Today, the linear economic reductionist *modus operandi* is one that extracts resources from Earth, converts them into products and services, and returns the residue (wastes) back to the biosphere (Figure 1). Figure 1 also suggests how this linear system can be modified toward a sustainable paradigm for business and local communities by also including the diversity, creativeness, and wisdom of local culture. Until recently, little thought was given to the impact of the linear-reductionist processes, particularly from a holistic perspective, on Planet Earth. All sectors of the global society have and are contributing to this ever-increasing problem. There is now substantial evidence that Planet Earth is not capable of sustaining the continued pace of technological-economic activity it has experienced over the past five decades, and we are losing the battle [30].

This paper offers a holistic schema with several models that address the environmental issue from a business perspective. I contend that adoption of a

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060011

holistic business philosophy has the potential to move the global economy toward a sustainable future using guidelines of the UN Global Compact. I critique the linear reductionist models that most business organizations utilize. Then I discuss a series of models with emphasis on the holistic, interconnectedness character of the biosphere (including Planet Earth and the cosmos) of which local socio-cultural stakeholders are significant contributors to attain holistic sustainable development as integral elements of business strategy.

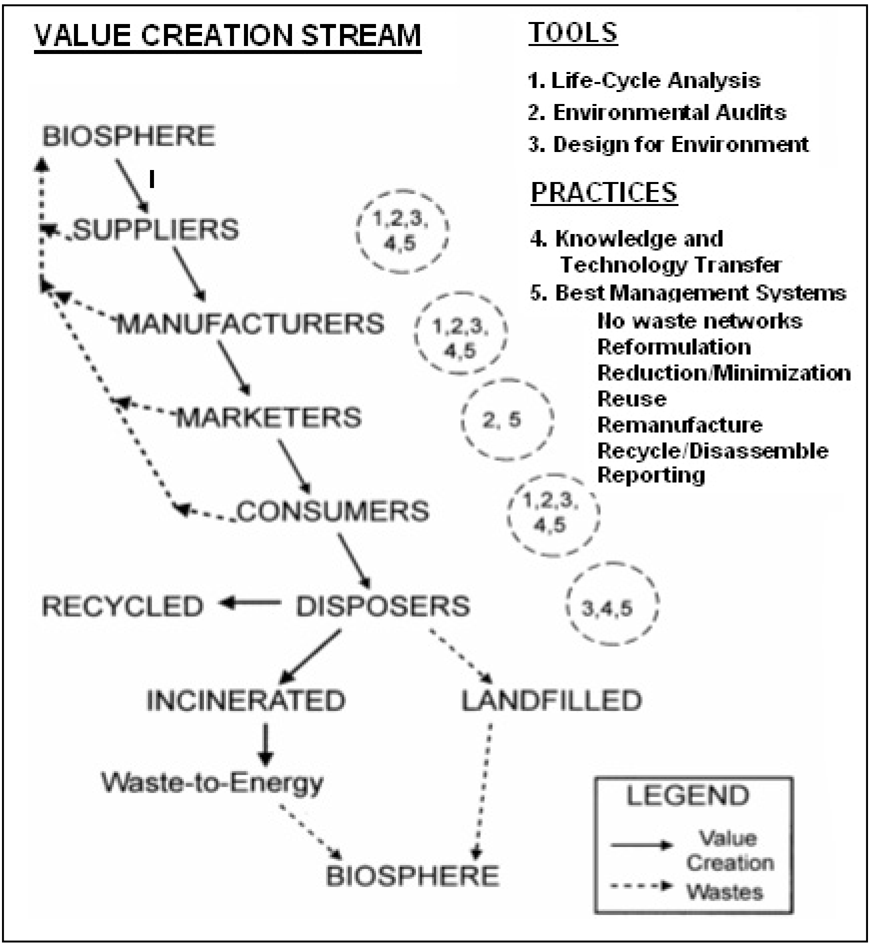


Figure 1: Linear reductionist system converted to an industrial ecology network.

2 A very brief history of man's view of the world

The linear-reductionist models that most of the economic-scientific world employs can trace their origins to the civilizations that flourished around the

Mediterranean basin from 500 B.C. up to the fall of the Roman Empire—Hellenism, Judaism, and Christianity. Support and development of this linear-reductionist model continued anew after Europe emerged from the Middle Ages. The Renaissance (14^{th}–17^{th} centuries) followed by the Enlightenment (18^{th} century) were periods of rich human development. Scientists, philosophers, and economists (e.g., Bacon, Locke, Descartes, Newton) flowered during this period, providing underpinnings for the industrial revolution of Europe [29]. Adam Smith showed clearly that production could be increased through specialization—a reductionist approach to manual labor. (See Figure 1 for an overview reductionist model with suggested modifications for sustainability.)

In reductionist models, man views the natural environment as a free resource to be exploited and used for his well-being. Man's control over the natural environment has risen to the level that the natural order of evolutionary processes is ever closer to being anthropogenic rather than biocentric. Continued application of linear models on par with the accelerated pace of the past five decades is unsustainable [3,28] (Figure 2).

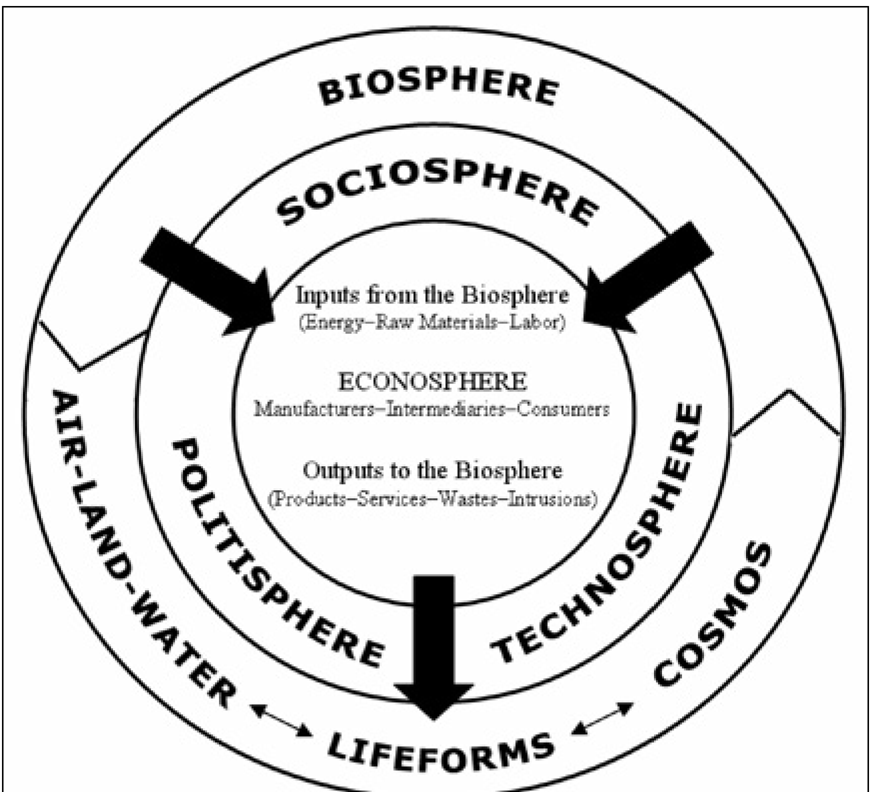


Figure 2: The human-biosphere interface.

3 Critique of the anthropocentric linear model

Development of the linear model has not gone unchallenged. Recently, a number of writers have raised serious questions about the limitations surrounding the application of classical Western scientific/economic traditions [6,8,9,12,15,22,24,25], but probably none more significant or impressive than Rachel Carson's *Silent spring* [5]. Her writings introduced ecology and its concept of interconnectedness of living things in nature to the public and spurred a heated debate and a spate of legislative activity in the Western World to protect the natural environment.

There are several reasons why reversing the rate of biospheric degradation has been elusive: (1) few national governments have had the political will to enforce environmental laws or make the necessary changes in their economic/social policies and structures; (2) the disconnectedness of production and consumption from the underpinning biospheric base; and (3) the lack of awareness and knowledge of the dynamic relationships between humans and Planet Earth [2,22]. What is lacking is a paradigm, both theoretical and operational, that reunites man to his biosphere [21].

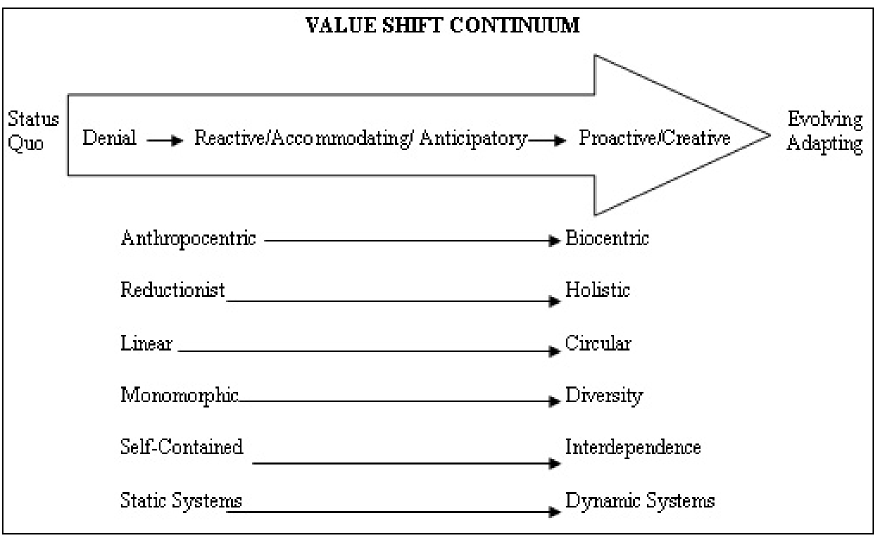


Figure 3: Transformation of values.

4 The transformation process

Response to environmental crises over the past three decades has evolved through a transformation process consisting of several stages: Beginning with denial of environmental problems and refuting the "questionable" scientific data base in the late 60s, to transforming values to reactive/anticipatory, and, now, to proactive/creative endeavours [21] (Figure 3). Sustainable development bridges

the gap between ecology and economics. Today, there is increasing support in the global business community for the UN Global Compact to close this gap (www.unglobalcompact.org).

Figure 3 illustrates the value transformation necessary to reach new levels of maturity, understanding, and potential for progress for all participants. Most writers on this subject agree that as transformation takes place and as business learns more about the workings of Planet Earth and its biosphere, its technologies and practices will be less damaging to the environment [18].

For the business community, this means de-compartmentalizing functions, practices, and supply chains and listening to fellow stakeholders outside the establishment [1]. Norgaard [23] describes these relationships as linkages, not a synthesis of unlike paradigms, but symbioses, whereby each unit is enriched because of its differences. No business is being asked to forego its efforts at satisfying consumers, shareholders, and other stakeholders. It will take a staggering amount of good business practices and cooperation to move in the direction of sustainable development—a holistic schema. That can only be attained by use of an inclusionary philosophy—all relevant stakeholders. Figure 4 expresses this integrated holistic concept.

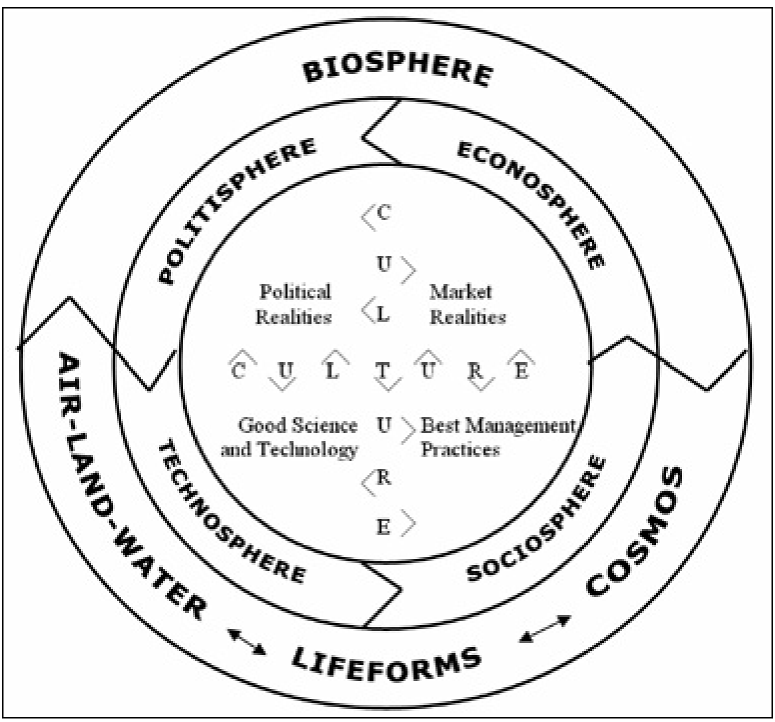


Figure 4: Integrated holistic paradigm.

Value shifting to the proactive/creative level leads naturally to synergistic activities, blending economics and ecology through combinations of good science, best management practices, and the inclusion of local stakeholders [30–32]. The global business community plays a key role in this process: they have the knowledge, they have the skills, they understand risk (the precautionary principle), and they have the ability to create win-win-win situations—a triple bottom line. Other stakeholders at the local level play pivotal roles in the development and application of sustainable initiatives, including research, education, and formulation of regulatory policies and transforming these preparations into good science and best management practices.

Building on Norgaard's [23] plea for pluralism, it is essential for the global business community to recognize and respect that different groups and cultures have unique approaches to their view of the world, applying different assumptions and methodologies that result in different models of the "same" phenomena. It is this very diversity and the synergy it creates that holds promise for progress toward a planetary system of sustainable development.

Those of us in the scholarly-intellectual world, because of our professional ethos, should foster creative opportunities for cross-functional, cross-organizational programs (teaching, research, management, policy formulation) within and among our institutions and with other stakeholders. Only in this way can we truly develop the holistic, interconnected, dynamic paradigms necessary for movement toward sustainable development [36].

5 Making sustainability operational

Limitations of present environmental policies and programs are well known, including bureaucratic ineptness and lack of will [14], short-term economic-political "remedies" at the sacrifice of long-term biosphere reality [13,27,32], and application of reductionist paradigms when increasing evidence points to a holistic dynamic general system [4,9,16,20,29,33–35]. Overcoming these issues is what the UN Global Compact is all about.

A holistic perspective is required. Norgaard [23], Rothschild [29], and Teilhard [35], among others, have developed holistic models that largely focus on "horizontal" dynamics and the integration of global enterprises, national economies, and political policies. These efforts have led to significant progress in our understanding of Earth. Also needed is a vertical dynamic that encourages discussion of concepts, policies, and supply chains throughout all levels of the global community—dynamic networking.

Twenty-five years of "command and control" policies targeted at the most obvious and egregious environmental problems are complex and, now, only marginally productive [22]. Continued progress in alleviating biospheric problems due to man's intrusions and consumption is proving to be difficult, and assuring a sustainable future requires global business leadership. Holistic and inter-vertical new initiatives, such as the UN Global Compact, are "forcing" the creative strategic plans of global enterprises to the local level; e.g., community pollution prevention programs, local government partnering, and business compliance assistance.

Many observers of the present environmental dilemma believe a significant number of undesirable outcomes of a monomorphic, elitist bureaucracy could have been avoided with local socio-cultural input and gradual withdrawal of central control [10,11,14,19,26]. No one is suggesting that government abandon its responsibility of serving its peoples. But a growing number of business leaders, professionals, and scholars are suggesting a gradual withdrawal of central control. In most well-educated communities with rising levels of human consciousness and awareness of the connectedness of all life forms [35], central control just doesn’t work; for example, the collapse of the Russian empire. Randhir and Lee [26] see the role of central government as an external observer, supplier of technical know-how, nurturer of mutual trust among the variety of stakeholders, and assistant to the development of well-structured incentive systems.

I. Vertical Interconnectedness

A. Responsibility to community/Individual and away from command and control

B. Interactive-networking – global enterprises, NGOs, local governments)

C. Results/Goal Oriented (Strategic Planning)

D. Series of local-based initiatives

Ecosystem Management

Place-based Management

Community-based Environmental Protection

Watershed Management

Cross-media Management

Figure 5: Making sustainability operational—missing links.

6 Conclusion

To approach a life-style that is sustainable, the global business community will have to expand its linear-reductionist orientation to encompass a holistic view of man, Earth, and the cosmos, similar to the view of the UN Global Compact. Are the global enterprises of the world ready to utilize these resources in their economic and strategic planning? Are global enterprises ready to embrace fellow local stakeholders as integral parts?

Several respected scientific journalists claim that the hard data says the battle is being lost [7,17]. For sustainability to move toward becoming a reality, global enterprises, global NGOs, and local socio-culture communities should implement general systems as a matter of general course. The global business community has the opportunity as well as the responsibility of encouraging this inclusiveness for the betterment of all, for creating triple bottom lines.

II. Local Socio-Cultural Participation

A. Sustainability requires input of local stakeholders

B. Power/Incentives to local stakeholders

1. Farmers/Ranchers
2. Small manufacturers
3. Community leaders
4. Consumers
5. NGOs

C. Gradual withdrawal of central government

D. Freedom, Diversity, Tolerance

E. Risk/Strength/Creativity

F. More decision makers

G. Successful ideas increase

H. Self-interest/Creativity

Figure 6: Making sustainability operational—missing links.

References

[1] Benson, William H. (1995). "Better science makes for better decisions," *Environmental toxicology and chemistry*, (Vol. 14, No. 11), 1811–1812.

[2] Berry, Thomas (1988). *The dream of the earth*. San Francisco: Sierra Club Books.

[3] Boulding, Kenneth E. (1966). "The economics of the coming spaceship earth," in *Environmental quality in a growing economy*, Henry Jarrett, ed. Baltimore: Johns Hopkins Press, 3–14.

[4] Burrows, Brian C., Alan J. Mayne, and Paul Newberry (1991). *Into the 21st Century*. Twickenham, U.K.: Adamantine Press Limited.

[5] Carson, Rachael (1962). *Silent spring*. Boston: Houghton-Mifflin.

[6] Clark, Mary E. (1989). *Ariadne's thread*. New York: St. Martins Press.
[7] Connor, Steve (2005). "The final proof: Global warming is a man-made disaster," *The independent*.
[8] Costanza, Robert (1991). *Ecological economics: The science and management of sustainable development*. New York: Columbia University Press.
[9] Daly, Herman E. (1991). *Steady-state economics*, 2nd ed. Washington D.C.: Island Press.
[10] DeGraaf, H. J., C. J. M. Musters and W. J. TerKeurs (1996). "Sustainable development: Looking for new strategies," *Ecological economics*, (Vol. 16, No. 3), 205–216.
[11] Enama, Mugangu T. (1994). "Culture: The missing nexus in ecological economics perspectives," *Ecological economics*, (Vol. X, No. 2), 93–95.
[12] Hardin, Garrett (1968). "The tragedy of the commons," *Science*, (13 December), 1234–1248.
[13] Hawken, Paul (1993). *The ecology of commerce*. New York: Harper Collins Publishers.
[14] Hess, Karl, Jr. (1992). *Visions upon the land*. Washington, D.C.: Island Press.
[15] Laborit, Henri (1985). "The Complexity of Interdependence in Living Systems" in *The Science and Praxis of Complexity*. Tokyo: The United Nations University, 146–169.
[16] Lovelock, James E. (1979). *Gaia: A new look at life on earth*. New York: Oxford University Press.
[17] McCarthy, Michael (2005). "Needed: A new Manhattan project," *The independent*.
[18] Nash, Roderick F. (1989). *The rights of nature*. Madison, WI: University of Wisconsin Press.
[19] Neace, M. B. (2005). "Adding co-evolutionary principles to a holistic paradigm of sustainable development," *Ecosystems and sustainable development V*, eds. E. Tiezze, et al. Southampton, UK: WIT Press, 2005, 355-364.
[20] Neace, M. B. (2003). "The catalytic role of the science community: Aking sustainability work," in *Ecosystems and sustainable development*, eds. Tiezzi, Brebbia, and Uso. Southampton, UK: WIT Press, 603 - 611.
[21] Neace, M. B. (1998). "Sustainable development: A critique and proposal," in *Ecosystems and sustainable development,* eds., J. L. Uso, et al. Southampton, UK: WIT Press, 105-113.
[22] Norgaard, Richard B. (1994). *Progress betrayed: The end of progress and a coevolutionary revisioning of the future.* NY: Routledge.
[23] Norgaard, Richard B. (1985). "Environmental economics: An evolutionary and a pleas for pluralism," *Journal of environmental economics and management*, (December), 382–394.
[24] Odum, Howard T. (1971). *Environment, power and society*. New York: John Wiley and Sons.
[25] Prigogine, Llya (1985). "New perspectives on complexity," in *The science*

and praxis of complexity. Tokyo: The United Nations University, 107–118.
[26] Randhir, Timothy O. and John G. Lee (1996). "Managing local commons in developing economics: An institutional approach," *Ecological economics*, (Vol. 16, No. 1), 1–12.
[27] Rees, William (1990). *Sustainable development and the biosphere*. Chambersburg, PA: ANIMA Books.
[28] Repetto, Robert (ed.) (1985). *The global possible*. Washington, D. C.: World Resources Institute.
[29] Rothschild, Michael (1990). *Bionomics*. New York: Henry Holt and Company.
[30] Royston, Michael G. (1979). *Pollution prevention pays*. New York: Pergamon Press.
[31] Saunders, Tedd and Loretta McGovern (1994). *The bottom line of green is black: Strategies for creating profitable and environmentally sound business*. San Francisco: Harper.
[32] Schmidheiny, Stephan (1992). *Changing course: A global business perspective on development and the environment*. Cambridge, MA: MIT Press.
[33] Stikker, Allerd (1992). *The transformation factor*. Rockport, MA: Element, Inc.
[34] Tarnas, Richard (1991). *The passion of the western mind*. New York: Harmon Books.
[35] Teilhard de Chardin, Pierre (1959). *The phenomenon of man*. New York: Harper.
[36] Vedeld, Paul O. (1994). "The environment and interdisciplinary ecological and neoclassical economical approaches to the use of natural resources," *Ecological economics*, (Vol. 10, No.1), 1–13.

Beyond predictability

W. Timmermans, J. Jonkhof, V. Kuypers & O. Braadbaart
Alterra, Wageningen-UR, The Netherlands

Abstract

This essay deals with the interrelationship between climate change, the flood, the Brent Spar, one hundred years of discussion about sewerage systems, the weather, the Anazasi Indians, Pim Fortuyn (a Dutch right-wing politician assassinated on May 6, 2002), soccer fans, ecosystems, and innovation.

1 Unpredictability

History is a succession of quiet and turbulent times. Developments that have been emerging gradually over the years may suddenly and apparently without warning start accelerating, assuming a dramatic or revolutionary character. It is of vital importance for any society to be able to counteract these changes by alertness and preparedness.

The biblical Noah was no fool when he obeyed God's command and built an Ark. His actions seemed absurd: who would be so foolish as to build an ark in the middle of a desert, miles away from the coast? Noah, however, built his ark with hazardous times in mind. He prepared himself for an event that no person in his right mind would ever consider to happen. Unlike so many others, he was determined not to be taken by surprise. The dinosaurs did not give a toss about comets, did they? The world could not imagine that the assassination of Franz Ferdinand in 1914 would lead to the death of hundreds of thousands of soldiers in muddy trenches a couple of years later.

In order to be able to understand how an event in a remote place can cruelly disrupt our peaceful little lives, it is necessary to study unstable dynamic systems. These are systems whose long term state of quietness and equilibrium suddenly changes due to seemingly minor marginal changes and turbulently transform into a new and unexpected state of equilibrium.

First, however, I will dilate upon the issue of the predictability of the future. Is it possible for us to know everything, anticipate everything, and prepare

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060021

ourselves for everything? In the following sections, a series of unpredictable and dramatic revolutionary events will be discussed. The systems and their developments will be investigated by means of theories in the field of complex unstable dynamic systems. The results of the investigation will form the basis for a discussion about the existence or non-existence of a universal principle.

2 Is the future predictable?

Intoxicated by the rapid development of mathematics, people have lived for a long time under the impression that everything in nature could be defined with exactness. Ever since the seventeenth century, the conviction has prevailed that a given initial state of a system determines its entire future development. Science is based on this fundamental principle, and to some extent present-day science still is. Natural phenomena are carefully observed and theoretically explained. The explanations are tested experimentally, accepted, adjusted, or rejected.

Newton developed the mathematical laws describing the orbital paths of two heavenly bodies in relation to one another. There seems to be a seamless explanation for everything. Voltaire saw the world as a mechanism, a timepiece. However, the state of euphoria was shattered when Poincaré, in an attempt to solve the mathematical three-body problem, concluded that it was mathematically impossible to predict the motion of three heavenly bodies in relation to one another on the basis of the initial conditions. As this is an all too complex problem for us, we have to conclude that the future development of a random complex system ranks highly among the issues that cannot be reduced or solved. Let us illustrate the matter of unpredictability by means of a number of examples.

3 The Lorenz Butterfly

Mathematician and meteorologist Lorenz carried out computerized simulations to predict weather conditions. One day, he wanted to continue a run he had done earlier so he put in the end variables from the day before. The new simulation yielded radically different weather results. He discovered that internally the computer worked with six decimal places but in the presentation with three decimal places, which caused a rounding error in the start of the simulation of less than 0.001%. This marginal difference in the values of one of Lorenz' variables resulted in an entirely different weather forecast.

This experiment gained fame as the metaphor of Lorenz' butterfly. A butterfly flapping its wings in one place on the earth could cause a hurricane in another. A minimal deviation in the initial variables caused by an unforeseen perturbation leads in a short time to completely different weather conditions. Within this context, a campfire on the Suburban Ark could affect the weather in the whole of the Netherlands in, for instance, a period of two weeks. This seems more spectacular than it is. The flapping of the butterfly's wing does not essentially change the weather and the climate will remain as it is now. All over the world, small perturbations are taking place continuously: countries are

teeming with butterflies; campfires are burning everywhere, and in many places, buildings are erected influencing turbulence.

In short, weather conditions are continually affected by all kinds of natural and anthropogenic interventions. The majority of disruptions do not affect the weather structurally. The weather conditions in Western Europe, for instance, will continue being determined by the low-pressure areas from Iceland. The system seems stable, yet there are indications that dramatic climate changes in the past were indeed caused by sudden unusual circumstances.

4 Ecosystem dynamics

A sudden event can also radically change an ecosystem. A system that looks very stable on a superficial level may turn out to be unexpectedly vulnerable. A striking example of such a rapid shift is the Sahara desert. Between ten to five thousand years ago, the area was a wetter place than it is now, with lush vegetation, lakes, and marshes. The transition from this area into the current desert took place very suddenly approximately five thousand years ago. The Dutch ecologist Scheffer has identified similar changes in lakes, coral reefs, forests, deserts, and oceans (Scheffer et al, [10]).

Figure 1: Fractal (frost-flower).

The shift takes place as follows. The environmental conditions in a system that at a given moment is in (an state of equilibrium) can change gradually. Increased pressure from the outside causes the environmental conditions of the system to change very slowly. As, inside the system, a large number of biological, physical, or chemical feedback mechanisms are still in action, the

situation remains as it is, i.e. stable. However, the resilience of the system undergoes a gradual deterioration. At a given moment, an impulse from the outside can lead to a catastrophic change. Such an impulse, which could be anything from a sudden drought to a flood, acts like the proverbial last straw. A catastrophic change always takes place unexpectedly and cannot be predicted with the current *early warning systems*. Restoring the fundamental necessary conditions of the system at issue will not reverse the change. In order to reach this, far greater efforts are necessary.

A small and simple system will change relatively rapidly. A large and particularly a complex system will offer more and longer-lasting resistance. Even so, the change will be sudden and unpredictable and the degree of irreversibility many times higher. Moreover, a system such as a lake seems to have two or more stable equilibrium states, which enable it to jump from one state to another. According to Scheffer, the changes are unpredictable because they are not linear in character.

5 Collapse of civilizations

In his book *Guns, Germs, and Steel: The Fates of Human Societies* [1], Jared Diamond lists a large number of crises that have led to the near-collapse of entire civilizations.

After Christopher Columbus had discovered America, Spanish armies comprising of no more than a few dozen soldiers managed to defeat tens of thousands of Indians. The deciding factor here was the fact that the Spaniards had horses, guns, and steel swords, objects with which the Indians were entirely unfamiliar. In the centuries following the 'discovery of America' and the first conquests, diseases that they lacked immunity to more than decimated many peoples of the new world.

Thus, a marginal development, i.e. the arrival of a small number of white people, led to a massive crisis in the existence of the Indians. In his book *Collapse: How Societies Choose to Fail or Succeed* [1], Jared Diamond answers the question why societies can collapse.

In his vision, the collapse of entire societies can be attributed to the combination and interrelation of factors such as environmental damage through human interventions, climate change, hostile neighbours, enemies of befriended societies, and the way in which a society's elite responses to its problems.

Janssen et al [4] have done research into the civilization of the Anazasi Indians. In the '*collapse literature*', they came across three factors that may lead to the collapse of a civilization: a) sudden negative events, b) the impressive size of the collapsed settlements at issue, and c) evidence of overexploitation. The researchers added a fourth element, based on psychosocial expertise: d) large investments in a certain area or in a certain way of living make it virtually impossible for a population to abandon the area or adopt another lifestyle in less favorable conditions. If the situation gets too unstable in terms of surroundings, circumstances or environmental conditions, a sudden event, a drought or a war, can lead to the collapse or fall of the civilization at issue.

6 Complexity theory

Since the seventies of the last century, science has adopted the view that the world is composed of systems sharing a number of qualities. These systems often seem to be in a seemingly stable situation. This stable equilibrium state is called an 'attractor'. The processes of development within a system continue in a linear way for a long time, remaining within the stability of the given attractor. Certain developments at the periphery of the system can lead to instability in the system, causing it to move into the sphere of influence of another attractor. This change can be extreme and often pass off chaotically: it is unpredictable what the new stable equilibrium state will be. Generally, the addition of extra energy or matter to a stable system will eventually lead to a phase jump.

Following Geldof [2] and others, the following characteristics play a role in the description of an unstable dynamic or complex system:

- Attractors. Complex systems are often experienced as stable. They usually are in a preference state. A lake, for instance, has an ecological equilibrium state. VINEX housing estates are generally developed according to largely the same procedures.
- Crisis. A crisis causes a complex system to jump from one system to another: the metaphorical flapping of the wings of Lorenz' butterfly or a sudden and unexpected drought can be the last straw.
- The border area between order and chaos (*the edge of chaos*). Intellectual, societal, and/or biotic processes in a system have changed to the extent that the existing equilibrium state is under pressure, but is not yet undergoing changes. The system is 'far from equilibrium'. In this state, a crisis, or a marginal event in the periphery of the system is likely to arise, forcing the system to move into the sphere of influence of another attractor.

Figure 2: World EXPO Hannover.

Gleick describes the gradual evolution of new insights in the work of scientists engaged in various fields of specialization such as philosophy, chemistry, ecology, communication science, management sciences, and archaeology (Prigogine en Stengers, [8]; Gleick, [3]; Lewin, [5]; Waldrop, [12], Janssen et al, [4]; Scheffer et al, [10]).

7 Back to the first question

What is the interrelationship between the flood, the Brent Spar, one hundred years of discussion about sewerage systems, the weather, the Anazasi Indians, Pim Fortuyn, soccer fans, eco-systems, and innovation?

A flood happens suddenly, unexpectedly, like a raid. In 1995, Shell Oil decided that from an environmental point of view, dumping the Brent Spar, a heavily contaminated oil installation, into the North Atlantic was the best solution. This decision met with well-founded global resistance, which Shell initially considered irrelevant and marginal. The installation of sewerage systems in the Netherlands has led to a heated debate about usefulness and necessity resulting in a plethora of systems. In the end, one specific system was applied in the greater part of the Netherlands. The Dutch 'purple' cabinet (a coalition of the labor party (PvdA), right-wing liberals (VVD), and liberal democrats (D66), which ruled from 1994 to 2002, thought to have matters under control when out of the blue former sociology professor and right-wing politician Pim Fortuyn raised a number of politically sensitive issues. Soccer fans massively started the wave or indulged in worse activities without incitement or stage direction. These examples are all indicative of a system, be it a group of people or a planning process, suddenly changing from one into another equilibrium state under the influence of a marginal or remote phenomenon.

8 Makeability

This approach has far-reaching consequences for our western way of thinking, which is strongly based on the concept of makeability. In his book *The End of Certainty* [6], Prigogine introduces the physics of non-equilibrium processes, a science that affects our way of thinking in two ways. Firstly, it proposes the fundamentally constructive role of 'the arrow of time', which expresses Prigogine's vision of the irreversibility of time. Even with limitless knowledge, it is fundamentally impossible to reverse time, which means that any process is irreversible.

Secondly, it proposes the principle of uncertainty. Traditional science relates limitless knowledge to certainty; however, limitless knowledge of a system or of its initial conditions does not imply that its development can be predicted. Prigogine claims that the laws of nature do not express certainties but possibilities. This vision is of major importance to systems innovations, which after all do not arise from thinking in terms of linearity and cognizance or the invention of a new system. System innovations come into being by developing the right idea, or new thing, at the right time in the right context. To attain that, one has to allow oneself to be lazy from time to time...

References

[1] Diamond, J., *Collapse,* North Point Press, 2004.
[2] Geldof, G.D., *Omgaan met complexiteit bij integraal waterbeheer,* TAUW, Deventer 1999.

[3] Gleick, J., *Chaos: making a new science,* Viking, Penguin 1987.
[4] Janssen, M.A., T.A. Kohler, M. Scheffer, 'Sunk-cost effects and vulnerability to collapse in ancient societies', in: *Current Anthropology* (2003), Vol 44, no 5.
[5] Lewin, R., *Complexity: life at the edge of chaos,* Macmilan, New York 1992.
[6] Prigogine, I., *Het einde van de zekerheden,* Lannoo, Tielt 1996.
[7] Prigogine, I., I. Stengers, *Order out of chaos,* Bantam Books, New York 1993.
[8] Prigogine, I. & I. Stengers, 1984. Order out of Chaos. Bantam books. New York.
[9] Scheffer, M., S. Carpenter, J.A. Foley, C. Folke, B. Walker, 'Catastrophic shifts in ecosystems', *Nature 413* (2001), pp. 591-596.
[10] Scheffer, M., F. Westley & W. Brock, 2003: Slow response of societies to new problems: Causes and costs. Ecosystems 6. pp 493 - 502.
[11] Waldrop, M.M., *Complexity, the emerging science at the edge of order and chaos,* Viking Books, London 1983.
[12] Waldrop, M.M., 1992: Complexity, the emerging science at the edge of order and chaos. Viking Books. London.

This page intentionally left blank

Sustainable global civilization

Y. Ishiguro
Instituto de Estudos Avançados, CTA, Brazil

Abstract

The fundamental cause of almost all problems facing humanity is the size of the human population, multiplied by the intense consumption. There will not be technological solutions unless the population and the consumption are reduced, because apparently the capacity of the planet has already been surpassed.

In order to enable a sustainable global civilization, some basic concepts regarding human existence need to be recognized or changed:
* The Earth is finite in space, resources and capacities of its systems;
* Man is adapted to the current equilibrium in the systems of the Earth;
* A healthy biosphere is essential for the equilibrium in the habitats for man;
* The biosphere is maintained by photosynthesis in plants;
* There is a limit to the size of the biosphere;
* An increase of human population means a reduction of those of other species;
* Perturbations in the biosphere as well as in other systems could break the equilibrium;
* In a closed system the effective fertility of each species must be that of replacement;
* There cannot be human rights that lead to the collapse of the community;
* The concept of community needs to include all humanity and the biosphere;
* Ultimately man has to live with renewable resources;
* The principles of economy need to be modified;
* The meaning of life must be found in something other than procreation or consumption;
* All ideals cannot be realized;
* Some new rationality needs to be added to the rules of community;
* The Earth can support indefinitely an abundant lifestyle in a limited global civilization;
* Human population must be reduced.

Keywords: sustainability, limits of the Earth, population, consumption, fertility control, irrational idealism, rational society, philosophical bases, capitalism.

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060031

1 Introduction

In the past centuries there have been remarkable progresses in various areas such as science, technology, medicine and agriculture and today some countries enjoy the highest standards of living ever achieved in history, with high levels of education, health care, long life, comfort, leisure and material consumption. The human population grew to a level never achieved by any animal species of a comparable size. Homo sapiens is the most successful species in the history of the evolution of life.

However, various difficulties for the continuity of human community have appeared in the past decades. Individuals, organizations and governments have been trying to resolve the current and perceived future problems and continue efforts to "save the Earth". In addition to environmental problems such as global warming, species extinction, deforestation and accumulation of wastes, there are diverse social problems, too, such as poverty, avoidable diseases and famine that affect a large part of the world population. Despite the well-known prosperity in industrialized countries, miseries and sufferings in less developed countries are increasing and social polarization is intensifying in most countries. International conflicts continue to occur and a considerable fraction of every nation's economy is directed to military spending [1].

The mode of living currently practiced in industrialized societies and aspired to by others cannot be continued for long nor extended to the whole world [2]. Competitions for limited resources and space could lead to social unrests, increased international conflicts and the destruction of the systems of the Earth. Moral restraints could be thrown away and a large part of the world population could be eliminated. However, there are technologies to assure the survival of a small group of elites. Homo sapiens will not be extinct. A new civilization will emerge, perhaps wiser and more rational after the trauma, but such a transition must be avoided.

The fundamental cause of almost all problems facing humanity is the size of the human population, multiplied by the intense consumption in affluent parts of the world. Man is adapted to the current equilibrium in the systems of the planet but he has abused the very systems he depends on. China is the only country that recognizes it and has adopted concrete measures. If the current trends continue, an environmental and/or social collapse in a global scale will be inevitable in near future and the whole world will have to adopt the one-child policy mostly criticized at present, or more restrictive policies, or face the collapse of the world civilization. But the world community continues to cling to impracticable ideals and declares with a consensus that a free choice of the number and spacing of children is a human right [3]. Humanity needs to recognize the reality of the world and be more rational in selecting ideals and rules of community. There cannot be human rights that lead to the destruction of the community.

Control of human population and fertility needs to be discussed openly and the current economic and political systems that emphasize continuous expansion of consumption need to be reviewed in global long-term perspectives. The current taboo needs to be broken and the myopia corrected.

A vibrant biosphere has been maintained for hundreds of millions of years. I believe that the Earth can support indefinitely a healthy, comfortable, abundant lifestyle that most people aspire to, if science and technology are applied wisely and if the population is controlled at a level compatible with the capacity of the Earth. But to achieve it, human mentality needs to change. A modification of human mentality will be difficult and it makes one pessimistic. But if we look at some current societies, we can be cautiously optimistic. I think that there still is time for the world community to change its ways and ensure its continuity. Man has to learn to take a global view and adapt to the inherent limitations of the planet. I propose a new ideal for humanity, that of building a sustainable rational global civilization that is in equilibrium with nature and in which all societies continue in peace, natural resources are shared by all, and a decent living is guaranteed for all who conform to the rules of community.

In order to realize a sustainable global civilization some basic concepts regarding human existence need to be recognized or changed:

- The Earth is finite in space, resources and capacities of its systems;
- Man is adapted to the current equilibrium in the systems of the Earth;
- A healthy biosphere is essential for the equilibrium in the habitats for man;
- The biosphere is maintained by photosynthesis in plants;
- There is a limit to the size of the biosphere;
- An increase of human population means a reduction of those of other species;
- Perturbations in the biosphere as well as in other systems could break the equilibrium;
- In a closed system the effective fertility of each species must be that of replacement;
- There cannot be human rights that lead to the destruction of the community;
- The concept of community needs to include all humanity and the biosphere;
- Ultimately man has to live with renewable resources;
- The principles of economy need to be modified;
- The meaning of life must be found in something other than procreation or consumption;
- All ideals cannot be realized;
- Some new rationality needs to be added to the rules of community;
- The Earth can support indefinitely an abundant lifestyle in a limited global civilization;
- Human population must be reduced.

2 Sustainable global civilization

The principle of evolution of life is the survival of the fittest. The evolution of human societies also followed the same principle. Many civilizations emerged, prospered and disappeared, some for natural or internal causes but many destroyed by others. Today, in a century of science, technology and globalization, international conflicts continue in various regions of the world.

Within societies, too, there are conflicts between social classes, ethnic groups, religious sects and groups of people with specific beliefs, ideals or interests.

The industrialized nations have achieved the highest standard of living in history but it was only possible at high costs to the global ecosystem, the biosphere and the rest of the human community. These nations continue taking advantage of the labor and natural resources of less developed countries and abusing the environment and the biosphere by their intense consumption, backed by strong economic and military forces.

The basic question for humanity now is whether it follows the path of natural selection or it builds a civilization equitable to all, including future generations and the biosphere. History and the current situation point to the first path, which would necessarily lead to the elimination of a large part of the world population.

To take the second path, some fundamental changes are necessary. Societies have chosen many ideals as their guides, some of which are impracticable and wishful but most are noble in themselves. These ideals cannot all be realized in a finite world and some compromises need to be made, including adoption of some rules that are contrary to the most basic principles of humanity. Fundamentally, basic limitations of nature need to be recognized, first of all the finiteness of the planet and, secondly, the principles of the biosphere, such as the workings of food chains, the low rates of survival of progenies and the finiteness of individual life, though these principles should be adapted to human society in order to enable a sustainable rational world community.

Recent advances in science and technology, and the enormous possibilities that they demonstrate, show that man can construct a sustainable global civilization and maintain a healthy biosphere if he adopts rational rules of community and controls his population and consumption. The collective human mentality needs to evolve and the only way to do so is education, formal as well as social. A small part of the huge annual military expenditures of the world could perhaps help finance universal education, creating much better prospects for peace than the escalating buildup of weapons. Humanity collectively needs to reflect on its value system and select the most essential ideals. One of the basic requirements is a refinement and extension of the concept of community, namely, the recognition of the mutual dependence including the biosphere and the elevation of community above individuals.

3 State of the global environment

Many problems have been recognized in the past decades and are well known to people concerned with the present and the future of the Earth. Principal difficulties include the following types.

- Exhaustion of accumulated resources, such as fossil fuels, soils and minerals;
- Use of renewable resources above the productive capacity, such as fresh water, forest and fishery;
- Demand on the systems of the Earth above their capacities, such as processing of refuses of human life and air pollution;

- Accumulation of man-made materials damaging to the biosphere, such as industrial wastes and toxic chemicals;
- Changes in the surface of the Earth, such as pavement, dams, deforestation and desertification;
- Perturbations of equilibrium in the global systems, such as CO_2, CFC, ozone and nitrogen cycle;
- Perturbations in the biosphere, such as species extinction and introduction of non-native species.

4 Philosophical bases

In addition to the technological progresses there have also been advances in philosophical and moral aspects that reduced abuses such as colonialism and slavery. However, humanity's philosophical bases continue mostly unchanged since the earliest civilization. Notable characteristics include unrealistic or impracticable idealism, shortsightedness, egoism, greed and brutality of individuals, social groups and nations.

It appears to me that current efforts to save the Earth suppose that technological solutions are possible without modifications of philosophical bases, rules of community, ideals and traditional customs, such as free procreation, growing population, unlimited accumulation of wealth, intense consumption of natural resources, abuse of ecosystems, interminable judicial processes, and infinite value nominally given to human life and that the effects of human activities on the biosphere, ecosystems and future generations are ignored. The current social and international rules allow a life of intense and ostensive consumption by some individuals and nations while billions of people live in misery and millions die every year of avoidable causes. Some ideals are taken to extremes by the interests of some sectors of society causing diversion of resources. Social and international conflicts are bound to occur in protest, threatening the core of social stability and the continuity of the world civilization.

The two fundamental causes are, first, the current philosophical bases, arrangements and workings of societies and, second, the finiteness of the Earth. The second factor cannot be changed in any significant ways. So the solutions must come from the first factor. There will be no technological solutions because apparently the limits of the systems of the Earth have already been surpassed.

5 Economy and politics

Capitalism is the only economic system that has passed the test of history, despite its well-known shortcomings, but it is one of the causes of the current situation of the world.

In most countries one of the primary policies is economic development, namely constantly increasing consumption, in order to assure employments for the growing number of people seeking means of survival. Houses, automobiles,

appliances and all sorts of consumer items are produced in increasing quantities and their sale is promoted by all conceivable means. Increasing demands on transport and other infrastructures mean constantly expanding constructions. Growing population squeezes the job market and further expansion of economy, further consumption, is promoted. Natural resources are extracted at the highest possible rates and refuses accumulate. Wealth is concentrated in the hands of a small number of people, huge tracts of lands are fenced off, small-scale farmers are shoed off their lands and together with jobless people seek subsistence farming in small plots of land cleared from virgin forests. The biosphere shrinks, resources are exhausted, environment deteriorates, and the quality of human life sinks.

The luxurious lifestyles in the industrialized countries are broadcast to the rest of the world where most people struggle to survive from day to day. These people would naturally feel envy, resentment and anger and their frustrations are bound to explode. The series of terrorist acts are inevitable consequences. I can only try to imagine the pains of starvation or the despair and hatred of a suicide bomber. The concentration of wealth, the buildup of military power and the polarization of technological capabilities enhance the disparity between the two worlds and further intensify the cycle of protests and suppression.

Humanity has experimented with various political systems and apparently selected democracy as the best. However, in the current world, it has an essential shortcoming in that the elected representatives are practically inhibited to have long-term visions. For most politicians the future is the next election and they would attend to the immediate wishes of the people they represent. Attempts to limit consumption, economic development or human rights would be a political suicide for them.

Some modifications of political, as well as economic, systems are necessary but at present decisions are in the hands of the people and nations with interests in the status quo and environmentalists and other visionaries who advocate changes have no real power. I cannot propose any concrete alternative systems except intensive education and the adoption of more rational rules of community. I hope that education can remedy myopia and other primitive aspects of human nature so that the common people would have long-term global visions and their elective power and demands for change would lead to more rational systems.

6 Human population and the limits of the Earth

The human population was a few hundred million at the beginning of the Christian era and increased slowly over the next 1800 years to less than 900 million. The rate of growth accelerated since then and the population reached six billion in 1999. The world population is increasing by more than 70 million per year and is projected to reach nine billion in the middle of the century.

Ecological footprint analyses show that the capacity of the Earth is adequate to sustain a population of a billion at the current rate of consumption in the United States and two billion at the rate in European countries.

7 Laws of nature and rational society

As the society guarantees a secure life for everyone, individuals need to assume their part of responsibilities and the definitions of human rights, individual freedom, crime and punishment need to be modified. In many societies ideals and interests of some individuals or sectors led to the definition as crime of some activities that are really a matter of personal choice without prejudice to others, resulting in much waste of resources that could be used for the benefit of the community. The most notable example is the current policy on narcotic drugs. Some people apparently choose to use them knowing that the consequence is much suffering and a shortened life. Liberalization of narcotic drugs would eliminate clandestine traffic and drug wars. The current prohibition is actually creating demands and violent organizations. The so-called drug war will never be won, just as the alcohol prohibition failed. As long as there is a demand, there will be supplies.

Most criminals are produced by the society, namely, many people are driven to activities defined as crime in order to survive in unjust societies that allow some individuals to live a luxurious life alongside slums and shantytowns. Once the stable equitable society is achieved there will be no need to commit crimes for survival. However, there will always be some individuals who commit acts of violence and deception. These people, a very small minority, need to be removed from the society. The concepts of human rights and justice need to be modified, with community placed above individuals. Communities and nations have, throughout history, resorted to war to guarantee their survival and peace, killing innocent people and destroying material bases of civilization. It would be much more rational to apply this principle internally in order to guarantee the continuity and peace of each community.

Abortion and euthanasia also need to be viewed more rationally. Humanity needs to recognize that life is finite, that life depends on the termination of other lives and that unlimited procreation, survival of all progenies and individual life beyond biological limits are impracticable ideals. The most fundamental ideal should be the continuity and peace of the community. In nature the rates of survival of progenies are quite small: it is the way the biosphere functions. Man alone cannot be an exception, unless he controls himself in some way. Human rights should include the freedom of individuals to choose how to live and die.

In most societies, judicial processes continue interminably in the name of human rights and justice, diverting resources that can be used for better ends. In the world where millions die of starvation and avoidable diseases annually, this is an example of irrational idealism.

Other examples of individual choice that many societies worry about or even criminalize include smoking, drinking, obesity, seat belt and motorcycle helmet.

In summary the sustainable global civilization should assure, on one hand, universal education, conditions for a decent life and freedom to choose how to live and, on the other hand, individual responsibility to ensure its continuity and equality of all who conform to rational rules of community.

8 Looking toward the future

Human history is full of conflicts since the first known civilization to this day. The future appears to be the continuation of this tradition, with larger scales and higher efficiencies in the killing and destruction. There are means to destroy large cities and kill millions of people in a matter of seconds or minutes. A large part of the world population lives in precarious conditions. Social and international conflicts continue in various parts of the world and will intensify with increasing shortages of essential resources and further polarization.

Science and technology have been trying to maintain a balance between increasing demands, production capacities, dwindling resources and preservation of the environment. However, the population increased to the limit at each moment and prevented advances in the living conditions of the majority, with benefits reaching to only a third of the world population at present. The possibilities of science and technology appear limitless but the planet is finite. The equilibrium of the systems is delicate and can be broken by small perturbations. Stored resources are being exhausted, some renewable resources are being used beyond replacement levels, species are vanishing and the environment is deteriorating.

Man has behaved till now like a child who disregards its future. After a few thousand years of civilization man learned ways to use nature to his convenience and, like an adolescent euphoric of his newly found power, ignored his future and abused his own home. The human population has exceeded the carrying capacity of the planet. To ensure the continuity of the global community man has to change his ways and an essential measure is a reduction of the population and consumption. It is time for man to grow and plan for a long mature life. Figure 1 shows a summary of the current situation and the two alternative paths to the future.

The most basic principle is that in a closed system the fertility must be that of replacement. In nature this principle is maintained by the food chain and natural constraints such as available space, variations of climate and limits of photosynthesis. Man, however, has no predator. Historically other forces such as diseases and food shortages, as well as wars, limited the population but man has strived to eliminate these factors one by one, except war, and today many diseases have been practically eliminated or can be cured routinely and the food supply is abundant overall with adequate reserves for any emergencies. Advances in medicine and sanitation, as well as in other areas, have reduced infant mortality and increased the average life span in most countries. Positive results have been glorified and negative consequences and their causes have been ignored.

Human history has been formed by power plays by nations, organizations and individuals, as in wild animal societies. The end result is the current situation that will most likely lead to the collapse of the global ecosystem or another world war. Humanity needs to recognize the inherent limitations of the planet and adopt more rational rules and a sense of global community. In place of trying to accommodate the needs of an increasing population, humanity can apply the

advances in science and technology to improve the living conditions of a limited constant population. Principles of economy also need to change, with characteristics such as: employments distributed to all in a basically constant and routine economy based on renewable resources, minimum utilization of stored resources, constructions limited mostly to maintenance, improvement and replacement. Some popular activities that involve wasteful consumption in the name of sports, leisure or economic development will have to be stopped. Life would be rather dull for some people, compared to that in current industrialized societies. But the only alternative, that I can see, is a global conflict with all its consequences.

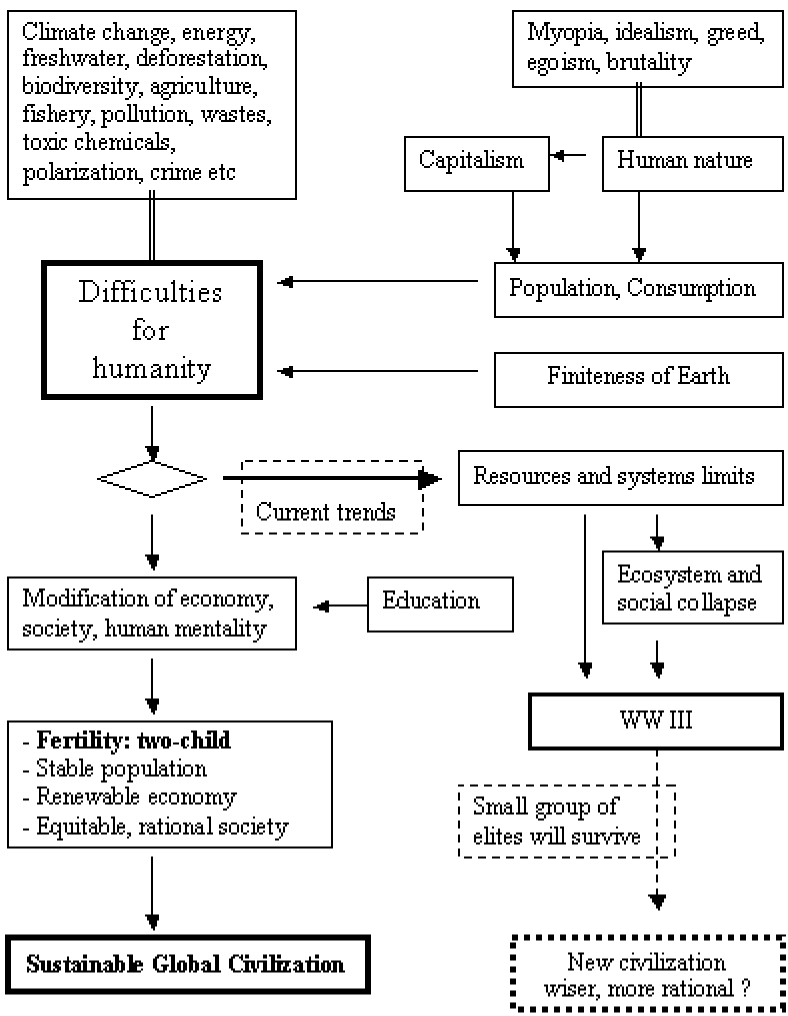


Figure 1: Summary of the current situation and the alternative paths to the future.

The relation of man with nature needs to be based on the recognition of his dependence on the equilibrium in the intricate systems of the Earth and in the biosphere. Man's position must be found within the limitations of the Earth. It is time to abandon the posit that a territorial conquest, a growing population and an expanding economy mean an advance of human existence. Control of the population is not a solution in itself but a necessary condition and the first step.

Man is the culmination of the evolution of life on this unique and beautiful planet and his intelligence is the most remarkable phenomenon other than life itself. With his logical mind, accumulated scientific knowledge and technological abilities man can construct a sustainable rational global civilization in harmony with nature. If he so chooses, man can assure the continuity of the biosphere, his own intellectual evolution and understanding of the universe.

Acknowledgments

The author is grateful to the Instituto de Estudos Avançados of Comando-Geral de Tecnologia Aeroespacial (CTA) for the permission to write a book that forms the basis of this essay. The author applauds the visions of the directors to allow studies of a subject beyond the normal activities of the institution. The opinions and views expressed are entirely of the author and do not necessarily reflect those of the institution or any person thereof.

References

[1] Ishiguro, Y, *Civilização Global Sustentável*, Booklink, Rio de Janeiro, 2005.
[2] Earth Council, Ecological footprint of nations, http://www.ecouncil.ac.cr/rio/focus/report/english/footprint/.
[3] Program of action of the United Nations International Conference on Population and Development, Cairo, 1994, Section 7.3, http://www.iisd.ca/Cairo/program/p00000.html.

Section 2
Political and social issues

This page intentionally left blank

Autonomy and heteronomy: the need for decentralization in a centralizing world

A. van Timmeren[1,2]
[1]*Climate Design & Environment (CD&E), Delft University of Technology (TUD), Faculty of Architecture, The Netherlands*
[2]*Atelier 2T, Haarlem, The Netherlands*

Abstract

There is growing heteronomy of the essential utilities, particularly energy and sanitation. The methods and techniques applied in present-day essential infrastructures for energy and sanitation supply may be considered traditional and driven by the separation and centralization paradigm. There is physical expansion, due to globalization combined with the liberalization of the energy market and – to a smaller extent – the solid-waste market. Other characteristics of development are specialization and segmentarization, with one or several dominant parties per sub flow or sector as major results. Convergence of the various technical infrastructures is a new characteristic. It results in greater complexity and more dependence on the structures as perceived by users.

For sectors that are left to market forces, positive effects are soon to be expected on the efficient use of the (infra)structures by oligopolistic market types, and, thus, on the affordability of the accompanying services. However, market participants have no interest in overcapacity, which puts pressure on the reliability of supply (by a maximum bid on the available capacity). Pressure can also be put on the other long-term interests, including maintenance of grids and investments in, research into or application of innovations, e.g. those that aim at sustainable development.

Main aspects for users are sustainability, a guarantee on supply and processing and affordability. Reliability and affordability gain more importance. At this moment, this is still at the expense of sustainability. There will be a (well-known) dilemma between the short term (economic efficiency) and the long term (sustainability and guarantee of supply). Where the essential infrastructures are concerned, the liberalization of the markets shows that the goals set cannot always be accomplished in an integral way. At a national level, there is (still) too little grip on the developments. The demand for supervision or rules at a supra-national level is being heard, and this causes one of the reasons for liberalization to be surpassed.

Keywords: heteronomy, infrastructures, autonomy, integration strategy.

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060041

1 Introduction

There is a common consensus in society about the necessity of fundamental facilities for meeting the most fundamental needs in the own living environment, viz. "Maintenance", the so-called primary necessity of life. The availability of energy and food, including clean drinking water, and the removal of waste are parts of it. It is no use trying to introduce sustainability measures that harm this fundamental need.

Many relevant participants however do not seem to realize that other, more sustainable alternatives can be found by abandoning the specific characteristics of the traditional paradigms rather than following them. The dominant participants have an interest in using existing structures as efficiently as possible and in developing them further with as few risky investments as possible. As yet, the cost of transport (distances) in most of these 'essential flows' (energy and sanitation, i.e. drinking water and waste) is not taken into account, and there is little product differentiation, while these aspects in particular offer possibilities for alternatives that support sustainable development.

Looked on from the aim of "sustainable development", the path of expansion selected (centralization) is not necessarily the optimum as perceived subjectively. A characteristic of expansion is the increasing importance of relocating the material and energy flows. In this, physical infrastructures play an important part. They are the bases for the supply of processes, products and services that meet the fundamental needs. Building infrastructure almost always implies slow and large-scale processes in the "underground" layer. For a structural solution and preservation, the technical infrastructure should be considered, as the lowest layer within a model of layers. It will be leading for the design and the allocation of the faster dynamics of the overlying layers: the layer of the "networks" and that of "occupation". It has turned out that the ongoing processes of liberalization have put pressure on the importance of the certainty of supply, and sometimes also removal. Working certainty of supply and independence out in further detail seems necessary, or even essential, not only for further development based on the future of scaling-up and heteronomy of different networks and users ("economies of scale"), but also for decentralization ("scale economy") [1].

2 Local autonomy or global heteronomy; a 'one town world'

2.1 Centralization and growing heteronomy

The increasing heteronomy observed in the so-called "essential" networks and accompanying managing parties for end users does not only hold for central networks, but also for decentralized or local systems. The difference lies in the consequences of misuse and catastrophes and the extent of visibility (the subjective perception) of the results of this heteronomy for the end user, and the possibilities for better involvement and understanding, and therefore sustainable improvements.

There is a subjective perception of the optimum scale of application for meeting fundamental needs. This is caused by different visions concerning the intended quality and quantity of the satisfaction of needs. It is of importance to analyse the differences and to optimize them to scale specific advantages.

The solutions as to the supply and processing of the essential sub flows and their infrastructures are the materialized reflections of necessary activities in order to be able to meet social targets and fundamental needs, in short: the 'suprastructure'. Dependence and (dominant) control play an important role in the relationship between infrastructure and suprastructure. The ongoing individualization more and more often leads to a pursuit of decreased independence on public infrastructures and the wish for decentralized utilities (connected or not), with autonomy of the individual or the household as an extreme version.

There are clear differences between the characteristics (or rather: advantages and disadvantages) of the various central networks, in the energy and sanitation sub flows each as well as between the energy and sanitation supply as a whole. They are caused by different "central scales" of application and different extents of visibility, but also by the management structure and the presence or absence of liberalization processes.

The infrastructure strongly correlates with production (supply as well as drainage). A change desired in the infrastructure, e.g. a bottleneck with respect to capacity, can be solved by investing in extending the infrastructure (now often accepted), but often also by adapting the "production" in strategic spots of the (central) grid. One possibility is connecting or disconnecting (decentralized, additional) sustainable sub production (generation or processing capacity). This may be overcome by including sustainability, via reliability, as an added value at relatively little cost, e.g. in the form of a decentralized (autonomous) utility and backup. Too little advantage is taken of this sub aspect of sustainability. For small-scale users, this results in a simpler arena, particularly where the number of parties is concerned with which contracts have to be signed. This may imply a gradual change of the paradigm, following a sliding time scale rather than a radical change at a certain, perhaps unexpected, point in time to come. Moreover, it may involve short-term interventions for long-term guarantees (sustainability, guarantees for supply or processing and in the end affordability). Such a principle may be useful as a kind of fall-back scenario for, for example, a serious and unforeseen dysfunction of the current process of further scaling up and liberalization of sectors.

In energy supply, there should be more emphasis on increasing the flexibility in the current (infra)structures, including Town and Country Planning in its entirety. The more so since it can be expected that there will not be only one decisive future technology to solve the coming problem(s) concerning sustainable development. At present, it could be maintained that for the infrastructures related to sanitation flows there is (in the developed world) already one applied (central) technology.

The "dialectics of progress" and the so-called "prisoner's dilemma" force themselves upon us: the deviation from this specific unsustainable (end-of-pipe)

type of solution(s) is so expensive and will involve such far-reaching social consequences that there seems to be no other choice than continuing with these (expensive) infrastructures and systems.

The distance created between the (environmental) problem and its solution leads to more and more complexity. The process of changing the interrelated public and private services, systems and infrastructures is becoming more and more complicated and less and less predictable. Together with the increased scaling, the convergence of utilities and the growing number of parties and techniques involved have increased the end users' (consumers') subjective dependence (heteronomy). This asks for a simplification of the processes, products (or rather: services) and parties involved. A larger concentration on integral provision of services, or, in other words, the supply and management of integral packages, offers possibilities. This seems to be reinforced by the ongoing liberalization processes. Another solution is having the level of application attune better to the lifestyle and direct surroundings of the users. Decentralized or local systems do respond to that demand.

2.2 Decentralization

Science, and increasingly the market too, bring up a rising number of solutions that imply possible smaller scales of implementation. The considered benefits are a possible reduction of infrastructure and better visibility and tuning into the demand and therefore more flexibility. Especially in the field of small scale Combined Heat Power generation and ecological sanitation systems important efforts have been made. The latter, so-called DESAR (Decentralised Sanitation and Reuse) systems, offer an alternative for the current status quo.

The idea behind these kind of smaller systems is their relative simplicity and adaptability, and therefore their possibility to create extra (sustainable) capacities in situations where:

- centralised systems have not been built yet,
- existing systems have reached the limits of their capacity and new buildings, districts and/or higher densities are planned; e.g. use as a (temporary) back-up,
- bio-climatical, geological or circumstantial characteristics make interventions (e.g. in the subsoil) difficult and/or expensive, and
- in case of desired improved sustainability or environmental performances e.g. through interconnections with other 'infra' systems.

There are still few examples of living and working environments with solely integrated systems concerning decentralised sanitation, energy and reuse. However in several developed and developing countries more and more examples are realised or close to completion.

A decentralized system must not be characterized as a static system, since there is an ongoing change of an existing situation. The scale level of a decentralized system is relatively fixed. It depends on the technique of the

administrative body itself, the context and the position of the observer. Technical (de)centralization concerns (a change of/in) systems. In the case of administrative decentralization, there is a distinction according to the nature of the administrative bodies: territorial decentralization (between/carried out by Government, Province and Municipality) and functional decentralization (e.g. within the Municipality).

As for technical decentralization, the various flows have different definitions of (the scale of) sub clusters and of "decentralized" sub networks and subsystems. Often, there is vagueness even within the various flows. The scale level is considered decentralized, but is defined in a relative manner too often.

Regarding technical decentralization, this study starts from the production and processing of the various flows closer to the users than is usually done, with the flows being fed back to the users in a direct way. Because of the relative new market of (technical) decentralization, "niches" can be created. It is possible that the creation of niches can also take place in a planned way. This is called "strategic niche management". The difference with the more familiar principle of "pilot projects" is that a shelter is built around the new technology in the case of strategic niche management, through which the technology can develop from prototype to an actually applicable technology. Eventually, the technology should work without any protective measures at all. It is of importance to find a strategy and a method that support such a process without relying on central authorities too heavily. With the aid of strategic niche management, innovations are implemented strategically in this type of "sheltered area", tested and evaluated.

The present-day competitive advantage of "sunk costs" for conventional (centralized) solutions should be avoided. Strategic niche management can be of help here. The strategic approach should focus on the higher dynamic efficiency of the decentralized systems: changed circumstances are easier to be anticipated with the help of decentralized systems. Investment risks may increase in this way, which is of more importance in the liberalizing markets. Nevertheless, this is also of importance in non-liberalized markets concerning sanitation flows. However, the advantages of privatization, as mentioned earlier, are reaped fastest in decentralized energy systems.

The use of new, sustainable technology and sub flows leads to larger quantitative fluctuations in supply, the peak load as compared to the average consumption (especially of the energy flows) and to the introduction of various (parallel) qualities (particularly of the sanitation flows), or, in other words, to a differentiation of products and services within the various technical (infra)structures. At the same time almost all decentralized sustainable energy sources have a low energy density, which, together with their variable character, will contribute to the obvious choice for a decentralized implementation. In the case of energy generation out of waste (water) flows, this particularly holds for systems based on natural techniques. For decentralized solutions to energy generation as well as decentralized sanitation systems, this leads to more use of space. This disadvantage is the reason why decentralized systems should be integrated with other architectural and/or natural facilities and functions as much

as possible. As a consequence, so-called “integrated systems” are preferred to “autonomous systems”. An important advantage is that the three-step approach (Reduce, Reuse, Recycle) is optimally facilitated by separation according to quality close to the source. Thus, the main demand of sustainability is met: a consistent quality of the flow (waste production and/or energy supply). In this way, cascading and high-quality recycling is easier to be accomplished in the waste and waste water flows. And in the energy flow, the same holds for the application of exergy.

Generally speaking, the two main problems in decentralized solutions are scepticism of the leading (often dominant) parties involved and the larger influence of a fluctuating flow size. The former is particularly caused by responsibility (certainty) and liability. This scepticism will increase because of the necessary transition of the market(s) from supply of products to supply of services.

The aspect of the flow size (in fact, the basis for the technical “economies of scale”) can be met locally by modern techniques of planning and tuning, the so-called “Real Time Control”, and the subdivision into parallel facilities. Thus, the remaining main points of interest for improving the competitiveness of decentralized systems and actually achieving the advantages for the environment and the users are the organization and implementation of maintenance, exploitation, provision of services and inspection of the various systems, together with the availability of backup provisions if necessary.

2.3 Local autonomy and autarky

This research takes a limited and so-called ecological interpretation of autarkic systems as a starting point: ‘systems that are closed for matter and energy, except for the continuous flow of solar energy’. Within this scope, the concept of autonomy is largely used as a synonym of autarky. However, autonomy cannot be considered a substitute, for autonomous concepts in the industries of environmental technology and building particularly deal with an autarkic ambition to sub aspects. The decentralization and, in some cases, even complete disconnection of central (infra)structures are at the centre of the developing emancipation of systems of which they are a part. In these cases, autonomous or possibly even autarkic systems emerge, that may be referred to as “local”.

A world consisting of autarkic “cells”, states or units produces a dilemma from a social point of view. This is an unattractive affair for many people, since methods of so-called “soft power” for reaching certain qualitative aims become less effective. They become more problematic if autonomy or autarky is implemented via (very) unsustainable techniques and concepts, at the cost of higher scaling levels. Thus, autonomy or autarky need not be identical to sustainability. Even more than with the present, conventional systems and structures, with autarky the scale of the energy and sanitation facilities always depends on the mutuality of the suprastructure (including society) and the infrastructure.

Worldwide, there are only a few genuinely autarkic concepts to be found. Most concepts striving for or claiming autarky mainly concern private initiatives for detached houses and buildings [2]. This is not only because they are often idealistic projects and designers (or "visionaries"), but particularly because of a lack of facilities, support and time for organized deviation from the usual ways of developing and building homes.

From a historical point of view, autarkic projects are quite possible at the level of living communities. Recent practical examples even prove that this is quite possible at various other levels nowadays, also thanks to several (existing) techniques. However, there have to be strong community spirit and – preferably – an egalitarian social structure. It is an important characteristic of autarky, in ecology as well as in economics, that it largely depends on the extent of its own inflexibility. In principle, a system based on independent, autarky-based cells, is more capable of absorbing change, but, at the same time, best flourishes in an environment of (economic) stability and slowly changing technology. The extensive integration of semi-autarkic decentralized systems within larger systems particularly offers a larger (market) control along with advantages resulting from the principle of the "economies of scale". The main rationale of looking for smaller scale levels of autonomy or even autarky is the bigger possibility of creating degrees of freedom in sub areas or at a smaller level, without having to adjust existing developments and investments in the larger structures. In such a system, changing conditions in environment or use of techniques may lead to being stuck to a technology or structure once chosen, and thus possibly to a restriction in the ability to adjust in time.

Connected small-scale "pseudo"-autonomous or autarkic entities, will be able to absorb the continuous transformations better, on account of their non-isolated character. In these situations, interdependence and heteronomy disassociate with the romantic ideal of individual autonomy, or even autarky, and form a better basis for spatially sustainable developments. However, the consequences depend on the extent to which the principle of the "economies of scale" has been applied. At the same time a structure of unconnected autarky is naturally unstable. A system based on a geographically clustered network of nodes that aim at autonomy offers possibilities for timely anticipation of changes that originate from technique, society or market conditions [3].

Two development processes concerning decentralized technology for the purpose of autonomy have come forward as topical: viz. first, the efficiency and improvements in the integration of sub techniques and co-ordinated, connected concepts, and, second, a better harmony between supply (input) and demand of the (different) sub flows. Additionally, there are two more general underlying development processes. The first is the environment-technical, environmental and, to some degree, also social optimization of decentralized systems within semi-autonomous projects. In spite of the potential of the underlying optimization principle of the "scale economy" claimed in much of the literature and projects, and in spite of its importance, which was also proven, it has only been applied to a small extent. Consequently, there still are not many "economies of scale" in this area. However, the sub aspects concerning the

application freedom and environmental integration (smaller sizes, fewer secondary demands, etc.) and user-related demands (comfort, ease of use, costs, etc.) do improve noticeably.

The second underlying development process concerns the link to economic applications related to the surroundings, often determined by soil or users, including taking nutrients back to agriculture and other lateral applications or possibilities, such as car-sharing systems. In addition to the possibility of other types of use of (agricultural) grounds (urban agriculture), the link to agriculture may not only lead to a structurally different infrastructure (aboveground and underground), but also to different country planning as a whole, when applied on a larger scale. It is an important link in the desired transformation of our society from one based on linear attitudes to resources and wastes towards a circular one, is a different way of handling sustainable energy and (ecological) sanitation. It is a ‘closed-loop approach’, in which for instance excreta are returned to the soil rather than to water. This implies a (better) formalization of the existing ‘leakage flows’ between the cultural world and the natural world, or ‘ecological field’ as a self-sustaining environment. Sound ecological sanitation, based on a closed cycle of nutrients as resources for food production, is always central to this [4].

Usually this implies the incorporation of, or the connection to, (types of) sustainable agriculture.

The added value is summarized below:

- It connects the closing of the water cycles and sustainable energy generation to the essential cycle of nutrients in environmental planning [5];
- it links up ecological solutions to economic developments;
- it initiates solutions for the increasing problem of (urban sprawl) urbanization and (agricultural) monocultures;
- it offers instruments for the connection of urbanization to greenbelt development;
- it contributes to more efficient sanitation systems based on separated waste flows and the accompanying energy generation and nutrients recycling: waste and pollution can be avoided and sustainable energy is relatively cheap and (locally) abundant [5].

There are also various disadvantages and potential problems connected to these decentralized systems of interconnected solutions of waste management, energy generation and nutrients recycling. As yet, there are no or few “economies of scale” in the production of components and the management of the technical units, due to the restricted number of (pilot) projects. The main problem is that sanitation is largely a social phenomenon, rather than a technical one. Additionally, systems based on natural technologies and natural processes are particularly vulnerable to incorrect use or sabotage, although possible negative effects will restrict themselves to a relatively small area or a small number of users, because of the decentralized character. Another problem is how reuse of nutrients from black water can be accomplished safely in practice.

3 Discussion

With respect to both extremes, globalization (heteronomy by interconnection) and striving to complete (ecological and/or economic) autarky cannot be seen as an optimal development for the suprastructure, or, in other words, a good, democratic basis for societies. And what is more, neither of them (in their specific pure form) is to be considered a good basis for further, sustainable development of the structures for those societies.

For the essential (technical) infrastructure, the dynamics of non-simultaneous, slow transformation necessary for attuning the complex structures of society, the essential "flows" and nature (or natural processes) implies that it is wrong to still think in separate systems within integral development processes. That is, since there is an increasing interconnection and interdependence in the technical infrastructure of the essential flows. Because of the fundamental need of protection of maintenance, autonomous or semi-autarkic projects should be able to meet such changes, either by means of a connection to a "backup" system (often on higher scale levels), or by means of parallel solutions (hence over dimensioning) within the system itself. In practice, we see far-reaching semi-autarkic projects being connected to central infrastructures. To be able to connect, to a larger extent than approximately 30% (electricity network) of the network capacity, projects (subsystems) based on autonomy and/or renewable (discontinuous) sources new network philosophies (or network geometry) of these centralized grids should be introduced. For the sanitation infrastructures alternative use of existing networks offers possibilities to cope with increasing costs due to aging and shortages on capacity. A changed network philosophy has far-reaching consequences for the way in which these infrastructures are designed and integrated [3]. For complex systems, the coherence with which and the way in which dynamic processes are dealt with determines the translation to physical "integralness". It is important to establish that the stability or resilience of networks is directly related to the their complexity. It is not the components of the various structures that matter, but the way they are organized together as intelligent structures. It is important to learn from the organization structure and topology of existing adaptive, complex structures. Recognizing the structures of each network is needed for combining their optimally ongoing development, possible decline and damage done to them, whether desired or not, with constant or increasing sustainability and certainty guarantees for users [6, 3].

Innovation, like the application and fitting in of new decentralized techniques and/or alternative network structures, does not suffice for the accomplishment of "sustainable development". Too often there is tension between the mechanisms and the institutions that regulate motivation on behalf of individual or joint wishes. In following the conventional centralization paradigm, this type of "ritualism" stands in the way of a development into a society with more opportunities for changes according to the principle of "conformity". It creates niches of "sustainable development" of all alternatives that do not comply with the centralization paradigm. This occurs in the shape of concepts that can be placed under "rebellion" and even "separation". Examples are to be found in

most of the Eco-villages and some Eco-districts, started by private – sometimes collective – initiatives and in some instances as individual projects or silent-green examples, as e.g. Ruigoord, near Amsterdam [7]. Although projects such as the Eco-villages are to be considered as the application typology of "conformity" according to Merton's definition [8], they are often placed under the application typology of "rebellion" or even "retreatism" by the dominant institutionalized authorities, looking at them from their own context on the basis of the current paradigm. Opportunities for a widely supported need for innovation are neglected here, and so is the chance of more significant "sustainable development", e.g. through scale invariance [3].

The problem of the directing centralization paradigm, which is even seen as imperative by some people, is often in the way of a more structural change. Nevertheless, the application typology of "rebellion", for example, which was started as a niche, can be taken as a method of allowing innovations to grow for the purpose of a more structural and large-scale use. In a way, Ruigoord is a spatial example of this, and the development of new technologies, like the 'Living Machine' – at first in Eco-villages – is a successful example of a developed innovation based on natural processes.

In current central infrastructures of energy as well as waste water flows, the possibilities of an alternative network layout are not or not sufficiently taken into account. More and more connections are made between the various (national) networks and sub networks in gas and electricity networks, but this occurs because of considerations of capacity and economic (business) perspectives, rather than on the basis of the principle of network geometry. Consequently, there is a direct interest for large-scale central networks to have subsystems as a decentralized cluster included into the complex network. Because of the principle of self-organization, it also offers the possibility and the guarantees for being able to make local decisions with respect to, for example, further-reaching sustainability without abandoning the principle of scale size ("economies of scale"). Procedurally, it implies that authorities and (public) grid managers may abandon policy aiming for a fixed ultimate goal. It eliminates the aspect of policy more and more lagging behind reality, which is common nowadays. Systems within decentralized planning concepts may lead to networks, complex or not, with a more strongly decentralized network structure with part of the networks performing relatively autonomously. These may support flexible and especially sustainable planning concepts in town and country planning. Moreover, the issue of a more precise attribution of (network) costs (use) and environmental effects to specific customers or transactions (which becomes more and more important as complexity decreases with ongoing liberalization) may be solved or may easier be solved.

References

[1] Timmeren, A. van. *Autonomie & Heteronomie. Integratie en verduurzaming van essentiële stromen in de gebouwde omgeving*. Delft University of Technology. Eburon Academical Publishers, Delft. The

Netherlands. 2006. (In Dutch only; English ed. expected at the beginning of 2007).
[2] Hasselaar, B.L.H., Graaf, P.A. de, Timmeren, A. van, Decentralised *sanitation within the built environment casu quo integrated in living environments*. The Architectural Annual 2005/2006. Delft University of Technology. Delft. The Netherlands. 2006.
[3] Timmeren, A. van, Kristinsson, J., Röling, L.C. *The interrelationship of sustainability and resilience- & vulnerability of networks, related to the critical flows in society; a future deadlock?* Proceedings International Conference Sustainable Building (SB05). Tokyo. Japan. 2005.
[4] Timmeren, A. van, Eble, J., Verhaagen, H., Kaptein, M. *The 'park of the 21st century': Agriculture in the city*. Proceedings of The Sustainable City III, Urban Regeneration and Sustainability. WIT press. Southampton. UK. 2004.
[5] Timmeren, A. van, Sidler, D. *Decentralised generation of energy and interconnection with treatment of waste and wastewater flows in an urban context*. Proceedings CISBAT CSFF Conference, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland. 2005.
[6] Timmeren, A. van *Heteronomy and (un)sustainability of essential technical infrastructures*. Proceedings 'Eco Architecture' Conference. Wessex Institute of Technology. WIT press. Southampton. UK. 2006.
[7] Timmeren, A. van; Seitz, V.; Heijligers, J.J.A.; Dorst, M. van *Case 'New-Ruigoord', Almere*. Proceedings 'Passive and Low Energy in Architecture' PLEA'03, Santiago de Chile. 2003.
[8] Merton, R. (1957) *Social Theory and Social Structure*. Glencoe, Free Press, in: Röling, N. *Gateway to the global garden; Beta/Gamma Science for Dealing with Ecological Rationality.* Eighth Annual Hopper Lecture, University of Guelph, Canada. 2000.

Towards a Sustainable World

H. H. Kleizen
IDEGO, Delft University of Technology, The Netherlands

Abstract

The world has a limited potential. Mankind has to balance its activities with those of the Geosphere (atmosphere, hydrosphere and lithosphere) and those of the Biosphere (from whales to archea). This balance leads to the concept of a Sustainable Technological World (STW) encompassing all artifacts (factories, houses, mobile phones) and their needs (air, water, soil, energy) and constraints (emissions). This STW is summarized giving estimates of its main parameters. A Sustainable World is more than the STW. In a Sustainable World people (present and future generations) have equal opportunities to take their share of the STW-artifacts. But societies (per continent, country, region, city, etc) have also different starting positions. First a novel method is presented to characterize the structure of the world population, revealing differences and similarities on the same level (country for example) or on different levels (continent and city for example). Secondly the Hofstede 5 cultural dimensions are analyzed and a new culture model is proposed based upon the sequence of 3 of the 5 dimensions. The world population has more than 6 cultures and for 88 countries the culture is given. Evidence is found that the population structure and culture model are linked. This opens a window to connect culture and the limited resources available to mankind.
Keywords: culture, population, society, STW, sustainability, technology.

1 Introduction

Not everybody likes to possess everything forever. According to Brundtland [1] sustainable development is development that meets the needs of the present without comprising the ability of future generations to meet their needs. One way to interpret this much quoted statement is that there have to be opportunities for those who populate the Earth nowadays and they should increase the opportunities for the generations to come. So indeed the very idea that some

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060051

entity claims all the artefacts on Earth is outrageous. But the question remains: how much is available for mankind and how do we safeguard the distribution of goods and services among the people of the Earth in time?

This paper will not solve this problem. It will contribute to a solution in three ways.

First it treats a Sustainable Technological World (STW), a concept of Kleizen [2,3], encompassing all material, energy, water and air that mankind has maximal available for technology and services. Fair distribution means simply that the amount available is divided by the 6.4 billion people living on Earth nowadays, as estimated in the CIA Factbook [4]. But large societies can create complex products and services, small societies have to keep these goods simple or compete by cooperation. A new model has been published recently by Kleizen [5,6] to break down the world population down in smaller societies (world, continents, countries, provinces, cities, neighbourhoods, streets etc).

How goods are appreciated depends on culture. To quote Hofstede 's favourite statement [7] of Pascal: " There are truths on this side of the Pyrenees that are falsehoods on the other". In his book Hofstede [7] recognizes 5 cultural dimensions. The third new element in this paper is to select only 3 of them giving rise to 6 world cultures as demonstrated by Kleizen [2]. So a new picture is made with fewer pixels but still supporting Pascal's statement.

The paper presents some evidence that the novel culture and structure model are linked [6] allowing one to address the distributional problem: A connection is made between the material dimension (artefact, air, water and energy), the size/complexity dimension and the cultural dimension.

2 Sustainable Technological World

The relevant parts of the STW concept as presented in [2,5] are summarized. Humans are part of the living matter on Earth together forming the biosphere. The biosphere is dispersed in the upper region of the Earth called the geosphere. The geosphere contains three subspheres: the atmosphere, hydrosphere and the lithosphere.

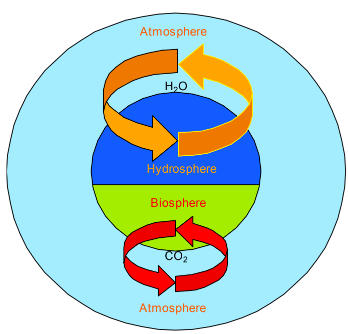


Figure 1: The 2 grant cycles on Earth.

The geosphere contains three subspheres: the atmosphere, hydrosphere and the lithosphere. Where the atmosphere meets a water surface mass exchange and life occurs: wet soil (the lithosphere) or seas (the hydrosphere). The two great cycles on Earth are depicted in Figure 1.

The greatest cycle is the mass exchange between hydrosphere and atmosphere. The 1.7 billion Gton (Lide and Frederikse [8]) heavy hydrosphere evaporates annually about 540 thousand Gton (Houghton [9], Grübler [10]) water and gets it back as precipitation from the 5.1 million Gton ((Lide and Frederikse [8]) weighing atmosphere. Meanwhile the biosphere, with an estimated mass of 5.7 thousand Gton (Smil [11]), inhales 300 Gton/yr (Smil [11], Dunn [12]) carbon dioxide from the atmosphere reduces it and sends it back after oxidation. The two cycles are coupled to energy exchange. The H-cycle between atmosphere and hydrosphere involves the vaporisation and condensation of water and the C-cycle between biosphere and atmosphere the carbon fixation in plants and carbon oxidation in plants and animals.

A third cycle involves the STW, the sustainable collection of technological artefacts on Earth. See Figure 2. The mass of the STW and its molar flux to and from the atmosphere is postulated making two assumptions about the position of the STW with respect to hydrosphere and biosphere. Oxygen is selected as the characteristic molecule shuttling between STW and atmosphere and water reduction is chosen to maximise the power associated with oxygen exchange. The major properties of the STW are collected in Table 1.

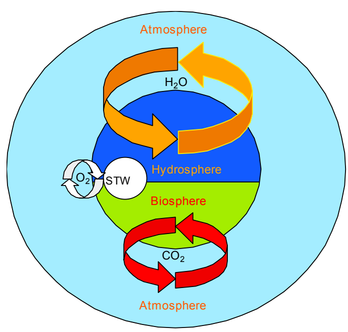


Figure 2: The STW the 3rd cycle.

Table 1: STW properties.

Basic properties		
Mass STW	110	Gton
Power STW	8800	GW
Life-span artifacts, average	25	yr
Benchmarks		
Annual production of artifacts	4.4	Gton/yr
Specific energy consumption producing artifacts	64	MJ/kg
Specific power consumption STW	80	mW/kg
Air and water flows		
Air	375	Gton/yr
Water	100	Gton/yr

3 Structuring societies

This model has been published in the same form elsewhere [5,6].

The population of a society can be classified using a logarithmic distribution like in filtration. The population is described by 2^N with N an integer and therefore a class width of 2. (In each class the number of people varies from $N/\sqrt{2}$ up to $N\sqrt{2}$.) In the CIA World Factbook [4] one counts 236 counties (nations, dependent areas and other entities) permanently populated by 6.446 billion people (July 2005 estimate). In Figure 3 the distribution over the World population over de 236 countries is sketched in terms of N.

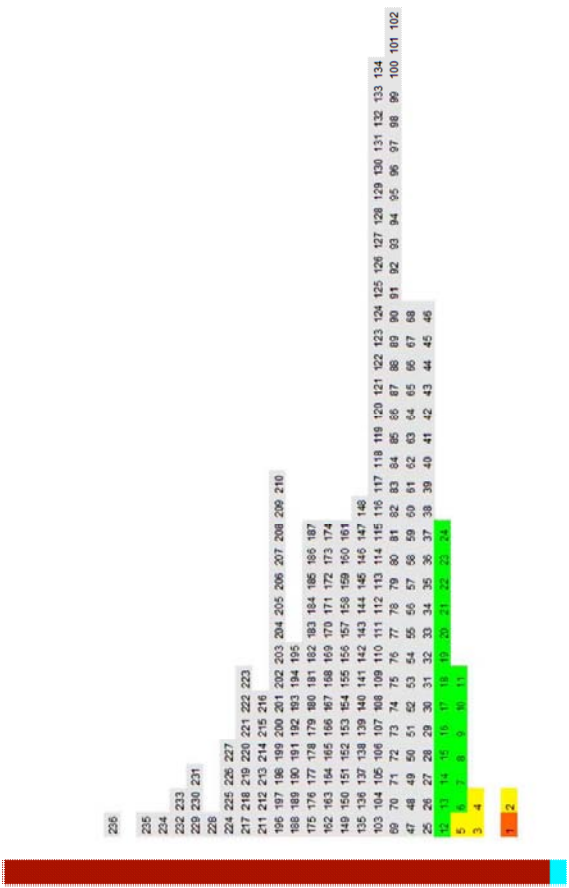


Figure 3: The World population distributed over 236 countries in terms of N.

The World population fells in class N=33. The largest countries China and India are in class N=30 and the smallest one – the Pitcairn Islands – with a population of 46 are in class N=6.

To describe the World top down a second variable is introduced: the order m of the distribution. The order m is equal to difference between the N-value of the society and the one of the largest subsociety. Thus the world has the order 3 (33-30=3). Too complete the topdown description three indices hkl are introduced, which satisfy the condition,

$$2^N = h \cdot 2^{N-m} + k \cdot 2^{N-m-1} + l \cdot 2^{N-m-2} \tag{1}$$

with l,k=0,1,2,3 and h=1,2,3,… since there is at least 1 largest subsociety. The number of hkl solutions is equal to 4^m as can be verified from Table 2. To avoid confusion the k-indices are underlined.

Table 2: The hkl solutions for m=0,1 and 2.

m=0	m=1		m=2			
hkl	hkl	hkl	hkl	hkl	hkl	hkl
100	200	120	400	320	240	160
		112		312	232	152
		104		304	224	144
					216	136
					208	128
						1110
						1012

The world with N=33, m=3, h=2 (China and India), k=0 (no country with N=29) gets then an l-value of 24 as calculated from equation (1).

The model can be applied to all kind of societies, from soccer teams to parliament seats occupied by political parties (Kleizen [2]).

4 World cultures

Hofstede [7] in his well-known study recognizes five cultural dimensions: The Power Distance Index (PDI), the Uncertainty Avoidance Index (UAI), the Individualism Index (IDV), the Masculinity Index (MAS) and the Long Term Orientation Index (LTO). A society has its own set of 5 values for these 5 indexes. Hofstede supplies data for 66 countries and ITIM [13] gives elaborated guesses for another 22, so that a list can be made (Kleizen [2]) of 88 countries of the 236 in previous paragraph. (Since LTO was discovered later, some countries are described by 4 dimensions.)

Kleizen [2,6] does make a distinction between the 5 dimensions. Sex came late on an evolutionary time scale and long term orientation is definitely not in Darwinism. So PDI (=B), UAI (=C) and IDV (=D) are considered more basic and applicable also to mammals and more simple living creatures. A second simplification is to look at the sequence of values not at the absolute values themselves. Than there are of course 3!=6 sequences: BCD, CBD, CDB, DCB,

DBC and BDC. In Table 3 the 88 countries are classified by their sequence (= World culture). 83 countries have unique sequences and 5 countries have double sequences because two indexes are the same.

Population averages can be calculated for each index and sequence using the general formula,

$$V = \sum_{i=1}^{i=j} f_i \cdot V_i \tag{2}$$

with V the population average of a given index of a given sequence, f_i the population fraction of country i, V_i the index of country i and j the number of countries with that sequence. Countries with double sequence participate in both sequences with half of the population.

Table 3: B,C,D, sequence of the 88 countries [2].

nr	BCD	CBD	CDB	DCB	DBC	BDC
1	Albania	Argentina	Austria	Australia	Sweden	Bhutan
2	Bangladesh	Brazil	Belgium	Canada		India
3	Burkino Faso	Bulgaria	Czech Republic	Denmark		Jamaica
4	China	Chile	France	Finland		Singapore
5	Dominican Republic	Colombia	Hungary	Germany		Slovakia
6	Ecuador	Costa Rica	Israel	Ireland		
7	Ethiopa	Croatia	Luxembourg	Italy		
8	Fiji	Egypt	Malta	Netherlands		
9	Ghana	El Salvador		Norway		
10	Hong Kong	Greece		New Zealand		
11	Indonesia	Guatamala		Switzerland		
12	Iraq	Iran		United States		
13	Jordan	Japan				
14	Kenya	Korea South				
15	Lebanon	Mexico				
16	Malawi	Pakistan				
17	Malaysia	Peru				
18	Morocco	Poland				
19	Namibia	Portugal				
20	Nepal	Russia				
21	Nigeria	Slovenia				
22	Panama	Spain				
23	Philippines	Surinam				
24	Saudi Arabia	Taiwan				
25	Sierra Leone	Turkey				
26	Sri Lanka	Uruguay				
27	Syria					
28	Tanzania					
29	Venezuela					
30	Vietnam					
31	Zambia					
nr	BCD	CBD	CDB	DCB	DBC	BDC
1	Romania	Romania				
2	Thailand	Thailand				
3			Estonia	Estonia		
4				South Africa	South Africa	
5				United Kingdom	United Kingdom	

The results of the calculations are collected in Table 4. Per index there are 2 maxima, 2 intermediate and 2 minimal values. The differences in each pair are for C smaller than for B and D. By a procedure mineralizing deviations it can be shown that the set {84,48,36} represent the best fit for maximal, intermediate

and minimal value for the 3 indexes. In Figure 4 these values are depicted in the B,C plane. (In the other planes D,C and D,B the figure the same but the position of the sequences changes with the axes.)

The lines connect the six cultures along the shortest path. Mankind started in Africa and strolling through Table 3 it seems that cultural change is quite slow. All phenomena could be explained by starting with a BCD culture in Africa, going East to find BDC (India) or walking West generating successively CBD, CDB, DCB and DBC. Whether or not these cultures were generated before migration or are the result of migration is lost in the mist of times.

Table 4: Population averages of B,C,D-sequences and their sum [2].

Sequence	Nr of countries	Population July2005est	B=PDI	C = UAI	D = IDV	B+C+D
BCD	32	2494209634	79.2	41.1	21.4	142
CBD	27	1322629339	67.0	81.8	33.5	182
CDB	8.5	107263665	56.7	83.5	68.6	209
DCB	13.5	589420094	39.5	51.6	83.0	174
DBC	2	61394571	39.5	39.2	77.7	156
BDC	5	1095089282	77.1	39.8	47.9	165
BCD	all 88	**5670006583**	**71.0**	**52.2**	**37.3**	160

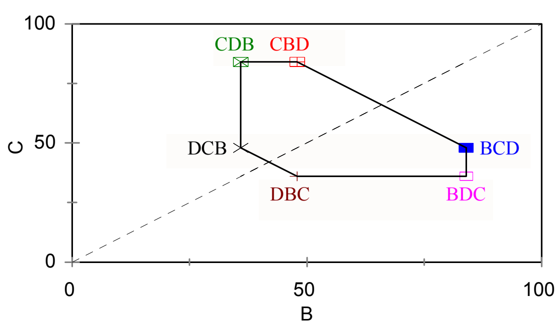


Figure 4: The 6 sequences in the B,C- plane allowing only the values {84,48,36} [2].

5 Linking the 2 models

The society structure models deals with the relative size of societies. The culture sequence model is based upon the sequence of 3 of the 5 Hofstede dimensions. The individualism index deals with the behavior of individuals within a group and cannot be related to different societies. The uncertainty avoidance index focuses on handling uncertainties and can not possible be related to other *known* societies. Therefore Kleizen [2] argues the link - if any – should be found in the PDI (=B) index. Hofstede [7] distinguishes 40 key differences (in the family, at

school, in the work organization) between low and high PDI values. One of them could make the connection. Flat organization pyramids have a low PDI, tall organization pyramids correlate with high PDI. The structure model cannot be flatter than hkl=h00. So flatness could bridge both models. The population fraction of the largest subsocieties r,

$$r = \frac{h \cdot 2^{N-m}}{2^N} = \frac{h}{2^m} \tag{3}$$

is introduced as a measure of flatness. If r=1 the society structure is flat and B (=PDI) is expected to be low and with decreasing r, B should increase.

First evidence is found in the calculated culture of the European Union up to its 11th lustrum. (No Hofstede data are available on Cyprus, Latvia and Lithuania so the comparison is based upon 98.6% of the EU-27 population.) From Table 5 it can be learnt that r and B behave in the way expected.

Table 5: EU evolution 1952-2007 in terms of r, B and sequence [2,6].

EU		Population est	r	B = PDI	Sequence
1952	EU-6	1.6E+08	0.5	50.2	CDB
1973	EU-9	2.6E+08	1	45.9	DCB
1986	EU-12	3.3E+08	1	48.3	DCB
1995	EU-15	3.7E+08	1	46.3	DCB
2004	EU-25	4.6E+08	0.5	49.2	CDB
2007	EU-27	4.9E+08	0.5	51.4	CDB

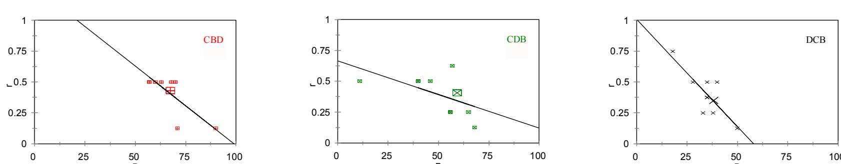


Figure 5: Regression line and population average (larger symbol) for the 3 major sequences of EU-27 countries in terms of their subsocieties (Amts, etc). After [2,6].

In the case of the EU the culture of the subsocieties is known and the culture of the society (the EU) calculated. Looking at countries in terms of Amts Bundesländer, counties, departments, krajs, provinces etc, the cultures of the larger subsocieties are unknown and are assumed the same as the country. In the EU the major sequences are CBD, CDB and DCB, see also Table 3. The dependence of r on B is depicted fin Figure 5 for these 3 major cultures of EU-countries.

It can be concluded that B and r correlate as expected upon the level supranational/national as well as national/regional.

6 Discussion

In this paper 3 models have been shortly presented. The STW is a concept of how large the Technological World might be without disturbing the balance with the Biosphere and Geosphere on planet Earth. The structure model of societies enabling the world population in nested societies (continental societies, supranational entities like Asean, EU and Nafta, countries, provinces, cities) but also in term of multinationals, national companies, LE's and SME's and non-profit organizations. The culture model as presented divides the World in 6 cultures. It is still incomplete as only data on 88 of 236 countries are known, but in these 88 countries 88% of the World population lives, so the lack of data is not a major concern.

Together the 3 models have the potential to become basic instruments in a toolkit to increase the rate of sustainable change and avoid durable excursions in the jungle left and right of this road to be paved.

An example to illustrate this point: China (BCD), India (BDC) and the USA (DCB) have different cultures. The EU may stick to the fourth culture CDB, but if it wants to beat the competition, it should first research why CDB is so innovative. And if isn't, the EU should reformulate it research frameworks and perhaps give credit to its major cultures CBD, CDB and DCB to take the driving seat.

The 3 models are models and could become stronger if they could be unified. The future will learn from the past [14].

7 Conclusion

The STW, structure and culture model are valuable instruments in the durable sustainable development theories of Life on Earth.

References

[1] Brundtland G., *Our Common Future*, The World Commission on Environment and Development, Oxford University Press: Oxford, 1987.

[2] Kleizen H.H., *Towards a Sustainable Technological World*, Thesis Delft University of Technology, in press, 2006.

[3] Kleizen H.H., *Air and Water Filtration in a Sustainable Technological World,* In: Höflinger W. (ed). Chemical industry and Environment V. Vienna (Austria), 2006: 785-792.

[4] CIA - The World Factbook, www.cia.gov/cia/publications/factbook/geos/in.html

[5] Kleizen H.H., *Structuring the World and Sustainable Air Sharing*, In: Mander U, Brebbia CA and Tiezzi E (ed) The Sustainable City IV, Proceedings 4th Int Conference Sustainability City 2006, Wit Press, Tallinn (Estonia), 2006: 303-311.

[6] Kleizen H.H., *A Novel Method to structure the World in Population Entities and Application to European Countries and Regions*, 46th ERSA Conference, Aug 30- Sep 3, Volos (Greece) 2006.

[7] Hofstede G., *Culture's Consequences, Comparing Values, Behaviors, Institutions and Organizations Across Nations,* Sage Publications, California, 2001.

[8] Lide, D.R. & Frederikse, H.P.R., (eds*), Handbook of Chemistry and Physics*, CRC Press: Boca Raton, Ann Archor, London & Tokyo, pp. 14-4 & 14-5, 1993.

[9] Houghton J., *Global Warming: The complete Briefing*, Cambridge University Press: Cambridge, 1997.

[10] Grübler, A., *Technology and Global Change*, Cambridge University Press: Cambridge, 1998.

[11] Smil, V., *Elementaire Kringlopen- Wisselwerking tussen biosfeer en beschaving*, Natuur & Techniek: Amsterdam, 1999. Dutch translation of: Smil., V., *Cycles of Life, Civilization and the Biosphere*, The Scientific American Library: New York, 1997.

[12] Dunn, S., Decarbonizing the Energy Economy, in: *State of the World 2001, A Worldwatch Institute Report on Progress towards a Sustainable Society*, W.W.Norton & Co: New York, Chapter 5, pp. 83-102, 2001.

[13] ITIM, *5-D Pocket Guide*, Brochure, 2001.

[14] Kleizen H.H., *S_j Particle Systems: Predicting Number of the Chemical Elements and Shape and Function of Crystals and Enabling Molecules,* Unpublished Work, 1970-1982.

Laboratories of democracy: the collision of federal and local global warming policy in the United States

J. Manko
Manko, Gold, Katcher & Fox, LLP, USA

Abstract

Scientists have reached a consensus that global warming is a reality and human emissions of greenhouse gases are affecting the climate. As the negative effects of global warming become more certain and more pervasive, the federal government of the United States has, to date, largely refused to take action. Hurricane Katrina (2005) and its aftermath not only provide a case study of the federal government's failure to take positive steps, but also its tendency to make regressive policy decisions on global warming. At the state, local and individual level, however, the problem of greenhouse gas emissions is beginning to receive the attention it deserves. As pressure for action mounts, the conflicting federal and local views have now reached the courts. The U.S. Supreme Court has decided to hear a case, brought by petitioning states, cities and public interest groups, demanding that the federal government regulate greenhouse gas emissions. Although this will be an important case, it is unlikely that the local interests will prevail and even a victory for the petitioners would have only a limited effect. The more important meaning of the case is an acknowledgement by at least one branch of the federal government that state and local pressures to take action to limit greenhouse gas emissions can no longer be ignored. Comprehensive action on greenhouse gas emissions in the U.S. must come from the executive and legislative branches.
Keywords: law, global warming, greenhouse gas emissions, United States environmental policy.

1 Introduction

The scientific case for the existence of human-made climate change has now been established (Karl *et al*. [1]). Scientists have catalogued a wide range of phenomena caused by global warming within the United States alone (Parmesan and Galbraith [2]). Although scientific efforts must continue to observe, define and quantify climate change, global warming can no longer be discounted as a purely scientific phenomenon or an academic question. Scientific and technological breakthroughs will be indispensable to combat its problems, but these breakthroughs may never be implemented, or perhaps even attempted, without sufficient support from governments throughout the world. In the twenty-first century and beyond, global warming cannot be stopped through science alone; it must be stopped by sound environmental policy on a global scale.

The Vermont Environmental Law Center describes good environmental policy as "represent[ing] affirmative decisions to protect natural systems from the impact of human activities. Such policy must be formed at the intersection of politics, law, science, economics and ethics." Thus, good environmental policy requires input not only from scientists; it also requires the work of lawyers, politicians, businesspersons and private citizens. This paper will investigate the political and legal struggles to develop good global warming policy in the United States.

2 Pre-Katrina U.S. federal policy on global warming

The defining element of the U.S. government's policy on global warming – both before Hurricane Katrina and at present – is the refusal to implement the Kyoto Protocol to the United Nations Framework Convention on Climate Change ("Kyoto Protocol"). While former Vice-President Al Gore symbolically signed the Kyoto Protocol in 1998, neither the Clinton nor Bush administrations presented the treaty to the Senate for ratification. President Bush has made it clear that he will not sign Kyoto because of its cost to the U.S. economy and its failure to regulate developing countries such as India and China, other major greenhouse gas emitters. Despite the U.S. refusal to ratify the treaty, the Kyoto Protocol came into effect in February 2005. The U.S.'s status as the largest emitter of greenhouse gasses makes its failure to participate extremely damaging to the overall goals of Kyoto.

Greenhouse gas emissions and global warming have not been entirely ignored by the federal government. The Energy Policy Act of 2005, signed into law by President Bush, contained funding for the development of renewable energy sources that do not emit greenhouse gases and an initiative to increase the amount of clean-burning ethanol used in the U.S. Perhaps more importantly, the enacted legislation deleted several provisions from the original bill that would have relaxed Clean Air Act regulations on power plants and permitted drilling for oil within the Alaskan National Wildlife Refuge.

Although the Energy Policy Act of 2005 did not ignore global warming, many contend that the provisions cited above merely paid lip service when taken in context with the bill as a whole. The funding for renewable energy sources in the bill was dwarfed by the funding provided for traditional energy sources such as oil and coal. Provisions calling for Kyoto-like regulation of greenhouse gas emissions in the U.S. in the original bill were completely excised from the enacted legislation.

3 Katrina as a case study of federal global warming policy

There is a scientific consensus that rising temperatures in the tropical Atlantic Ocean are responsible for the increased violence of hurricanes over the past decade. A recent study supported by the U.S. National Science Foundation finds that global warming, and not natural fluctuations, is causing this temperature increase (Trenberth and Shea [3]). If human activity is indeed raising ocean temperatures, then the United States is already suffering the first adverse affects of global warming.

The 2005 Atlantic hurricane season had several record-setting storms, including the most destructive storm of all time, Katrina, and the most intense storm ever to hit the Gulf of Mexico, Rita (Nesmith [4]). Many hoped that Katrina's devastation would serve as a "canary in the mineshaft," alerting the federal government that it must take immediate action to limit greenhouse gas emissions. The reality of the situation is the exact opposite – the federal government has ignored the warning and its actions in the wake of Katrina have actually been counterproductive. The federal government's response to Katrina provides a case study of regressive policies on global warming.

It is important to note here that this critique of current U.S. federal policy does not suggest that the current administration, or even the U.S. as a whole, is in some way "responsible" for Hurricane Katrina. It is impossible to assign global warming as the cause of any individual hurricane. Furthermore, the causes and possible consequences of global warming existed and had been noted scientifically by the end of the nineteenth century. Assigning blame to an administration that has been in power for roughly six years, regardless of its policy decisions, would be unfair.

Although the federal government's inadequate emergency response to Katrina has been strongly criticized, the regressive response reflected in federal energy policy is equally troubling. If the harrowing images of destruction provided by Katrina result in the government actually *decreasing* regulation of greenhouse gas emissions, it is hard to imagine the evidence necessary to spur responsible change.

In the aftermath of Katrina and Rita, federal policymakers were not concerned with the possible *causes* of the unusually destructive hurricane season; they were concerned with its *effects* on traditional, petroleum-based U.S. energy supplies. Katrina and Rita damaged the U.S. capacity to refine oil, and this and other factors resulted in a huge increase in energy prices, especially prices at the pump for gasoline consumers.

As a result, the immediate response of the U.S. Environmental Protection Agency ("EPA") to Katrina was the temporary waiver of gasoline and diesel fuel standards in Louisiana, Mississippi, Alabama and Florida. More limited waivers were issued to the other 46 states. A bill was introduced in the Senate that would have granted the EPA administrator authority to waive any regulation for purposes of disaster recovery or if he considered it to be in the public interest. The bill never reached the Senate floor for a vote.

Similarly, the U.S. House of Representatives' immediate response was to pass the "Gasoline for American Security Act," commonly called the GAS Act, to encourage increased refining capacity in the U.S. Opponents alleged that the bill was a handout to gasoline companies, would result in more emissions and dirtier skies, and was essentially an attempt to push through, under the auspices of national security, policies rejected from the aforementioned Energy Policy Act of 2005. The House narrowly passed the bill over vigorous opposition, but the Senate has yet to take action on the bill.

In a surprising about-face, President Bush declared in his 2006 State of the Union Address that the country was "addicted to oil" and promised to increase efforts to find renewable sources of energy. His specific pledge to cut imports of foreign oil by 75 percent before 2025 was most striking. The very next day, however, the U.S. energy secretary and national economic adviser announced that the President had not really meant what he said – rather, the United States would always import oil from the Middle East and the President's statements had been "merely an example" (Neikirk and Martin [5]).

More recently, the Bush administration has congratulated itself on its efforts to limit greenhouse gas emissions. The cause for celebration was not a decrease in the *amount* of U.S. emissions, but merely a decrease in the *rate* at which emissions were increasing. Such a position ignores the international consensus, enshrined in the Kyoto Protocol, that stabilizing the global climate requires a *reduction* in the amount of greenhouse gas emissions far below current levels.

There are some sources of hope within the federal government. Most notably, the U.S. Senate passed a resolution in 2005 calling for a program to regulate greenhouse gas emissions. A resolution, which has no binding effect, is far from the comprehensive regulatory program some would desire, but it is far better than the increasing deregulation of emissions seemingly favored by the U.S. House of Representatives.

4 State and regional responses

The U.S. is often seen by the rest of the world as a monolithic political entity embodied by its President. The political reality of the U.S. federal system is far more complex than this view suggests. Many of the foregoing examples of regressive policy were vigorously opposed by numerous states and cities. As an example, perhaps the least responsible bill, the GAS Act, was objected to strenuously by the National League of Cities, the United States Conference of Mayors and the Environmental Council of the States – all national organizations representing state and local interests.

While the international community attempts to exert pressure on the U.S. from *without*, the most likely source of change will be the pressure from *within*. Some leverage derives from traditional public interest groups with environmental agendas, but increasingly it comes from individual states and cities. To some extent, states and cities have sought to pressure the federal government by filing lawsuits discussed later in the paper. In addition, states and cities have begun to step into the regulatory role that the federal government has largely declined to play.

Perhaps the most important regulatory effort is an agreement between several eastern states that has been described as a "mini-Kyoto." (Globe Editorial [6]) The Regional Greenhouse Gas Initiative ("RGGI") seeks to create a cap-and-trade market for carbon dioxide allowances starting in 2009. Disagreements over implementation of the program have slowed its progress and caused two states (Rhode Island and Massachusetts) to drop out. Critics cite as major problems the RGGI's failure to regulate greenhouse gasses other than carbon dioxide, its sole focus on power plants, and the possibility of leakage (i.e. states importing "dirty" energy from outside the region). Despite these real concerns, the RGGI remains the most serious attempt to regulate greenhouse gas emissions in the U.S.

On the west coast, California is leading efforts to limit greenhouse gas emissions. Although some plans in California may be further from implementation than the RGGI, they exceed the eastern states' efforts in both breadth of regulation and public support. California has already demonstrated its commitment to this issue by passing restrictive regulations on automobile emissions. Spurred by California Governor Schwarzenegger's 2005 executive order establishing targets for emissions reduction, state lawmakers have introduced comprehensive legislation on greenhouse gas emissions. The program would create an emissions trading program similar to RGGI, but not limited to power plants. It would also create mandatory emissions reporting and a requirement that all new energy generated in the state come from sources with greenhouse gas emissions less than or equal to efficient modern power plants. While public support for the legislation appears to be high, there is some concern that the bill will be severely modified before passage due to industry pressure.

California has also stepped onto the international stage on global warming issues. In August 2006, Governor Schwarzenegger met with British Prime Minister Tony Blair to announce an agreement between the two governments. The collaboration's practical extent is limited to the exchange of information. While the move is viewed as largely symbolic, the agreement between a U.S. state and a foreign government is an unprecedented attempt to circumvent the federal government's lack of engagement on this issue at the international level.

In America's federal system, states are often described as "laboratories" for federal policy. If the experiments conducted by California and the RGGI are successful at curbing emissions while not unduly damaging state and regional economies, their example would undercut much of the opposition to a national regulatory scheme. Although state and regional actions can never be satisfactory

substitutes for federal action, they can serve as both proof-of-concept and a model framework for a comprehensive national program.

5 Business and industry responses

One surprising source of positive change has been industry itself. In the past, efforts from some industry sectors appeared to focus on minimizing the climate change problem and maintaining a positive public image. For instance, during the Kyoto negotiations, an industry group known as the Global Climate Coalition spent millions of dollars on advertising downplaying the effects of global warming. In 2002, the coalition was disbanded after major members such as Ford, General Motors and Texaco became convinced that global warming was a reality and left the group. One former member, BP Amoco, has become an industry leader in reducing greenhouse gas emissions and joined the Pew Center on Global Climate Change Business Environmental Leadership Council ("BELC"), a group of businesses committed to fighting global warming.

The members of BELC are taking steps to address global warming of their own accord, without prompting from the federal government. Because its members are multinational companies, BELC can take its efforts beyond the U.S. and onto the international stage. While the U.S. and the European Union have so far been unable to reach an understanding on global warming, many U.S. and European companies have bridged the ideological gap. As the U.S. companies take it upon themselves to comply with international regulations, European companies have begun to seek an international market-based approach with which some European nations take issue, but the U.S. would prefer.

Efforts to decrease greenhouse gas emissions have not been limited to BELC. Many more companies have entered into compliance with voluntary programs administered by the federal government. In the newest data from these efforts, the 226 companies participating in one program lowered their 2004 emissions by three percent from the previous year. While the information from these voluntary regulatory programs has been positive, some continue to cite the obvious problems with any voluntary system: participation is limited and some participants do not rigorously follow the program.

The realities of regional and international regulatory programs force many companies to adopt emissions-limiting policies above and beyond those suggested by voluntary federal programs. As regional emissions restrictions emerge in the U.S., companies that do business on the national level are forced to consider programs to limit emissions for at least some of their operations. Multinational companies face another source of regulation to the extent their international operations are regulated by the Kyoto Protocol. Some companies, such as 3M, respond to these dual pressures by implementing company-wide – and therefore global – policies on greenhouse gas emissions.

The shift in industry attitude about global warming is a welcome development. Although many businesses may initially engage in these efforts to improve their public image and, for large companies, to comply with international regulations, they are also driven by the simple economic motive

that drives most business decisions: making a profit. As companies have accepted the reality of global warming as a concern for their financial future, they have begun to adopt methods for dealing with it as they would for any other threat to business (Pew Center on Global Climate Change-BELC [7]). Against this backdrop, the continual downplaying of global warming by the federal government and its characterization of emissions regulations as "bad for the economy" appears even more insupportable when contrasted with the acceptance of its reality by these profit-driven entities.

6 Collision in the Courts

Louis Brandeis, the late U.S. Supreme Court Justice, introduced the metaphor of states as laboratories. Justice Brandeis wrote "a single courageous state may, if its citizens choose, serve as a laboratory; and try novel social and economic experiments without risk to the rest of the country" (New State Ice Co. v. Liebmann [8]). This metaphor for America's federal system has been an enduring one. Today, however, some states are dissatisfied with their role as laboratories and instead intend to serve as soldiers on the front line, attacking the federal government's interpretation of its own responsibilities under federal law. The final battleground for these disputes will almost certainly be the U.S. Supreme Court.

The Supreme Court recently announced it would hear an appeal in a case brought by several states, cities and public interest groups seeking to force the federal government to regulate greenhouse gas emissions. In *Massachusetts v. EPA*, the U.S. Circuit Court of Appeals for the District of Columbia found that EPA acted within the law in refusing to regulate these gases, but the Supreme Court's decision to hear the appeal means the high court will have the final word in this dispute (Massachusetts v. EPA [9]).

This decision to hear the case provides a reason for those concerned about global warming to take heart – the federal courts have previously been responsive to concerns about EPA shirking its duty to regulate emissions. For example, in *New York v. EPA*, a coalition of 14 states and several cities challenged EPA's attempt to loosen regulations on old power plants (New York v. EPA [10]). In March 2006, a federal appeals court ruled that EPA's position violated the Clean Air Act. The court delivered a stinging rebuke to EPA, declaring the agency's interpretation reasonable only in a "Humpty Dumpty world" [10]. This was a major victory for states, cities and public interest groups.

On the other hand, there are also reasons to be extremely concerned about the possible outcome. The makeup of the U.S. Supreme Court changed recently by President Bush's appointment of two conservative judges to fill vacant seats. The new Court has already handed down its first environmental decision in a case regarding the protection of wetlands, in which both new Justices joined the opinion that would have severely limited wetlands protection. The Court overall, however, was so divided that it could not issue a majority opinion and simply remanded the case for further consideration by the lower courts. Although it is

difficult to characterize the Court generally or to predict its decision in the upcoming case, it would certainly be a mistake to view the current Court as pro-environment.

The resolution of the lower court case from which the *Massachusetts v. EPA* appeal springs should also be cause for concern. The three-judge panel hearing the case issued three separate opinions [11]. One judge would have dismissed the case on its merits, finding EPA's decision not to regulate greenhouse gases legally permissible. A second would have dismissed the case for a lack of standing – a jurisdictional issue – because the plaintiffs had not been "injured" by EPA's decision. The final judge found EPA's interpretation of the Clean Air Act impermissible and would have required EPA to regulate greenhouse gas emissions. The result of these separate opinions was a dismissal of the case and the subsequent appeal to the Supreme Court. What concerns environmentalists is the existence of two legal theories upon which to dismiss the suit, and only one on which to find for the plaintiffs. If the Supreme Court splits along the same lines the suit will be dismissed.

In the end, the entire suit may be dismissed for a lack of standing. Standing is often thought of as a legal technicality, but here it provides insight into the nature of the conflict the U.S. faces. Under the U.S. Constitution, a lawsuit may only be brought if the plaintiff has suffered a "particularized" injury. As the appeals court in *Massachusetts v. EPA* noted: "[A] plaintiff raising only a generally available grievance about government – claiming only harm to his and every citizen's interest in proper application of the Constitution and laws, and seeking relief that no more directly and tangibly benefits him than it does the public at large – does not [have standing]" [10].

The doctrine of standing defines the proper role of the courts in the U.S. – to redress injuries suffered by individuals. A rejection of this claim based on a lack of standing is not a statement that no problem exists. Instead it means that a problem of this magnitude, affecting every individual in the U.S. and on the planet, should not be redressed in a court of law, but rather must be addressed by the Congress as a political matter. Common-sense takes us one step further than the doctrine of standing: when a problem affects everyone on the planet, it must be addressed by the world as a whole.

Even a victory by the plaintiffs in this case would have a relatively small impact on greenhouse gas emissions in the U.S. The plaintiffs' claim is limited to one section of the Clean Air Act dealing with automobile emissions, and thus would not affect power plants or other sources of emissions. It also seems certain that today's EPA would, if forced, adopt the least stringent restrictions on emissions that would satisfy the Court, and those restrictions would likely be slow in coming.

7 International legal responses

Some environmentalists view international law as another avenue for legal challenges to U.S. policy. One example is a recent petition filed with the Inter-American Commission on Human Rights. The petition, filed by a group

representing the inhabitants of the Arctic Circle, seeks a declaration that U.S. refusal to regulate greenhouse gas emissions represents a violation of their human rights (Choo [11]). The likely impact of the petition is limited, as it does not seek damages, although a favorable ruling could theoretically ground tort claims against the U.S.

At the present time, international legal action remains, like the petition, a largely theoretical method of influencing the federal government, and there are serious questions about its possible efficacy. One reason for this is that the U.S. has not ratified the treaty creating the International Court of Justice. As such, an opinion from that court could only express international displeasure at U.S. policy. Most international legal efforts will suffer from the same limitation, and the U.S. has made clear in the past it is willing to buck international opinion on the issue of global warming.

8 Conclusion

The hope inspired by the Supreme Court's decision to hear *Massachusetts v. EPA* is not misplaced, but it may spring from the wrong sources. It is true that sound global warming policy is not purely scientific, and its implementation will not be achieved solely through technology. Supreme Court watchers must realize, however, that good environmental policy is also not purely legal and no court can ever provide *the* solution.

The Supreme Court's decision to hear a case involving greenhouse gas emissions is a reason to hope because it represents an acknowledgement, by a branch of the federal government, that state, local and public opinion can no longer be ignored on this issue. Any real step towards good environmental policy on global warming in the U.S. will have to come from the U.S. Congress and the president. But as pressure mounts to do something about greenhouse gas emissions, the political resistance to regulation will become weaker and weaker. While the federal courts have provided a venue for state and local interests to make their demands felt, the final battleground on this issue will lie within the legislative and executive branches. Thus, the congressional elections in November 2006, and the presidential election in November 2008, may answer the question of how long it will take the U.S. federal government to join with the rest of the world to address the problems created by global warming.

Acknowledgements

The author would like to acknowledge the assistance of Rodd Bender and Chris Havener in researching this paper.

References

[1] Karl, T., Hassol, J., Miller, C. & Murray, L. (eds.). *Temperature Trends in the Lower Atmosphere: Steps for Understanding and Reconciling*

Differences, Climate Change Science Program and the Subcommittee on Global Change Research: Washington, DC, 2006.
[2] Parmesan, C. & Galbraith, H, *Observed Impacts of Global Climate Change in the U.S.*, Pew Center on Global Climate Change: Arlington, VA, 2004.
[3] Trenberth, K. & Shea, D., Atlantic hurricanes and natural variability in 2005. *Geophysical Research Letters,* 33, L12704, 2006.
[4] Nesmith, J., Global warming blamed for hurricanes. *Philadelphia Inquirer*, 23 June 2006, http://www.philly.com/mld/philly/14881262.htm, 2006.
[5] Neikirk, W. & Martin, A., Energy-saving technologies are years away. *The Chicago Tribune*, 2 Feb. 2006, p. 4, 2006.
[6] Globe Editorial, Capping the greenhouse. *Boston Globe*, 4 June 2006, http://www.boston.com/news/globe/editorial_opinion/editorials/articles/2006/06/04/capping_the_greenhouse, 2006.
[7] Pew Center on Global Climate Change, Business Environmental Leadership Council, http://www.pewclimate.org/companies_leading_the_way_belc
[8] *New State Ice Co. v. Liebmann,* 285 U.S. 262, 311 (1932) (Brandeis, J., dissenting).
[9] *Massachusetts v. EPA*, 415 F.3d 50 (DC Cir. 2005).
[10] *New York v. EPA*, 443 F.3d 880 (D.C. Cir. 2006).
[11] Choo, K. Feeling the heat. *American Bar Association Journal*, July 2006, pp. 29-35, 2006.

Economic development and *colonias* in Texas

C. Giusti
Department of Landscape Architecture and Urban Planning,
Texas A&M University, USA

Abstract

This paper concentrates on economic, institutional, financial and legal matters that have a direct impact on sustainable planning and development. It reflects the research that is being developed in *colonias* on the Texas side of the US - Mexico border region. *Colonias* in Texas are defined as areas with substandard housing, inadequate plumbing and sewage disposal systems, and low-income residents. Three case-studies are presented: the first refers to micro businesses and their potential to promote local economic development; the second is related to the "legalization" of "informal" developments; and the third refers to the sustainability of micro-credit for house improvement. *Colonias*, mainly isolated from urban areas, are starting to enjoy a relatively stable population, and regardless of the many difficulties they face, *colonias* are showing signs of being livable communities. This paper discusses these low-income communities from the point of view of promoting local economic and community sustainable development.
Keywords: economic development, micro credit, land tenure, housing.

1 Introduction

When approaching communities lagging economic dynamism, with no defined or stable institutions, with unclear legal framework, there is a concern on how sustainable are these communities. This paper will deal, specifically with three research projects in *colonias* along the Texas-Mexico border region related to their land, legal, and economic concerns. The first case study refers to a research done on the contribution of micro-businesses to local economies; the second one related to the effects of micro-credit for house improvements on the well-being con *colonias*; and the third one is related to a lot-titling regularization program and its effects on local residents.

www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060071

Colonias, distinct to most of the United States, resemble much of the characteristics of developing countries: institutional instability, uncertain legal regulations, low education levels, poor healthy and safety conditions, and so on. These are communities not only with low income population, but also with very minimal infrastructure and high levels of isolation. The purpose of this paper is to discuss ways to promote local economic development from different perspectives, and try to assert which programs may be suitable, in our quest for improving quality of life in a sustainable way. Specifically, the three case-studies point on issues related to housing, economic activity and land tenure and they will be discussed as they are related to sustainable local economic development.

The specific location of the *colonias* in our research is shown in Figure 1. The land and credit improvement studies were mainly made in Starr County, while the study on microbusinesses was done in Webb, Hidalgo and Cameron Park counties.

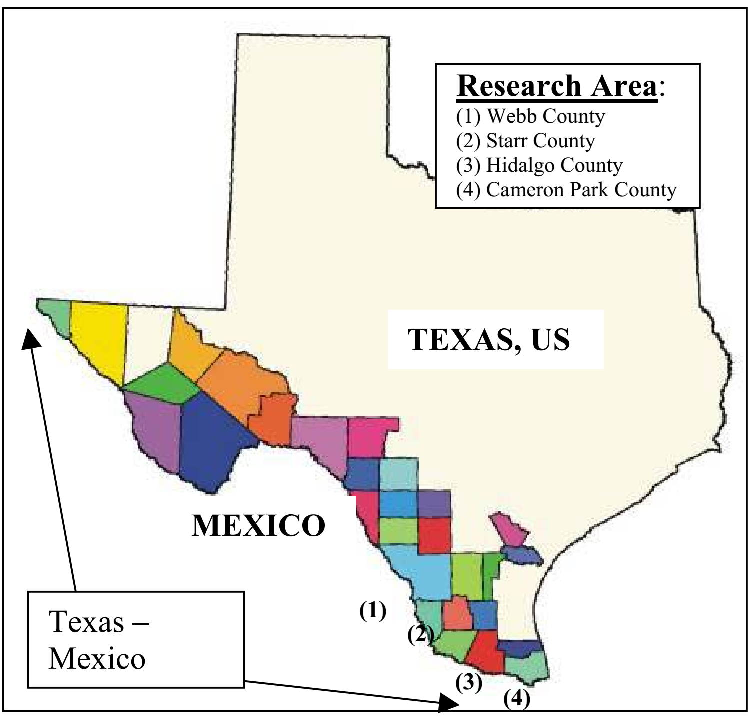


Figure 1: Map of Texas border counties and location of the case-studies.

Local economic development is defined, in this paper, as a process of creating assets to "generate improved and broadly shared economic well-being and quality of life for a community or region" [13]. Our approach is holistic, combining human, social and economic, legal and institutional aspects in the

development process. The paper is organized as follows: first we present the main characteristics of *colonias* in Texas; then we present the three case-studies, and then we compare them and outline our conclusions as well as few policy suggestions.

2 *Colonias* in Texas

Following the Office of Attorney General definition, *colonia* in Texas refers to a residential area along the Texas-Mexico border that "may lack basic water and sewer systems, electricity, paved roads, and safe and sanitary housing" [11].

Table 1: Basic characteristics of *colonias* in Texas.

	Rio Bravo Colonia	Webb County	Texas	US
% Hispanic Pop.	97.7%	94.3%	32.0%	12.5%
% High school graduate or higher/ >25	21.5%	53.0%	75.7%	80.4%
% Bachelor degree or higher/ 25 >	2.0%	13.9%	23.2%	24.4%
Same house 1995	70.3%	58.8%	49.6%	54.0%
Born in US	60.3%	69.7%	85.0%	87.7%
Spanish at home	95.3%	91.3%	27.0%	10.7%
Speak English less than "very well"	58.2%	44.2%	12.3%	5.2%
% Unemployed	6.9%	4.9%	3.8%	3.7%
% Of individuals below poverty level	58%	31%	15%	12%

Colonias can be found in the US-Mexico border counties of Texas, New Mexico, Arizona and California, but Texas has both the largest number of *colonias* and the largest *colonia* population.

According to the 2000 census, Texas contains 1,450 *colonias*, a 21.5 percent increase from 1,193 *colonias* recorded in 1992 that are located along the state's 1,248 miles border with Mexico. Currently, 350,000 residents (280,000 in 1992) are concentrated in these *colonias*, an increase of 25 percent in eight years.

The development of Texas *colonia*s dates back to at least the 1950s, but the 1980s and early 1990s showed the largest growth. Developers, taking advantage of scarce housing resources, created unincorporated subdivisions using agriculturally worthless land, land that lay in floodplains, or other rural properties. They divided the land into small lots, put in little or no infrastructure, and then sold them to individuals seeking affordable housing. For a variety of reasons, regulations operating in most urban and rural developments in the U.S. were not mandatory, and *colonia* residents originally had very little control over the basic services provided to them or over their legal status as landowners [8, 18, 20]. Approximately half of the population in *colonias* does not have an adequate water supply and most do not have wastewater services. Most *colonias*

have dirt roads without even gravel surfaces and have no surface drainage systems. Incidence of health problems is high.

Colonia residents are mainly of Hispanic/Mexican origin (97 percent, according to the latest census) [17]. Educational levels are very low in *colonias*. In Rio Bravo, one of the *colonias* studied, only 2 percent of the population has a bachelor's degree or higher. This is much lower than for Webb County at 14 percent, Texas at 23 percent, and the national rate of 24 percent. For illustration purposes table 1 shows some basic characteristics of one "typical" *colonia* compared with the rest of the county at the county, state and national levels.

Median annual household income in Rio Bravo is about one-third of the median household income in the U.S. and about a quarter of that in Texas. Average family size is 4.8 people, much higher than the 3.14 in the nation.

Nevertheless, despite all these problems, *colonias* show more stable population than other communities in the US. An interesting census variable is "residence in the last five years". Whereas in the U.S. as a nation, 54 percent of individuals have lived in the same house for the last five years, in the case of Rio Bravo *colonia*, more that 70 percent declare living in the same house for the last five years. This statistic tells us that population in *colonias* is more steady than in the rest of the country, contrary to the widely held belief that they are "temporary" communities. *Colonias* population is more established than anticipated. This is indeed, what we have observed in our case-studies and is presented next.

3 Case-studies

We will present next the summary of three case-studies related to three aspects of development: economic activity, housing, and land tenure. Each of them represents a variety of topics related to our holistic concept of local economic development in a sustainable way.

3.1 Economic activity in *colonias* – Microbusinesses

Colonias microenterprises will typically be family businesses with very few employees. A microbusiness is defined here as a business with less than 10 employees or a one-person only (also considered as self-employment). Furthermore, in our case we will also add a SBA [14, 15] definition of "disadvantage" entrepreneur: an owner who is denied access or has no access to credit due to his/her lack of income.

Colonias, against prior belief, enjoy a relatively stable population, and regardless of the many difficulties they face, are showing signs of being livable communities. As they are small and un-connected communities their provision of goods and services is limited. To cover these needs many businesses have been established even without financial support. Our observation in *colonias* show the existence of a variety of local shops and business of a wide range of

scope: from basic plumbing and housing maintenance services to tax advise, beauty salons, day-care services, and most of anything, small restaurants and / or grocery stores.

Microbusiness, then, emerge to satisfy the needs of a growing (however small) and dynamic population. Additionally, most *colonia* resident have limited technical skills, lack of language and many of them have no English proficiency limiting their access to labor markets [1, 4, 9]. Self-employment and microbusiness become the alternative of low-income individuals who are in need of increasing family incomes.

Microbusiness owners and self-employment are very similar concepts in the context of *colonias*, and in many instances the concepts are used indistinctively. Self-employed men and women in *colonias* respond to two basic requirements: they serve a need from the community and they pursue their businesses as a means to surviving and as a complementary way to increase their family income.

3.1.1 Financing microbusinesses

The majority of business owners in *colonias* are out of the realm of formal banks and financial institutions. They concentrate in economic sectors that have easy access as they require limited initial investment, therefore the risk is minimized.

Most of the initial capital, as well as their working capital or improvements done, are financed with personal and / or family savings. Our research even shows that in most cases business owners decided not to apply for any loans. This is observed as the rate of loan approval of the few that actually apply is not that low. They self-exclude themselves from applying. Why?

We propose that business owners are rational economic agents, and as any other entrepreneur they maximize profits given particular restrictions: budget, knowledge, network, time, and human capital, we also propose that they minimize risk. Even in the cases they could receive loans, by not entering formal financing, business owners are minimizing the risk connected to such loans.

3.1.2 Microbusiness and local development

The contribution of microbusiness in *colonias* is more than just serving community needs in the form of goods and services [3, 12]. First of all, they allow local residents generate most-needed income. Second, we observed that many business owners pay some kind of taxes or fees. They are very small but still need to comply with basic regulations, especially those in the food or restaurant business – a large majority. Third, although not in big numbers, still they generate some level of employment. In general this is more part-time jobs, but is exactly the type of jobs that are capable of absorbing the type of labor available in *colonias*: low-skill, minority population in need for complementing family income [6, 7, 16]. Any local job creation results on more cash circulation within the community.

3.2 Housing and land market in *colonias*

Housing condition in *colonias* is poor. In imprecise lot plans and unclear legal regulations, houses have emerged without following any government control or

regulation. As in many cases no basic infrastructure was available in *colonias*, many houses were build without water, wastewater or electricity. Even today not all houses have basic services. The majority of the housing stock, however, has improved since.

3.2.1 Incremental construction

It is common to find, as in developing countries, many self-built houses and very little or no-at-all government control to ensure minimum construction standards. As houses started with one or two rooms, with the passing of time and as income is available, they have expanded to accommodate the needs of growing families.

This practice, while common and many times encouraged in low-income neighbourhoods in developing countries, is not common in other parts of the US nowadays. Strict regulation in the construction industry is prevalent and safety and security controls are being tightly enforced in every new building in the US. Nevertheless this is prevalent in *colonias*.

3.2.2 Financing and housing in *colonias*

Once more we observed that also in the case of housing financial institutions have very little presence in *colonias* [19, 20]. As local residents are low-income individuals with poor or inexistent financial history, very likely they will not be on the radar of local banks. The result has been a lack of access to financial credit by the majority of *colonia* residents who have built their basic houses and improvements, with their own limited personal and family savings.

One of the few institutions that actually lend money for this purpose in *colonias* is the Community Resources Group, a no-profit organization who has programs in several counties along the Texas border. They have implemented a micro-credit program assisting home-owners with their improvements. Current research on the benefits and/or problems of this program is undergoing by the author of this paper. Results of this research should be published at the end of this year. However, our initial perception is that very small loans have a dramatic impact on the well being of local residents. First of all, the small amount of the loan make it affordable within limited family budget; second these loans allow homeowners undertake works that are necessary to improve the quality of their houses; third, with that loan residents typically local labour (not only unpaid family members but also full time or part time jobs); fourth, they will also need to buy materials to do the improvement, and we observed that most of the supplies come from local vendors; fifth, as property values increase due to these loan, taxes on these properties increases. All these elements give a more comprehensive view of this program within these communities. A final observation on this microcredit program is that it seems that payment rates are not very different than in formal banks.

3.3 Land tenure

Colonia residents bought their lots through a *contract for deed*, a property-financing method whereby developers typically offer a low down payment and low monthly payments but no title to the property until the final payment is

made [8, 18]. Although this arrangement is very dubious, it has been an affordable alternative for thousands of families who could not afford formal financing systems [10]. Payment was often not a fixed amount and a common practice was that residents paid "as money was available."

The informal, unplanned way in which these communities emerge has been changing and improving since the late 1990s, partly due to deliberate policies enacted by the Texas legislature. Land title has been a key element in this reform. One of the immediate objectives of such legislation was the conversion from *contract for deed* to a more formal legal instrument called *warranty deed*. The objective was to reduce uncertainty and minimize tenure conflicts assigning legal ownership and responsibility for property taxes.

When analyzing the effects of such title program we found that most residents of *colonias* believed themselves to be "owners" of their lots even before receiving a deed. This is similar to what is reported happening in developing countries [5]. As long as they were making payments they had a "sense of ownership". This explains why residents in these *colonias* were making home improvements even without a proper legal deed on their properties. This perception, however, was not validated by the "legal" system.

An important consequence of lot ownership is the fact that the new lot owner must start paying taxes. It was observed that most of them are not reluctant to do so. As taxpayers, residents have changed their perception of their "right" to request basic infrastructure from local officials to improve the provision of services as water, sewage, paved roads. The fact of having legal also meant an "external" validation of their status as owners, and has also had an empowering effect as members of their communities.

In terms of market value, our research shows that in real terms the price per square foot has not changed significantly due to legal ownership. Getting proper title did not change market conditions and the isolation of *colonias* from main cities and their lack of basic services.

Another element that was also explored in our research was to what extent new owners were willing to use this new asset as a backup for entering the financial system. Much is being proposed in terms that "informal" occupation of land is equivalent to having useless assets. Because there is legal rights over the land they occupy residents cannot use this asset in the financial market [2]. Following this argument, residents were expected to use their new lot title as collateral for loans. However, this is not what we mainly observed. Few residents are currently using or expressed their intention of using their loans as collateral. In this case the explanation has two sides: from the "demand" point of view, in most cases new owners are very reluctant to jeopardize their only asset and do not want to risk losing it. From the supply point of view, financial institutions are not actively pursuing these "clients", as the status of new owners has not changed dramatically just by the fact of owning their land. They are still characterized as low-skilled workers who do not have stable and well-paid jobs. As the market value of these lots is not growing dramatically, it is not a "profitable" business for the banks to give loans to individuals who, in their observation, are still high-risk clients.

Therefore this titling program is observed as positive in terms of enhancing quality of life in *colonias*, it has not resulted on a more active participation of the new owners in the financial market.

4 Summary and discussion

We have briefly presented three case-studies related to *colonias* during the last four years on three aspects of development: economic activity, housing, and land tenure. We introduced some of the many complex issues facing *colonias*: from the results of a titling program, to the existence of microbusiness, and the special characteristics of land and housing markets. Next we frame all these elements from the view of local economic and sustainable development.

First of all *colonias* are the result of "free" market forces without government controls. This particular outcome of supply and demand forces did not bring very positive results in terms of quality of life in these communities. And the data collected suggest the need for some type of intervention.

Second, *colonia*, are starting to enjoy a relatively stable population, and regardless of the many difficulties they face, *colonias* are showing signs of being livable communities. As dynamic communities, they are in constant need of more goods and services.

Third, *colonia* residents are rational agents subsisting within very limited budgets, a fragile built environment and unclear legal setting. We found sluggish land and housing markets where residents value their properties more for their use-value than their exchange-value. A house is the place to raise their families and they are not risking this valuable (in personal not in market terms) asset.

Fourth, *colonias* residents do pay taxes: both business and property taxes. On the one side this validates their ownership on their property and business; and on the other side, as tax payers, they can demand more services from local authorities.

Fifth, self-employed men and women in *colonias* respond to two basic requirements: on the one side they serve the growing needs for goods and services from the community; and on the other side they pursue their businesses as a means to surviving and as a complementary way to increase their family income.

Sixth, to promote local development it is not enough to secure property titles. This is a necessary but not sufficient condition. It is also essential to provide basic infrastructure like water, electricity, paved roads; a better education level, access to employment (through self-employment or jobs) and mainly access to credit.

Seventh, an urgent need in *colonias* is credit. As financial institutions are reluctant to give credit to low-income residents with unsecured jobs, it is necessary to delineate more creative ways to provide access to credit. We have briefly presented a successful case of microcredit for house improvement. Very small loans have many advantages: they are affordable and at the same time they have a broader impact in the community in terms of living conditions and

community development. House improvements need labour and materials, and residents find them within their communities, contributing with the local economy. As they are rational agents we should expect that they will maximize the benefits of such loans and this will have positive direct impact in the quality of life in *colonias*.

Finally, financial credit for house improvements can be differentiated from business credit. In the first case it has an immediate impact on health, safety, and security of families. In the case of business loans the effects are more indirect, but it does not mean they are not important. We observed that these businesses make an important contribution to local economic development.

References

[1] Bates, T. *Race, Self-Employment, and Upward Mobility: An Illusive American Dream.* Washington, DC: The Woodrow Wilson Center Press, and Baltimore: John Hopkins University Press. 1997.

[2] De Soto, H. *El misterio del capital: por qué el capitalismo triunfa en Occidente y fracasa en el resto del mundo*. Lima: Empresa Editora El Comercio, 2000.

[3] Glazer, N. Affirmative Discrimination: Ethnic Inequality and Public Policy. New York: Basic Books. 1975.

[4] Glazer, N. & Moynihan, D. Beyond the Melting Pot: The Negroes, Puerto Ricans, Jews, Italians, and Irish of New York City. Cambridge: M.I.T. Press. 1970.

[5] Hall, P. Microenterprise. *World Watch* 9 (March-April):10-19. 1996

[6] Harrison, R. *Houston Hispanic Entrepreneurs: Profile and Needs Assessment.* New York: Garland Publishing. 1995.

[7] Headd, B. The characteristics of small-business employees. *Monthly Labor Review* 123, 4: 13-18. 2000.

[8] Larson, J. Free markets in the heart of Texas. *Georgetown Law Journal* 84: 179-260. 1995.

[9] Light, I. & Rosenstein, C. *Race, Ethnicity, and Entrepreneurship in Urban America*. New York: Aldine de Gruyter. 1995.

[10] Mettling, S. *The Contract for Deed.* Chicago: Real Estate Education Company. 1997.

[11] Office of the Attorney General of Texas (OAG). *Texas and Mexico Border Issues.* http://www.oag.state.tx.us/index.shtml 2006.

[12] Raheim, S. Microenterprise as an approach to promoting economic development in social work. *International Development* 39, 1: 69-82. 1996

[13] Seideman, K. Economic Development Finance. Sage publications: Thousand Oaks. 2001.

[14] Small Business Administration. Characteristics of small business employees and owners. SBA, Office of Advocacy. http://www.sba.gov/advo/stats/ch_em97.pdf, 1998.

[15] Small Business Administration. Small business by the numbers. SBA, Office of Advocacy. http://www.sba.gov/advo/stats/sbfaq.pdf 2003
[16] Servon, L.& Bates T. Microenterprise as an exit route from poverty: recommendations for programs and policy makers. U.S. Census Bureau, Center for Economic Studies, CES 98-17. Accessed through EconPapers. http://econpapers.hhs.se/paper/wopcenses/98-17.htm 1998
[17] U.S. Census Bureau. *Census 2000.* http://www.census.gov. 2002.
[18] Ward, P. *Colonias and public policy in Texas and Mexico: urbanization by stealth*. Austin, TX: University of Texas Press. 1999.
[19] Ward, P., De Souza, F., Giusti, C., Larson J., & May, M. *An evaluation of the Community Resources Group (CRG) colonia lot titling program in Rio Grande City, Starr County, Texas.* Report presented to the CRG. 2003.
[20] Ward, P. De Souza, F., & Giusti, C. (2004). 'Colonia' Land and Housing Market Performance, and the Impact of Lot Title Regularization in Texas. *Urban Studies*, 41 (13), 2621-2646.

Policies and their enforcement in the process towards sustainability

H. J. Holtzhausen
Faculty of Art, Design and Architecture, University of Johannesburg, South Africa

Abstract

The dilemma of sustainability is that it is difficult to enforce, and in general its enforcement is perceived to be accompanied with a reduced quality of lifestyle and profit. The majority of people agree that sustainability is important, hence the development of a number of policies and protocols that advocate a reduction in the exploitation of natural resources. The huge problem faced by these initiatives is that it is not sufficiently enforced and that the main initiators for the environment, the United Nations (UN), are seen to be insufficiently empowered to use coercion.

This paper discusses existing policies such as the Maastricht Treaty, The Brundtland report, The Kyoto Protocol and Agenda 21, to examine strategies to improve the earth's environment. The influence of major agencies that assist with sustainability such as the United States' Environmental Protection Agency (EPA), The World Energy Council and Major Financial institutions are discussed briefly.

The enforcement of environmental issues remains a complex concern that is best achieved through participation and co-operation rather than coercion. The difficulties in enforcing environmental policies are discussed, as are ways of overcoming possible pitfalls.
Keywords: social issues, environmental policies.

1 Introduction

Sustainable development is a complex concept that aims to reconcile two apparent contradictory processes: economic development and environmental protection (Baker [1]). Policies that promote sustainable development will make

www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060081

all stakeholders aware of, and accountable for the environmental costs of their actions. As stated by Baker, this is bound to be deeply unpopular with consumers and producers because of the significant impact it will have on profitability and quality of life. It is the aforementioned fact that makes sustainability hard to enforce, and has provided a lot of summits, advisers, and policies that few are prepared to follow at the risk of disapproval. Most organisations are more prepared to donate money to development initiatives than to work towards their own accountability. The general attitude of 'pay cash and your sins are absolved' prevails in the arena of sustainable development. This paper aims to investigate a variety of policies available, and the way in which these policies can be enforced or implemented. This paper is will not cover the problems of the exponential expansion of the global population, which is the major source of a non-sustainable future, but will instead look at theories on methods of how sustainable development can be achieved.

2 Existing models

Many policies regarding Sustainable Development have been created. The most important and far-reaching of these are the Maastricht Treaty, the Brundtland Report, and Agenda 21. The focus of these policies is to define sustainable development, and devise methods according to which the world can work towards a Sustainable Future.

2.1 The Maastricht Treaty

According to the Center for International Earth Science Information Network (CIESIN) [2] The Maastricht treaty (also known as the Treaty on European Union) is an amendment of the Treaty of Rome, and includes specific references to environmental policy and sustainable development, and was signed in 1992. This treaty in incorporated into the governing of the European Union (EU) and from a variety of source the EU seems to be globally the most progressive in achieving sustainability.

2.2 The Brundtland Report (known as "Our Common Future")

The publication of the Brundtland Report, also known as "Our Common Future" in 1987 under the initiative of the United Nations stimulated a range of actions, including the Earth Summits in 1992 and 2002, and the Agenda 21 programmes. The Brundtland report focuses on strategies to improve short-term human well being without threatening the environment in the long term.

Sustainability defined by the Brundtland report is "the ability of current generations to meet their needs without compromising the ability of future generations to meet their own needs."

The three main components of sustainable development are addressed in the Brundtland Report are social equity, environmental protection, and economic growth.

2.3 Agenda 21 and the Rio Principles

Signed in Rio in 1992, Agenda 21 is a global plan of action to be taken in every area in which human impact on the environment. It is a global partnership with the aim to protect the integrity of the earth, and to recognise the interdependence of existence on earth. The Rio Declaration consists of twenty-seven principles that support sustainable development and the importance of involving all stakeholders (from governments to the children in school) in sustainable development.

At the World Summit on Sustainable Development (WSSD) held in Johannesburg, in 2002, the full implementation of Agenda 21, and the Commitments to the Rio principles was re-affirmed. The list of participating and signatory governments to agenda 21 is not available.

2.4 Kyoto Protocol

The Kyoto Protocol is an agreement between the developed nations to limit their greenhouse gas emissions in comparison with the levels emitted in 1990. According to the Energy Information Administration, the US agreed to reduce their emissions by 7% in the period 2008 to 2012 [4].

3 Agencies for sustainability

3.1 EPA (United States' Environmental Protection Agency)

The EPA [9] is a group of professionals such as engineers, scientists, and policy analysts employed by the president of the United States, working for a cleaner, healthier environment for the American people since the 1970s. Their primary functions involve the development and enforcement of regulations, the performance of environmental research, the sponsoring of voluntary partnerships, and furthering environmental education. The EPA has full law enforcement authority since 1988, and work closely with the other law enforcing authorities of the US to bring environmental criminals to justice. The author suggests that an international EPA is needed to aid global sustainable development.

3.2 Financial institutions

As stressed by Dibble [3], Institutions could do a number of things to promote sustainability. Among these are the promotion of transparency by making information available, promote co-operation between the IMF (International Monetary Fund), the World Bank, the WTO (World Trade Organisation) and the UN.

The World Bank is in a position of power where it can dictate conditions under which it can lend money to governments. Some of its conditions under the "Structural Adjustment Programs" on recipient countries seem to be against sustainability. Downsizing public programs for social programs in developing

countries is not a good approach towards “fighting poverty”. The author believes that the World Bank has a hidden agenda just as the US has a hidden agenda in its “war on terrorism”. It is indeed worrying that one of the most powerful institutions in the world is not geared towards sustainability on earth.

One of the issues that came from the Rio+10 summit in Johannesburg is that there is not sufficient corporate accountability, and that multinationals should improve their practices. This could only be achieved through pledged commitments by the multinationals, and proper follow-up mechanisms.

3.3 World Energy Council

The World Energy Council (WEC) [11] is an independent (non-governmental, non-commercial and non-aligned) charity registered in the UK. It covers all the types of energy available today, including renewables. As stated on their web site: “WEC is well known on the global energy scene for its authoritative reports, analyses, research, case studies, medium and long-term energy projections, and policy and strategy recommendations”. Over ninety of the world’s largest energy consuming and producing countries have member committees of the WEC. In this light, the WEC is a good resource for obtaining support for renewables. The WEC focuses on the following aspects in their work regarding energy policies:

- Market restructuring and liberalisation
- Energy pricing and subsidies
- Sustainable development
- Energy poverty
- Assisting developing countries
- Enforcing and implementing policies

The difficulty of sustainable development lies not in creating policies for its existence any more, but in the execution of the policies. The very nature of Sustainability is extremely unpopular with traditional expansionist economic views, because the principles go against private selfishness whether it is a corporation or a single person. For this reason it is important to find ways of implementing Sustainability.

3.4 Rio+5

In 1997, the Rio+5 global campaign was launched by several organisations that support the principles of sustainability. This was done to translate the Earth Summit Commitments into reality.

Rio+5 identified key issues hindering the implementation of sustainability. People feel inadequate and that an individual can have no impact on the environment, Municipalities are not involved in the decisions that affect them, and local governments expect handouts from federal governments, and feel powerless to influence national policies. The population’s mistrust in local politicians also inhibits meaningful participation.

Issues that drive the implementation of sustainability were perceived as using a Local Agenda 21 as guidance, combined with political willingness for a

successful Local Agenda 21. Rio+5 emphasise the importance of the exchange of information among local governments regarding the implementation of Agenda 21, and the facilitation of civilians to take the initiative in sustainable development, as well as good communication between local authorities and stakeholders. Stress is also placed on allowing for development from local level first, then moving towards national levels, and to acknowledge that poverty issues are related to the environment.

The mechanisms for implementing sustainability identified by Rio+5 are campaigns to educate the public, using peer tutors for specific projects and data collection, and a village-based approach to participation – even in cities. A neighbourhood would be treated as a village where its residents get "empowered". Another important mechanism is for local politicians to work at gaining the trust of people for meaningful participation.

3.5 Rating systems

Rating systems are one of the most effective ways of influencing the consumer into sustainable development. By educating the consumer, a demand is created for more energy-efficient appliances, buildings and vehicles. Positive results have come from rating systems in California. With the Californian federal standard for refrigerators' use of electricity, sales of refrigerators in 1980 were not allowed if they use more than 45 units of electricity. By 1997 the use of electricity by the average refrigerator sold was down to 25 units (Hawken et al. [6]). Developed by members of the U.S. Green Building Council, LEED (Leadership in Energy and Environmental Design) [10] is voluntary, national standard for creating sustainable buildings. The rating system was created to promote integrated building design practices, raise consumer awareness of green building benefits, and to create a common standard of measurement. Australia is also progressive towards Green Building Rating systems, and under the guidance of the Green Building Council of Australia, a number of rating systems are currently in use.

3.6 The Tragedy of the Commons applied to co-operation

In order to develop a sustainable outcome for the future, it is important to get the co-operation of all stakeholders. This in itself provides a social dilemma called The Tragedy of the Commons (Hardin [5]) where a shared finite resource is over-utilised. When a small percentage of users over-utilises the resource, it is to their benefit, and against the common good. When more users over-utilise the resource, it gets destroyed, and nobody benefits from it anymore. Nickerson [8] suggest the best outcome for sustainability is privatisation combined with socialism, where small, manageable groups are held accountable for their own resources. With smaller groups, freeloading is easier to identify.

Nickerson [8] states that there are a number of suggested approaches by about eight different experts. These approaches include governmental initiatives such as laws to regulate commons uses and the enforcement of compliance, the use of education and persuasion, and appeals to peoples' moral convictions and

religious beliefs. Privatisation and the provision of incentives for co-operative behaviour are also suggested. Nickerson further states that laboratory research showed that coercion is more acceptable when there is clear lack of co-operation and the gains from collaboration are high. Offsetting short-term benefits to long-term negative consequences can also modify behaviour. The behaviour's immediate cost can be increased, and the long-term consequences can be made obvious to encourage behaviour with positive long-term consequences.

Nickerson found that the specifics influencing the ease with which a resource is managed are:

- The resource's natural rate of renewal
- The reversibility of the resource's depletion
- The value of the resource to its users.

3.7 Axelrod's 'The Prisoner's Dilemma' and sustainability

As stressed by Dibble [3], co-operative players of the prisoner's dilemma can only expand in a world of cheaters (competitive players) if they get to play against other co-operative players most of the time. When equating co-operative players with those attempting sustainability, and competitive players with those benefiting from co-operation without contributing, the theory brings to mind the success of Eco-villages where the majority of players are perceived to be altruistic, and the community is small enough to identify competitive players. The other way of overcoming the problem of dispersement among competitive players is to be able to accurately identify other co-operative players through signature marks such as clothing, handshakes and customs (Dibble [3]). The importance of reputation also becomes significant, such as among traders on eBay.

Nickerson's [8] comments suggest that an important finding with the prisoner's dilemma situation, is that when it is played continuously with the same players, and one player is playing co-operatively or competitively, it evokes the same response with the other player. These "tit for tat" results are believed to be the origins of co-operative or competitive behaviour.

With the aforementioned theories in mind, it is easy to see that governments and especially the media have a big responsibility to advocate correct behaviour. It is suggested that media watchers be positioned to report on environmentally and behaviourally unsound media projections.

4 Conclusion

There are a number of policies in place to aid sustainable development. The technologies for sustainable use of resources and development exist, and should be used to aid developing countries to achieve sustainability. This is an opportunity for developed countries to prevent the absolute destruction of natural resources that was the result of their development. In spite of the disappointing rate of ratification of the Kyoto Protocol, developed countries do seem to place a big emphasis on environmental issues, as seen with the Environmental Protection

Agency (EPA), and the World Energy Council. The most promising factors coming out of environmental consciousness are the energy ratings policies for appliances and buildings that are created by local governments.

The Rio+5 document identified issues that can aid and hinder the implementation of environmental policies. The best way forward is to achieve sustainability through clear guidelines and policies, and through empowerment and accountability of all stakeholders. According to Hawken et al. [6], the World Bank's 1995 Wealth index found human capital to be three times greater than financial capital reflected on global balance sheets. Humans are the most important feature in sustainable development, and should be sufficiently empowered to create a turn-around in global results towards sustainable development with the aid of clear guidelines and accountability of all stakeholders.

References

[1] Baker, S., The evolution of European Union environmental policy, *The politics of sustainable development: theory, policy and practice within the European Union,* Ed. S. Baker, M. Kousis, D. Richardson and S. Young. Routledge. London, 1997

[2] Center for International Earth Science Information Network (CIESIN), Columbia University. *The Maastricht treaty*, http://www.ciesin.org/TG/PI/TRADE/maastric.html, 1996

[3] Dibble, C., Lectures on Environmental Geography, *Sustainability,* http://www.geog.umd.edu/homepage/courses/cdibble/geog303/lectures/geog303_10_x3.pdf, 2003.

[4] Energy Information Administration, *Summary of the Kyoto Protocol,* http://www.eia.doe.gov/oiaf/kyoto/kyotorpt.html, 2002.

[5] Hardin, G., The tragedy of the commons, *Science*, (162), pp1243-8, 1968.

[6] Hawken, P., Lovins, A.B., & Lovins, H.L. *Natural Capitalism: Creating the Next Industrial Revolution*. Rocky Mountain Institute, Snowmass, CO, 2004.

[7] Korten, D.C., *Sustainability and the Global Economy: Beyond Bretton Woods*. Available from: http://www.pcdf.org/1995/bretton.htm, 1994.

[8] Nickerson, R. S., *Psychology and Environmental Change*. Mahwah: Lawrence Erlbaum Associates, Inc., 2003.

[9] United States Environmental Protection Agency, *About EPA*, http://www.epa.gov/epahome/aboutepa.htm, 2006

[10] US Green Building Council. *LEED (Leadership in Energy and Environmental Design)*, http://www.usgbc.org/leed/leed_main.asp. 2006.

[11] World Energy Council. *About WEC*, http://www.worldenergy.org/wec-geis/wec_info/about_wec/about_wec.asp, 2005.

This page intentionally left blank

Environmental communication strategies: when is what appropriate?

B. Tyson & C. Unson
Connecticut State University, USA

Abstract

How do advocates for the environment best communicate the need to properly manage natural resources? Under conditions of high scientific and high social consensus, a reinforcement strategy would be suitable (i.e., incentives or laws). Under conditions of low scientific consensus and low social consensus, dialogic strategies ought to be followed (i.e., participatory decision making). Under conditions of high scientific consensus but low social consensus, communication could be educational and/or persuasive depending on the immediacy of the situation. This paper provides justification for the above thesis and an analysis of the facilitating factors and barriers affecting each of the four communication strategies listed above. Questions are raised about what constitutes consensus vis-à-vis the controversy inherent among scientists, policy makers, industry, and advocates over many environmental issues. Examples of industry versus advocacy controversy and confusion generated by ideological differences and conflicting scientific findings are discussed. Conclusions are suggested that treat the four communication strategies as phases that lead to support for environmental policy.
Keywords: environmental education, environmental communication, environmental social marketing, environmental behavior change.

1 Introduction

The criteria used to select environmental communication strategies are somewhat unique when compared to other forms of public communication. Environmental issues often generate abnormally high amounts of controversy. Issues are often emotionally charged and the rationale for solutions are frequently subjective. It is often difficult to choose when reinforcement (incentives or laws), persuasive, educational or dialogic (participatory) strategies are the most appropriate.

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060091

According to Focht [1], environmental issues can be assessed in terms of the degree that a) the scientific community and b) the general public agree amongst themselves about the cause/effect and solutions to an issue. He separates environmental issues into one of four categories.

- High scientific and high social consensus (e.g., the need to protect forested areas to control runoff and preserve water quality)
- Low scientific and low social consensus (e.g., ways to control nonpoint source pollution)
- High scientific and low social consensus (e.g., the need to control land fragmentation for protecting wildlife habitat).
- Low scientific but high social consensus (e.g., paper versus plastic as environmentally preferred packaging).

Focht [1] posits that under conditions of high scientific and high social consensus, environmental communication can ethically mandate a reinforcement-based course of action (i.e., incentives or laws). Under conditions of low scientific consensus and low social consensus, he suggests that dialogic communication strategies ought to be followed in which there are few preconceived outcomes (i.e., stakeholders discuss issues until they can agree on a course of action). Under conditions of high scientific consensus but low social consensus, he believes communication should be educational and/or persuasive to encourage action consistent with scientific knowledge. And under the unusual condition of low scientific consensus and high social consensus, any communication that might take place would be primarily ideological.

Focht's thesis is interesting and potentially helpful. Yet, a key question remains unanswered: what constitutes high and low consensus? True consensus will be all but impossible to achieve given the amount of controversy inherent in environmental issues. Much of the public sees environmental and economic issues as polarized. Answers derived from scientific methods of inquiry can be questioned. The validity and reliability of scientific methods are not absolute. Questions are often raised about the basic assumptions scientists make when testing their ideas. Science is susceptible to subjective assumptions, questionable interpretations, and controversy.

2 Environmental communication strategies

2.1 Reinforcement strategies

As Focht [1] states, under conditions of high scientific and high social consensus, it is ethical to mandate a required course of action. Change agents need to impress on individuals that negative consequences will be realized if they do not cease an undesirable behavior. A desirable remedial behavior is elicited and then followed with a reward (e.g., financial subsidy) or an undesirable behavior is exhibited and followed by a punishment (legal penalty). Extrinsic rewards can motivate an individual to perform the behavior repeatedly. Punishments will discourage the performance of a behavior.

Incentive based and regulatory strategies are often the strategy of choice when faced with symptoms of a “social dilemma”. Individuals strive to use their limited time and money in a manner that will enable them to get ahead. But sometimes this self-interest can lead to less than optimum management practices. The same principle can be applied on a larger scale. Companies and municipalities are also driven by self-interest. The deleterious practices by just a few can endanger the natural resources of many. Situations like these can potentially become social dilemmas.

Social dilemmas can be defined by two characteristics: (a) each individual, company or municipality receives a higher payoff for a socially noncooperative choice than for a socially cooperative choice no matter what others do, but (b) all are better off if all cooperate than if all do not (Dawes [2]). Common reasons for noncooperation include (a) the perception that the costs outweigh the benefits, (b) fear of sacrificing options or income by cooperating when few others actually do, and (c) the belief that a critical mass of people are already cooperating and one can "free ride" on these efforts with little notice or effect (Wiener and Doescher [3]).

Incentive based and regulatory strategies are frequently the choice of governmental agencies charged with protecting society against the potentially destructive actions of self-interest. Though these strategies can elicit a fairly rapid change in behavior and are therefore good when problems need a quick fix, the sustainability of incentive based and regulatory strategies is questionable because they are so cost and/or labor intensive. In addition, there is evidence to suggest that reinforcement strategies do not lead to permanent behavior change; i.e., individuals are apt to revert to their original behavior once the reinforcement is removed (Bettinghaus and Cody [4]).

Recently, an article in The Economist [5] suggested that a new evolving form of reinforcement may be the more efficacious route to environmental protection. The article lead with the belief that “today’s environmentalism is just another special interest”… where “mandate, regulate, litigate has been the mantra”… and goes on to say that “if environmental groups continue to reject pragmatic solutions”… “they will lose the battle of ideas”.

The reinforcement strategies that the article refers to employ market-based incentives that may be more sustainable then other types of incentive-based strategies (e.g., subsidies) as long as the correct market forces are in place. Examples include: assignment of property rights over commons such as fisheries, tradable emissions quotas, efforts to value services such as water filtration and flood prevention, or where a water utility might charge more per liter as consumption increases, therefore rewarding conservation behavior and penalizing over-consumption.

There is a current proposal under consideration that suggests that users of the Panama Canal should pay surrounding landowners to reforest the watershed to control siltation and nutrient overload that threatens the ability of commerce to move through the canal. If implemented, the scheme will have environmental, social and economic benefits. Another example using market-based incentives was the decision by the City of New York to pay farmers to protect their

wetlands whose filtering potential was considered a valued ecological service. This decision was significantly less expensive than building a multi-billion dollar filtration plant.

The challenge, according to The Economist [5], is to do good science so that good information can be used to set realistic prices that can lead to realistic cost:benefit analyses. Driving these initiatives is the realization that the environment can no longer be treated as a "free good", that society needs a better understanding of what the environment does for it, and that there is a need to accept that the marginal costs of improvement may not be worth the cost (e.g., the incremental cost of removing the last percentage of a pollutant may not make sense).

2.2 Dialogic strategies

The effectiveness of dialogic strategies rests with the power of social norms that are created through open participatory discussion and serve as a model for individual behavior. This supports Ostrom's [6] finding that successful resource management groups have strong norms that define proper behavior that are reinforced through observable actions. Focht [1] asserts that dialogic strategies are useful when there is low scientific and low social consensus – i.e., when no clear solution is evident and all concerned parties must participate in open discussion to arrive at an acceptable plan of action.

Considerable controversy can be expected under these circumstances. Ironically, economics and ecology, derived from the same root concept implying mutual reliance, are all too often polarized in today's world. Advocates for industry and advocates for the environment frequently find themselves in adversarial positions portrayed as playing the role of defender or aggressor depending on the perspective.

An industry-under-siege/environmental advocate-as-aggressor perspective is reflected in a handbook still considered the primary guide to industrial environmental public relations (Harrison [7]). In this publication, Harrison offers the following comparison of factors affecting environmental communication by industry and environmental advocacy groups.

- Industry favors industrial growth while environmental advocates oppose industrial growth.
- The public has a general mistrust of industry while the public perceives environmental advocates to have high credibility.
- Relationship to government industry is regulated at many levels while environmental advocates are not regulated.
- Industry is not aggressive and seeks thoughtful coverage from the media while environmental advocates are aggressive and seek dramatic coverage.

In contrast, an environment-under-siege/industry-as-aggressor perspective is another viewpoint and one commonly exhibited by environmental advocates

such as participants of a listserv linking leading environmental communication academics in the U.S. For example, an announcement posted on the listserv about an upcoming International Greening of Industry Networking Conference titled *Sustainability: Ways of Knowing/Ways of Acting e*licited the following exchange (COCE [8]):

- Person 1: "Why are we posting this on our listserv? I would hope that we would not lend our efforts to help business and industry communicate their "green image." I think it is relatively safe to assume that the purpose of conferences like this is to help business face environmental challenges more effectively, not develop more "sustainable practices" – or at least this is the role environmental communication would end up playing. Certainly the bridges built there are more likely to help businesses operate more efficiently, on their terms, not gain access for environmental advocates to the decision-making channels of industry".
- Person 2: "There is a lot of greenwashing going on and I disagree with it, but unless we begin to work with industries and businesses that really are trying to green up their systems we are always going to be faced with confrontational issues. As communicators it is up to us to begin the process and begin going to these business conferences to learn how to deal with and understand the "other" side as we begin the process of getting them to become green".
- Person 3: "Let's face it, environmentalism is cluttered with ironies, and, in my presumptuous mind, any discourse that involves the threads of economics and environmentalism, woven and spun together by clout from business and industry, is not only ironic, but I have to presume that the economic thread will win out. Am I to believe that business will cut into its bottom line simply because it wants to be perceived as "sustainable" or that it's "smart business." I'm sorry but I have to be critical of such ironic discourse".

In a recent keynote speech to the Society of Environmental Journalists (Moyer [9]), Bill Moyer, a leading American Journalist stated "our government and corporate elites have turned against America's environmental visionaries". "They have set out to eviscerate just about every significant (environmental) gain of the past generation and while they are at it they have managed to blame the environmental movement itself for the failure of the Green Revolution". He goes on to provide evidence that the Bush administration has staffed key environmental positions with skeptics of environmental science and that industry uses Public Relations strategies to discredit hard science findings. The result, "in July of this year, ABC News reported that 66% of the people in a new survey said they don't think global warming will affect their lives"…and "45% of Americans hold a creational view of the world discounting Darwin's Theory of evolution". "I don't think it is a coincidence that in a nation where nearly half

our people believe in creationism, much of the populace also doubts the certainty of climate change science". As a suggestion to journalists, Moyer states: "I wouldn't give up fact based analysis – the ethical obligation of journalists is to ground what we report in evidence. But I would tell some of my stories with an ear for spiritual language, the language of the parable, for this is the language of faith".

As described, when there is low scientific and low social consensus there will be considerable controversy, and dialogic strategies in which all concerned parties participate in open discussion to arrive at an acceptable plan of action will be necessary. The TAIERI Trust (TT) project, funded by the New Zealand Ministry for the Environment's (MfE) since 2001, is a good example of this strategy in action. The setting is the Taieri River, the third longest river in New Zealand, that travels 318 kilometres and drains 5,650 square kilometres before it enters the sea. The project is an effective vehicle for dealing with growing water quality problems in the Taieri River catchment associated with polluted runoff from farms, septic systems, and urban storm water. There is considerable pubic and scientific uncertainty surrounding these issues.

Project management is by a committee of agency resources managers, community members, university faculty, Iwi (indigenous people) and a full time salaried project coordinator and assistant. Representation from the various geographic areas of the catchment and the varied interests/motivations of committee members ensures that many views are considered.

The management committee's policy of advocating a single position on an issue only when there is total consensus among committee members is important for preserving perceptions of the TT as a neutral body. TT's role in information dissemination, public education and stakeholder communication, allows the project to fill a critical gap in the catchment – a gap that can only be filled effectively by a neutral body such as the TT that has no regulatory function.

A significant number of residents in the catchment believe that the TAIERI Trust has been highly successful in their efforts to improve working relationships among stakeholders. Efforts to establish an information exchange system have included the development of a project website, newsletters, workshops, agriculture show exhibits, and extensive media coverage. A considerable amount of effort was also spent working with primary school students and teachers including the development of a curriculum kit and video on the Taieri River. The effectiveness of these activities is reflected by the sizeable number of people in the catchment that believe the TAIERI Trust has helped raise awareness and understanding of environmental issues in the catchment (Tyson [10]).

Actions for environmental improvements have included prioritization of catchment areas/issues, development of model restoration sites, field days and planting days, and university research on riparian management. Annual reviews of project efforts have been conducted via community surveys and interviews with key stakeholders and results have been widely disseminated. In 2003, the TT project received a special Green Ribbon award from the MfE recognizing the project's national leadership role in Integrated Catchment Management, a strategy premised on dialogic participatory decision making.

2.3 Educational and persuasive strategies

Focht [1] posits that educational and persuasive strategies may be well suited when there is high scientific consensus but low social consensus. Educational methods are designed to promote changes in environmental awareness, knowledge, and skills. According to Archie et al. [11], educational approaches equip audiences with the background needed to make informed decisions about their own choice of behavior. The goal is to build capacity and commitment to engage in problem-solving and decision-making to assure environmental quality. The audience is usually a significant portion of the population (including youth). Outcomes may include environmental sensitivity and changes in knowledge and skills. The time frame is generally long-term because of the emphasis on broad changes across an extensive social framework. A wide range of issues lend themselves to this strategy, particularly those that are not immediate or acute. Communication channels generally rely on print media and written educational supplements.

The other strategy suited to conditions of high scientific consensus and low social consensus can be termed "social marketing". Social marketing, a fitting label for persuasive communication strategies that promote ideas the same way marketers promote products, is a research based, audience focused approach to changing the way people act. Social marketing often starts with educational objectives (e.g., awareness and knowledge of an issue) and once this foundation is laid, shifts to a focus on motivational objectives (e.g., attitude and behavior change) (Tyson [12]).

Audiences need to perceive that the "benefits" associated with a proposed behavior exceed the "costs" if the new behavior is to be adopted. This supports Ostrom's [6] contention that successful resource management groups perceive that the benefits of the resource cannot be discounted and costs of cooperation are low. The challenge is to identify pertinent benefits and costs so rewards can be optimized and barriers minimized.

Additional insight concerning social marketing is offered by Archie et al. [11]. The goal they say is behavior change. The audience is generally a specific target audience that shares common values, access points, or obstacles. Social marketing strategies are good when change is needed in the short-term and are therefore well suited to issues considered acute or critical. Social marketing employs all forms of marketing/advertising tools (i.e., interpersonal, group and mass communication channels).

The Eight Mile River Watershed Project conducted by the University of Connecticut Cooperative Extension System is a good example of a campaign that blended educational and social marketing strategies. The campaign addressed issues facing the conservation of forestland and wildlife in an area recognized by the Ramsar Convention as a wetlands system of international importance (Tyson and Worthley [13]). The campaign dealt with both acute and nonacute issues that generated a lot of varied public opinion. Yet, there was considerable agreement among project technical advisors concerning the required changes and the science supporting these interventions. The objectives of the campaign were for landowners with ten or more acres to do the following:

- Become knowledgeable about the importance of their land management practices in the context of the long-term health of the watershed.
- Become knowledgeable about resource inventory and stewardship planning, and the benefits of those activities.
- Show positive pre- to post-campaign changes in attitudes that are key predictors of stewardship behavior (anticipated personal and community consequences, perceived threats to self and community, and family and community norms).
- Assess conditions and compile an inventory of forest and wildlife resources.
- Formulate specific stewardship goals and develop forest stewardship plans.

Research was initially conducted to define the primary target audience, messages for this audience, and their preferred communication channels and sources of information. Those who showed strong inclination to protect land from development and develop forest and wildlife stewardship plans were chosen as the primary target audience (35% of the population). Findings showed that this group, compared to other segments, believed strongly that watershed resources were indeed at risk. They thought that rivers and streams, trees and plants, and production of forest products were at greatest risk; and they perceived that the top three benefits of forest stewardship planning were preserving natural beauty, insuring that heirs will be able to enjoy the land, and keeping drinking water safe. These factors became the content of campaign messages that were passed through mass channels for achieving educational objectives and personalized channels for achieving motivational objectives. Audio-visual channels were selected for conveying emotions associated with risk and written channels were selected for conveying detailed information. The primary sources of information were state and university specialists who the target audience indicated during initial research were the most credible.

An evaluation of project processes and outcomes was conducted at then end of the five year project that identified which messages and channels worked best and the extent to which project objectives were realized. The project turned out to be moderately successful in increasing landowner knowledge of watershed issues and resource evaluation and stewardship planning strategies. The success the project had in increasing knowledge was due mostly to field demonstrations and tours. The personal contact that landowners had with professional foresters at these events was key to teaching about complex tasks. The project was particularly successful in changing attitudes associated with the impact forest stewardship had on the community. This was an important factor because forest stewardship is inherently about caring for resources that extend beyond individual property boundaries. In the end, the number of completed resource inventories and stewardship plans increased significantly during the project. Nearly half the landowners that were surveyed reported completing some form of assessment and plan during the time of project.

3 Conclusion

Focht's thesis is interesting and potentially helpful. His thoughts on using reinforcement strategies when there is little controversy make sense (i.e., when there is high scientific and public consensus). His thought on using dialogic strategies when there is considerable controversy makes sense too (i.e., when there is low scientific and public consensus). Building on Focht's thesis, it is suggested that environmental advocates treat the four communication strategies that have been discussed as phases that eventually lead to support for environmental policy. If an issue has no evident scientific support, change agents should start with dialogic strategies. Once a critical mass of experts agree on the issues, change agents can pursue education or persuasive strategies. If the issue already has scientific support, change agents can start with an education or persuasive strategy. Once a reasonable degree of scientific and public agreement is achieved, reinforcement strategies can be used to ensure consistency in audience behavior and safeguard against free rider and social loafing tendencies.

Where Focht's thesis becomes questionable is deciding when exactly educational and/or persuasive strategies are called for. These are important and frequently used strategies by environmental advocates. As Archie et al. [11] suggest, educational strategies are suitable when issues are not immediate or acute and when the ability to think critically is the goal. Social marketing strategies are useful when issues are believed to be more acute or critical and targeted behaviour change is the goal. Focht's thesis for when to use educational or persuasive strategies is premised on defining what constitutes high scientific consensus. Yet, as mentioned earlier in this paper, scientific methods of inquiry can generate a lot of controversy.

A good illustration of this was displayed in a National Public Radio broadcast entitled *The Economy and Emissions* (Baron et al. [14]). Three economists from the Economic Strategy Institute (conservative), the Department of Energy (moderate), and Harvard University (liberal) debated the potential impacts of a potential treaty to prevent climate change. When asked about the effects on GDP, responses ranged from 2.5 to 3.0 percent below what it would be without the treaty, to no impact, to a gain of .69 percent. When asked about the effects on unemployment, responses ranged from a 1.8 million job loss, to a "net gain", to an increase of 1.2 percent. When asked about the effects on gasoline prices, responses ranged from 50 cents per gallon, to 6 to 12 cents per gallon, to about five cents. The models these economists use to make their predictions depend on the assumptions that are built into the models in the first place. For instance, regarding the issue of technological change, many studies assume companies will develop new, energy-efficient cars, appliances, and power plants at a steady rate. Other models assume the rate of innovation will accelerate.

A more recent example of how scientific findings can be controversial is evident in bestselling author Michael Crichton's new novel *State of Fear* (Crichton [15]) in which he brings into an imaginary setting the, what he considers, factual ideas that he has shared in congressional testimony and several speeches, including one to the Commonwealth Club in 2003 (Crichton [16]).

Both Crichton's nonfiction and fiction claim that environmental advocates base their opinions on unfounded religious-like myths and beliefs and that global warming concerns are overemotional, unfounded and need to be supported by objective science. In a Hartford Courant editorial (Thorson [17]), Professor of Geography Dr. Robert Thorson criticized Crichton for blurring the link between his fiction and nonfiction and doing what he considered a public disservice by minimizing the threat of global warming in his widely sold novel. Thorson emphasized what a broad consensus of world scientists believe, that high quality science has confirmed beyond any doubt that global warming is a real threat that requires cooperation from all countries to remedy. A recent special report in Rolling Stone states that the novel has been "roundly discredited by the scientific community" and named Crichton one of the top six public "misleaders" in this regard (Little [18]).

Lack of consensus about scientific findings and the resultant controversy that this generates weakens the potential strength of educational and persuasive messages. We must try to minimize the controversy surrounding scientific findings generated by economic versus environmental and industry versus advocacy interests.

The three most important variables that affect the efficacy of persuasive messages are source credibility, message quality and message discrepancy (Hamilton and Thompson [19]). All three need to be optimized to maximize chances of a persuasive strategy working. Source credibility is apt to be a function of two factors: perceived expertise and perceived trustworthiness. Complex issues demand a high level of expertise. Risky issues demand trust.

As stated earlier in this paper, campaign planners seeking to maximize the quality of their messages tailor their messages based on their audience's cost and benefit perceptions. The degree that the quality of these messages can be increased by stressing factual versus emotional elements is a topic of debate for Moyer (who claims environmental messages need more spiritual language) and Crichton (who claims they need less emotion and more fact) (see Moyer [9] and Crichton [16]).

Message discrepancy is the difference between the campaign's position on the issue and the audience's initial position. The audience of a campaign message is more likely to argue against a message and less likely to change when message discrepancy is great. This highlights the need for careful audience analysis early in the campaign planning process so that messages can be designed within an audience's latitude of acceptance. If the campaign is attempting to move the audience great psychological distances, then campaign designers are best to plan change in small minimally discrepant increments; perhaps beginning with educational strategies (informational objectives) and later evolving to social marketing strategies (motivational objectives).

References

[1] Focht, W., A Proposed Model of Environmental Communication Ethics, *National Association of Professional Environmental Communicators Quarterly,* Spring issue, pp.8-9, 1995.

[2] Dawes, R., Social Dilemmas. *Annual Review of Psychology*, 31, pp. 69-93, 1980.
[3] Wiener, J. L. & Doescher, T. A., A Framework for Promoting Cooperation. *Journal of Marketing*, v.55, pp. 38-47, 1991.
[4] Bettinghaus, E.P. &, Cody, M.J., *Persuasive Communication* (5th Edition), Fort Worth: Harcourt Brace College Publishers, 1994.
[5] The Economist, *Rescuing Environmentalism* and *Environmental Economist, Are you being served?* April 23, 2005.
[6] Ostrom, E., G*overning the Commons - The Evolution of Institutions for Collective Action*, Cambridge University Press: New York, 1990.
[7] Harrison, B., *Environmental Communication and Public Relations Handbook, 2nd Edition,* Government Institutes, Inc., Rockville, MD, 1992.
[8] COCE (Conference on Communication and the Environment) listserv, 1997.
[9] Moyer, B., A Question for Journalists: How do We Cover Penguins and Politics of Denial? *Common Dreams Newscenter* (www.commondreams.org), October 1, 2005.
[10] Tyson, C.B., *Evaluation of the TAIERI Trust Project*. Research report for The Taieri Trust. Department of Geography, University of Otago, Dunedin, New Zealand and Department of Communication, Central Connecticut State University, 2004.
[11] Archie, M., Mann L., & Smith, W., *Environmental Social Marketing and Environmental Education,* Academy for Educational Development, Washington, D.C., 1993.
[12] Tyson, C. B., *Strategic Environmental Communication: Communicating Strategies for Influencing Environmental Behaviors.* Xanedu Publishers, 2003.
[13] Tyson, C. B., & Worthley, T. E., Promoting Basic Forest Stewardship, A Model for Watershed Management. *Journal of Forestry*, Vol. 99, No.8, August, 2001.
[14] Baron, D., Siegel, R. & Wertheimer, L., Economy and Emissions, *National Public Radio, Washington, D.C., 1996.*
[15] Crichton, M., *State of Fear,* Harper Collins, 2004.
[16] Crichton, M., Remarks to the Commonwealth Club, September 15, 2003.
[17] Thorson, R.M., *The Fearful Practice of Treating Global Warming as Fiction*. Hartford Courant Editorial. January 14, 2005.
[18] Little, A. G., The Misleaders. *Rolling Stone,* v.987:83 November, 2005.
[19] Hamilton, M. & Thompson, W., Testing an Information Processing Model of Message Intensity Effects, *World Communication,* v.23, pp 1-14, 1994.

This page intentionally left blank

E-waste transboundary movement violating environmental justice

J.-h. Kim
Department of Public Administration, Seo-Kyeong University, South Korea

Abstract

This study examined the problems in the transboundary movement of e-waste with respect to the process of the importation and treatment of e-waste (electronic waste), from the point of view of environmental justice. If the relationships between e-waste-exporting countries and -importing countries is considered in terms of their economic interests, various positive factors surface. If the problem is considered from the standpoint of environmental justice, however, both countries are guilty of complicity. In the short-term economic interests, if these are understood in terms of economic justice, e-waste importation is still not recommended; and if it is considered in the context of environmental justice, it is clearly an international criminal act. In this context, this study also suggested why this problem requires another policy, in response to the WTO system that controls international trade.
Keywords: electronic waste, hazardous wastes, environmental justice, human right, global environmental governance, basel convention, WTO.

1 Introduction

Recently, the UN Commission on Human Rights addressed the links between the environment and human rights, by concluding that everyone has the right to live in a world free from toxic pollution and environmental degradation. Due to economic reasons, however, many small and significant cases violating such primary human rights occur. Constructing or operating environmentally hazardous sources or facilities in economically poor areas is an environmental injustice and cannot simply depend on the market mechanism. This is because the act of discriminately violating man's primary environmental right can never be accepted by reason of unequal economic status.

www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060101

2 Environmental justice versus economic injustice

2.1 Environmental justice

Environmental problems between humans and nature can be understood in the same context as other problems of human society, such as inequality, racial discrimination, and sexual discrimination [1]. As with other social problems, solutions to environmental problems in the context of justice could be proposed. In this context, environmental justice could be understood as a basic human right, or in other words, calling for fair treatment and meaningful involvement [5].

2.1.1 Fair treatment

The object of fair treatment can be seen from the personal point of view. This means my social and economic group, to which I belong, such as my race and my nation, should not be treated inappropriately in any occasion. Likewise, this norm may be taken to mean that I should not experience discriminatory treatment with respect to environmental human rights due to the characteristics of my private position. It may have come to mean that it is unjust to undermine environmental rights due to race, nationality and other standards of one's socio-economic group and position.

2.1.2 Meaningful involvement

Meaningful participation may be approached from the 'relative' point of view. This means relationships with other parties should be open so that the parties' mutual interests will be appropriately satisfied. In other words, benefits and losses of one party must be compared with those of the other party, and both parties must participate in decision-making to maximize their benefits and minimize their losses. This is to say that, to realize environmental justice, environmental information in mutual relationships should be open, the interested parties must fairly and substantially give inputs to the decision-making process to come up with relevant environmental regulations, and the decision-makers should be required to monitor the impacts of their decisions.

2.1.3 Environmental justice as human right

Therefore, environmental justice can be regarded as a goal of the entire human races. Environmental rights should be secured without any form of discrimination based on people's nationality, race or socio-economic level. Moreover, people should be able to access information when their environmental rights are in danger of attack by other parties, and there must be a system that would prepare them to make a significant response against the threat.

2.2 Economic injustice

The ideal scenario of environmental justice painted above is in crisis, however, due to the today's economic injustice. For example, consider cases involving hazardous waste. According to the United Nations Environment Programme

(UNEP), about 0.3 to 0.5 billion tons of hazardous wastes are produced throughout the world every year and about 80-90% of these are produced in industrial countries [4]. Under these circumstances, the Basel Convention widely limited the free trade of hazardous waste through the Basel Convention on the Control of Transboundary Movements of Hazardous Waste and Their Disposal, and clearly stated that rich countries have a responsibility not to externalize their pollution treatment cost to poor foreign countries through imports. Due to the economic benefits of the transboundary movement of hazardous waste to both developed and developing countries, however, it persists to this day and is in fact becoming more serious. This scenario is analyzed below from the point of view of economic injustice.

2.2.1 Injustice of developed countries

Most developing countries support the restriction of transboundary movement of hazardous waste based on the Basel Convention, but some developed countries still oppose its ratification. For example, the U.S. signed the Basel Convention in 1989 but has not ratified it yet. This is because if a developed country exports its hazardous waste to developing countries, it can sidestep its strict domestic regulations and standards, which will lead to huge savings on disposal costs of hazardous substances and increased economic benefits. At the very reason, however, transboundary movement of hazardous waste is not actively regulated on the grounds of ‘recyclability,’ although this still reeks of economic injustice.

2.2.2 Injustice of developing countries

In the transboundary trade of hazardous waste, the economic benefits of developing countries cannot be denied as well. In the process of recycling the hazardous waste as resources, even in a poor condition, many jobs can be created for simple labourers who have been unemployed, and other economic benefits can be reaped. In particular, developing countries that lack resources to manufacture products are strongly tempted to take such direct and indirect economic benefits. In the same manner that there are many reasons that make it difficult to forbid international trading of illegal drugs, there are also, for developing countries, two sides to the thorough restriction of imports of hazardous waste. Although the importation of hazardous waste is legally forbidden, it is difficult to expect execution of the law considering the blind points in the system, namely, the acceptance of bribes by the officers who are supposed to control such trade, and the economic benefits derived by the nation. Therefore, the temptations posed by the short-term economic benefits to developing countries of the importation of hazardous waste result in economic injustice, which undermines environmental justice even as it obstructs long-term development and brings with it environmental damages and dangers to people’s health.

2.3 Injustice of e-waste trade

In effect, the economic injustice of the transboundary movement of hazardous waste is the result of the improvement of the total amount of mutual economic

benefits between the developed country and the developing country, which are difficult to separate between the assailant and the victim. In particular, if there is a way for a developed country to export its harmful waste to a developing country, the developed country will issue a disincentive to the developing country it will effect as a "disincentive" which discourages the developed country to invest in new technology to reduce such hazardous waste or a recycling technology, so that it would have to produce more hazardous substances. From the point of view of environmental justice, however, the transboundary movement of hazardous waste is the violation of the environmental rights of developing countries to maximize the benefits of developed countries. Even domestically, it could be argued that the economically strong violate the environmental justice of the economically weak. Therefore, this study aims to consider the transboundary movement of e-waste from the point of view of environmental justice, and, in particular, to further study the current status of imports and treatment of e-waste in China from the same standpoint.

3 Current imports and treatment status of e-waste in China

The problem of importation and treatment of e-waste in China has already been pointed out by Green Peace, a representative international environmental NGO, and BAN (Basel Action Network) with respect to the Guiyu area in Guangzhou Province, as earlier mentioned. Not only CCTV in China but other media outfits all over the world have started to publicize the seriousness of the problem. This study focuses, however, on the fact that this problem is not that of a limited area in Guiyu in southern China, but is already a critical problem that widely occurs in various regions in Asia. From an in-depth field research in Taizhou, China which is fast growing as an e-waste treatment hub from its previous virtual anonymity, the issues discussed further below have arisen.

This survey team was accompanied by the Green Peace activist who raised the problem. Although there was a possibility that problems would arise in the team's search for accurate spots in the field and in its interviews with the workers, which could have raised sensitive issues, the team was able to finish the survey within the limited six days from October 14, 2004 it had, with the help of a local high school teacher and some students.

3.1 Introduction to Taizhou, China

Taizhou, a small port city on the coast of the China Sea five hours away from Shanghai, has rapidly developed, in the last 10 years, into an industrial city amidst rural communities. The rapid growth of this city was due to the presence therein of many recycling companies utilizing waste resources imported from all over the world, which formed a huge recycling industry cluster that is expected to expand even more in the future. This area already has a serious air and water pollution problem, though. Moreover, at night-time and on weekends, when government officials do not work, large volumes of waste are illegally incinerated. In many areas of the city, residuals of illegal incinerations and

landfills can be found. Taizhou is a very active region, however, not only in its metropolitan area but also in its farming area. Many of its residents work hard to make a living, and many people move to the city from inland China to work in Taizhou's labour-intensive recycling industry.

3.2 Current status of e-waste in Taizhou

3.2.1 Amount of e-waste imports

Since the importation of waste resources started in the ports of Taizhou in 1992, Taizhou's volume of imports has drastically increased from less than 10,000 tons to 860,000 tons in 2001, 980,000 tons in 2002, and 1,350,000 tons in 2003. The said volume is expected to grow further to 1,500,000 tons at the end of 2004. According to relevant statistics pertaining to Taizhou for 2003, of the total volume of its imports through its ports, approximately 85% is known to account for imports of waste resources. It can be said, therefore, that the recycling of waste resources in Taizhou is becoming its leading infrastructure industry. Therefore, although it is difficult to confirm Taizhou's actual volume of e-waste imports, it can be assumed that a considerable volume of e-waste would be included in its annually imported 1,500,000 tons of waste resources. Assuming that the ratio of e-waste to the total volume of Taizhou's waste resource imports is about 4.5% based on the findings from the inspection conducted by Taizhou's customs authorities, it is possible to estimate that Taizhou imported around 67,500 tons of e-waste in 2004.

3.2.2 Scale of revenues from e-waste and regulatory difficulties

It is normally estimated that about 0.1 g of gold is extracted from 1 kg of waste electronic goods. Considering that Taizhou imports about 67,500 tons of e-waste a year, the economic value of the gold that could be extracted from such e-waste would be around 6 billion Yuan in Chinese currency. Among other major valuable metals extracted from waste, the combined economic value of palladium (1 g = 60 Yuan) and silver (1 g = 20 Yuan) reaches up to about 10 billion Yuan, or an estimated 1.3 billion US dollar. This huge scale of revenues has become the fundamental ground that makes actual control ineffective, although the Chinese government has pronounced it as illegal. Moreover, since most e-waste is illegal imports, it is impossible to make appropriate investments for their proper treatment, so that they are eliminated through illegal incineration and illegal landfills. On the whole, this illegality has resulted in the more serious release of hazardous substances and worse air pollution not only in places in China but also in the neighbouring countries, as it is carried in the air and covers it with hazardous substances. Further, it adds to the environmental pollution cost of all Asian societies.

4 Critical factors in the movement of e-waste

4.1 Crises in the e-waste-exporting countries (developed regions)

The e-waste exporting country can misconceive that it is reaping economic benefits with the increase in the profits of relevant exporters within the country.

The exporting country should not overlook, however, possible crises it could experience in terms of the following three aspects.

4.1.1 Obstruction in the development of environmental technology

Export of e-waste significantly obstructs the development of relevant technologies in industrial countries. In other words, it makes it difficult for environmentally friendly production technologies to be developed, which weakens industry productivity and obstructs its consistent growth. In particular, as developed countries transfer their treatment of wastes to developing countries as waste exports, the appropriate treatment regulation standards become helpless (i.e., loopholes concerning the regulations increase), making environmentally friendly companies bankrupt and possibly resulting in obstructing further environmental technology development or investments.

4.1.2 Reduction of domestic jobs

Exporting e-waste to developing countries instead of treating them means taking jobs away from the domestic low-wage working class. Since most e-waste disassembly processes are labour-intensive, they usually pay relatively low wages. Therefore, the export of e-waste to developing countries ultimately results in the closing of workplaces in waste treatment companies and recycling companies, which employ many domestic workers. Although revenues would increase from exports of e-waste, the serial bankruptcy of recycling companies will result in a more vicious cycle of negative socio-economic repercussions.

4.1.3 Spread of environmental pollution and the boomerang effect

In the process of developed countries exporting e-waste to other countries, they could misjudge that their environment has become better since they have removed e-waste from their country. As their neighbouring countries, however, which imported their e-waste, treat the e-waste using very dangerous methods under poor conditions, secondary- and third-level pollution arises in series. Since such series of pollutions may contaminate the e-waste-exporting countries' skies and seas, export of e-waste can bring about the boomerang effect of environmental pollution. For example, as in the case above, the e-waste flown in to the agricultural communities of China cause serious water pollution as they are reprocessed as large volumes of hydrochloric acid and chemical substances, which will contaminate coastal areas and pollute the ocean, contaminating the fish on Japanese dinner tables with heavy metals and exposing the Japanese to Itai-Itai disease. The diverse dangers that e-waste-exporting countries may face are summarized in Table 1 below.

4.2 Crises in e-waste-importing countries (developing regions)

On the part of e-waste-importing countries, they could misapprehend that e-waste importation would benefit them in terms of huge economic benefits, more jobs, and improved supply and demand of raw materials. These countries should also bear in mind, though, that they will also suffer very difficult losses.

Table 1: Costs of e-waste exporting countries.

Obstruction of the development of environmental technologies	With the increase in loopholes in the appropriate treatment regulations due to waste exports, disincentives to technology development and investment increase.
	Weakening of industrial productivity due to the obsolescence of environmentally friendly production technologies
	Despite the increased use of hazardous substances, relevant regulations and policies would fail.
Shrinkage of domestic recycling businesses	Increase in unemployment in recycling companies and of low-wage workers
	Increase in waste factors of resources
	Harm to infrastructure for the construction of the base for a sustainable zero-waste society
Spread of environmental pollution and boomerang effect of such pollution	Increase and spread of air pollution
	Harm to health due to the spread of various hazardous substances
	Rapid increase in respiratory diseases such as asthma in neighbouring e-waste-exporting countries such as Japan

4.2.1 Problems with local residents' health

E-waste-importing companies cannot process e-waste under good conditions and through proper processes. As they extract valuable substances under poor conditions and produce large amounts of hazardous substances in the process, not only the workers in their area but also the local residents face the risk of health problems due to serious environmental pollution. Diseases would occur due to land and underground water pollution, and food shortage would get worse due to the shrinkage of agricultural activities. For example, according to the results of the survey conducted and presented by Shantou University Medical School in Shantou, Guangzhou Province, China in February 2004, sample deposits collected in the Guiyu region, where the e-waste industry started in 1990, earlier than in Taizhou, showed heavy metal pollution 1,338 times higher than chrome standards and 212 times higher than lead standards.

Huge amounts of carcinogenic substances have already contaminated the water and land in those areas, and directly affected some 311 people with e-waste-industry-related jobs, 34.1% of them now with diseases in the nervous system, 24.9% with digestion problems, and 15.7% with respiratory diseases. The ratio of people experiencing headaches and dizziness has already exceed 35.7% (data from Green Peace).

Table 2: Costs of e-waste-importing countries.

Problems with the local residents' health due to environmental pollution	Occurrence of diseases due to land pollution and underground water contamination
	- More serious food problems due to the shrinkage of agricultural activities - Lack of drinking water
	Occurrence of diseases caused by environmental pollution such as Itai-Itai disease
Natural disasters due to irrecoverable environmental conditions	Expansion of barren areas where people cannot live
	Geometrically progressive increase in future environmental restoration costs due to environmental pollution
	Possibility of unpredictable aftermath of pollution from hazardous substances
Social unrest	Resistance of the local community to serious environmental pollution
	Tightening of external international environmental regulations such as trade restrictions
	Large-scale unemployment and occurrence of various political instabilities due to the sudden environmental regulatory measures

4.2.2 Natural disaster due to irrecoverable status of natural environment restoration

The mortal pollution in the natural environment due to e-waste impoverishes the neighbouring areas within 10 years and makes ecological restoration impossible. Concerning the Guiyu region, where e-waste imports began earlier than in Taizhou, there is already a report that there are areas therein that have become barren, i.e., where people can no longer reside due to serious environmental pollution [3]. If restoration of the polluted area is attempted in the future, the restoration cost would reach astronomical figures; and if restoration is not undertaken, an unpredictable aftermath could occur due to the constant outflow of hazardous substances.

4.2.3 Social unrest

If this level of environmental pollution continuously occurs due to imports of e-waste, it will give rise to the resistance of local residents and to international pressure for environmental regulation by neighbouring companies. Although these kinds of dangerous factors were but anticipated in the researcher's field survey in October 2004, they have already come true only six months later in April 10, 2005, when a large-scale riot occurred in Huasi village in Dungyang city near the Taizhou region, involving 30,000 local residents who opposed the increase of environmentally polluting factories in the village and 3,000

policemen who attempted to control them (Chosun Daily Newspaper, April 13, 2005). Moreover, in Zhansha village near Taizhou, a fierce demonstration and collision were reported in June 2005, which were related to the issue of environmental pollution (Kookmin Daily Newspaper, July 4, 2005).

5 Conclusion

This study examined the problems in the transboundary movement of e-waste with respect to the process of the importation and treatment of e-waste, especially focusing on Taizhou, Zhejiang Province, China, from the point of view of environmental justice and economic injustice. If the relationships between e-waste-exporting countries and -importing countries in considered in terms of their economic interests, various positive factors can be surfaced. If the problem is considered from the standpoint of environmental justice, however, both countries are guilty of complicity.

On the short-term economic interests, if these are understood in terms of economic justice, e-waste importation is still not recommended; and if it is considered in the context of environmental justice, it is clearly an international criminal act. This is also why this problem requires another policy, in response to the WTO system that controls international trade. The WEO (World Environmental Organization), which leads in raising public awareness of international environmental matters on behalf of entire societies worldwide, in response to the private principle of maximizing economic benefits, can push for Global Environment Governance to realize environmental justice in response to economic injustice, and this is expected to become a major issue in relevant domestic and international environmental policies in the future.

Therefore, future Global Environmental Governance have to be undertaken to restrain the economic benefits maximization theory in the world trade order that stimulates globalization and openness. Such policy would make it possible to expect mandating of active regulations in some areas where mutual economic interests of developed and developing countries are met. The interests-oriented relationship of e-waste-exporting and -importing countries could be beneficial to both in the short term. Since it could impose significant costs on them all in the long term, however, the environmental regulations of each country on this matter should be approached not in the context of domestic private interests but from the standpoint of primary human rights; and if these regulations are violated, such violations should be treated as international criminal acts.

All humans should be fairly treated regardless of their nationality or socio-economic group. To ensure meaningful response to the relevant environmental pollution problem, each government should approach and design their environmental policies and systems in a new light. The starting point in overcoming this world's environmental crisis is the renewed protection of environmental rights, which are primary human rights regardless of nationalities and socio-economic status.

References

[1] Han, Myeon-hee, *A Theory of Green Civilization*, Dongnyeok Press: Seoul. 2004.
[2] Kim, Jung-hoon, Sustainable Urban Waste Management System in Metropolitan Seoul. *The Sustainable City III*. WIT Press: UK. 2004.
[3] Lee, Sherry, Ghosts in the Machines in South China Morning Magazine. May 12, 2002: HK.
[4] McGinn, Anne P., Reducing Our Toxic Burden, *State of the World*. WWI. Norton Co.: NY. 2002.
[5] Rawls, John, *A Theory of Justice*. Revised edition. Harvard University Press: MA. 1999.
[6] Tieenberg, T., *Environmental and Natural Resources Economics*. Addison Wesley Co.: NY. 2002.

Time series analysis of the ISEW at the local (regional) level: method and intent

F. Ciampalini, R. Ridolfi, F. M. Pulselli & N. Marchettini
Department of Chemical and Biosystems Sciences, University of Siena, Siena, Italy

Abstract

This paper presents a time series analysis of the Index of Sustainable Economic Welfare (ISEW). It is a measure, alternative to GDP, that in accordance with the principles of sustainability takes into consideration environmental and social aspects within the traditional economic accounting system. The Index is calculated for an Italian Province from 1971 to 2003 and it will be compared with the provincial GDP during the same period. A growing gap between the two monetary values (GPI grows faster then ISEW) has been occurring consistently with the national trend. This means that part of the wealth which GDP asserts does not correspond to welfare. The ISEW monetizes this gap in order to orient decisions for a more sustainable way of life and more correct policies.
Keywords: GDP, ISEW, welfare, sustainability, ecological economics.

1 Introduction

This paper represents one of the many examples, in this case an ecological economics initiative, where it's possible to find a political planning where economy and ecology speak the same language and rely on the research and university for building their future society in which the unique development to wish is the sustainable one. Ecological Economics offers this potential because it is built around a more complex understanding of human society that transcends the traditional economic vision [1].

The economic system is an open subsystem of a larger but finite and non-growing ecosystem; so no infinite economic growth is possible in a finite world. Society is nowadays moving from a relatively empty world to a relatively full world, and an exclusive emphasis on economic growth could produce serious, and possibly irreversible, ecological damages.

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060111

Now, it is easy to understand the necessity of a shift from the growth model to the development one. The process of economic development could improve people's standard of living not only in monetary terms but also in terms of well-being (in the broader sense of quality of life). In brief, GDP is not more the right index to study and plan long term policies, because it tells us only part of the story. On the contrary, it is not possible to have a sustainable development without transparency. GDP is implicitly or explicitly used as a measure of wealth and welfare, of development and progress, of good policy and management. Actually, it was not introduced by Simon Kuznets for measuring welfare.

Its formula (GDP = C+I+G+NX) represents the value of all final goods and services that are treated via markets. For our purposes, some weaknesses of GDP can be put on evidence: for example, it does not reflect the value of no market goods and services which are, instead, an important part of our welfare; it grows when a forest is cut; it would be higher if a person dies in a hospital than if s/he dies at home; car accidents increase GDP; it does not take into consideration household labor. GDP just adds up goods and services produced by the economic system and bought in the market without making a distinction if those market transactions contribute positively or negatively to human welfare.

In 1989, Daly and Cobb proposed the Index of Sustainable Economic Welfare (ISEW) as an alternative measure to the conventional reliance on GDP. It is an indicator of economic welfare, and represents an attempt to measure the underlying economic, social and environmental factors that create real progress. The ISEW represents an important index of underlying long term trends in real welfare. It measures and makes understandable something that is considered important [2] for human well-being and it informs policy-makers and the general public about the factors that affect welfare [3].

Creating awareness at the local level is important. The local (regional or provincial) level often represents the optimal dimension for implementing policies, especially in certain fields. This is mainly due to a major attention paid by local institutions for peculiar problems and to the administrative decentralization [4]. For this reasons, the ISEW has been calculated at the local level, in order to give local authorities a further instrument for taking political decisions consistently with sustainable viewpoint.

This paper shows some results from a time series analysis of ISEW (1971-2003) for the Province of Modena, an Italian province. This is one of the first attempts to perform such a calculation in time series at the local level.

2 Methods and results

The ISEW is a measure of economic "performance", alternative to GDP, with the purpose of assessing the level of sustainable welfare of a population living in an area. It takes into account the traditional dimensions of sustainable development, embodying the social, economic and environmental components of well-being. The ISEW is based upon a portion of economic wealth (GDP), relevant to welfare, but various monetary adjustments (additions and subtractions) are made in order to take into account environmental and social items, not currently considered in the GDP framework.

Table 1: ISEW components and reasons for inclusion/exclusion.

	Sign	**ISEW components**	**Reasons for inclusion/exclusion**
A		Year	1971-2003
B		Personal Consumption Expenditure	The basic variable directly affecting economic welfare
C		Index of inequality distribution	Connected to social aspects, it is necessary to have a more realistic starting point
D		Weighted Personal Consumption	The basis to which all other positive and negative modifications are applied
E	+	Services of households' labour	An activity outside the market process that however leads to improvements in welfare
F	+	Consumer durables services	Services connected with goods are benefits
G	+	Services from Public Infrastructure	In general people use infrastructures without paying any costs but in reality lots of money are spent for citizens' welfare
H	+	Public expenditure on health and education	It is only a portion, the non-defensive one, of public spending that is expected to add to welfare
I	-	Expenditures on consumer durables	They must be subtracted because they are not strictly connected to welfare, but only the services from them
L	-	Defensive private expenditures on health and education	The total value is already present in B or D. Here the defensive portion, not affecting welfare, is removed
M	-	Local advertising expenditure	This component has been removed from this ISEW exercise because of the lack data at local level
N	-	Cost of commuting	It is included in the consumption but it is a defensive cost related to traffic and overcrowding
O	-	Cost of urbanization	Urbanization implies the increase in the price of houses. This increase in wealth is fictitious then it must not be considered.
P	-	Cost of car accidents	Defensive costs, not related to welfare, that have to be removed from consumption
Q, R, S	-	Cost of water pollution Cost of air pollution Cost of noise pollution	These costs represent the contribution a society must pay to solve environmental problems. They are not related to welfare because defensive.
T	-	Loss of wetlands	Wetlands are important natural habitats, whose protection is one of the keys of sustainability
U	-	Loss of agricultural land	Agricultural land is threatened by urbanization and soil erosion, involving costs that must take into consideration
V	-	Depletion of non renewable resources	The consumption of non renewable natural capital, undermine the possibility for future generation to meet their needs
W	-	Long-term environmental damage	Long term damages caused by greenhouse gas emissions.
X	+	Net capital growth	To ensure an economic sustainability, the level of capital per worker must be maintained.
Y		ISEW	The final value to be compared with the GDP trend

The ISEW is roughly defined by the following formula:

ISEW = personal adjusted consumption + non defensive public expenditures – defensive private expenditures + capital formation + services from domestic labour – cost of environmental degradation, pollution and damage – cost of depletion of nonrenewable resources

The main works that have been taken as reference points are the ISEW calculations for USA [2], for Italy [5] and for the local level [4].

The following paragraphs will explain each component of the ISEW as shown in Table 1, and their characteristics and trends. Furthermore, some variations with respect to the examples in literature are described too. They are mainly due to the differences between the calculation at the national and the local level. Finally, the resulting trend of the ISEW with respect to GDP is shown.

2.1 Personal consumption expenditure

The unique data source available for all the time series analysis was from ISTAT (the Italian institute of statistical analysis). It is a document called "Consumption for Italian families" where it is possible to find the consumption for family o for family's component at regional level, and other several data as shown afterwards. This figure has increased by 54% from 1971 to 2003, with an increasing of population nearly 17%.

2.2 Index of inequality distribution

The Gini coefficient has been used here. It is a measure of inequality of a distribution, defined as the ratio of the area between the Lorenz curve of the distribution and the curve of the uniform distribution. It is often used as measure of income inequality. Its value can vary from 0 to 1, where 0 means everyone has the same income, so called "perfect equality"; 1 means that one person owns all the income, that is "perfect inequality". The Gini trend for the Province of Modena shows an improvement in the equal distribution of income, jumping from 0.656 in the 1971 to 0.314 in the 2003.

2.3 Weighted personal consumption

This adjustment is obtained by dividing the consumption expenditure (item B in table 1) by 1 + Gini coefficient (Fig.1).

2.4 Services of households' labour

Following the methodology used for Italy [5], three classes of people have been considered: housewives, unemployed persons and students spend time in housework of 8, 4 and 2 hours/day, respectively, for a total of 330 days/year. ISTAT publishes annual data on working and non-working population, and the price per hour is based on previous studies [4]. Services of households' labour decrease from more than 2,730 millions €/2003 in 1971 to 1,381 millions €/2003 in 2003.

2.5 Services from consumer durables

A positive element of the ISEW are the services from consumer durables (durable goods, cars and houses). The calculation method is the same of Daly and Cobb [6] with a little modification in the percentage assigned to the value of services from a house, which is 5% instead of 10%.

The final value shows a shift from a nearly thousand million of euros/2003 in 1971 to more than 5 thousand millions of euro/2003 in 2003.

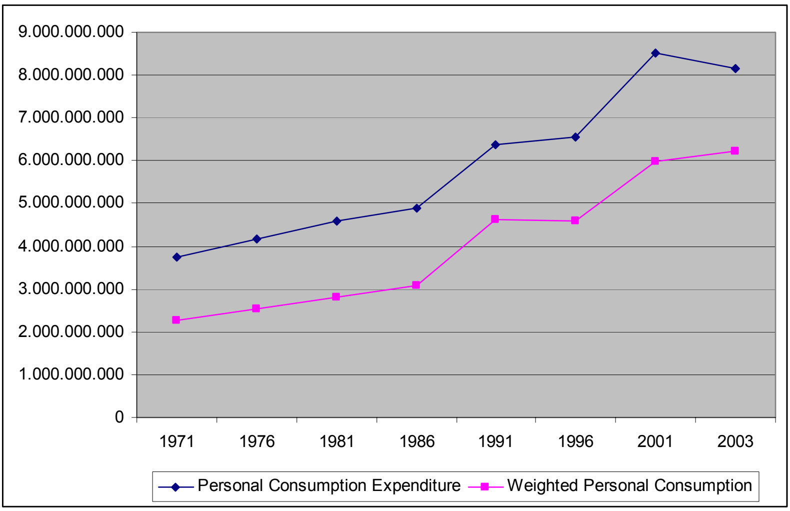


Figure 1: Time series of adjusted consumption for the Province of Modena.

2.6 Services from public infrastructure

Daly and Cobb [6] generally exclude government expenditure on public services because they are considered as defensive, with few exceptions. Bad quality roads decrease our welfare for many reasons, such as high maintenance expenditure, more probabilities to have a car accident, etc. Generally, people use the services from the road network without paying, and, in order to embody it in the welfare calculation, this portion of public expenditure must be considered.

This value is not relevant, but it increases from 1,483 thousand of euros/2003 in 1971 to a value of more than 10 millions of euros in 2003.

2.7 Public expenditure on health and education

According to the methodology by Guenno and Tiezzi [5], public expenditure on education are entirely considered as not defensive, whereas the 50% public expenditure on health are considered as defensive. Hence the 100% public expenditure on education and the 50% of public expenditure on health must to be added to Personal Consumption.

This portion of public expenditure was around € 822 millions (€/2003) in 1971, and around € 973 millions in 2003.

2.8 Expenditures on consumer durables

The consumption expenditure on durable goods (televisions, personal computers, cars) does not represent the welfare derived from their use. These goods provide utility (and thus welfare) along their life time. This aspect must be counted in the ISEW, so this figure must be subtracted from consumption (note that the services from consumer durables – item F in tab.1 or Par. 2.5 – are added to consumption). These expenditures were around € 343 millions in 1971 (€/2003) to reach a thousand million euros in 2003.

2.9 Defensive private expenditures on health and education

Consumption also includes private expenditures on health and education. According to Daly's work [3], only the 50% of the total is related to welfare; the rest is defensive (sustained by families not for improving welfare but for restoring or compensating the welfare lost due to modern life-style such as pollution, damages, competitiveness, etc.).

Defensive expenditures on health exceed those for education since 1991, reaching in 2003 a gap of nearly 80%. The total value increases a lot during time, from 88 millions (€/2003) in 1971 to 217 millions € in 2003.

2.10 Local advertising expenditure

No data on local advertising costs were available, this item was thus omitted.

2.11 Cost of commuting

The formula used to value this item is a further development of the ones used by Guenno end Tiezzi [5] and Pulselli et al. [4].

The formula is:

$$C= 0.3A+0.3B+0.3C_1$$

where
0.3 is the estimated portion of the total car use directly related to commuting;
A is the value of services derived from a car and not its total value;
B is the expenditure for tickets on public transportation;
C_1 is the sum of costs for public and private vehicle maintenance.

The cost of commuting trend increases from €116 millions (€/2003) in 1971 to nearly €440 millions € in 2003. This is strictly dependent on the excessive use of car instead of other means of transport, especially in city centers. However, in this case, the value is not so high thanks to the fact that most of the territory is situated on the plain and the use of bicycle is encouraged.

2.12 Cost of urbanization

This cost was not considered because, despite buying a house at prices always rising does not produce a correspondent increase in economic welfare, it gives a high level of satisfaction and an equal or greater monetary feedback at the moment of its selling is often assured.

2.13 Cost of car accidents

This cost is a part of consumption that is naturally defensive and should be subtracted from the index. The cost of an accident is composed by two factors:

the personal injury costs which have already been counted among "Defensive private expenditures on health and education" and the non-injury costs. The latter is calculated here, by multiplying the number of car accidents per year by the social costs of an accident.

The trend of this cost is fluctuating and linked to the problem of urban mobility. It is about € 7 millions in 1971 (€/2003), up to 2001-2003 when it reaches about € 13 millions (€/2003).

2.14 Cost of water pollution

This is an attempt of including environmental aspects within the ISEW framework. In literature, there are several examples to calculate this cost. In this paper, the E.I. (Equivalent Inhabitants) method has been used. The E.I. of the Province of Modena are multiplied by the cost for purifying. Data source for the cost is [4]; E.I. are estimated on the basis of the number of inhabitants, industrial workers and animal productions. The value fluctuates around 70 millions (€/2003), reaching its maximum value during the '80s with 88 millions (€/2003).

2.15 Cost of air pollution

The estimation for this environmental cost is the same of [5] and [4]. Major pollutants are taken into consideration such as CO_2, NO_X, SO_X and PM10 and then associated to the abatement cost per ton of emission, taken from [4].
For this figure it seems to be a decrease of all pollutants except CO_2.

SO_X have been decreasing since 1980 because of planning policies adopted by the local government, which decided to move to low environmental impact fuels like methane and discouraged the use of carbon and fuel oil. The cost of air pollution was around € 78 millions (€/2003) in 1971. The highest value was in 1991 (€ 132 millions), while in 2003 it was 90 millions (€/2003).

2.16 Cost of noise pollution

There is not so much information to evaluate this item. People are affected from noise in different ways, and there are problems in measuring noise levels on local/national level. The most important source of noise is urban traffic, then the final value was obtained from the relationship between the total number of cars on the provincial territory and an estimation of the value of noise damage produced by a car. The trend of this figure, being strictly connected with territorial car pool, is increasing. It was 18 millions in 1971 (€/2003) and more than 53 millions (€/2003) in 2003.

2.17 Loss of wetlands

Wetlands are considered to be among the most threatened ecosystems because of their susceptibility to both human and natural stressors. Their importance was highlighted by Daly when he decided to insert them within the ISEW calculation. The variation in wetlands in the Province of Modena was assigned a monetary value according to the paper by Guenno and Tiezzi [5].

The results are positive, meaning that for the Province of Modena there was an increasing in wetlands from 500 hectares in 1971 to 1,470 hectares in 2003. The equivalent monetary value was 6,386 in 1971 (€/2003) and 119,766 in 2003.

2.18 Loss of agricultural land

The economic growth has generated social costs in terms of the loss of farmlands. This adjustment reflects the human colonization of land (road construction or urbanization) and the deterioration of soil quality and erosion (intensive agriculture). The value of this item increases along all the period, from more than 12 millions (€/2003) to around 23 millions in 2003. This growing scarcity is due to a bad management of natural capital and it's over exploitation.

2.19 Depletion of non renewable resources

El Serafy's formula [8] has been used for the ISEW of the Province of Modena:

$$R - X = R[1/(1+r)^{n+1}]$$

where:
X annual income;
R revenue from extraction net of extraction costs;
r discount rate;
n residual life-time of the stock of resources.

The value of this figure is limited to quarry materials and data necessary for finding the final value is related to: type of materials, quantity extracted per year, market prices and extraction costs.

This variable, so important under a sustainability viewpoint, is not so relevant in monetary terms: it was 67 millions in 1971 (€/2003) and 112 millions in 2003. This result is because since the '70s the Province of Modena have been paying attention to the mining industry, aware of all the problems and damages produced by this kind of activity on the surrounding environment.

2.20 Long-term environmental damage

The long-term environmental damage is related to the greenhouse gas emissions. We used the calculation method proposed by Daly for USA [6] and we consider all the consumption of fossil fuels, methane and electricity. They are converted in equivalent barrel of oil and multiplied by a tax of 1.9 euro/2003 for each barrel used [7].

This item gives an idea, in monetary terms, of the damage we are producing that will be paid by future generation. For the Province of Modena we calculated a kind of environmental tax of 20 euro/2003 per-capita in 1971, and 50 euro/2003 per-capita in 2003. The monetary value of this item increases from 124 millions in 1971 (€/2003) to more than 324 millions in 2003.

2.21 Net capital growth

To ensure an economic sustainability, the level of capital per worker must be maintained. Due to the lack of data it was not possible to use the formula from [5]. Hence, in order to estimate trend, gross fixed investments were used weighted by employed persons in the industrial sector. The trend appears to be

constant for all the period analyzed, around €1,400 millions (€/2003) with the lowest value in 1986 of about € 996 millions (€/2003).

2.22 ISEW vs. GDP

A large number of adjustments to consumption base have been done in order to estimate the ISEW. This result can help to understand the criticisms about GDP dealing with its inability to distinguish among transactions which increase or decrease welfare. GDP has to be used only as expression of total business volume, which sums costs and expenditures instead subtracting them. On the contrary, the ISEW methodology aims to make a distinction between positive and negative economic operations. The ISEW, as many said, is not an indicator that shows how far a system is from a sustainable condition, it orients society towards the right direction. Furthermore, the ISEW reflects the influence of policies designed to affect social progress, economic growth, environmental protection and prudent use of natural resources. In doing so, it allows to present a systematic assessment of domestic progress towards sustainable development over long period of time, and to compare this against GDP. Figure 2 shows the comparison between ISEW and GDP for the Province of Modena.

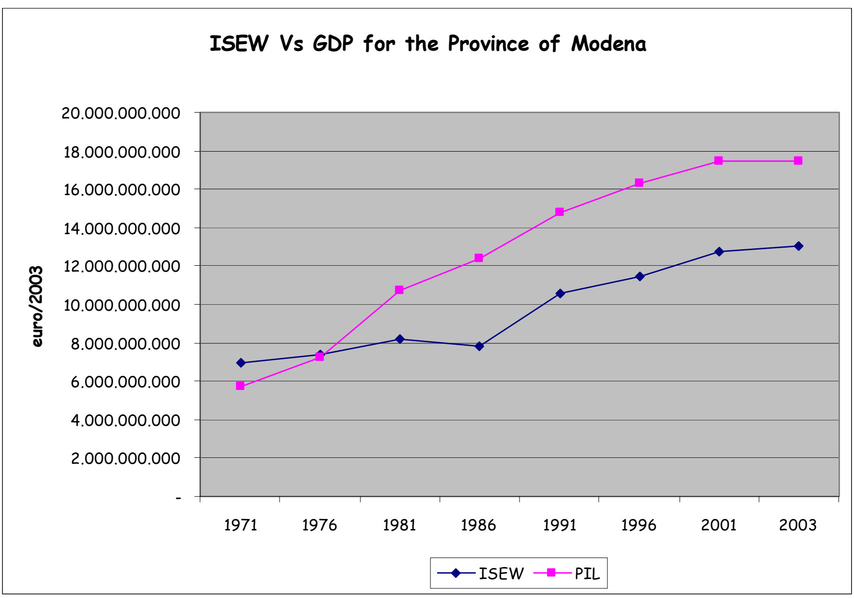


Figure 2: The trend of ISEW and GDP for the Province of Modena.

3 Conclusion

The ISEW offers some advantages with respect to other aggregate sustainability indicators:

- it tries to include social and environmental aspects of sustainable development

- unlike more usual measures of progress, it considers the distribution of resource in a society, and takes into account the treatment of natural capital
- it tries to estimate the net benefits of a society which produces and consumes, showing business people and local leaders the impacts of their activities
- ISEW reflects the influence of policies designed to affect social progress, economic growth, environmental protection and prudent use of natural resources. In doing so, it enables policy and decision makers to present a systematic assessment of domestic progress towards sustainable development over a long period of time, and compares this to GDP.

This paper, from a project in collaboration with the Province of Modena, is an experiment oriented to assess the sustainability at the local level. The assessment of the sustainability of economic and social activities at the local level is certainly a great challenge for local authorities. Their power to take decisions, though regulated by laws, should be put into action on the basis of accurate and comprehensive information. Any decision, plan or project to be implemented should be evaluated both in economic terms and considering social and environmental aspects.

References

[1] Kysar D.A., *Sustainability, distribution and the macroeconomic analysis of Law*. Boston College Law review, Vol XLIII No.1, pp. 1-71, December 2001.

[2] *Sustainability Indicators for waste, energy and travel for Scotland*, Scottish Executive Central Research Unit 2001, Chapter two.

[3] Daly H.E., Cobb, J.B. *For the common good: redirecting the economy towards community, the environment, and a sustainable future*. Beacon Press, Boston. 482 pp., 1989.

[4] Pulselli et al., *The Index of Sustainable Economic Welfare (ISEW) for a Local Authority: A Case Study in Italy.* Ecological Economics, in press, 2005.

[5] Guenno G., Tiezzi, S., *The Index of Sustainable Economic Welfare (ISEW) for Italy*. Worknote 5.98. Fondazione Enrico Mattei, Milano, Italy, 1998.

[6] Daly H.E., Cobb J.B., *Un'economia per il bene comune. Il nuovo paradigma economico orientato verso la comunità, l'ambiente e un futuro particolarmente* sostenibile. RED edizioni, Como, Italy, 1994.

[7] Anielski, M., Rowe, J., *The Genuine Progress Indicator - 1998 update.* Redefining Progress. San Francisco, CA, 1999.

[8] El Serafy, S., *The proper calculation of income from depletable natural resources.* In: Ahmad, Y.J., Serafy, S.E., Lutz, E., (Eds) Environmental Accounting for Sustainable Development. An UNDP-World Bank Symposium, World Bank, Washington, DC. 10-18, 1988.

Sustainability and urban regeneration: the community and the city

E. Douvlou[1] & A. Ryder[2]
[1]*School of Architecture and School of Environmental Design and Management, University of Portsmouth, UK*
[2]*Department of Geography, University of Portsmouth, UK*

Abstract

Successful sustainability requires a re-evaluation of social paradigms. Portsea, in Portsmouth UK, is located in a prime location. It is adjacent to the city's Historic Dockyard, the University, a shopping centre and expensive private sector housing, but it remains socially and physically isolated from the city as a whole. Extensive analysis has shown that, despite apparent pathologies, there is a strong sense of community, as residents take pride in and identify with their neighbourhood. The prospect of additional development and redevelopment in the area suggests that it might lead to another case of gentrification. However, our work in Portsea suggests that apparently deprived districts may be far more "sustainable" than is thought. Policy makers need to tap into this community spirit, to activate it, and to enhance it. It is hoped that this work can contribute to the debate related to social perspectives in historical districts among policy makers, experts, communities, non-governmental organizations, industry, and the general public.
Keywords: community, city, development, sustainability, regeneration.

1 Introduction

In urban regeneration, sustainability is increasingly defined in social as well as environmental terms. However, practitioners and academics appear to distinguish just two types of ideal communities: those, which are successful, despite poor physical surroundings, and those, which are failing, often despite an above average physical environment.

doi:10.2495/RAV060121

Since the 1960s, the ideal of the successful community in a poor physical environment has been popularised by Jacobs [1] and Gans [2] in the US and Wilmott and Young [3] in the UK. They wrote about "Urban Villages" and the social capital embedded in existing communities, arguing that although the physical environment might be poor, community structures were often robust and worth retaining. This contrasted with the then traditional view that poor environments created poor communities. More recently, architects and urban designers such as Oscar Newman [4] have suggested that communities which suffer from social pathologies may be victims of poor design. They have promoted the re-design of neighbourhoods and urban districts to promote "defensible space", the growth of community-based management and control, and reterritorialising urban space.

At the same time, a large body of literature suggests that social integration is a cure to some community ills, calling for the construction of new, market rate housing to alter the social mix of an area. A contrasting view holds that some communities, measured in terms of crime, social pathologies, and indices of deprivation, are best broken up or diluted. However, it is not always the case that high levels of crime and poor social indicators mean that a community has broken down.

We show this by assessing the results of a questionnaire and associated research addressed at the residents of Portsea, a socially isolated community in Portsmouth, UK. Although urban growth and change is altering some of the socio-economic characteristics of neighbouring areas, within Portsea there is strong support for maintaining the social and physical fabric of the district, and local residents' views of what makes a good community appear to be substantially different from those of local government decision makers.

2 Portsea: physical qualities and social characteristics

Portsea has changed significantly since it first became a dockyard in 1194, on the orders of King Richard. It was, and still is, home to the Royal Navy and was once home to heavy industry such as shipbuilding, a busy dockyard and a brewery. Until recently, the navy controlled until recently much of the land in the area. What was left was used for high-density terraced housing which housed the low paid industrial workforce, which worked in the adjacent naval dockyards. As recently as the 1940s, thousands of employees worked in the yards, and many of them lived nearby. Due to heavy bombardment in WW2, many of the early and historic buildings were damaged and then demolished during the post war period.

The area was rebuilt in the 50's and 60's, borrowing heavily on the ideals of the Bauhaus and the "Garden City" ideal. However, policies of the time aimed to retain the working class character of the area. Reconstruction led to a "mish-mash" of architectural styles, textures, and materials. Virtually all the new housing was "council", that is, publicly owned housing. Moreover, redevelopment significantly altered the physical layout of the community. Terrace houses with individual gardens (the traditional form of English housing)

were replaced by flats surrounded by extensive public open space. Traditional streetscapes were destroyed, and buildings set back from roads. Densities were substantially reduced, and industry was banished from the area. Today, the majority of buildings in Portsea are residential, mainly consisting of social housing. Throughout the 1960s and 1970s industry and employment in the dockyards was drastically reduced. More recently redevelopment has occurred in on old naval property nearby. Gunwharf Quays is a new development of housing, shops and entertainment facilities.

Figure 1: Portsea in a postcard from 1906.

Figure 2: Portsea as it is today viewed from the Spinnaker tower.

Figure 3: A typical pub in the Hard.

Figure 4: A tower block.

Figure 5: Typical council housing in Portsea.

Figure 6: A model of the Portsea area. The darker blocks denote housing.

Figure 7: A poster of the admiralty-quarter development visualizing how it will look in a few months when the housing is completed.

Portsea is adjacent to the University of Portsmouth, major tourist attractions such as the Historic Dockyard, a new shopping centre and expensive private sector housing. However, it remains socially and physically isolated from the city as a whole. It has one of the highest levels of deprivation in the United Kingdom Indices of Deprivation [5] and National Statistics, Crime Levels [6], despite being central to Portsmouth, and crime levels are several times the city average. Unemployment is quite high and the residential values of flats are less than one-third of the houses and flats in the adjacent new development of Gunwharf Quays, which has created 2,000 jobs. Recently, a new community centre with contemporary design and a wide choice of leisure facilities as well as a small library and a nursery has fulfilled an important role in the social rejuvenation of the area. However, this is still a relatively deprived area, characterized by a high number of people living in social housing, high unemployment rates, and low incomes. The original local employer, the Dockyards, which dominated community life and determined the social structure of the neighbourhood has vanished, undermining much of the rationale behind the original reconstruction of the area, and leaving behind a concentration of low skilled, low waged or unemployed residents. Moreover, he process of

reconstruction destroyed much of the physical fabric of what was one of the oldest settled areas in Portsmouth. Segregation of the community may lead to social exclusion of existing residents as redevelopment in the area takes place, including luxury apartments with more affluent people moving in, while not providing inputs to the wider local community.

The current research was motivated by a desire to assess the views and opinions of the local population about their community as well as how they perceive the current changes in the area and what that might mean for the future.

3 The survey

To gain an understanding of community spirit in the area, we undertook a survey and meetings with several focus groups. As well, we “surveyed” the Portsea area through the eyes of 16 final year undergraduate students taking a course on sustainable design and environmental management. Copies of the survey were left in the leisure centre, and residents were sampled on the streets and open spaces of the neighbourhood as well. The combination of questionnaire and focus group was decided on after the trials of the questionnaire prompted a fair amount of discussion on the limitations of a simple question-answer format. It was felt that the range of likely opinions could not be covered within the limits of a simple question-answer format.

The questionnaire was administered to over one hundred local residents in the area of Portsea, with equal representation from both genders. The questionnaire comprised 23 closed questions divided in eight thematic areas assessing local peoples’ opinion on transport, education and leisure facilities, public spaces, safety, as well as their views on the current development in the area and their perception on how their community will be affected in the future. A five-point scale answer was employed ranging from “strongly agree” to “strongly disagree” and from “excellent” to “very poor”.

The majority of the interviewees were in the 21-35 age group. A large percentage of the participants were unemployed and many had low skill jobs. The majority had a low level of qualifications and as was indicated in the questionnaire, most have struggled to find a job in the area: 22 % of the people interviewed were unemployed or retired.

The questionnaire included several questions designed to assess resident satisfaction with the area. Residents were asked to assess both the safety and the friendliness of the area. Interestingly, they viewed Portsea as friendly, but also as unsafe (Figure 8). In fact, it has one of the highest crime rates in the city of Portsmouth particularly regarding list crimes.

Moreover, despite the view that Portsea is friendly, only 31 % of respondents felt that Portsea was better than other areas in the Portsmouth (Figure 9).

The questionnaire indicated that overall, people are satisfied with schools and shopping facilities in the area but think that public transport and leisure facilities could be better (although a new community centre was just about to open as the questionnaire was distributed).

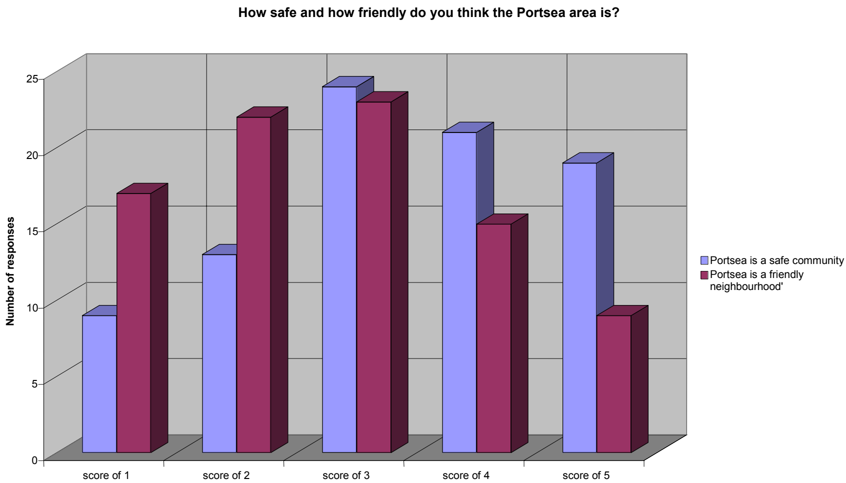


Figure 8: Residents' view on how safe and how friendly the Portsea are is.

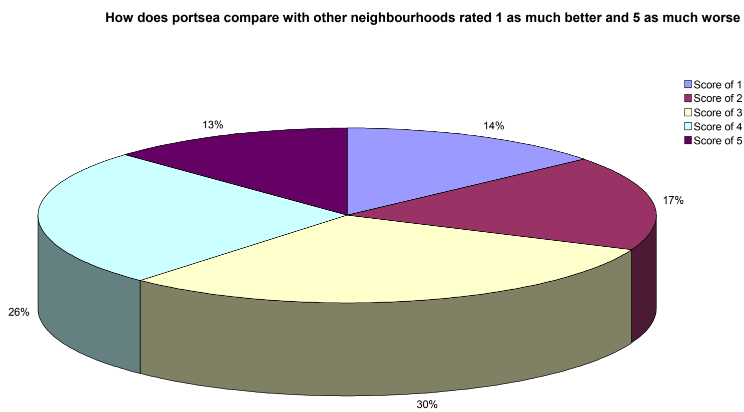


Figure 9: A comparison between Portsea and other neighbourhoods.

It also appears that the built environment has significantly affected the way residents view the neighbourhood, and also significantly affect the way public spaces are used. Figure 10 shows the results of how often people use the open spaces/play areas in Portsea and how do the rate them. Most people only use the open/public spaces occasionally, and view them as being unsatisfactory. The general consensus from the focus groups was that the spaces were adequate. It is apparent from the responses that there is a correlation between the negative opinion and the low rate of usage. Dissatisfaction was attributed to several causes. Almost all respondents thought that most spaces were not designed to welcome people spending time in them. Families and children did not feel safe. There were no places to sit and watch children, the elderly felt excluded due to an absence of benches, it was difficult if not impossible to supervise children from a flat, and there were no designated areas for smaller children to play.

Inadequate lighting was also a factor, followed by the lack of presence of some kind of canopy to protect from weather elements and the lack of a coffee/snack bar.

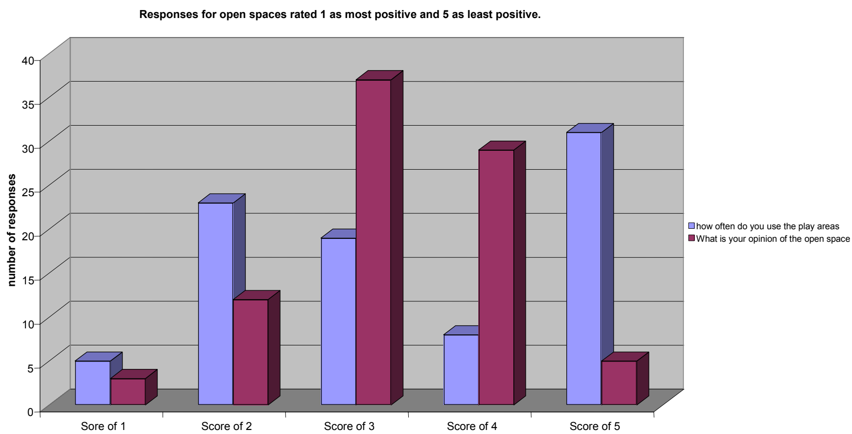


Figure 10: How residents rated the public/open spaces available and the frequency they use them.

The new transformation of an old naval dock and helicopter landing pad into the "trendy" Gunwharf Quays, which comprises a shopping outlet, a range of bars and restaurants and other leisure facilities, and expensive (for the area) accommodation triggered a lot of discussion among respondents. Most people had a very strong opinion about what this new development "should" have contributed to their area and what the reality was (Figure 11). Although the development has generated over 2,000 permanent jobs, it was a common view that developments such as the above should have created more job opportunities for the area and aid in regenerating the whole of Portsea. Going by questionnaire responses and focus group discussions, it seems however, that most residents of Portsea feel that it has only helped to emphasise social exclusion. Many argued that although jobs were created, they were not accessible to local residents, who were not qualified enough (although this may be more myth than reality, since many of the jobs are entry level posts. However, it is very likely that the competition for employment is high, since the area has a large number of full and part time students.)

On the other hand, the new community centre, completed a few months ago symbolises local government commitment to the area. It houses a variety of leisure facilities as well as a library and a nursery. It provides a new and welcomed meeting point for the community and is part of a larger redevelopment project for the area. It is however, quite different from the development in Gunwharf Quays. It is directly related to the community and was designed with the community's input. More importantly, most interviewees talked with enthusiasm about the project and how more similar projects could really help regenerate the area, suggesting that they would welcome a chance to have a greater say in the organization of their physical environment.

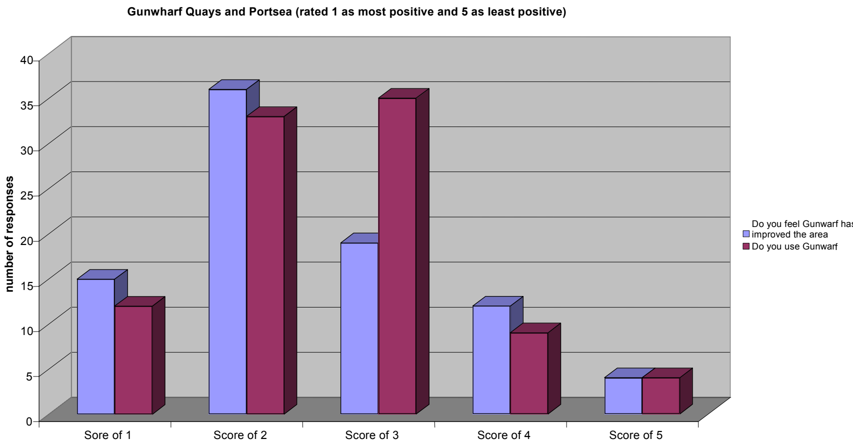


Figure 11: Portsea residents' views on Gunwharf Quays.

4 Discussion

The analysis of the questionnaire indicates that the community of Portsea is satisfied with their neighbourhood, but realizes that there is potential to improve in many areas. In particular, there is concern over the crime rate, and over the environmental quality of the area, particularly as it relates to the physical environment. For example, while Portsea is fortunate enough to have several small parks dotted throughout the area, the perception is that they are not used for safety reasons. From the assessment of the built environment and the results of the focus group it is apparent that respondents do not always find the area to be aesthetically pleasant. Green areas and open spaces are not inadequate but it was widely agreed that lack street furniture and lighting compromise their use by all ages and at all time. Moreover, a large number of derelict and uninhabited buildings do not help improve a built environment, which is landscaped by council terraced houses, and tower blocks.

The focus groups explored in greater detail responses to the built environment, in particular how residents viewed and used public open spaces. This turned out to be an important issue for local residents, generating a great deal of comment and discussion. It became apparent that local residents people felt relatively powerless to influence or control the arrangement of public space. They complained about the lack of gardens, the lack of street furniture, the difficulty of supervising pay areas, and poor lighting. Interestingly, despite the area's high crime rate, vandalism is virtually non-existent, and there is little graffiti on inhabited buildings.

Residents expressed not just frustration poor quality of the physical environment, but also at their lack of control in the management process. Extremely low levels of home ownership and the high level of council housing means that most residents have very little control over the area in which they live. They control neither their personal living space nor the broader public

spaces, which make up “their” neighbourhood. Housing is managed by the Portsmouth City Council’s Housing Directorate, and open space is managed by the Parks Directorate or even by the Education Directorate.

A key finding was that although residents were unhappy with the physical environment, they were *not* unhappy with the built environment. Although there was not a question asking about housing satisfaction on the questionnaire, it was not an issue in the focus groups. As noted above, graffiti is almost non-existent in the area. A generation of residents have grown up in dwellings, which are now considered socially unhealthy. Although resistance to living in flats was widespread after 1945, residents of Portsea now take this kind of living for granted. Their taste in housing, their views of what constitutes a “normal” residential environment, have been shaped by their housing experience. However, it was evident from the responses of our students that housing in the area no longer conforms to best practice paradigms, and even among city council planners there are still suggestions that the neighbourhood should be substantially rebuilt. (For an extensive discussion of the broader implications of these changing paradigms, see Douglas [7] or Muthesius’ comprehensive discussion of the utopian design aims of new universities In the 1960s and 1970s and the response of users to those utopian goals (Muthesius [8]).

Residents are twice removed from the management of the area in which they live and in which they are legitimate stakeholders. Because they rent rather than own their housing, they have little say or control over it’s appearance or management policies: they are supplicants rather than participants in the management process, which is top down, paternalistic, and often condescending in its well meant intentions. They are additionally removed from control because there is little private space in the area. There are relatively few private gardens, and the traditional streetscape does not exist in most of the area. They have no control over building colours, gardens, landscaping, plantings, or street furniture. Again, they are clients rather than decision makers, and decisions are made by an often remote and benevolent city council bureaucracy. This stands out in contrast to neighbourhoods in most British towns and cities, where local residents influence streetscape through their choice of house colours, plantings, and other forms of personal expression (Muthesius [9]). Residents cannot even park their cars outside their flats. Thus, while the area suffers from high social exclusion, it also suffers from a high level of social disengagement. One could argue that residents can express their views through municipal elections. However, their “electoral ward” has three council members who belong to a minority party on the city council, and have little voice in shaping policy. Moreover, in Portsmouth as in most of the UK, respect for local government is limited, and turnout in local elections is low.

At the same time, the views of our students are instructive. They saw the neighbourhood as amorphous, unfriendly, and uncomfortable; open, chaotic, and unlit. They did not like visiting the area, which reflected both their personal living experiences and the things they had heard about the neighbourhood while living in Portsmouth, despite the fact that they work alongside it, and often pass through it on their way to and from University. There are no landmarks, there

are few old buildings, and there is no sense of “place.” Moreover, despite the strong local identity of residents, there is no sense of a district or neighbourhood in the built environment. Portsmouth City Council has invested large sums of money in creating neighbourhood identities in other parts of the city through the use of distinctive streetlights and paving. The area is characterized by what is it not rather than what it is. There are few shops, few pubs, and no restaurants. Despite extensive open space, there are no seating areas. This is not a place to sit and relax or enjoy a meal or a drink. There are no through streets or main pathways with the exception of Queen Street, which far from being a “path” is a heavily trafficked artery which divides the neighbourhood in two and at first glance appears to be lined with derelict buildings or open space.

5 Conclusions

Portsea was first established as a dockyard and is still the home of the Royal Navy. The Portsea area gradually grew up around the dockyard to provide housing, entertainment and other facilities for the dockyard personnel. Today, Portsea is considered amongst the most deprived areas in the south of England although it is adjacent to historical attractions, a university and new shopping and leisure facilities. However, extensive analysis has shown that despite apparent pathologies, there is a strong sense of community. Residents take pride in and identify with their neighbourhood.

Urban regeneration should embrace successful social sustainability. This requires a re-evaluation of social paradigms. Our work in Portsea suggests that apparently deprived districts may be far more “sustainable” than is thought. System built housing, and utopian houses in the sky surrounded by public open space are not always a sign of social decay. Moreover, high levels of crime and poor social indicators do not mean that neighbourhoods lack social capital or “friendliness”. It would be a mistake to destroy the social fabric of the community while redeveloping the built environment. Policy makers need to tap into the community spirit, to activate it, and to enhance it. It is hoped that this work can contribute to the debate related to social perspectives in historical districts among policy makers, experts, communities, non-governmental organizations, industry, and the general public.

Today, Portsea is at a crossroads. The construction of a new shopping district and residential complex on former naval land adjacent to the area, and pressure to construct new market rate housing to meet local demand means that the area could be substantially rebuilt over the next decade. A system-built tower block is approaching the end of its design life span, and its demolition will almost certainly lead to other changes. Already, new market rate housing is being built for sale or rent at far higher prices than local residents can afford. The centrality of the area, its access to good rail and bus services, it closeness to the university, which is a major employer and, thanks to students and staff, a major source of housing demand, all place pressure on the area to change.

However, the city council “owns” the area, and can legitimately claim the right to manage and redevelop Portsea with the needs of the broader civic

community in mind, particularly given its accessibility and its relatively low density, and particularly since the number of people who pass through or near to it is immense, and each visitor is a form of “stake-holder” as well. Moreover, since the area was rebuilt in the 1950s and 1960s, planning and urban design paradigms have changed, and moved away from the Bauhaus ideal. The “city in the park” has been replaced by a desire to recreate traditional busy streets and streetscapes, characterized by a mix of uses and a mix of housing types and social classes. This is perhaps best epitomized by the 1987 Berlin International Building Exhibition (IBA), which can be contrasted with the 1957 IBA. (See, for example, the discussion about the reconstruction of Marzahn, in Berlin, and the influence of the 1987 IBA in Young, [7].) These changes are also reflected in the Portsea area. In the 1980s, new hosing construction moved away from flats to terraced houses. However, new construction followed a suburban pattern of clustered housing, surrounded by extensive lawns, and densities remained low. More recently, in the early 2000s, new housing has been built which follows the street line, and abuts right against the pavement.

While there is little doubt that the area must change, and it has already started doing so, we argue that whatever physical changes take place; the social fabric should be preserved. We recognize that in a dynamic and rapidly changing world, the concept of stakeholder is a complex one, and admit that it is difficult to determine who “owns” a neighbourhood. While existing residents may view Portsea as “their” neighbourhood, new residents can make a similar claim, as can those who work nearby, or live there as students. Therefore, perhaps more importantly, we argue that social engagement is an important part of any regeneration process, and should precede any such process. In fact, we argue that social engagement is as important a goal as social inclusion, and that it should be a central part of any reconstruction programme.

References

[1] Jacobs, Jane (1961) The Death and Life of Great American Cities

[2] Gans, Herbert (1962) The Urban Villagers, Free Press

[3] Wilmott and Young (1957) Family and Kinship in East London

[4] Newman, Oscar (1973) Defensible Space London: Architectural Press

[5] Indices of Deprivation (2004), Department for Communities and Local Governmenthttp://www.odpm.gov.uk/

[6] National Statistics, Crime Neighbourhood Statistics (2004) Charles Dickens Ward http://neighbourhood.statistics.gov.uk/dissemination

[7] Young, Douglas (2005) “Still Divided? Considering the Future of Berlin Marzahn” Berichte zxur deutsche Landeskunde v. 79, no. 437-456.

[8] Muthesius, Stefan (2000) The Postwar University New Haven/London: Yale University Press.

[9] Muthesius, Stefan (1982) The English Terrace House New Haven/London: Yale University Press

Legal and health aspects of sanitary waste management in Italy

M. Di Basilio & M. Papacchini
ISPESL – National Institute for Occupational Safety and Prevention, DIPIA – Department for Production Plants and Interaction with the Environment, Italy

Abstract

Sanitary wastes represent an important health aspect of extreme complexity on account of their heterogeneity, their quantitative and typological growing trend and of their simultaneous belonging to different hazardous classifications. The purpose of this study is to review the occupational exposure and risks for workers involved in the sanitary waste collection and management industry. Sanitary Wastes (S.W.) can be categorized as infectious or non-infectious. Workers of the sector considered are exposed to the most common etiological agents such as the Hepatitis A, Hepatitis B, Hepatitis C and HIV viruses, bacterial infections caused by the *Bacillus, Staphylococcus* and *Mycobacterium* varieties and fungal infections from the *Aspergillus* variety, mainly due to accidental events or faulty application of safety procedures. The nature of the relevant biological agent depends on the processing unit in question. Referring also to current Italian legislation and to risks connected to S.W. management, this study will consider tools to put in practice a rigorous hygienic management of S.W., in order to minimize all risks connected with their production, collection and disposal. An important sector of this study will be dedicated to S.W. classification, differentiating infectious from non-infectious waste, in order to choose the most acceptable treatment method for the various types of waste, reducing risks for workers due to the presence of non-infectious hazardous substances. All actions taken in order to protect persons handling S.W. will be analysed for each phase of their packaging, handling, transportation and storage.
Keywords: sanitary wastes, management, professional exposure, Italian legislation.

www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060131

1 Introduction

Various aspects contribute to the great interest aroused by sanitary wastes. One of these concerns health and hygiene. Other aspects regard the management and heterogeneous characteristics of waste, the growth trend in type and quantity of wastes and the fact that they can belong contemporaneously to more than one risk category. According to the World Health Organisation, potentially infectious sanitary wastes represent a risk for the environment, the health of patients, health facility staff and public health in general (on account of risks connected with transportation and final disposal).

2 Aim

The aim of this study is to analyse occupational exposure and risks in spinnaret workers during sanitary waste collection and management. These wastes can be classified basically as infectious and non-infectious. Infectious wastes usually include those of human, animal or of biological origin generally, and all elements that may have been contaminated by pathogens. Non-infectious wastes include toxic chemical substances (solvents, oils, disinfectants, etc.) and cytotoxic substances (medicines). Numerous data on exposure are available for the latter sector.

Workers in the sector examined in this study are exposed to etiological agents such as: hepatitis A, hepatitis B, hepatitis C and HIV virus, bacterial infections caused by the *Bacillus, Staphylococcus* and *Mycobacterium* varieties, and fungal infections from the *Aspergillus* variety. In most cases occupational exposure to these agents occurs accidentally or is due to incorrect application of safety procedures.

By examining existing Italian legislation in this sector and the risks involved in sanitary waste management, this study analyses regulations currently in place for stringent hygienic management of sanitary wastes with a view to reducing the risks involved in each of the required phases.

An important part of the study is devoted to the classification of sanitary wastes where a clear distinction is made between infectious and non-infectious wastes, since this subdivision is important in identifying adequate methods of treating different types of waste and of reducing risk.

3 The size of the problem

In recent years, the problem of wastes, especially sanitary wastes (SW), has taken on considerable importance due to a continual increase in the use of throwaway materials, to the growing presence of dangerous or potentially hazardous substances and to their infectious nature. Even though infectious SW represent only a small proportion of the overall volume of wastes produced, they are nevertheless a problem on account of their special characteristics and the lack of facilities for adequately treating this type of waste.

By way of example, a hospital produces from 1.5 to 15 kg of waste per patient per day, whose composition (usually made up of 75% "urban-type waste" and 25% special hospital or potentially infectious waste) depends not only on the activities performed in the different hospital departments, but also on the type of hospital involved (Angelini *et al* [1]).

4 Occupational exposure and associated risks

Many studies have demonstrated that sanitary waste disposal workers can come into contact with biological agents more frequently than other occupational categories (Canegallo *et al* [2]). As regards occupational exposure (biological, infectious, chemical risk), the productive sectors mainly involved in the management of sanitary wastes are: hospitals, diagnosis and research laboratories, *stables*, activities connected with disinterment and exhumation. The commonest carriers are biological materials such as blood, urine, cerebrospinal and seminal fluids, faeces, sputa, organs and/or tissue fragments, animal carcasses and chemicals such as disinfectants (aldehydes, alcohols, phenols, etc.), detergents, medicines (chemotherapies, antibiotics, hormones). Moreover, we must remember that even small quantities of pharmacological substances usually have a high active capacity and this increases the risk for those handling them. In this case eventual pathologies are caused by genotoxicity (chromosome aberrations, mutagenic effects on DNA), carcinogenesis, teratogenesis and a reduction in fertility. *Stables* deserve a special mention since, like laboratories, they are defined according to the risk category of the biological agents studied there. There are four levels of biosafety that take into consideration different classification factors. These include the pathogenicity of microorganisms, methods of transmission and possible hosts, the availability of effective preventive measures and treatment. For example, in the veterinary field, waste disposal workers can come into contact with and be infected by endo- and ectoparasites that are the natural hosts of stabled animals (infections from caseworm and *cryptosporydius*); moreover, some animal zoonoses can be transmitted to humans (brucellosis, leptospirosis, hemarrhagic fever, etc.). Immune-type reactions can be caused by the dissemination of hairs or other components of animal origin that act as allergens. In the hospital sector, the commonest etiologic agents workers are exposed to are: hepatitis A, hepatitis B, hepatitis C, HIV viruses, bacteriological infections due to many different germs such as *Bacillus, Staphylococcus, Mycobacterium*, fungal infections from the *Aspergillus* variety, parasitosis; most of these occur accidentally or through the faulty application of safety procedures (Trevisan *et al* [3], Franchi *et al* [4], Lopez Hernandez *et al* [5]). The type of biological agent involved depends on the process units in which exposed workers are employed, but in the case of exposure to non-infectious agents, studies usually reveal exposure to bio-aerosols containing live bacteria, moulds, bacterial mycotoxins and endotoxins. In this case reactions are prevalently immunologic (bronchial asthma, dermatitis, allergic rhinitis, alveolitis) or toxic (absorption of endotoxins and ß 1-3 glucans) (Barth *et al* [6]).

5 Legal framework

5.1 National legislation on hospital/sanitary wastes up to 1997 ("Ronchi" Decree)

Up to 1997, national legislation concerning hospital/sanitary wastes was often influenced by urgent or contingent circumstances. Consequently it was not always introduced in an organic way into what was already a disjointed and complex national legal framework on waste materials.

The Legislative Decree 5 February 1997, n. 22 [7] made a significant attempt to redress this situation.

With regard to sanitary wastes, after reviewing all the regulations, the Ronchi Decree

- confirms the fact that wastes derived from sanitary activities are to be considered special wastes;
- classifies sanitary wastes as dangerous in compliance with the European list of dangerous wastes;
- lays down rules concerning the temporary storage of sanitary wastes, the responsibility of hospital managers and the final destination of sanitary waste;
- introduces a register for the loading and unloading of wastes;
- defers the definition of further sanitary waste management methods and the identification of sanitary wastes requiring special disposal systems to later legislation

On the basis of the definition of sanitary waste in code 18 of the European list of dangerous wastes, this Decree refers first of all to an approach that requires only potentially infectious wastes that represent a real hazard for public health to be subject to special laws.

The law requires sanitary wastes to be incinerated; disposal in a rubbish tip is allowed only in the case of urgent and contingent circumstances, and constitutes an exception.

In point of fact, a distinction has been made between dangerous wastes that must be incinerated and all other wastes (kitchen, office waste, etc.) that must be disposed of according to the category to which they belong. In short, this special law has been applied only to the field of sanitary wastes (Apat [8]).

As for the intermediate phases of sanitary waste management, the Decree lays down rules for temporary storage and sterilisation, deferring regulation of further phases to subsequent administrative action.

In the case of temporary storage, the law provides for a maximum of 5 days for quantities exceeding 200 litres, and 30 days in all other cases, provided temporary storage takes place on the production site and under conditions that cannot cause alterations that could involve health risks.

In the case of sterilisation of dangerous sanitary wastes off the site of the hospital that produced them, the Decree states that this is subject to authorised procedures similar to those required for other types of waste.

5.2 Ministry Decree 26 June 2000, n. 219

This regulation [9] sets out all the specific technical rules required under art. 45 of the Ronchi Decree, i.e.:

- collection, disinfection, sterilisation, transportation, recovery and disposal of dangerous sanitary wastes;
- identification of cemetery wastes and the technical rules for its management;
- identification of elements of sanitary wastes included as urban waste matter;
- identification of elements of sanitary wastes that require special disposal systems

This regulation includes important references to true sanitary wastes, i.e. dangerous and potentially infectious wastes. It also regulates the following types: wastes requiring special management methods (unusable medicines or those that are past their use-by date, including antineoplastic medicines, laboratory animals and drugs), exhumation and disinterment wastes, non-hazardous sanitary wastes, included in the definition of urban waste, and dangerous non-infectious sanitary wastes (Angelini *et al* [1]).

On the subject of the handling of potentially infectious sanitary wastes, the MD carefully states that from the very first stages of waste production, appropriate and easily identifiable strong disposable packaging must be used for collection, storage and removal so as to prevent persons or things coming into contact with these wastes before final disposal.

The decree no longer requires wastes to be disinfected before removal. As regards the creation and management of sterilisation processes for potentially infectious wastes, the MD states: “ The sterilisation of dangerous sanitary wastes performed off the production site is subject to authorised procedures”.

On the other hand, the MD explicitly states that sterilisation plants located on hospital premises need not be authorised (articles 27 and 28 Legislative Decree 5 February 1997, n. 22) provided these plants treat only the sanitary wastes produced in that particular hospital.

In any case, this concessionary regime for the auto-disposal of wastes falls outside the jurisdiction of “dangerous” waste disposal (art. 32, comma 6 and art. 3, comma 2 of Directive 91/682/EC). The MD also aims at reducing waste by separating it at source and thereby diminishing the quantity of waste sent for disposal, and also at reducing the hazardous nature of wastes by separating out the most contaminating parts.

5.3 Italian Presidential Decree 15 July 2003, n. 254

This Decree [10], which lays down new regulations concerning the management of sanitary and special wastes that are equally as dangerous as potentially infectious ones, abrogates all former regulations, even rules of law, on the subject of sanitary wastes (including art. 45, Legislative Decree 22/97) and

unconditionally addresses all waste producers. This legislation is still in force even following the publication of the new Consolidation Act on the environment; an Act that includes a section on wastes, but that does not change the validity of the regulations of the Decree in question.

The Decree defines as sanitary wastes all those listed in its enclosures I and II, and divides up relative case records more than in the past. They thus include those produced by private and public organisations (referred to in Leg. Decree 509/92) engaged in medical and veterinary prevention, diagnosis, treatment, rehabilitation and research, and also special wastes produced outside health facilities (with the exception of sanitary napkins), that are nevertheless classed as potentially infectious dangerous wastes.

These wide-ranging new regulations apply to all health facilities (hospitals, etc.) and all public and private health workers (general practitioners even when operating outside their own surgeries, dental surgeons, professionals, etc.) as they all they produce special wastes. Thus microbiological water and food testing laboratories, and also beauty parlours and the like form part of the latter categories since throwaway sharps come into contact with vascularised skin. This means that even small facilities must now follow all the requirements necessary for the management of this type of waste: temporary storage methods, a loading and downloading register, annual MUD declaration.

The Presidential Decree therefore sets out rules for waste management that are again based on the type of wastes involved, but it includes a few exceptions compared to previous legislation.

Non-hazardous sanitary wastes can be temporarily stored on the production site for up to a year provided they do not exceed 200 mc. However, larger quantities of waste must be disposed of every three months.

Potentially infectious sanitary wastes can be temporarily stored for up to thirty days if they do not exceed 200 litres, or for up to a maximum of five days after the container is closed; they must also be in compliance with hygiene and safety requirements and under the responsibility of the producer. These wastes must be kept in special packaging suitable for sharps or non-sharp material. Disposal must be by incineration in authorised plants and, if wastes pose a further hazard (dental amalgam with biological fluids, etc.), they should be disposed of in plants reserved for dangerous wastes. The Presidential Decree also states in this type of disposal, there must be no direct handling of waste during furnace loading.

Wastes can also be sterilised so as to simplify the disposal process, but this involves very complex and burdensome procedures that are rarely available in small health facilities. In any event, the temporary storage, collection and transportation of this type of waste are regulated by technical provisions for the management of special non-hazardous wastes.

As to other wastes, under this provision, stone and inert materials used in cemetery constructions are exempted from the need for authorisation, whereas management rules regarding sanitary wastes assimiliable to urban wastes, remain basically unaltered.

6 Conclusions

Given that infection may occur after a single exposure, it is vital to minimize risks by using a control plan to prevent exposure. In fact current engineering know-how, up-to-date safety apparatus (double containers, puncture resistant *polyethylene* containers for sharps), the implementation of good laboratory practices (use of gloves, protective clothing) and staff training now enable us to minimize risks. Our review of recent legislation in this field shows that regulations have now been directed at containing risks associated with the management of sanitary wastes. This has been done by tackling important questions of management such as the identification of waste categories in which there is an unequivocal risk of infection, the packaging of wastes for storage and transportation, the sterilisation of sanitary wastes in compliance with validated technical regulations; and by attempting, when this is feasible and safe, to encourage recycling.

References

[1] Angelini, p., Meneghini, F., Gataleta F., *La gestione dei rifiuti sanitari*, Centro Editoriale e Librario: Università degli Studi della Calabria, 2000.

[2] Canegallo, C., Ciliberti, R., Melioli, R., Il rischio da agenti biologici nel laboratorio di ricerca e di analisi cliniche. *Ambiente e Sicurezza sul Lavoro*, **11**, pp 57-62, 1997.

[3] Trevisan, A., Stocco, E., Fanelli, G., Bicciato, F., Paruzzolo, P., Seroprevalence of hepatitis A markers in subject exposed to biologic risk. *Internal Archives of Occupational and Environmental Health* **72(2)**, pp 125-127.

[4] Franchi, A., Amicosante, M., Rovatti, E., Bonini, R., Marcheggiano, P., Girardi, E. et al., Evaluation of a western blot test as a potential screenig tool for occupational exposure to Mycobacterium tubercolosis in health care workers. *The Journal of Occupational and Environmental Medicine* **42(1)**, pp 64-68, 2000.

[5] Lopez Hernandez, B., Almagro Nievas, D., Cabrera Castillo, MJ., Seroprevalence of brucelosis in the workers of a plant of treatment of sanitary wastes. *Med Clin*, **120(10)**, pp 376-7, 2003.

[6] Barth, E., Talbott, N., Gable, R., Richter, S., Reponen, T., Evaluation of bioaerosol exposures during conditioning of biofilter organic media beds. *Appl. Occup. Environ. Hyg*. **17(1)**, pp 10-4, 2002.

[7] Decreto Legislativo 5 febbraio 1997, n. 22, Attuazione delle direttive 91/156/CEE sui rifiuti, 91/689/CEE sui rifiuti pericolosi e 94/62/CE sugli imballaggi e sui rifiuti di imballaggio, G.U. n. 38, 15 febbraio 1997 - Supplemento Ordinario n. 33.

[8] APAT, Il sistema di contabilità dei rifiuti sanitari: una indagine conoscitiva, Rapporti*,* **33**, Roma, 2003.

[9] Decreto 26 giugno 2000, n. 219, Regolamento recante la disciplina per la gestione dei rifiuti sanitari, ai sensi dell'art. 45 del decreto legislativo 5 febbraio 1997, n. 22, G.U. n. 181, 4 agosto 2000.
[10] Decreto Presidente della Repubblica 15 luglio 2003, n. 254, Regolamento recante disciplina della gestione dei rifiuti sanitari a norma dell'art. 24 della legge 31 luglio 2002, n. 179, G.U. n. 211, 11 settembre 2003,

Section 3
Planning and development

Public action and social capital in rural areas

C. Cecchi & E. Basile
SPES Development Studies,
Research Centre at the University of Rome "La Sapienza", Rome, Italy

Abstract

This paper deals with the process of building social capital as constrained by the availability of public and private services. The purpose is to show that social capital shapes, both in positive and negative terms, the capacity of local institutions to support development projects and strategies. The analysis shows that a higher level of social capital gives to communities a stronger planning capacity; a lower level of social capital implies that new initiatives are implemented only if they give private individual benefits; a higher level of social capital generates actions that increase and improve social capital itself.

The first section outlines the logical links between the concept of social capital and the availability of "services". Social capital is seen here as a public good. Section 3 focuses on the relationship between the supply of *public* services and the level of social capital. Section 4 compares the Italian case with other cases in Northern Europe.

The concluding section shows the relationship between marginality and the availability of public goods. It contributes to the analysis of the relationship between social capital and development, showing that there is both a theoretical framework and evidence to justify the "principle" that public action supporting social capital has a major role in determining the performance of development strategies and actions.
Keywords: social capital, rural marginal areas, planning systems.

1 Introduction

Social capital is frequently considered as a public good. For example the UK National Statistical [19] agrees with Putnam [21] and adopts this definition:

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060141

> 'Social capital is generally perceived to be a private and public good (Putnam, [22]) because, through its creation as a by-product of social relations, it benefits both the creator and bystander. It is a classic public good because of its non-exclusivity - its benefits cannot be restricted and hence are available to all members of a community indiscriminately' (Woolcock, [29])

This definition emphasises the possibility for all members of a community of gaining from *the use* of the social capital. But also the state is a subject that actively participates to the process of creating social capital. The state sustains private action because it contributes in creating the social and political environment; the important role of the state in this action is justified because of the *public good* character of the outcomes of social capital.

Which are the ways the state uses for the creation of or for the support to social capital?

Taking social capital as an example of the basis for the production of public goods, one can easily observe that there is not any single firm interested in producing that type of public good, which is non-excludable and non-rival in consumption. On the contrary, it will be easy to find people who over-consume the public good or who have no hindrance to deplete it, if its use gives them some satisfaction, utility or cost reduction in some economic activity.

A broader interpretation of public goods can be taken to include goods that are not *per se* public goods but goods that entitle people to exercise a public good. For example, railways can be considered as public goods, "if one assumes that the right to freedom of movement means that one citizen cannot be excluded by virtue of an appropriation by another citizen" (Bianchi [5], p. 111).

Public goods are therefore a consequence of a "natural" or an "institutional" market failure, which prevent the supply of meeting the needs of the population and of the firms.

The shortage of a public good can therefore be faced by means of action implemented by the state or by some group. The difference between the former cases refers to the resources used to produce the public good and to the people who are entitled to use the newly produced public good.

The state – and, in general, public administration – chooses to supply a public good using public resources coming from direct or indirect taxation. The state decides to produce it by means of state owned structures or to buy it from a private enterprise. The state decides to make a consumer pay a "contribution" – not a price – for the use of that public good, or not.

Some groups – private associations, formal and informal networks – can choose to supplement the shortage of a public good by means of a "private" supply. The group chooses the resources to be used for the production of the public good: from the pocket of a single member or from the pockets of all members. The group decides how to produce the public good: by means of structures owned by the group, or buying it from a private producer. The group decides whether it supplies a proper public good (with non-excludability) or it supplies a quasi-public good (reserving the use to the members of the group), which represents a substitute of the lacking public good.

As a summary, social capital represents for a community also a stock that provides public goods, which can be increased or consumed by means of the action and of the choices of the state, of groups and of individuals.

2 The *Maremma* district as a rural area

With particular reference to the area we are now considering – the *Maremma* rural District in Southern Tuscany (Italy) –, it is necessary to emphasise the fact that we are dealing with a rural area. Therefore, we must take into account some specificity that distinguishes this type of territory.

With reference to the evolution of the Italian economy during last decades, rural and agricultural local system analysis gives an important contribution to the understanding of the development path, because this kind of territory cover more than a half of the national surface and contain more than 40% of the Italian population.

The rurality that emerges in recent times appears as very different from the pre-industrial one. Basile and Cecchi [2] have shown that this difference is as relevant at the theoretical level as it is on the basis of the evidence coming from the Italian countryside. Post-industrial rurality is characterised by a complete integration within the contemporary economic and social organisation of the capitalistic world. This organisation is different from others' local systems because *i.* production is differentiated between and within sectors; *ii.* the dispersal of economic activities generates a rarefied social fabric that is, under many circumstances, the antithesis of the industrial district; *iii.* finally, the presence of agriculture, amongst other activities, completely integrated represents a characterising element, which contributes to build the singularity of the economic system and the qualification of the social system.

Before the end of the Nineties, economists have been able to show two different situations for the countryside. The first one relates to areas where agriculture's modernisation has been able to specialise resource with an agricultural use. On the contrary, where the modernisation process has failed, a large part of local resources remained under-employed, and available for a different use. In this second situation, "modern" rurality emerges because rural resources have attracted the urban ones.

Modern rurality is then characterised by rarefied social fabric, economic differentiation and the presence of agriculture. This is the consequence of the interaction between urban and rural resources.

Modern rurality is frequently considered as a positive situation, because it represents a new vitality for declining social organisations. Nevertheless, a limit has been shown to the process of development and of growth for rural areas. This limit is represented by the potential shortage of services, which are needed for industrial growth, for resident families and communities.

In other words, it is true that urban areas expel human and financial resources, and it is true that rural areas are suitable for new activities. But, the availability of services generates new competitive advantages for each different type of territory.

Traditional services, financial services and "modern" services, when suitable for the economic rural local system, can sustain growth and development or can constrain them when not integrated in the economic and social organisation.

The research we have carried out in Southern Tuscany has analysed the role of services within the development processes, which involve rural and agricultural local systems, in three different directions. The first direction concerns the definition of specific characters that distinguish rural and agricultural development processes. The second direction concerns the definition of a typology of services, suitable for the support of the growth process of an economy, which is differentiated by sector and locally integrated. The third research direction concerns the identification of the subjects of the governance of rural development processes. The research has concentrated on the multi-dimensional composition of the development process. The dimensions are: the planning procedure, the management and the monitoring practice and, finally, the control and the accountability of actors.

Within the theoretical framework described above, public goods are strictly linked to public action and to the action of civil society. Our attention to public goods is confined to the local dimension; this assumption implies that we are not interested in public goods such as defence, which is usually considered as an emblematic kind of public good. Our focus is mainly on services – within the study area – that constitute the way in which the state contributes to the building up and to the increasing of public goods. But we are also interested in understanding the ability of the local community to face and to supplement any lack of state action. Therefore, we looked for plans implemented by the civil society – networks, association etc. – that produce or develop local public goods.

We focus on three types of services, which refer to different dimension of the public good component of social capital. In addition, we focus on the creator's dimension of those services; in other words, we have tried to highlight the subject that has the power of supplying the service, without any reference to the subject that actually produces it.

3 The *Maremma* backwardness in terms of services

The combined action of the structural transformation and of the Italian welfare system has shaped the initiatives enhanced by the population in the area. On the one side, the change that has affected the economic structure of the area has destroyed a large part of the traditional jobs: the mine sector has definitely closed down during the 1970s, and agriculture has witnessed a massive introduction of technical change, which has caused an astonishing reduction of labour requirement. On the other side, the State has mainly concentred its action on the agrarian reform, during the 1950s; after that action, the welfare state prevailing in Italy has largely failed to meet the needs of a population who lived in an area with a declining level of opportunities. People in the area – as it was in any Italian agricultural and marginal areas – obtained public support only from the agricultural policy, and obtained a lower welfare support that the one granted to industrial workers and to people in towns. Under these conditions, the private

initiatives have been mainly oriented towards profit generating activities; migration to growing industrial areas represented the most used alternative for the population who lacked resources.

The lack of public services has been faced more by means of "exits" from the community than by the search of other people who shared the same need; as an alternative, private individual choices has supplemented the shortage due to the little state action. The population continues to migrate from the area, even if there is some significant amount of immigrants. These immigrants apparently requires a lower level of public services or – as wealthy as they are – they can afford to buy private services – mainly outside the area. The most recent concern for the environment in rural areas should promote collective actions, but instead the prevailing trend shows many forms of privatisation also of the environmental dimension of the community.

The interviews we have carried out in the area show that development actions improve and increase, in the opinion of the people, more the private earnings than the quality of life of the community as a whole.

4 Welfare state, public services and social capital

This section aims to show now how communities participate to the process of building new and structured initiatives as part of some network or some association. In other words, we link the degree of participation to common local life to the process of building social capital.

In our research, the study of social capital has focused on the analysis of networks and the participation of the members of the local community to networks. Attention has been paid to networks that have been build on the basis of the public action and of the functioning of public administration; but the greatest attention has been paid to networks that are "privately" managed and those that people build as a "place" where to share common interests and common concern toward some problem.

In Sotkamo (FIN), all the respondents belong at least to one (formal) association and, amongst them, there is a group – that Lehto and Oksa ([16], p. 31) have named as *super-active* – composed by people who are not only members of many groups but that play also an active role at least in four networks. Respondents to questionnaires consider as successful the development actions, implemented during the recent years, more because of the degree of their participation than for the economic aim of the actions themselves.

Also in the Norwegian Mountain Region (N), the degree of participation appears as very high because all the interviewees are member of at least one network/group. The degree of information and of participation to public action in the area is quite surprising.

> '… the general impression is that the various state or semi-state agencies are perceived to play the most important role.' (Rye and Winge, [24], p. 12)

In Leksand and Rättvik (S), networks play a very important role because they represent the basis for the construction of trust of the citizens towards public administration. In Sweden too the participation to networks and, mainly, the voluntary work of the members show the old tradition of "village democracies". The role of network in building plans and other development initiatives is emphasised by the Swedish team noting that the success of those plans derives from the secondary position played by key actors and from the capacity to involve in actions also the large number of entrepreneurs who are not embedded in networks of civil society (Tillberg Mattsson and Stenbacka, [27]).

At Skye and Lochalsh (GB), whilst the number of associations and networks in lower than in the other Northern areas, the degree of participation to the activities of those groups is particularly high. This high degree of participation is associated with a diffused success of public initiatives that have also involved citizens and associations.

> '… the results suggest that respondents are less than content with those areas of development that are perceived to be the responsibility of the national governments.' (Lee and Árnason, [15], p. 28)

A similar situation is observed in the Lake District (Irl). Also in this area networks and groups exist, but the degree of participation appears as a non-relevant aspect. The link between association and the public administration appears as vague, and the associations do not see any consequence between their action and development (Kinlen, [14], p. 18). The Irish report highlights that people appear as not being in condition to describe their role in development 'despite mentioning their involvement in various community and voluntary associations in the area' (*ibidem*).

The case of Maremma (I) appears as significantly different. The degree of participation to associations and networks is lower than in any other area. One third of the sample does not give any answer to the question related to the type and number of association the respondents belongs to. In relation to the counterpart of the public action for development, it appears that only few type associations play some role. The evidence given by respondents shows that people in the sample are not in condition to identify a limited group of specific subjects that link the administration with citizens.

The major consequence of the weakness of the links between planners and citizens is represented by the low level of awareness of what development actually means for a citizen and about the ordinary ways of planning development actions.

As a summary, we can observe that there is a relationship between the quantity and quality of formal and informal networks of citizens and of associations, and the capacity of the public administration to face the needs of the population or, at least, to make the population aware of the public action. The consequence of this observation is that networks of relations increase and improve their way of functioning by means of public support. For this reason, we must emphasise the necessity to pay attention to social capital in all its components in order to understand the process of change in rural areas.

5 Social capital and planning

The first consequence of our research on restructuring of rural marginal areas relates to an alarm bell that signals some dangerous weaknesses of the rural development policy of the European Union. This observation comes from the analysis on the relationships between social capital that characterises a community and the development actions that the same community is in condition to plan and to implement.

The bell rings in a room that hosts the more general theme of public policies for rural development and their funding. On the one hand, the concern comes from the reduction of public expenditure, as a consequence of the commitment of the European Union members' governments to maintain the so-called Maastricht parameters within the agreed range of variation, which has its major consequences on the welfare state and on the provision of public services that should be granted to the population. On the other hand, the enlargement towards East of the European Union raises problems linked to the risk of another change of the emphasis of Common agricultural policy (CAP) in favour of the support of agriculture instead of economic differentiation in rural areas, because agriculture, in these countries, is still a very important activity from many points of view.

The consequences of this twofold recent threat concern the entire functioning of the CAP, from the growing importance given to role of private initiatives – the supremacy of market forces – to the types of strategies to be implemented for rural development support. The regional and social equilibrium, which was at the basis of the TREATISE OF ROME since the 1960s, has been cancelled because of the failure of the mechanisms introduced with the CAP tools to support the agricultural sector and because of great change that has involved the countryside at the end of the 20th century. The CAP of the 1990s has promised to give support only to farmers who were becoming true entrepreneurs showing their attention to market forces; the same policy has also promised support to rural communities that were able to plan development actions by means of internal forces – the so-called endogenous development strategies.

This change of the aims of the EU's policy has been based on the assumption that social and regional problems might be solved by the private action and that the only role of the government is to supplement resources to those who lack them.

Our research has shown that this assumption is completely wrong. In other words, it is wrong to assume that farmers and the rural population behave always in order to maximise their profits and it is wrong to assume that the local communities have always had a strategy that allows the filling of the gaps that separate them from the richest areas.

Focusing on the social capital that characterises a number of rural communities in marginal areas, we have shown that the capacity of defining and of implementing development strategies highly depends of the social capital availability, which in turn depends on the history of the community in terms of the results reached by means of the public action. Moreover, the success of the public action depends on the type and on the level of services that the state has

granted to local population. This means that the greater effectiveness of rural policies does not depend on the target of the public funds transfers – farmers or other components of the community; it does not depend on the specific strategy that each single community is able to plan; it rather depends on the capacity of public action to create the conditions that make any action – and the resources that are necessary for their implementation - efficient and productive.

The social capital that characterises a community significantly influences: i. the capacity of the governance of the community to project new initiatives; ii. its management capacity of dealing with the implementation of the initiatives; iii. the capacity to link top-down actions "suggested" by the central government to local private personal or collective actions.

From the analysis of the link between the needs of the population and the way in which each single citizen can meet those needs, it emerges that in those localities where the "normal" answer to needs comes from public services, there people trust collective action as a means of solving both private and common problems. In other words, communities with a good level of public services have a high level of awareness of the role of collective action and of the significance of public services as public goods. Therefore, one may expect that, as a consequence of the reduction of the supply of services by the state, the network of local relations creates the condition that allows for the substitution of the state action with a local collective action. This causal relation appears to be as stronger as higher is the level of participation of single citizens to the decision process – the level of democracy. This "tradition" of democratic participation to the decision making process – in other words, the degree of coincidence between local government and local governance – explains the reason why development programmes represent, in these areas, the result of answers to local needs by means of local resources and of external public funding.

References

[1] Arrow K. (2000), Observations on Social Capital, in P. Dasgupta and I. Serageldin, eds., *Social Capital: A Multifaceted Perspective*, World Bank, Washington, D.C.

[2] Basile E. and Cecchi C. (2001), *La Trasformazione Post-Industriale della Campagna. Dall'agricoltura ai Sistemi Locali Rurali*, Rosenberg & Sellier, Torino.

[3] Becattini G. (1999), Introduzione, in *Modelli locali di sviluppo*, edited by G. Becattini, Il Mulino, Bologna.

[4] Becattini G. edited by (1987), *Mercato e forze locali*, Il Mulino, Bologna.

[5] Bianchi P. (1998), *Industrial Policies and Economic Integration. Learning from European Experiences*, Routledge, London.

[6] Cecchi C. (2001), *Rural Development and Local Systems. The Case of the "Maremma Rural District"*, Department of City and Regional Planning, University of Wales, College of Cardiff.

[7] Cecchi C. (2002), Sistemi locali rurali, e aree di specializzazione agricola, in *Sviluppo rurale: società, territorio, impresa*, a cura di E. Basile e D. Romano, Franco Angeli, Milano: pp. 90-115.

[8] Fine B. (2001), *Social Capital versus Social Theory. Political economy and social science at the turn of the millennium*, Routledge, London.
[9] Ferrara M. (1998), *Targeting Welfare in a "Soft" State, Italy's Winding Road to Selectivity*, Paper prepared for the ISSA Conference, Jerusalem, 25-28 January.
[10] Grootaert Ch. and van Bastelaer Th. (2002), Social Capital: From Definition to Measurement, in *Understanding and Measuring Social Capital. A Multidisciplinary Tool for Practitioners*, edited by Grootaert Ch. and van Bastelaer Th., The World Bank, Washington D.C.
[11] Hillmert S. (2003), Welfare State Regimes and Life-Course Patterns: An Introduction, http://www.mpib-berlin.mpg.de/en/institut/dok/full/e2001.0232/frames/paper.htm
[12] Jonsson I., Rydén G., Tillberg K. (2002), *Innovativeness and traditionalism: Rural development in Leksand and Rättvik*, Sweden, Swedish context report, Restructuring in Marginal Rural Areas (RESTRIM) - Dalarnas forskningsråd (DFR), Dalarna Research Institute, Falun, Sweden.
[13] Kinlen L. (2002a), *The Lake District, County Mayo*, Irish context report, Restructuring in Marginal Rural Areas (Restrim) - The International Centre for Development Studies, National University of Ireland, Galway, Ireland.
[14] Kinlen L. (2002b), *RESTRIM Project Analysis of Questionnaire,* Irish questionnaire report, Restructuring in Marginal Rural Areas (Restrim) - The International Centre for Development Studies, National University of Ireland, Galway, Ireland.
[15] Lee J and Árnason A. (2003), *Questionnaire Report for Skye and Lochalsh*, Scottish questionnaire report, Restructuring in Marginal Rural Areas (RESTRIM) - Arkleton Centre For Rural Development Research, University Of Aberdeen, Scotland (GB).
[16] Lehto E. and Oksa J. (2002a), *Rural development in Kainuu region in Finland: A Case of Sotkamo,* Finnish Context Report, Restructuring in Marginal Rural Areas (RESTRIM) – Research and Development Centre of Kajaani, University of Oulu and University of Joensuu, Finland.
[17] Lehto E. and Oksa J. (2002b), *RESTRIM Questionnaire Study of Sotkamo, Finland,* Finnish Questionnaire Report, Restructuring in Marginal Rural Areas (RESTRIM) – Research and Development Centre of Kajaani, University of Oulu and University of Joensuu, Finland.
[18] Lowe, P., Ray, C., Ward, N., Wood, D., Woodward, R. 1998, *Participation in Rural Development: a Review of European Experience*, Centre for rural economy – Research report, Newcastle upon Tyne.
[19] National Statistics (2001), *Social Capital. A review of the literature*, Social Analysis and Reporting Division, Office for National Statistics, October.
[20] Nightingale A. (2002), *The Magic of Skye: Innovation, Growth and Rurality in Skye and Lochalsh*, Scottish context report, Restructuring in

Marginal Rural Areas (RESTRIM) - Arkleton Centre For Rural Development Research, University Of Aberdeen, Scotland (GB).
[21] Putnam R. (1993), *Making Democracy Work: Civic Traditions in Modern Italy*, Princeton University Press, Princeton.
[22] Putnam R. (2000), *Bowling Alone - The Collapse and Revival of American Community*, Simon & Schuster, New York.
[23] Rye J.F. and Winge A. (2002a), *The Mountain Region*, Norwegian context report, Restructuring in Marginal Rural Areas (RESTRIM) - Centre For Rural Research, Norwegian University of Science and Technology, Trondheim, Norway.
[24] Rye J.F. and Winge A. (2002b), *A survey of development priorities, Projects and networks in the Mountain Region*, Norwegian questionnaire report, Restructuring in Marginal Rural Areas (RESTRIM) - Centre For Rural Research, Norwegian University of Science and Technology, Trondheim, Norway.
[25] Rye J.F. and Winge A. (2002c), *The role of Networks in Promoting New Usage of Farmland in the Mountain Region, Norway*, Norwegian report on networks, Restructuring in Marginal Rural Areas (RESTRIM) - Centre For Rural Research, Norwegian University of Science and Technology, Trondheim, Norway.
[26] Sabatini F., (2003), *Capitale sociale e sviluppo economico*, Serie Working Papers dei dottorandi Dipartimento di Economia pubblica – Università degli studi di Roma ‘La Sapienza’, No. 12 – luglio.
[27] Tillberg Mattsson K. and Stenbacka S. (2003), *The Role of Social Capital in Local Development: The Case of Leksand and Rättvik*, Draft National Report – Sweden, Restructuring in Marginal Rural Areas (RESTRIM) - Dalarnas forskningsråd (DFR), Dalarna Research Institute, Falun, Sweden.
[28] van Staveren I. (2000), *A Conceptualisation of Social Capital in Economics: Commitment and Spill-Over Effects*, Institute of Social Studies, The Hague - The Netherlands, Working Paper Series No. 324 – November.
[29] Woolcock, M. (2001) The place of social capital in Understanding Social and Economic Outcomes, *ISUMA Canadian Journal of Policy Research,* 2 (1), pp. 11-17.

The needs of sustainability: the problem of data availability for calculating indicators

F. M. Pulselli, S. Bastianoni, V. Niccolucci & E. Tiezzi
Department of Chemical and Biosystems Sciences, University of Siena, Siena, Italy

Abstract

New instruments and scientific tools are necessary to improve our knowledge about the state and the possible evolution of complex systems, both natural and anthropic, from a sustainability viewpoint. In order to achieve the assessment of environmental sustainability, several holistic methods have been developed in the last two decades that are often applied worldwide. If holistic methodologies are to be implemented for the study of anthropic systems, such as territorial systems or production processes, in order to calculate synthetic sustainability indicators, then the availability of the most reliable and consistent data set is a *condicio sine qua non* to obtain a verisimilar representation of reality. Nowadays, databases created and usually provided by institutions seem to neglect important aspects of the relationship between human life and the biophysical bases which it is founded on. In other words, while it is relatively easy to obtain information about economic aspects of human activity (in monetary terms), it is arduous to appreciate the physical essence of flows and stocks of natural resources directly and indirectly involved in the same processes. This paper shows the application and comparison of more than thirty indicators applied to territorial systems, in order to highlight their complementarities and overlapping and the most relevant data to be collected.

1 Introduction

Today, the progressive spreading of the concept of sustainable development is leading people to look for solutions from a political, economic, social, institutional and environmental point of view. This new tendency, that may be defined as "therapeutic" (with respect to various problems that have arisen in

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060151

recent years), is often partial and only effective in the short run, contravening the basic element of sustainability: time. In our opinion, before any "sustainability therapy", it is necessary to implement a "sustainability diagnosis" on the basis of an objective analysis of the environmental platform which human activities are based upon. Currently, the lack of diagnostic instruments is often mentioned by different documents. Article 40 of the Agenda 21 (UNCED, [10]) deals widely with sustainability indicators: "Indicators of sustainable development need to be developed to provide solid bases for decision-making at all levels and to contribute to a self-regulating sustainability of integrated environment and development systems". From this statement, it is clear that a serious diagnostic analysis is a necessary and preliminary condition for designing consistent sustainable policies. Measuring sustainability is an ambitious task, that is different from assessing how clean or polluted a given system is, and implies several levels of knowledge of the system under study. For this reason, advances in scientific research play a key role in supporting decision making. The territorial context is conceived as a complex, dynamic and open system in which several components develop and interact with each other. The choice of methods and indicators as instruments for assessing sustainability is a key point of such an analysis. The holistic vision, that considers the world as an integrated ensemble rather than a set of parts separated from each other, seems to be the most appropriate approach to adopt for the study and analysis of anthropic systems.

2 Methods

Sustainability is a state of the ecosystem that consists of interrelated components such as biotic elements and economic systems, in which the structure and the functions of each component should be maintained for a long period of time. A sustainability assessment can be performed on the basis of the flows of energy and matter and deviations can be corrected by means of socio-economic measures (Ruth, [6]). Some methods are presented in order to evaluate different aspects of the impacts of human activity and assess the territorial sustainability: Emergy evaluation, Ecological Footprint analysis, Greenhouse gas inventory and Remote sensing analysis. They are chosen in order to represent, under different viewpoints, the behaviour of a population that is in connection with the environment. The joint use of different methods gives rise to a holistic view of the system in order to implement suitable policies at different levels: territorial planning, resource and energy management, land use, etc. For each cited indicator the correspondent symbol (then used in Figure 1) is reported between brackets.

2.1 Emergy evaluation

To take into account all the flows of resources (natural and manufactured) supporting a system (national, regional or productive) and to evaluate the environmental work necessary in order to make them available, the concept of Emergy was introduced by Howard Odum [4,5]. He defined emergy as the quantity of solar energy necessary (directly or indirectly) to obtain a product or

an energy flow in a given process. Emergy is the memory of all the solar energy that has been necessary to feed a given process. Hence, emergy is the common basis which a system of environmental accounting can be built on. Its unit is the solar emergy Joule (sej). In order to convert all the inputs supporting a system into sej, the concept of Transformity is introduced. Solar Transformity is defined as the emergy required per unit of product or service (Ulgiati et al. [8]) or as the solar energy directly or indirectly necessary to obtain one unit of energy or mass (Joule or gram) of another type of energy.

Once all the inputs are classified in renewable (R) and non-renewable (N) resources and local (L) and imported (F) inputs, some synthetic indicators of the environmental performance can be calculated. For example, the Environmental Loading Ratio (ELR) is the ratio of non-renewable (local and imported) emergy to renewable environmental emergy; the Emergy per Person (EpP) is the ratio of total emergy to the inhabitants of the system; the Empower Density (EmD) is the ratio of total emergy to the area (expressed in sej per m^2); EYR is the ratio of the total emergy flow (Em) to the imported inputs (F).

2.2 Ecological Footprint

The Ecological Footprint (EF), introduced by W. Rees and M. Wackernagel [9], is defined as the total area of ecologically productive land (forests, arable land, pasture, built-up area, etc.) and water ecosystem required to produce the resources and services consumed by a given population, as well as to assimilate wastes generated by that population. The ecological footprint formulation starts from the consumption of goods and services by the population, and is based on the hypothesis that every unit of energy and matter derives, directly or indirectly, from a certain extension of land. The conversion of all consumed inputs into equivalent hectares of land necessary to produce the resources and to absorb the emissions is possible by means of suitable conversion factors provided on the basis of the global average productivity of land. The analysis also defines a bio-productive capacity of a certain region (local, national or global), and comparisons between the EF and the biocapacity (BC) of an area determine the so-called ecological deficit or surplus. The ecological deficit provides an evaluation of the local overload, revealing how much a region depends on extra-territorial productive capacity.

2.3 Greenhouse gas inventory

The anthropic emissions of greenhouse gases have global implications in terms of sustainability. According to the IPCC guidelines [3], the emissions in the energy sector, agriculture, land-use change and forestry and waste are monitored. The main result is the comparison between the emission of equivalent CO_2 (Eq CO_2) and the absorption capacity of the ecosystems inside the territory (Abs CO_2): their difference gives the net CO_2. The inventory includes emissions of CO_2, CH_4, N_2O, NO_x, CO, NMVOC (non-methane volatile organic compounds), SO_2, HFC, PFC, SF_6 and the absorption of CO_2. It is consistent with the policy of emission reduction and is a useful way of assigning their responsibility among different sectors and areas.

2.4 Remote sensing analysis

Remote sensing means that instruments or recorders are not in direct contact with the objects under investigation, and observations and measurements of those objects occur from a distance. This method (Focardi et al., [2]) is implemented by using satellite images, and focuses on the measurement of energy at different wavelengths that is emitted, transmitted or reflected from an object in order to determine certain properties of it. The results of the analysis enable one to gain information on the spatial characteristics of a territory by combining a series of multitemporal and multispectral images, and to determine, for example, land cover (ESI), biomass cover (NDVI) and emission characteristics (RTI).

Table 1: Data necessary for the Emergy Evaluation of a territorial system (data expressed in unit per year).

Input	Unit	Input	Unit
Local Renewable Emergy (R)		**Imported Goods and Services (F2)**	
Sunlight	J	*Breeding, Hunting And Fishing*	
Rain	g	Products of breeding	g or J
Wind	J	Products of forestry	g or J
Geothermal heat	J	Products of fishing and hunting	g or J
Rivers	g		
Tidal energy	J	**Imported Goods and Services (F2)**	
Seawater	g	*Minerals*	
		Metallic minerals	g
Local Non-Renewable Emergy (N)		Non metallic minerals	g
Soil erosion	J		
Water consumption	g	**Imported Goods and Services (F2)**	
Extracted minerals	g	*Manufacturing And Craft*	
		Food industry	g or J
Imported Energy Sources (F1)		Tobacco industry	g
Electricity	J	Leather industry	J
Gasoline and diesel	J	Textile industry	J
Fuel oil, LPG, lubrificants	J	Furniture and fixtures industry	J
Natural gas	J	Wood industry	g or J
		Paper industry	g or J
***		Industry of graphics	g or J
		Metallurgical industry	g
Imported Goods and Services (F2)		Mechanic industry	g
Agriculture		Industry of minerals	g
Crops and legumes	J	Chemical industry	g
Fruits, seeds, spices and tobacco	J	Industry of rubber	g
Plants and flowers	g	Other industries	g

3 The needs for a sustainability "diagnosis"

a) the availability of the most reliable and coherent data is a condicio sine qua non to obtain a verisimilar representation of reality and apply methods

Synthetic indicators usually result from a translation of different elements and phenomena into a common unit. The procedure starts by identifying those elements relevant for the analysis and collecting raw data about them. The tables (1, 2 and 3) summarize the detailed raw information that would be necessary to implement the methods described above.

It is clear that the reliability of a method is proportional to the availability of data. Without a systematic recognition of statistical information and measurement of local realities, no verisimilar picture of the state of the system or prediction for the future is possible. Territorial planning activity is difficult if a Public Administration does not know what is happening within its territory. Usually, socio-economic information is the easiest to find, whereas environmental aspects are often neglected.

Table 2: Data necessary for the Ecological Footprint analysis of a territorial system (emissions, wastes, tourists, consumption and production are expressed in unit per year).

Input	Unit	Input	Unit
General Data		*Chemical Products*	t
Area	m^2	*Metallic Products*	t
Population	n. of inh.	*Non metallic Products*	t
Land use: cropland, built-up, grazing, forest, marine lands	m^2		
Emiss. of pollutants: CO_2, CH_4, other gas., liq. and solid poll.	t or m^3	**Productions**	
Prod. and recycling of wastes	t	*Agricultural products*	
Tourists	n.	Products of Cultivation	t
		Derivative of products	t
Consumption			
Products of animal origin		*Zootechny*	
Food	t	Number of heads	n.
Other products	t	Derivatives (food or other goods)	t
Products of vegetal origin			
Food	t	*Energy*	
Other products	t	Production of electricity	GWh
		Autoproduction of electricity	GWh
Energy		Incineration of wastes	GWh
Coal	t		
Fossil fuels	t	*Chemical Products*	t
Methane	m^3	*Metallic Products*	t
Electricity	GWh	*Non metallic Products*	t

For example, the number of employees in a given sector is well known and long time series are registered, while the extracted quantity of sand (in m^3, not in monetary terms!) or the urban or industrial consumption of water is not so accessible. Currently, there is a growing tendency towards the collection of data in economic terms, or relevant for economic purposes, rather then information on the biophysical basis which every economic, social or urban system is founded on. In general, before any investigation into the sustainability of a territorial system, institutions should verify if they are prepared to plan suitable investments in information and knowledge.

b) sustainability indicators must be calculated by computing raw data in order to obtain synthetic information on wide complex phenomena

Depending on the availability of data, it is possible to describe various phenomena. As it has been shown, all the items are expressed in different units that must be converted and integrated together in order to calculate one or more synthetic indicators representative of a more complex picture. Indicators are thus

the results of mathematical combinations of elements and conversion factors, and the purpose of such a calculation is to highlight the most important mechanisms characterizing the territorial system, most of them stimulated by human activity. The physical-based methods described above imply that all the information is converted into a common unit (sej for emergy evaluation, global equivalent hectare for ecological footprint, equivalent CO_2 for Greenhouse gas inventory) on the basis of objective and not arbitrary conversions. This avoids the use of incommensurable indicators, that make it difficult to rank systems when they are compared with each other, as in the case of some multicriteria analyses. We have used a series of indicators that can be used within a sustainability analysis. In particular, we present some selected data together with most of the indicators from the methods presented above. They contain information that is often easily understandable by policy makers, and that can steer planning and programming policies as well as give information on the current and optimal level of urbanization, industrial development, energy and resource management, etc.

Table 3: Data necessary for the Greenhouse Gas Inventory of a territorial system (production, consumption, rice and forest growth, combustion and wastes are expressed in unit per year).

Input	Unit	Input	Unit
Agriculture		**Energy Sector**	
Zootechny	n. of heads	LPG	t or m^3
Rice growing	t	Natural Gas	m^3
Combustion of residual	t	Wind	t or m^3
Fertilizer	t or l	Gasoline	t or m^3
		Kerosene	t or m^3
Wastes		Fuel oil	t or m^3
Solid wastes production	t	Electricity	GWh
Commodity composition	-		
Sewage	m^3 or l		
		Forestal Sector	
		- Wooded areas (ha) and growth (m^3/ha) of:	
Industrial Sector (productions)		Chesnut	
Concrete	t	Beech	
Mass concrete	t	Hop-hornbeam	
Limestone	t	Downy Oak	
Sodium Carbonate	t	Turkey Oak	
Asphalt	t	Holm-Oak	
Glass	t	Silver fir	
Ammonia	t	Stone pine	
Nitric acid	t	Austrian pine	
Iron	t	Robinia	
Steel	t	Cypress	
Aluminium	t	- Wood Use	
Paper	t	Working timber	m^3 or kg
Liquors	l	Firewood	m^3 or kg
Food	t	- Blazes: burnt-out areas	ha

c) the integration of methods and indicators is possible through statistical instruments, in order to efficiently provide the most reliable and usable information - A case study.

A research project, called "Spin-Eco" (The SPIn-Eco Project (the acronym means "Sustainability of the Province [of Siena] by means of Ecodynamic

Indicators", that explains the need to relate ecosystems to thermodynamics, ecology and the social sciences in order to arrive at satisfactory solutions for sustainable development and environmental management) is a multi-year (2001-2004) research program proposed by the Provincial Administration of Siena and funded by the Monte dei Paschi Foundation.), has been developed by a team led by the University of Siena, Italy, with the purpose of assessing the environmental conditions and the sustainability of the Province of Siena by means of indicators. Different methods and all the indicators presented have been calculated for the whole territory of the Province and the 36 Municipalities that compose it. Finally, the results of the project (the results of this project will be published within a Special Issue of the Journal of Environmental Management in the course of 2006) consist in a series of indicators (26) calculated for a series of comparable territorial systems (36).

In this context, the Principal Component Analysis (PCA) has been applied in order to analyse relationships, as well as possible overlapping and complementarities among indicators. PCA is a mathematical transformation in which linear combinations of the input variables (here indicators) are created; the new variables, called principal components (PCs), explain as much of the variation as possible of the original data.

PCA has a twofold function: 1) it provides a way of reducing the dimensionality of the data; 2) it is a powerful visualization tool that enables graphic representation of intersample and intervariable relationships for exploratory data analysis. In other words, it plays both a methodological role, as a statistical computation method, and a practical one, presenting an overview of the contribution of each indicator in terms of information. However, this does not mean that PCA is a further instrument of synthesis, because combining information is often tantamount to a loss of information. Notwithstanding, it can be convenient in order to optimize, even from an economic point of view, the combination of indicators necessary to know some aspects of an anthropic system because "a large number of original variables are reduced to a small number of transformed variables" (Saisana and Tarantola, [7], p. 12).

The results of the PCA within the SPIn-Eco Project (see Bastianoni et al., [1]) can be summarized in the following figures that show the Principal Component Loading Plot and a Principal Component Biplot. They represent the first two PCs with a cumulative explanation of 55.9%. Figure 1 shows the eigenvectors for the indicators used within SPIn-Eco Project and Figure 2 shows the plot of the 36 municipalities according to their similarity, utilizing combined information of the applied indicators. The characteristics of different sites can be highlighted by their position: according to the PC1, on the left side (high value of P, PD, I, CO_2 and F) we can find the largest urban areas, for example the city of Siena (32), whereas the municipalities on the left side (high value of BC, ITS, NDVI and CO_2 absorbed) are characterized by agriculture, forests and low population density (Monticiano, 18; Radicondoli, 25). PC2 isolated three sites with high value of N, EYR, ELR, EpP, Em and EmD, because of the exploitation of local non-renewable environmental resources, namely yellow marble (Sovicille, 34), clay (Trequanda, 36) and travertine (Rapolano, 26).

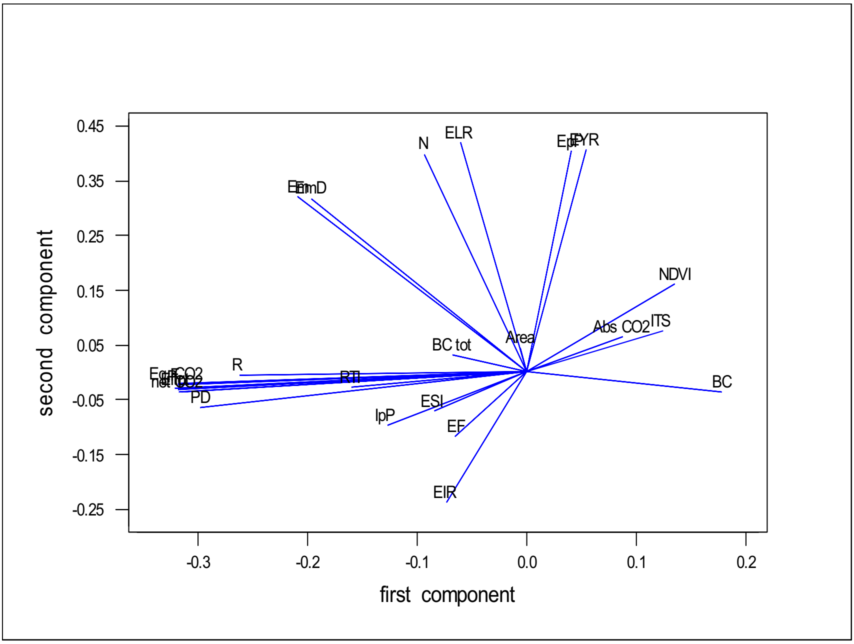


Figure 1: Principal Components Loading Plot. On the left side, where several lines and symbols overlap, we can find: Population (P), Population Density (PD), Total Income (I), Local Renewable Resources (R), Purchased non-local Resources (F), Equivalent CO_2, Net CO_2, Total Ecological Footprint (EFtot) (source: Bastianoni et al., [1]).

This analysis can be a reference point for planning (or orienting) human activities within a territory, since some new elements emerge such as homogeneity and diversity. In particular, the application of these results to the network of subsystems managed by the Province of Siena have shown that a certain degree of diversity among the sites is due to different environmental conditions or to different economic/urban/industrial characteristics. These diversities need to be preserved in order to reduce the fragility of the whole system with respect, for example, to possible economic crises in one particular sector.

From a methodological point of view, correlations and anti-correlations among indicators can be represented, on the basis of the position of the lines in the figures: at least for what concerns the first two PCs, the closer the direction of the lines, the more the indicators are correlated within this data set; a 90° angle means total independency, a 180° angle anti-correlation. High level of correlation exists among CO_{2eq} emissions, EFtot, F and I, typical expressions of the presence of humans (P). At the same time, the low level of correlation between EF and EpP is quite surprising, since the methods of Emergy and Ecological Footprint seem similar. Actually, they calculate the environmental cost of human consumption, the former measuring it through the direct and indirect solar energy requirement, the latter through the amount of land needed.

However, the low level of correlation is explained by the fact that, while emergy evaluation computes the non-renewable materials extracted locally (N), EF does not consider these types of materials, preferring to use renewable substitutes when possible. Furthermore, while Emergy takes into account the total consumptions of goods within the area, Ecological Footprint includes only the household consumptions, without considering industrial consumptions and similar.

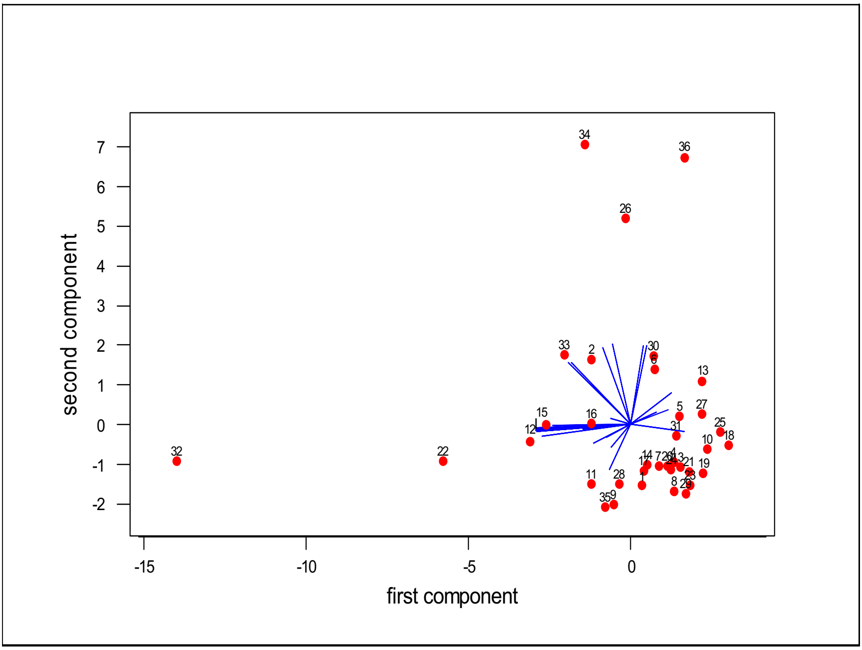


Figure 2: Principal Components Biplot. The numbers represent the Municipalities of the Province of Siena according to the following alphabetical order: 1 Abbadia S. Salvatore; 2 Asciano; 3 Buonconvento; 4 Casole; 5 Castellina in Chianti; 6 Castelnuovo Berardenga; 7 Castiglione d'Orcia; 8 Cetona; 9 Chianciano; 10 Chiusdino; 11 Chiusi; 12 Colle Val d'Elsa; 13 Gaiole in Chianti; 14 Montalcino; 15 Montepulciano; 16 Monteriggioni; 17 Monteroni; 18 Monticiano; 19 Murlo; 20 Piancastagnaio; 21 Pienza; 22 Poggibonsi; 23 Radda in Chianti; 24 Radicofani; 25 Radicondoli; 26 Rapolano; 27 San Casciano dei Bagni; 28 San Gimignano; 29 San Giovanni d'Asso; 30 San Quirico d'Orcia; 31 Sarteano; 32 Siena; 33 Sinalunga; 34 Sovicille; 35 Torrita; 36 Trequanda (source: Bastianoni et al., [1]).

The application of the PCA to a relevant number of results enabled this project to understand the degree of congruence of the indicators and the possibility of recognizing patterns or clusters in the description of the subsystems that compose the Province of Siena.

4 Conclusion

The use of synthetic indicators is a fundamental step within a sustainability assessment. They are the results of the application of suitable methods useful for the analysis of some aspects of the complex relationship between human behaviour and the environment. A reliable and available set of data must be collected and then computed according to rigorous procedures, preferably based on physical measures that are objective and not arbitrary. Further statistical computation (for instance PCA) may be of help in order to select the best ensemble of indicators to calculate. Statistical information plays an important role in articulating a research project whose purpose is to assess the sustainability of one or more systems. For this reason, public administrations should be requested to invest more resources in acquiring information on the state of the system as well as in scientific means in order to understand the sense of their policies from the point of view of sustainability.

References

[1] Bastianoni, S., Gramatica, P., Pulselli, F.M., Focardi, S., Tiezzi, E.B.P., 2006. Correlations and complementarities in data and methods through Principal Components Analysis (PCA) applied to the results of the SPIn-Eco project. Journal of Environmental Management, in press.

[2] Focardi, S., Loiselle, S., Mazzuoli, S., Bracchini, L., Dattilo, A.M., Rossi, C., 2006. Satellite based indices in the analysis of land cover for municipalities in the Province of Siena. Journal of Environmental Management, in press.

[3] IPCC, 1996. Greenhouse Gas Inventory. Reporting Instructions, 1996 (revised). IPCC, Intergovernmental Panel on Climate Change, 1,2,3.

[4] Odum, H.T., 1988. Self organisation, transformity and information. Science, 242, 1132-1139.

[5] Odum, H.T., 1996. Environmental Accounting. Emergy and Environmental Decision Making. John Wiley and Sons, New York.

[6] Ruth, M., 1993. Integrating economics, ecology and thermodynamics. Kluwer Academic Publishers, Dordrecht, Netherlands, 251 pp.

[7] Saisana, M. and Tarantola, S., 2002. State-of-the-art report on current methodologies and practices for composite indicator development. Report EUR 20408 EN. European Commission-Joint Research Centre, Ispra.

[8] Ulgiati, S., Odum, H.T. and Bastianoni, S., 1994. Emergy use, environmental loading and sustainability. An emergy analysis of Italy. Ecological Modelling, 73, 215-268.

[9] Wackernagel, M., Rees, W., 1996. The Ecological Footprint: reducing human impact on the Earth, New Society, Gabriola Island, BC, Canada.

[10] UNCED, 1992. Agenda 21, Chapter 40, Information for decision making. Rio de Janeiro 3rd-14th June, 1992.

A study to designate a National Land Sustainable Plan in Japan: focus on land recycling

M. Taniguchi
Graduate School of Environmental Science, Okayama University, Japan

Abstract

Although expansion of urban areas has been promoted by rapid population growth in Japan, the lands that have been deserted as a result of decreased population and office space have been increasing lately in various locations throughout the country. Urban layouts with scattered disused lands are not only inefficient, they are also undesirable in terms of sustainability. Furthermore, issues of how to take measures to meet such a decreasing societal capacity as Japan is facing will be shared among more and more countries. To cope with that decreasing trend appropriately, Japan has proposed a guideline, called the "National Land Sustainable Plan", for nationwide space management. The present study elucidates phenomena that reveal themselves now mainly from the viewpoint of land recycling promotion, actual conditions in the "National Land Sustainable Plan", and similar future plans.
Keywords: National Land Sustainable Plan, urban space recycle, ecological footprint.

1 Introduction

The policy motivation of space use of national lands should vary generally depending on a country's economic development level. During early stages, with low economic activity, a nation should first establish plans to promote economic activity to make the local community affluent. On the other hand, as the society matures as a result of their promoted economic activity, quality of life issues and concern for sustainability should dominate the planning of policy-makers. Japan has experienced exactly such a change of planning ideas over the past five decades.

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060161

This study is intended to: 1) outline the process of establishing the "National Land Sustainable Plan" in Japan; 2) summarize matters that are being considered in the "National Land Sustainable Plan" at the present stage; 3) examine the actual conditions of present space recycling in Japan; and 4) evaluate environmental independence with particular emphasis on space use on a local scale with the ecological footprint index. Finally, based on the findings obtained during the process of 3) and 4), this study will elucidate how the "National Land Sustainable Plan" might be improved.

2 Process of establishing the "National Land Sustainable Plan"

After Japan's defeat in World War II, the national government has promoted the development of industry and national lands under the "Comprehensive National Development Plan." That plan's contents have been revised five times since the first plan was formulated in 1962. It has indicated basic directions of policies to resolve problems facing the Japanese nation in each era, especially urban and local problems resulting respectively from overcrowding and depopulation. It has achieved results related to decentralization of the location of the manufacturing industry and the narrowing of income disparities among regions.

The following are inferred to be the results of the "Comprehensive National Development Plan":

1) Decentralization of industry, education, and other activities
2) Growth of central and core cities
3) Reducing the influx of population into metropolitan areas
4) Preventing pollution and reducing congestion in cities
5) Improvement of living environments in local areas

The following, in turn, are pointed out as issues that remain unsolved:

1) Centralization of urban functions and population in the Tokyo Metropolitan Area and the Tokaido region
2) Increase of depopulated areas
3) Hollowing of central urban areas in local cities
4) Overpopulated areas in large cities left unimproved
5) Lack of concern for a good landscape
6) Soil pollution, water pollution, and illegal dumping as social problems

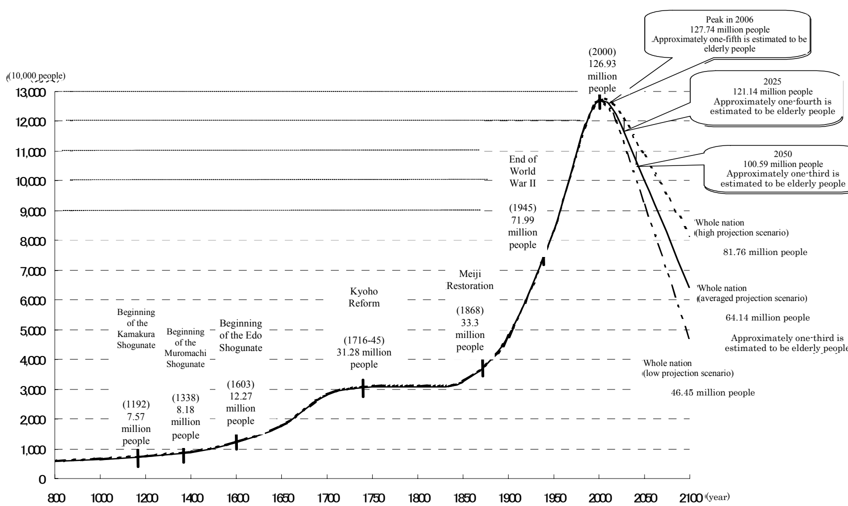
Viewing the long-term change of Japanese population, as shown in Fig. 1, the population has increased continuously since the "Comprehensive National Development Plan", mentioned above, came into force. Nevertheless, the Japanese population apparently peaked in 2006, and has already started to decline.

3 Issues for examination in the "National Land Sustainable Plan"

Taking into account the background described above, Japanese national and local administrations have been working cooperatively to promote a new land plan for

sustainability instead of the development-oriented "National General Development Program." The "National Land Sustainable Plan" is a comprehensive and basic plan to promote utilization, improvement and conservation of national lands (hereafter referred to as the sustainable formation of national lands), concerning the following items:

1) Use and conservation of land, water and other natural resources of national lands
2) Use and conservation of water areas
3) Prevention and reduction of earthquake disasters, water damage, wind damage, and other disasters
4) Coordination and development of the scale and location of cities and agricultural, mountainous and fisheries villages
5) Proper location of industries
6) Use, development and conservation of transportation facilities, information and communication facilities, research institutes concerning technology and other important public facilities
7) Protection of resources related to culture, welfare, and tourism, and the use and development of related facilities
8) Creation of a good environment, the conservation of other environments, and the creation of beautiful scenery



Source: Prepared by the National and Regional Planning Bureau, Ministry of Land Infrastructure and Transport, according to "National Census of Japan," "Annual Report on Current Population Estimates" by the Ministry of Internal Affairs and Communications, "Overview of Future Trends in Population in Japan" (conducted January 2002) by the National Institute of Population and Social Security Research, and "Long-term Time Series Analysis on Changes in Population Distribution throughout the Japanese Archipelago" by the National Land Agency (1974).

Figure 1: Long-term trend of the total population in Japan.

Above all, the following are shown as examples of problems that are now regarded as important in terms of national land policy and are being investigated.

a) Expansion of areas that have difficulty in providing fundamental social services

It is presumed that the percentage of small towns with populations of fewer than 5,000 will increase in the future. The data are illustrative: there were 722 towns and villages of less than 5,000 population in 2000, accounting for 1.7% of the entire population and 22.9% of the total area of the nation. In 2050, those figures are estimated to leap respectively to 3.3% and 43.2%. In other words, half of all of the national land is expected to be occupied by those small towns and villages. In that situation, it might be difficult to provide even such basic social services as water supply and sewer systems, elementary schools, fire fighting, medical care, etc.

b) Increased cultivation of abandoned lands

Abandoned arable lands increased by 1.4 times during the five years of 1995–2000. The area of such lands is greater than 340,000 ha, equivalent to 1.5 times the area of the Tokyo metropolitan area. Simultaneously, Japan depends on other countries for many agricultural items, which would require 12 million ha of arable land for their production. In addition, regarding forestry, six of ten forestry households own private forests that must be thinned out by owners. Those forests are considered to be virtually abandoned. Moreover, 80% of the timber consumed in Japan is imported from overseas.

c) Increase in environmental loads

Greenhouse gas emissions in Japan are 1.3 times those of ten years ago. Dumped wastes have actually increased. The loss of biodiversity also presents a serious issue. It is said that 2,663 species inhabiting Japan are in danger of extinction. Therefore, it is necessary to devote greater attention to future development of urban areas.

4 Actual conditions of space recycling

4.1 Preconditions of analysis

Although a "National Land Sustainable Plan" is being produced on the basis of the various data mentioned above, few sufficient examinations into the issues have been advanced as to how space is reused in an area that has already been urbanized. For that reason, this study shows concretely that the generation of such wasted space is not recycled in the actual case of an urban area. As analyses, both micro and macro methods of examination are used, which respectively address the lot level and grid levels. In addition, no standard of efficiency to assess deserted lands exists because no study has adopted such a viewpoint. Therefore, the micro examination presented in this study originally establishes a standard to evaluate unused land, as shown in Fig. 2, by which a lot that is left as a vacant site or a vacant house is defined as a deserted space.

4.2 Space recycling at the micro level

As a concrete sample area, this study presents examination of a suburban arterial road running in Okayama City, a typical district capital. To ascertain the actual conditions of urban space waste and recycling, the author followed up the change

of land use of the area along the road closely on a micro scale from 1980 to the present. Concretely, the survey was conducted over the entire area of a section 4.5–18 km from the center of the city along national Route 2. For measurement, the land along the road was divided into lots as the minimum unit of a land with a building, to illustrate land use (classified into 18 kinds) of each lot using past housing maps and present findings of on-site studies.

Classification 1(Under construction)

Classification 2(For sale and rent)

Classification 3(Vacant houses)

Classification 4(Vacant sites)

Figure 2: Definition of urban space waste and recycle in unused land.

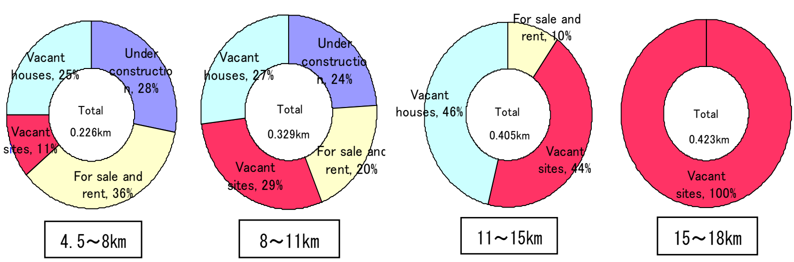


Figure 3: Waste and recycle states of “unused lots”.

The tendency throughout the sample area is that the amount of unused land decreased substantially during 1980–1995, meaning that development caused by suburbanization had advanced. Since 1995, in contrast, unused lands have increased. Next, from the viewpoint of space recycling, Fig. 3 depicts the number of present disused spaces such as “vacant lands” or “vacant houses,” by

applying the classifications shown in Fig. 2 exclusively to the lots in which facilities used to exist but which are now disused. Then, the classified total was computed by distance from the city center. Results showed that the absolute quantity of disused land that must be managed from the viewpoint of space recycling is greater in distant suburbs than in nearer suburbs. In other words, the further one travels from the city center, the greater the ratio of space that is not recycled, but is instead left as disused.

4.3 Space recycling at the macro level

A macro-level examination was conducted of all commutation areas of Okayama City including results of the analyses presented above (5 cities, 14 towns, and 2 villages, with Okayama City and Kurashiki City as the core cities). Main data used here include population according to the national census and the number of employees according to business and enterprise statistics in the local mesh statistics, and land use area of land-use grid data provided by national land numerical information as the land use index. Furthermore, in this study, the sum of the population and the number of employees is designated as the "active total population." The land use classifications of land use mesh data are aggregated into natural land use and urban land use; they can be compared to estimate the actual conditions of space recycling.

To verify at the macro level whether space recycling has been actually promoted in the meshes where the activity has been on the decline, the change of land use was analyzed in districts where the "active total population" had decreased during 1985–1995. Concretely, the grid areas showing substantial decreases (decrease of "active total population" of more than 50) in the surveyed areas are taken as samples. Fig. 4 shows the spatial distribution of the change of natural land use in districts whose "active total population" has been declining. Results shows that the districts with decreased "active total population" are 422 grid areas, of which four grid areas are the districts with increased natural land use, 294 show no change, and 124 are the districts with decreased natural land use. In other words, although the "active total population" has decreased in most districts, the decrease of "active total population" is prominent in the developed districts, where land is simultaneously left as designated for urban use and abandoned.

5 The possibility of environmental independence in space use (from ecological footprint)

5.1 Preconditions of analyses

Results described in the previous passage indicate that urban land use is not converted into natural land use at all in a non-growing society. If so, then how much of an environmental load attributable to human activity can be absorbed? The ecological footprint (EF) index is one expression of the area that people exploit within (or beyond) biological productive space, such as farms and forests

providing food and timber, sites for social infrastructure like roads, forests that are converted into absorption of CO_2 emission promoting global warming, etc. The EF varies depending on the population scale, the average amount of consumption per capita, and the resource-intensive level of used technology. Although WWF [2] has already promoted examination of the EF index worldwide, there have been few accumulated studies that have examine it specifically at the local level. In this study, referring to the method of WWF, to calculate the environmental loads attributable to human activity, the land areas necessary in the following five items were determined in an original manner.

(1) Arable land to cultivate crops for food and feed (x=1)
(2) Pasture to feed livestock for meat and dairy products (x=2)
(3) Forest to provide timber and materials for paper manufacture (x=3)
(4) Land necessary to provide housing, industry and transportation facilities (x=4)
(5) Forest area necessary to fix discharged carbon dioxide (x=5)

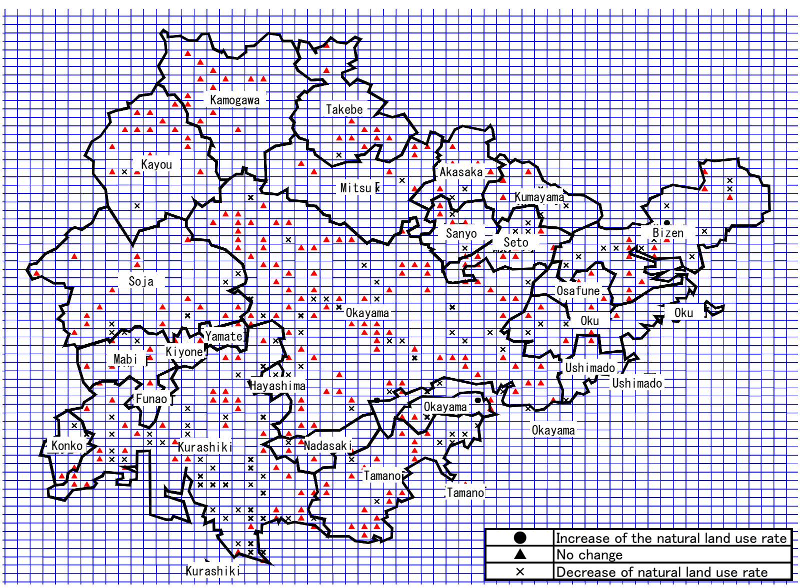


Figure 4: The change of land use at the macro level in grid areas in which "active total population" is on the decline.

5.2 Formulation

For this study, land areas that provide food and necessary amenities to live in each city, town, and village, are called "demand areas," whereas the land area producing food and other materials in this country are "supply area." The sum of the demand areas in each item is called the "environmental load"; that of supply areas in each item is called an "amount corresponding to the environment." The difference between both is called the "exceeded footprint." The exceeded

footprint represents the amount of environmental burden imparted by human systems. The exceeded footprint is defined using eq. (1) in terms of its total amount and net amount.

$$S^k = \sum_{x=1}^{5} S_{dx}{}^k - \sum_{x=1}^{5} S_{sx}{}^k \tag{1}$$

S^k: the total exceeded footprint of municipalities
$S_{dx}{}^k$: (x=1–5) is the total amount of environmental loads of k municipality
$S_{sx}{}^k$:(x=1–5) is the total amount equivalent to environment of k municipality

In addition, the amount of imported food and other materials from overseas, as converted into land area, is called the "import area," and the amount of that exported overseas, as converted into land area, is the "export area." The demand area from which the import area is subtracted is called "net environmental load"; the supply area from which the export area is subtracted is called "net amount equivalent to the environment," as defined in eq. (2).

$$S_d{}^{k\prime} = \sum_{x=1}^{5}\left(S_{dx}^k - A_{Ix}^k\right) \tag{2}$$

$$S_s{}^{k\prime} = \sum_{x=1}^{5}\left(S_{sx}^k - A_{Ex}^k\right) \tag{3}$$

$S_d{}^{k\prime}$: Net amount of environmental loads of k municipality (without import area)
$S_s{}^{k\prime}$: Net amount equivalent to environment of k municipality (without export area)
$A_{Ix}{}^k$: The import area of k municipality (x=1–5)
$A_{Ex}{}^k$: The export area of k municipality (x=1–5)

This total exceeded footprint represents the loads not only within the nation but also those extrapolated overseas. This is the land area including the amount that is allocated to producing exports to overseas, as converted into area, which is represented with eq. (1). The other net exceeded footprint represents only the loads within the nation, from which the amount exported to overseas is subtracted, defined as eq. (4).

$$S^{k\prime} = S_d^{k\prime} - S_s^{k\prime} \tag{4}$$

$S^{k\prime}$: Net total exceeded footprint of k municipality

5.3 Results of examination into the ecological footprint

Total and net EF excess rates are shown in Fig. 5. It is clarified that the total amount equivalent to the environment exceeds the total amount of the environmental load only in depopulated Kamogawa town, and that the other municipalities effectively *import land area* from overseas. As shown in Fig. 5, although the net EF excess rates distribute under 5.0 commonly among countries, total EF excess rates are substantially different. Although the total EF excess

rates in Okayama City, with a large population, and Kurashiki City with developed industrial sites are large, it is the smallest in Kamogawa town. Municipalities with net EF excess rates greater than 1.0, in other words, those that depend on other national municipalities, such as Okayama City, Kurashiki City, and others, account for some 70% of the areas targeted for the present analysis. The analytical results suggest that it is actually difficult to balance the occurrence of ecological loads with the absorption of ecological loads. To bring ecological accounts more closely into balance, as in Kamogawa town, it is necessary to advance the ideas of space recycling.

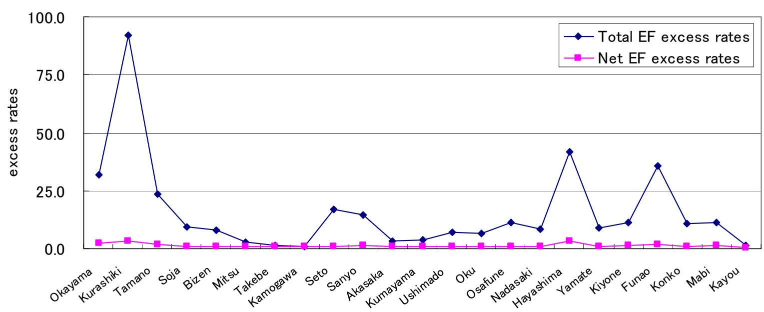


Figure 5: The total Ecological Footprint excess rates and the net Ecological Footprint excess rates by municipalities.

6 Conclusion

To realize sustainability in the "National Land Sustainable Plan" that is being made now, it is worth considering that the plan should have some legal force. For that purpose, it is desirable to have a set of convenient indexes by which to judge the degree to which the purpose has been achieved. To reduce external ecological loads to the greatest amount possible, it is extremely important to demand that each local administrative area calculate its EF index. Simultaneously, it might be also important to measure the effects of space recycling and thereby establish a target value of space recycling in each region using the EF index. A system of monitoring the state of achievement of a national land plan is necessary from the viewpoint of sustainability.

References

[1] National Planning Division, National and Regional Planning Bureau, Ministry of Land, Infrastructure and Transport, Japan: National Land Sustainability Planning System Reform, 2006.

[2] WWF Japan: Living Planet Report 2002 http://www.wwf.or.jp/lib/publication/downloadfiles/lpr/Final_LPR_2002_pp_01-36.pdf.

[3] Taniguchi, M., Abe, H. and Shigekane, K.: Prefectural Balance Sheet of the Environment: Study Based on the Ecological Footprint, Studies in Regional Science, Vol. 34, No. 1, pp. 23–36, 2004.

Application of Environmental Management Systems (EMS) to Natural Parks and reserves

F. Ardente, G. Beccali, M. Cellura & M. Fontana
DREAM (Dipartimento di Ricerche Energetiche ed AMbientali), Università degli Studi di Palermo, Italy

Abstract

Natural parks and reserves have to face problems concerning the conservation of their natural resources. Citizens and firms often perceive the institution of new protected areas more as a bind or a restriction to their activities than an added value of the territory. Following the principles of sustainable development, it is necessary to integrate the environmental protection with the needs of neighbouring urban areas and production sites. The implementation of an Environmental Management System (EMS) according to international standards (ISO 14001, EMAS) can support the Park administration in their complex management activities.

This paper analysed the criticisms related to the application of EMS to parks, with particular regards to the main difficulties during the certification process and the benefits and drawbacks related to the environmental certification.
Keywords: Environmental Management System (EMS), Natural Parks.

1 Introduction

Natural Parks represent an "atypical" company characterized by various and complex functions and aims, mainly related to:

- the saving, protection and management of natural resources;
- the offer of cultural and social services promoting scientific, educational and recreational activities addressed to different users.

It is then essential that the park's management will correctly define and apply adequate procedures to monitor and control the "state of health" of its territory, making it transparent to the citizens.

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060171

The introduction of an Environmental Managements System (EMS) to natural parks represents an engagement to grant a correct and sustainable fruition of the natural heritages. An EMS is defined as "the part of the overall management system that includes the organisational structure, planning activities, responsibilities, practices, procedures, processes and resources for developing, implementing, achieving, reviewing and maintaining the environmental policy" [EU, 2001]. EMSs have been regulated by the international standard ISO 14001 [1] and the European Regulation CE 761/01 "EMAS" [2].

The efficacy of an EMS can be improved also integrating it to other public initiatives for the sustainable development (as local Agenda XXI projects and the environmental accounting tools).

2 The role of the park's Responsible Body

The preservation of valuable natural subjects represents the main reason to the establishment of parks and reserves. The protected areas can also contain historical and artistic heritages or can involve standing or neighbouring build-up areas or productive plants. Depending on the dimensions, a park can involve a large and variable number of institutional subjects and public authorities. They cannot operate separately, but they have to state a common and synergic way to proceed and administrate the territory.

These elements make particularly complex the management of the park needing to meet the various social and economic interests of stakeholders.

For all these reasons is generally instituted a supervisor public body (the park's Responsible Body, successively named as RB) to manage the park [3]. The RB is generally directed by an administration council which includes political, scientific and technical staff.

The competences of the RB can include:

- *planning activities* to regulate the human activities and to control the evolution of natural habitats;
- *design activities*, including the protection of natural areas, the environmental recovery of degraded areas, the creation of new service structures;
- *watch*, to prevent and contrast environmental accidents and problems;
- *repressive actions* against outlaws activities;
- *issue* of new licences and authorisations;
- *administrative actions,* including management of the personnel;
- *management of park's facilities, services and buildings;*
- *formation and information* of citizens;
- *scientific researches*.

2.1 Implementation of an EMS to the activities of a park's RB

An EMS can support the management of a park, giving to the administration a structured, transparent and organised system to individuate, monitor and improve its significant environmental aspects. The implementation of an EMS becomes complex when it involves a large territory and a lot of different stakeholders (local authorities, towns or smaller settlements, industries, tourists, NGO, etc.).

In particular, the activities of private subjects could influence sensibly the quality of the park's environment and could be responsible of significant impacts. The RB cannot directly attend and modify these private processes but can influence indirectly their behaviours by means of control actions and restrictions. The relationships with external subjects are the crucial point for the success of the entire system [4].

In order to achieve the protection of the environment it is necessary to create a positive and collaborative atmosphere among all the subjects that operate inside the park, avoiding conflicting relationship and planning common shared objectives. Every stakeholder can be actively involved in the park's management system and they could positively contribute to the environmental improvement by providing new data and information, sharing the environmental targets and moving in a proactive way towards such objectives.

The EMS works at two different levels:

- the inner improvement of the environmental management of the RB, that has to reduce its impacts and to control its activities;
- the external improvement of all the controlled territory, inducing also external actors to a more environmentally friendly behaviour.

The key points about the application of an EMS to natural parks are following summarised:

- The EMS has to be applied to the park's RB or an equivalent directive structure. This body can use EMAS logo;
- The environmental certification should involve all the activities carried out by the RB that can directly or indirectly impact on the managed territory;
- The RB has to carefully evaluate the sensible and vulnerable natural targets, evaluating the quality status of the habitats and ecosystems. The study should also include a socio-economic analysis regarding both the protected and the neighbouring areas in order to avoid that adjacent productive activities that could threaten the park with their environmental impacts;
- The improving actions shall focus on the RB's competences but should indirectly involve all the activities insisting into the park;
- Targets and aims have to refers to the entire protected territory;
- People that operate in the park have to modify their actions in order to accomplish with the park's guidelines;
- The RB can influence external subjects, i.e. by supporting and financing environmentally friendly activities, by fixing restrictions to companies that operated in the areas or granting special licenses and authorisations only to "green" companies;
- The RB can qualify the external subjects by means of an environmental park's logo that points out and awards the eco-compatible companies. It is however desirable that these subjects will successively join the EMAS certification themselves [5].

2.2 EMS applied to Natural Parks: state of art

Nowadays in Italy, the certifications of natural parks include 11 organisations certified by ISO 14001 [6] and 6 organisations certified by EMAS scheme [7].

We can observe a great incidence of EMAS registrations. In fact, this scheme grants visibility and large image benefits and, compared to the ISO scheme, it could be considered as more suitable for the certification of an extended territory.

In Europe, we assisted to a generalised delay in the EMAS certification of parks: only four parks have been certified, one in Denmark and Spain and other two in Germany [7].

The greater diffusion in Italy can be probably related to the Italian "Law on the protected areas" (law n° 394/91) that regulates the management of parks and set up the special competences, powers and responsibilities of RB. As requested by the law, the park's management has to accomplish to several planning activities (including the plans for the socio-economic development of he area) that partially satisfy to the EMAS requirement of management of indirect and planning aspects [8]. It is then advisable that in future such typology of laws would be strengthened and boosted.

However, the legislation generally does not prescribe the direct involvement of other external subjects. This could instead represent the added value related to the implementation of an EMS.

Table 1: Environmental aspects.

Environmental aspects	**Related improving environmental objectives**
Effects on the biodiversity	Maintenance and protection of the habitats
	Sustainable exploitation of the natural heritage; Location of great and old trees; Increase of surfaces maintenance; Forestry; Study of the composition of the floristic and faunal heritage (electronic databases); Protection of pasture areas (delimitation and regulation); Protection of coastal and marine bands; Reduction of the coastal erosion; Incentive to the substitution of old plants.
Administrative and planning decisions	Survey of concessions of the state demesne; Analysis of commercial and productive activities/infrastructures in the territory; Monitoring of territory.
	Partnership of RB with other public authorities; Establishment of autonomous systems of surveillance and vigilance; Keeping of registers (tourist presences, sanctions, concessions, fauna); Regulation of sustainable tourism
	Electronic cartography; Creation of Geographical Information System (GIS)
Risk of environmental accidents	Fire and flood prevention plans; Maintenance of woods; Reduction of hydrogeology risks; Shared procedures with Forest Bodies and Civil Protection
Socio-economic promotion	Promotion of a Park quality mark; Certification of the local firms; Consultation of local authorities and stakeholders to define shared initiatives.
Mobility	Sustainable mobility plans for the tourists and villagers; diffusion and incentives to low-impacts transportations.
Diffusion of environmental information	Environmental communication for tourists and residents; Organization of events, publications, brochure, internet sites, videos, posters, etc.); Establishment of information points; Creation of naturalistic paths; promotion of environmental certifications and of the EMAS logo;
Waste management	Waste recycling; Installation of "ecological islands"
Waste water	Analysis and monitoring of surface and deep water bodies; Purification of waste waters with natural systems

2.2.1 Environmental aspects, objectives and targets

From the analysis of the Environmental Statements of registered parks [9–12], it has been selected a list of significant environmental aspects, objectives and targets. These data, summarised in Table 1, can represent a useful guideline for parks that would undertake the route to the environmental certification.

3 Peculiarities and criticises in the implementation of an EMS

The management of a large territory has peculiarities that differ from the normal management of a company. Besides the grant of services, the RB has specific planning responsibilities, granting the quality improvement of the environment and the conservation of natural resources. It is then necessary that the park's EMS would take into consideration all the environmental aspects of the territory, both caused by the direct park's management or by other external subjects [13]. The RB has to improve the efficacy and the efficiency of its planning and to control all the actions that can have an environmental influence on the territory.

It is desirable that the RB would implement an EMS to its management and, contemporary, would drive the private subjects to develop an own EMS. However, this approach is not easily practicable because some companies could perceive the certification as a compulsion and a restriction of their free enterprise. Table 2 describes the elements of the EMS and the criticisms related to the application to a natural park [14].

Stakeholders have the twofold role of controller and controlled: they, in fact, verify that the territory management would be eco-compatible and respectful of their necessities; on the other side, private subjects have to follows the park's recommendations and guidelines and to modify their activities in order to reduce their significant environmental impacts.

Being the EMS a voluntary tool, it is important that the improvement actions would be discussed and approved by all the interested subjects. Cooperation among stakeholders has to be activated just in the early stages of the certification, during the individuation of environmental significant aspects. Similarly to the Agenda 21 approach, the creation of forums and public meetings can assist the RB in the preliminary stage of certification [4]. These forums are not mandatory but they are desirable in order to share a common way to proceed and to avoid future contrasts and hostilities.

4 Benefits and drawbacks

A natural park can largely benefit about the implementation of a certified EMS. First of all, the data survey helps the organisation to acquire awareness of its activities and their repercussions on the environment as much as the vulnerable environmental elements of the territory.

Second, the EMS helps to systematically organise and control the park's competences and the decisional processes. The detailed IEA gives a figure of the effects and the consequences of park's action and planning activities. The EMS allows also to identify key roles and hot spots and to clearly assign the management responsibilities [15].

Table 2: EMAS requirements for EMS and related critical aspects.

Requirement	Critical aspects
Environmental Policy	The Policy has to include two mandatory engagements: the legal compliance and the undertaking to the continuous environmental improvement. The Policy can then refer to future general engagements as: protection of habitats and biodiversity, reduction of environmental impacts in the territory, sustainable fruition of citizens and tourist. The park's RB ratifies this document without necessarily involving citizens and external actors.
Environmental aspects	The EMAS regulation fixed that the company's top management would individuate every direct and indirect activity that could produce significant environmental impacts. The initial data acquiring process is summarised in a document named *Initial Environmental Analysis* (*IEA*). Being the preservation of natural habitats the main scope of park's administration, the collection of update data and information about the status of the environment is an initial prerogative as much as a future target. The environment monitoring reveals itself as a critical matter, a more critical point compared to the registration process competing to other private organisations. The IEA has to regards not only the RB but all the activities insisting in the park and in the neighbouring areas. The IEA aims to individuate: Natural characteristics (insisting flora, fauna, geomorphology, hydrology, landscape, etc.); Sensible and vulnerable targets; Activities and responsibilities of the RB; Productive and tourist activities; Socio-economic status of the studied area. It is import to underline that the control of indirect aspects is generally more important and complex than the direct impact produced by the RB. During the IEA it is possible to select significant environmental indicators and indexes that allow a continuous monitoring of the environment, its time evolution and the efficacy of undertaken improving actions. Aspects to take into account are: air and water emissions; waste management; natural resource depletion; use of the soil; use of chemicals; noise and vibrations; biodiversity; quality of natural habitats and ecosystems; flora and fauna conservation; the monitoring and the control of productive activities and urbanized areas.
Legal and other requirements	The EMS has to be centred to the respect of all the compulsory laws and regulations: general government laws, regional and local directives, special laws (as the special directives that set up the parks and define the competence of the RB). It is suggested to develop procedures and instructions in order to regularly update the register of official instructions and in order to control the respect of deadlines.
Objectives, targets and environmental programme	The definition of targets is probably the most important and critical matter. It applies not only on the inner activities of the RB, but also to activities of external subjects operating in the controlled territory. Objectives and targets have to be in accordance with the policy's statements. The targets involve also the design and planning activities of the RB. It is necessary that this stage would be responsibly carried out, evaluating in detail the effects of planned actions and the spin-off on the territory. By means of feasible and suitable plans the RB can control indirect environmental aspects related to the citizens and private company's activities. The efficacy of plans is also related to a strict control of the respect of the fixed requirements. Objective and targets are summarised in the environmental programme that includes a precise description of times, responsibilities and employed resources.
Structures and responsibilities	The managerial structure of the park should clearly define competences of managers and workers regarding their role and responsibilities in the EMS. Furthermore, the RB should designate intermediary bodies for the relationships with stakeholders to assist the RB during: data survey, location of environmental criticises, definition of environmental programmes, communications, monitoring and auditing activities.

Table 2: Continued.

Training, awareness and competence	The park's management should foresee training courses on those subjects that could influence the park's environmental management. First of all, the training has to concern workers. They have to acquire awareness of the environmental impacts related to their activities, the best practises and procedures to minimise impacts and their competences. The training has to be yearly updated, checked and documented. Contemporary, the RB has to improve the environmental awareness of external subjects that live or operate in the park's territory. They should be adequately informed and awaken about their role in the EMS.
Communication	The RB shall foresee apposite procedure to control of communications addressed to internal (workers, persons in charge, contracting firms) and external subjects (public authorities, private companies, citizens). Communication can regard: modifies of park's policy and programmes, new laws and regulations, environmental criticises, internal and external procedures, publication of data and environmental statements, emergencies, etc. The communication channels are also useful to acquire information from external subjects as: data, improvement suggestions, complaints and problems.
Document control	The control applies to internal and external documents. It aims to grant that all the significant records would be correctly renewed, stored and available where necessary. Due to the large number of documents and in order to increase the efficiency of the control, the park's RB should computerise this process.
Operational control	The operational control regards the monitoring of the environment quality by means of Environmental Condition Indicators (ECI) about the entire territory, and by means of Performance Indicators (PI) concerning the management of the park and of the other private companies operating in the territory. This phase is carried out by means of suitable procedures and responsibilities in order to control all the significant environmental aspects identified during the IEA The EMAS regulation also foresees that the organisation would fix and periodically update procedures for the selection of suppliers and contractors. These subjects should be qualified on the basis of environmental criteria and requisites fixed in collaboration with all the stakeholders.
Emergency preparedness and response	The RB has to locate possible states of emergency. They include natural accidents (as fire, flood, earthquake, etc.) as much as anomalies of the normal activities. Emergences belonging to the first type have to be faces thanks to a detailed analysis of the territory, with a constantly update data acquiring system; to face emergences of the second type it is necessary the collaboration of private subjects operating in the territory, in order to avoid environmental problems due, for example, to accidental releases of toxic substances or the halt of purification plants. Each emergency has to be monitored and apposite procedures are required to face the emergency and to alert citizens and other involved subjects.
Monitoring and measurement	This phase aims to acquire data useful to the selection of significant environmental aspects, to control the evolution of habitats and to the accomplishment of targets and objectives.
Non-conformance and corrective and preventive action	Nonconformances represent anomalous situations detected during audits and monitoring. Nonconformances require immediate actions, in order to prevent accidents or serious environmental problems. The RB should also undertake initiatives in order to prevent nonconformances.
Records	All the significant activities have to be recorded in order to grant an adequate evidence of the park's data and activities.

Table 2: Continued.

Environmental management system audit	During this phase, the EMS is checked in order to locate the non-conformances. The audits are also important to locate strong and weak point of the systems and the sectors with the higher improvement margins. External accredited verifiers control the activities of the RB. The park's management has also to plan internal audits in order to check its internal activities as much as the behaviour of other subjects operating in the territory.
Management review	This phase is carried out by the park's management. It is required a detailed review of all the previous steps and, in particular, of nonconformances in order to verify the adequacy and efficacy of corrective actions.

Other benefits related to environmental certification are:

- to improve the environmental efficiency (reduction of consumption of energy and natural resources, improvement of the environment quality, employment of plants and products with lower environmental impacts, diffusion of environmental information, realisation of actions for the environment recovery);
- to accomplish with the laws in force and the deadlines;
- to assign organically authorisations and concessions to private companies, in order to globally minimise environmental impacts;
- to monitor the quality and the evolution of natural habitats;
- to control the accomplishment of the environmental programme and the efficacy of environmental actions and plans;
- to set environmental indexes and indicators (useful to study, compare and communicate the park's performances);
- to improve the park's image [4].

The RB will also improve its credibility, giving to citizens and other organisations a transparent figure of the park's management. In this way, the undertaken actions and the imposed restrictions will be more easily accepted. In particular, the publication of the Environmental Statement increases the visibility of the park, showing the engagement for the environment and the obtained results.

The involvement of other stakeholders, such as public authorities and companies, allows one to avoid conflictual relationships among private and public subjects. The development of shared action programmes and planning make possible to companies to perceive the park's EMS as an added value for the territory. The certification could have a spin-off effect on the companies that could so link their name and production to the environmental quality of the territory [16].

Firms could then have economic benefits due to: resources saving, the improvement of waste management, the purchasing of sensible clients that will reward their ecological productions.

5 Conclusions

The management of a natural park includes a lot of different and difficult activities. The development of an EMS in accordance with the international

standards (EMAS, ISO14001) allows one to organically and systematically manage and control all these activities. In particular, the added value of the EMS is related to the detailed selection of environmental criticises and the clear definition of rules and responsibilities within the park. The development of an internal monitoring system allows one to control the evolution of environmental performances and the efficacy and efficiency of undertaken corrective actions and the accomplishment of environmental targets and objectives.

Finally the publication of the Environmental Statement grants the transparency of park's activities and improves the relationships with citizens, NGO, public authorities and other stakeholders.

The main difficulties of the certification process are related to the role of private subjects that the park's management cannot directly control. It is then necessary to involve these actors just during the planning activities, in order to define shared objectives and procedures.

Benefits are both environmental (due to the correct management of significant environmental aspects) and economic (due to the optimisation of resource use). All the territory could benefit of an image improvement with a spin-off effect to the tourist activities in the area and the promotion of local productions.

References

[1] International Standard ISO 14001 "Environmental Management systems — Requirements with guidance for use". Second edition 2004.

[2] Regulation (EC) No 761/2001 of the European Parliament and of the Council allowing voluntary participation by organizations in a Community eco-management and audit scheme (EMAS).

[3] Province of Turin, WWF Piemonte. "Implementation of EMS to Protected Areas - Guidelines for the application of ISO 14001 to the Provincial Park of the Lake of Candia -. Turin. (In Italian).

[4] Naviglio L., Adamoli R., Bruzzesi F., Castorina M., Minciardi M. R., Morgana J. G., Rossi G., Paci S. "The EMS as a tool for the promotion and the development in the parks and reserves", ENEA, Second National Conference of Natural Protected Areas, 11-13 October 2002 Turin. (In Italian).

[5] Molinas P. –" EMAS registration of "Territory": The sample of Parks and natural reserves" - Agenzia per la Protezione dell'Ambiente e per i Servizi Tecnici (APAT), Ambiente e Sviluppo 2002. (In Italian).

[6] SINCERT website www.sincert.it.

[7] European register of EMAS website www.europa.eu/emas.

[8] Italian Law 6.12.1991 n. 394 *"Legge quadro sulle Aree Protette"*.

[9] Environmental Statement 2004 -Protected Marine Area "Tavolara – Punta Coda Cavallo"- Olbia (SS) Italy. (In Italian).

[10] Environmental Statement 2003 "National Park of Dolomiti Bellunesi"- Feltre (BL) Italy. (In Italian).

[11] Environmental Statement 2004 -Protected Marine Area and Natural Reserve "Torre Guaceto"- Brindisi, Italy. (In Italian).

[12] Environmental Statement 2003 –Natural Park of Monte Avic-Champdepraz (AO) Italy. (In Italian).
[13] ENEA-SINCERT-"Application of ISO 14001 in the protected areas", Environmental management: Guideline 1, Milan 1-109. (In Italian).
[14] Paci S., Naviglio L. "The problem of application of an Environmental Management System in protected natural areas. ENEA October 2000 (In Italian).
[15] Naviglio L. "The certification in parks and natural reserves". Valutazione Ambientale n. 4, December 2003. (In Italian).
[16] Naviglio L. ENEA "Parks in quality: The Environmental Management System in protected areas". Parco Produce 2001- 24 November 2001 Ancona, Italy. (In Italian).

Using the full potential: regional planning based on local potentials and exergy

A. van den Dobbelsteen[1], R. Roggema[2], K. Stegenga[3] & S. Slabbers[4]
[1]*Delft University of Technology, The Netherlands*
[2]*Province of Groningen / Groningen University, The Netherlands*
[3]*Stegenga, Workshop for Urban Planning, Tilburg, The Netherlands*
[4]*Bosch Slabbers Landscape Architects, The Hague, The Netherlands*

Abstract

Climate change and the depletion of fossil energy resources pose a serious threat to many parts of the world. The northern region of the Netherlands, partly below sea level and currently depending importantly on the exploitation of natural gas, is a clear example of this. This paper discusses the studies undertaken to propose a radical change to the region, leading to a different approach to spatial planning, based on local potentials and the exergy principle. So far, energy has hardly been an influence in regional and urban planning. Nevertheless, smart deployment of local potentials for energy provision, based on the vernacular characteristics, will mean a new paradigm for spatial planning. Thorough analysis of the local climatic, landscape, geophysical, cultural, and technical characteristics with a focus on energy potential leads to better insight into the most effective energy provision opportunities in the areas of a region. In addition, the concept of low-exergy design – based on using high-quality energy for high-grade processes, and re-using the residual waste energy from these in lower-grade processes – will lead to serving more processes and functions by the same amount of primary energy, implying significant benefits for sustainability. Converting the deployment of local potentials and low-exergy design to regional planning, the spatial arrangement of functions will differ from the conventional approach. The new approach will be exemplified by the Grounds for Change project and its mapping results for the Northern Netherlands.
Keywords: climate change, local potentials, energy, exergy, spatial planning.

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060181

1 Introduction

1.1 Climate change

Climate change is a hot issue, literally. The most probable scenario for the global climate is based on the widely supported findings and expectations by the International Panel on Climate Change [1]. In this scenario, within this century, the earth heats up a few degrees. Eternal snow, glaciers and icecaps will melt. Just as by the expansion of the warmer sea water, the sea levels will rise due to this increased run-off of water on the land. This will lead to more and heavier storms. There will be more clouding. A greater difference between wet and arid areas will evolve.

There has been wide discussion about to which extent human actions are to be blamed for climate change. Evidence is stacking up, but whether or not caused by mankind, the developments described are already taking place, and a quick response is necessary.

1.2 Energy depletion

While dramatic changes to our climate are taking place, one of its probable causes, consumption of fossil fuel, will become uncertain in the near future. Natural oil and gas reserves are depleting or at least their extraction is becoming more and more complex and expensive. The tendency to again use coal for the generation of electricity will have an even more devastating impact on greenhouse emissions, if not to mention pollutions that seemed to have been almost banished. Many countries now seek refuge to nuclear power but as long as safety issues and the processing of nuclear waste have not been tackled properly, questions may be raised about this. Moreover, sustainable energy is abundant: based on the potential for energy from the sun, wind and water, Jong *et al.* [2] calculated that the Netherlands alone could already provide the global economy with sufficient energy. However, at this instant local potentials are still insufficiently seized.

1.3 The Northern Netherlands

The problems discussed above in particular relate to the Northern region of the Netherlands. The three provinces of Frisia, Groningen and Drenthe are partly below sea level and rely significantly on the exploitation of natural gas, of which the reserves are expected to deplete within 25 years. In spite of the merits of gas trade, the region is economically relatively weak. This problem and the threats of climate change and energy depletion make the Northern Netherlands an interesting test case for drastic measures needed for a sustainable future.

2 Climate change and the Northern Netherlands

2.1 Dramatic developments

As in most places elsewhere, the most probable scenario of climate change for the Northern Netherlands implies a temperature rise by a few degrees. It will also

become wetter. The most important impact to region is the rise of the sea, most probably around 60 cm within this century. This means that in cases of storm and spring tide the level might be three to four times higher, implying a greater risk of flooding. This sea level rise adds to the increased run-off of water from the mainland, building up the pressure from fresh and salt water onto lower areas, especially those below sea level. For a specific ecologically unique area, the Wadden Sea, the sandbanks may never run dry again, diminishing the habitat of seals and foraging birds.

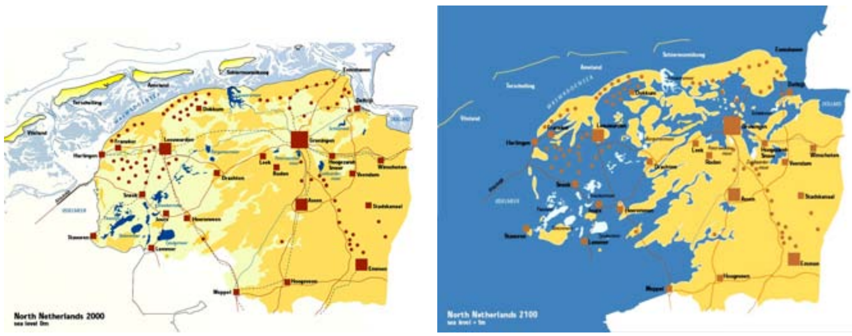


Figure 1: The Northern Netherlands at present (left) and around 2001, sea level risen by 1 meter, after a dike breakthrough (right).

As other climate scenarios are imaginable – for instance, an ocean current inversion leading to local cooling down instead of warming up – the safest policy seems to design a robust plan that can withstand different scenarios.

2.2 Possibilities to respond

The increased threat from the sea should be averted, or one should react to it actively. This can be established by making the sea dikes even higher, but the salty seepage will not be reduced by this, and the draining of polders will only aggravate this seepage, apart from the increased demand for energy for draining pumps. Another and perhaps effective strategy will be a layered defence partly interacting with the natural developments: additional defensive banks at sea and flooding areas behind the present dikes, using the deepest polders permanently for water retention. Furthermore, the Northern Netherlands can profit from climate change by touristic-recreative developments. Due to the rise of temperature, comparison with a more southern climate is realistic. Therefore, in the future less energy will be required for heating and hot water (heat), and more for cooling (electricity).

2.3 Grounds for Change

The Grounds for Change project was initiated to explore the possibilities of establishing a desirable yet imaginable future image of the Northern Netherlands.

It led to a new sustainable approach to regional planning and specific measures related to, for instance, water management, energy provision and habitats for living, working and leisure. The approach can be applied to any other region in the world. Based on the boundary conditions chosen, different images may evolve, which is demonstrated by the two reports published after Grounds for Change, Roggema *et al.* [3] and Noorman *et al.* [4]. There is not one single vision on future developments... This paper reflects the ideas of the regional design team of Grounds for Change.

3 A different approach

3.1 Another attitude

Rather than fearing the doom of climate change and energy depletion, they can be used also as a catalyst for sustainable redevelopment. Huge efforts can be put into the mitigation the effects expected, and this is very necessary indeed. However, one can also accept the most probable forecasts and use these in spatial planning. For the Northern Netherlands, prolonging 'business as usual' (enforcing the existing coastal defences and draining all polders) would cause too many problems to cope with; therefore, a new paradigm was found in cooperating with the sea rather than fighting it, and basing spatial planning on local energy potentials and in accordance with the exergy principle rather than extrapolating current developments.

3.2 Forecasting, backcasting and backtracking

First, in order to set the path for future sustainable developments, there are different time-based approaches.

- *Forecasting* is needed when we want to estimate the consequences of current developments and our own intervention on long-term effects, such as the issues of climate change and depletion of resources.
- *Backcasting* involves the description of a desired future state (sustainable and based on the needs that need to be met), translating this state back to strategies and measures we need to develop now.
- By means of *backtracking*, solutions are based on historical circumstances at the time when there still was a sustainable equilibrium. This sustainable past may be an instance for planning directions, for which perhaps certain valuable historical, natural or cultural features or circumstances can be brought back to new design.

Figure 2 clarifies the three approaches.

Forecasting is useful to predict trends that have developed over some time already, but is not very effective to establish a great leap toward a sustainable future. A better, more effective method for substantial change is backcasting, but one could consider this method rather detached from real life and how it evolved throughout the centuries. Therefore, backtracking can be used to link the present

to qualities of the past. None of the methods should be applied alone. We found that the simultaneous use of all three of them means synergy, picking the best solutions from history, the present, and the imagined and desirable future.

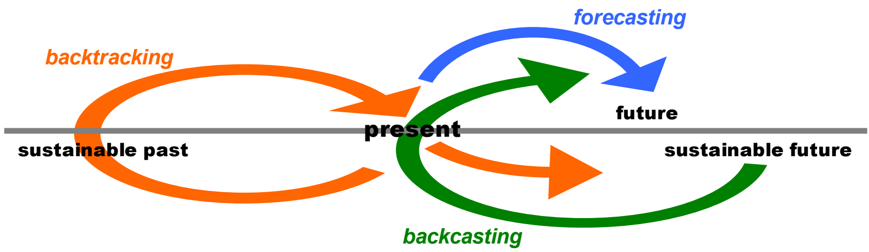


Figure 2: Graphical explanation of the forecasting, backcasting and backtracking methods.

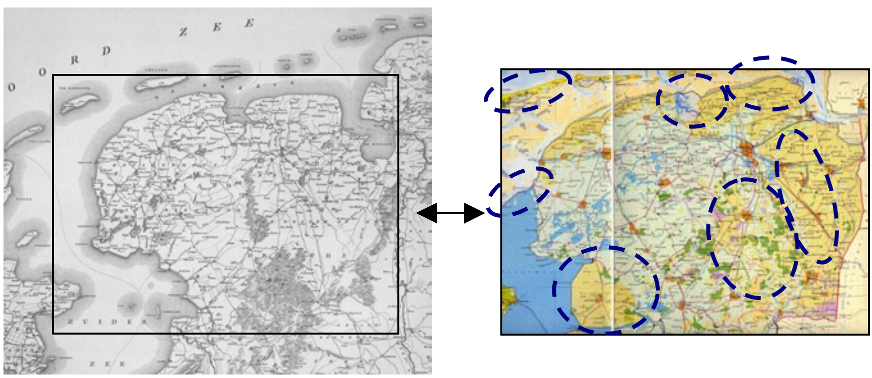


Figure 3: Map of the Northern Netherlands in 1810 and at present with some major changes.

3.3 Major changes to be expected

Around 1810, the landscape of the Northern Netherlands had hardly altered since centuries (Figure 3, left). After 1810, some major changes altered the regional landscape: containment of sea inlets and protective measures against the sea, polder extensions into the Wadden Sea, exploitation of the higher peat areas (the Peat Colonies), irrigation of the higher plateau through the introduction of artificial fertilisers, taking a whole new area from the sea (the North-East Polder), and the closure of the Southern Sea (IJssel Lake) that made this possible (Figure 3, right). Less visible is the exploitation of natural oil and gas.

4 Local potentials

Local strengths are usually neglected in regional and urban planning, especially with regard to local self-provision of energy. A different approach is supported

by energy potential maps. Analysis of the climatic conditions of the region, its geographical features and technical-cultural usage are among the issues to be studied to draw these maps. Some examples will be given here.

4.1 Energy potential maps for the Northern Netherlands

The wind potential map of Figure 4 shows the average wind force in the region (in Beaufort), indicating high potentials near the coast line. The sunshine potential map of Figure 4 gives mean sunshine hours, pointing out the areas most suitable for active or passive generation of solar energy.

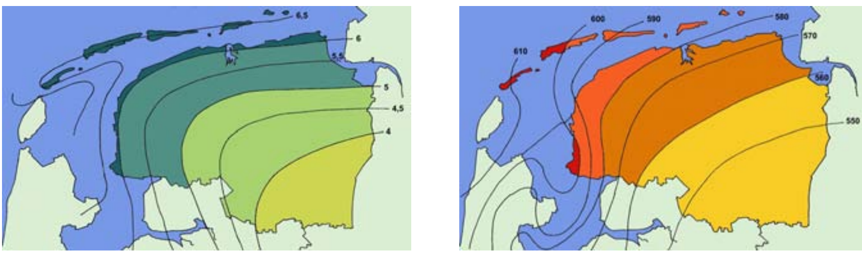


Figure 4: Average wind force (left) and average sunshine hours (right).

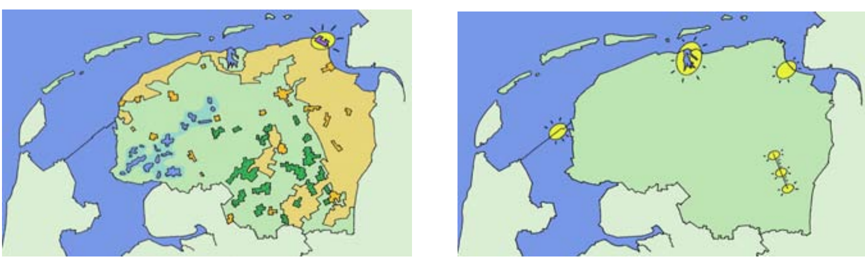


Figure 5: Location of biomass production (left) and water potentials (right).

Figure 5 gives the potential map for biomass, an interesting patchwork of agriculture, forests, lakeshore plants and domestic waste (produced in living areas), which can be made useful as local energy resource. Figure 5 depicts a few of the many opportunities for using water in the energy cycle: a drain plant at the IJssel Lake closure dike, a tidal plant in the former Lauwers sea inlet, osmosis plants at the borders of salt and fresh water, and plants and heat pumps making use of the great difference in height (seepage) at the edge of the Drenthe plateau.

Not presented here is the underground potential for gas and oil, hot geothermal water or storage of CO_2. An interesting opportunity is related to former gas drill holes, by means of which geothermal heat can be extracted, and around which therefore new areas for living may evolve in order to reduce heat transportation.

4.2 The energy mix map

When all potential maps are combined, an energy mix map evolves (Figure 6), indicating the energy resources most logical when optimally deploying the local potentials. This map can also be used for choosing new areas of development, to be situated at the places where energy is abundant.

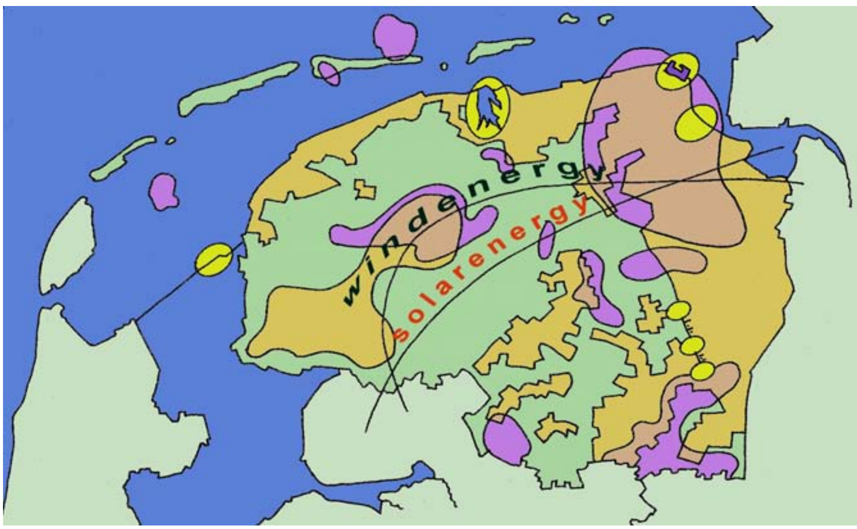


Figure 6: The energy mix map, evolving from the overlapping of all energy potential maps.

5 Exergy

5.1 What is exergy?

An extra dimension can be added to the method of energy potential maps by introducing the concept of exergy. As opposed to entropy, which can be considered the non-useful waste energy, exergy is the useful part in energy, the part that can be used to perform work, a measure of energy quality. In accordance with the Second Law of Thermodynamics processes constantly develop towards a state of increasing entropy (and decreasing exergy). The quintessence of this is that energy of a high-quality level (great exergy) should be used for high-grade functions before it transforms into a lower-quality state.

5.2 Low-exergy approach

Nowadays, households of the colder and temperate Northern hemisphere use high-caloric heat from the combusting of gas, gasoline or coal (at least 1200^{o}C), to heat up their houses to the pleasant level of 20-22^{o}C. This process would be much more effective if steps were inserted in-between, as in a cascade: the high

temperature would be used only in heavy industrial processes, of which waste heat could be used in lower-grade functions such as manufacturing processes, horticulture, and consequently for residential heating (Figure 7).

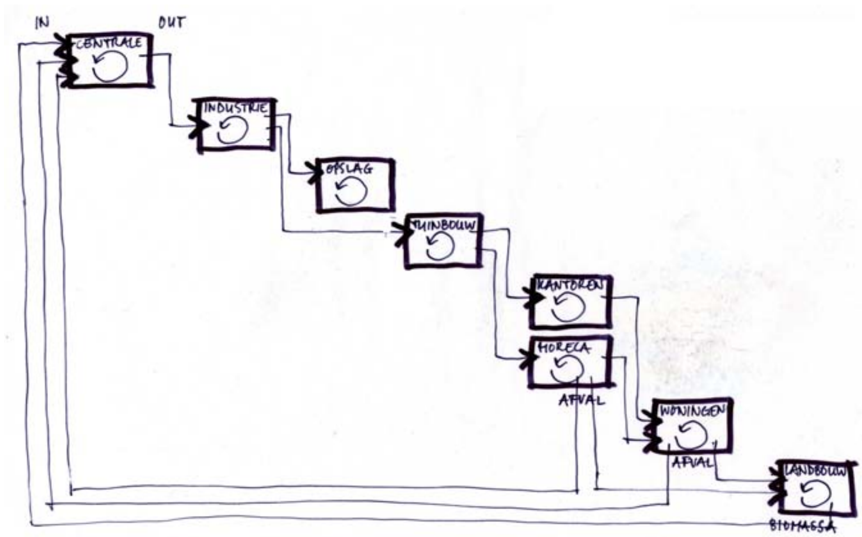


Figure 7: Graphical example of a heat cascade.

Thus, four instead of just one function would be served by the same amount of primary energy, a very significant contribution to a sustainable energy system.

5.3 The meaning for spatial planning

The exergy principle can be introduced to many scales of spatial design: region, city, district and building. It implies the concentration or mixing of functions (industry, agriculture, horticulture, commercial buildings, housing, etcetera), each embodying one of the steps in the cascade. Low-caloric heat should not be transported over long distances: heat losses would be too big. This would mean a radical shift from the separation of functions by present-day planning: horticulture near industry, residences near horticulture, or even integrated with it.

For the Northern Netherlands this means that the high-graded industry and power plant near Eemshaven and Delfzijl (province of Groningen) could be a starting-point for heat cascades, but this would also be possible on a smaller scale, starting from small industries in cities, districts or villages.

6 New images

6.1 The regional plan

Based on the energy potential maps and the exergy principle discussed in this paper, the current topography and not least a landscape analysis (Figure 8, left), a

new map (Figure 8) could be drawn for the Northern Netherlands. This map depicts a possible future in which some connections to the sea are opened again and utilised for energy generation, in which a larger part of Frisia and the lower parts of Groningen have become lakes again, and in which the entire spatial planning is based on an effective use of local energy potentials (sun, wind, water, biomass, geothermal heat, etcetera) and exergy (heat cascading).

Figure 8: Landscape analysis (left) and a new map of the Northern Netherlands (right).

Not directly visible in the right map of Figure 8 is the usage related to the new order of the landscape. It will be obvious that by flooding certain parts, introducing salty water, coupling urban settlements or detaching them for a certain degree of autonomy may seem threats to the present situation, especially when they relate to specific people, their homes and incomes. However, the analysis resulting from the climate change and energy depletion urge, offer various opportunities for new lifestyles (dynamic living, living on water, self-sufficiency) and economies (energy production, trade and knowledge, active and natural tourism, e-commerce, senior living).

6.2 Visiting cards

In order to better illustrate the consequences of the new regional map on a lower scale, so-called *visiting cards* were drawn for specific areas in the region. These are maps meant to inspire, tempt and invite. Figure 9 illustrates two of them. The Eemshaven-Delfzijl area (left), which may be energetically based on a multi-fuel plant (powered by domestic waste, biomass, residual fossil fuel), heat cascades, geothermal heat, an osmosis plant, and solar and wind power. This is perhaps the area richest in energy potentials, and therefore suitable for various developments in the near future. The Emmen area (right) is a much more poor area in energy terms. It forms a link to the peat colonies to the East and higher sand plateau to the West. The peat remaining, industrial and greenhouse developments may be the basis for spatial planning here.

7 Conclusion

The approach discussed in this paper implies a new mind-set on spatial planning based on sustainable development and effective use of local energy potentials.

By clearly translating local features and opportunities into maps, it enables visualisation of sustainable ambitions for the future and the consequence of it to planning. Thereby it proved to be an ideal catalyst for discussions, as opposed to written policy plans. The new approach can be used anywhere in the world, most probably leading to different solutions. A sustainable future has different faces.

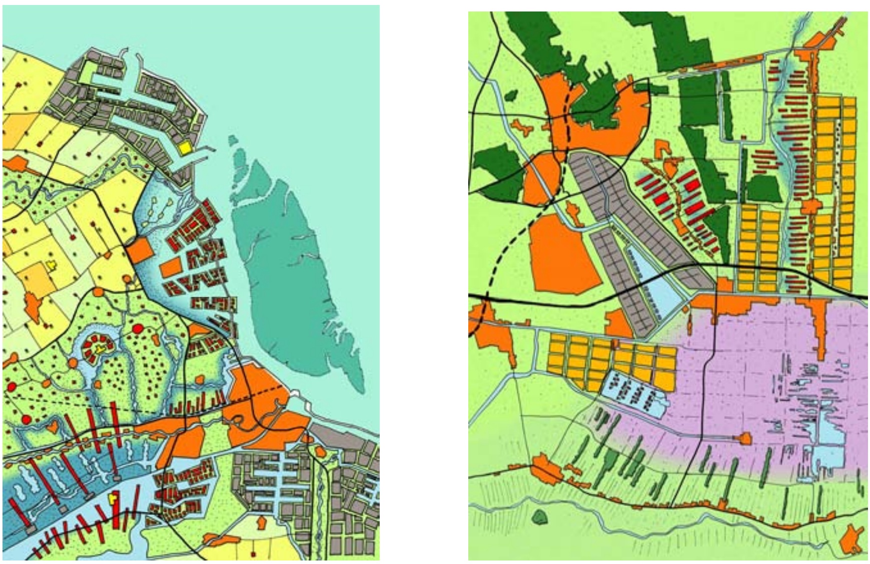


Figure 9: Visiting card for the Eemshaven-Delfzijl area (left) and the Emmen area (right).

References

[1] IPCC; Climate Change 2001: Synthesis Report (Technical Summary); IPCC: www.ipcc.ch, 2001.
[2] Jong T.M. de (ed.), Moens M.J., Akker C. van den & Steenbergen C.M.; Sun Wind Water, Earth Life and Living – Legends for design; Delft University of Technology, 2004.
[3] Roggema R., Dobbelsteen A. van den & Stegenga K. (eds.); Pallet of Possibilities; Groningen, 2006.
[4] Noorman K.J. (ed.) et al.; €nergie(k) Noord-Nederland (in Dutch); Groningen, 2006.

Environmental management in the Port of Koper and neighboring urban settlements

O. Bajt[1], I. Jurincic[2] & B. Marzi[3]
[1]*Marine Biology Station Piran, National Institute of Biology, Slovenia*
[2]*TURISTICA – College of Tourism Portoroz, University of Primorska, Slovenia*
[3]*Port of Koper, Slovenia*

Abstract

Service quality and appropriate environmental management in the Port of Koper are a constant concern for the administration. The port is located at the northern and eastern borders of the town of Koper, therefore, the administration has always paid careful attention to limiting and controlling its potentially negative environmental impacts. All its spatial interventions have therefore been subject to thorough spatial and environmental assessment. The withdrawal of some potentially inappropriate projects in the history of the Port was an obvious result of detailed assessments. During the preparation of the development and spatial plans, significant corrections and moderating measures were introduced in the majority of the proposed projects due to the port's location close to the urban environment. Permanent environmental monitoring of noise, air and water in the Bay of Koper was also introduced during the investments of the past twenty years, which made auditing of the Port's impacts on the environment possible.
Keywords: environmental management, Port of Koper, ISO 14001/2004, environmental impacts, spatial planning, monitoring.

1 Introduction

The Port of Koper is located in close proximity of urban areas and for that reason a concern for appropriate environmental management has always been present in the administration of the port. To ensure the safety of the environment, a number of measures have been taken according to independent environmental studies. A

www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060191

special investment service was founded in the beginning for the inspection and execution of these measures, and an environmental manager was named in 2004 in compliance with the environmental protection law.

Pretentious urban planning of the port has always been carried out in tight cooperation of the administration with local and government authorities. Reasons for such a method lie in the very foundation of the company. In the beginning, it was a huge government project which resulted in a rise of employment in the area. Soon the rapid growth of the port called for effective spatial planning measures. These were supervised by those competent for development and spatial growth. The port itself is now bigger than it's nearest neighboring town, Koper where it has been located by since 1957. It has tremendous economic value of which a major portion is still owned by the government. Yet an important part is also the property of the community of Koper.

The Port of Koper is run in connection with the marine traffic corridor and is one of the entry points of the fifth pan-European corridor. It therefore exceeds the significance of Koper, Primorska region or Slovenia. European documents encourage the stimulation of intermodal transport chains where seaports are very significant and have a key role. The Port of Koper is a link between the marine traffic corridor and the fifth pan-European corridor and thereby the largest intermodal junction in Slovenia. A consistent growth of the commercial Port of Koper is crucial as more and more merchandise is transported by marine-traffic.

The governmental concern for the only commercial port can be seen in the careful monitoring programs of the Slovene sea. This has been executed by an independent organization National institute of biology - Marine biology station Piran.

2 Spatial planning of the Port of Koper in the past

By reviewing the history of spatial planning in the area of the port it is clearly seen that the local community shows a high responsibility for economic development of the port in accordance with spatial possibilities. Its effects on environment must not exceed the limits dictated by law, nor can they have a long-term impact on natural resources. Every major development of the port area has been the subject of careful evaluation involving thorough study of possible negative effects on the surrounding areas and their inhabitants [1].

In the 1970s there was proposed an intensive economic development in the port. It was a mega industrial project which envisaged cheaper import and export of power supplies, ironworks, thermal power stations and other manufacturing industry at the port. This project, however, was terminated. It was decided that such an industrial development would exploit the advantages of a location close to the supply of resources, and would be a threat to the environment.

Similarly, a project in the 1980s, the off-loading of oil derivatives, suffered a similar fate. Its location was outside the port's borders on the Debeli rtic peninsula which lies between Koper and Trieste. In addition, it could have had a potentially huge negative environmental impact on the surrounding area in the case of an accident. It was taken into consideration that the peninsula is the most

preserved coastal and tourist region of the Slovenian sea. After a broad public discussion, the great flysch cliff region was declared a natural monument and thereby the oil terminal project ultimately rejected.

The unloading and storage of liquid petroleum gas (LPG) in the 1990s hit upon disapproval by local habitants as well. Despite good results from environmental studies, proving a very small potential for disasters, the port had to withdraw the project under public pressure.

A number of planned projects have also been partially or substantially modified due to the close proximity of urban areas. The question of environmental impact has often been the major determinant of these decisions. A terminal for unloading scattered cargo in 1980s was reduced by half. The constant moistening of the dumping ground was also introduced. At the same time the port began regular monitoring of the European energy terminal (EET) and its effects on the environment. The monitoring led to other modifications as well. At first a green curtain had been installed to reduce micro-particles in the air. This was later improved by raising a high enclosure that provides much better results.

3 Environmental impact of the Port of Koper

For one year, April 18, 2005 to April 19, 2006 a constant recording of total concentrations of inhalable dust particles PM10 was enforced in the immediate vicinity of the European energy terminal. This was executed by University of Primorska - Primorska Institute of Natural Sciences and Technology.

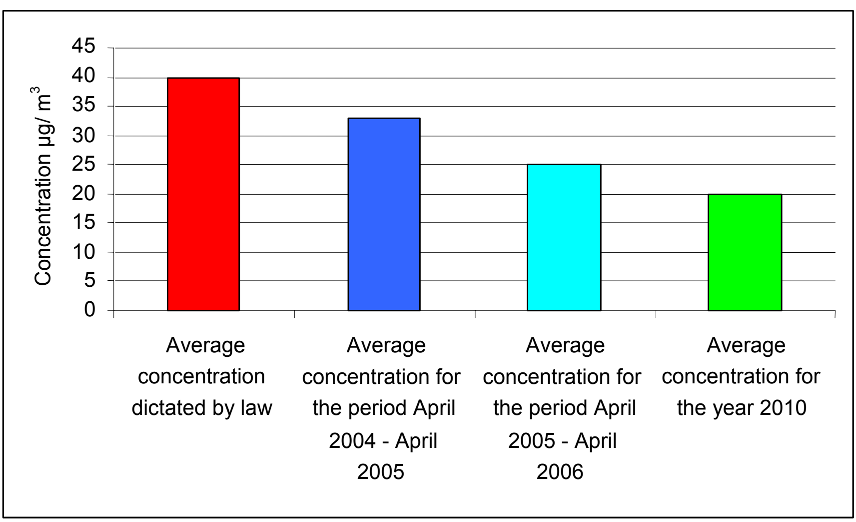


Figure 1: Concentration of dust particles PM10.

By EU guidelines the highest permitted concentration of dust in the air is 40 $\mu g/m^3$ [2]. According to previously mentioned recordings the concentration of

dust particles PM10 in the port was PM10 25,1 μg/m^3. In comparison to the average concentration recorded a year before, the recorded value was reduced by about 23%. It is important to note that off-loading activities actually increased during this period of time.

In accordance with the EU Environmental Code of Practice for seaports, namely co-existence of the port with the environment, and increasingly lower limit for concentration levels of individual elements in dust particles, lesser than 10 μm in size, the port has invested in environment-friendly equipment. Thus, by setting up fences that protect neighboring settlements from dust and noise, the port has also improved its visual appearance.

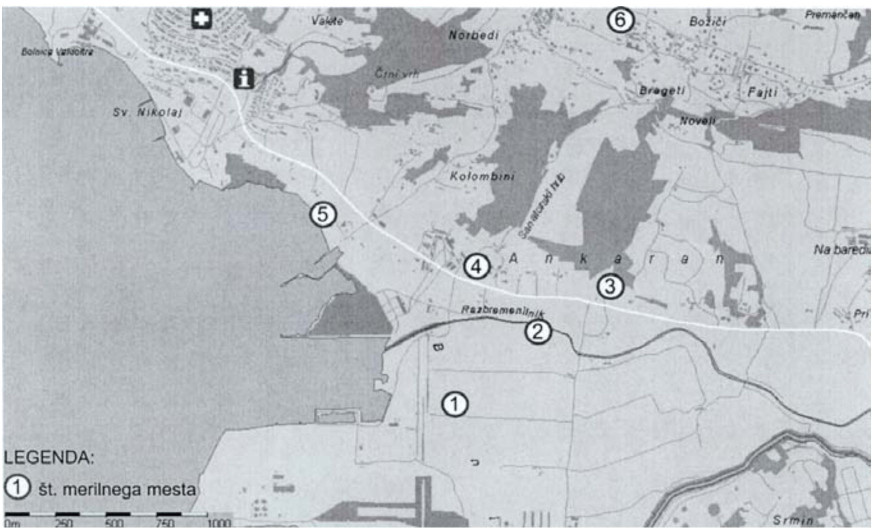


Figure 2: Measuring points for air pollution by dust sediments.

In October 2005 a project was initiated with the objective of constant development to ensure higher environmental responsibility and to introduce a system of environmental regulation to a wider area. The project is called "Inspection of air pollution by dust sediments in the area of local communities Ankaran and Hrvatini". It includes yearly analysis of assessments from the regions impacted by dust particles. This takes into account the wind and distance from the iron-ore and coal dumping grounds at the EET region. Samples are taken from six points (figure 2), from which the measuring point 6 is a checkpoint used for comparison between dust concentration from direction of Trieste and of Koper. The results will be known by the end of the year 2006.

4 Impact of the Port of Koper on the sea

Bay of Koper is a part of Gulf of Trieste (north Adriatic Sea), on its south-east section. It is a rather shallow bay with its maximum depth below 20 m. Its water-

mass circulation is complex and depends upon various factors, such as tide, winds and influx of freshwater from the land. There are two freshwater rivers flowing into the bay of Koper, namely river Rizana and river Badasevica. River Rizana is somewhat bigger and has its outlet at the port's area. These two rivers supply the bay of Koper with a significant amount of nutrients (eutrophication processes). They also contain some pollutants such as heavy metals, pesticides and hydrocarbons. An important potential source of pollution is the city of Koper with its sewage treatment plant (only the primary step of purification) and waters released into river Rizana just before it flows into the sea, as well as its pollution of the atmosphere. There is also a coastal road between Koper and Izola at the south section of the bay. This road is highly traveled, especially in the summer, with the influx of tourists. The runoff from this road is washed away directly into the waters of the bay of Koper.

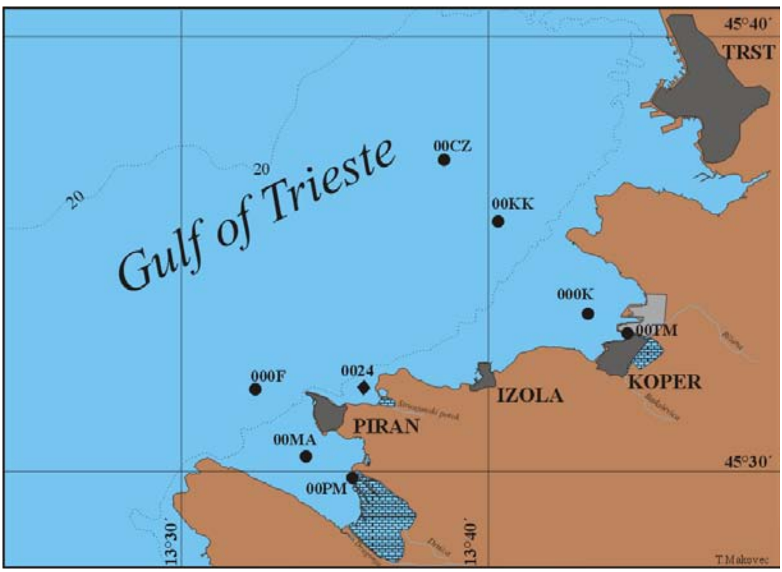


Figure 3: Sampling sites for monitoring of the quality of Slovenian sea.

The quality of Slovenian sea is constantly being monitored by the Republic of Slovenia [3]. Main sampling sites are shown in figure 3, but samples of water, sediments and marine organism are taken elsewhere as well. It varies depending on the specific purposes of the sampling.

The appointed parameters are mostly basic physical-chemical parameters (temperature, salinity, density, pH, alkalinity, nutrients…) and some specific pollutants, chosen according to the expected origins of input (heavy metals, hydrocarbons, pesticides).

The impact of the Port of Koper on the sea is most evident with the level of hydrocarbons in samples of sediments and certain mussel - *Mytilus Galloprovincialis.* Sediment samples were taken because hydrocarbons are hydrophobic substances with low solubility in water and they adsorb on

suspended matter in the sea. In this way they accumulate in sediments. Mussels create a similar effect, as they are filtering organisms which accumulate various pollutants in their tissue. This particular species is quite common in this region and is, therefore, a good indicator organism.

The analyses have been performed according to instructions by UNEP [4, 5]. The level of aliphatic hydrocarbons (C14-C34, Phytane, Pristane) in the sediments is shown in table 1, while the level of polyaromatic hydrocarbons (PAH) in table 2.

Table 1: Content of aliphatic hydrocarbons in sediments of the Bay of Koper (ng/g of dry sediment) in 2003.

Sampling site	00TM	000K	00KK	000F
Aliphatic hydrocarbons	9951	2149	1404	956

Table 2: Content of polyaromatic hydrocarbons in sediments of the Bay of Koper (ng/g of dry sediment) in 2003.

	SAMPLING SITE			
HYDROCARBONS	00TM	000K	00KK	000F
Naphthalene	<1	9	<1	<1
1-methyl naphthalene	<1	3	<1	<1
1-ethyl naphthalene	<1	<1	<1	<1
Acenaphthene	<1	<1	<1	<1
Acenaphtylene	<1	<1	<1	<1
Phenanthrene	38	38	11	28
Anthracene	39	20	17	35
Fluorene	213	<1	<1	<1
2-methyl phenanthrene	12	5	5	5
1-methyl phenanthrene	6	4	4	6
Fluoranthene	41	75	25	46
Pyrene	47	110	57	57
3,6-dimethylphenanth.	<1	<1	<1	<1
1-methyl pyrene	<1	<1	<1	<1
Perylene	169	<1	<1	<1
Chrysene	351	74	47	57
Squalan	<1	<1	<1	<1
Benzo(a)anthracene	54	44	19	30
Benzo(b)fluoranthene	119	93	50	51
Benzo(k)fluoranthene	49	48	25	47
Benzo(e)pyrene	46	58	47	47
Benzo(a)pyrene	15	41	23	53
Indeno(1,2,3-c,d)pyrene	103	50	33	55
Dibenzo(a,h)anthraceen	42	17	30	15
Benzo(g,h,i)perylene	39	77	138	74
Total	1383	766	531	606

In both cases we can see a higher level of concentration at the 00TM site which lies at the port's entrance. Concentration level of a site further into the bay is much lower and hardly any higher than the reference site 000F, located at a more open section of Slovene sea. Regarding the structure of PAHs present, we can conclude that the hydrocarbons are of various origins – a prevailing pyrogenic origin with some smaller amount of the petrogenic origin.

The impact of Port of Koper on marine environment is, as said before, also indicated by results from analyses of mussels which are presented in table 3. In the case of PAH we can see an increased concentration level in the mussels sampled near the port's entrance (00TM) in comparison to those sampled in the bay of Strunjan (0024, reference site), where there are no significant signs of pollution. The difference is not as noticeable in the case of aliphatic hydrocarbons, which is most likely a result of additional natural inputs of aliphatic hydrocarbons in the sea.

On the basis of presented results, we can clearly see that the activities of the Port of Koper have a definite impact on the pollution of the Bay of Koper, but this impact is relatively limited to the port's vicinity. We can therefore conclude: the marine environment of the Bay of Koper is only moderately polluted with hydrocarbons, as the concentration levels of PAHs in the surroundings of the Port of Trieste are several times higher [6]. It would be necessary, however, to initiate several more analyses of other possible pollutants in the bay of Koper, which may be a product of activities in the Port of Koper.

Table 3: Hydrocarbon level (ng/g dry weight) in *Mytilus Galloprovincialis* tissue in 2004.

Sampling site	0024	± SD	00TM	± SD
Aliphatic hydrocarbons	2563	464	3701	834
Polyaromatic hydrocarbons	771	183	2759	303

5 Planning of new terminals and environmental management

The Port of Koper knows about the environmental impacts of their activities. That is why their business politics are directed towards a rational management of the environment, to preserve it for the generations to come. Assessment and control of the environmental impacts has therefore become regular practice. The port was among the first in Europe and the very first in the Adriatic Sea to initiate an environmental management system according to ISO 14001 standard for all seaport activities in year 2000. In May of 2006 it was upgraded to ISO 14001/2004. When it comes to regular assessment and control of environmental impacts, the port cooperates with competent professional institutions. The Port of Koper handles the environmental management thoroughly. The fact that the company, Luka Koper d.d., manages the whole port area, enables the protection system to be present at all terminals and at all activities.

A need for new off-loading and storage areas has lately appeared, due to the changes in off-loading structures. Cargo ships are getting constantly bigger,

hence the demand for accommodation of operative coasts, entire docking infrastructure, and compatible technologies. The goal is to turn the extensive spatial usage to a more intensive or specialized construction and restructuring, which will also solve many ecological questions.

The Port of Koper is situated near a densely populated region – Koper, Ankaran and other settlements in the area. For that reason further development of the port must be carefully planned. Thus taking into consideration the boundaries set by an existing concept of urban development for wider area. It is possible to plan further spatial development of the port in Koper by expanding the existing terminals. The operative coasts of piers I and II can be extended to the west, and a 3rd pier can be constructed. This includes the area between the port and the coast of Ankaran. The port is also planning to build a green line and a water line of separation. In addition, an infrastructure of the passenger terminal in Koper area is going to be improved.

The traffic infrastructure of the Port of Koper is of national importance. The Ministry of Environment and Spatial Planning and the community of Koper are preparing a long-term detailed plan to specify spatial possibilities for economic development of the port. The detailed plan is to determine the borders of the port on its south and especially on its northern end. For this reason, representatives from the bordering communities are to be present at the planning stages (figure 4). The inhabitants of Ankaran, for example, are either opposing the expansion of the port towards their tourist settlement, or demand the 3rd pier's activities not to be aggravating and disturbing. The situation to the south is not as problematic. Some of the valid spatial plans are already being executed, withdrawing some of the fist docks of the port from the north part of city of Koper. After fifty years the north part of Koper is now opened to the sea again, and a bypass that enables circular traffic has been built around the city. The coast cleared by the port will be used for a passenger terminal, a suitable infrastructure.

There are various activities performed throughout the entire rejoin of the port and the impact of their synergy can be quite problematic. In order to best monitor spatial usage, and suitably distribute various activities, the international port of Koper, and its contact points (residential, tourist, administrative areas, and protected zones), must be developed holistically.

The new detailed plan of national importance, now being prepared, envisions a docking infrastructure in accordance with the developmental projects of the Port of Koper. Spatial activity planning is founded on the basis of the existing port's structure, where individual functional complexes are connected with the infrastructure. The plans for functional complexes are [7]: reinstatement of a passenger terminal with suitable infrastructure, expansion and widening the activities on piers I and II by building a new infrastructure and extending the operative coasts, construction of a multipurpose terminal on pier III and development of a new entrance into the port at Srmin.

Bearing in mind the programme, defined by the plan for preparation (based on references from first spatial conference, where representatives from neighboring communities also participated) of spatial analysis, analysis of guidelines (acquired with analysis of spatial conditions, tendencies and

possibilities), as well as study of vulnerability, the Ministry is putting up a public tender to gain a holistic professional solution regarding urban, landscape, architectural, and constructional-technical needs. To evaluate these solutions, spatial development (rational spatial usage), its functionality, and safety will be taken into consideration. Based on the chosen solution, the planner has to propose and substantiate the best possible variable, direct its optimization, create professional basis, and form a proposition for the detailed plan of national importance.

Figure 4: Area of the detailed plan of national importance for the Port of Koper.

6 Conclusions

The Port has shown awareness of the advantages and disadvantages brought about by its close proximity to the urban center. In accordance with the EU Environmental Code of Practice for seaports, that is, co-existence of the port with the environment, the port has invested in environment-friendly equipment and appearance, which enhances the contact area of the Port of Koper with the neighbouring settlements and its visual appearance. In spite of the observed impacts of the Port we can still say, however, that the marine environment of the Bay of Koper is only moderately polluted. Due to its need for new development area, the Port has been moving away from Koper and come closer to another urban and tourist settlement, Ankaran. This is also the reason why the latest plans for the construction of a new container terminal are subject to detailed assessment and the local community actively involved in the process. The port

has to constantly give additional reports to the neighboring inhabitants of current events and plans of any new activities. Public relations need to be widespread such as regular updating of the homepage, forwarding information through local and national media, and opening doors for public visits. Some of the most urgent public questions can be answered by an open and transparent politics of the port.

References

[1] Jurincic, I. & Bratoz, D. Koper. In Sarec, A., Sasek-Divjak, M., Mandic, S., Vitorovic, Z., Zavodnik Lamovsek, A. (Ed.). *Habitat II: Slovenian national report*, pp. 55-78. Ljubljana: Ministry of the Environment and Physical Planning, 1996.

[2] Zitnik, M., Jakomin, M., Pelicon, P., Rupnik, Z., Simcic, J., Budnar, M., Grlj, N., Marzi, B. Port of Koper - elemental concentrations in aerosols by PIXE. *X-ray spectrometry* **34**, pp. 330-334, 2005.

[3] Turk, V., Malej, A., Bajt, O., Mozetic, P., Ramsak, A., Horvat, M., Milacic, R., Scancar, J., Bosnjal, D. *National monitoring programme of Slovenia: programme for the assessment and control of pollution in the Mediterranean region (MED POL - PHASE III): report 2004, Reports MBS - Marine Biology Station 70.* Piran: National Institute of Biology, Marine Biological Station, 2005.

[4] UNEP/IOC/IAEA. Determination of petroleum hydrocarbons in sediments. Reference Methods for Marine Pollution Studies No. 20, 1992.

[5] UNEP/FAO/IOC/IAEA. Guidelines for monitoring chemical contaminants in the sea using marine organisms. Reference Methods for Marine Pollution Studies No.6, 1993.

[6] Adami, G., Barbieri, P., Piselli, S., Predonzani, S. & Reisenhofer, E., Detecting and characterising sources of persistent organic pollutants (PAHs and PCBs) in surface sediments of an industrialized area (harbour of Trieste, northern Adriatic Sea). *Journal of Environmental Monitoring*, **2**, pp. 261-265, 2000.

[7] Ministry of Environment and Spatial Planning, Spatial Planning Directorate, www.gov.si/upr/aktualno.php

Sustainable resources management in the context of agro-environmental EU policies: novel paradigms in Thessaly, Greece

E. Koutseris
Department of Planning and Regional Development,
University of Thessaly, Volos, Greece

Abstract

Primarily, this is a discussion and review of the general European policies of management protection in the perspectives of the new CAP (Common Agriculture Policy). There are some technical proposals in Thessaly for agriculture and environmental management with the contribution of a voluntary group such as the Working Team "Natural environment" (WTNE) in Natura 2000 and SPA (Special Protection Areas) by following a strategic PLAN. The fundamental problem and concept that this article examines is if the sustainable confrontation or approach can be better improved via means of the effectiveness and the short-term "valorization / utilization" of natural resources or if solutions should be found or proposed via means of long-term "conservation" of resources. Consequently, the completion of works, as paradigm, of 'Karla' reservoir dam's plain (collectors, water reservoir, etc.) and the beginning of projects for exploitation – valorization, created a reflection in regard to the type and manner of water consumption and the more general development and the valorization of level areas in Thessaly, Greece. The choice between different sustainable management solutions, except the solely "economic effects", will be supposed to include the parameter "environment" and it approaches the question of "exploitation – valorization" of resources in a context of effectiveness and "conservation" of resources.
Keywords: environmental EU policies, sustainable development, natural resources "valorization / utilization" versus "conservation", regions of eco-development or controlled activities area.

1 Introduction

The most important measures for the environment were taken in the European Community, with the Single European Act in 1987 and the treaty on the EU in 1992. They addressed important environmental problems caused particularly by modern agriculture [1, 2]. At the same time, two new protection issues arose: The network "Natura 2000", SPA and the management arose for Greece and Thessaly of water [3, 4]. In the EEC in 1992 an important step was made towards "*sustainable development*", the object of which is not only the protection of the environment but also sustainable future development. A new policy, focusing on preventive action, placed particular emphasis on the more rational use of resources, as well as on the reduction and the more effective use of the products, themselves [5–7].

"The new policy for rural space encouraged *planning for the comprehensive development of the environment.* Financial incentives were established in the form of various policies, which were aligned with the objectives of 'Natura 2000 - SPA' and would be able to contribute to its success". In this way, safeguards were placed on the areas where the certain types of EC Habitat and species of flora and fauna are found, while at the same time taking into consideration economic, social and regional requirements [8–10].

With the reform of CAP 1992, traditional developmental objectives were enriched with three new horizontal objectives, incorporating the environmental objectives into the CAP (Regulation 2078/92 and Reg. 2092/91): (a) the protection of environment, (b) the protection of human health and (c) the rational and viable exploitation of natural resources [3, 11, 12].

Regulation 1257/99 and the new CAP 2003 further promote the environmental regard in the rural sector connecting the issuing of aid with the observation of rules "equitable agricultural practice". This concerns mainly: the protection of waters, the controls in the use of fertilizers and pesticides, the protection of regions "Natura 2000" and the incorporation of environmental methods of production in the quality of products [3, 6, 9, 10, 13].

In this context, two new plans, as paradigms, must be harmonized with the previously mentioned provisions for the agricultural development aiming at the same time to protect the environment. This approach forms the future of sustainable development in agriculture concerning: a paradigm for hydro-environmental planning (management of water resources) and sustainable development in Thessaly and a novel control policy paradigm of sustainable management in Thessaly (eco-development of Natura 2000 plain-mountainous area). It will be supposed is given particular weight in the more rational use of resources, aiming at the reduction of environmental impacts and more effective use of produced products [7, 10, 12, 13].

The Author as Scientific in charge of Working Team WTNE, approaching the aspect of natural resources from the sustainability point of view, of these regions he attempt to contribute first in proposals concerning the minimisation of intensiveness of the flat regions under examination and afterwards in proposals

concerning the appointment, projection and in general exploitation of mountainous regions to be contribute [7, 10, 11, 13, 14].

2 New regional agro-environmental policies, proposed by a voluntary group in Thessaly, Greece face to the new CAP

Two important regional programs, from EU policies, Greece following. The first 'Natura 2000' Program covered three phases in Greece: Phase A (1992–1995), during which the inventory and evaluation of regions was accomplished; Phase B (up to 1998), during which a list of regions of community importance was created; and Phase C (up to July 2004), during which the 'Special Areas of Maintenance' were to be officially delimited and classified as habitat areas and the priorities for their conservation and the rehabilitation determined. In addition, 'regions of special protection' (SPA), contained in the Natura 2000, list was delimited, though these areas were focused on the protection of birds, based on Directive 79/409/EEC [14, 15]. The second program of the EU, the 'Corine Biotopes Project', led to the recording of the most important biotopes in Greece [6, 10, 13–15].

Unfortunately, however, the expected consultations with local communities on the terms and the restrictions of 'Natura 2000' did not take place, when the first (scientific) list was drawn up in 1995. In many cases, lines had already been drawn on the map without due consideration of local needs and insights [11, 6, 14, 10]. The extent of the regions covered by "Natura 2000" in the region of Thessaly is roughly 220,000 Hα (percentage 16–20%). As is apparent there are several regions of ecological interest that have been studied or are being studied. The creation of a European ecological network of special zones which will be named "Natura 2000" is anticipated for these regions. Those, according to the criteria of the 92/43 directive were classified by the Greek Centre of Biotopes - Wetland in three categories of A., B., and C., significance [3, 6, 14].

During 2003 and 2004, the Working Team tried to cover the current lack of briefing and consciousness-raising through multiple meetings both in the capital of Thessaly, Larissa, and in 15 regions of special interest to environmental planning. Simultaneously, e-mails concerning general and specific proposals on environmental issues were sent to various institutions and to all the municipalities of Thessaly [4, 6, 7, 10, 11, 13, 14].

Apart from briefing-informing, *the goal* of the Working Team was also the determine the activities within the framework of the 'Strategic Regional and Environmental Planning for the Sustainable Development of Thessaly' (PLAN) for the "Natura 2000" regions and for the aquatic resources of Thessaly. Planning focused more on the sectors that have an impact on the natural environment, while keeping in mind all of the diagnoses and regional policies, the Land-Planning Plan (LP/P), and the policies for rural space regarding the management of water and soil, as well as urban planning in rural areas [10, 11].

The proposed PLAN was not an end in itself, but a means for the intensification and future guarantee of *development* · This however must be supported by theories of sustainability. *A general objective is that the*

development and the protection of the natural, social and cultural environment should become coordinated, so that, over time, the greatest possible economic profits for the region would be achieved. In other words, the main objective of the plan must be the minimization of negative impacts along with a parallel maximisation of advantages for the local economy, society and environment [7, 10, 13, 14].

Priorities of the measures and actions of the 3rd Community Structural Funds (CSF) of Thessaly were the promotion of "regions of natural beauty" and their integration into the "Natura 2000" network, to also include sensitive ecosystems and biotopes along rivers, lakes and coastal regions that were not included before. This is to be accomplished through the development and implementation of studies and rehabilitation of these regions, the promotion and management of their landscapes, the encouragement of visitors through the creation of centers and kiosks for Information and Environmental Education, the publication of briefing materials, the mapping out and organizing of paths and of other activities, sign-posting, fencing and land use planning, and the promotion of products of the 'Natura 2000' regions [10, 13–16].

Certain new EU policies have been presented in the so-called *Agro-foodstuff Sector and Globalization*: This labeling of organic and environmentally friendly products would take place under the auspices of the certification systems organized by ISO 14000 and EMAS (Environmental Management and Audit Scheme), as well as POP, E, CE - HACCP, which would insure competitive advantages in the international market for quality products and processes [4].

Taking into consideration developments and tendencies in the WTO, EU regulation 1257/99 and the new intermediary revision of CAP 2003 promote environmental considerations in the rural sector even more, connecting the provision of aid with adherence to the rules of 'equitable agricultural practice' which mainly concern: *the protection of water, controls in the use of nitrate fertilizers and of pesticides, the protection of the 'Natura 2000' regions and of habitats, and the incorporation of environmentally friendly methods of production into the quality of products* [10, 12].

But this would require a reorientation in EU spending in favor of viable labor-intensive agriculture and know-how, which is at the same time environmentally friendly. In addition, there must also be a drastic reinforcement of laws concerning the commercial labeling of quality products, promotion of institutions for interdisciplinary and inter-professional regional collaboration and programs for the environmental consciousness-raising of consumers [14, 17].

However, all this would require the strategic reformation and reconstruction of the European agricultural model from individual production to a more qualitative, i.e., more collective and environmentally friendly system of production and consumption, all of which would require an increase in EU funding [7, 10, 13, 14].

The new CAP (June 2003 and April 2004) *takes into consideration both the consumer and tax payer-citizen, while at the same time, giving European farmers the capacity to produce whatever the market demands. In the future, financial aid to agriculture will be given in part irrespective of the volume and type of*

production. In order to avoid the abandonment of production, member states can retain certain types of aid under specific conditions [11, 10, 13, 14, 18].

3 One paradigm for hydro-environmental planning and sustainable development in Thessaly

The above project PLAN constitutes an important contribution in the environmental planning based on the network of regions of protection, conservation (maintenance) “NATURA”, ‘SPA’ and the water resources management (Directive 2000/60 EU) in the new prospect of rural space. This project was drawn up by the Working Team “WTNE” of the region of Thessaly within the Program of local events of raising the awareness of the rural population in three phases. Phases A and B of PLAN were completed in 2004. Simultaneously, it constitutes (a) an evaluation of the participation in the Phase C of a Plan for Thessaly, (b) an evaluation with the title “Environmental social consensus in Thessaly”, (c) an ex-post evaluation of acceptance and impacts of the application of the “Natura 2000” – “SPA” regions and management of aquatic resources, in the new prospect of rural space’, which is briefly presented here, was recently published in the form of a Monograph. It constitutes the conclusions of all the activity of WTNE, in the form of the composition and synopsis of the Environmental Strategic Planning mentioned above [6, 10, 11, 13, 14, 18]. From the analysis PLAN the need of a ‘New Strategic Management Natural and Water Resources and Work of Thessaly’ results, which will include the following priorities from the special occasions of the PLAN.

A. Priorities in the water management until the year 2010 [10, 11, 13, 14]:
1st Sufficient and Qualitative Water supply of urban centers and settlements of Thessalian countryside by “Institution of Management of Waters Thessaly”
2nd Guarantee of Minimal necessary Environmental Benefit for Viability in the natural system of Pineios river and its tributaries by “Barrier in the extension of Private Drillings and in the over pumping of underground waters” and substitution of >1000 millions of m^3 of underground reserves of over pumping in the period 1985–2000.
3rd Maintenance of viability of Rural Exploitations.

The means for above reported are:

- The water management as collective natural and social resource in insufficiency.
- Support of planning and subsidy of action of substitution in regions of mono-culture of cotton with subsidized not irrigated cultures.
- Restriction of mono-cultures, and reduction of 10% and increase only livestock-farming which does not overload the environment.
- Use of veterinary surgeon flows and the exploitation of organic fertilizers in replacement of chemists, the creation of model veterinary surgeon settlements, the biological Agriculture and the extension of regulation 1257/99 (2078/92) EU.

- ❖ Proportional changes are included in regulation 1257/99 on the creation and operation irrigatory and land reclamation work in the frames of 3rd CSF. The integration of each irrigatory project will be supposed to observe measurable criteria and certain beginnings that were cited already.

B. Intervention for the creation of work of storage of surface waters
As fundamental region of action of PLAN entire the Region Thessaly is considered. The PLAN forecasts the concretisation of two pilot drawings and the further choice of certain regions of action. This because its application in a relatively big region is a particularly awkward undertaking and so that are achieved concrete results which both they will contribute in the improvement of prevailing situation and they will function positively in its distribution. These regions, which can have geographic or thematic determination and content, will be offered for tourist and sustainable development and consequently they correspond in the criteria that were formulated. In the base of this aspect as regions of action they are selected.

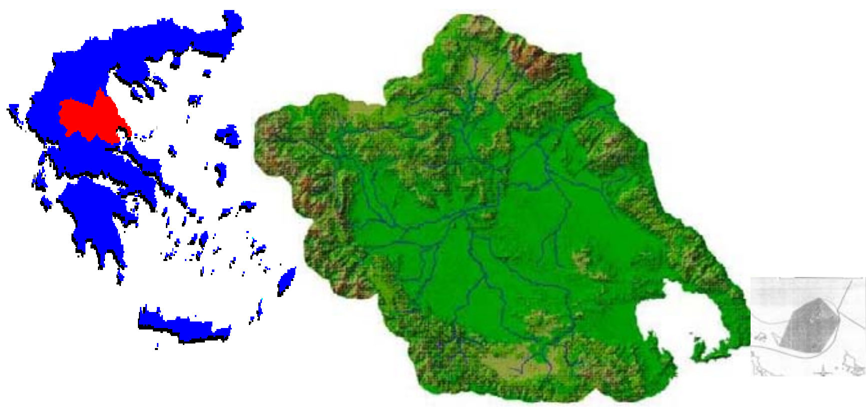


Figure 1: Region of Thessaly, Greece and Pineios river basin, with the Natura 2000 and SPA islands (Total surface area: 10.550 km²). Source: [13, 19].

Adversative places of prospect of development were presented in WTNE laboratory, which are influenced immediately by the political CAP. Concisely, the new proposals of the PLAN are as follows:
1. Sustainable tourist development in Thessaly. 2. Sustainable hydro-enviromental development. 3. A hierarchy of environmental planning within the Prefectures .4. Choice of the regions for action. 5. Relating regional action priorities to the strategy of the plan. 6. Zoning. 7. The need for a national zoning and energy policy [11, 13, 14, 17, 18].

4 A novel control policy paradigm of sustainable management in Thessaly

The region of Thessaly allocates a large variety of eco-hydrological reserve that covers on the one hand all the natural particularities of Greek space and on the other a large breadth of human activities. Particular gravity was given by the Working Team of the Thessaly region, as model can be what is described in the above monograph the Eco-development region of "Karla – Mavrovouni – Kefalovriso", as a controlled development area and the installed Committee of Management, which region is located in the south - eastern part of Thessaly [11, 17, 20, 21].

This was characterized, via Special Environmental Study (SES) and according to the articles 18 par 3 and 19 par. 2 and par. 5 of Greek National Low 1650 / 1986 as *region of Eco - development Region* (or pSCI, [4, 21]). The above also applies for other regions within the framework of Directive 92/43/EEC, which have been evaluated as essential for their integration into the Natura 2000 network. These areas have also been presented by the appropriate authorities to the European Committee, along with Regions of Special Protection (REP), as "proposed Regions CC" (pRCC), which will constitute the Network Nature 2000 (or pSCI), [21, 4]. The regions listed below include all the above characteristics and are included under the Ministry of the Environment and Public Works plan [22]:

A) Mount Mavrovouni and the marine region surrounding it, which has been included in the European network of the protected region "Natura 2000" (see Figure 2, A and B in [13, 14]).

B) The region affected by the work of re-creating Lake Karla which consists of the region of the re-flooding of the lake and the immediately adjacent region (Figure 2, B in [13, 14]).

C) The wider region, flat and hilly, which is directly connected to the lake and constitutes its drainage basin (Figure 2, Γ in [13, 14]).

For the completed and effective protection and conservation of itself and its values it was separated in individual areas. These concern extensive regions that include villages or settlements that present on one side, problems of intensive agriculture and de-valorization, so that protection is required and on the other side, particular value and interest because of ecological geomorphologic, biological, scientific importance that require maintenance [4, 13, 14].

The proposed new policies are related with the observed lack of regulations for certain questions that concern mainly the improvement of management of natural resources. This lacks in the sector of maintenance (or even reduction) the levels of consumed natural resources (energy or other) concern (a) regulations in local level for the important repercussions of environment, (b) regulations for the under protection regions, (c) regulations on restriction of consumption of energy (electric and humid fuels) and natural resources, (d) recycling and re-use of waste, (e) recording of levels of pollutants, (f) water management in the total of region or even in the wider Thessaly region and (g) systematic follow-up of habitat and not enacted interventions in cases of problems [7, 14].

5 New proposals of sustainable resources management in Thessaly, Greece and conclusions

The result of the rational management the surface and underground waters imposes the restriction of use of underground waters and the exploitation of surface waters of Karla area. This choice coincides with the general environmental choices of EU and strength then the economic development of region, which as mentioned before, presents economic and social degradation.

We approaches the question of “exploitation – valorization” of resources in a context of effectiveness and “conservation” of resources. The new European policies incorporate the environmental policies in the economic development and concern more the development of territorial entities despite the development of a sector of economy, such as agriculture. The “regions of controlled activities” – with one extension of protected regions and one reformation of cultivation land – are promoted, simultaneously with the national policy and the protected areas “Natura 2000, SPA”, which in certain cases concern mainly level regions. The new rural and environmental rules, which are imposed parallel by the new CAP, lead to this direction more and more [7, 10, 14].

In this context, the article deals with the development of the above work taking into consideration the new policies, it categorizes and rates the environmental problems territorial and quantitative and proposes rules. The article shapes on the one hand proposals for non-intensive exploitation of examined level areas, and on the other hand, proposals of appointment, projection and in general exploitation of the mountainous areas. All these are examined in a context of sustainable development of natural resources of these regions. These opposite opinions, with regard to the prospect of development of rural regions, were formulated in recent meeting that the Authorities of the region of Thessaly organized, where relative proposals for sustainable development of the Thessaly were proposed [14].

From the *first part* of the PLAN [11, 13] or the monograph [6], mentioned before, the results are: under the title ‘Conclusions from the Evaluation’ we pointed out the main problems that must be faced in Eastern Thessaly. These include: the management of solid waste, the intense problem of overexploitation of water resources, the need for controlled development of expanding settlements in the Eco-development regions, human intervention in nature, and the need for partial zoning and regulating of areas under intense use from human activity.

In Chapter Four, ‘Examination of Methodology’, certain conclusions were reached. These referred to principles on the general direction of PLAN, based on the ecological reserves, which led to two plans for the PLAN, including a new strategy for managing natural and aquatic resources, as well as construction projects and regulations in Thessaly, such as water supply for cities, restrictions on drillings, alternate sources for thousands of cubic metres of water, maintenance of the viability of agricultural production, and, finally, the creation of an Institute for the Management of Resources and the Environment in Thessaly.

In Chapter Five under 'Final Conclusions – Priorities and Proposals', three general conclusions were reached. These referred to the need to establish priorities for the use of water up to the year 2010, interventions needed to regulate the demand, use and management of water, and planning for the creation of projects to collect surface waters. Finally, it was proposed that the number of areas and resources to be protected be increased [7, 10, 13, 14].

From Part Two seven proposals were made, relating to:

1. *Sustainable tourist development in Thessaly*
2. *Sustainable hydro-enviromental development*
3. *A hierarchy of environmental planning within the Prefectures.*
4. *Choice of the regions for action.*
5. *Relating regional action priorities to the strategy of the plan.*
6. *Zoning*
7. *The need for a national zoning and energy policy*

References

[1] A. Kungolos, P. Samaras, E. Koutseris. 2003, 'Using bioassays for testing seawater quality in Greece': In *Journal of Environmental Science and Health*, eds Marcel Dekker, Inc, Reprint Program, NY, Vol. A38, No. 533-544, 2003M

[2] Podimata, E. Koutseris and N. Tsiropoulos. 2004, 'Water Quality Assessment by Ecotoxicological and Chemical Methods in Magnesia, Greece': *In Water, Air, and Soil Pollution:* Focus, October 2004, vol. 4, iss. 4-5, pp 179-187 (9) Kluwer Academic Publisher

[3] Koutseris E.: 2003, General Directions of the Natura 2000-SPA regions and water resource management in Thessaly (dir.2000/60/EU): Management Committees and Regional – Agro- Environmental Measures in the New CAP''. In *Newspaper 'Neos Tipos'*. July, extensive publication, Volos

[4] Koutseris E.: 2003 a, 'General Directions of NATURA 2000-SPA regions and water resource management in Thessaly (dir.2000/60/EU): Management Committees and Regional – Agro- Environmental Measures in the New CAP': In *Proceeding*. 3rd Congress of the Development of Thessaly, ETHEM, Un. Thessaly, Technical Education Foundation (TEF), Larissa, 12-14/12/03, pp 375-393

[5] EU directive/92/43.: 1992, 'Natura 2000', «on habitats», *EEEK,* Brussels.

[6] Koutseris Ẽ: 2004 a, *Environmental social consensus in Thessaly*: An ex post evaluation of the acceptance and impacts of the application of the 'Natura 2000 - SPA' regions, and water management resources, in the new prospect of rural space''. Eds. Erodhios. Thessaliniki, p. 180 (in Greek)

[7] Koutseris E., Papavassiliou: 2005, 'Water as social and economic commodity: Efficacy or conservation sustainability methods'. *In Conference Proceedings, EWRA69C.* 6th International Conference 7-10/9/05, Menton, France

[8] Kotios A.: 2002, ‘Restrictions and capacities of development activities in the ‘Natura regions’ in *Conference Paper 26-4-02*, dep. of P and RD, Un. Thessaly, Volos (in Greek)

[9] Koutseris E.: 2002, ‘The integrated reconnaissance in environmental protection: A new approach to planning and design’. In *Proceedings of the International Conference* on ‘Protection and Restoration of the Environment VI’ (eds. A.G. Kungolos, A.B. Liakopoulos, G.P. Korfiatis, A.D. Koutsospyros, K.L. Katsifarakis, A.D. Demetracopoulos) Skiathos, pp. 1377- 84, July 1-5, 2002

[10] Koutseris E., Polysos S: 2005, ‘Water resources exploitation-valorization in the context of environmental policies: The case of Smokovo's dam in Thessaly’. *In Conference Proceedings, EWRA69A.* 6th International Conference 7-10/9/05, Menton, France

[11] Koutseris E. et al (Scientific in charge): 2004, *Regional Environmental Strategic Plan of Sustainable Development of Thessaly (RESP of SD in Th)*: Based on the ‘Natura’ area and water researches management (dir. 2000/60/EU), WTNE of Region Thessaly, Larissa, p. 301 (in Greek)

[12] Koutseris E.: 2003 b, ‘The new water policy: control and location of territory units with methodologies ‘for a special or multiple impacts’ in the small scale’. *In Proceedings.* 3rd Congress of the Development of Thessaly, *ETHEM*, UTH, TEF, Larissa, 12-14/12/03, pp 269-291 (in Greek)

[13] Koutseris, E.: 2006, ‘Application of E.U. Environmental Programs in Thessaly, Greece: Problems in management - protection’, submitted in International Journal of sustainable development

[14] Koutseris E.: 2004 b, ‘Agriculture ou Environnement ? La région d’ex lac KARLA en Grèce’, *Discussion Paper* Series, 10 (16): 407-438, Un. Thessaly, Volos

[15] Ministry of Greek Environment, EP.PE.R. 2001, ‘Environmental Operational Programme’, *Internet,* http://www.minenv.gr and http://www.minenv.gr/4/41/4108/g410801.html

[16] P.E.P.: 2000, *Regional Operational Program of Thessaly 2000 – 2006*, Special Management Service P.E.P., Larissa (in Greek)

[17] Koutseris E.: 2005 a, ‘Methodologies of reconnaissance, with evaluation of environmental planning impacts and one new model of protection: For the protection or conservation’. In HELECO *Symposium Proceedings* for the technology of the Environment: In O-B 374 doc or abstract book: B 58: 71, TEE: HELECO ’05, 3-6/2/2005, Athens

[18] Koutseris E.: 2004, ‘Methodologies of reconnaissance and evaluation of the environmental planning impacts: for protection or conservation’. In *EWRA Symposium Proceedings*, on ‘Water Resources Management: Risks and Challenges for the 21st Century’, Vol II, pp 903-914, Sept 2-4/9/04, Ismir, Turkey

[19] Panagopoulos A.; Evagelopoulos A.; Alexiou J.; Kalfountzos D. and Kotsopoulos S. (2001): Environmental Actions of Reg. 2078/92/EEC -

State of Art Knowledge of Water Resources of river Pineios estuary, Thessaly, National Agricultural Research Foundation, Athens, 2001
[20] Koutseris E.: 2005 b, 'Climate and two-type torrential physiography: New Interdisciplinary Approaches and Discussion of Processes from Mediterranean Environmental Planning'. *In Conference Proceedings, EWRA69B*. 6th International Conference 7-10/9/05, Menton, France
[21] ETHIAGE: 2000, Special *Environmental Study and plan of management.* ETHIAGE, Special Protection Area of mount Mavrovouni of the LIFE program (in Greek)
[22] Ministry of Greek Environment (Y.PE.HO.DE.): 2002, Special Environmental Study, *Conclusions* of Dep. of Natural Environmental Management (in Greek)

Intelligent flexible disassembly and recycling of used products to support sustainability and total quality management in modern manufacturing industry

P. H. Osanna[1], M. N. Durakbasa[1], J. M. Bauer[2], H. S. Tahirova[1] & L. Kraeuter[1]
[1]*Department for Interchangeable Manufacturing and Industrial Metrology, Vienna University of Technology, Wien, Austria*
[2]*National University of Lomas Zamora, Buenos Aires, Argentina*

Abstract

Important changes in the environment, reduced availability of natural resources and the increasing growth of waste means that new concepts and strategies to recycle technical consumer goods, such as household instruments, consumer electronics and passenger cars, are required instead of land filling, burning or steel production. In view of the large quantity and high personnel costs, an advanced disassembly technique is needed using more rational than traditional manual processes and cheaper than highly sophisticated fully automated high technology machines and devices. One of the biggest problems for the recycling or disposal of products besides the growing mass of waste is their complex composition. An economic, sustainable and ecological procedure is necessary to prepare and to refurbish products, therefore the disassembly is the first step to achieving a higher rate of recycling.
Keywords: disassembly, waste, life cycle assessment, sustainability, planning, development, recycling, consumer products.

1 Introduction

In modern industrial production the very important issues of the protection of environment and sustainability must be taken increasingly into consideration [1, 2]. Manufacturers of products for every day life – consumer electronics,

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060211

automobiles, house hold devices – are facing increasing demands from consumers, public opinion and governments to minimize the pressure on environment, decrease consumption of raw materials and energy, and reduce pollution and waste [3, 4].

The complexity of scrap recycling from the above mentioned consumer products is a strongly interdisciplinary field where there are items such as collection logistics, disassembly, components re-use, recovery of precious and rare materials such as copper, recycling of non-hazardous materials and disposal of hazardous and toxic substances. All these require the coordinated work of interdisciplinary teams [5, 6].

European Union Environment Council approved in 1990 the Commission's Strategy for Waste Management, which included ELVs (End of Life Vehicle) as a Priority Waste Stream.

Concerning of the Waste Strategy Priorities the generally accepted hierarchy of priority with respect to waste is:

- prevention,
- recovery,
- disposal.

Within the term recovery the concept includes the re-use of parts, the recycling of material and the recovery of energy.

The Objectives of the European Commission Proposal for automobiles are the following:

- avoidance of waste,
- reduction of landfill demand, and
- reduction of toxicity.

European "Take-Back Law" requires automobile manufacturers to take back all vehicles which were not sold in that country. Voluntary agreements have been widely accepted by industry and the threat of legislation has subsided slightly.

Essential advantages to application of old automobiles are established by the following directives of EU and according to their transformation to the national right:

- The directive 75/442/ European Economic Community of advice from July 15, 1975, about waste.

- The directive 96/61/ European Economic Community of advice from September 24, 1996, about the integrated evasion and reduction of environmental contamination.

- The directive 1999/31/ European Economic Community of advice from April 26, 1999, about gathering waste.

- The directive in 2000/76/ European Economic Community of the European parliament and advice from December 4, 2000, about combustion of waste.

- The directive in 2000/53/ European Economic Community of the European parliament and advice from September 18, 2000, about old automobiles.

The overall targets for the improved process, which are in line with those proposed by the EU, are the following:
• for a car being scrapped in 2002, a maximum of 15% of initial weight to go to landfill
• for a car being scrapped in 2015, a maximum of 5% of initial weight to go to landfill.

Economic viability is a necessity for the implementation of the improved disposal process under market conditions. If analysis of detailed monitoring information indicates that an action is viable, but direct financial support is required for it to be sustainable, then the means of providing such support will be considered.

Currently, the European Commission is proposing a directive in which the vehicle manufacturer has the prime responsibility for the product. However, the complex nature of vehicles, the length of life, and the other industrial sectors involved subsequently, point to a shared responsibility between all affected parties.

Essential targets in accordance with item 7 of the EU instructions 2000/53/EG about old vehicle are [4, 7]:

- Target from 2006:
  85% recycling
  80% material recycling
  that is a maximum of 15% disposal.
- Target from 2015:
  95% recycling
  85% material recycling
  that is a maximum of 5% disposal.

2 Life cycle costs

In the past the cost structure of a product contained only the cost of development, production including quality assurance, marketing, sales and service. As this is mainly influenced by the manufacture of the product it is important to take into consideration the life cycle costs meaning that the real costs in each life stage are added (see Figure 1).

For this purpose a detailed quantitative assessment of expenses for a product must be carried out taking into account the cost contributions.

2.1 Development costs

Roughly 5% of the life cycle costs originate from this stage. On the other side to this small part of the costs between 60% up to more than 80% of the total life

cycle costs are fixed during the development stage. The influences of other stages on the cost level are small compared with those of research and development.

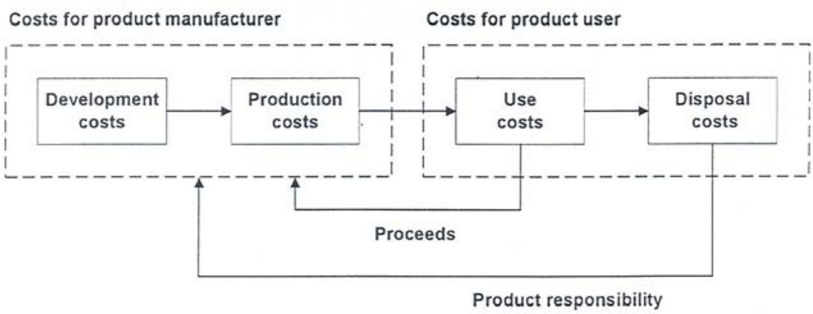


Figure 1: Influences on costs during product life cycle.

2.2 Production costs

In any case one can operate on the rough rule which allows the determination of production costs based on costs of different materials used for engineering purposes

2.3 Use costs

The costs of the product use phase can be calculated on the basis of the energy costs per time unit that the product is in use.

2.4 Disposal costs

The costs of take-back, recycling and final disposal are comprised by the costs of take-back systems, recycling and final disposal processes as well as the returns for the gained secondary raw materials.

In order to account for the different points of time when costs occur, the costs for usage and recycling have to be discounted over the life span of the product.

3 Life cycle costs

The challenges for innovative product design are two fold: on the one side having to contain material and energy flows within the product life cycle to close the loop, on the other side reducing the overall consumption of materials and energy to make the loop thinner (see Figure 2). These both should be achieved whilst at the same time the consumer needs are to be identified and met.

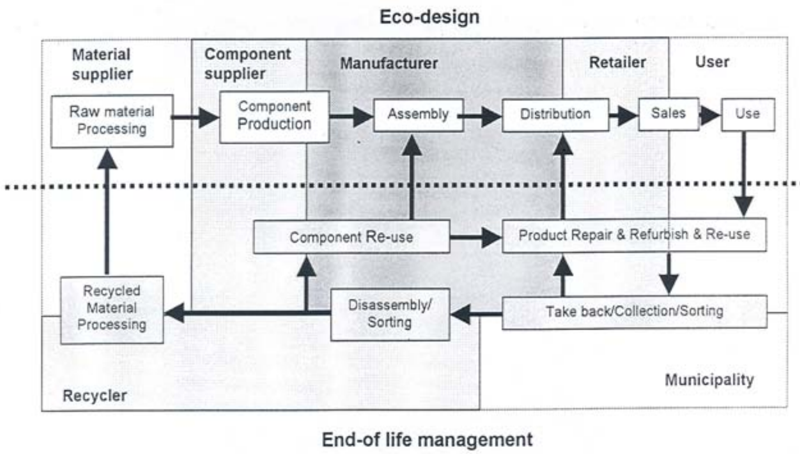


Figure 2: Closed life cycle of consumer products.

The closed life cycle therefore consists of two main streams:

- The "traditional" product life chain (such as production, distribution and use). Introducing environmental concerns in this stream is called "Eco-design".

- Re-using and recycling products, components and materials (such as take back, re-use and recycling). Operation of this stream is called "End-of-life Management".

By disassembly of further non usable parts a considerable reduction of use costs and disposal costs can be achieved. In this respect it is very important that different branches of industry collaborate insofar that different experiences are exchanged as far as different products and different industrialisation are concerned [9].

4 Principles of a system for intelligent flexible disassembly

A fixed automated single purpose disassembly system cannot operate economically for the time being [6, 8]. So it is very important to design such a system with high as possible flexibility. Such a system can be called "Intelligent flexible disassembly cell" (see Figure 3) and it consists of different main modules:

- Disassembly Robot or handling device with special features like path and force control

- Robot Gripper for a wide spectrum of parts with different geometry and dimensions

- Disassembly Tools especially developed for robots

- Components Data Base including data of re-usable and re-manufacturable parts

- Storage Device for tools and parts
- Transport System and feeding system for products to be disassembled
- Clamping Device and fixture system for parts with different geometry and dimensions
- Manual Disassembly Station for specific tasks
- Quality assurance and quality control module [10]
- Sensors for force, torque, visual recognition, position and distance
- Intelligent Cell Control Unit able to process information from extended sensors.

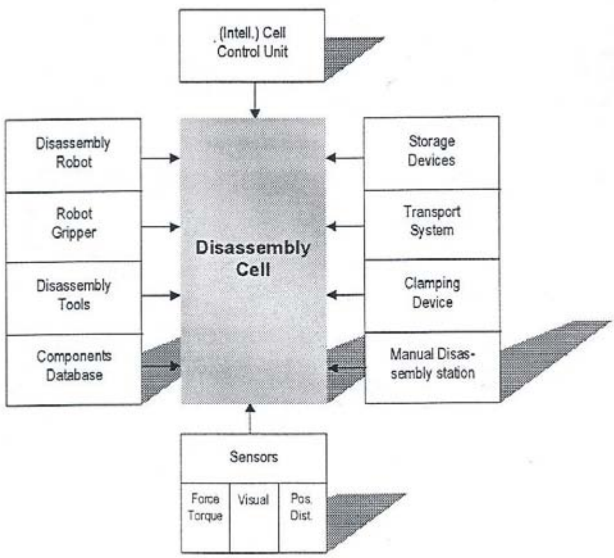


Figure 3: Modules of an "Intelligent flexible disassembly cell".

Flexible semi-automated disassembly can only be carried out economically when the conditions are satisfactory. There must be enough products available at the disassembly cell and the separation technique must be appropriate.

5 Conclusions

It can be summarized that the modern society begins to realize that it is necessary to find economical and ecological solutions in order to reduce industrial waste and to preserve natural resources. It has been proved that the only solution is to recycle technical consumer products, as there are household devices, consumer electronics and passenger cars. In this connection, disassembly of products can be referred to as a first step towards efficient recycling and supports strongly sustainability and TQM in European industry.

References

[1] EN/ISO 14001: Environmental Management Systems - Specification with Guidance for Use (ISO 14001: 1996).
[2] ISO 14001: 2004: Environmental Management Systems - Requirements with Guidance for Use. ISO Standard, 2004.
[3] Perlewitz, H., Seliger, G., Tomiyana, T., Umeda, Y.: Activities of Japanese Industry for Product Recycling. Proceedings of 2nd International Seminar on Re-use, Eindhoven, NL, 1999.
[4] Directive 2000/53/EC of the European Parliament and of the Council of September 18, 2000, on end-of-life vehicles
[5] Penev, K.D.: Design for Disassembly Systems - A Systematic Approach. PhD Thesis, Eindhoven, NL, 1996.
[6] Broete, S.: Disassembly Systems - Process Analysis and Strategic Considerations. PhD Thesis, Linköping, DK, 1998.
[7] Kopacek, B.: The European WEEE Concept - A Contribution to the Recycling and Re-use of Waste from End-of-life Electrical and Electronic Equipment in Europe. PhD Thesis, Vienna-Wien, A, 1999.
[8] Daichendt, K., Kopacek, P., Zebedin, H.: A New Strategy for a Flexible Semi-automated Disassembling Cell of Printed Circuit Boards. Proceedings of International Conference ISIE'2001, Pusan, Korea, 2001.
[9] Tahirova, H.: TQM, LCA und nachhaltige Entwicklung unterstuetzt durch intelligente flexible Metrologie und Demontage zur Wiederverwertung von gebrauchten Produkten in unterschiedlich industrialisierter Umgebung. PhD Thesis, Vienna-Wien, A, 2006.
[10] Bauer, J.M.: Qualitaetssicherung und die Erstellung von Qualitaetssicherungssystemen an Organisationen fuer die universitaere Ausbildung von IngenieurInnen. PhD Thesis, Vienna-Wien, A, 2003.

This page intentionally left blank

Carrying capacity assessment in tourism: the case of the Dodecanese archipelago

B. S. Tselentis[1], D. G. Prokopiou[1], M. Toanoglou[2] & D. Bousbouras[3]
[1]*Department of Maritime Studies, University of Piraeus, Greece*
[2]*University of Surrey, England*
[3]*Hellenic Ornithological Society, Greece*

Abstract

Greece depends heavily on the tourist trade, as tourism is the main economic activity. Carrying capacity assessment has become an indispensable tool for formulating policy and strategies in the tourist industry worldwide. The Dodecanese archipelago has 19 islands; mainly concentrated over than 250.000 beds. The tourist product is a blend of ecological, social and economic sub-systems, operable in the area of interest. For the Greek Islands, environment, both natural and man made, plays a leading role in the sustainable development of the industry. It is the purpose of this paper to apply the principles of carrying capacity assessment to the Dodecanese islands, differing in their tourist development, in an effort to highlight the importance of such a tool in developing long-term sustainable policies for such communities.

1 Introduction

The World Tourism Organisation (WTO) proposes the following definition of carrying capacity: "The maximum number of people that may visit a tourist destination at the same time, without causing destruction to the physical, economic, socio-cultural environment and an unacceptable decrease in the quality of visitors' satisfaction." [1].

Today, controlling tourist growth has become a central policy issue for the tourist trade [2], and it is noteworthy that carrying capacity assessment has

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060221

become an important tool for facilitating planning and developing policy in the industry [3]. The Dodecanese islands have a concentration of over 2 million tourist arrivals per year.

Relevant carrying capacity indicators have been developed and categorised into three groups: 1. physical-ecological, 2. infrastructural and 3. Sociological-psychological [4]. It is highly topical to develop and integrate such indicators [5] into the planning process of the Greek tourist business.

2 Basic population characteristics figures for Dodecanese

Table 1: Population in the Dodecanese: Islands with a population greater than 50 inhabitants.

POPULATION CHARACTERISTICS (2001)				
Island	Population	Area (km^2)	Density inhabitants / km^2	Municipality status
AGATHONISI	158	13,5	11,70	MUNICIPALITY OF AGATHONISSI
ARKI	54	6,7	8,0	MUNICIPALITY OF PATMOS
ASTYPALAIA	1.238	96,85	12,7	MUNICIPALITYY OF ASTIPALAIA
FARMAKONISI	54	3,9	13,8	MUNICIPALITY OF LEROS
HALKI	313	28,12	11,13	MUNICIPALITY OF HALKI
LEROS	8.133	52,95	153,60	MUNICIPALITY OF LEROS
LIPSI	600	15,97	37,57	MUNICIPALITY OF LIPSI
KALYMNOS	13.257	110,8	148,38	MUNICIPALITY OF KALYMNOS
KARPATHOS	5.908	301,17	19,62	MUNICIPALITY OF KARPATHOS, MUNICIPALITY OF OLYMPUS
KASOS	990	65,67	15.10	MUNICIPALITY OF KASOS
KOS	30.947	290,2	106,64	MUNICIPALITIES OF KOS, DIKAIOS, AND HERACLEIDES
MEGISTI	430	9,12	47,15	MUNICIPALITY OF KASTELORIZO
NISYROS	948	41,4	22,90	MUNICIPALITY OF NISIROS
PATMOS	2990	34,05	87,81	MUNICIPALITY OF PATMOS
PSERIMOS	130	20	6.50	MUNICIPALITY OF KALYMNOS
RHODES	117.007	1.398,07	83,65	10 MUNICIPALITIES: ARHAGELOS, ATTAVIROS, AFANDOU, IALYSOS, KALLITHEA, KAMIROS, LINDOS, SOUTH RHODES, PETALOUDES, RHODES TOWN
SIMI	2.606	58,1	44,85	MUNICIPALITY OF SIMI
TELENDOS	54			MUNICIPALITY OF KALYMNOS
TILOS	533	62,82	8,48	MUNICIPALITY OF TILOS

3 Tourist indicators

3.1 Tourist development

Tourist development in Kos started in the early '70s and that in Rhodes started in the 60's. On the other hand, tourist development in Kalymnos, Patmos and Karpathos started in 1989. All development data are presented in table 5.

Table 2: Bed capacity in the Dodecanese related to area and local population.

BED CAPACITY IN DODECANESE (2006) [6]*Islands with population over than 50 inhabitants*						
	Hotels	Hotel's beds	Other hospitality enterprises	Beds in hospitality enterprises	Total beds	Density beds/ km^2
AGATHONISI	-	-	3	34	34	2,52
ARKI	-	-	?	60	60	
ASTYPALAIA	8	270	91	1.073	1343	13,87
HALKI	5	89	38	192	281	9,99
LEROS	26	1.149	60	846	1.995	37.68
LIPSI	2	75	22	234	309	19,35
KALYMNOS	44	1918	102	1.216	3.134	28.29
KARPATHOS	99	4.547	93	1.323	5.870	19,49
KASOS	2	27	5	60	87	1,33
KOS	371	38.947	472	10.436	49.283	169,82
MEGISTI	3	73	18	200	273	29,93
NISYROS	5	233	10	150	383	9,25
PATMOS	43	1826	107	1.091	2917	84.55
PSERIMOS	-	-	?	24	24	1,20
RHODES	511	73.420	906	15.378	88.798	63,51
SYMI	13	481	84	569	1.050	18,07
TELENDOS	1	39	?	56	95	
TILOS	5	234	29	419	653	10,39

Table 3: Closed hotels in the Dodecanese.

CLOSED HOTELS IN DODECANESE [7]	
ASTYPALAIA	1
LEROS	5
KALYMNOS	9
KOS	20
PATMOS	3
RHODES	76
SYMI	1

Table 4: Airports and ports in the Dodecanese.

AIRPORTS AND PORTS IN DODECANESE					
	International airports	National airports	Ports	Organized marinas	Shelters for fishing boats and little ports
AGATHONISI	-	-	1	-	-
ARKI	-	-	1	-	-
ASTYPALAIA	-	1	2	-	1
FARMAKONISI	-	-	-	-	1
HALKI	-	-	1	-	-
LEROS	1	-	-	-	-
LIPSI	-	-	1	-	1
KALYMNOS	-	1	3	-	7

Table 5: Bed capacity history in the Dodecanese: Islands with a population greater than 50 inhabitants.

BED CAPACITY HISTORY 1989-2006									
	Hotel's beds			Beds in other hospitality enterprises			Total beds		
	1989 [8]	1999	2006	1989	1999	2006	1989	1999	2006
AGATHONISI	0	0	0	0	0	34	0	0	34
ARKI	0	0		0	0		0	0	
ASTYPALAIA	128	126	270	326	621	1.073	454	747	1343
HALKI	-	88	89	85	141	192	85	229	281
LEROS	541	1.187	1.149	330	766	846	871	1.953	1.995
LIPSI	28	82	75	0	107	234	28	189	309
KALYMNOS	1.172	1.928	1918	594	1.928	1.216	1.766	3.856	3.134
KARPATHOS	603	3.823	4.547	427	977	1.323	1.030	4.800	5.870
KASOS	32	32	27	8	34	60	40	66	87
KOS	16.227	31.272	38.947	7.760	9.480	10.436	23.987	40.752	49.283
MEGISTI	32	32	73	0	176	200	32	208	273
NISYROS	85	254	233	0	138	150	85	392	383
PATMOS	1.280	1.799	1826	185	776	1.091	1.465	2.575	2917
PSERIMOS	0	0	0	-	1993: 37	24	0	?	24
RHODES	45.059	65.500	73.420	13.127	13.489	15.378	58.206	76.989	88.798
SYMI	220	473	481	113	493	569	333	996	1.050
TELENDOS	0	96	39	0	0	56	0	96	95
TILOS	75	0	234	17	141	419	75	141	653

For tables 5, 6, 7 and 8 we have not any data for: Agathonisi, Arki, Astpalaia, Lipsi, Pserimos and Telendos. Economic crises have taken place in tourist enterprises; a lot of hotels have stopped working (Table 3). As indicated by Table 7, the island of Kos has the most arrivals and overnight stays per 100

inhabitants compared to the other islands. As indicated by Table 8, total airport arrivals do not compare to hotel arrivals, indicating that other lodgings are operable in the islands, such as non-licensed hotels and rooms to let, as well as camping facilities that are not taken into account in the official statistics.

3.2 Tourist arrivals, related to local population, during high and low season

Table 6: Overnight stays.

OVERNIGHT STAYS IN HOTELS AND OTHER ROOMS IN 2003		
	Low season, overnight stays in October	High season, overnight stays in August
LEROS	201	2.345
KALYMNOS	944	14.599
KARPATHOS	4.736	46.358
KASOS	12	59
KOS	359.636	1.077.136
MEGISTI	53	288
NISYROS	529	1.883
PATMOS	991	13.613
RHODES	949.104	1.906.729
SYMI	2.294	1.518
TILOS	15	1.889

Table 7: Tourists related to residents.

TOURIST ARRIVALS IN HOTELS PER 100 INHABITANTS IN 2003				
	Low season, October airport arrivals per 100 inhabitants	High season, August airport arrivals per 100 inhabitants	Low season, overnight stays per day/ 100 inhabitants	High season, overnight stays per day/ 100 inhabitants
ASTYPALAIA	24,31	78.19	-	-
LEROS	10,11	15	2.47	28,83
KALYMNOS	1,95	118,4	0,1	1,95
KARPATHOS	57,74	265,25	80,16	815,88
KASOS	24,04	41,31	1,21	5,96
KOS	403,37	484,54	1.162.10	3480,58
MEGISTI	71,86	187,21	12,33	66,98
NISYROS	-	-	55,80	198,63
PATMOS	-	-	33,14	455,28
RHODES	122,67	255,34	811,15	1.629,58
SYMI	-	-	88,03	58,25
TILOS	-	-	2,81	354,41

Table 8: Comparison of airport arrivals to hotel arrivals.

AIRPORT AND HOTEL AND OTHER ACCOMMODATION ENTERPRISES ARRIVALS IN 2003[9]				
	Arrivals in the airports, October	Arrivals in the airports, August	Arrivals in hotels and other rooms, October	Arrivals in hotels and other rooms, August
ASTYPALAIA	301	968		
LEROS	822	1.221	95	422
KALYMNOS	-	-	503	2.426
KARPATHOS	3.234	15.130	584	5.503
KASOS	238	409	9	32
KOS	124.832	149.952	53.000	137.000
MEGISTI/KASTELORIZO	309	805	17	48
NISYROS	-	-	112	302
PATMOS	-	-	402	2.966
RHODES	143.528	298.752	117.000	256.000
SYMI	-	-	329	1.518

Table 9: Beach impact factor for Rhodes.

BEACH IMPACT FACTOR IN DODECANESE							
Municipalities	Beach length (m)	Inhabitants	Hotel beds	Rooms to let (beds)	Total beds	Seasonal population	Beach impact factor (people/km of beach)
AGATHONISI	0,25	158	-	34	34	192	768
ARKI	-	54		45	45	110	
ASTYPALAIA	3	1.238	270	1.073	1343	2.581	860,33
FARMAKONISI	-	54	-	0	-		-
HALKI	0,6	313	89	192	281	594	990
LEROS	10	8.133	1.149	846	1.995	10.128	1012.8
LIPSI	0,8	600	75	234	309	909	1136,25
KALYMNOS	8	16.441	1918	1.216	3.134	19.575	2,68
KARPATHOS	22	5.908	4.547	1.323	5.870	11.778	535,3
KASOS	0,6	990	27	60	87	1.077	1795
KOS [10]	38.9	30.947	38.947	10.436	49.283	80.230	2,18
MEGISTI	0,3	430	73	200	273	703	234,33
NISYROS	5	948	233	150	383	1.331	266,2
PATMOS	7	2990	1826	1.091	2917	5.907	843.86
PSERIMOS	4	130	-	24	24	154	38,5
RHODES [11]	151	117.007	73.420	15.378	88.798	205.805	2,8
SYMI	4	2.606	481	569	1.050	3.656	914
TELENDOS	2	54	39	56	95	149	74,5
TILOS	5	533	234	419	34	567	113,4

4 Environmental indicators

The Psalidi wetlands at Kos and south of Rhodes are considered as the most environmentally sensitive region of the Dodecanese.

4.1 Beach impact factor

With this indicator we analyse the pressures facing the coastal environment, as they describe the concentration of people visiting and using the facilities of the coastal area, and especially beaches.

Rhodes (2.68 persons per metre of beach) and Kos (6.5) do not face the same pressure on their coastal regions as Lipsi (1136.25) and Leros (1012.8). Rhodes and Kos seem to be the islands with the highest number of beaches with Blue Flag certification (10 in Kos and 31 in Rhodes), indicating that serious attempts have been made to protect the environment and possibly increase competitiveness in the tourist services offered. All the other islands do not have any beaches with Blue Flag certification, a result that agrees well with other indicators that they have slow tourist development.

4.2 Natural environment

In the Dodecanese there are only two established national parks; one is in Tilos and the other is in Kos. To conclude, we also have fourteen environmentally protected areas, which belong to the “Natura” network. These environmentally protected areas, based on the “Natura” program, do not, as yet, have a management scheme and are not governed by a specific establishment or organization. Two of these areas are in Rhodes and one is in Kos [12]. The Natura area in Kos island houses an important bird area of Greece, the Psalidi wetland area. This wetland area is under threat as it is situated in the suburbs of the town of Kos, which is a main tourist area. The park of Tilos has different species of eagles. Many different migratory birds visit Psalidi in Kos and Tilos island within the year.

The regulations that comprise the management options adopted have been developed and implemented by the Ministry of Environment. In Rhodes, one of the protected areas is the well-known area called the “Butterfly valley” which is managed by the local municipality. Rhodian deer and little ponies from the Archagelos are species of special interest to the island of Rhodes and many efforts are underway to ensure that both survive excessive human intervention. The indicators presented above indicate the number of threatened species in proportion to the total number of native species (Table 10) [13].

4.3 Garbage and waste management

Urban waste management (solid and liquid) on all the islands is characterized by the lack of efficiency. Only the large urban centers fulfill the basic requirements of modern waste management installations. It is interesting to note that other smaller settlements do not even have a complete urban waste collection network.

5 Conclusions

In this paper we present selected carrying capacity assessment indicators for Dodecanese Archipelago islands that differ in their tourist development, as well as draw some conclusions as to how these indicators can assist in developing sustainable tourist development policies, in island settings.

Table 10: Protected areas (Natura network) in the Dodecanese related with protected local fauna.

SITE_NAME	**AREA (ha)**	**Manage-ment body of Protected Area**	***Monachus monachus***	***Puffinus yelkouan***	***Larus audouinii***	***Hieraaetus fasciatus***	***Buteo rufinus***	***Falco peregrinus***	***Falco eleonorae***	***Emberiza caesia***
KASOS AND ISLATS	13452,39		v							
CENTRAL KARPATHOS AND ISLETS	9321,90	National Level	v		v	v			v	
NORTH KARPATHOS	11297,96	National Level	v		v	v			v	
KASTELLORIZO, RO, STRONGYLI	1769,68									
RODOS: AKRAMYTIS, ARMENISTIS, ATTAVYROS	27514,59		v							
RODOS: PROFITIS ILIAS - EPTA PIGES - PETALOUDES	11184,40									
NOTIA NISYROS KAI STRONGYLI	4055,74		v						v	v
KOS: AKROTIRIO LOUROS - LIMNI PSALIDI - OROS DIKAIOS - ALYKI-PARAKTIA THALASSIA ZONI	10138,24									
EAST ASTYPALAIA END ISLETS	7027,21		v							
ARKOI, LIPSOI, AGATHONISI AND ISLETS	12407,03		v							
SOYTH AEGEAN ISLETS	4568,46									
PATMOS ISLETS	62,00								v	
AGATHONISI, AND ISLETS	1419,00			v	v		v		v	
LIPSI ANS ISLETS	876,00			v	v		v		v	
ARKI AND ISLETS	502,00			v	v		v		v	
LEROS ISLETS	62,00				v			v	v	
KALYMNOS ISLETS AND TELENDOS	528,00						v	v	v	
KINAROS KAI LEVITHA KAI NISIDES LIADA, PLAKA, GLAROS, MAVRIA	1457,00									
EAST ASTYPALAIA AND ISLETS	1459,00								v	

Table 11: Urban waste and garbage management in the Dodecanese islands.

URBAN WASTE AND GARBAGE MANAGEMENT				
	Inhabitants	Urban waste treatment plant	Percentage of waste treated	Garbage management
AGATHONISI	158	NO	0	Place of sanitary burial
ARKI	54	NO	0	?
ASTYPALAIA	1.238	NO	0	Scrap heap
FARMAKONISI	54	NO	0	?
HALKI	313	NO	0	Dump place
LEROS	8.133	1	0	Dump place
LIPSI	600	1	100%	Place of sanitary burial
KALYMNOS	16.441	NO	0	Dump place
KARPATHOS	5.908	NO	0	Dump place
KASOS	990	NO	0	Scrap heap
KOS	30.947	1	72%	Place of sanitary burial
MEGISTI	430	NO	0	Dump place
NISYROS	948	NO	0	Dump place
PATMOS	2990	1	0	Dump place
PSERIMOS	130	NO	0	Dump place
RHODES	117.007	5	70%	Place of sanitary burial
SYMI	2.606	2	25%	Dump place
TELENDOS	54	NO	0	Dump place
TILOS	533	NO	0	Place of sanitary burial

- The data, presented in the form of tourist infrastructure indicators, highlights the fact that the islands of Kos 169,82 (beds/ km^2), Patmos (84.55 beds/ km^2) and Rhodes (63,51 beds / km^2) have higher tourist concentrations in terms of accommodation. Comparing these three islands to other municipalities, such as most of the Dodecanese archipelago islands, we conclude that the former ones could already be saturated, and have no real capacity for further development in accommodating a greater number of visitors except some municipalities at south Rhodes and west Kos; taking into account other indicators the conclusion is reached that these saturated islands can progress towards developing quality and alternative tourist services. Areas that have not been developed, should not necessarily develop along the same lines as the major tourist attraction areas, but should, at this stage of their development, plan ahead in order that planning and policy development and implementation lead to a truly competitive and environmentally sound business [14].
- Hotel occupancy indicators for Kos and Rhodes, showing 50% during low season and 100% in the peak periods, combined with the fact that the tourist season on the other islands is mainly during the months of August and July, as hotel occupancy in October in Karpathos is 10% of the occupancy in August, lead us to conclude that the tourist industry should aim at extending the tourist

season to include more months, and probably visitors with varied interests and expectations of the beautiful islands [15].

• Kos, through local tourism organizations, has achieved a better balance between high and low season tourist figures compared with tourism in Rhodes and all the other Dodecanese and seen the need for policies and measures aimed at alternative tourist attractions, thus achieving a better quality and extended tourist season [16]. In Karpathos October arrivals (3.234 passengers) are only 20% of the August arrivals (15.130).

• Environmental indictors indicate that the transformation from a low quality, high numbers tourist trade, to an alternative high quality one will not be easy, since waste management systems are insufficient in dealing with urban and solid waste. It is well known that such inadequacies have serious environmental consequences, and hinder any attempts towards developing a tourist industry based on quality. Urban waste treatment plants operate only at Rhodes, Kos, Lipsi and Symi.

• Calculated coastline impact indicators indicate that the island of Kalymnos (with a total of 8 km of beach), Kos 38.9 Km of beach and Rhodes 151 Km of beach, does not face the same pressure on its coastal regions (data in people/km of beach) as Lipsi 1.136,25, Leros 1.012, Astypalaia 860,33, Simi 914, Halki 990 and Patmos 843.86

• The Dodecanese has 14 environmentally protected areas. The increasing public interest in nature and landscape preservation is, today, considered a major positive factor in the tourist development process. It is true that the growing influx of visitors can exert strong pressures on fragile ecosystems [17]. Environmentally sensitive areas should have an effective management scheme [18] and be run under a modern and highly sophisticated regime.

Acknowledgements

We would like to thank for their help and collaboration: Dimitra Tselou, economist, University of Piraeus, Spyros Gavanozis Hellenic Civil Aviation Authority, Loucas Mastis, Chief Editor “Rodiaki” newspaper, Mr Petsis, Department of Tourism, Periphery of South Aegean, Rhodes and Paraschos, Vasiliki, Maria and Matia Zouglas, Rhodes and Nikos Kadarzis, Kos.

References

[1] UNEP/MAP/PAP, 1997

[2] Coccosis H and Mexa A, ‘The challenge of Tourism Carrying Capacity Assesment’, Ashgate 2004

[3] Fernando J Garrigós Simón, Y. Narangajavana, and D. Palacios-Marqués. ‘Carrying capacity in the tourism industry: a case study of Hengistbury Head’

[4] Jurincic I 'Carrying capacity assessment of Slovene Istria for tourism', 2nd International Conference on sustainable planning and development Bologna 2005, Wessex Institute of Technology

[5] Zannou V, 'Guide of Socio-Economic Studies for the Integrated Management of the Water Environment' 1999

[6] Direction of Tourism, Periphery of South Aegean, Rhodes

[7] Mr Petsis, 'Department of Tourism, Periphery of South Aegean, Rhodes'

[8] Logothetis M. Dodecanese economy 1988-89, Athens 1990

[9] Hellenic Civil Aviation Authority, Air traffic statistics for 2003

[10] Prokopiou DG 'A study of environmental impact assessment in Kos and Rhodes coasts' Third Conference in Coastal Management, National Technical University of Athens, Athens 2005(in Greek)

[11] Prokopiou DG – Tselentis BS ' Environmental Impacts of Development in Rhodes' International Conference of sustainable planning and development Skiathos 2003, Wessex Institute of Technology

[12] Hellenic Ornithological Society, www.ornithologiki.gr

[13] Indicators for the sustainable development in the Mediterranean region, Blue Plan – PNUE 2000

[14] 'Rodiaki' newspaper 2006-02-12

[15] Prokopiou DG and Tselentis BS 'Regional development and the islands of Kos and Rhodes - a study of sustainable financial opportunities' 2nd Conference on sustainable planning and development Bologna 2005, Wessex Institute of Technology

[16] Prokopiou DG, MBA Dissertation: 'Regional Investment Study for Kos and Rhodes islands', University of Piraeus, 2005(in Greek)

[17] Papayanis T, 'Tourism carrying capacity in areas of ecological importance'

[18] Prokopiou DG and Tselentis BS 'Proposals for sustainable development and Environmental protection for the island of Rhodes', Rhodes 2003 (in Greek)

This page intentionally left blank

Strategy guidelines for the Croatian petroleum industry based on geopolitical risk assessment

D. Karasalihovic
Department of Petroleum Engineering,
Faculty of Mining Geology and Petroleum Engineering, Croatia

Abstract

Over the last 50 years 79 oil and gas fields were discovered in Croatia with more than 460 million m^3 of oil equivalent (OE), which sets the petroleum industry as a very important part of the Croatian economy. Proven oil and gas reserves were estimated to be 198 million m^3 OE. In the same period Croatia produced more than 155 million m^3 OE. Today there are 43 oil and 13 gas fields. Until the 1970s domestic oil production was sufficient for satisfying the entire Croatian demand for liquid fuels. The fact that current domestic production meets only 20% of oil and around 50% of natural gas demand, implies the necessity for acceptable strategy implementation that will perceive domestic market demands along with global market conditions. Since one company represents the largest part of the Croatian petroleum industry, the basic problem represents compliance of small and medium petroleum companies to heavy business conditions in the global market. This is a consequence of petroleum companies' high sensibility to geopolitical risks. Therefore, the precise evaluation of geopolitical risks in time and the assessment of the impact of the creation of a company business strategy could be crucial for the entire Croatian petroleum industry. The Croatian petroleum industry expressly needs business expansion in the markets of neighbouring countries and recovering oil and natural gas reserves from areas with a higher potential in reserves exploration but also with higher political instabilities. Also, it is pressurized by the necessity of business rationalization and the improvement of profitability. Therefore, geopolitical risk assessment is of great importance for the creation of strategic guidelines that should show a way of meeting domestic hydrocarbon consumption along with achieving a desirable business position especially in the regional market.
Keywords: Croatian petroleum industry, strategic planning, risks assessment, geopolitical risk.

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060231

1 Introduction

Croatian Energy consumption in 2004 was in total 412 PJ and since 1992 there was constant increase in consumption. Annual consumption rate in last few years averaged around 2.2%. The liquid fuels have the highest share of 44%, while the natural gas share was around 25%. The Croatian energy consumption per capita is only 55% of EU15 average. Energy production from primary energy sources in 2004 was in total 204.4 PJ with 1.9% of average annual increase since 1999. In energy structure in 2004 the highest share of 38% was in natural gas production, 38% in hydropower and 21% in crude oil production. Therefore, the primary energy production rate was lower than rate of total energy consumption rate and much lower than, rate of direct energy consumption. These issues represent strategic importance of petroleum industry as part of integral Croatian economy sector [1].

2 Croatian petroleum industry

Croatian petroleum industry is represented by one national oil company, which is mainly state owned in strategic partnership with Hungarian petroleum company MOL. INA is vertically integrated oil and gas company operating in oil and gas exploration and production, refining and marketing of oil products. Its subsidiaries are engaged in LPG business, natural gas transportation and providing integrated oilfield services. INA is a medium-sized European oil company and a significant regional player. The mother company INA Group consists of a few subsidiaries with main activities in exploration and production of oil and natural gas in Croatia and abroad, natural gas import and marketing of domestic and imported gas for industry and gas distributors, refining, petroleum product marketing in country and abroad, marketing of crude oil, diesel and gasoline, services including drilling, over haul and other activities during hydrocarbon exploration and production onshore and off shore, pipeline construction, process facilities construction and maintenance, petroleum lubricants production and marketing and LNG production and marketing.

Concerning company size, market dimension, dispersion and range of business activities, INA represents typical national company on middle and southeast European market. The fact that INA is national company refers to mostly state ownership, and INA has the superior position in oil product trade at domestic market, highly above 50%, with the smaller share at the neighbour markets at the same time.

Strategic goal for Croatian petroleum industry represents annual production of minimum 4 million m^3 OE (oil equivalent) and 100% reserve replacement ratio which is possible only along with development and exploration activities of domestic and foreign exploration areas. Although Croatian areas with commercial discoveries are highly explored, it is important to stress justifiability of further investment into development of new areas, regarding necessity for hydrocarbon reserves replacement, but also investment into brown fields. This is

important considering constant increase of oil prices. Therefore, reserve replacement should be included into business strategy of company whose business success is based on exploration and development of oil and natural gas fields. The minimum annual reserves amount should equal annual production when business performances are measured by reserves replacement ratio (Croatian Ministry of Economy, Labour and Entrepreneurship [2]).

Based on production and consumption data, energy structure and hydrocarbon share have been predicted for the future 20 years [3]. In addition hydrocarbon import projections have been made for the future period (figures 1 – 4).

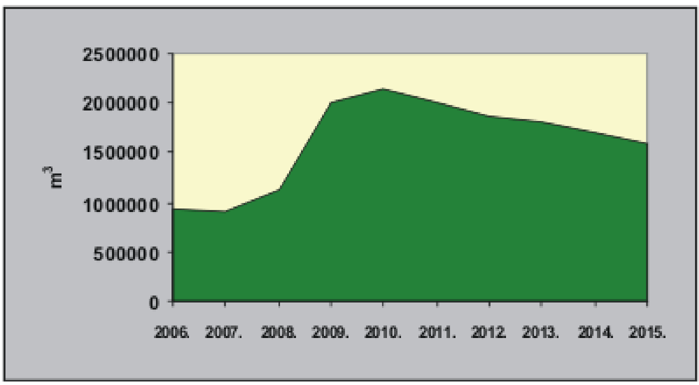


Figure 1: Prediction of oil and condensate production in Croatia in period 2005–2015.

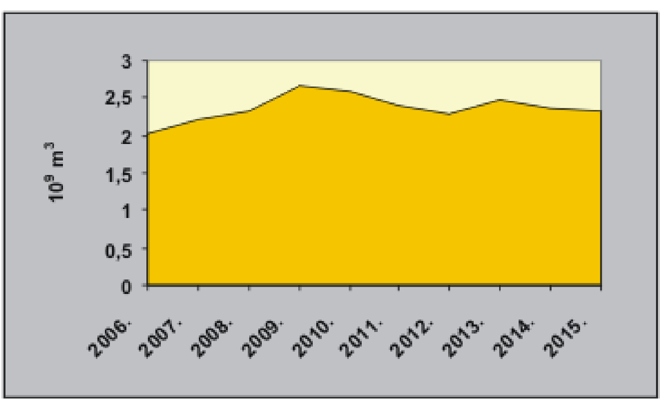


Figure 2: Prediction of natural gas production in Croatia in period 2005–2015.

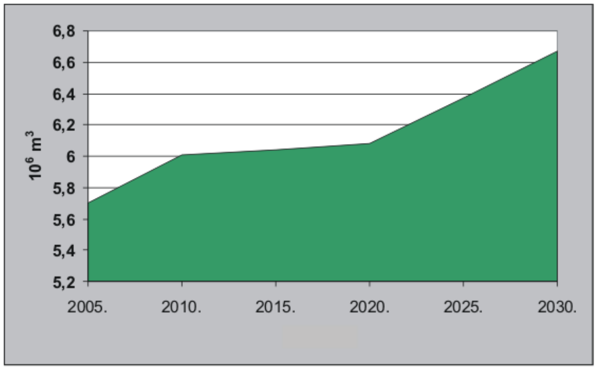


Figure 3: Prediction of liquid fuels consumption in Croatia in period 2005–2030.

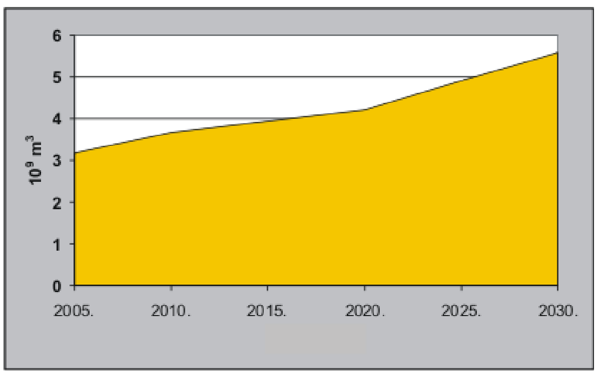


Figure 4: Prediction of natural gas consumption in Croatia in period 2005–2030.

Shown data should be included into supply directions concerning national energy demand. Besides, it is necessary to include consumption increase patterns into strategy guidelines with review of current and planned supply directions for oil and natural gas, concerning Croatian geostrategic position in European supply network, along with its political aspect. It should be emphasized that petroleum industry mid-term investments are made for sufficiency insurance reasons, but also for supply safety reasons. Concerning GDP growth, which was 4–5% annually between 2000 and 2005, the same expectations are that GDP could come to € 7000 per capita. Consequently future economic growth and market development will be depending on energy sufficiency insurance. Predicted energy demand requires significant energy projects planning (Beck [4]).

3 Strategic planning and risk assessment in petroleum industry

Oil companies were the first in world in implementation of long-term strategy planning. During history oil companies were developed as capital intensive with need for very huge investments. Relatively high exploration risks resulted in business practice, which includes mainly or fully proper financing of oil and natural gas project (Seba [5]).

The petroleum exploration, production and refining industry is characterized as a “risk business”. The usual reference is to the geological risk of drilling non-productive wells. With the growing volatility of oil and gas prices, financial risk becomes increasingly important factor, which includes also geopolitical risks. The traditional method of coping with risk has been through methods such as diversification, sheer size and vertical integration of oil and gas production and downstream refining and marketing. There are many varieties of risk including technical, economic and political risks. Based on analysis of oil prices and oil and natural gas reserves distribution in the world the category of geopolitical risk should be introduced into enterprise risk management. Furthermore, it is common practice to weight risk mathematically in order to obtain a proper overall approximation of risk impact, but risks are generally considered as three broad categories namely, technical, economical and political.

Economic risk covers a very broad range of potential situations. The prime economic concerns involve drilling and operating costs, inflationary effects and interest rates, as well as product prices and demand.

Technical risks involve the operational nature of the project and may include engineering capability and experience, hydrocarbon volumes that actually exist underground, producing rates and ultimate recovery.

Political risk can be defined as any political change that alters the expected outcome and value of a given economic action by changing the probability of achieving business objectives. Politics influences how markets operate. Often the most unpredictable economic events are political in origin, the result of flagging political willingness or capacity to maintain a consistent and predictable economic environment. Furthermore, political risks involve uncertainty arising from possible changes in the policies of regulatory authorities and degree to which such changes may affect the project revenues. Regulatory considerations can be subdivided into fiscal and non-fiscal considerations. The fiscal aspects primarily include continuity in the levels of local currencies, changes in levels of customs duties on imported equipment and supplies and possible imposition of locally denominated prices for the production. Non-fiscal political risks may relate to possible interruptions by regulatory authorities over environmental meters, disagreements over hiring or firing of local personnel, determinations of commerciality or outright nationalization. Meters such as the provisions for transfer of operator ship to the national oil companies and the potential for political unrest in the host country also fall under this category (McAfee [6]).

Geopolitical risks is the new category that implies uneven distribution of oil and gas reserves and their situation in regions without strong control of western

industrialized countries which are the main oil and gas consumers. Furthermore, geopolitical risks involve international stability and possibility of terrorist attacks, international sanctions, production output blackmails, fears of shortages but also wide range of other geopolitical issues that can directly impact oil and gas supply, demand and prices and consequently impact companies business performances. This category also relate to unavailability of exploration areas with higher discovery potential to small and medium petroleum companies without internationally strong political backgrounds. There is also different probability for e.g. medium company originated from USA or Croatia for exploration concession gaining in Iraq.

Therefore, risk assessment as important part of strategic planning represents a continual process and way of doing business that helps organizations take more proactive approach to identifying and eliminating exposures and comprehensively blocking attacks while planning and implementing business strategies. It is specifically designed to help define critical assets, assess threats and vulnerabilities, comprehensively block attacks and implement the security measures required to minimize risks.

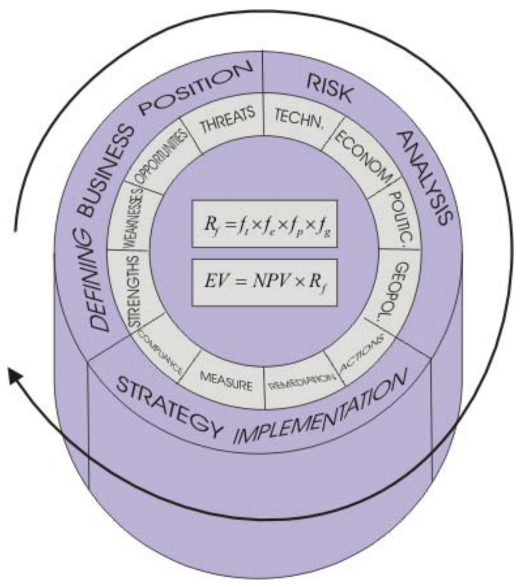


Figure 5: Strategic planning process with geopolitical risks assessment.

3.1 Strategic planning with introduction of geopolitical risk into risk assessment

Based on risk assessment, including all types of risk, management should answer some key questions like: What policies have company established, what assets does company have, and how important are those assets to business goals? What

risk does company face given the vulnerabilities on business critical assets, the threats that could exploit those vulnerabilities, and the risk company is willing to accept? What can management block while planning and implementing strategies based on risk assessment? What can be measured and how to check compliance against original policies to determine the progress making with implemented strategy? (EurasiaGroup [7])

Introduction of new category of geopolitical risks represents incorporation of recommendations that enhance a company's internal capacity to manage global risks, as well as to identify and capitalize on unexploited opportunities. Geopolitical risk assessment helps business leaders to understand the nature of global political risk and its impact on their international investments, and to seize the opportunities it affords. Evaluating company's exposure to risky political events, and assessing their impact, should be key components of any company's strategy (Figure 5).

Conducting a risk analysis turns uncertainty into calculable risk. Because businesses are often affected by political decisions in the countries where they operate, at home and abroad, all companies factor the political environment into planning scenarios. However, geopolitical risk can seem so amorphous that many business leaders lack a framework for evaluating their exposure.

Since security concerns rise, understanding of geopolitical risk becomes increasingly important as terrorism and conflicts in the Middle East and Northeast Asia generate new security policy concerns. The United States is now a major driver of international risk, and United States Government new willingness to pre-empt threats to American security and national interests has changed risk calculations everywhere. Companies must identify whether domestic, regional, or global security threats will affect the cost of doing business. How will those costs compare with doing business elsewhere?

Also energy dependence shows that all energy-importing countries share an interest in diversifying their oil supply, both away from an increasingly unstable Middle East and toward alternative sources of fuel. Common objectives in energy coordination include increasing efficiencies in energy transfer and use and promoting infrastructure efficiencies that avoid bottlenecks and diminish regional variation in energy costs. Understanding how local, regional, and global energy concerns can affect investment decisions requires country-specific knowledge of how political actors will respond to energy shortages.

Furthermore, geopolitical risks include terrorism, energy-price volatility, political instability in the Middle East, weapons proliferation, Northeast Asian security instability, and the role of India and China in the global marketplace. Geopolitical risk analysis informs company how previously mentioned issues directly affect financial markets and long-term foreign investments.

According to PricewaterhouseCoopers and Eurasia Group consultants, political stability is linked to economic vitality. Looking at a map highlighting many of the world's developing nations, the countries with the strongest economies also tend to be the most politically stable (figure 6). This is true even where political stability is a result of a strong, yet market-oriented, autocratic regime (PricewaterhouseCoopers [8]).

By their nature, emerging markets are places where political decisions have a greater effect on markets than economic trends, thus diminishing the value of employing economic guideposts to investment decisions. In politics, risks are more difficult to identify, to measure, and to hedge. Consequently, investors from petroleum industries are extremely concerned with the risks of nationalization, weak legal systems, corruption, and regulatory stability.

By considering political-science theories and financial markets expertise these, business leaders can adjust strategy to deflect adverse affects on operations or take advantage of opportunities. Corporations and governments concerned with the impact of an international issues, such as terrorism or energy supply, need systematic analysis to complement country-specific coverage. By integrating global political risk into the company's strategy, executives can better understand the global exposures and balance the company's risk appetite against achievement of corporate objectives.

The geopolitical risk assessment must have systematic approach for understanding and anticipating how current and future political events could materially affect a company's organization, and thereby helps the company better manage its international exposures.

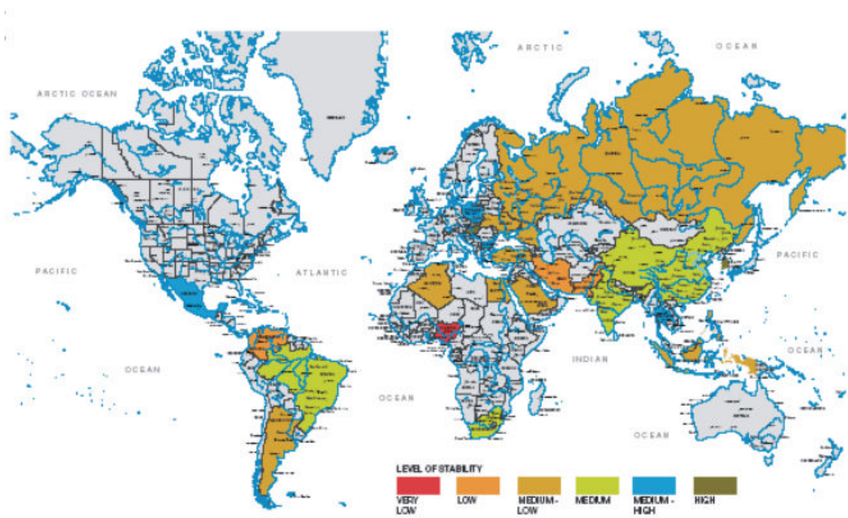


Figure 6: Level of stability in the world according to PricewaterhouseCoopers.

4 Defining strategy guidelines for Croatian petroleum industry

Process of defining strategy guidelines for business strategy of INA should start with enterprise risk management, process that has entered the mainstream of

corporate consciousness over the past decade. Corporations and financial institutions globally have spent a great deal of money to develop and implement systems and processes to assess and manage risk more effectively.

Geopolitical analysis as part of risk assessment should give standpoints of INA's current and future international investments, global supply chains, and key foreign commercial relationships. These issues should be related to global trends, macro-level country risks, and petroleum industry-specific risks to create a comprehensive picture of risk exposure. Second stage of strategy description should include impact analysis that will assess the INA's vulnerability to risks and the potential economic and strategic impacts of risks on company costs and revenues. Furthermore, it is necessary to test qualitative and quantitative risk scenarios and strategic responses. The last strategy faze will give recommendations that will include a plan for mitigating identified risks, pursuing potential opportunities, or seeking alternative strategies with ongoing monitoring of political risks and business compliance issues.

The strategic importance of Croatian national petroleum company and its business rationalization are the reason for implementation of geopolitical risk assessment into company strategy guidelines with the main object of meeting domestic consumption along with achieving desirable business position on regional market.

Therefore, Strategy Guidelines should include further issues:

- Safe supply and production of oil and natural gas,
- Supply directions diversification and adoption of new technologies for exploration and production of hydrocarbons,
- Affordable oil and natural gas prices,
- Increase of energy efficiency,
- Increase of natural gas share in total energy consumption,
- Environmental protection,
- Easy access to regional and European energy markets,
- Development of a clean and efficient exploration and production technologies,

The basic principles of INA's enterprise strategy guidelines for achieving long-term sustainable development of Croatian petroleum industry include implicitly of:

- Increase of income and maximization of the company profit,
- Realizing desirable business performances,
- Achieving strong company competitiveness and export orientation,
- Systematic opening and access to international business environment,
- Further exploration and production of domestic hydrocarbon reserves,
- Firming company position in national social and economic environment,
- Self-financing and decrease of company indebtedness,
- Reserve replacement and investment into new exploration and production technologies,
- Acquisition of foreign concessions,
- Achieving strong position among international petroleum players.

5 Conclusion

In past, business strategy of INA was related to wider market which included the whole area of former Yugoslavia, afterwards it was only oriented to domicile market, and now in present circumstances INA is re-entering neighbour markets. Furthermore, INA's business strategy involves characteristics, typical for any other transitional European company. INA has accomplished business strategy modification with partial restructuring for business improvement under market principles and has started privatisation process. In mentioned circumstances INA's business strategy becomes characteristic for medium-size petroleum company with the careful adoption of global oil market influences, introduction into global oil market through assurance of foreign oil and natural gas exploration and production projects, and finally defending and firming of leading position in refining and marketing of petroleum products at domicile market.

References

[1] www.eihp.hr
[2] Croatian Ministry of Economy, Labour and Entrepreneurship; Energy in Croatia 2003, Croatia, 2003.
[3] www.mingorp.hr
[4] Beck, R.J., Worldwide Petroleum Industry Outlook, 20th Edition, Penwell, US, 2004.
[5] Seba, R.D., Economics of Worldwide Petroleum Production, OGCI Publications, Tulsa, Oklahoma, US, 1998.
[6] Minimizing Business Risk; McAfee, www.mcafee.com
[7] www.eurasiagroup.net
[8] www.pwc.com

About how 'was' 'becomes': emergence of a sustainable spatial-energy system

R. Roggema[1] & A. van den Dobbelsteen[2]
[1]*Province of Groningen / Faculty of Spatial Sciences, University of Groningen, Groningen, The Netherlands*
[2]*Faculty of Architecture, Urbanism & Building Sciences, Delft University of Technology, Delft, The Netherlands*

Abstract

In this paper the relationship between the complexity of a spatial system, sustainability and the type of design principles used is researched. In an analytical framework design projects are phased in timeframes and are placed in a complexity-sustainability field. The conclusion is that we are living in a timeframe in which complex systems are most common and that a new planning phenomenon might emerge based on the right impulse at the right time: swarm planning. The aim to reach a fundamental change in our spatial system, leading to a jump ahead in sustainability, is possible by adjusting our designs and planning system to the new laws of swarm. To find the right sustainable impulses to meet these laws an image of a backtracked landscape is useful. Not taking into account the complexity context we are in and not adjusting our planning system accordingly means heading for failure.
Keywords: planning and design, sustainability, complexity, swarm planning, energy.

1 Introduction

As the World population continues to grow as rapidly as it does nowadays and the growth of wealth in developing countries holds up with the standards of western countries we might need more natural resources then ever will be available on one earth. One side of the solution is that the depletion of resources by the rich communities should be diminished by a factor 20 as Ehrlich and

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060241

Ehrlich [1] and Van den Dobbelsteen [2] describe. In spatial planning this challenge is not yet on the agenda. And the spatial-energy system is not sustainable yet.

As the world becomes rapidly more complex the urge to find simple, one-dimensional, solutions is huge amongst designers. The focus on only one aspect of the whole neglects the complete picture. The effect of this way of looking at things is most likely a non-sustainable situation on other aspects involved. A complex world, a complex society and a complex system require multi-complex principles and solutions to reach a sustainable future.

The relationships between the kind of design principle, the complexity of the system and sustainability are the key issue of this paper.

2 Sustainability

The main aim of sustainability as described by Brundtland *et al.* [3] is to fulfil the demands of the current generation, without making it impossible for future generations to fulfil their demands. If there is a balance between the amount of people in a certain area and the availability of natural resources in that same region for now and in the future, one can speak of a sustainable situation. In an historical perspective a changing impact of sustainability in planning and design processes is described by Roggema *et al.* [4]: the 4 waves of sustainable planning. The solution in the fourth wave is that the design in itself is sustainable. Main characteristic of these design processes is that they do not focus especially or only on environmental issues, but show an integrated development towards a sustainable future. By taking care of the beauty of the site, the social and economic development and the living conditions for man and nature an entire new perspective for the area opens up: the system changes.

3 Complexity

Complexity theory is looking at the functioning of systems. Systems that are more or less complex. If there is minimal complexity we call it order, if there is (too) much complexity we call it chaos. A complex system, that is not chaotic, but at the edge of chaos, is adaptive and self-organising. Mitchell Waldrop [5] describes the characteristics of complex systems. Interesting aspect of this theory for sustainability and planning and design is that a complex system is able to change into a new order if very simple dynamic rules lead to extraordinary behaviour and the edge of chaos is reached.

Based on this it is logical that if a simple dynamic design principle is used at the edge of chaos, extraordinary behaviour takes place and the complex system can change at once: into a new sustainable complex system.

3.1 Typology of complex systems

Which kind of systems are the 'playground' for edges of chaos and the right design principles? To find that, a division in four types of systems might help.

Wolfram [6] distinguishes closed system, linear feed back systems, systems randomly open to assimilation and non-linear adaptive systems. De Roo [7] describes the characteristics of class I-III systems and class IV systems. Mitchell Waldrop [5] points out that behaviour in Class IV systems enables entities to maximise benefits of stability while retaining a capacity to change.

The question is which connection between design and complex systems is appropriate? Design projects contain a large number of interactions, in the design projects simple rules underpin complexity, adaptation, self-organisation and co-evolution is apparent, the design transforms and retains the project and design principles are characterized by robustness, emergence and fitness for purpose. The same characteristics de Roo [7] describes for class IV systems. Above that, the subject of design is often sensitive to impulses, tipping points.

3.2 Swarm planning

The question is which planning approach can be useful if the future consists of mostly Class IV systems (problems, projects). The insights of organisation dynamics can be useful here. Homan [8] describes different types of conditions to improve the overall fitness of the system.

The common characteristic seen in the conditions described before is that they are divers, contain large numbers and have many interactions. There needs to be some kind of large pool of elements. Then, the chance that things interact is larger. Eventually processes like collective patterns, influence on large parts, auto catalytic processes, developing of new structures, etcetera emerge. So, some large pool of elements (there has to be a lot of something), leads to a certain process, which increases the overall fitness of the system.

What seems to be missing is a certain trigger, a focal point that enforces the pool of elements to interact and starts the process of changing the system. These points, the 'flutter the dovecotes' kind of ones, make things happen. Every element in the system orientates itself at these points, and by doing so the system as a whole changes. Exactly this makes from a bunch of ideas an innovation. An impulse needs to be added to the scheme (figure 1).

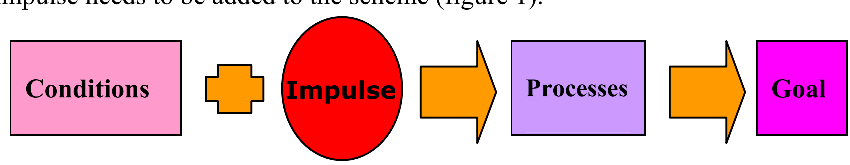


Figure 1: An impulse added.

Translated in planning design terms, the right design principle should be used to orientate all spatial, societal, political elements in such a way that the entire region or city is changing. This requires a new way of planning. Roggema [9] describes how spatial planning can use impulses that will shape the patterns in the future city: swarm planning. The role of spatial planning is to introduce small but essential impulses, hereby influencing the whole system, like a swarm of birds is reshaping itself constantly, depending on current impulses. Spatial

planning is no longer concerned about the whole picture, but needs to focus on those essential design interventions, that enforce the region to reshape itself. No Blueprint, but Acupuncture.

3.3 Tipping points

The tipping point is that magic moment when an idea, trend or social behaviour crosses a threshold, tips, and spreads like wildfire. The possibility of sudden change is at the centre of the idea of the tipping point. Big changes occurring as a result of small events. Gladwell [10] defines three rules of epidemics:

1. The law of the few, a small part of the whole is doing all the work (80/20)
2. Stickiness factor: the message makes an impact. It is impossible to forget.
3. Power of context: sensitivity to the environment. Influence of the surrounding.

If we try to transform these rules to planning and design, the question when a design becomes a success and enforces the required changes can be answered. First of all the law of the few tells us that a design needs to be originated by a small group of individuals. The design is not what the common people expect. To change things the design should be away-from-the average as Ridderstråle and Nordström [11] and Roggema [9] describe.

Secondly, the stickiness factor shows that a design should stick in ones heads. Once you saw the image of the design, you never forget it. A good example of this is the design for Almere Poort, the Wall by MVRDV [12].

Finally, the power of context in relation to design processes will tell us that the design should give the solution to a commonly felt problem. If a fundamental change is required, like the urge for a sustainable system is, the context better be in deep trouble. A sense of urgency is required to change fundamentally. A crisis could give the energy to jump to the new situation. Or as Geldof [13] shows us in a graph: if the existing system dissatisfies, a crisis is required to jump to the next level of complexity and the system is upgraded. These crises can also be seen as the tipping point in design processes.

To condense this: a design group has to come up with a sticky idea, which is/can be used as a tipping point in times of crisis. Or: a crisis has to be arranged to make the time right for a sticky idea, to be designed by a couple of designers.

This approach is not yet common, but the first evidence can be seen growing in the field of spatial planning and urban development. The way injections, like the Blauwe Stad, are planned in the poorer areas of the Groningen province [14], the impact the Öresund-bridge has on the popularity, economic welfare and image of Malmö and Copenhagen or the way Mendini [15] changed the entire inner city in Groningen with the Groninger Museam are early examples of swarm planning.

4 Design principles

For designers and planners it is important to discover the working and effects of design principles. Design principles are general measures that can reach a certain

goal, which can be used by designers in most of the design projects, but is working differently in every situation, depending on the specific characteristics of the project.

Some of the principles are working straight forward on one or a small number of issues: single dimensional design principles. Other design principles are able to change a certain area completely: the multi-dimensional ones. The principle Mendini [15] used to create a good connection between station and city-centre not only aims to stimulate the traffic flow here, but makes a complete change of functional patterns in the area.

The change into a sustainable future, as described in the introduction, enhancing a reduction of the use of natural resources by a factor 20 [1,2], requires design principles that are able to initiate this change.

Which design principles are useful to meet this aim? Principles that:

- Are able to support the sustainability of a spatial design. Or: Using these principles in spatial designs create a desired sustainability.
- Have the power to enforce the change of the entire system (tipping points, change of colour)
- Have an impact on ecological, economic and social functions of the system.

Here 10 principles are (preliminary) selected to measure designs.

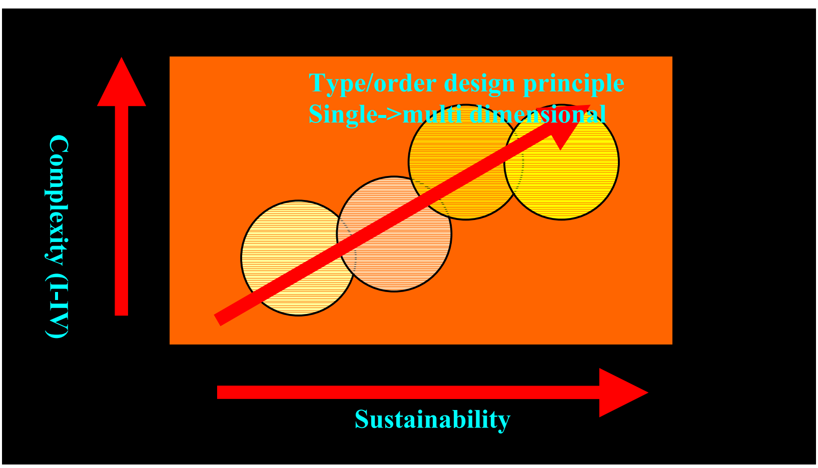


Figure 2: Sustainability, complexity and dimensionality of design projects.

5 Analytical framework

Based on the different analyses on sustainability, design and complexity it is now possible to give an insight in the relative sustainability of different design projects. The question what the sustainability of a certain design is and how several designs can be positioned can be answered by placing them in a graph. In the graph the axes are complexity and sustainability. If a project is complex (class IV) swarm planning can be used to make a sustainable design for it. In

swarm planning we are searching for those design principles that are able to enforce a breakthrough. These design principles are mostly the multidimensional ones. If in a design project more multidimensional design principles are used, these projects are highly complex and highly sustainable. Analysis of the design projects thus, tells us something about the sustainability of them.

6 Design analyses

Three design projects are deeply analyzed. How and if the 10 design principles are used in the different projects is assessed. Roggema *et al.* [16] describe the Grounds for Change project, Bing Yu [17] the Longhu project in Chongqing and the Jinze town project is described by Roggema and van den Dobbelsteen [18]. Beside these thorough analyses 13 detailed designs are roughly analyzed. They are also positioned in the graph.

Design principle	Elucidation Grounds for Change	
1. Backtracking	Historic sustainable functioning was the base for the design for the future.	+
2. Multi energy & LowExergy	In every area of the region several energy sources were combined and used. LowExergy principle is used in appropriate area.	+
3. Multifunction & densification	Combination of energy-functions and urban functions was introduced. Densification was not specifically done (point for further elaboration).	+/-
4. Acupuncture, nodes	Choice of specific projects in the region to catalyse developments was done.	+
5. Landscape a living machine	The capacity of the natural system is a limiting condition for developments. Natural system delivers the energy typology.	+
6. Chaos aesthetics	Regional differentiation is based on energy and landscape typology. Climatic conditions (of the future) are part of the design.	+
7. Visualisation of processes	Process of LowExergy and cycles like tides are made visible (could be extended).	+/-
8. Niche planning	Unexpected chances for specific places (Grolloo, Hoogezand, Lauwerslake, Leeuwarden North, Peat colony) emerge in the design.	+
9. Habitat approach	Dominance of sustainable energy and climate change, effects on social-economic and living conditions.	+/-
10. Symbiosis city-landscape	City and landscape strengthen each other. The city is networked in the landscape and the landscape steers urban developments.	+

Figure 3: The use of different design principles in Grounds for Change.

6.1 Grounds for Change

In the Grounds for Change project a design approach for the Northern Netherlands is undertaken, where-in energy-potentials in the region steer the spatial differences [16].

To find the different energy typologies a new method was developed. The usual negligence of local strengths – climate, landscape, nature, cultural-

technical elements – in regional and urban planning, especially with regard to the potential for local self-provision of energy triggered the emergence of this method. The paper 'Using the Full Potential' by van den Dobbelsteen *et al* [19] describes the Grounds for Change project more extensively.

6.2 Design principles used

In the project the functioning of the natural system was taken as a starting point. Sustainable functioning in history, with tidal influence, natural flows of water, living on the higher grounds, was taken as the base for up to date and futuristic designs. Doing so, future developments have a sustainable base as well.

The conclusion on the Grounds for Change project is that most of the multi dimensional design principles were used. Energy is here the trigger that steers the design and gives the whole system, including sub-designs, a reason for change into a sustainable energy region.

6.3 Chongqing - Longhu

The Longhu project, in the outskirts of Chongqing, China, is one of the many urban developments around the city. In the existing proposal for the site the urban design restructures the whole area. Existing height-differences were diminished, the functioning of the water-system was re-ordered and natural energy-principles were neglected (ventilation, orientation). In the design-method used to review the design, the natural qualities of the site were the base for the redesign [17]. An integrated model was presented in which the different driving forces were combined with each other and with traffic structure, building typologies and parking solutions.

6.4 Design principles used

The redesign for the Longhu project pays a lot attention to the natural dynamics of the site, the landscape and the climate. Most of the multidimensional design principles were used in the redesign. Water is the trigger that steers the possible redevelopment of the design. Using the principles and rules of water direct the development in a complete different direction than planned: The way the water is treated in this project, functions as the tipping point for the change. The same kind of table can be developed for Longhu as shown earlier for Grounds for Change.

6.5 Jinze town

Jinze Town is a small historic village at the edge of Shanghai province. The enormous developments in the metropolis will have their impact on the future developments in Jinze. A regular Chinese approach would diminish existing culture, landscape and qualities. In the design for the Jinze area the historic fundaments were used to create a new future. The region once was part of the delta of the Jangze river. Living on islands, producing just enough food to feed

everyone and transporting people and goods over the water, the area was in a natural balance. This balance is transformed into a modern and future sound design. Large parts of the area, in use as fish ponds are transformed into wetlands or open water, to clean polluted water: the clean machine. Through traffic is abandoned to the outer parts of the area and around it and water transportation takes the most important place in the traffic system again. Existing resources are used to produce energy; wind, solar and biomass supply most of the demand.

New houses are build in an ecological way and eco-tourism becomes the main economic driver in the region. Several new tourist routes are proposed through the area, over the lakes and besides the historic cultural centre of town and the newly introduced shipyard and ecovillage.

6.6 Design principles used

Instead of just another fast growing neighbourhood of Shanghai this becomes an ecological and water-rich recreational and restful space in the region. The principle of the delta functions as the trigger to change. By putting specific new functions at exact locations (shipyard, eco-village) developments are catalyzed. The spatial system can be turned into a system which stays close to its origins, uses the natural potentials, lives in balance with their surroundings and develops a different kind of prosperity. For the Jinze project the same kind of table van also be developed as shown for the Grounds for Change project.

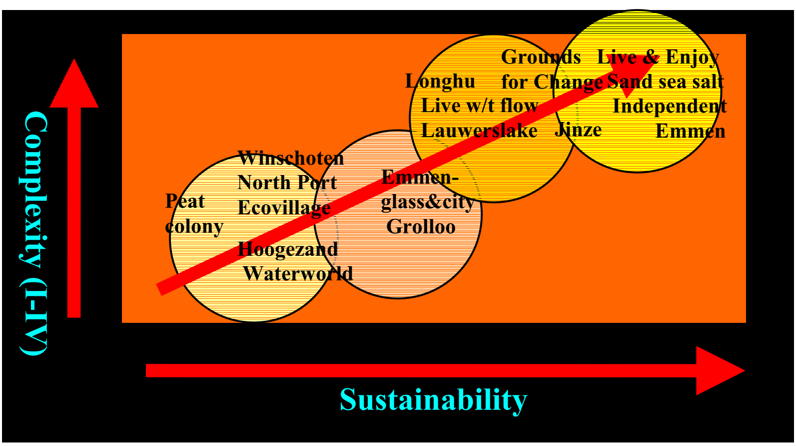


Figure 4: The designs placed in the complexity-sustainability field.

6.7 The designs placed

Figure 4 shows that designs with certain characteristics are found in the top right corner. They all have a base in the natural functioning of the landscape. Secondly, they all show an integrated approach. And the last factor they have in

common is that there is a specific trigger that is able to change the entire system. These acupuncture markers, all multi-dimensional, are all letting the system 'flip' to a next phase in history, reaching a higher level of sustainability. They prove that giving at the right time it is possible to let the swarm adjust to a new situation: swarm planning can be a successful approach.

7 About deltas, energy-mixes and slope-casting

Bringing together theories of sustainability, design and complexity offers a serious chance to influence the spatial(-energy) system of a region, a landscape, an area.

- There is a linkage between the use of multidimensional design principles, the purpose of their usage and the context they are used in. The chance at reaching sustainability in a complex context is higher with multi dimensional design principles.
- Influencing the context to some form of crisis can be useful to enforce the desired changes of the system.
- Multi dimensional design principles can become tipping points if the timing is right (i.e. the system urges for change, in times of crisis).
- If the conditions are right, an impulse is at hand and a process starts unfolding, the system increases its overall fitness (becomes more sustainable).

Finally: What do the three design projects (Grounds for Change, Longhu and Jinze) have in common? They all picked something out of the natural characteristics of the region and upgraded it to a steering principle. By doing so the rest of the spatial system is orienting itself to these principles. So, if we pick the right characteristic, aiming to create a sustainable future, it must be possible to enhance the emergence of a fundamentally more sustainable spatial-energetic system. They all gave the area their reason to exist, separate them from their surroundings and create new and richer spatial typologies. Last but not least they took care of the desired increase of sustainability in the region.

This leads to the following principles of swarm planning:

1. Sufficient conditioning forces (networks, elements, people) have to be available or need to be organised
2. The right impulse has to be picked. This is a multidimensional design principle originated in the natural landscape. Using the backtracking principle these principles can be discovered.
3. A crisis needs to be present, enforced or stimulated. Only then it is the right time to come up with an impulse.
4. Let the process emerge. The system will find its new structures, it will urge for interventions, newly developed areas or retrofits.

The result will be a much fitter and better-adapted system than before so that we are prepared for the future.

References

[1] Ehrlich P. & Ehrlich A. The population explosion, Hutchinson, 1990
[2] Dobbelsteen, A.A.J.F. van den; The sustainable office, TU Delft, 2004
[3] Brundtland G.H. et al, Our common future, Oxford University Press, 1987
[4] Roggema, R.; H. Hofstra; W. Timmermans, From Ecolonia, across the Waalsprong towards the Sun (in Dutch), Handboek ROM, 2002
[5] Mitchell Waldrop, The edge of chaos, Uitgeverij Contact, 1994
[6] Wolfram S.; A new kind of science, Wolfram Media, 2002
[7] Roo G. de; Understanding planning and complexity - a systems approach, AESOP-working group complexity and planning, 2006
[8] Homan, T.; Organisational dynamics (in Dutch), Sdu uitgevers, 2005
[9] Roggema R. Hansje Brinker take your finger away, Oxford Futures Forum, 2005
[10] Gladwell M., The tipping point, Back bay books, 2000
[11] Ridderstråle J.; K. Nordström, Karaoke capitalism, Pearson education limited, 2004
[12] MVRDV, KCAP, R. Roggema; Working City, Development Plan Working Locations, Almere Poort, Department of Urban development, Municipality of Almere, 1999
[13] Geldof, G.; Coping with complexity in integrated water-management (in Dutch), Universiteit Twente, Tauw, 2002
[14] Karelse van der Meer, Urban design Blauwe Stad, Province of Groningen, 2003
[15] Atelier Mendini, Groninger Museum, Fabbri editori, 1994
[16] Roggema R.; A. van den Dobbelsteen; K. Stegenga, Pallet of Possibilities, GfC, 2006
[17] Bing Y. (ed.); TOR-report, Chongqing Longhu Real Estate Development Inc., 2006
[18] Roggema R.; A. van den Dobbelsteen; A blue heart in a red belt, Charette report, Jinze town, Shanghai province, 2006
[19] Dobbelsteen, A. van den; R. Roggema, K. Stegenga, S. Slabbers, Using the Full Potential - Regional Planning based on local potentials and exergy, Paper to the Ravage of the Planet Conference, 2006

Section 4
Ecology

This page intentionally left blank

A sensitivity analysis of the European Union coastal zone based on environmental and socio-economic sustainability indicators

A. Kull, T. Oja & Ü. Mander
Institute of Geography, University of Tartu, Estonia

Abstract

Coastal zones, which have a very high biodiversity value and a long history of being impacted by human activities, are among the most vulnerable ecosystems on our planet. Moreover, some scenarios predict that within 50 years more than 75% of the global human population will live in coastal zones. Although Europe's coasts are among the most altered, there are significant regional differences within Europe in terms of naturalness and human pressure. During the integrated SENSOR project of the European Union's (EU) 6^{th} Framework Programme, we have analyzed the sensitivity of the EU's coastal zones using data on 191 coastal cells at The Nomenclature of Territorial Units for Statistics (NUTS) level, based on information from the Statistical Office of the European Community (EUROSTAT) and other Europe-level sources. To determine sensitive coastal areas, multidimensional clustering was provided. The NUTSx level (a combination of NUTS2 and NUTS3 level cells in order to achieve better spatial coverage) clustering resulted in seven clusters which were not, however, well distinguished in terms of the SENSOR sensitivity areas concept. Therefore clustering on the NUTS2 level was provided using 15 parameters characterizing economic, social and environmental aspects, yielding six clusters that were logically distinguishable from each other. The cluster characterized by a rapidly growing economy combined with relatively high but decreasing unemployment, low but increasing income, a low rate of investment in research and development (R&D), a low share of arable land but a high rural population, moderate length of coastline with the highest share of coast exposed to coastal erosion and a highly variable level of environmental protection, including 23 coastal NUTS2 areas plus five islands, has to be considered sensitive and needs further attention from the European Commission.
Keywords: coastal zone, designated areas, European Union, GDP, sensitivity analysis, SENSOR, sustainability assessment.

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060251

1 Introduction

Coastal zones, which have a very high biodiversity value and a long history of being impacted by human activities, are among the most vulnerable ecosystems on our planet [1]. The expansion of economic activities, urbanization, increased resource use and population growth continuously increase the vulnerability of the coastal zone. This vulnerability is now raised further by the threat of climate change and accelerated sea level rise [2]. Moreover, according to some scenarios, in 50 years more than 75% of the global human population will live in coastal zones [3], while six billion people are already expected to live in the coastal zone by 2025. Although Europe's coasts are among the most altered ones in the world, there are significant regional differences within Europe in terms of naturalness and human pressure [4]. The sensitivity (vulnerability) of coastal areas has been recognized by many authors and authorities. The reason for this is the high pressure from a developing economy (tourism and recreation, rising GDP, traditional coastal fisheries etc.), increasing human population and the environmental sensitivity of the coast. The last issue is related to problems such as the dramatic loss of biodiversity [5], climate-change-driven sea level rise [6, 2], storm and tsunami hazards, increasing erosion and seawater intrusion into freshwater sources [7].

The problems of assessing coastal areas in the development of planning and policies are highly relevant. As presented by several authors [8, 9], integrated coastal zone management (ICZM) is recognised as the most effective tool to incorporate the conservation and sustainable use of marine and coastal biodiversity aspects into the planning of coastal areas. Therefore the growing concerns about the deterioration of the European coastline, environmentally, socio-economically and culturally, have prompted the European Commission (EC) and Member States to introduce a range of measures since 1996. It is intended that these measures will lead to the sustainable development of the whole European coast in the future. The first of these was the Commission's Demonstration Programme. This 3 year programme [10] included 35 individual projects and six thematic studies, embracing the Baltic Sea, the North Sea, the Atlantic seaboard and the Mediterranean Sea, and was launched in 1996. It was a joint programme of the three Directorates General, viz. Environment, Regional Development and Fisheries. Its aim was to test co-operation models for the integrated management of the coastal zones and to stimulate a broad debate among the various stakeholders involved in coastal planning and the management or use of the coastal zones. It was also intended to provide the technical results necessary to foster dialogue between the European Institutions and coastal stakeholders. Based on the results of this programme, the European Commission has subsequently produced two important documents on the subject of the ICZM. The first of these is a strategy for the EC [11] concerning the implementation of the ICZM throughout the EU coastal states. This 38-point strategy consists of a series of concrete actions building upon existing instruments, programmes and resources and is a flexible, evolving instrument designed to cope with the specific needs of the different regions and conditions.

One of the activities (no. 27) calls for the improvement of data provision and the use of this data to produce indicator-based assessment reports. The second document is a recommendation [12], which was called for as the first point of the strategy. This recommendation, although not legally binding, has now been adopted for implementation by all member states. The EU 6th Framework integrated SENSOR project can also be added to this list [13].

The sustainability of coastal areas has been addressed by many authors. Researchers have analysed general questions of the sustainability of coastal zones and options to evaluate sustainability in different regions of the world [14, 15] or compared to the effectiveness of policies in different regions [16]. In some cases specific indicators have been offered for the measurement of sustainability, and applied to certain regions [9, 17, 18]. A number of studies deal with different aspects of coastal zone management [ˇ19, 20], including economic aspects [21] and legislative measures [22]. In particular, in many cases integrated means for management are seen as a useful tool [23–25].

Planning as a tool for management is also found to be useful [26], the importance of public participation in the planning process is underlined [27], and the development of a vision for the future [28] is accentuated. The management of knowledge [29], the resolution of conflicts between different interest groups [30], capacity building in the local community [31] and a clear understanding of the role of stakeholders [32], is found to be critical to achieve sustainability in social aspects.

Numerous papers address methodological questions or specific methods for coastal zone management and research [33], including tools like environmental impact assessment [34], remote sensing and GIS and mapping in coastal zone management [35]. In addition, the methodological aspects of upscaling – downscaling have been addressed in the comparison of local solutions and European perspectives [36].

The impact of tourism and urbanisation has been found to be a very high-pressure factor in coastal areas [37]. More specific aspects of sustainability have also been addressed, for instance water quality [38] and resource management in coastal waters – aquaculture [39], fishing or mussel cultivation [40]. The results of the quantitative assessment of habitat and nutrient fluxes in coastal areas as an environmental factor were analysed [41, 42]. The increased amount of research on coastal areas is demonstrated by many papers reporting different research projects directed towards coastal areas [43, 44]. Also, options for the measurement of sustainability and different indicators useful for management schemes and policies are addressed. For example, Spangenberger and Hinterberger (as presented by Spangerberger [18]) have offered a two-layer indicator system for the European Union, based on the concept of environmental space and four different types of capital. The first layer includes unidimensional key indicators and the second adds indicators linking the key indicators. In another study, Shi et al. [17] have calculated the sustainability of coastal areas in Shanghai, China, using a similar set of indicators.

The objectives of the study are to compile an overview of the location of sensitive regions in European coastal areas and define a methodology for the

identification of key environmental, social and economic sustainability issues. This overview should help test the SENSOR project's Sustainability Impact Assessment Tool (SIAT) [13] and be linked to a methodology and map for European Integrated Spatial Reference Framework and sustainability issues identified in the extensive and intensive social studies.

We consider sensitivity to be a ratio of conflicts between different aspects of development and conservation issues [45]. An example of possible conflicts is given in Table 1.

Table 1: Matrix of possible conflicts between selected sustainability factors. The numbers in the table show the tension level between the factors. NH – natural hazards, NC – nature conservation, D – development, R – restoration, TEU – traditional economic use.

	NH	NC	D	R	TEU
NH	x				
NC	1	x			
D	3	3	x		
R	2	1	2	x	
TEU	2	1	3	2	x

2 Material and methods

2.1 Coastal NUTSx areas

The analysis was performed on the basis of NUTSx regions that have a shoreline. NUTSx regions located no more than 10 km from the shoreline and having access to the sea via a river (Antwerpen, Oost-Vlaanderen, Comunidad Foral de Navarra) were also included. The exploration and evaluation of existing data, information and expert knowledge, results of ongoing and targeted research, and the generation of complementary information on social, economic, and environmental issues were gathered in an iterative way.

2.2 Data sources

Statistical data/indicator values were obtained from EUROSTATS for the NUTS2 level and from national statistics for the NUTS3 level. In addition, available Map Data such as CORINE Land Cover (CLC) 1990 and CLC 2000 were used. CLC changes indicate possible endangered areas (not all areas of change are necessarily sensitive). Available CORINE Biotopes layers were also analysed. Two European R&D projects – LACOAST and EUROSION – have addressed the environmental sensitivity of coastal zones in the 10 km strip along almost the whole coast of the EU. Data from these two projects was used in the study.

LACOAST (a General Directorate Environment of the European Commission in 2002-2004). The analysis was performed on the basis of CLC1990 and

comparative Landsat imagery from 1970, in the case of AC 3 (with the exception of Turkey), Belgium, Denmark, Estonia, Finland, France, Germany, Greece, Ireland, Italy, Latvia, Lithuania, the Netherlands, Poland, Portugal, Slovenia, Spain and the United Kingdom.

EUROSION (a General Directorate Environment of the European Commission project in 2002-2004). The project geocoded CORINE Biotopes for AC 3 (with the exception of Turkey), Austria, Belgium, Denmark, Estonia, Finland, France, Germany, Greece, Ireland, Italy, Latvia, Lithuania, the Netherlands, Poland, Portugal, Spain and the United Kingdom), and actually covers a broader area than just the coast. The data available on the EEA site (http://www.eea.eu.int/main_html) forms the background for the analysis of changes that have taken place after 1990. UNDP Human Development Indicators were used (http://hdr.undp.org/statistics/data/).

2.3 Statistical analysis of the data

The coastal areas were analysed through combined indicators. The data available for NUTSx areas was preferred [47], and in cases where direct data for NUTS3 areas was not available, data from the relevant NUTS2 area was used instead. The indicator values were calculated for all NUTSx cells. Maps were produced using ArcGIS 9.1 and the ETRS89 Lambert Azimuthal Equal Area projection.
Multidimensional clustering (k-means statistics) with selected indicators was provided in order to highlight possible NUTSx areas sensitive to several factors using Statistica 7.

3 Results and discussion

3.1 Environmental and socio-economic factors characterizing sustainability

Environmental values and the loss of highly valued areas characterize sustainability issues. The number of designated areas can be considered to be an indicator of ecological value and awareness (Fig. 1). Also, the total length of the coastline and the proportion of the coast affected by erosion appeared to be important for clustering.

Population density and the changes therein is another important factor that characterizes the pressure and affects sustainability. In multifactorial clustering, population density appeared to be an important factor. A change in land use in the area is one of the leading sustainability factors, with the share of agricultural land and forested land having the greatest effect (Fig. 2). Other aspects of land use (the share of arable land and grasslands and changes in land use as demonstrated by a comparison of land use in 1990 and 2003, also show the sustainability of the cell. Purchasing power and the amount of money circulating in the area are also definitely important factors characterizing the pressures on the environment and thus the sustainability of the area. The increase of GDP in coastal areas is presented in Fig. 3 and 4. The rapid growth of GDP may be considered more dangerous, as the formation of stability requires time, and a

rapid increase in financial means may initiate a development that could endanger sustainability. The rapid growth of GDP is characteristics of the Baltic countries, Ireland, Poland and parts of the UK. Portugal and Greece also have a relatively high increase in GDP.

3.2 Multidimensional clustering

Clustering has been widely used to classify environmental, social and economic phenomena and combinations thereof [47].

In order to determine clusters of similar NUTSx areas, two attempts at cluster analysis were provided using selected indicators characterizing different aspects of sustainability. The unevenness of the data at the NUTSx level caused very high variability, and the clusters distinguished are not informative for the further selection of sensitive clusters. Therefore clustering was performed at the NUTS2 level using the same k-means clustering option.

3.2.1 Clustering at the NUTS2 level

At NUTS2 level we changed the selection of sustainability factors on the basis of our experience from the first clustering attempt. The results of this exercise are presented in Table 2 and Figure 5.

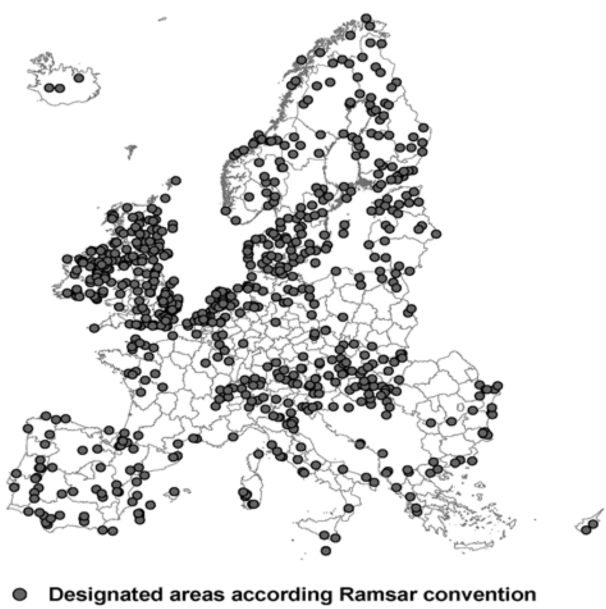


Figure 1: Ramsar sites in the European region. Source: Ramsar Convention Centre.

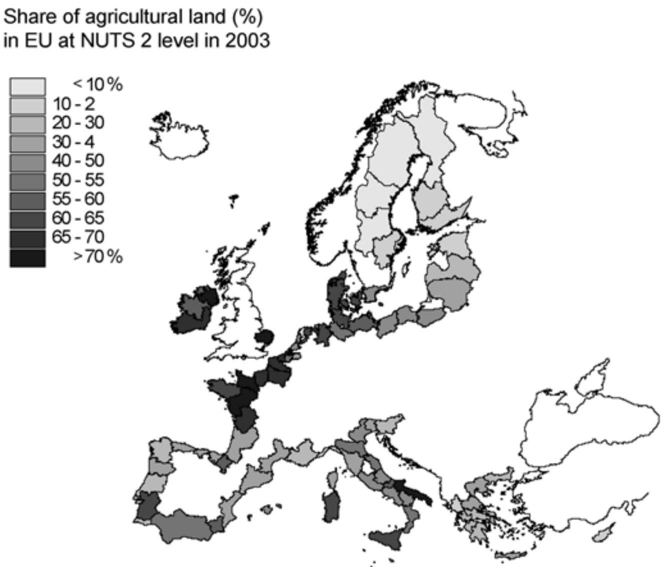


Figure 2: The proportion of agricultural land in the European coastal region. Source: EUROSTAT.

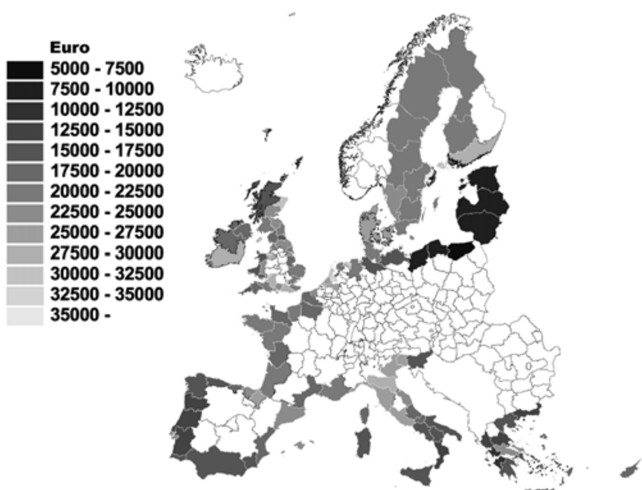


Figure 3: GDP per capita in 2002 (in Euros) by purchasing power standard. Source: EUROSTAT.

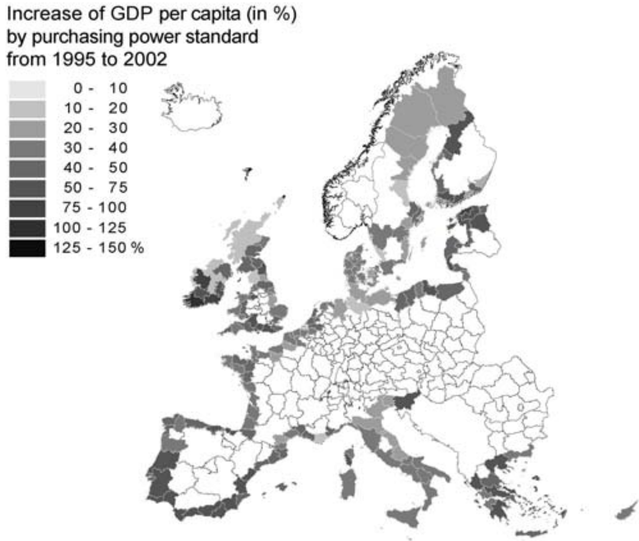


Figure 4: Increase in GDP per capita. Source: EUROSTAT.

The clusters are characterised by the following parameters.

Cluster 1. Economically and socially strong, the rate of development is moderate, but GDP is above the EU25 average. A high proportion of urban population combined with low unemployment, and environmental awareness is high.

Cluster 2. An economy that is stable and developing relatively slowly, with a high share of agricultural activities, a high share of rural population and relatively low GDP. This also causes a high rate of unemployment. Environmental awareness is marginal.

Cluster 3. A moderately developing economy which is slightly influenced by a high unemployment rate, partly caused by the high proportion of urban population. GDP is slightly below the EU25 average.

Cluster 4. Economically and socially very strong, a high proportion of urban population combined with high population density, low unemployment, a long coastline and well-developed environmental awareness.

Cluster 5 (Sensitive, Table 3). Rapidly growing economy combined with relatively high but decreasing unemployment, low but increasing income, and a low rate of R&D investment. A low share of arable land but a high rural population. Despite the moderate length of the coast line, the highest proportion of coastline exposed to coastal erosion. Environmental protection has a high variability within the cluster (from 14% to 2% of area for designated areas). However, in the Baltic Sea coastal zone of this cluster, environmental and nature protection issues are of high importance. Therefore the conflicts between development and nature conservation issues are extremely relevant.

Table 2: Clustering at NUTS 2 level. Parameters used for clustering. POP_DENSIT – Population density; UNEMP_2004 – unemployment rate, %_MEUR – change in GDP between 1995 and 2002 in millions of EUR; 2002_EU25_% – relative GDP compared to EU 25 average; 02_95_CH% – change (%) in GDP per inhabitant by Purchase Power Standard compared to the EU25 average in % from 1995 to 2002; 2002HABEUR – GDP (in EUR) per inhabitant by Purchase Power Standard compared to the EU25 average in %; Agri%_2003 – the percentage of agricultural land; Forest%_2003 – the percentage of forest land in the cell; RAMSAR – the number of Ramsar sites in the cell; R_1000SQKM – the number of Ramsar sites per 1000 sq km; UrbanPop% - the percentage of urban population in the area; Designated% – the share of nationally designated ecologically valuable areas in the cell; R&D2004mEUR – spending on research and development in 2004; CLENGTH_KM – the length of coastline in the cell; COAST_EROD – the percentage of coastline exposed to erosion.

Indicator	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
POP_DENSIT	182.6	161.3	127.5	279.8	183.3	207.1
UNEMP_2004	6.3	13.6	12.1	5.3	10.1	9.5
%_MEUR	43.9	52.1	58.6	55.8	76.5	30.7
2002_EU25_%	109.6	77.3	82.9	142.3	68.8	97.3
02_95_CH%	-2.4	-2.5	6.9	9.1	6.1	-3.7
2002HABEUR	24,014.8	15,635.4	15,072.3	34,006.9	10,834.0	22,009.1
Agri%_2003	36.2	56.6	37.0	40.1	31.8	57.5
Forest%_2003	31.9	18.3	39.6	24.8	26.0	18.2
RAMSAR	5.6	1.9	2.7	8.4	2.0	3.2
R_1000SQKM	0.5	0.09	0.1	0.7	0.1	0.2
UrbanPop%	77.2	67.0	76.0	78.0	57.9	78.4
Designated%	7.1	8.0	8.0	9.4	4.7	15.1
R&D2004mEUR	9999.8	6319.8	8946.0	14,769.0	817.7	20,495.4
CLENGTH_KM	967.5	688.3	679.3	1882.0	723.4	640.2
COAST_EROD	50.6	167.3	80.0	153.1	173.7	151.2

Cluster 6. A traditionally strong and stable economy with a high proportion of agricultural activities. Population density is high and most of the population lives in urban areas. Despite the high living standard, a high level of R&D expenditure is needed to avoid economic stagnation.

For our further activities we can consider Cluster 5 to be the most sensitive, one because of the highest degree of potential conflicts between development and environmental and social factors. Therefore the number of sensitive NUTS2 cells for further SENSOR project activities in coastal areas will be 23 (Table 3).

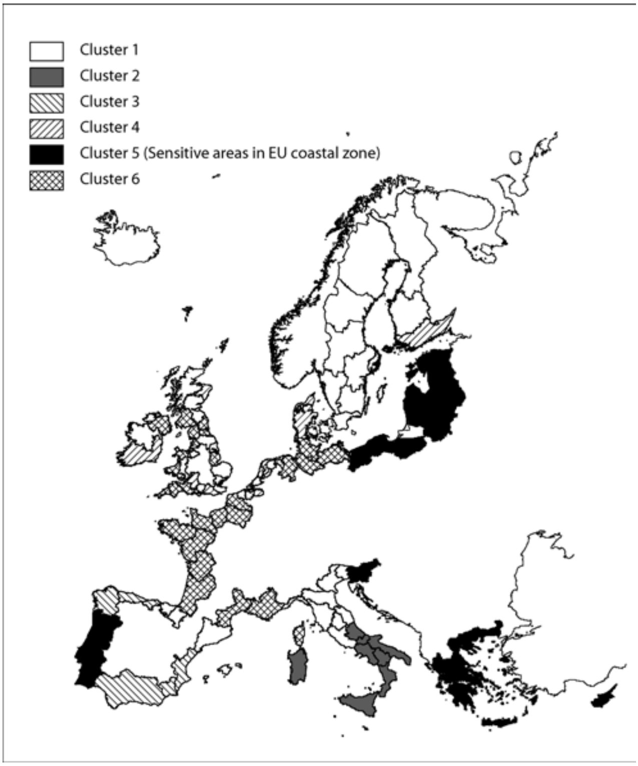


Figure 5: Regional clusters of European coastal zone based on economic, social and environmental indicators.

Table 3: Sensitive coastal NUTS2 cells.

Code	NUTS2 region	Code	NUTS2 region
EE00	Eesti	LT00	Lietuva
GR11	Anatoliki Makedonia, Thraki	LV00	Latvija
GR12	Kentriki Makedonia	PL42	Zachodniopomorskie
GR14	Thessalia	PL62	Warminsko-Mazurskie
GR21	Ipeiros	PL63	Pomorskie
GR22	Ionia Nisia	PT11	Norte
GR23	Dytiki Ellada	PT15	Algarve
GR24	Sterea Ellada	PT16	Centro (P)
GR25	Peloponnisos	PT17	Lisboa
GR30	Attiki	PT18	Alentejo
GR41	Voreio Aigaio	SI00	Slovenija
GR42	Notio Aigaio		

The other 5 NUTS2 areas also belonging to the sensitive cluster 5 are islands and should be handled considering specific island issues. These are CY00 – Kypros / Kibris; GR43 – Kriti; MT00 – Malta; PT20 – Regićo Autonoma dos Açores; PT30 – Regićo Autonoma da Madeira.

4 Conclusions

The NUTSx level clustering resulted in 7 clusters which were not, however, well distinguished in terms of the SENSOR sensitivity areas concept. Therefore we provided the NUTS2 level clustering using the following indicators: population density, unemployment rate, change in GDP between 1995 and 2002 in millions of EUR, relative GDP compared to the EU 25 average, the change (%) in GDP per inhabitant by the Purchase Power Standard compared to the EU25 average in % from 1995 to 2002, GDP (in EUR) per inhabitant by Purchase Power Standard compared to the EU25 average in %**,** the percentage of agricultural land; the percentage of forest land in the cell; the number of Ramsar sites in the cell, the number of Ramsar sites per 1000 sq km, the percentage of urban population in the area; the share of nationally designated ecologically valuable areas in the cell, spending on research and development in 2004, the length of the coastline in the cell and the percentage of coastline exposed to erosion as parameters that gave six clusters that are logically distinguishable from each other. As a result, 6 clusters have been distinguished, among which Cluster 5 was found to be the sensitive one. The cluster consisting of the Baltic Sea coast of Estonia, Latvia, Lithuania and Poland, the Atlantic coast of Portugal, the Greek Mediterranean coast and the Slovenian coast is characterized by a rapidly growing economy combined with relatively high but decreasing unemployment, low but increasing income and a low rate of investment in R&D. There is also a low proportion of arable land, but a high rural population. Despite the moderate length of the coastline, this cluster has the highest proportion of coastline exposed to coastal erosion. Environmental protection has a high variability within the cluster (from 14% to 2% of area for designated areas). In this area the potential risk of conflicts between development and conservation issues, which is the criterion for sensitivity, is highest. This region needs further attention from the European Commission.

Acknowledgements

This paper has been supported by the EU 6th Framework Integrated Project "Sustainability Impact Assessment: Tools for Environmental, Social and Economic Effects of Multifunctional Land Use in European Regions" (SENSOR).

References

[1] Turner, R.K., Subak, S., Adger, W.N., Pressures, trends, and impacts in coastal zones: Interactions between socioeconomic and natural systems. *Environ. Manage*. 20(2), 159-173, 1996.

[2] Vellinga, P., Klein, R.J.T., Climate change, sea level rise and integrated coastal zone management – an IPCC approach. *Ocean Coast. Manage*. 21(1-3), 245-268, 1993.
[3] Small, C., Nicholls, R.J., A global analysis of human settlement in coastal zones. *J. Coast. Res*. 19(3), 584-599, 2003.
[4] Loveland, T.R., Reed, B.C., Brown, J.F., Ohlen, D.O., Zhu, Z., Yang, L., Merchant, J., Development of a Global Land Cover Characteristics Database and IGBP DISCover from 1-km AVHRR Data. *Int. J. Remote Sens*. 21(6-7), 1303-1330, 1998.
[5] Shi, C., Singh, A., Status and interconnections of selected environmental issues in the global coastal zone. *Ambio* 32(29, 145-152, 2003.
[6] Fletcher, C.H., Sea-level trends and physical consequences – applications to the United States shore. *Earth Science Reviews* 32(2), 73-109, 1992.
[7] Michael, H.A., Mulligan, A.E., Harvey, C.F. 2005. Seasonal oscillations in water exchange between aquifers and the coastal ocean. *Science* 436(7054), 1145-1148.
[8] Fabbri, K.P., A methodology for supporting decision making in integrated coastal zone management. *Ocean Coast. Manage*. 39(1-2), 51-62, 1998.
[9] Pickaver, A.H., Gilbert, C., Breton, F., An indicator set to measure the progress in the implementation of integrated coastal zone management. *Ocean Coast. Manage*. 47(9-10), 449-462, 2004.
[10] European Commission. Demonstration programme on integrated management of coastal zones, 1996; document XI/102/96.
[11] Commission of the European Communities. Communication from the Commission to the Council and the European Parliament on integrated coastal zone management: a strategy for Europe, Brussels COM 547, final, 2000.
[12] Commission of the European Communities. Council Recommendation of the European Parliament and of the Council of 30 May, 2002 concerning the implementation of Integrated Coastal Zone Management in Europe, Brussels L 148/24, 2002.
[13] Helming, K., Kuhlman, T., Wascher, D.M., Sieber, S., Tabbush, P., Dilly, O., Firbank, L., Bach, H., Andersen, E., Tscherning, K., König, B., Müller, K., Wiggering, H. Sustainability Impact Assessment of multifunctional land use in the European policy context – the SENSOR project. In: Mander Ü., Helming. K., Wiggering, H. editors. Multifunctional Land Use: Meeting Future Demands for Landscape Goods and Services. Springer, Berlin, 200X Accepted.
[14] Davos, C.A., Sustaining co-operation for coastal sustainability. *J. Environ. Manage*. 52(4), 379-387, 1998.
[15] Söderqvist, T., Eggert, H., Olsson, B., Soutukorva, Å., Economic valuation for sustainable development in the Swedish coastal zone. *Ambio* 34(2), 169-175, 2005.
[16] Juhasz, F., An international comparison of sustainable coastal zone management policies. *Mar. Pollut. Bull. 23*, 595-602, 1991.

[17] Shi, C., Hutchinson, S.M., Xu, S., Evaluation of coastal zone sustainability: an integrated approach applied in Shanghai Municipality and Chong Ming Island. *J. Environ. Manage*. 71(4), 335-344, 2004.
[18] Spangenberger, J.H., Environmental space and the prism of sustainability: Frameworks for indicators measuring sustainable development. *Ecol. Indicators*, 2(3), 295-309, 2002.
[19] Anker, H.T., Nellemann, V., Sverdrup-Jensen, S., Coastal zone management in Denmark: ways and means for further investigation, *Ocean Coast. Manage*. 47, 495-513, 2004.
[20] Irtem, E., Kabdasli, S., Azbar, N., Coastal zone problems and environmental strategies to be implemented at Edremit Bay, Turkey. Environ. Manage. 36(1), 37-47, 2005.
[21] Ehrlich, Ü., Oja, T., Planning of protection zones and management costs in protected coastal areas. In. Brebbia, C.A., Saval Perez, J.M., Garcia Andion, L., Villacampa, Y. editors. Coastal Environment V incorporating Oil Spill Studies. WIT Press, pp. 27-34, 2004.
[22] Dauvin, J.C., Lozachmeur, O., Capet, Y., Dubrulle, J.B., Ghezali, M., Mesnard, A.H., Legal tools for preserving France's natural zone management. *Ocean Coast. Manage*. 47(9-10), 463-477, 2004.
[23] Christie, P., Is integrated coastal management suitable? *Ocean Coast. Manage*. 48, 208-232, 2005.
[24] Laine, A., Kronholm, M., Bothnian Bay Life: towards Integrated Coastal Zone management. *Env. Sci. Policy* 8(3), 259-262, 2005.
[25] Stojanovic, T., Ballinger, R.C., Lalwani, C.S., Successful integrated coastal management: measuring it with research and contributing to wise practice. *Ocean Coast. Manage*. 47, 273-298, 2004.
[26] Davis, B., Regional planning in the US coastal zone: a comparative analysis of 15 special area plans. *Ocean Coast. Manage*. 47(1), 79-94, 2004.
[27] Morf, A. 2005. Public participation in municipal planning as a tool for coastal management: case studies from western Sweden. *Ambio* 34(2), 74-83, 2005.
[28] Geskou, I., Developing the vision: Managing the coastal zone of Magnesia through information and consent. *Coast. Manage*. 31(2), 187-194, 2003.
[29] Barde, J., A metadata service for integrated management of knowledge related to coastal areas. *Multimed. Tools Appl*. 25(3), 419-429, 2005.
[30] Bruckmeier, K., Interdisciplinary conflict analysis and conflict mitigation in local resource management. *Ambio* 34(2), 65-73, 2005.
[31] Barker, A., Capacity building for sustainability: towards community development in coastal Scotland. *J. Coastal Res*. 75(1), 11-19, 2005.
[32] Buanes, A., Jentoft, S., Karlsen, G.R., Maurstad, A., Søreng, S., In whose interest? An exploratory analysis of stakeholders in Norwegian coastal zone planning. *Ocean Coast. Manage*. 47(3), 207-223, 2004.
[33] Sarda, R., Avila, C., Mora, J., A methodological approach to be used in integrated coastal zone management processes: the case of the Catalan coast (Catalonia, Spain). *Estuar. Coast. Shelf Sci*. 62(3), 427-439, 2005.

[34] Matishov, G.G., Denisov, V.V., Kirillova, E.E., Role of a procedure of environment impact assessment (EIA) in elaborating the integrated project of managing the Barents Sea coastal zones. *Ocean Coast. Manage*. 41, 221-236, 1998.
[35] Chen, X.W., Using remote sensing and GIS to analyse land cover change and its impacts on regional sustainable development. *Int. J. Remote Sens*. 23(1), 107-124. 2002.
[36] Cave, R.R., Ledoux, L., Turner, K., Jickells, T., Andrews, J.E., Davies, H., The Humber catchment and its coastal area: from UK to European perspectives. *Sci. Total Environ*. 314-316, 31-52, 2003.
[37] Burak, S., Dogan, E., Gazioglu, C., Impact of urbanization and tourism on coastal environment. *Ocean Coast. Manage*. 47, 515-527, 2004.
[38] Bogdanowicz, R., Drwal, J., Maksymiuk, Z., Osinski, A., Water quality management in the period of considerable water consumption decrease. *Water Sci. Technol*. 44(2-3), 337-342, 2001.
[39] Buck, B.H., Krause, G., Rosenthal, H., Extensive open ocean aquaculture development within wind farms in Germany: the prospect of offshore co-management and legal constraints. *Ocean Coast. Manage*. 47, 95-122, 2004.
[40] Krause, G., Glaser, M., Co-evolving geomorphological and socio-economic dynamics in a coastal fishing village of the Bragança region (Pará, North Brazil). *Ocean Coast. Manage*. 46(9-10), 859-874, 2003.
[41] Laane, R.W.P.M., Brockmann, U., van Liere, L., Bovelander, R., Immission targets for nutrients (N and P) in catchments and coastal zones: a North Sea assessment. *Estuar. Coast. Shelf Sci*. 62(3), 495-505, 2005.
[42] Valiela, I., Tomasky, G., Hauxwell, J., Cole, M.L., Cebrian, J., Kroeger, K.D., Operationalizing sustainability: Management and risk assessment of land-derived nitrogen loads to estuaries. *Ecol. Appl*. 10(4), 1006-1023, 2000.
[43] Carlberg, A., The SUCOZOMA program: results and challenges. *Ambio* 34(2), 62-64, 2005.
[44] Pacyna, J.M., Barrett, K., Namiesnik, J., The RU contribution to global coastal zone research: an ELOISE (European Land-Ocean Interaction Studies) project. *Estuar. Coast. Shelf Sci*. 62(3), 387-389, 2005.
[45] Brody, S.D., Highfield, W., Arlikatti, S., Bierling, D.H., Ismailova, R.M., Conflict on the coast: Using geographic information system to map potential environmental disputes in Matagorda Bay, Texas. *Environ, Manage*, 34(1), 11-25, 2004.
[46] Impact Assessment Guideline (2005). http://ec.europa.eu/governance/impact/docs/SEC2005_791_IA%20guidelines_annexes.pdf
[47] Arhonditsis, G., Karydis, M., Tsirtsis, G., Integration of mathematical modeling and multicriteria methods in assessing environmental change in developing areas: A case study of a coastal system. *J. Coastal Res*. 18(4), 698-711, 2002.

Promoting solar thermal design: the Mechanical Engineering building at the University of New Mexico

A. A. Mammoli[1], P. Vorobieff[1] & D. Menicucci[2]
[1] Department of Mechanical Engineering, The University of New Mexico, USA
[2] Sandia National Laboratories, USA

Abstract

Thermal conditioning of buildings accounts for about half of the energy use in industrialized countries. Thus there is a considerable incentive to use renewable energy, especially solar, to heat and cool buildings. For small residential applications, 'passive' designs often suffice to cover most needs, however for larger, institutional buildings more technological 'active' measures are often required to provide thermal comfort. The Mechanical Engineering building at the University of New Mexico is an example of the latter class of building, in which a thermal solar panel array in combination with an absorption chiller and various energy saving measures results in a substantial reduction of energy requirements. At the same time, the building serves as an educational facility and a research laboratory for training future engineers in the field of renewables. The design of the building is presented, both in terms of its educational purpose and its energy saving measures. The performance of the building is also discussed.

1 Introduction

There is now widespread consensus in the scientific community that human activity (burning of fossil fuels) has already influenced Earth's climate [1], and that severe consequences are likely to result unless corrective action is taken [2]. At the same time, fossil fuel reserves are dwindling, with possibly disastrous effects, based on historical experience [3]. Despite this, political action to alleviate these

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060261

problems is either non-existent (i.e. voluntary restrictions on greenhouse emissions) or largely ineffectual (Kyoto protocol). The reasons for this lack of action are complex and beyond the scope of the present work, however the best hope for a change is simple: education. There is an alarming amount of ignorance and misconception on the topic of energy in general, and renewable energy in particular. This ranges from total pessimism (hot showers will not be possible without burning fossil fuels) to overt optimism (solar powered cars). In reality, the switch from the fossil economy to a renewable (principally solar) economy will require major adjustment in both technology and economic infrastructure [4], and the time available to do this is limited [5]. Arguably, mechanical engineering students are the most effective target for an education in renewable energy, since it is they who will design energy systems of the future.

The transition to a renewables-based economy will necessitate serious changes in socio-economic systems: transportation, energy production and distribution, building design standards, and food production, to name a few. Some of these changes are difficult (e.g. transportation) while others, notably heating and cooling of buildings, are within easy reach of existing technology. In industrialized countries, a large fraction of the energy (approximately 50%) is used in buildings. Of this fraction, a large proportion is consumed by heating, ventilation and air conditioning (HVAC) systems. A relatively small investment in proper design (architectural and mechanical) of buildings of all sizes will bring about massive returns. The design philosophy of the ME building at UNM is based on the principle that students who see that renewable energy is an effective substitute for fossil energy will contribute to the widespread use of renewables more than any energy policy debate.

2 Historical background

The University of New Mexico Mechanical Engineering building was designed and built towards the end of the energy crisis in the 1970s. The building envelope is highly insulated, with small windows (see Fig. 1), so that the load is internally dominated. In addition, massive internal concrete structures (stairwells, slabs) provide very high thermal inertia. A total of eight 50,000 liter water tanks serve as thermal storage, depending on seasonal requirements. In the warm months, the electric chillers 'charge' the tanks by cooling them to near freezing during the night. The cold water is circulated in the cooling coils during the following day, while the electric chiller is not operated. This procedure has a twofold advantage: the chiller COP is higher during the night due to the increased cooling tower performance, and electricity rates are lower, the latter reflecting the power utility company's base load requirements. A solar collector array of approximately 300 m^2 (Energy Design Corporation XE-300 compound parabolic collector) was originally installed on the roof. The heat from the solar array was able to cover approximately 50% of the total building load in the coldest months. In the warm months, heat from the array was used to drive a Rankine cycle system (Barber-Nichols) which in turn drove a generator.

Figure 1: The Mechanical Engineering building at the University of New Mexico in Albuquerque.

As the Albuquerque Journal of October 12, 1980 put it, "The $5.1M structure, UNM's first especially designed to conserve energy, is not only a teaching and office complex, but serves as a built-in laboratory used to educate students in techniques for analyzing energy use in buildings." At the inauguration ceremony in October 1980, Senator Pete Domenici (R-N.M.) said the University of New Mexico's new mechanical engineering building "represents a true step in the future of energy in our nation." Funding for monitoring and operation of the thermal system was provided by PNM (New Mexico's power utility company) and by Oak Ridge National Laboratory, for a period of two years.

Despite the promising building performance and the good intentions manifested in the building, the system was not maintained after the initial funding ended. It experienced a series of problems due to the lack of maintenance, and in combination with the waning interest in energy conservation and renewables beginning in the early 1980s, primarily as a consequence of the availability of extremely cheap fossil energy, it was progressively abandoned. Today, the Rankine engine has been salvaged, the solar collectors have deteriorated beyond repair and the electric chillers are used directly against the cooling load, bypassing the thermal storage.

Until recently, most ME students were completely unaware of the existence of energy conservation measures and solar energy collection which were an integral part of the building they inhabited. This ignorance can be considered as benign, as had the state of disrepair of the system been generally known, it might have contributed to the notion that renewable energy is impractical, expensive and unreliable. Sadly, this largely undeserved perception already extends to the wider community.

3 Modernization and refurbishment

The first fossil energy crisis was largely a product of politics. The impending-current one, on the other hand, is a result of resource depletion. Although there is still debate on the issue, there is mounting evidence that Hubbert's peak, where demand for oil overtakes supply, is either imminent or has already taken place [5]. As a consequence, energy conservation and renewable energy is again a topic of high interest. Funding from the Clean Energy Grants program, managed by the New Mexico Energy, Minerals and Natural Resources department, allowed the authors to begin a project to refurbish and modernize the solar-thermal system in the ME building. The decision to deviate from the original design was based on a combination of factors, first among which the requirement to avoid the equipment failures which were partly responsible for the demise of the original system. In addition, the system would be modernized to reflect of changes in technology that have taken place over the past 25 years. A simplified schematic of the thermal systems in the building is shown in Fig. 2.

One of the major problems with solar thermal installations is that they are generally used to provide a heat source for cold months, while they are unused in warmer months, making the economics less favorable. The original installation made use of the Rankine cycle unit to produce electricity, which could be used for a wide variety of applications including powering an electric chiller. However the Rankine engine has a number of disadvantages, principally complexity (leading to unreliability and vulnerability of the entire system), and low cycle efficiency (partially offset by the high COP of electric chillers). In the new design, in the summer the heat from the collectors is used to drive a single-stage absorption chiller (Yazaki), which will provide 20 tons of cooling capacity for approximately 6 hours / day.

The perception in the engineering community of absorption chillers merits a few remarks. Absorption chillers, single stage ones in particular, are often dismissed as low-efficiency machines, to be avoided if possible. When looked at in isolation, this is true: their typical COP of 0.7 is low compared to the typical COP of 4 for electric chillers. However, electric chillers use electricity, a high grade energy form which is produced predominantly by coal-fired power stations, which have an overall efficiency of 0.3-0.33, not counting losses in the distribution grid. On the other hand, absorption chillers use low-grade heat (moderately hot water possibly by the sun or by waste heat). Viewed in combination with the source of power, the apparently huge thermodynamic efficiency advantage of the electric chiller is reduced considerably. If the impact of the energy source is accounted for, the absorption chiller becomes advantageous. It is hoped that the installation described here will serve to make students aware of the holistic nature of thermal system design.

4 Building performance

The thermal performance of the building was monitored immediately after the beginning of its occupancy, in 1981 and 1982, the effort being funded by a grant

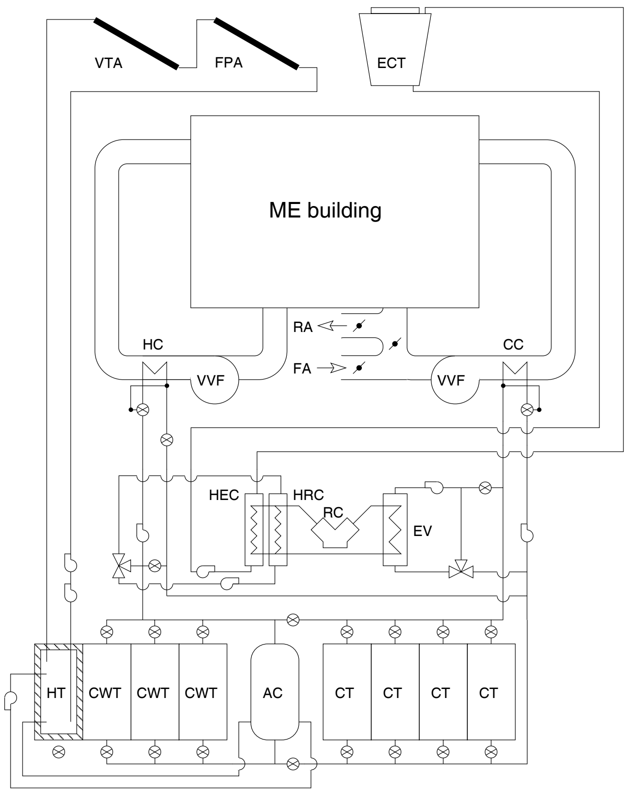


Figure 2: Schematic of the UNM-ME building thermal system. Key: VTA - vacuum tube array; FPA - flat plate array; ECT - evaporative cooling tower; HC - heating coil; CC - cooling coil; VVF - variable volume fan; RA - relief air; FA - fresh air; HEC - heat extraction condenser; HRC - hear recovery condenser; RC - reciprocating compressor; EV - evaporator; HT - hot tank; CWT - cold / warm tank; CT - cold tank; AC - absorption chiller.

from Oak Ridge National Laboratory. Detailed results of the monitoring are available in the report by Wildin [6]. The principal results are briefly summarized here, focusing on the solar system performance and energy consumption of the building.

Data were recorded twice per minute, on a continuous basis, by a computer-based data acquisition system, for 20 months beginning in February 1981. Instrumentation included temperature sensors at the liquid inlets and outlets of most heat exchangers in the system, and a large number of temperature sensors in the stratified thermal storage tanks. Flow rates were also recorded for water, glycol and air in various parts of the system. Electrical energy use was monitored by watt-hour

meters equipped with pulse generators. Finally, precision spectral pyranometers were used to monitor incident solar radiation.

A contractual requirement of the current refurbishment project is a minimum reduction of 20% in the heating and cooling costs for the ME building. Currently, many of the original energy conservation measures, such as the thermal storage tanks, are inactive. The heat pumps are used directly against the cooling and heating loads. The refurbishment will re-activate the original energy conservation systems. and the solar energy collected will be used directly to cool the building. Because the building characteristics have not changed substantially in the past 25 years, and because more recent data are not available, the 1981 data are assumed to be representative of the true building performance. Of particular interest are the heating and cooling loads for the building. These are shown in Fig. 3. The total cooling load is clearly much larger than the heating load. As a consequence, the decision to favor summer heat gain over winter heat gain in the refurbished system was made. The flat plate collectors (124 m^2 absorber area, Lennox LSC18-1) will be mounted at an angle of 25° to the horizontal, while the vacuum tube collectors (108 m^2 absorber area, Sunda Seido 1-16) will be mounted at 35° from the horizontal, equal to the latitude of Albuquerque, NM. This is in contrast to the previous arrangement, in which summer cooling was not implemented, and as a consequence the collectors were mounted at 45°, favoring heat collection in the winter.

Using a purpose-written computer code, the performance of the flat plate collectors and the vacuum tube collectors was estimated as a function of time of year, for both cooling conditions (average collector temperature of 95°C) and for heating conditions (average collector temperature of 60°C). The performance estimate takes into account the daily 30-year averaged temperatures for Albuquerque, as well as the position of the sun and the transverse and longitudinal incidence angle modifiers. In the cooling mode, a COP of 0.7 was assumed for the absorption chiller, while in the heating mode a heat loss of 10% was assumed in the heat storage system.

The predicted cooling and heating capacity of the solar system are plotted along with the measured heating and cooling loads in Fig. 3. From January to March, the solar collectors can provide most of the heat required by the building. In April and part of May, the solar-fired chiller can supply the required cooling to the building. In the hottest months, from mid-May to August, solar cooling accounts for approximately one third of the total cooling load. The fraction of cooling provided by solar energy increases starting in September, until in mid-October the system is switched to winter operating mode.

In 1981, the thermal performance of the building was quite exceptional: its annual purchased energy coefficient of 95.9 $kWh/m^2 \cdot yr$ was approximately one third of the coefficient of the other buildings on campus, and approximately one half of the coefficient of an energy-conservative office building occupied at approximately the same time in Albuquerque. By today's standards, the building is still an excellent performer, using the same energy per unit area as the Adam Joseph Lewis Center for Environmental Studies in Oberlin, Ohio, occupied in 2000 [7].

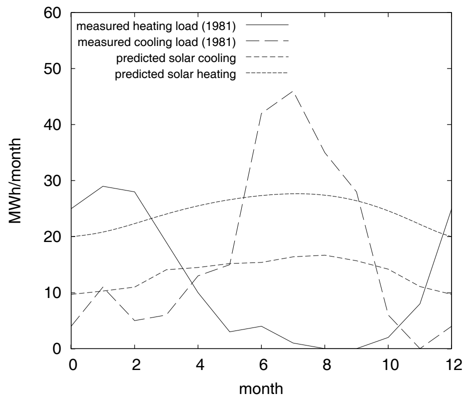


Figure 3: Heating and cooling load for the building measured in 1981, and estimated cooling and heating capacity of the solar thermal system.

The national average energy consumption for academic buildings is currently approximately three times that of the ME building.

The energy consumption of the electric chiller was metered for a period of 48 days beginning in late May 2006. The metered energy is approximately linear in time at about 43.86 kWh/h. The chiller is rated at 110 tons cooling capacity. The COP was measured in the 1982 report by Wildin at 3.8, so that the monthly chilled water energy can be estimated at 124,000 kWh. This is far in excess of what was measured in 1981 for the same period, approximately 42,000 kWh/month. The average energy consumption for 1981 is shown in Fig. 4 for comparison. The discrepancy could be caused by a variety of factors: the management of the HVAC system, in particular the fresh to recirculated air ratio; the increased cooling load of the building, due to the presence of many more personal computers; reduced efficiency of the chiller; meter calibration. The real cause is likely to be a combination of the above, and will be traced, and rectified, over the course of the system refurbishment.

5 Curriculum development and graduate projects

Undergraduate and graduate courses offer the best possibility of challenging conventional wisdom (or lack thereof), namely the mindless reliance on fossil energy. The audience is generally interested, and open to learning new concepts. Incorporating renewable energy concepts into ‘core’ courses such as Heat Transfer, Thermodynamics, and Heating, Ventilation and Air Conditioning Systems, serves the dual purpose of making the courses more interesting, and of introducing students to concepts to which usually they have only had peripheral exposure, perhaps through popular science broadcasting. For example, in the heat transfer course, the solar system, if properly instrumented, can be used as a laboratory to cover the standard

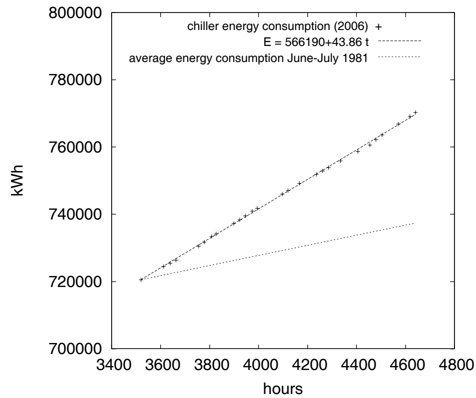


Figure 4: Current chiller energy consumption, compared to the energy consumption in 1981.

experimental program, including conduction, convection (internal and external) and thermal radiation. In each case, the 'real life' nature of the experiment is much more likely to stand out in the students' experience than the conventional 'canned' experiments. ME students are also required to attend a final year design course, where they must design and build a piece of equipment that performs a particular task. The solar system will provide an ideal workshop for this type of activity. Possible projects include night-time radiant coolers, combined PV-thermal panels, and various types of solar collector. A similar situation exists for the Thermal System Design & Optimization course, and in fact the system has already served as the basis for one project in this course.

The ME department also offers three graduate-level courses directly related to the solar thermal system: Conduction, Convection, Radiation, and thermal Systems Design & Optimization. Although these are theoretical in nature, the instrumentation of the solar system will make it possible to conduct experiments to compare theoretical predictions with measured data in great detail. Again, while this is standard practice, generally experiments are conducted on highly idealized laboratory apparatus, while it is not usual for the experiments to occur on a real working system of arguably great current interest, which would both give relevance to the theory associated to the experiment and expose students to the technical intricacies of thermal system design. Two new courses related to the system are planned: Sustainable Energy, and Solar System Design. By the time the curricula have been modified and new courses introduced, it will become possible to offer a certificate program in Solar Design.

Currently, three students are pursuing projects connected with the system: measurement and optimization of solar-assisted heating and cooling (MS), building performance simulation (MS), Holistic economics of solar thermal energy (PhD). The scope of the system is broad enough for several other graduate and under-

graduate level projects, and an unusually high number of students have expressed interest in working on projects related to the system, demonstrating that the lack of research in the area of renewables in not due to lack of public interest, but to scarcity of resources.

6 Perspectives

The solar thermal system in the ME building is a stepping stone towards an engineering program in tune with the energy needs of the 21^{st} century The future of energy production and distribution will see a shift towards localized energy production and distribution (microgrids) as well as towards better load management for the current centralized grid. The U.S. Department of energy has provided funding to build a new control system for the ME building HVAC system, making it possible to respond to requests from the power generation utility to reduce demand in real time. The ME building will be the first of a series of buildings at UNM, and will demonstrate a web interface, automatic demand and response, use smart metering for load modeling and shifting. While load shaving is not entirely a new concept, the Gridwise concept introduces innovation in the sense that the load reduction will not necessarily result in deteriorating thermal conditions in the building, given the thermal storage ability and the solar assisted cooling, and improved forecasting abilities for weather conditions and electricity demand. In addition, real-time negotiation will take place, so that demand reduction in peak times can be offset with future rewards. When this strategy is applied to many building, it will become possible to offset the construction of new power stations intended to satisfy peak demand, and which remain idle for all but a small fraction of the year.

Another possibility is the installation of a photovoltaic (PV) array for the production of electricity to power the lighting, computers and experimental equipment. There are several parts of the building which could be used for such a purpose, including al large patio area, which could easily accommodate 400 m^2 of cells, at the same providing a shaded area for students and faculty. In conjunction with energy-saving measures which will be installed during the course of regular maintenance, the PV-generated electricity could account for a good portion of the energy consumption for the building, it could become an active part of a microgrid, so that the building could become a net energy producer.

A third possibility for energy saving is to use night-time radiant cooling. This concept is gaining recognition in areas of the world where the majority of nights is clear and cool. Also because of the altitude (1618 m a.s.l) the night-time sky temperature is generally very low, and as a consequence it is possible to dissipate heat efficiently. In addition, the radiators can be located in places not suitable for either solar thermal or PV collectors.

Finally, research in the area of heat storage using phase change materials is made simple by the distribution of the thermal storage into several tanks. One of these could be isolated for research purposes without affecting the overall performance of the system.

References

[1] Committee on Surface Temperature Reconstruction for the Last 2,000 years, (ed.) *Surface Temperature Reconstruction for the Last 2,000 years*. National Academies Press: Washington, DC, USA, 2006.
[2] Watson, R. & the Core Writing Team, (eds.) *IPCC, 2001: Climate Change 2001: Synthesis Report. A Contribution of Working Groups I, II, and III to the Third Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press: Cambridge, UK and New York, NY, USA, 2001.
[3] Diamond, J.M., *Collapse: how societies choose to fail or succeed*. Penguin Group: New York, NY, USA, 2005.
[4] Scheer, H., *The solar economy: renewable energy for a sustainable global future*. Earthscan: London, UK, 2002.
[5] Goodstein, D.L., *Out of gas: the end of the age of oil*. Norton and Co. Inc.: New York, NY, USA, 2004.
[6] Wildin, M.W., Results from use of thermally stratified water tanks to heat and cool the mechanical engineering building at the University of New Mexico. Technical Report ORNL/Sub/80-7967/1, Oak Ridge National Laboratory, 1982.
[7] Pless, S. & Torcellini, P., Energy performance evaluation of a low-energy academic building. Technical Report NREL/CP-550-38962, National Renewable Energy Laboratory, 2005.

Sensoristic approach to biological damage and risk assessment

L. Campanella & C. Costanza
Department of Chemistry, University of Rome "La Sapienza", Rome, Italy

Abstract

A review is proposed of various types of sensoristic approaches to biological damage and risk assessment. The traditional approach to the toxicity evaluation of a matrix based on the determination of toxic concentration is replaced by a sensoristic one. Some sensors that are able to detect risky and dangerous situations and to evaluate them by screening and marker procedures are presented. The marker indexes we propose are total organic carbon, total radicalic concentration, integral toxicity and ecopermanence.
Keywords: TOC analysis, ecopermanence, TiO_2 photosensor, biosensor, integral toxicity assessment, human tissue biosensor.

1 Introduction

There are a continuously increasing number of compounds, both newly discovered and already known, which are disposed of into the environment and then pass into the food chain. Very complicated controls and analysis are required to ascertain whether there are any risks to citizens and users.

These analyses are generally very costly so they are not always completed – and if they are completed, it may not be as frequently as is necessary – and they are not permitted everywhere. It is also to be observed that sometimes the obtained results show that the fear of pollution of the considered matrix is disproportionate, as the pollution level is insignificant and below dangerous levels, so that analysis could have been avoided, but obviously this cannot be foreseen.

So a new approach must be more opportunely adopted: the approach of the "crossing light", preliminary screening and markers. The first step of this

www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060271

approach is a simple test which entails waiting for a red or green light that means a negative or a positive sample: if it is green, a detailed and specific analysis can be avoided; if it is red, analyses must be carefully performed in order to find the cause of the negative indication. These analyses can be carried out with a preliminary screening that limits the analysis to samples characterised by values of the key indexes close to the limit values (± 20% can be considered a reasonable range). Finally, when the screening step shows the need for a more detailed investigation, markers can be adopted giving information and capable of ordering the positive samples according to their toxicity. The choice of these markers depends on the type of matrix, but the research advancements able to be dynamically considered individuate continuously new markers for the fundamental role of integral parameters. Finally when the marker test also confirms the danger and the risk of significant toxic levels a complete full analysis has to be performed, generally by hyphenate methods such as GC-MS, HPLC - MS, CE - MS, LCP - MS.

In this paper we want to propose some new instruments for the first steps of the suggested approach. Therefore some biosensors that are able to detect risky and dangerous situations (red crossing light) and to evaluate them by screening and marker procedures are presented. The markers we propose are total organic carbon, total radicalic concentration, integral toxicity and ecopermanence.

The sensors are tools substantially born for analytical aims, the most interesting characteristics of which are the simple operating procedure, the rapidity, the contained costs, the possibility of use in flowing systems, both in situ and with automatic engines.

The analysis of inorganic compounds received particular attention from the scientific community whilst with the organic compounds, the sensor is very often coupled with a biological system, so that a biosensor is being produced. This biological system (enzyme, tissue, micro-organism, whole cells) is sometime immobilised in a membrane. On passing from synthetic to real samples, problems can sometimes arise, as these can contain compounds that are able to deactivate the biological systems (for instance due to enzyme denaturation). The use of biosensors is continuously increasing, especially in the fields of food, chemical, biological, environmental analysis.

1.1 TiO_2 sensors for TOC (Total Organic Carbon) analysis

TOC is an instrumental index, the determination of which consists of the catalytic oxidation of total organic carbon to CO_2, the produced amount of which is assumed to calculate the TOC index. TOC is a primary marker of the healthy or unhealthy state of an ecosystem and a very widely-used datum for monitoring organic pollutants, both natural and synthetic.

Luckily, most of the former are not noxious to humans, but the latter, on the contrary, are generally heavily involved in problems of safety and hygiene. Generally, national or international laws do not give any limit value due to the fact that the same TOC values have different significance according to the compounds accounting for them. So in many cases the profile is more relevant within the time scale: sharp upward variations are surely to be feared and

contrasted. The industrial analysers are based on combustion at high temperatures or on the coupling between a UV reactor and persulphate treatment with conductimetric or infrared determination of produced carbon dioxide.

There is an increasing demand for the measurement of organic carbon in solutions in both industry and environmental research for the purpose of continuous water quality monitoring. Practically all the methods used are based on a catalytic system in which a metal oxide plays a major role. The development of the new TOC measurement system we performed uses a mixed anatase-rutile form TiO_2 as a catalyst to the photomineralisation process, and direct measurement of the CO_2 produced, by a gaseous diffusion electrode. Titanium dioxide is a well-known photocatalyst. Less well-known are its characteristics as a sensor material. In this paper we describe some recent experimental results concerning the adoption of TiO2 (in its prevailing anatase form). The research is based on heterogeneous catalysis using an immobilised catalyst, which can offer a considerable advantage over other methods of catalysis.

Four catalysts were tested [1], all based on titanium dioxide in its mixed rutile-anatase form. In three cases, TiO_2 was used in the P25 nanoparticled form produced by Degussa, often described in literature [1] as an excellent catalyst for photodegradation. Most of the oxide was in the crystalline anatase form. It was tested both supported on a glass or metal grid and suspended in solution. In one of the tested catalysts, the titanium grid was oxidised directly in order to enhance TiO_2 adhesion to the metal support and to obtain a more uniform surface distribution. A highly pure submillimetric titanium grid with suitable mechanical properties was supplied by Delker (CT, USA) [2].

According to the chemometric treatment of the results using PCA (principal component analysis), the catalytic efficiency scale we obtained confirms these expectations and despite the compounds used, it nevertheless demonstrates the superiority of TiO_2 in suspension. Therefore, TiO_2 in the form of anatase nanoparticled suspension was used, in the presence of a phosphate buffer, which was unable to interfere with the TOC value, and operating in such a dimension model that enabled the acquisition of a portable TOC monitor. The cell is Plexiglas and the UV radiation corresponds to 350nm. An auxiliary cell allows a good mixture. GC-MS was applied to determine intermediates and finally check the complete photomineralisation of ions in the organic compounds.

1.2 TiO_2 photosensor of the ability of a compound to be mineralised and of environment assessment

The difference between degradation and mineralization must be emphasised. The former brings a compound to a lower molecular weight, the toxicity of which can also be higher than that of the original compound, and the latter ensuring the total production of CO_2, H_2O and the mineral acids corresponding to the elements present in the mineralized molecule. The potential for the pollutants to be degraded down to mineralization is an important characteristic for the protection of the environment. To evaluate this can be useful for environmental quality assessment.

1.2.1 TiO_2 as a sensor of the potential for a compound to be mineralised

During mineralization, acidification occurs. TiO_2 is a photocatalyst of the process [3], but also behaves as a pH sensor [4] so, at the same time as activating a process, it is also possible to monitor proceedings. The time needed in order to record a pH shift to more acidic values (or a potential shift to more positive values) can be assumed as a delay proportional to the time of recalcitrance of the tested compound (Figure 1).

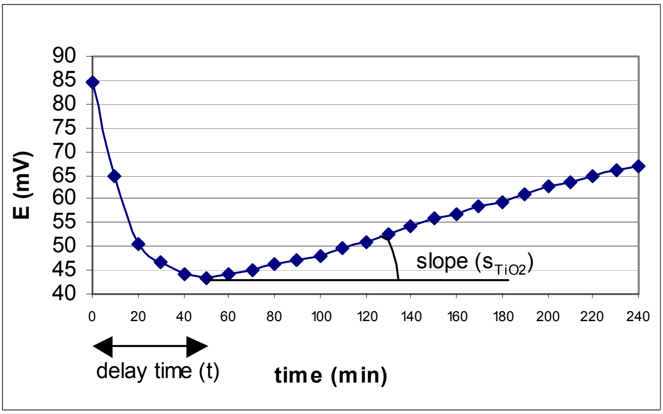


Figure 1: Curve E=f(t), E=TiO_2 potential; t=irradiation time.

Table 1: Delay time and slope of pH decrease (potential E_{TiO2} increase) vs. time of some common pollutants.

Compound and concentration	Delay Time (min) (RSD%=±6)	Slope (ΔpH/Δmin)
p-chlorophenol 10^{-2} mol/L	30	0.15
m-chlorophenol 10^{-2} mol/L	30	0.31
o-chlorophenol 10^{-2} mol/L	60	0.22
Hydroquinone 10^{-2} mol/L	50	0.08
p-quinone 10^{-2} mol/L	20	0.04
Glucose 10^{-2} mol/L	60	0.04
Carbaryl 10^{-2} mol/L	50	0.12

These values agree with other molecular properties related to the stability of the considered compounds.

Further information was obtained by the slope of the pH decrease following the beginning of the mineralization process denounced by the acidification after the delay time. The obtained values are also shown in the last column of Table 1. It can be tentatively assumed that the ratio between the values of the second column and the corresponding last column are an index accounting for the environmental permanence of a compound, that we can reasonably assume to be inversely proportional to mineralizability.

This parameter could be of great help in the case of an unknown or non-characterised compound: a White Book of the European Community invites the scientific community to make the most of efforts, in order to set up chemical

tests that are able to give information – especially alarm advice – in real time (or almost real time) about the toxicity of a compound, the state of a workplace or an environment.

1.3 Biosensors to evaluate free radical scavenger properties

Radical reactions are involved in the processes leading to the breakdown of the lipids contained in most foodstuffs and to their rapid deterioration. They are responsible for ageing processes and cause numerous diseases from which considerable economic loss and potential health problems may follow.

As a consequence of the increasing number of biochemical problems involving free radicals, the use of scavenging and antioxidant compounds in the fields of health, drugs and food has grown. Different screening tests have been developed to determine the antioxidant properties of natural and synthetic antioxidant compounds.

To this end, several studies have been reported involving organic solvents, micelles and liposomes, sometimes using the rate of oxygen uptake via the pressure transducer method, or measures by an oxygen electrode. Successively, methods based on conjugated diene formation, which have a higher sensitivity than oxygen uptake and allow spectrophotometric or fluorimetric measurements are proposed, such as RANSOD, in which the superoxide radical reacts with a derivative of phenyltetrazolium chloride to form a red formazan dye, or those that measure the absorbance of the reduced nitro blue tetrazolium compound. Several in vitro tests have been recently proposed for measuring the antioxidant properties of food products, based on the inhibition of human low-density lipoprotein oxidation, or by using the oxygen radical absorbance capacity (ORAC), used to determine the total antioxidant activity.

Our research group recently approached this problem, beginning with the determination of oxygen free radicals, in particular superoxide radical [5], and assembling several new kinds of electrochemical sensors and biosensors suitable for this purpose; firstly a voltammetric system based on the detection of reduced cytochrome c; this system was also applied to develop a suitable amperometric carbon paste electrode; secondly two potentiometric sensors (one classical selective membrane sensor and the other a solid state field effect transistor sensor), based on selective membrane entrapping benzylidenephenylnitrone with potentiometric detection. More recently, we studied two different kinds of biosensors to determine superoxide radicals obtained by coupling a transducer consisting of an amperometric gas diffusion electrode for the oxygen, or another amperometric electrode for hydrogen peroxide, with superoxide dismutase enzyme immobilized in kappa carrageenan gel. Both the sensors showed a suitable response to the superoxide radical. We think that the second type of biosensor is now mature, both from an engineering and an operative point of view, to check the antioxidant properties of several compounds comparing the response of the biosensor both in the presence and in the absence of the considered scavenger compound.

At this end using the SOD/H_20_2 biosensor set up in our laboratory, we recently evaluated the scavenging properties in vitro of important molecules such

as cysteine, melatonin, β-carotene and acetyl salicylic acid. In addition, we addressed this problem through the study of scavenging properties of fresh vegetal and fruit tissues, or extracts of the same fruits and vegetables present on the market and available in chemists' or herbalists' shops; lastly, the investigation was extended to include several specific active principles contained in the examined vegetal tissues.

We also addressed the problem of the experimental evaluation of any difference in the defence against free radicals (that is, in practice, the antiradical properties) of human kidney tissues both healthy and diseased, following a completely original approach, starting from our experience in designing and building a superoxide dismutase biosensor. In order to apply it, devoted to superoxide radical determination, radicals must be produced in a quantitative way. So we produced radicals by xanthine oxidation (to uric acid) in the presence of xanthine oxidase and measured the scavenging action of human kidney tissue both from a healthy and a cancerous organ, concluding that the antiradical capacity of the former is much higher than the latter.

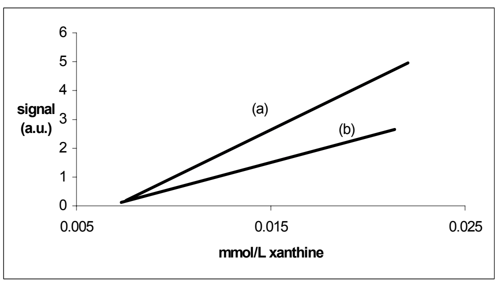


Figure 2: Typical (SOD) biosensor calibration curves, (a) in absence and (b) in presence of antioxidant matrices.

1.4 Biosensor for integral toxicity assessment

For some years, the toxicity of many chemical substances has been studied by means of living organisms used as biological indicators. Higher organisms are generally used, such as the guinea-pig, as the results they produce are often extremely reliable, even though the response times may be lengthy (ranging from a few days to several months, or even years in tests for chronic toxicity). Nevertheless, by exploiting unicellular organisms, in particular yeasts, it is possible to reduce the testing time considerably. Furthermore, by using suitable biological systems, such as immobilised yeast colonies, it is possible to implement the method even in-situ, with the advantage of obtaining a value of integral toxicity as due neither to this nor to that compound but to all these together with the eventual antagonisms and synergisms too.

An immobilised yeast cell biosensor [6, 7] has been developed for the total toxicity testing of a sample that may contain a number of different pollutant species; the biosensor uses amperometric gas diffusion oxygen as an indicator electrode. The method is based on the perturbation of the respiratory activity of a yeast, *Saccharomyces cerevisiae*, immobilised on an agar medium containing the

culture medium (i.e., 'agarised medium'), due to the toxic tested substance. Glucose is used as a substrate while the tested species consisted of several metallic ions, phenols, cationic and anionic surfactants, pesticides and other toxics.

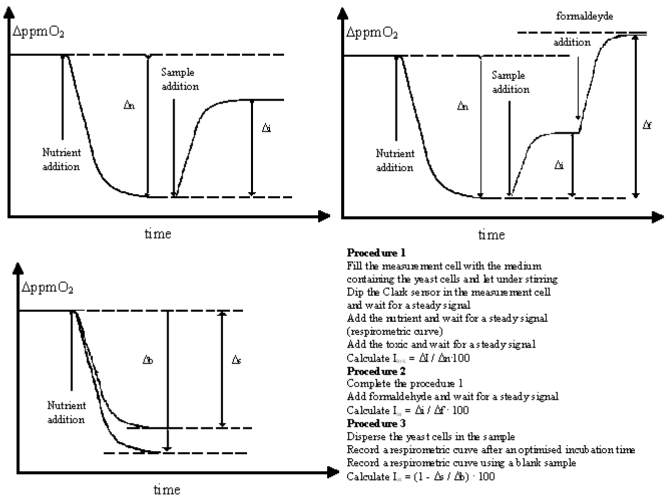


Figure 3: Various respirometric curves recorded in both normal conditions and after exposure of the system to the tested toxic substance.

The measurements are performed by allowing the biosensor to stabilize under magnetic stirring, in the glass cell thermostated at 25° C, containing a fixed electrolytic solution (15ml), isotonic with the yeast cell plasma. The next step is to add a standard glucose solution (final concentration 0.1% w/v), and the decrease of dissolved oxygen concentration caused by the increase in the yeast cell respiration process is recorded as a function of time; in practice, we obtained what is commonly called the 'respirometric curve', which shows the behaviour of oxygen consumption as a function of time. From an examination of the various respirometric curves recorded in both normal conditions and after exposure to the system of toxic substances at different concentrations, we obtained a calibration curve to refer in order to determine the toxicity of the selected sample. In practical terms, for each measurement, two different respirometric curves are recorded: one in the presence of the substrate alone and the other in the presence of the substrate but after an initial incubation with the toxic substance (8 min.); the results can by processed using one of the two following sets of experimental data (Figure 3):

a) difference in final current intensity (δ), which is proportional to the 'distance' between the respirometric curves, at a fixed time (30 min.) after the addition of the glucose;

b) the difference in amplitude of the angles (α - α_0) referring to the slopes by which the two respirometric curves attain a stationary state, which is measured 25min after the addition of glucose.

In this way it is possible to monitor integral toxicity of water from STP, especially if it is destined for drinking water, of soils incidentally polluted and followed during bioremediation of sediments or of industrial media. Two very important applications were performed on radon pollution and on the effects of cellular phones on biological systems, both are the object of controversies due to the contrasting conclusions reached by different authors, very often basing their conclusions on unclear and less accurate measurements. Particularly for cellular phone field it has been concluded that 900 and 1800 MHz cellular phones, corresponding to first and second generation phones, can be considered dangerous if the user is exposed to the effects for longer than 5 hours. In any case, total reversibility was observed after 48 hours since the last exposure.

1.5 Algal biosensor

Eukaryotic and prokaryotic microalgae are regarded as relevant indicators in the field of environmental monitoring and risk assessment. As a result of their easiness of culture and sensitivity to a number of pollutants, they are frequently used in ecotoxicological screening of contaminated fresh and seawater. In toxicity tests, several parameters can be measured to assess the effects of toxicants on microscopic algae. Growth and photosynthetic activity are the most commonly monitored factors [8-9].

Many different chemical pollutants have been assayed for their effects on algal species. Particularly, the interactions between microalgae and trace metals have been extensively studied, with reference to both bioaccumulation [10-11] and toxicity [12]. Microalgae are also used as test species in ascertaining the environmental effects of a variety of synthetic chemicals for pesticide use, primarily herbicides, but also insecticides and fungicides [13]. Most of the proposed devices are based on the electrochemical detection of the inhibiting effect on the photosynthetic activity of algae and cyanobacteria exerted by some toxicants.

Algal biosensors may provide a successful solution to the need for automatic devices and alert systems allowing real time detection and on-line monitoring.

Our analytical device is represented by an amperometric biosensor able to monitor oxygen consumption and evolution of respectively respirometric and photosynthetic algal activity, obtained by coupling a suited algal bioreceptor with a Clark electrode. The selected algal species is the marine filamentous cyanobacterium *Spirulina subsalsa*, which exhibited positive responses in all our preliminary tests and, owing to its morphological features, does not require any immobilization procedure.

Measurements were performed as follows: the sample is allowed to flow in the cell under irradiation cycles (dark/light rate = 1/4), a saw-toothed signal is registered corresponding to the succession of photosynthesis (light) and respiration (dark) processes of the algae; a random variation not exceeding 5% can be obtained in the signal height when unpolluted samples are tested (fig. 4a) during at least 48 hours whilst, in the presence of added pollutants, it decreases as shown in fig. 4b; the time ($t_{1/2}$) required to reduce the height to a half can be related to the toxicity (and of course concentration) of the considered pollutant.

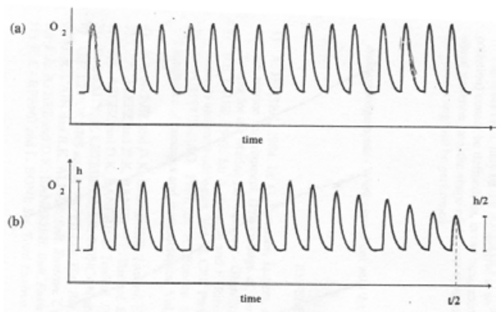


Figure 4: Experimental plot obtained in the analysis of blank (a) and sample (b).

Table 2: Data obtained using algal biosensor in the analysis of organic pollutants.

Pollutant	Concentration	$t_{1/2}$ (min)	RSD %	Pollutant	Concentration	$t_{1/2}$ (min)	RSD %
Atrazine	20 μg/L	690	5.8	p-chlorophenol	$1 \cdot 10^{-4}$ mol/L	1080	0.5
	50 "	290	8.6		$1 \cdot 10^{-3}$ "	312	2.9
	100 "	130	7.7		$1 \cdot 10^{-2}$ "	89	11
	150 "	90	4.4				
Terbutilazine	0.5 μg/L	76	6.5	Trifluoracetic acid	$1 \cdot 10^{-4}$ mol/L	528	0.7
	0.5 "	48	14		$1 \cdot 10^{-3}$ "	147	4.8
	1.0 "	42	7.1		$1 \cdot 10^{-2}$ "	91	7.7
	2.5 "	51	11				
Parationmetil	1.7 mg/L	817	0.1	Banzalkonium chloride	$1 \cdot 10^{-4}$ mol/L	507	2.2
	2.5 "	554	1.0		$1 \cdot 10^{-3}$ "	188	3.7
	3.5 "	325	138		$1 \cdot 10^{-2}$ "	161	0.6
	4.0 "	310	170				
	5.5 "	250	240				
Carbaryl	0.5 mg /L	740	3.9	Tween 80	$1 \cdot 10^{-3}$ mol/L	463	1.1
	1.0 "	370	5.4		$5 \cdot 10^{-3}$ "	415	0.7
	2.5 "	155	6.5		$1 \cdot 10^{-2}$ "	399	1.7
	5.0 "	70	7.1				

Table 2 lists all the obtained results. It can be seen that, as expected, terbutilazine and atrazine were the most toxic compounds respectively (shorter half time at lower concentration); moreover, terbutilazine toxicity remains higher than some of the other considered pollutants that, in turn, result in being more toxic than the other compounds, which means that synergetic and antagonist effects occur. The results show the suitability of the proposed biosensor in the evaluation of sea water quality; today its use can be suggested for the integral toxicity tests of estuaries and bays that are close to doubtful quality matrices such as industrial or civil effluents. It has to be pointed out that the parameter $t_{1/2}$ (time to reach half height of the saw-toothed signal), even if widely used in this kind of test and for this reason, can today be adopted by us in order to compare our results with bibliographic ones, may be not meaningful in an "in situ" analysis where a zero time (starting time of the pollution) cannot be defined.

Therefore in our next research, alternative indexes will be looked for, i.e. the signal variation rate, the shape of the saw-toothed signal and of the whole plot. Chemiometric data processing will also be performed.

References

[1] Campanella L., De Angelis G, Visco G., Chemometric investigation of efficiency response of different TiO_2 based catalysts as principal component of TOC photochemical sensor. *Anal. Bioanal. Chem.*, **376**, pp.467-475, 2003.

[2] DELKER Corporation 14 Commercial Street - PO Box 427 - Branford, CT 06405, www.delkergrid.com

[3] Mills A., Le Hunte S., An overview of semiconductor photocatalysis. *J. Photochem. Photobiol. A: Chem.*, **108**, pp. 1-352, 1997.

[4] Campanella L., Battilotti M., Costanza C., Studies on simulated ageing of paper by photochemical degradation. *Ann. Chim,* **95,** 2005, pp. 727-740.

[5] Campanella L., Favero G., Tomassetti G., Superoxide dismutase biosensors for superoxide radical analysis. *Anal. Lett*, **32**, pp. 2559-81, 1999.

[6] Campanella L., Favero G., Mastrofini D., Tomassetti G., Further developments in toxicity cell biosensors. *Sens. Act. B Chem.*, **44**, pp. 279-285, 1997.

[7] Campanella L., Cubadda F, Sammartino M.P, Tonnina D., Visco G., Integral toxicity test of sea waters by algal biosensor.*Ann.Chim.***92**pp.477-484, 2002.

[8] Jensen A., Marine ecotoxicological tests with phyto-plankton. *Proc. of the International Symposium on Ecotoxicological Testing for the Marine Environment-Ghent, Belgium, 12-14 September 1983*, eds G. Persoone, E. Jaspers, C. Claus, pp. 195-213, 1984.

[9] Puiseux-Dao S., "Phytoplankton model" in ecotoxicology. *Aquatic Ecotoxicoology: Fundamental Concepts and Methodologies*, Vol. II, eds A. cology. A. Boudou and F. Ribeyre, pp. 163-185. CRC Press, lnc., Boca Raton, Florida, 1989.

[10] Azeez P. & Banerjee D., Effect or copper and cadmium on carbon assimilation and uptake or metals by algae. *Toxic. Environ. Chem.*, **12**, pp. 77-86, 1986.

[11] Azeez P. A. and Banerjee D. K., Effect of chromium on cyanobacteria and its accumulation. *Toxic. Environ. Chem.*, **16**, pp. 229-240, 1988.

[12] Cain J. R. & Allen R. K., Use of a cell wall-less mutant strain to assess the role of the cell wall in cadmium and mercury tolerance by *Chlamydomonas reinhardtii. Bull. Environ. Contam. Toxic.* **25**, pp.797 801, 1980.

[13] Padhy R., Cyanobacteria and pesticides. *Residue Reviews*, **95**, pp.1-44, 1985.

Design for human and planetary health: a transdisciplinary approach to sustainability

D. C. Wahl
Centre for the Study of Natural Design, University of Dundee, Scotland, UK

Abstract

This paper explores various integrative frameworks that are contributing to an emerging transdisciplinary meta-perspective on sustainable development. It proposes a holistic/integral strategy based on scale-linking design for human and planetary health: First, ‘Integral Theory’, ‘Spiral Dynamics’ and ‘Integral Ecology’ are briefly reviewed as dynamic mapping methodologies to structure, facilitate and mediate between diverse value systems and perspectives of multiple stakeholders and disciplines. Changes in worldview, value system, and intentionality are crucial to the emergence of a sustainable civilization. Second, design is described as a transdisciplinary integrator and facilitator of informed decision making in the face of uncertainty. Design for systemic health can catalyse the sustainability transition. Third, the paper outlines how complexity theory, combined with a holistic conception of health, informs a scale-linking approach to sustainable design. Systemic health is a scale-linking, emergent property of healthy interactions and relationships within complex dynamic systems. The health of human beings, societies, ecosystems and the planetary life support system is fundamentally interconnected and interdependent. Sustainability, as a process of community-based learning, is expressed through design that is informed by ecological principles and adapted to local, regional and global limits and opportunities. In general, sustainable design is synergetic, symbiotic, scale-linking, salutogenic and sacred. There is a need to integrate ecological, social, cultural, economic and psychological (spiritual) considerations into a flexible and responsive strategy to facilitate the sustainability transition. Design for human and planetary health requires a transdisciplinary dialogue aiming for appropriate solutions and community-based visions of sustainability.

Keywords: scale-linking design, complexity, integral ecology, salutogenic design, transdisciplinary integration, health, sustainable civilization, vision.

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060281

1 Introduction

> "While design became a professional practice with the rise of the industrial culture, more fundamentally, as elemental to mind, it is and always has been, one of the designations of what it is to be human. In this frame, everyone is a designer" [1].

The complexity of interrelated psychological, social and ecological problems that dynamically interact to drive the growing crisis of an unsustainable human civilization cannot be understood or responded to appropriately by compartmentalised, specialized, piece meal thinking. There is a clear need for broadly integrative frameworks that help to create a meta-level synthesis which draws on insights from a diverse range of disciplines and worldviews, and bridges theory and practice. Analytical, reductionist, objectivist, and quantity-focussed perspectives have to be contextualised through integrative, holistic, participatory, and quality-focussed perspectives.

What will affect the transition towards a sustainable human civilization even more profoundly than the necessary changes in our energy and resource use, settlement patterns, production and transport systems, and the local and global economies, are the underlying changes in worldview, value systems, life styles, and intentionality. There is a material (biophysical and ecological) dimension as well as an immaterial (psychosocial and conscious) dimension to the fundamental changes that will steer us towards sustainability.

This paper can only provide the faintest of outlines of the transdisciplinary synthesis attempted in the author's doctoral research. It is but a short summary of a two-volume thesis, entitled *Design for Human and Planetary Health: A Holistic/Integral Approach to Complexity and Sustainability*. The paper introduces some of the central concepts and strategies reviewed, developed, and integrated by this research. It sketches out a tool and a map to promote and structure a transdisciplinary dialogue about sustainable development and the participatory creation of a collective vision of a sustainable human civilization.

2 Mapping and integrating diverse stakeholder perspectives

> "Briefly what I am proposing is that the psychology of the mature human being is an unfolding, emergent, oscillating, spiralling process marked by progressive subordination of older, lower order behavioural systems to newer, higher-order systems as an individual's existential problems change. Each successive stage, wave, or level of existence is a state through which people pass on their way to other stages of being. When the human is centralized in one state of existence, he or she has a psychology which is particular to that state. His or her feelings, motivations, ethics and values, biochemistry, degree of neurological activation, learning system, belief system, conception of mental health, ideas to what mental illness is and how it should be treated, conceptions of and preferences for management, education, economics, political theory and practice are all appropriate to that stage" [2, pp.5-6].

The psychologist Clare Graves first began in the 1950s to map the complexity of human perspectives, worldviews, or dynamics of psychological development. He proposed a rough, but extremely useful, map of different human 'biopsychosocial systems' [3]. The fundamental validity of Graves' cartography of human consciousness has been tested internationally, through countless independent studies involving more than 50,000 people [2, p.6].

Graves' map has been developed further by his students, Don Beck and Christopher Cowan [3] into a framework and methodology for mediation and conflict resolution among diverse stakeholders with often drastically different points of view and value systems. The 'Sprial Dynamics' approach has been used in a wide range of circumstances, including the South African post-Apartheid reconciliation process. It is a tested tool for the integration of and mediation between diverse stakeholders and perspectives. Beck and Cowan argue:

> "Spiral Dynamics apply to a single person, an organization, or an entire society. Since it describes human nature in a universal sense rather than through personality types or racial, gender, and ethnic traits, the model provides a common language for grappling with both local and global problems. It offers a unifying framework that makes genuinely holistic thinking and actions possible" (p.30).

The philosopher Ken Wilber has incorporated, and developed the spiral dynamics approach into an even more encompassing and detailed integrative framework under the name of "integral theory". According to Wilber, the integral approach aims to "include matter, body, mind, soul and spirit as they appear in self, culture, and nature" [2, p.xii].

Wilber's 'integral theory' is trying to make sense of the individual *and* collective, as well as, the exterior *and* interior, aspects of humanity's conscious and co-creative participation in a continuously transforming *kosmos*. Just like the reductionistic, dualistic, materialistic map that defines modernity's dominant scientific and mechanistic worldview, Wilber's integral "theory of everything" is only a *map* albeit much more encompassing and inclusive. Maps remain extremely useful as long as we do not confuse them with the territory [4]. Integral theory can be effectively employed to situate different worldviews, value systems, and disciplinary perspectives and to acknowledge their validity and respective contributions to the dialogue on sustainability.

> "Integral Ecology weaves together the myriad approaches to the natural world in an effort to respond as effectively and timely as possible to the complex ecological problems that face ourselves, our communities, and our world in an evolving universe. In effect, Integral Ecology unites consciousness, culture, and nature in service of sustainability. People who are utilizing the Integral ecology framework recognize that it is not enough to integrate ecosystems and social systems. Instead, what is needed is an integration of subjective (e.g. psychology, art, phenomenology), intersubjective (e.g. religion, ethics, philosophy), and objective" (e.g. behavior, science, systems analysis) realities" [5].

The transdisciplinary scholars Sean Esbjörn-Hargens and Michael E. Zimmermann [6] are among the initiators of this emerging transdisciplinary research initiative which applies integral theory to the sustainability transition. An Integral Institute and an Integral University have been established in Colorado. Integral theory and integral ecology are promising emerging research agendas that are complementary to, and significantly overlap with, the scale-linking, systemic health promoting, design, and metadesign framework described by the research summarized in this paper.

> "Integral Ecology takes a participatory approach to the environment by recognizing that ecological phenomena are the result of an interaction between the knower, what is known, and how it is known. By acknowledging and honoring the multivalent nature of ourselves, our communities, and our environment, we can as global citizens, embedded in local eco-social systems, work effectively together towards sustainable solutions. A premium is placed on solutions grounded in mutual understanding between divergent viewpoints and understandings. By cultivating the capacity to inhabit other perspectives and hold multiplicity, we will be able to respond more adequately than current, less comprehensive approaches to the complex problems that currently face our bioregions" [7].

From within the social and ecological sciences, as well as the arts and humanities there is an ever-stronger call for transdisciplinary integration. The severe challenges of climate change, resource depletion, environmental and social disintegration, and national and international inequality, are converging into a global crisis that confronts humanity as a whole. The common purpose of health, and well-being, in full awareness of global-local interdependence, will stimulate transdisciplinary and transnational cooperation in the creation of more inclusive, multi-perspective based, decision-making processes that steer us towards the vision of a sustainable human civilization.

3 Design as transdisciplinary integrator and facilitator

Design occurs at the nexus between theory and practice. It can be used to integrate divergent value-systems and worldviews. How to meet true human needs within the ecological limits of the planetary life support system is a question of appropriate design. Broadly defined, design is the expression of intentionality through interactions and relationships. Design, as transdisciplinary integrator and facilitator will take a leading role during the 21st century to structure cooperation within the context of envisioning sustainability at a local, regional and global scale. With such multifaceted, multi-scale visions of sustainable futures, we can collectively create the strategies working toward such futures.

The Nobel laureate Herbert Simon proposed in his seminal book *The Science of the Artificial* [8], first published in 1969, that "the proper study of [hu]mankind is the science of design, not only as a professional component of a technical education but as a core discipline for every liberally educated person" (p.138). Simon saw design as special kind of science that is informed by the

natural sciences and deals with the artefacts and processes created by humans. He made the important distinction that while the natural sciences "are concerned with how things are," (p.114) and try to make nature more intelligible, most "design solutions are sequences of actions that lead to possible worlds satisfying specific constraints" (p.124). Design is based on human intentions and goals and therefore "concerned with how things ought to be" (p.114). It is this visionary and creative character of design, which gives design its central role in envisioning and creating a healthier and more sustainable future.

Scale-linking design for systemic health offers a holistically informed response to changing circumstances in the complex dynamic system that unites nature and culture, as well as mind and matter. It can help to create the political, social and economic institutions that are relevant and appropriate to changing circumstances within the interconnected and interdependent complexity of the real world. Issues like climate change, poverty, resource depletion, terrorism, inequality, and global environmental degradation can only be tackled through such a concerted response. Richard Buchanan [9] writes:

> "There is no area of contemporary life where design – the plan, project or working hypothesis which constitutes the 'intention' in intentional operations – is not a significant factor in shaping human experience. Design even extends into the core of traditional scientific activities, where it is employed to cultivate the subject matters that are the focus of scientific curiosity" (p.6).

In the material dimension the intentionality behind design is expressed through the interactions and relationships formed by products, transport systems, economies, systems of governance, settlement patterns, and resource and energy use, with the complexity of social and ecological processes. In the immaterial dimension our organizing ideas, worldviews, and value systems affect how we make sense of our experience of reality through metadesign. Metadesign, the psychological, epistemological and ontological aspect of design affecting human experience, has to be recognized as a crucial catalyst in the transition towards a sustainable human civilization. Buchanan [10] argues:

> "Design is a discipline where the conception of the subject matter, method, and purpose is an integral part of the activity and the results. On the level of professional practice, the discipline of design must incorporate competing interests and values, alternative ideas, and different bodies of knowledge".

At the nexus between theory and practice, between worldviews, value systems, and diverse stakeholder interests, and faced with the need to maintain an effective planetary life support system, design can fully step into its crucial role as interdisciplinary integrator and facilitator. Sustainable decision-making and design processes have to be open to contributions from diverse disciplines and perspectives, and at the same time, conscious of the epistemological and ontological metadesign that defines the perspective of each discipline.

There is an important visionary element to design that affects how we experience and shape our environment. "Designers deal with possible worlds and with opinions about what the parts and the whole of the human environment should be" [10]. Creating an inclusive vision of a globally sustainable human civilization, expressed through a diversity of locally adapted communities,

requires the integration of multiple worldviews and value systems. This integration can be facilitated through a dialogue-based, transdisciplinary process that applies future state visioning methodologies to the collective design of a sustainable human civilization at local, regional, and global scale. John Todd, one of the pioneers of integrative design for sustainability believes that ecologically informed design can help us to create such a civilization:

> "…through ecological design, it is theoretically possible to have a high civilization using only one tenth of the world's resources that industrial societies use today. We can reduce the negative human footprint by ninety percent and thrive as a culture. We do not have to destroy the Earth. Ecological design allows us to link human life support systems in a symbiotic way to the rest of the biosphere" [11].

4 Scale-linking design for systemic health

We are participants in a fundamentally interconnected physical, chemical, biological, ecological, social, and psychological process. The complexity of interactions and relationships between diverse agents makes this process fundamentally unpredictable and controllable. The appropriate way to come to terms with this fundamental unpredictability and uncontrollability is to remain constantly flexible, and increase resilience, adaptability and health on all scales throughout the holistic hierarchy - or holarchy [12] - of holons within holons, or networks within networks. Through *appropriate design* the negative human impact on the planetary life support system could be drastically reduced, and ecological and social integrity, resilience and health can be restored.

The constitution of the 'World Health Organization' (WHO) defines the concept of health as "a state of complete, physical, mental and social well-being and not merely the absence of disease or infirmity." In 1986, the WHO's 'Ottawa Charter' added the following "fundamental conditions and resources for health: peace, shelter, education, food, income, a [dynamically] stable ecosystem, sustainable resources, social justice and equity." In 1991, the WHO's 'Sundsvall Statement' emphasized the "way forward lies in making the environment – the physical environment, the social and economic environment, and the political environment – supportive to health rather than damaging to it" [13]. This implies a salutogenic design approach that contextualises and promotes individual, community, societal, and ecosystems health.

Design for human and planetary health aims to explore strategies that sustainably integrate humanity into the health maintaining and life-supporting processes of the biosphere. It responds to the challenges and opportunities mentioned in the WHO Commission on Health and Environment report [14]:

> "There is a powerful synergy between health, environmental protection, and sustainable resource use. Individuals and societies who share the responsibility for achieving a healthy environment and managing their resources sustainably become partners in ensuring that global cycles and systems remain unimpaired" (p.*xxx*).

Broadly conceived, if there was a more salutogenic (health-generating) intention behind all acts of design, humanity could greatly improve community, societal, ecosystems and planetary health, thereby driving the sustainability transition. The report argues: “Health depends on our ability to understand and manage the interaction between human activities and the physical and biological environment.” It concludes: “We have the knowledge for this but have failed to act on it, although we have the resources to meet current and future needs sustainably” (p.*xiv*).

Humanity’s failure – up to now - to engage in globally and locally cooperative salutogenic (health-generating) design aimed at the creation of a sustainable civilization is predominantly due to inappropriate cultural metadesign. We are culturally trapped in a mindset focussed on the individual rather than the collective, competition rather than cooperation, quantitative rather than qualitative growth, and a reductionistic rather than holistic understanding of our participatory and co-creative involvement in the complex dynamic process that unites nature and culture into a global community engaged in what Alfred North Whitehead called ‘life’s continuous exploration of novelty’.

> “Complexity theory is becoming a science that recognizes and celebrates the creativity of nature. … it opens the door to a new way of seeing the world, recognizing that these complex dynamic systems are sensitive to initial conditions and have emergent properties. We have to learn to walk carefully in relation to these complex systems on which the quality of our lives depends, from microbial ecosystems to the biosphere, because we influence them although we cannot control them. This knowledge is new to our western scientific mentality…” [15].

Brian Goodwin explains: “Emergent properties are unexpected types of order that arise from interactions between components whose separate behaviour is understood. Something new emerges from the collective – another source of unpredictability in nature.” He continues: “The complex systems on which our lives depend – ecological systems, communities, economic systems, our bodies – all have emergent properties, a primary one being health and well-being”(p.27).

Most broadly, sustainable design can be defined as appropriate (salutogenic) participation in social and ecological process. Appropriateness should be judged by the extent to which a certain design maintains the overall dynamic stability, resilience, flexibility, adaptability, or health of the system as a whole. In order to create sustainable designs we will have to learn to reintegrate social and ecological processes. This will require us to consider insights from many different disciplines through trans-disciplinary co-operation and dialogue. Designers will also have to become more conscious of the way that a particular design may participate in various, interconnected scales of natural process at one and the same time. Furthermore, the role of conceptual metadesign (epistemological and ontological assumptions) has to be considered explicitly.

Complexity theory, health, symbiosis, synergy, appropriate participation, and integrative design are related scale-linking concepts and frameworks. They can help to structure an integrated strategy to maintain human and planetary health

and achieve sustainability. Bryan Norton offers a definition of sustainability within the context of human, community, ecosystems and planetary health:

> "Sustainability is a relationship between dynamic human economic systems and larger, dynamic, but normally slower-changing ecological systems, such that human life can continue indefinitely, human individuals can flourish, and human cultures can develop – but also a relationship in which the effects of human activities remain within bounds so as not to destroy the health and integrity of self-organizing systems that provide the environmental context for these activities" [16].

From this perspective that aims to integrate social and economic realities into their wider ecological context, the notion of sustainability and the notion of maintaining and restoring a healthy and therefore resilient environment – at the community, ecosystem, and the planetary scale – are inextricably linked. Ecological and societal health, as a system-wide emergent property, facilitates healthy human development, and allows for healthy and diverse cultural expressions. Systemic health emerges as locally adapted communities learn to co-create sustainable modes of interaction and relationships within the limits and opportunities set by the ecological and social conditions of their local bioregion within a global context. In a continuously changing, complex system, the promotion of health and sustainability requires constant learning in order to adapt appropriately to such change.

Haskell *et al.* [17] emphasize that ecosystem health "cannot be defined or understood simply in biological or ethical or aesthetic or historical terms. Many approaches must be used in clarifying the goals of environmental protection." The concept of ecosystem health is best understood from a "pluralistic, multidisciplinary collection of perspectives … covering a broad spectrum of ideas from philosophy, science, and management" (p.3). The concept of "protecting and restoring health to ecological process at all levels" may help us in maintaining "the autonomous, self-integrative processes of nature as an essential element in a new ethic of sustainability" (p.4). Haskell and his co-authors understand ecosystem health as a characteristic of complex natural systems. They explain: "Since fast-changing human cultures are embedded in larger scale, slow-changing ecological systems, we must develop policies that allow human cultures to thrive without changing the life support functions, diversity, and complexity of ecological systems" (p.4).

Robert Costanza [18] reviewed a number of conceptual definitions of 'ecosystem health' based on health as: homeostasis, absence of disease, diversity or complexity, stability or resilience, vigour or scope of growth, and as balance between systems components (p.239). All of these conceptualisations of health have a valid perspective and can be informative, but they also have their limitations. Costanza calls them "pieces of the puzzle." He proposes that ecosystem health should be understood "as a comprehensive, multiscale, dynamic, hierarchical measure of system resilience, organization and vigour," and argues: "These concepts are embodied in the term 'sustainability', which implies the system's ability to maintain its structure (organization) and function (vigour) over time in the face of external stresses (resilience)." Costanza

emphasizes the important holarchical, scale-linking aspect of health: "A healthy system must also be defined in the light of both its context (the larger system of which it is part) and its components (the smaller systems that make it up)" (p.240).

David Brunckhorst [19], head of the UNESCO Institute for Bioregional Resource Management, emphasized that "resilience, like sustainability, has multi-faceted elements effecting it through scales of space and time – it does not simply occur at a local or global scale." He explains: "To sustain and restore resilience in ecological and social systems for long term sustainability, we must begin to integrate our planning and operate our management across multiple scales…". According to Brunckhorst, we may be able to do so by "nesting functional requirements of ecological systems and social systems for an enduring future" (p.16). He writes:

> "Sustainability implies not challenging ecological thresholds on temporal and spatial scales that will negatively affect the resilience or adaptive capacities of social and ecological systems. … Resilience within and across systems operates at multiple temporal and spatial scales. Loss of resilience undermines the ecosystem's capacity to continue to deliver life-support and other ecological services to humanity under a wide range of environmental conditions" (p.15).

Just as design can serve as an integrative concept for trans-disciplinary cooperation in the creation of more sustainable solutions, health can provide the integrating concept that unites social, ecological and economic needs across all scales (*and cultures!*). Such integrative concepts are crucially important in motivating individuals, societies, cultures and humanity collectively to collaborate in the creation of more holistically considered, sustainable solutions.

The emerging transdisciplinary strategy to integrate sustainable development through ecologically informed, salutogenic, and scale-linking approach to design has been heralded over the past century by the work of such pioneers like, Patrick Geddes, Lewis Mumford, Ian McHarg, John and Nancy Todd, Bill McLarney, Bill Mollison, John Tillman Lyle, Gregory Bateson, Victor Papanek, Seaton Baxter, Sim van der Ryn, Stuart Cowan, David Wann, Daniel Chiras, Robert Costanza, Janis Birkeland, Paul Hawken, Hunter and Amory Lovins, William McDonough, Michael Braungart, and David Orr [20].

Many, but not all, of the approaches promoted through the important work of these people, are explicitly stating the improvement of individual, community, ecosystem, and planetary health as their central aim. Their lowest common denominator is the intention to contribute to the creation of a more sustainable human civilization through design and metadesign that leads to appropriate participation in social, economic, and ecological process. Professor Orr writes:

> "The etymology of the word 'health' reveals its connection to other words such as healing, wholeness and holy. Ecological design is an art by which we aim to restore and maintain the wholeness of the entire fabric of life increasingly fragmented by specialization, scientific reductionism and bureaucratic division. …The standard for ecological design is neither efficiency, nor productivity, but health, beginning with that of the soil and

extending upward through plants, animals, and people. …It is impossible to impair health at any level without affecting it at other levels" [21].

This perspective recognizes the complexity of health as a scale-linking emergent property of the complex dynamic system that unites ecological and social processes into a continuously co-evolving and transforming whole. It contextualises an approach to sustainable design that is synergistic, symbiotic, scale-linking, salutogenic, and sacred.

Within the context of a fundamentally interconnected complex whole, design is a co-creative act that affects all life. As human beings we both shape, and are shaped by life's evolutionary process. In this frame, the evolution of life and consciousness is the sacred ground of our being. As such, all acts of sustainable design that preserve the health and integrity of the community of life and the planetary life support system are also sacred acts of appropriate participation in the wider process that gives us identity and meaning.

5 A holistic/integral approach to sustainability

"The new science keeps reminding us that in this participative universe, nothing lives alone. Everything comes into form because of relationship. We are constantly called into relationship – to information, people, events, ideas, and life. Even reality is created through our participation in relationships. We chose what we notice; we relate to certain things and ignore others. Through these chosen relationships we co-create our world. If we are interested in affecting change, it is crucial to remember that we are working within webs of relations, not with machines" [22].

Ultimately, the shift towards a sustainable human civilization and increased human and planetary health will require a majority of global citizens to assume full responsibility for their co-creative involvement in shaping humanity's and the planet's future. To a greater or lesser extent, we are all designers of this future. The author's doctoral research concluded that if the basic intention behind all human design was salutogenesis - the improvement of health throughout the wider system that contains us - we would be able to facilitate a drastic shift towards more sustainable practices at the local, regional, national and international scale.

Salutogenic design aims to facilitate the emergence of health at *and* across all scales of the whole. It recognizes the inextricable link between human, ecosystem, and planetary health. Rather than primarily focussing on the relief of symptoms of disease or ill-health, this approach tries to promote positive health and a flourishing of the whole by altering underlying relationships and interactions in such a way that health can emerge as a systemic property on all scales of the whole. In other words, the aim of salutogenic design is to create healthy individuals in healthy communities that act as responsible participants in healthy societies [23], ecosystems, bioregions, and ultimately a healthy biosphere *and* noosphere.

Valerie Brown and her colleagues list two criteria that should guide human behaviour if we hope to avoid serious damage to the biosphere and the natural

processes that maintain its health. The first strategy is "to consume nature's flows while conserving the stocks (that is, live off the 'interest' while conserving natural capital." The second strategy is "to increase society's stocks (human resources, civil institutions) and limit the flow of material and energy" [24].

Nature's processes are fundamentally scale linking and unite the nanometer scale of photosynthesis to the macro scale of atmospheric composition and climate change. Maintaining and improving systemic health across all scales of this dynamic whole requires an intentionally salutogenic attitude to be widespread throughout society. The global shift towards improved health and sustainability involves everyone, everywhere. To turn the vision of diverse, sustainable communities and cultures and a sustainable human civilization into reality, not only professional designers, everybody needs to act consciously and responsibly. Sustainability depends on the intention to collectively envision and create a sustainable future through the daily interactions and the relationships we form in our human and ecological communities.

This holistic and participatory perspective on sustainability proposes that the fundamental intentionality guiding us in the uncertain and uncontrollable journey towards a sustainable future should be to improve the overall health of the whole system. Through a deeper understanding of the relationship between eco-systemic, biospheric, and human health, an integrative framework for a holistic/integral approach to sustainability and complexity is emerging.

If all our actions are considered in the light of how they might affect the health of the local and the global environment as well as the health of human communities and individuals, both in the short- and in the long-term, appropriate participation, and thus sustainability, will cease to be an elusive concept and become a tangible strategy. Truly everybody is a designer, whether we are professionally engaged in the design industry, or whether we contribute to cultural, societal, and biological evolution through the way we relate to nature and culture and express these attitudes through our actions and lifestyles. Salutogenic, symbiotic, synergistic, scale linking, and sacred design can promote sustainable development, and help to structure transdisciplinary integration in a globally and locally cooperative effort to create a sustainable human civilization.

References

[1] Fry, T., *Remakings: Ecology, Design, Philosophy*, Envirobook: Sydney, p.113, 1994.

[2] Wilber, K. *A Theory of Everything: An Integral Vision for Business, Politics, Science and Spirituality*, Gill & Macmillan: London, 2001.

[3] Beck, D. & Cowan, C., *Spiral Dynamics: Mastering Values, Leadership and Change*, Blackwell Business Publishers, 1996.

[4] Bateson, G., *Steps to an Ecology of Mind: Collected essays in Anthropology, Psychiatry, Evolution, and Epistemology*, Jason Aron, 1972.

[5] Esbjörn-Hargens, S., 'Guest Editor's Introduction', *World Futures*, Vol.61, No.1-2, Special Issue on Integral Ecology, p.1, 2005.

[6] Zimmerman, M.E., 'Integral ecology: A Perspectival, developmental, and Coordinating Approach to Environmental Problems', in *World Futures: The Journal of General Evolution*, Vol.61, No.1-2, pp.50-62, 2005.
[7] Esbjörn-Hargens, S., 'Integral ecology: The *What, Who,* and *How* of Environmental Phenomena', *World Futures*, Vol.61, No.1-2, Special Issue on Integral Ecology, p.36, 2005.
[8] Simon, H., *The Science of the Artificial*, 3rd edition, The MIT Press, 1996.
[9] Buchanan, R., 'Wicked Problems in Design Thinking', in Margolin & Buchanan edits, *The Idea of Design*, The MIT Press. 1995.
[10] Buchanan, R. 'Rhetoric, Humanism and Design', in Buchanan & Margolin edits. *Discovering Design*, The University of Chicago Press, p.26, 1995.
[11] Todd, J. 'Ecological Design in the 21st Century', Annual Schumacher Lecture, Schumacher Society, UK, (see www.oceanarks.org), p.3, 2000.
[12] Koestler, A., *The Ghost in the Machine*, Arkana Books, 1967.
[13] Waltner-Toews, D. *Ecosystem Sustainability and Health: A Practical Approach*, Cambridge University Press, p.90, 2004.
[14] World Health Organization, *Our planet, our health: A Report of the WHO Commission on Health and Environment*, WHO, Geneva, 1992.
[15] Goodwin, B. *et al.*, 'Participation in a living World', in *Revision: A Journal of Consciousness and Transformation*, Vol.23, No.3, pp.27, 2001.
[16] Norton, B.G. 'A New Paradigm for Environmental Management', Costanza, Norton & Haskell edits. *Ecosystem Health*, Island Press, p.25, 1992.
[17] Haskell, B.D., *et al.* 'What is Ecosystem Health and Why should We Worry About It?' in Costanza *et al.* edits. *Ecosystem Health*, Island Press, 1992.
[18] Costanza, R., 'Towards an Operational Definition of Ecosystem Health', in Costanza *et al.* edits. *Ecosystem Health*, Island Press, pp.239-256, 1992.
[19] Brunckhorst, D.J., *Bioregional Planning: Resource Management beyond the New Millennium*, Routledge: London, 2002.
[20] Wahl, D.C., *Design for Human and Planetary Health: A Holistic/Integral Approach to Complexity and Sustainability*, Ph.D. thesis, School of Design, University of Dundee, Scotland, 2006.
[21] Orr, D.W. *The Nature of Design: Ecology, Culture, and Human Intention*, Oxford University Press, p.29, 2002.
[22] Wheatley, M.J., *Leadership and the New Sciences: Discovering Order in a Chaotic World*, Berrett-Koehler Publishers, p.145, 1999.
[23] Wilkinson, R.G., *The Impact of Inequality: How to make sick societies healthier*, Routledge: London, 2005.
[24] Brown, V.A., *et al.*, *Sustainability and Health: Supporting Global Ecological Integrity in Public Health*, Earthscan: London, p.45, 2005.

Technology and sustainability for the development of the Caldenal

H. E. Laborde, R. E. Brevedan & M. N. Fioretti
Departamento de Agronomía,
Universidad Nacional del Sur and CERZOS, CONICET, Argentina

Abstract

The Caldenal is a temperate semi-arid region located in the central part of Argentina and comprises an area of about 40,000 km^2. It is located in the ecotone between the cultivated humid Pampa to the east and the arid Monte to the west.

The region is characterized by a highly variable annual and seasonal distribution of rainfall (400-600 mm) concentrated in spring and fall. Average annual potential evapotranspiration is 800 mm. The average annual temperature is 15-16°C. Soils are mainly Calciustolls with a petrocalcic horizon at a depth varying from 0.5 to 2 m.

Cattle raising, based on cow-calf operations, is the most important economic activity in the Caldenal where stocker calves are raised to be fattened later in the more humid Pampa. The ranchers in Caldenal practice continuous grazing with very high grazing pressure, that has resulted in a severe overgrazing. Stocking rates are approximately 7 ha cow^{-1} yr^{-1}.

Currently the area supports an increasing pressure over its grass herbaceous resources. On the other hand, the humid Pampa is increasingly devoted to cash crop cultivation, mainly soybean. This causes a displacement of cattle breeding and fattening enterprises to more marginal areas. As a result, some feed-lot businesses are being located in the Caldenal region nowadays.

The Caldenal region appears suitable for a more intensive agricultural use. In an attempt to find a cropping strategy that reduces the constraints of climate and soil a non-tillage, legume ley farming system based on improved high-yielding pasture legumes is proposed in the Caldenal.

Keywords: Caldenal, rangeland, livestock production, medics, ley farming.

1 Introduction

The Caldenal region included in the rangeland area of Argentina occupies about 40,000 km^2 between approximately 32° and 40°S latitude. The region is a transition between the Monte desert to the west and the cultivated humid Pampa to the east. The Caldenal is part of the semiarid central Pampa ecoregion [1]. The humid Pampa is increasingly devoted to cash crop cultivation, mainly soybean, wheat and corn. This causes a displacement of cattle breeding and fattening enterprises to more marginal areas where non-irrigated crops are not possible. So far, the grassland ecosystem of the Caldenal is devoted to cow-calf enterprises (approximately 7 ha cow^{-1} yr^{-1}) where stocker calves are raised and then exported to be fattened in the more humid Pampa. The Caldenal along other semiarid regions of Argentina is expected to receive, in the near future, a large pressure to use its natural resources for breeding cattle.

While a system of extensive animal production is widespread in the Caldenal and has persisted for 100 years or so, its profitability has often been marginal depending on the size of the enterprise. The sustainability of this system is now being questioned, given the increased capital expenditure required. Attempts to replicate the cash crop system common in the humid Pampa have been done in the Caldenal. In fact, some feed-lot businesses are being located in the region nowadays.

The eastern part of the Caldenal is at the fringe of the arable humid and sub-humid Pampa. Plowing of this marginal land for crops (wheat or oat) is a permanent temptation for land owners. Most of the time this is translated into poor grain crop, leaving behind an impoverished ecosystem in terms of forage plant quality and soil erosion.

The climate of the region poses constraints to agriculture. Low and erratic rainfall, severe cold or high temperature, resulting in high season-to-season variability. In dry areas, measures to utilize natural precipitation more efficiently *in situ* must have high priority, since this will not only help to increase production and decrease risk, but also help to reduce erosion.

Land degradation problems have also emerged and include soil salinization and compactation. When soil is lost from fields it disrupts water courses and damage natural resources.

Overgrazing led not only to changes in the vegetation but also in the structure of the soils. There is an increase in bulk density and a reduction in depth of the top soil horizon of the grazed areas. The negative effect of grazing on the physical properties of soil can be observed down to a depth of 0.10 m. The water retention of grazed and ungrazed areas reflect the changes in pore spaces distribution. Total porosity in the top few centimeters are lower in grazed than in ungrazed areas (17%) due to the collapse of large pores. There is a marked reduction in the quantity and density of roots in the top soil horizon and there are also changes in the mean diameter of roots [2].

Water conservation and erosion control are central to dry farming objectives. Significant advances have been made in water conservation techniques through

improved methods of weed control, surface residue management and minimum tillage practices.

Wind erosion is almost always associated with cultural practices that leave the soil unprotected when high winds occur. Most control practices are based on establishing adequate surface cover and by maintaining a rough soil surface. Water erosion is much less of a problem than wind erosion.

Many semiarid zones are now being severely overused. The tendency is for individuals to over-use and under-invest in them. As a consequence many rangelands are degrading rapidly. Dryland degradation is a widespread phenomena which affects grazing lands around the world. To help prevent further over-utilization and degradation the vegetation and land must be used within its capability to support grazing and dryland agriculture.

Forage from natural rangeland provides a major component of the animal feed resources in the Caldenal. However, pressure from the increasing animal populations has led to over-stocking and over-grazing. One solution to these problems is to improve the productivity of the natural rangeland. The strategies for the development of natural rangeland like the Caldenal should include: the appropriate distribution of water points as a mean of control of grazing activities, the establishment of feed reserves to be utilized during the period where vegetation is scarce, the establishment of range reserves on order to secure an abundance of fodders during the critical periods and the appropriate fencing of too large paddocks.

The Caldenal region appears suitable for a more intensive agricultural use. For instance, alfalfa crop for hay and livestock feeding could be successfully cultivated using some kind of irrigation. Water could be derived from rivers but besides the Colorado River the only surface water is what drains from rainfall forming lagoons and salt pans in some depressions. Shallow ground-water is the main source of water in the region but its quality is poor, the conductivity ranges from 0.74 to 16.00 dS m^{-1}. The high salinity content of ground water is mainly a consequence of overgrazing and deforestation [3].

Viglizzo *et al.* [1] suggested that land use is the principal factor driving the environmental behavior of a region. In the Caldenal contamination or erosion risk indicators are particularly sensitive to land use. Also agriculture has had a negative impact on natural habitats, wildlife and biological diversity.

Minimum tillage or non-tillage practices compensate in part the more intensive use of the land. Such practices can determine an increased productivity without increasing the soil erosion risk [1].

2 Dryland farming in the Caldenal

In an attempt to find a cropping strategy that reduces the limitations of climate and soil it is proposed a non-tillage, legume ley farming system based on improved high-yielding pasture legumes in the Caldenal. The key feature of the system is the use of biologically fixed N, minimal soil disturbance with mulch cover, self-regenerating legume ley pastures, and integration of cropping with a livestock grazing enterprise. There is a rotation in which legume pastures grazed by sheep and cattle are alternated with a cropping phase.

The system is a highly flexible low input agriculture which permits a wide range of different grain and pasture/animal options according to the duration of the ley and cropping phases. There is now overwhelming evidence to show the advantages of this ley farming system over the wheat-fallow agriculture which is supplanted in terms of productivity and the preservation of soil resources. However it should be recognized that is the legume pasture, not the animal which makes the main contribution to increased wheat yield potential and the restoration of soil structure and fertility during the intercropping period, livestock being the most profitable way of converting the pasture forage to food and fiber. The use of pastures containing subterranean clover (*Trifolium subterraneum*) or annual medics (*Medicago spp*.) to improve soil nitrogen fertility for following crops was considered to be a successful model of sustainability and profitability in Australia. However, since the 1970s there has been a trend towards shorter pasture-crop rotations and continuous cropping in many parts of southern Australia. This in part reflects the relative economic returns from livestock enterprises and crop production, and the perceived shortcomings in the ley system associated with problems of pasture decline and soil acidification.

The availability of N, more than any other mineral element, is a critical determinant of both the structure and productivity of grassland in plant communities. Besides, mineral N supplies to the plant community largely determine the value of the biomass produced as forage for consumers. Most efforts aimed at improvement of rangeland under the assumption that animals fed on nitrogen-rich forage such as legumes increase their consumption of ligneous vegetation.

Legume crops increase land use, permitted alternative weed control and reduce the amount of fertilizers used. Because of their ability to fix atmospheric nitrogen, can be important in nutrient cycling and soil development processes on N-poor soils. Improved fertility provided by the legume would increase water-use efficiency.

In that situation legumes might be expected to have a significant role. In terms of their ability to exploit stored soil moisture, legumes are unlikely to fare very well in competition with grasses. Legumes seldom dominate natural ecosystems.

Pasture legumes are considered in light of their agricultural contributions to increase livestock production, to fix N from the atmosphere leading to an increase in soil N and improvement in soil structure, to increase the availability of some nutrients other than N and raise the cation exchange capacity by increasing the levels of organic matter [4]. The wider use of pasture legumes is seen as the most viable option for increasing crop yields in dryland areas, such as the Caldenal.

Management studies must define the critical range of grazing pressure beyond which persistence is threatened. Persistence is sensitive to both under-grazing and over-grazing. On the other hand management of diversity is more complex and more species can increase production costs on the farm.

What is needed is a good understanding of the ecophysiological requirements of targeted habitats and a knowledge of what species likely to fit these habitats will be found.

Alfalfa fixes more N than medics even under drought conditions. However the shallow soils and the shortage of rain in the Caldenal are not suitable for alfalfa except in lowlands close to rivers or streams. Medics have strategies to avoid drought and produce forage and plenty of seeds in those environments.

Medic pastures with proper grazing management regenerate from their own seed in subsequent seasons, even after a cereal crop has been grown on the same land.

Medicago minima (L) Bartal. var. minima is an annual legume species that was introduced from Europe and became naturalized in the Caldenal. During spring, this species can be an important forage resource for cattle in the extensive over-grazed areas of this region. Under these conditions, *M. minima* can constitute more than 60% of the total herbaceous production. This is in contrast to most rangelands which suffer a scarcity of natural leguminous species [5].

That annual legume is common in several other arid and semiarid rangelands and can become a dominant species under conditions of aridity and over-grazing [6].

Initiation of its cycle occurred in autumn and appeared to be associated with soil water availability. The end of the growing season happened during late spring, concomitant with high maximum air temperatures.

The species hasten its development when air temperature increases and soil water availability decreases. This could be an important strategy in the species which allows them to persist as seeds, and produce a new generation under favorable environmental conditions [7]. *M. minima* shows a high phenotypic plasticity under different levels of soil water availability which may enhance its capacity to survive and reproduce. It can survive under severe drought conditions and exploit the environmental resources when they are available. Their persistence in the region appears to be ensured by their capacity to colonize open, degraded areas as well as to grow in association with perennial grasses [8]. When exposed to water stress allocated a higher proportion of the total plant dry weight to fruits than those grown under irrigated conditions during early spring. Under drought conditions this is a beneficial feature because it enables plants to complete their life cycle and ensure seed production [8].

Plants survived even those exposed repeatedly to very severe water stress (-6.0 MPa or less) [9].

Medics showed a fairly good dry matter production (Table 1) mostly in winter time, in spite that Caldenal has not a strictly Mediterranean climate [10]. *In vitro* dry matter digestibility is over 70% during winter (Table 1) [11].

Associations of oat and legumes are frequently utilized in the semiarid regions of Argentina to extend forage production from autumn to late spring season. Dry matter yield of those associations are shown in Table 2.

Weeping lovegrass and other warm season forage grasses were introduced in the Caldenal to improve forage production. There was an increase in dry matter yield of those grasses when sown with alfalfa compared with the grass alone, and

also an increase in the wheat grain yield sown after the pasture (Table 3). These results show the good potential to get higher forage and subsequent grain yields with grass-legume mixtures compared with grass pastures alone.

Table 1: Dry matter (DM) and digestible dry matter (DDM) yield (in Mg ha^{-1}) of annual medics (*Medicago spp*) and vetch (*Vicia sativa*). Two-year average, SW Buenos Aires province.

	DM				DDM
Species	Autumn	Winter	Spring	Total	Total
M. truncatula cv Jemalong		4.823 a	1.555 a	6.378 a	4.520 a
M. truncatula cv. Cyprus	0.658	4.958 a	0.625 c	5.583 ab	4.120 a
M. rugosa cv. Paraponto		4.819 a		5.477 ab	4.289 a
M. littoralis cv. Harbinger		4.507 ab	0.315 c	4.822 bc	3.670 ab
M. rugosa cv. Paragosa		3.727 b	0.997 b	4.724 bc	3.582 ab
M. tornata cv.Tornafield	0.801	2.909 c	1.146 b	4.055 c	3.070 b
Vicia sativa cv. Common		2.587 c	0.495 c	3.883 c	3.041 b

(Adapted from [11]).

Table 2: Dry matter (DM) yield of annual medics and vetch as pure stands or associated with oat. Two-year average, SW Buenos Aires province.

Species	Cultivation	Total DM yield (Mg ha^{-1})
M. truncatula cv Jemalong	Associated with oat	5.301 a
	Pure stand	4.077 ab
M. truncatula cv Paraggio	Associated with oat	5.353 a
	Pure stand	3.770 b
Vicia sativa cv Conmon	Associated with oat	4.908 a
	Pure stand	1.818 c

Means of associated crops yields between species (P > 0,05).
Means of pure stand yields between species (P < 0,05).
(Adapted from [12]).

Table 3: Dry matter (DM) yield (Mg ha^{-1}) of *Digitaria eriantha, Eragrostis curvula* and alfalfa as pure stands or associated and subsequent grain yield (Mg ha^{-1}) of wheat. Three-year average, SW Buenos Aires province.

Treatments	DM	Wheat yield
1. Pure alfalfa	4.880 ab	3.344 a
2. Pure digitaria	0.787 c	0.713 c
3. Pure weeping lovegrass	3.047 b	0.639 c
4. 1 & 2 seeded in different rows	4.406 b	2.127 ab
5. 1 & 3 seeded in different rows	6.601 a	1.659 b
6. 1 & 2 seeded in the same row	5.042 ab	2.181 ab
7. 1 & 3 seeded in the same row	7.180 a	1.425 b

(Laborde *et al*., unpublished data).

The association of *Digitaria eriantha* with alfalfa increased the forage (217%) and the subsequent wheat grain yield (260%). In case of weeping lovegrass the increases were of 236 and 223%, respectively.

In recent years, there has been a growing awareness world-wide of the general multipurpose value of forage trees and their specific role in cropping-livestock systems. However, farming of fodder shrubs is virtually unknown in the Caldenal. Species of *Prosopis* (*P. flexuosa*) are browsed regularly and play a significant role as a reserve and supplementary fodder [13].

3 Conclusions

Some areas crucial to dryland agriculture need to be understood before a great increase in agricultural productivity in the Caldenal could be achieved. For example, the livestock-cropping interface must be clarified. Farmers are not simply cash-crop producers or livestock producers-they are both. Decisions relating to one activity also have implications for the other. Secondly, risky production and risk attitudes are critical to dryland agriculture such as Caldenal's, particularly when producers are presented with new and uncertain technologies. Income stability may be as important as income levels and consequently, this affects production decisions regarding inputs, new techniques and even choices of farming activities. Finally, learning is a process that improves farmers' perceptions of available choices y consequently can reduce errors which farmers make when experimenting with these new technologies. If the farmer understands a new alternative better, he will be more likely to make a considered decision about it, with a greater chance of adopting it [14].

The development of more sustainable systems in the Caldenal can only be planned with the full involvement of communities who take them seriously, participate in the testing of options and take responsibility for the outcomes.

References

[1] Viglizzo, E.F., A.J. Pordomingo, M.G. Castro and F.A. Lértora. 2002. La sustentabilidad ambiental del agro pampeano. INTA. 84 pp.

[2] Villamil, M.B., N.M. Amiotti and N. Peinemann. 2001. Soil degradation related to overgrazing in the semi-arid Southern Caldenal area of Argentina. Soil Sci., 166: 441-452.

[3] Paoloni, J.D., M.E. Sequeira, C.E. Fiorentino, N.M. Amiotti and R.J. Vázquez. 2003. Water resources in the semi-arid Pampa-Patagonia transitional region of Argentina. J. Arid Environments, 53: 257-270.

[4] Fitzpatrick, E.N. and I.C. Rowland. 1980. The role of legumes in dryland agriculture. Int. Congress Dryland Farming, Adelaide, Australia: 10-11.

[5] Fresnillo Fedorenko, D.E., O.A. Fernández and C.A. Busso. 1995. The effect of water stress on top and root growth in *Medicago minima*. J. Arid Environments, 29: 47-54.

[6] Volaire, F., M. Gordon and F. Lelievre. 1990. The natural herbaceous grasslands of Corsica. I. Construction of a typology by the characterization of both the environmental factors and the flora of the types. Agronomie, 10: 163-174.

[7] Fresnillo Fedorenko, D.E., O.A. Fernández, C.A. Busso and O.E. Elía. 1996. Phenology of *Medicago minima* and *Erodium cicutarium* in semiarid Argentina. J. Arid Environments, 33: 409-416.

[8] Busso, C.A., O.A. Fernández and D.E. Fresnillo Fedorenko. 1998. Dry weight production and partitioning in *Medicago minima* and *Erodium cicutarium* under water stress. Ann. Bot., 82: 217-227.

[9] Fresnillo Fedorenko, D.E. 1990. Estrategias ecológicas de *Medicago minima* (L.) Grufb. var. minima y *Erodium cicutarium* (L.) L'Herit., dos anuales de valor forrajero en el Caldenal. Ms. Tesis, Universidad Nacional del Sur, Argentina. 132 pp.

[10] Laborde, H.E., R.E. Brevedan and M.N. Fioretti. 1997. Sustainable development of the Caldenal region of central Argentina. En: J.L. Uso, C.A. Brebbia and H. Power (eds.). Ecosystems and Sustainable Development: 363-372.

[11] Gargano, A.O., F.E. Mockel, M.A. Adúriz and M.A. Cantamutto. 1986. Evaluación de leguminosas anuales del género *Medicago* comparadas con *Vicia sativa*. Rev. Arg. Prod. Anim., 6: 573-580.

[12] Gargano, A.O, M.A. Adúriz and M.C. Saldungaray. 1990. Rendimiento de forraje y semilla de *Medicago truncatula* cv Jemalong y cv Paraggio, y *Vicia sativa*. Rev. Arg. Prod. Anim., 10: 19-26.

[13] Campos, C.M. and R.A. Ojeda. 1997. Dispersal and germination of *Prosopis flexuosa* (Fabaceae) seeds by desert mammals in Argentina. J. Arid Environments, 35: 707-714.

[14] Nygaard, D.F. 1980. Socio-economic constraints to the future development of the dryland farming regions of the world. Int. Congress Dryland Farming, Adelaide, Australia: 39-40.

Coral reefs: threats and future focusing in over-fishing, aquaculture, and educational programs

L. Molina Domínguez[1], F. Otero Ferrer[2] & M. Izquierdo López[1]
[1]*Grupo de Investigación en Acuicultura,*
Instituto Canario de Ciencias Marinas,
Universidad de Las Palmas de Gran Canaria, Spain
[2]*Musée Océanographique de Monaco, Monaco*

Abstract

Coral reefs cover less than 0.25% of the marine environment; however their importance is not only justified by their beauty but also because coral reefs provide food and livelihood to millions of people located in communities around the world especially in developing countries. They are considered the "rainforest of the sea" (Spalding et al., 2001) because of their high productivity and specially their biodiversity. In fact, coral reefs support approximately 4000 species of fish, about 800 of reef-building corals and more than a thousand other invertebrates and sponges. Besides, a great number of species use these ecosystems as reproduction or husbandry areas. 20% of the world's coral reefs have been effectively destroyed and show no immediate prospects of recovery (Wilkinson 2004). The global threats to coral reefs began to raise alarm after the Rio World Environmental Summit in 1992 and the 7th Symposium of Coral Reefs, starting the first calls in order to stop the decline and to provide effective protection for these ecosystems. The threats affecting coral reefs are great and the degree of damage will depend on the different areas of coral reefs locations. The threats can be classified into three groups (note that two are caused directly or indirectly by human activity):

- Natural disasters, such as tropical storms, inundations, earthquakes, volcanoes, etc. To a great extent coral reefs are generally considered to have a strong potential to recover.

- Global climate change, such as increments of sea surfaces temperature, rising CO_2 levels and other pollutants.

- Direct human pressure, such as over-fishing, not only for food consumption but also for aquariology purposes, development of coastal zones (urban, industrial, tourism, and transport coral reefs uses).

Against these threats some possible solutions can and must be developed. This paper focuses on over-fishing and the relationship between aquariology and aquaculture.

Keywords: coral reef, aquariology, aquaculture.

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060301

1 Introduction

Although reefs cover less than one quarter of 1% of the marine environment, they are considered "the rainforest of the seas" [2] because of their biodiversity, and they are considered to be the most biologically and productive rich ecosystems throughout the world. Coral reefs support over 4000 species of fish (which represent one third of the total marine fish species), about 800 species of coral [3] and a great number of other invertebrates and sponges. Coral reefs host an extraordinary variety of marine plants and animals and it has been estimated so far that only about 10% of these species have been described by scientists.

Reefs support 39% of the world's population living within 100 km of coast. They provide important resources for hundreds of millions of people through both local sustenance and commerce: food, tourism revenue, coastal protection and new medications. For example, AZT (a drug for people with HIV infections) is based on chemicals extracted from a Caribbean reef sponge and more than half of all new cancer drugs are based on marine organisms.

Within these activities, collecting marine ornamental organisms provides one of the few potentially sustainable local industries in many coastal communities with limited resources, and few other options for generating income. At stake is the employment of thousands of people, especially in source nations, and the high incentives for coral reef stewardship which the marine aquarium trade is capable of providing.

Worldwide, coral reefs face anthropogenic damage and many are affected by the same economic activities they support [4]. They face a diversity of risks, such as pollution, sedimentation, bleaching, global warming, tourism, and over-fishing not only for ornamental purposes, and also tsunamis in the last years. Each of these pressures is bad enough in itself, but together the cocktail has become lethal. Besides, human-perturbed reefs show a prolonged failure to return to their former state compared with perturbed reefs by natural disturbances [5].

Between these threats ornamental fisheries should not be forgotten. The marine ornamental trade began as long as ago the 1930s, and it has now expanded into a multi-million dollar industry, between 28 and 44 millions dollar annually and a global annual catch of about 14 million to over 30 million fish [6], although these figures should be considered with caution due to the lack of contrasted data. .

It must be underlined that nearly all tropical marine fish and invertebrates are taken from the wild coral reefs and adjacent habitats. On the other hand, severe decrease or local extinction of reef organisms might cause, directly or indirectly, deterioration in coral abundance [5]. Reefs of Southeast Asia are the most important source of the aquarium trade, and because of that they are particularly at risk [7] taking into account that damaging techniques of fishing are sometimes used and target species over-exploited in these areas.

Unlike freshwater aquaria species, where 90% of species are currently farmed, the great majority of marine aquaria are stocked form wild caught species [8]. Aquaculture has experienced a great development of culture

technologies in the last decades, related to the increase of the world production. All these technologic improvements could be applied to the reproduction and larval culture of species with ornamental interest, with a market already established. Overexploitation of some species fisheries could be diminished, and their is the possibility of repopulation of natural ecosystems in some highly affected cases [9].

2 An overview of present status of coral reef

The *World Atlas of Coral Reefs* prepared by the United Nations Environment Programme World Conservation Monitoring Centre (UNEP-WCMC) provides a new global estimate for coral reefs world-wide: 284 300 sq km, an area just half the size of France. For the first time, it also provides reef area estimates for individual countries (Fig.1.) and includes detailed maps and statistics for all the world's coral reef nations.

Indonesia, followed by Australia and Philippines are the largest reef nations, while France comes in fourth (14,280 km ^{2}), United Kingdom is 12th (5500 km ^{2}), and United States 16th (3770 km ^{2}), most located in overseas territories.

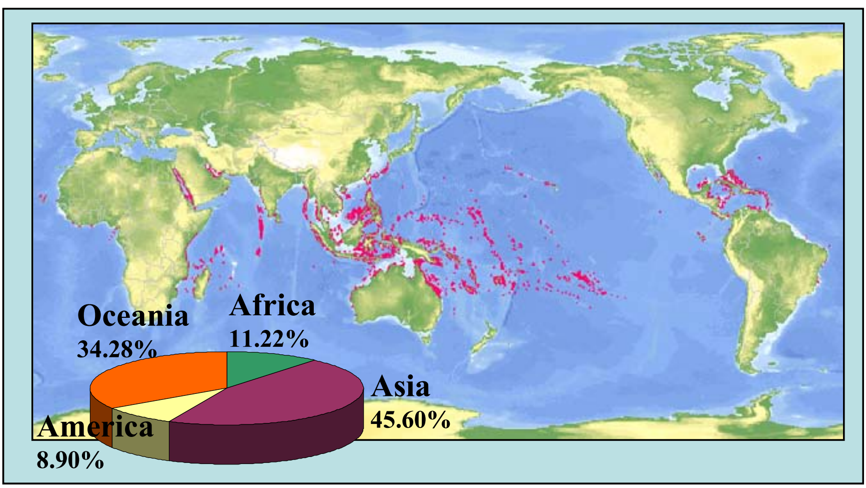


Figure 1: Coral reefs of the world (adapted from www.reefbase.org).

The threats to coral reef and their associated ecosystems vary widely between countries. Most of them are related to the lack of sustainable income-generating opportunities and it must be addressed to achieve sustainable conservation of coral reefs.

Nearly 60% of the earth's coral reefs are threatened by human activity, they are among the least monitored and protected natural habitats in the world (Fig. 2).

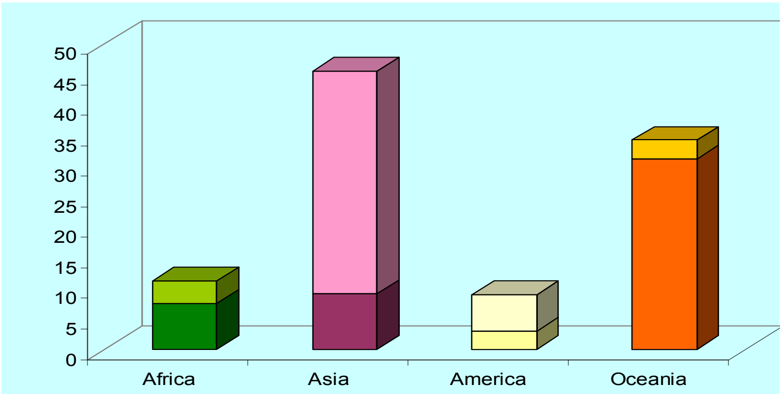


Figure 2: Percent of affected coral reefs within the different continents referred to the global contribution of each one. Solid colours: unaffected area; soft colours: affected area.

- Coral reefs of Southeast Asia, the most species-rich on earth, are the most threatened of any region. More than 80% are at risk, primarily from coastal development and fishing-related pressures.
- Nearly two-thirds of Caribbean reefs are in jeopardy. Most of the reefs on the Antilles chain, including the islands of Jamaica, Barbados, Dominica and other vacation favourites, are at high risk. Reefs off Jamaica, for example, have been ravaged as a result of overfishing and pollution.
- Approximately 30% of East African reefs are at risk. Although levels of estimated reef destruction range widely within the region. Coastal population and coral bleaching are the two primary threats.
- Most United States reefs are threatened. Almost all the reefs off the Florida coast are at risk from a range of factors, including run off of fertilizers and pollutants from farms and coastal development. Close to half of Hawaii's reefs are threatened, while virtually all of Puerto Rico's reefs are at risk.
- Most pacific coral reefs are healthy, less than 10% are described as threatened, because of effective management or because they are remote from the continent and have low numbers of inhabitants. Generally, there was more damage to the coral reefs through natural events such as cyclones and coral bleaching, than by direct human activities, although climate changes could have an effect.

3 Conservation measures

A number of measures can be taken to reduce the threats affecting coral reefs. Focusing in ornamental industry, the lack of contrasted data is one of the main

problems in order to conserve the population of ornamental organisms living in coral reefs. An effective way of monitoring the fisheries is as simple as to collect data on catch by collectors, including fish that die during the catch. These are in use in Australia, Palau, Cook Islands, Maldives Islands, and Sri Lanka.

Marine ornamental fisheries need to be managed in such a way that they are biologically sustainable and do not conflict with other resource uses. The aquarium trade if managed properly could support long-term conservation and sustainable use of coral reefs. In order to achieve this objective some arguments must be emphasized.

3.1 Developing sustainable fishing and husbandry practices

Destructive fishing techniques include the use of sodium cyanide and other chemicals, to stun and catch fish. Estimation of its use is difficult to obtain due to the clandestine nature of these practices. Other collection techniques such as dynamite and explosives are also commonly used for food fishing. During the coral collection for trade, many more colonies may be damaged or broken. All these methods cause terrible damage to organism populations but also to the reef habitat itself. Obviously, these practices must be overcome; some other collection methods have minimal impacts on coral reef. Cyanide fishing is illegal in most countries, but more efforts in producer countries are needed: changing their policies, promoting awareness campaigns, regulating the importation, distribution and use of the cyanide. It is necessary to limit the collecting effort by controlling the number of collectors or establishing quotas on numbers of fish captured or exported [3] based on scientific research carried out in order to decide the appropriate numbers. Consumers also have an important role to play: if the demand of sustainable fishing organisms increases it is likely that this will affect producer countries. One of the best examples is Marine Aquarium Council (MAC) Certification launched in late 2001 that provides aquarists with the only comprehensive “reef to retail” system, a tool to identify marine aquarium organisms that originate from managed reefs, are caught using non-destructive methods, and are handled and transported according to best practices. An increasing number of collectors, wholesalers and retailers in Asia, the Pacific, North America and Europe are now MAC certificated.

3.2 Promoting the implantation of locally based projects of ornamentals cultivation

This action is very important in order to sustain the trade and to present an alternative livelihood to local people. Mariculture facilities must be designed in order to create job opportunities for coastal populations, especially fishermen. Some initiatives are already working (i.e. Lombok Frags, Indonesia) and it could be extended. In this way a new basis of subsistence may reduce the anthropogenic pressure and increase awareness protection in their reef areas. However, the intensive use of net penned fish could affect corals, although there are some discrepancies about this subject [5, 10–12].

3.3 Promoting the research and the culture of ornamental species

The limited number of fish species bred in captivity can be pointed out. Due to the increasing regulations or restrictions about the trade of some species, without aquaculture to provide specimens for exhibits, many species may not be accessible for display and for exchange with other institutions [13]. Moreover, the knowledge and technology of establishing and maintaining the live coral reefs exhibits is important for the conservation and the restoration of the wild coral reefs. Besides, as controllable facilities, they are powerful tools for the research of global environment change on coral, both at the community and the organism levels [14]. Public aquariums have a great opportunity for future conservation and that also means a great responsibility.

3.4 Coral reef conservation and restoration

Coral reef conservation focuses on “passive” measures in order to allow the ecosystem to recover with minor human actions. Coral reef restoration is based on artificial manipulations in order to improve the coral state. Various methodologies for coral restoration have been proposed [15], such as coral culture, chemical products, and artificial reefs. Farmed coral could be used to restore damaged reef areas and aquarium trade or the pharmaceutical industry. Not only asexual but also sexual reproduction must be used in order to preserve genetic diversity [16]. An artificial reef is a method not used specifically for this purpose, but with interesting potential at least in boosting the availability of species. It must be pointed out that it must be a temporary solution with the possibility of the generation of new problems [17]. Because of that, a great effort to improve this tool must be made, for example by specific design. Artificial reefs must be carefully sited and sized in order to provide abundance and diversity of specimens. More research will be necessary to promote the improvement of these techniques.

3.5 Domestic or public aquaria to educate conservation

Traditionally public aquarium’s tasks are to communicate, instruct and educate people on aquatic issues, but today’s aquaria could have an important role in applied or basic scientific research associated with other institutions, providing scientific publications, connecting different stakeholders (experts, politicians, industry, non-governmental associations, and public in general), and creating exhibits in order to show environmental concerns. New modern aquarium must play an important role in conservation and education programs and those also related to the research projects. As our planet becomes more threatened by social impact, the main objectives of an educational promotion at a public aquarium must be to engage visitors to pollute less, support wise political practices in conservation (catch limits, certification, better identification of fishing areas etc), and engage the next generation in attitudinal readiness. By making these changes, ecosystems that have been threatened because of bad practices can be allowed to recover [18]. Moreover, live coral reef exhibits are one of the most

powerful tools to hold people's attention and provide an understanding of the beauty, uniqueness and importance of coral reefs. The popularity of these exhibits could be used to raise public awareness.

4 Conclusions

The most important actions for promoting healthy coral reef ecosystems depend largely on the efforts by local governments, community groups, environmental organizations and the private sector. More accurate trade data and more information about target species are needed to establish regulations, marine reserves, quotas and maximum yield productions, in order to promote the ornamental trade as a sustainable basis. Finally, further research into ornamental species technologies in order to diminish the pressure off wild stocks and increase the effectiveness of aquaculture facilities. Some of these projects should be locally based in source countries as an alternative livelihood for fishing communities, integrating conservation and progress.

References

[1] Status of coral reefs of the world. Australian Government. Australian Institute of Marine Science. Online. http://www.aims.gov.au/pages/research/coral-bleaching/scr2004/.

[2] Spalding, M., Ravilious, C. and Green, E. *World Atlas of coral reefs.* University of California Press, Berkeley, USA.pp.11-13, 2001.

[3] Paulay, G. Diversity and distribution of reef organisms. *Life and Death of coral reefs*, ed. C. Birkeland, Chapman and Hall, New York, USA. pp. 298-353. 1997.

[4] Wielgus, J., Chadwick-Furman, N., Dubinsky, Z. Shechter, M. & Zeitouni, N. *Coral reefs* **21**, pp.253-259.2002.

[5] Loya, Y. The coral reefs of Eilat -Past, Present and Future: Three decades of coral community structure studies. Coral health and disease. Springer, pp. 1-34. 2004.

[6] Wood, E. Global advance in conservation and management of marine ornamental resources. *Aquarium Sciences and Conservation,* **3 (1-3)**, pp. 65-7. 2001.

[7] Burke L., Selig, E. & Spalding, M. *Reefs at risk in Southeast Asia.* World Resources Institute, Washington DC, USA, 2002.

[8] Andrews, C. The ornamental fish trade and fish conservation. *Journal of fish biology,* **37** *(Supplement A)*, pp. 53-59, 1990.

[9] Molina Domínguez, L. & Ounaïs, N. The sustainability of Aquarium trade. *Proc. of XI World Aquaculture International Congress*, Nusa Dua, Bali (Indonesia) p. 53. 2005.

[10] Bongiorni, L., Shafir, S. & Rinkevich, D. Effects of particulate matter released by a fish farm (Eilat, Red Sea) on survival and growth of *Stylofora pistillata* coral nubbins. *Marine Pollution Bulletin*, **46**, pp. 1120-1124. 2003.

[11] Rinkevich, B. Nutrient enrichment and coral reproduction: between truth and repose (a critique of Loya et al.). *Marine Pollution Bulletin*, **50**, pp.111-113. 2005.
[12] Rinkevich, B. What do we know about Eilat (Red Sea) reef degradation? A critical examination of the published literature. *Journal of Experimental Biology and Ecology*, **327**, pp. 183-200. 2005.
[13] Page, G. Gomezjurado, J. & Ashley, A. Ecological Aquaculture a new paradigm for conservation. *Proc. of the VI International Aquarium Congress*, Monterey, California, USA, p. 53. 2004.
[14] Fan, T., Kuo, F. & Fang, L. The coral reef Mesocosms in the national Museum of Marine Biology and Aquarium, Taiwan. *Proc. of the VI International Aquarium Congress*, Monterey, California, USA, p. 31.2004.
[15] Rinkevich, B. Conservation of coral reefs through active restoration measures: recent approaches and last decade progress. *Environmental Science & Technology*, **39**, pp. 4333-4342. 2005.
[16] Petersen, D., Laterveer, M. & Schuhmacher, H. Methods to enhance sexual recruitment for restoration of damaged reefs. *Marine Biology*, **146**, pp. 937-942. 2005.
[17] McAllister, D.E. Is mariculture the remedy to problems of coral reefs of coastal communities? *SPC Live Reef Information Bulletin* 5 pp. 47-48, 1999.
[18] Sonnenschein, L. Aquariums reaching the public with Conservation Campaigns. *Proc. of the VI International Aquarium Congress*, Monterey, California, USA, p. 53. 2004.

Maximizing storage rates and capacity of carbon dioxide sequestration in saline reservoirs

A. Abou-Sayed[1], Q. Guo[2], A. L. Graham[3], L. A. Mondy[4], M. S. Ingber[5] & A. A. Mammoli[5]
[1]*Advantek International, USA*
[2]*M-I Swaco, USA*
[3]*Los Alamos National Laboratory, USA*
[4]*Sandia National Laboratories, USA*
[5]*Department of Mechanical Engineering, University of New Mexico, USA*

Abstract

The Kyoto Accords have been signed by 140 nations in order to significantly reduce carbon dioxide emissions into the atmosphere in the medium to long term. In order to achieve this goal without drastic reductions in fossil fuel usage, carbon dioxide must be removed from the atmosphere and stored in acceptable reservoirs. Research has been undertaken to develop economical new technologies for the transfer and storage of carbon dioxide in saline aquifers. In order to maximize the storage rate, the aquifer is first hydraulically fractured in a conventional well stimulation treatment with a slurry containing solid proppant. Well fracturing would increase the injection volume flowrate greatly. In addition, there are several ancillary benefits including extension of the reservoir early storage volume by moving the carbon dioxide further from the well. This extended reach would mitigate the problems with the buoyant plume and increase the surface area between the carbon dioxide and the formation facilitating absorption. A life-cycle cost estimate has been performed showing the benefits of this approach compared to injection without fracturing.
Keywords: carbon dioxide sequestration, reduction of green house gases, hydraulic fracturing.

1 Introduction

Substantial reduction of carbon dioxide (CO_2) emissions to the atmosphere will require economical and long-term stable sequestration in geological formations. Of the three possible types of formations, oil fields, coal beds and deep saline aquifers, the latter offer the most attractive long-term solution for several reasons.

- The capacity of saline aquifers is sufficient to accomodate disposal of CO_2 for the foreseeable future [3, 6].
- Aquifers are close to a large percentage of the CO_2 production sources in North America [1].
- Brackish waters in these formations are an ideal storage medium, since they are not suitable for potable water usage or even irrigation [2].
- Sequestration in these formations does not present health risks [4].

Several studies have been performed to establish the feasibility of CO_2 sequestration in geological formations [7, 1, 5]. These studies have identified several issues as impediments to the economical viability of sequestering CO_2 in deep saline aquifers and other geological formations. These issues include the injection rate and pressure required to achieve the required throughput and the long-term containment of CO_2. In particular, in these previous studies, the injection pressure was limited by the fracture pressure of the formation, thus severely limiting the injection rate. In the current research, both the injection rate issue and the containment issue are addressed in a unified manner by intentionally fracturing the formation. Not only can the injection rate be substantially increased but the CO_2 is spread further from the well-bore early in the injection process and the surface area between the injected supercritical gas and the formation is also substantially increased. These last two factors significantly reduce the problem of any buoyant plume of CO_2 forming near the well-bore, and hence, seeping out of the formation. A life-cycle cost estimate of sequestration with and without hydraulic fracturing is performed to show the advantages of the current approach.

2 Estimate of the relative costs for injection of CO_2 into saline aquifers

The current analysis of the life-cycle cost estimate to sequester CO_2 from a coal-burning power plant is based on a 800 MW plant for a period of 20 years. The separation costs of the CO_2 from the waste stream, the transportation costs, and the compression costs are not included. It is tacitly assumed that the saline aquifer will be in the proximity of the power plant so that transporting the CO_2 to the injection site will be minimal. An analysis of acceptable aquifer locations was performed and several candidate formations were identified including several in the Central Appalachian Basin of western Pennsylvania, eastern Ohio, and eastern Kentucky and the San Juan basin in New Mexico. It is estimated that a CO_2 injection rate of 12,600 m^3/day will be required by the 800 MW power plant. An analysis is performed to determine the number of required injection wells

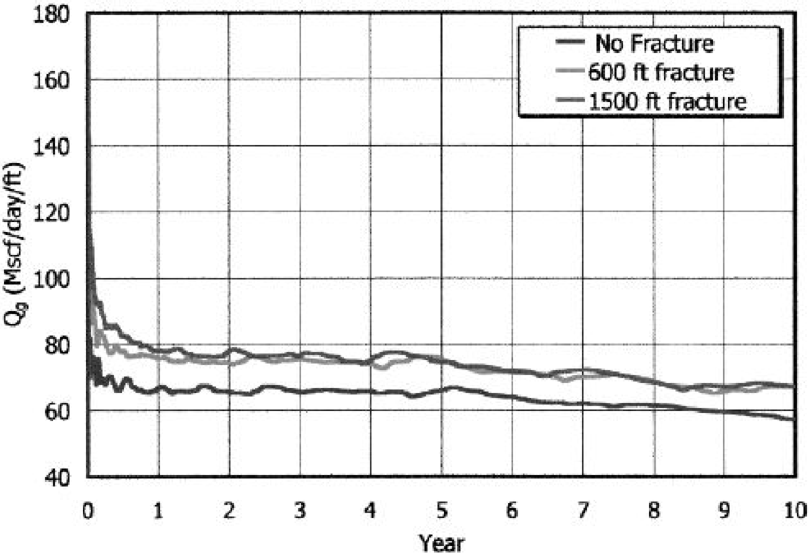


Figure 1: Injection flow rate as a function of time for fractured and non-fractured injectors.

depending on process including injection from non-fractured vertical wells, fractured vertical wells, non-fractured horizontal wells, and fractured horizontal wells. As an example, an analysis has been performed to measure the injection rate as a function of time for both a hydraulically fractured and non-fractured injector. Fig. 1 shows the injection rate as a function of time at a given injection pressure (equivalent to the constant compression pressure at the surface). Not only does the injection rate increase with fracturing, but also, the life of the well is also increased from approximately 5 years to approximately 20 years.

In the current analysis, the existence of an impermeable stress barrier is assumed, and hence, the CO_2 will only flow through the sequestration zone and the bottom-hole injection pressure is limited at a value below the minimum horizontal in situ stress of the barrier. The required number of wells was determined from either radial flow equations or the flow equations of fractured injectors. The drilling, completion and fracturing capital expenditures (CAPEX) are considered as a lump sum that is accounted for in the year the operation takes place only. A more thorough analysis would take into account both the tangible and intangible value of the asset and spread it over a given number of years. Both the maintenance and man power costs, operating expenditures (OPEX) are considered to be fixed amounts that are deducted on a yearly basis. At the end of a well life, an abandonment cost capital expenditure is considered. Another operating expenditure associated with the process is power, which makes up the largest part of the operating costs. In cases where a hydraulic fracturing stimulation is

Table 1: Input parameters considered for the cost model. All dollar values are in millions of dollars (mm$).

Operations	Value	Unit
Drilling	2.00	mm$
Hydraulic Fracturing	0.30	mm$
Abandonment	0.40	mm$
Compressor	3.00	mm$
Retrofit	0.30	mm$
Pipeline	1.00	mm$/mile
Distance	2.00	mile
Power	10.00	cents/kWhr
Maintenance	0.05	mm$/year/well
Labor	0.12	mm$/person/well

required an extra cost is added to the drilling cost for the fracture job, which again comes under capital expenditure. Finally all the costs are transferred their respective years and a discount rate is considered to establish the Net Present Value of the option. Table 1 shows the input parameters considered in the model for cost estimation.

In the first scenario, the viability of using a vertical well operating under matrix conditions is considered. The initial skin in this case is present due to formation damage from the drilling process and the type of completion used, and has an adverse effect on the injection process. It was found that a minimum of four wells were required to completely dispose of the CO_2 generated from the 800 MW power plant. Furthermore, it was concluded that vertical unfractured wells would only operate under these conditions for five years at which point another four wells would be drilled, at another remote location, requiring a new compressor and pipeline connections. With eight wells now operational, both the maintenance and manpower costs are increased. One benefit of that scenario is that the injection rate can be decreased significantly among the eight wells, and hence, the wells would last for another seven years at which point another four wells would need to be added, and the process is repeated. No further wells would need to be added for the twenty year period of analysis. In the second scenario, the case of a vertical well that has initially been stimulated using hydraulic fracturing that creates a limited length contained fracture is considered. The effect on the injection process is notable and the previously adverse skin effect is overcome by the fracturing. The well's daily capacity is significantly increased for the same injection pressure. In this case, only two wells are needed and the wells will operate for the

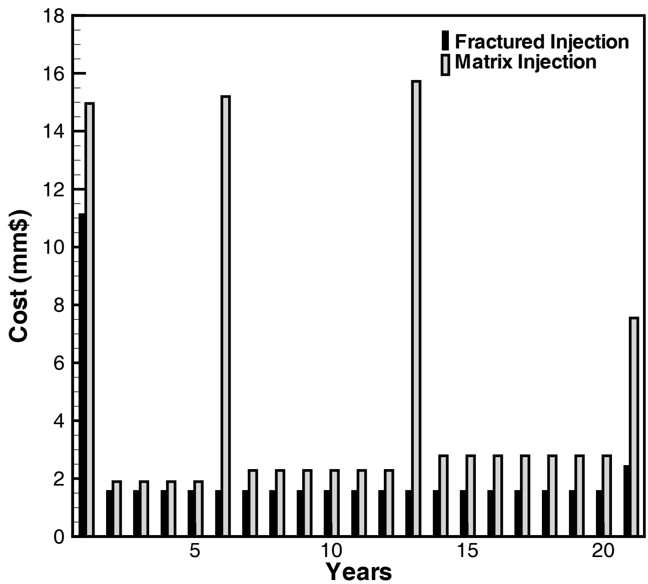


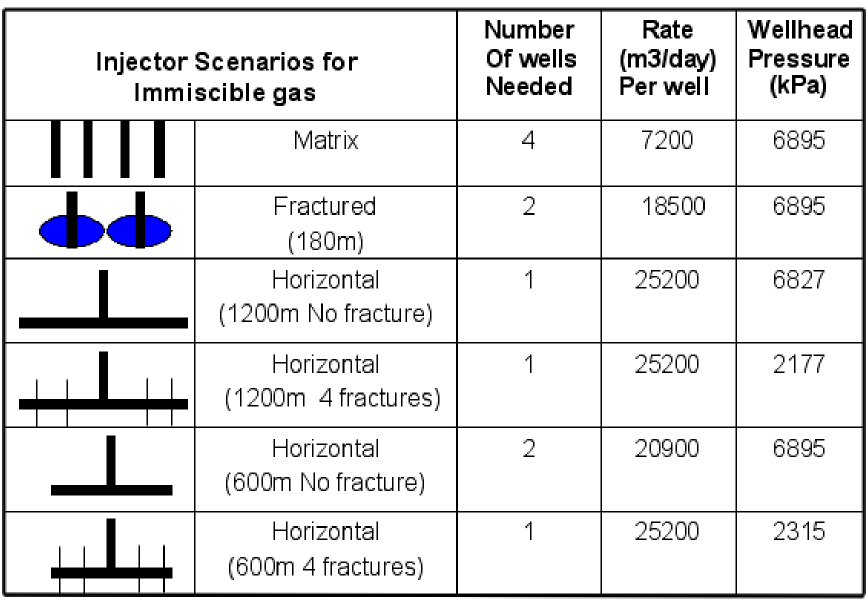
Figure 2: Yearly expenditures in millions of dollars for fractured injection and matrix injection.

duration of the twenty years. The total yearly expenditure for both the matrix injection and fractured injection are shown in Fig. 2. For the matrix injection scheme expenditure, there are several spikes denoting the repeated drilling phases discussed earlier and the final abandonment cost is much greater than that for the fractured injector scheme since the number of wells to be disposed of is considerably larger. Operating costs are increased for the matrix injection after each set of new wells is drilled caused by the increase in both man power and maintenance costs for the added wells. A break down of the costs per year of the two options (matrix injection and fractured injection) is shown in Table 2. The cost per year of matrix injection is over twice the cost per year of fractured injection. The Net Present Value (NPV) of each option was also determined. The NPV of fractured injection was calculated to be -25mm$ whereas the NPV of matrix injection was calculated to be -50mm$, thus demonstrating that fractured injection is a far more economically viable option.

Several additional cases were considered for horizontal injectors although a complete economic analysis was not carried out. From an engineering point of view, horizontal injectors would improve injection dramatically, and, in most cases, only one horizontal well with fractures is required. Consequently, this can greatly improve the economics of CO_2 sequestration. The number of required wells and the associated flow rate and pressure is shown for a variety of injector geometries in Fig. 3. As seen in the figure, a horizontal well with 4 fractures can achieve the required daily flow rate at the minimum pressure.

Table 2: Cost of the fractured and matrix option per year.

	Cost (mm$)/Year	
	Fractured	**Matrix**
Drilling	0.2	1.2
Hydraulic Fracturing	0.03	0
Abandonment	0.04	0.24
Compressor	0.15	0.45
Pipeline	0.1	0.3
Power	1.5	1.5
Maintenance	0.11	0.46
Labor	0.13	0.55



Injector Scenarios for Immiscible gas		Number Of wells Needed	Rate (m3/day) Per well	Wellhead Pressure (kPa)
	Matrix	4	7200	6895
	Fractured (180m)	2	18500	6895
	Horizontal (1200m No fracture)	1	25200	6827
	Horizontal (1200m 4 fractures)	1	25200	2177
	Horizontal (600m No fracture)	2	20900	6895
	Horizontal (600m 4 fractures)	1	25200	2315

Figure 3: Injection rates and pressures for a variety of injector geometries.

3 Conclusions

Saline aquifers have adequate capacity to serve as reliable sites for carbon dioxide sequestration for the foreseeable future. However, for low permeability formations, hydraulic fracturing of the injectors is necessary to achieve reasonable sequestration rates. The appropriate fracture length depends on the formation permeability and geomechanics moduli, stresses, and strengths of the different vertical lithologies of the subsurface column at the injection site. The effectiveness of vertical fractures in enhancing carbon dioxide injectivity increases as the reservoir absolute permeability decreases. The full life-cycle economic analysis indicates that costs can be greatly reduced by using fractured injectors compared to matrix injection. In addition, the use of properly designed fractured injectors, instead of injection in naturally fractured formations provide an assurance process for the fate and extent of the injected CO_2 plume.

Acknowledgments

This work was partially supported by the U. S. Department of Energy (DOE) under grants DE-FG03-97ER14778 and DE-FG03-97ER35332. This financial support does not constitute an endorsement by the DOE of the views expressed in this paper. This work was supported by Los Alamos National Laboratory for the DOE under contract DE-AC04-94AL85000 and by Sandia National Laboratories for the DOE under contract DE-AC04-94AL85000.

References

[1] Bergman, P. D. and Winter, E. M., "Disposal of carbon dioxide in aquifers in the US," *Energy Convers. Mgmt.*, **36**(6-9), 523–526, 1995.

[2] Gunter, W. D., Perkins, E. H., and McCann, T. J., "Aquifer disposal of CO_2-rich gases: Reaction design for added capacity," *Energy Convers. Mgmt.*, **34**(9-11), 941–948, 1993.

[3] Hendriks, C. A. and Blok,K., "Underground storage of carbon dioxide," *Energy Convers. Mgmt.*, **34**(9-11), 949–957, 1993.

[4] Koide, H, Tazaki, Y., Noguchi, Y., Nakayama, S., Iijima, M., Ito, K., Shindo, Y., "Carbon dioxide injection into useless aquifers and recovery of natural gas dissolved in fossil water," *Energy Convers. Mgmt.*, **33**(5-8), 619-626, 1992.

[5] Law, D. H.-S. and Bachu, S., "Hydrogeological and numerical analysis of CO_2 disposal in deep aquifers in the Alberta sedimentary basin," *Energy Convers. Mgmt.*, **37**(6-8), 1167-1174, 1996.

[6] Reichle, D., Houghton, J., Kane, B., Ekmann, J., Benson, S., Clarke, J., Dahlman, R., Hendrey, B., Herzon, H., Hunter-Cervera, J., Jacobs, G., Judkins, R., Ogdent, J., Palmisano, A., Socolow, R., Stringer, J., Surles, T., Wolsky, A., Woodward, N., and York, M., "Carbon Sequestration - State of the Science," Working Paper on Carbon Sequestration Science and Technology, U.S. Department of Energy, Office of Science, 1999.
[7] Van der Meer, L., "The CO_2 storage efficiency of aquifers," *Energy Convers. Mgmt.*, **36**(6-9), 513-518, 1995.

Effects of a thinning regime on stand growth in plantation forests using an architectural stand growth model

Y. Chiba
Forestry and Forest Products Research Institute, Japan

Abstract

The architectural form of a tree results from a combination of its physiology and mechanical support requirements. Using mathematical models of tree architecture that describe the mechanical relationships between different tree organs, a forest growth model was developed to predict the effect of thinning operations on tree/stand growth. The model is based on the height growth of trees, which is strongly correlated with forest site condition. The crown form of a tree alters as a result of available growing space within the stand, which, thus, affects crown depth. The parameter crown depth is a key characteristic for quantifying individual tree growth in order to reflect the effect of the thinning regime in plantation forests. The applicability of the model was verified using real stand growth data under various thinning regimes. Since the model includes not only the crown form but also the stem form of individual trees, we can predict tree size (clear bole length, stem diameter, ring width), stem shape, tree weight and stand biomass over the course of stand development for a variety of thinning operations. Without any thinning, the forest canopy will close at an early growth stage, there will be a high stand density, and the crown depth will be restricted and will exhibit poor growth. In contrast, forest plantations that are exposed to a number of thinning operations, at an appropriate intensity and timing, will exhibit the required biomass growth and wood quality. The effects on CO_2 fixation were also discussed with reference to thinning regimes in plantation forests.
Keywords: biomass increment, NPP, stand density control, biomass allocation, tree architecture.

1 Introduction

It is assumed that forested areas contribute to the fixation of atmospheric carbon dioxide, which would otherwise promote global warming [3,9]. Much effort has been exerted in the monitoring of CO_2 budgets, including eddy covariance measurements in a variety of forested areas across the world, and quantifying the potential CO_2 sequestration of various forest functional types and climatic conditions [2]. Forests in a natural environment are likely to be affected by a variety of factors such as climatic conditions, stand density, transitional stages and soil conditions. We, therefore, need mechanistic and process-based simulation models [2,6] in order to evaluate CO_2 fixation by forests and the effects of both global environmental change and human activities, including forest management.

The CO_2 budget of forest ecosystems remains unclear even in man-made forests where tree size distribution, forest structure, and environmental gradients in the canopy appear to be considerably simplified. In this paper, a stand growth model is presented; it simulates the growth processes of forest trees in even-aged plantation forests with different thinning treatments. This model is based on the architectural development of trees with changing crown structure over the course of thinning. Architectural models of tree form [3–5] have been formulated to describe quantitative patterns of branching structure and stem formation. The objectives of this study were to quantify the architectural development of forest structure as affected by the forest management associated with various thinning regimes.

2 Data used

In order to understand and compare the effect of thinning treatments on timber production in plantation forests, many reports relating to forest biomass were investigated particularly those about sugi (*Cryptomeria japonica* D. DON) plantations in Japan [1,8]. Of the reports linking productivity with stand density, Ando et al. [1] explored the relationship between biomass yields and tree density in several thinned sugi plantations in some traditional forest regions (Yoshino, Nishikawa, Obi, National Forests in Fukushima prefecture). They provide data sets comprising tree sizes (tree height, diameter at breast height, clear length below crown, dry weights of each organ) and tree densities before and after thinning operations. These data were used to characterize general thinning regimes for sugi plantations in order to validate the model presented herein.

3 Model description

3.1 Crown form

Since tree density (number of stems per hectare) determines the average growing space per tree, the average crown diameter in a stand can be determined using

this figure. Crown shape is specific to each conifer species [7]. Denoting crown width and crown depth for isolated trees as *CW* and *CD*, respectively, the following reciprocal equation was defined for conifers:

$$\frac{1}{CW} = \frac{A}{CD} + \frac{1}{C_{max}} \tag{1}$$

where A is a constant that represents the angle at the apex of a tree crown, and C_{max} is the maximum crown diameter. The constant C_{max} is dependent on the distance between adjacent trees in the stand, i.e. stand density.

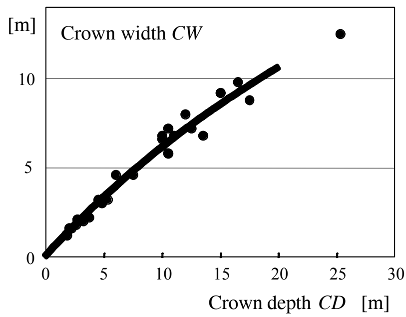


Figure 1: Relationship between crown width and crown depth for conifer crowns.

3.2 Tree height growth

It has been reported that tree height is strongly influenced by forest site conditions, including nutritional and water status of the soil. Many researchers have also pointed out that stand density exerts little effect on height growth [8] in plantation forests. In this paper, height growth in sugi plantation forests was approximated by the following formula, which was determined from the data measured in the different stands, irrespective of stand density:

$$H = H_{max}\,(1 - \exp(-a_H / H_{max}\; x_A)) \tag{2}$$

where H_{max} denotes the maximum tree height [m] specific to the site, x_A is stand age [yr] and a_H is the initial height growth [m yr^{-1}] specific to the site.

3.3 Stem form

Stem form can be approximated by a hyperbolic function with the two asymptotes of the exponential function representing the upper and lower parts of the stem, respectively [5]. Denoting z as the distance from the apex of a stem, the S-curve [5] is given by

$$\frac{1}{S(z)} = \frac{1}{S_0 \exp(\, z/\alpha\,)} + \frac{1}{S(z_B) \exp(\,(\, z - z_B\,)/\beta\,)} \tag{3}$$

where $S(z)$ is the stem density (stem weight per unit length) at position z [g m^{-1}]; S_0 is the top stem density [g m^{-1}]; z_B is the distance from the apex to the crown base [m]; α is a constant specific to each individual tree, termed the "top taper length" [m]; and β is a constant specific to each individual tree, termed the "specific stress length" [m]. By definition, z_B corresponds to the crown depth.

Of the constants in eqn (3), S_0 and α can be regarded as being independent of tree size, and $S(z_B)$ and β can be approximated by an allometric function based on crown length (= z_B). Therefore, stem form eqn (3) can be represented by the values of crown length of each tree. In addition, each crown length is likely to be determined mainly by available growing space for the tree, so that average crown length is a function of stand density.

3.4 Tree weight

Since leaves and branches of an individual tree are restricted to the tree crown, some attributes of the crown could be employed to estimate the weights of leaves and branches of each individual tree. In this paper, the following allometric relationships between the weights of leaves and branches and crown depth were adopted:

$$w = a\, CD^{\,b} \tag{4}$$

where a and b are constants, CD is the crown depth and w is the weight of an organ (branches or leaves). Stem weight of each tree was approximated as follows:

$$w_S = g\,(D^2H)^{\,h} \tag{5}$$

where w_S is stem weight, D is stem diameter at breast height, H is tree height and g and h are constants.

3.5 Thinning effects on a tree crown

Just after tree planting, the canopy is unlikely to be closed and the tree crowns will not be overlapped so the height of the lowest branch could be zero. Once the canopy closes, lower branches gradually wither and are shed, and the height at the crown base will increase as the tree increases in height (fig.2). Provided that the stand density remains constant after canopy closure, the crown depth may remain almost constant. However, once the canopy is opened up by thinning, the height at the crown base may remain unchanged until the canopy closes again.

After the canopy closes again, the height at crown base may increase along with the tree height. In other words, crown depth will increase stepwise with every thinning operation. Thus, the process of crown development could be

linked to thinning treatments. In the present study, crown depths before and after thinning were calculated as follows. An average crown width *CW* was determined on the basis of stand density. From eqn (1), crown depth *CD* was determined from the *CW* for closed canopy. After thinning, the comparison between the current crown depths and the potential crown depth for the current stand density was used to determine whether the canopy had already closed.

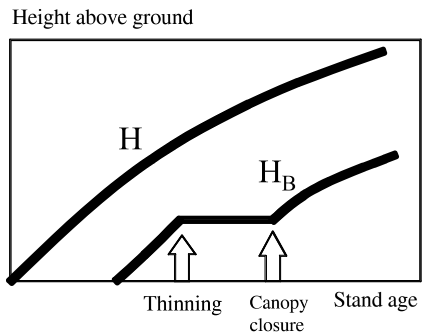


Figure 2: Schematic illustration of tree heights H and height at crown base H_B against stand age.

4 Results and discussion

4.1 Typical thinning regime in sugi plantations in Japan

For the purpose of timber production for general use, between 2,500 and 4,000 sugi cuttings or seedlings are planted per hectare. The plantations are thinned three or four times, reducing the stocking to a final level of about 700 stems per hectare. Thinning operations are usually completed by ca. 40 years of age. The trees are harvested at about 60 years. This pattern of tending work is widely used in Japan. In contrast, some private foresters undertake specific labor-intensive and carefully planned forestry in order to produce timber for specific uses such as alcove posts or “sake” barrels. They establish dense plantations with over 10,000 trees per hectare, and conduct frequent thinning operations to produce timber from appropriately sized trees, until the final harvest when the trees are around 100 years old. The chronology of this silvicultural system is summarized in table 1.

4.2 Tree size development

The results of the simulation are shown in fig. 3, as applied to typical thinning regimes for sugi plantation forests (table 1). Height at the crown base H_B and diameter at breast height D were based on actual tree height growth for the sites considered. At the initial planting stage, H_B was almost zero because the canopy had not yet closed. Although the stand age when canopy closure first occurs is

dependent on stand density, H_B appears to start increasing at a stand age of about 10 years. After this, the crown length (= H - H_B) shows a gradual increase due to repeated thinning treatments in the course of stand development. Figure 3 demonstrates the close correspondence between the simulation and the real data. Furthermore, stem diameter at breast height, D, is also closely approximated by the present model, which uses stem form as presented in eqn (3).

Table 1: Typical thinning regimes for sugi plantations in the National Forest and Yoshino.

Stand age (year)	National Forest (stems/ha)	Yoshino (stems/ha)
0	3,000	14,000
20	1,700	3,500
30	1,200	2,000
40	850	1,500
50	750	1,000
60	590	900

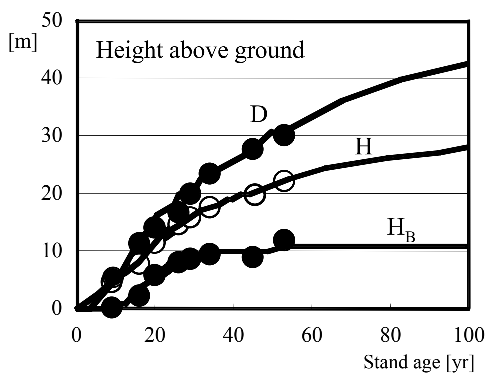


Figure 3: Simulation of tree sizes in a plantation forest with thinning. D, H, and H_B denote diameter at breast height, tree height and height at crown base, respectively.

4.3 Biomass growth

As mentioned above, the mean crown depth, *CD*, of an average tree in a stand can be calculated with a considerable level of accuracy for each stand age. Using these *CD*s, the weights of leaves and branches were calculated using eqn (4), and the stem weight was obtained from eqn (5). Finally, we simulated biomass development of sugi plantation forests as shown in fig.4. The saw-tooth appearance of the growth pattern for each organ is the result of the thinning treatments.

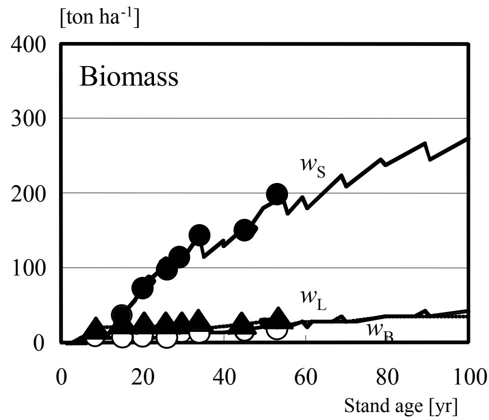


Figure 4: Simulation of biomass growth in a plantation forest with typical thinning regimes. Solid and dashed lines: results of the simulation; symbols: actual data of the biomass of stem w_S, branch w_B and foliage w_L, respectively.

It is likely that foliage and branch mass reach their maximum levels before the stand is mature. In particular, foliage mass tends to reach its highest level just after canopy closure in the model, this corresponds to the real data from the various sugi stands investigated to date. In contrast, stem mass exhibits continuous incremental growth throughout the life of the stand. Recent reports on biomass of old growth plantation forests demonstrate similar results, with higher stem growth than previously expected, indicating that even in stands over 70-80 years of age, the stem volume of the stand does not decline.

Using the simulation model presented here, CO_2 or carbon fixation in forest plantations was evaluated. Much higher CO_2 fixation was observed in the early stage of development of dense stands compared to other tending regimes. However, such dense stands are considered at high risk of abiotic disturbances, such as wind throw and snow damage, because of the very slender stems. In addition, it has been demonstrated that CO_2 fixation by these dense stands declines rapidly after canopy closure. In contrast, very low density plantations cannot always increase CO_2 fixation because they do not fully occupy the available space. Taking these results into account, it is possible that plantation forests managed in the most typical way, like the National Forests, maximize the sequestration of CO_2, providing effective and stable conditions.

References

[1] Ando, T., Hatiya, K., Doi, K., Kataoka, H., Kato, Y. & Sakaguchi, K., Studies on the system of density control of sugi (*Cryptomeria japonica*) stand. Bulletin of Government Forest Experiment Station, 209, pp. 1-76,1968.

[2] Baldocchi, D.D., Wilson, K.B. & Gu, L., How the environment, canopy structure and canopy physiological functioning influence carbon, water and energy fluxes of a temperate broad-leaved deciduous forest-an assessment with the biophysical model CANOAK. Tree Physiol, 22, pp. 1065-1077, 2002.
[3] Chiba, Y., Plant form analysis based on the pipe model theory. I. A statical model within the crown. Ecological Research, 5, pp. 207-220, 1990.
[4] Chiba, Y. & Shinozaki, K., A simple mathematical model of growth pattern in tree stems. Annals of Botany, 73, pp. 91-98, 1994.
[5] Chiba, Y., A quantitative analysis of stem form and crown structure: the S-curve and its application. Tree Physiol, 7, pp. 169-182, 1990.
[6] Farquhar G.D. & von Ceamerrer, S., Modeling of photosynthetic response to environment. Encyclopedia of plant physiology, New Series, 12B: Physiological Plant Ecology II, ed. O.L. Lange, P.S. Nobel, C.B. Osmond & H. Ziegler, Springer-Verlag: Berlin, pp. 549-587, 1982.
[7] Kobayashi, S., Studies on the simulation model of stand growth of Japanese larch (Larix leptolepis GORD.) plantation. Bulletin of the Hokkaido Forest Experiment Station, 15, pp. 1-164, 1978.
[8] Tadaki, Y., & Hatiya, K., Forest ecosystem and its dry matter production. (Forestry Research Exposition Series 26). Ringyo Kagaku Gijutsu Shinko-sho: Tokyo, pp.1-64, 1968.
[9] Thornley, J.H.M. & Cannell, M.G.R., Modeling the components of plant respiration: representation and realism. Annals of Botany, 85, pp. 55-67, 2000.

Section 5
Design and sustainability, organised by W. P. de Wilde

Adaptable versus lightweight design of transitory dwellings

W. Debacker[1], C. Henrotay[2], W. P. de Wilde[1] & H. Hendrickx[2]
[1]*Department of Mechanics of Materials and Constructions (MeMC), Vrije Universiteit Brussel, Faculty of Engineering Sciences, Belgium*
[2]*Department of Architecture (ARCH), Vrije Universiteit Brussel, Faculty of Engineering Sciences, Belgium*

Abstract

Optimisation of structures and materials is a justifiably popular engineering topic. Contemporary research is concentrated, among other things, into cost minimisation, structural efficiency and intelligence, in compliance with environmental and social preservation. This paper puts the accent on the time dependent aspect of constructions, such as the life cycle cost, the possibility to make (non-) structural changes and recycle or reuse of building materials. In search of an optimisation of this ‘dynamic’ efficiency of constructions, a design strategy has been developed at the Vrije Universiteit Brussel (dept. MeMC & ARCH). A logical question arises: can both types of optimization coexist? The answer is given through a realistic case study: transitory dwellings.
Keywords: adaptable, reuse, lightweight, optimization, transitory dwelling.

1 Introduction

A typical Western dwelling has an average life span of about a hundred years; a period which even the best futurologist won’t burn his fingers on [1]. During this lapse of time households and habitats can change radically and repeatedly. Technical and social revolutions have already modified the meaning and the shape of the dwelling in the past decades. Which unforeseen defining moments shall housing deal with in the coming years?

Nowadays designers are faced, more than ever before, with the difficult task of giving form to dwellings for a period during which the composition of the

www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060331

household and the associated spatial rites will go through major changes. ‘Flexibility’, ‘adaptability’ and ‘versatility’ are the key elements to confront this unpredictable future! Hence a sustainable built environment requires dynamic design strategies in which no end states or final goals can be defined.

The “4Dimensional Design Strategy” (4D), invented by H. Hendrickx and H. Vanwalleghem includes this dynamic view on the built environment [2]. 4D is a guide to design multiple construction systems, all compatible to each other, by which a variety of adaptable and reusable construction elements can be composed. Each construction system is made of a minimum number of basic elements and a set of combination rules. They allow the conversion of each artefact to a different configuration, by means of adding, removing or transforming the basic elements which it is made of. It offers a high potential of recycling and (direct) reuse. The outcome can be compared with the ‘Meccano’ building set, which, in this view, encloses all materials and techniques, and is applicable to all scales. Hendrickx and Vanwalleghem proposed a set of standardisation rules, which they called a “generating form and dimensioning system”. The generating system is a central concept in the design strategy, in the sense that it ensures full compatibility of form and dimensions between all basic elements. The rules are translated into a *fractal model*, based on basic forms, such as the square, the inner circle and its diagonal, and a dimensional range using the operator “multiply or divide by 2”.

2 Flemish social housing: a face lift!

Similar to our neighbouring countries, Belgium is experiencing a shortage of adequate (social) habitats. Many dwellings do not meet the contemporary standards anymore. Subsequently the Flemish government currently invests in building a thousand new houses [3]. It seeks ecological solutions with the intention of lowering the “total life cycle cost”. Flexible building methods are highly coveted, since the dwellers group is very heterogeneous (culture, composition, age and social background) and the demand of a social habitat is still growing. Reuse of elements and the possibility to make structural and non-structural changes is essential to enhance the possibilities of the social dwellings.

In anticipation of the construction of new dwellings and the reinstatement of existent ones, qualitative and low cost temporary housing is required. Transitory habitats will be offered to (re)allocate underprivileged people for a span of 3 to 6 months – exceptionally 1 year [4]. Since the Flemish Ministry of social housing is carrying out a big scale reconditioning of its entire heritage – an ongoing project – the idea has popped up to make use of re-movable transitory habitats that are easily transported from one place to another. Flexible building methods, which are compatible to the existing transportation standards, are thus also indispensable for short term solutions.

Transitory housing will not offer any structural answer to the Belgian housing problematic. Nevertheless it is a way to counter the process of social disintegration which comes along in becoming homeless [4].

3 The design of transitory dwellings

3.1 Design constraints

In partnership with the steel producer Arcelor and thanks to the input of the Flemish Ministry of Social Housing, the departments of Architecture (ARCH) and Mechanics of Materials and Constructions (MeMC) of the Vrije Universiteit Brussel have designed a construction kit for temporary constructions. All base elements are interchangeable and resizable, making them the construction elements for a multitude of configurations, such as various dwelling types. A list of the main design constraints is given below:

- The temporary necessity of the dwellings imposes a *reversible impact* on the environment. The structure and its components have to be assembled, dismantled and carried away without leaving permanent traces which can harm the construction site.
- An *effective transport* method in compliance with (inter)national transport regulations is required. Unused transport space is to be minimised and 'intermodal' possibilities is to be maximised.
- The structure of the transitory dwelling has to be *easily assembled and dismantled* in a short period of a few days to cope with changing situations, such as a succession of dwellers and reconditioning of an existing habitat, or unforeseen events, such as a flood or a fire which makes an existent dwelling temporary inhabitable.
- Frequent assembling, dismantling and transporting will make some parts vulnerable to accidental charges and deterioration, such as corrosion and fatigue. Besides an intelligent choice of the building material, several *maintenance* procedures (during a life cycle of 20 years) should be taken into account.
- A minimum of *comfort and safety* has to be provided. Performances concerning fire safety, thermal and acoustic insulation, hygiene, heating, cooling and ventilation are well defined by the Flemish Housing Society [3].

3.2 Composition of the construction kit

Based on the previously mentioned constraints a construction kit was designed. It is composed of 4 major types of simple construction elements:

1. *bearing frames* (end and intermediate frames)
2. *horizontal girders* (base structure and girders)
3. *sandwich panels* (floor, roof, side and end walls)
4. *dry connections* (bolts and twistlocks)

Each of them is made of 'neutral' base elements such as *profiles* (one-dimensional), *plates* (two-dimensional), hollow and plain *volumes* (three-dimensional). Once they are assembled together they acquire a

constructive meaning! Thanks to the generating system, each base element is compatible with any other – and thus interchangeable. The above mentioned construction components are hereby resizable and transformable as required using a minimum of base elements.

Since sanitary facilities and a kitchen had to be incorporated into the dwellings, a *fifth* hybrid construction element type is introduced, i.e. the *wet cell*. It is composed of the 4 former types of construction elements and encloses the necessary services to allow the daily rituals. These (small) prefabricated spaces will move up the construction speed, since they only need to be plugged into the interior surrounding where electricity and water supply is available.

The proposed dwellings (Fig. 5) are too big to transport them in one piece. To allow 'intermodal' transport (train, truck and inland waterway) the design was oriented to a 20' ISO container wherein the bearing frames, horizontal profiles and sandwich elements of 5 equivalent 20' units are compacted to each other (Fig. 1). Two wet cells fit likewise into a 20' ISO container. The ISO container can be composed of elements of the kit or a second hand cargo container can be reconditioned. Following these design tips transport and storage cost will drop significantly.

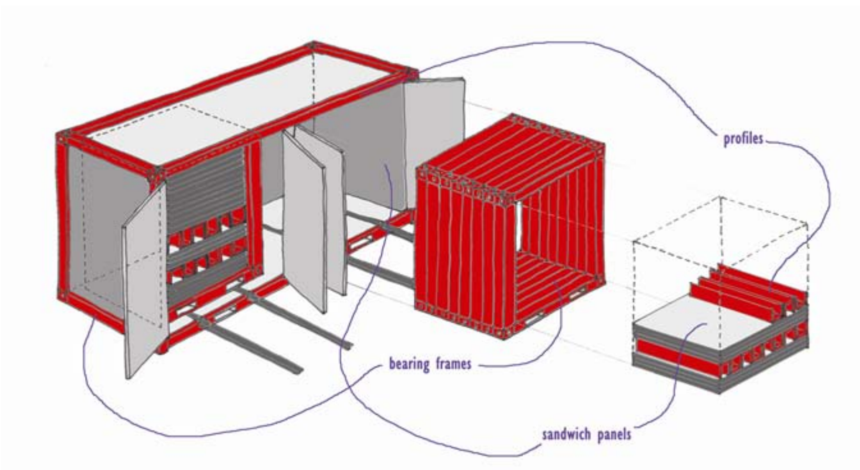


Figure 1: Composition of the construction kit during transport.

3.3 Design catalogues

An aid to the development of the construction kit was to draw up design catalogues for the 4 major types of construction parts. Each of them was decomposed into design variables:

1. *bearing frame*: thickness of plate; length, width, depth and sectional form of profile, material type; form of ISO corner element; way of composition of the frame (monolithic – composite); way of connecting.

2. *horizontal girder*: thickness of plate; length, width, depth and sectional form of profile, material type; way of composition of girder (monolithic – composite); way of connecting.
3. *sandwich element*: thickness and form of faces, material type of faces and insulation, thickness of insulation; length and width of the sandwich panel; opening type and form; way of composition of sandwich element (monolithic – composite); way of connecting.
4. *dry connections*: connection type; bearing section (diameter) and length of connection, number of elements, material type.

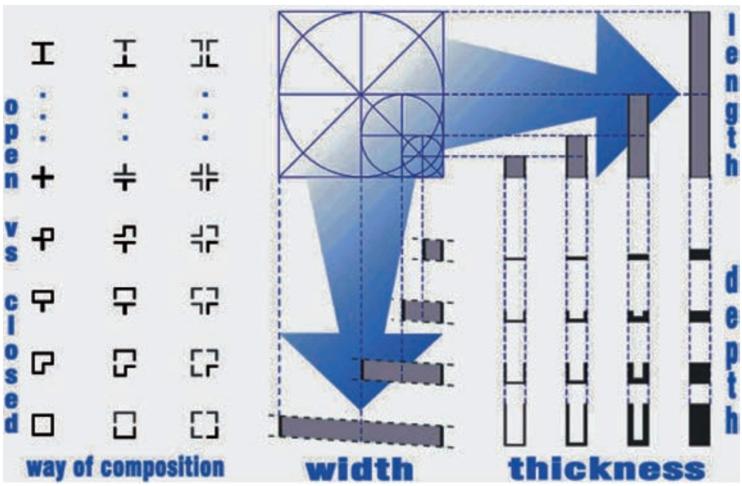


Figure 2: A partial design catalogue of steel horizontal girders.

Once the variables determined, series were drawn up for each of them. First extreme 'values' were considered, after which intermediate values were depicted through interpolation and/or combination of these outer points. This action was done with the help of the fractal model of the generating system. An example is illustrated in *Fig. 2.* A theoretical design catalogue is thus established, combining and juxtaposing elements. The emphasis has been put on 'theoretical', as in practice not all combinations are possible or technically sound. This means that they are erased in the practical catalogue.

3.4 Dry connections

The construction kit includes two kinds of dry connections, i.e. "*bolts et al*" and "*twistlocks et al*". The idea is to use the former for the composition of the major construction parts, such as the girders and the frames. The latter are used for frequent combinations between elements, such as the assembly of construction elements on the field. The twistlock is a well known device in the field of international transportation, since it is employed to secure containers during stacking (in cargo ships or at the docks) and the connection with the bodywork of trailers. In these cases *vertical* connections are made with either the top or bottom corner fitting of containers.

It is our intention to improve the existing twistlock system by adding *horizontal* connection possibilities. This means however that the corner fittings and twistlocks are to be reinvented. The existing bottom and top corner fittings are integrated in one volume with multiple connection possibilities: twistlocks, bolts, plugs et al (*Fig .3)* Combinations with other base elements are possible via the 6 faces. The 'revisited' twistlock (*Fig. 4*) is composed of a steel wire whereon different plates are fixed. The head of the twistlock is removable and can be replaced by an other one that fits in the desired hole of the polyvalent ISO corner element. By doing so, vertical and horizontal connections are made possible using the same device. To open or close, the handle must only turn a quarter of a circle.

Figure 3: Polyvalent corner fitting.

Figure 4: Adaptable twistlock.

The features of these dry connection elements endorse the '4dimensional aspects' of the construction kit. Besides the ease of assembly and dismantling, the maintenance of construction parts and the reuse of components in other constructions are made possible. Damaged basic elements can be reused allowing small transformations, even in fields where the structural, i.e. load bearing properties, are less important (furniture, window frames…). Hence the "total life cycle cost" of the elements will decrease. A good design of the connection is thus imperative.

4 Structural analysis

During the design of the transitory dwellings the question popped up if the 4D strategy could be coupled with structural optimization? Is it possible to present the user of the kit a minimum of compatible basic elements which can be assembled into multiple lightweight components and still meet the normative strength, stiffness and stability regulations? The weight of the components will determine the way and the facility of handling during assembly and dismantling.

A catalogue of structural U-profiles generated by the fractal model (Ugen) was entered into finite elements software. By using this tool the minimal

dimensions of a steel skeleton of 4 different dwelling configurations (Fig. 5) were determined. The study was restricted to static analysis and characteristic local charges.

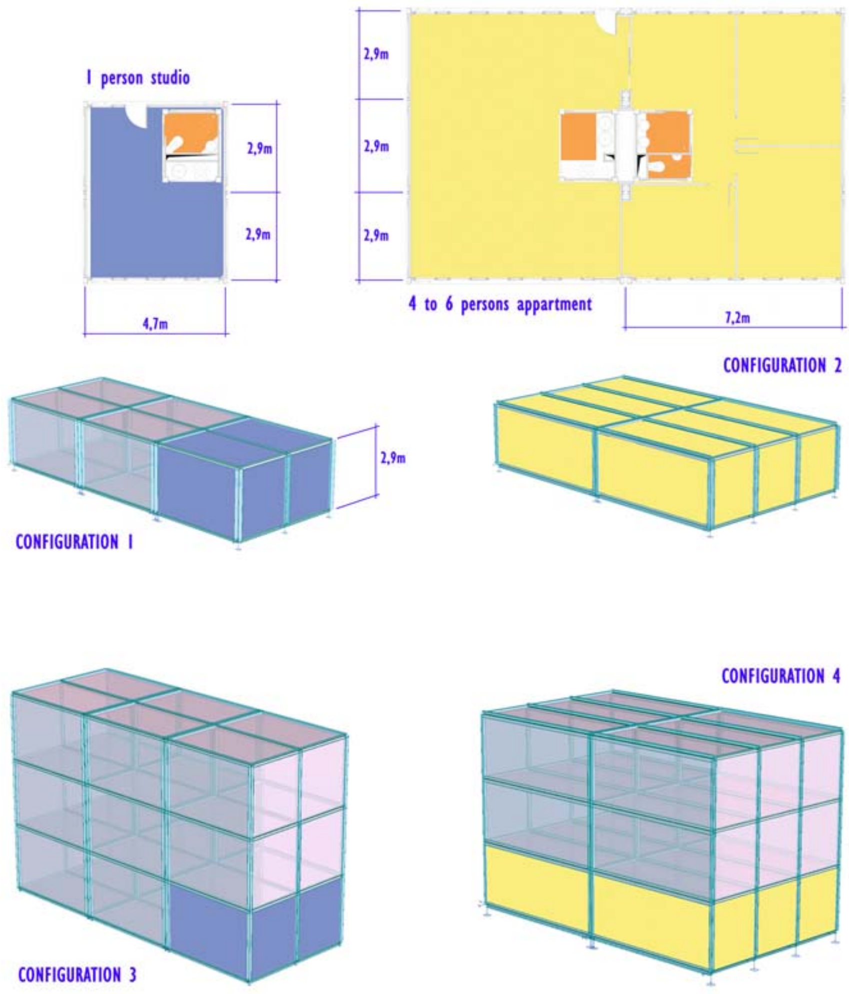


Figure 5: Four different dwelling configurations made with the same construction kit.

A comparison between the Ugen-profiles and typical U-sections available on the market (UAP, UPE, UPN) gives an idea about the relative performance of the structural properties of the construction kit. A summary is given in the following table.

The above mentioned U-profiles fulfil all structural constraints on the construction. The weight of (too) slender sections was penalised with the extra weight of the compulsory stiffeners to prevent local buckling.

Table 1 shows that for the 4 scenarios the Ugen-sections compete with the U's currently available on the market. The choice of only considering open sections was strictly a technical /logistical one. Compared to tubular sections, which have better structural properties, U-profiles are easier to connect and disconnect.

Table 1: Optimal U profiles for the 4 dwelling configurations.

CONFIG. 1	length (m)	Ugen type	Ugen mass (kg)	UAP type	UAP mass (kg)	UPE type	UPE mass (kg)	UPN type	UPN mass (kg)
FRAME: vertical	2,9	153x153x3,6	40,0	150	51,9	160	49,2	160	54,3
FRAME: horizontal	2 x 1,98	153x153x4,8	38,1	150	35,5	160	33,7	160	37,1
GIRDER	2,5	153x153x3,6	36,8	150	45,4	160	43,1	160	47,5

CONFIG. 2	length (m)	Ugen type	Ugen mass (kg)	UAP type	UAP mass (kg)	UPE type	UPE mass (kg)	UPN type	UPN mass (kg)
FRAME: vertical	2,9	153x153x4,8	55,6	150	51,9	160	49,2	160	54,3
FRAME: horizontal	3 x 2,0	2 x [153x153x3,6]	2 x 27,6	250	68,1	270	69,7	260	74,7
GIRDER	2,5	153x153x3,6	36,8	150	45,4	160	43,1	160	47,5

CONFIG. 3	length (m)	Ugen type	Ugen mass (kg)	UAP type	UAP mass (kg)	UPE type	UPE mass (kg)	UPN type	UPN mass (kg)
FRAME: vertical	2,9	229x229x3,6	63,4	175	61,5	180	54,4	180	63,3
FRAME: horizontal	2 x 1,98	153x153x4,8	38,0	150	35,5	160	33,7	160	37,1
GIRDER	2,5	153x153x3,6	36,8	150	45,4	160	43,1	160	47,5

CONFIG. 4	length (m)	Ugen type	Ugen mass (kg)	UAP type	UAP mass (kg)	UPE type	UPE mass (kg)	UPN type	UPN mass (kg)
FRAME: vertical	2,9	229x229x4,8	84,2	200	72,7	200	66,0	200	72,8
FRAME: horizontal	3 x 2,0	2 x [153x153x3,6]	2 x 27,6	250	68,1	270	69,7	260	74,7
GIRDER	2,5	153x153x3,6	36,8	150	45,4	160	43,1	160	47,5

5 Conclusions

The combination of structural optimization and the 4D design strategy offers a double benefit! Considering the fact that the volume of the structural components is minimised, the initial material cost will drop. In addition the construction, maintenance and demolition cost will be minimal since the principle of reusing, easily assembling and dismantling is incorporated into the design. The result is highly interesting: a minimisation of the construction cost and the construction

energy over the whole life cycle of the artefact and its components. Although this study has been done for transitory dwellings, this design philosophy can be extrapolated to other constructions as well (fig. 6).

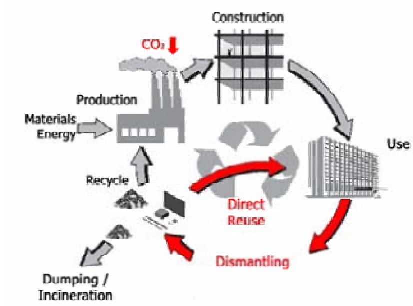


Figure 6: Proposed life cycle model of a construction.

References

[1] Leupen B, *Kader en generieke ruimte – Een onderzoek naar de veranderbare woning op basis van het permanente*, Uitgeverij 010, Rotterdam, 2002, 254p.

[2] Hendrickx H., Solutions derived from natural processes harmonising nature and material culture. *Proc. of the 1st Conf. on Design and Nature 2002. Comparing Design in Nature with science and engineering*, ed. C.A. Brebbia & L.J. Sucharov, Wessex Institute of Technology, United Kingdom and P. Pascolo, Universita degli di Udini, Italy, 10p. 2002.

[3] WTCB / VHM, *ABSW 2003: Algemeen Bestek voor de Sociale Woningbouw: Prestatiebestek,* Vlaamse Huisvestingsmaatschappij, Brussels, december 2003, 152p.

[4] Roggemans M-L et al, *Transit-huisvesting: Vangnet en Springplank: Evaluatie van en perspectieven voor transit-huisvesting in het Brusselse Hoofdstedelijk Gewest*, Koning Boudewijnstichting, Brussel, 1994, 82p.

This page intentionally left blank

4-dimensional design: a design strategy for efficient shelter and sustainable housing after conflict-based and natural disasters

C. Henrotay[1], W. Debacker[2], M. Mollaert[1], W. P. de Wilde[2]
& H. Hendrickx[1]
*[1]Department of Architectural Engineering Sciences (ARCH),
Vrije Universiteit Brussel, Faculty of Engineering Sciences, Belgium*
*[2]Department of Mechanics of Materials and Constructions,
Vrije Universiteit Brussel, Faculty of Engineering Sciences, Belgium*

Abstract

Catastrophes like wars and natural disasters have occurred over the ages. However, in the last decade the number of emergency situations has increased in a frightening way. Unsustainable human behaviour causing environmental changes and social conflicts undoubtedly contributes to this increase. A global approach for shelter intervention may contribute to efficient relief while supporting a transition from aid to development and helping to prevent future catastrophes. In this paper one investigates how the "4-dimensional design strategy" developed at the Vrije Universiteit Brussel (department of Arch & MeMC) can be implemented to reach this goal.
Keywords: disasters, emergency situations, shelter, reconstruction, adaptability, reuse, design strategy, design catalogue, construction system, construction kit.

1 Shelter intervention and reconstruction after natural and conflict-based disasters

1.1 Shelter intervention in emergency situations

The latest natural disasters have shown that catastrophes – predictable or unpredictable – and the resulting emergency situations may occur all over the world, in developing as well as in developed countries. Unfortunately, the international aid community still seems insufficiently prepared for these

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060341

incidents. Assistance in emergency situations, especially shelter intervention, is extremely complex. Relief workers have to manage quick interventions, a great amount of varying parameters and the unknown course of the situation.

Research has indicated that the current international relief actions are deficient partly due to inappropriate infrastructure [1]. An important cause of this deficiency can be found in the lack of adaptability and polyvalent use of the existing shelter material, which is often poorly adapted to the local situation. In addition the potential for adaptation or expansion during later phases of assistance is limited or inexistent and less attention is paid to reuse of shelter material within later reconstruction phases of humanitarian assistance.

1.2 Reconstruction and affordable housing

Although natural hazards can threaten anyone the poorest communities, especially in the developing countries, are usually hurt most. These communities are vulnerable because they tend to live in greater density in badly-built housing on land at risk [2]. Even before being hit by a catastrophe most developing countries are facing important socio-economic problems among which the lack of appropriate housing. While UN and nongovernmental organizations tend to make housing a basic right; adequate housing is still unaffordable for a large part of the world population. The UN Human Settlements Program (UN- HABITAT) estimates that 600 million urban residents and 1 billion rural dwellers in developing countries live in inadequate housing [3]. In addition, according to ECHO (European Community Humanitarian Office) the economic loss suffered by developing countries after natural catastrophes (as a percentage of GDP) far exceeds that in developed countries. Disasters reduce the output of the poorest nations by around 3%, depriving them of resources needed to escape poverty. As a consequence the poorest communities affected by a catastrophe are often unable to manage the consequences of a hazard and are exposed to an increasing poverty and vulnerability.

1.3 Construction industry and the over consumption of material and energy

Besides producing unaffordable buildings the building industry is responsible for generating an enormous amount of waste. According to the Statistical Office of the European Commission construction and demolition waste correspond to 31% of the total waste generated in Western European Countries for the period 1990 - 2001 [4]. The current build environment is designed in terms of end state and thus unable to support the quickly evolving trends. This waste represents an inefficient use and an enormous loss of resources in the form of material and energy. Moreover the waste and the endless production of new construction material is an important cause of pollution.

It is important to emphasis that the increasing amount and impact of natural hazards can be related to the increasing poverty but also to environmental changes caused by pollution and the exhaustion of natural resources. On the other hand, beside the political, ethnical and religious factors, the (over)consumption of resources plays an important role in conflict-based disasters. According to a report of the Belgian Ministry of Foreign Affairs and

Foreign Aid and Development 25% of the current armed conflicts are caused by quarrels regarding valuable natural resources or are financed by them [2]. Reducing the consumption of natural resources by increasing the efficient use of material means may contribute to prevent future catastrophes.

1.4 Design for adaptability and efficiency

In this research project one strives for efficient shelter and housing interventions after natural or conflict-based disasters while anticipating and possibly preventing future hazards by efficient use and reuse of material and energy.

Quick and efficient shelter intervention in emergency situations is only possible if one is prepared for new catastrophes to come. The worldwide installation of *national and international stockpiles* may contribute to this preparedness. However, this requires a global and universal approach for the design, the production and the use of shelter material resulting in adaptable, adequate and compatible shelters and shelter components. The *design of adaptable and polyvalent building systems* that offer the opportunity to support different shelter strategies, to be easily adapted to the local factors and to be upgraded from an emergency shelter to a medium or long term habitat may support this approach.

The design of adaptable construction systems may also offer the opportunity to create affordable and adequate minimal housing that can be moved, upgraded and extended according to the needs and the financial means of the inhabitants, and thus reduces the vulnerability of the poorest populations.

The ability to transform, to dismantle, to move and to reuse relies on a dynamic and process-based behaviour of constructions. Hence, *a dynamic and global concept* is required to support a step-by-step redesign process of gradual changes in which no end states or final goals can be defined.

2 The 4-dimensional design strategy: a global approach

The “4-dimensional design strategy” proposed and developed by professor Hendrickx (Vrije Universiteit Brussel) and architect Vanwalleghem proposes such a dynamic and global approach enabling the design of adaptable and generally usable construction systems. This strategy aims to design adaptable and compatible building systems consisting of a minimum amount of different building elements and their combination rules. Therefore the strategy encloses two important design tools: the use of a “generating form and dimensioning system” and the set up of design catalogues [5].

2.1 Design tool 1: generating system

The “generating form and dimensioning system” encloses a set of standardisation rules regulating different construction systems and their components by defining the shape and the dimension of the basic elements. While ensuring full compatibility of form and dimensions between all the simple basic elements, an infinite variety of building systems can be generated using the proposed system.

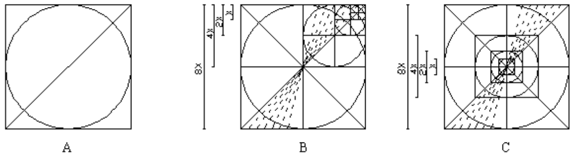


Figure 1: Generating form and dimensioning system.

The system rests on the fractal principle and a mathematical function guaranteeing compatibility and the capacity to adapt. The generating system relies on the 3 simple basic forms: the square, the inscribed circle and the triangle formed by the diagonal (Fig 1 – A). The dimensions are regulated by simple mathematical rules being halving and doubling (Fig 1 – B, C).

2.2 Design tool 2: design catalogues

Hendrickx and Vanwalleghem presuppose every artificial material solution can be considered as the result of combinations and can be objectively described based on its characteristics, qualities and defaults [6]. Each feature can then be decomposed into his single characteristics or parameters as a counterpart, all bracketed between predefined limits. As a result series of variants are drawn up for each single characteristic. Stepwise variations of the series are preferred since they reduce the number of values and the use of a minimal number of standardised basic elements can more easily be reached. Discrete series regulated by a set of standardisation rules, in this case the generating system, may guarantee the required compatibility. Through interpolation and/or extrapolation of the outer elements of each series all variants can be achieved. Recombining these variants results in a large variety of material solutions and the set up of theoretical design catalogues. The emphasis has been put on "theoretical", as in practice not all combinations are possible or technically sound. Practical design catalogues on the other hand, encloses only functionally and technically relevant material solution distilled from the theoretical design catalogues.

3 The set up of design catalogues

The design of adaptable construction systems relies on the development of theoretical and practical design catalogues. The compatibility is guaranteed by the standardisation rules defined by the generating system. Two different kinds of catalogues can be distinguished: the catalogues of basic elements and components and the catalogues of composed constructions.

3.1 Design catalogues of basic elements and components

3.1.1 Basic elements

The basic elements can be defined as the smallest independent mono-material part of a building system. Three types of basic elements can be distinguished: *linear elements* (one-dimensional), *plane elements* (two-dimensional) and *volume elements* (three-dimensional). All relevant characteristics of every single

basic element can be described verbally and represented graphically. After which series of variants - regulated by the generating system – can be drawn up, creating a theoretical design catalogue. Practically the different characteristics are interrelated depending on the material characteristics and the strength, stiffness and stability equations. Once a specific characteristic is defined, e.g. the length of a bar, the freedom regarding the variants of the other characteristics, e.g. the dimensions of the section, will be restricted.

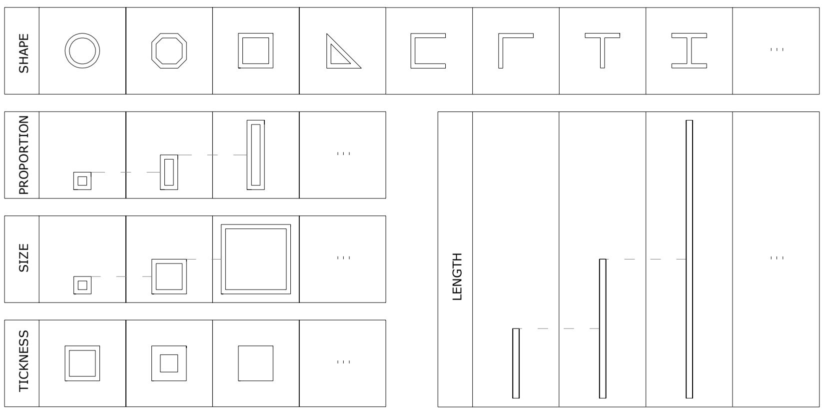


Figure 2: Series of variant characteristics of linear elements.

A variety of elements is created by defining their material and geometrical properties. At this stage the elements are defined as *basic sub elements* since they are completely neutral and unprocessed. Series of processing activities can be drawn up in which the nature of the manufacturing – perforating, incising, bending, etc. – the position and the amount of manufacturing points may vary. The different series may be applied to all basic sub elements upgrading them to *basic element*s with an encreased potential.

The basic elements can easily be standardised by projecting the model (Fig.1) on every single basic element. As a result beside the dimensions and the shape of the elements, the position and size of existing or potential manufacturing and connection points are defined by the model, guaranteeing compatibility between all basic elements. According to the material properties and building standards (e.g. Euro Code) the practically sound and possible basic elements are selected.

It is important to emphasis that every single basic (sub) element is neutral. This means that no function or semantic meaning is linked to the elements until they are implemented or combined in a specific situation. Depending on the degree of processing the element becomes more and more specific and the degree of neutrality will decreases.

3.1.2 Components

The basic elements, performing a wide variety of functions, can be combined to create a more complex unity or *component*. These components are defined as an

independent entity – assembled out of basic elements or smaller components – that have to be combined to a larger construction in which it will fulfil a specific function. A distinction can be made between *linear components*, *plane components* and *volume components.*

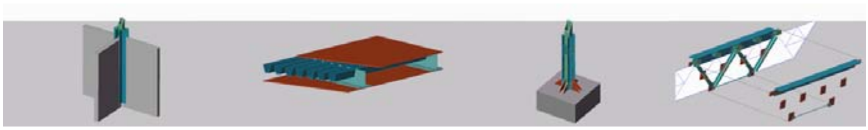


Figure 3: Components.

The basic elements can be combined according to series of combinations creating a wide variety of components that can be drawn up in the design catalogue. Since the spectrum of possible combinations is endless the design catalogues has to be considered as open design catalogues.

3.2 Design catalogues of composed constructions

Due to the compatibility in form and dimension the basic elements and components can easily be combined to form a variety of constructions and artefacts. Every composed artefact can in turn be considered as a component of an artefact of a higher level of complexity. As a consequence the design of the composed structures and constructions is also dictated by the generating system to achieve compatibility at all levels. However, since one strives for adaptable and thus reversible artefacts, other composed constructions may be possible, as long as they are made up by the basic elements and components of the design catalogues.

When drawing up the series of variants a first distinction can be made between open constructions – shading, windscreen, etc. – and closed constructions enclosing an entire space. In both cases their global characteristics – typology, structural nature, erection mode, anchoring and foundations, etc. – can be objectively described and analysed.

Due to the quick interventions and the relative short term of use the global characteristics of emergency shelter, especially the structural nature, the erection mode and the anchoring will differ considerably from the global characteristics of permanent constructions. Consequently depending on the mobility and the duration of use series of variants can be drawn up for the above characteristics.

In the following paragraphs only the design catalogues of the closed constructions and especially the mobile shelters (emergency shelter and transitional shelter) will be discussed. The series of variants for the typology, the structural nature and the erection mode are illustrated in (Fig. 4). Six basic typologies are defined. These will be used to organize the different shelter systems in a morphological and structural overview, where the basic typologies are contained in the first row and the structural nature is contained in the first column. For each structural type a subdivision is made enclosing the erection mode and a distinction between the internally or externally positioned structures. Combinations of the basic typologies, creating alternative typologies are not included.

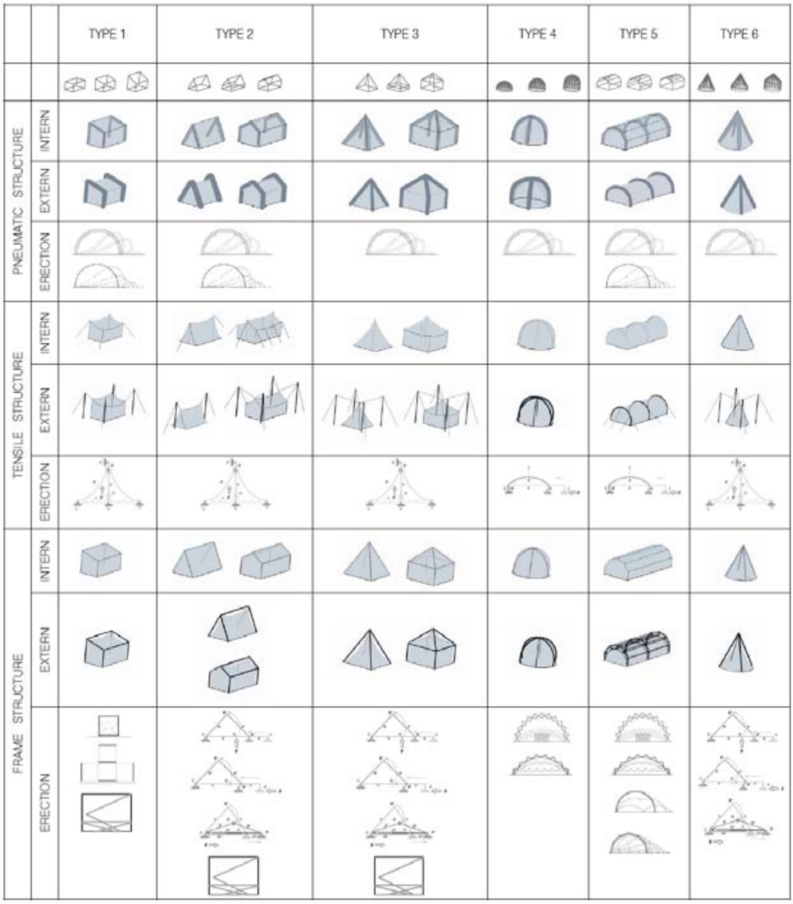


Figure 4: Morphological and structural overview of mobile shelter.

The erection mode is related to the typology and the structural nature but also to the mode of transport and the time span of the assemblage. The shelter may be transported in separated pieces, later on assembled in the field. On the other hand the shelter may be transported in a pre-assembled way enabling a quicker erection. According to Gengnagel and Burford [7] the possible erection modes of pre-assembled constructions are represented in Figure 4.

For each intersection of the table represented in Figure 4 new series of variants can be drawn up in regard to the type of anchoring, the size, the climate related characteristics (insulated, ventilated, etc.), the amount and position of internal subdivisions, etc. As the design catalogues of basic elements and components the *design catalogues of composed constructions* has to be considered as open design catalogues since the spectrum of possibilities regarding present and future constructions is endless. They can be completed by all composed artefacts as long as they are drawn from the standardization rules.

4 The design of adaptable building systems and construction kits

4.1 Adaptable construction systems

Construction systems are made up of a number of basic construction elements and components and their combination rules. These rules define how to combine the basic elements and components to become more complex components, structures and constructions. Depending on the material used and the nature of the structural and connecting system a variety of construction systems can be distilled out of the design catalogues of basic elements and components. Since all the basic elements and components are designed according to the same standardisation rules and the combination laws of the different systems rely on the same generating system, the elements and components of the different construction systems are compatible and can easily be combined.

The adaptability of the construction system depends on the nature of the connecting system and the degree of composition. Construction systems using dry connections are preferred since they ease the dismantling and the reuse. A high degree of composition increases the reversibility and adaptability.

4.2 Design of an adaptable shelter kit

A number of basic elements and components belonging to one or more building systems can be selected to constitute a *construction kit*. The elements of the construction kit can be combined to one or more, adaptable structures or constructions.

The process from sub element to shelter kit and adaptable construction is illustrated by the set up of an adaptable shelter kit (Fig. 5). The presented shelter kit [8] consists of aluminium bars and plates and pieces of canvas of different sizes and shapes, selected in the design catalogue of basic elements and components. All elements and components are designed for a maximal polyvalence. This means the elements and components can be used and combined in different ways to realise an infinit variety of solutions. Therefore the basic sub elements are upgraded by processing and/or combining them to become polyvalent basic elements and components.

The linear and plain aluminium sub elements are processed by adding perforations and incisions. The resulting elements (Fig. 5 – Elements A & B) may then be combined to create a variety of adaptable components – corner joints, protal frames, etc. (Fig. 5 – Components A).

The canvas sub elements are improved by adding perforations at the edge of the fabric (Fig. 5 – Elements C). A second improvement can be made by adding a strip of looped and perforated canvas at the edge of the fabric (Fig. 5 – Components B). The canvas is then upgraded from sub element to element and component.

The resulting elements and components constitute the proposed shelter kit. They can easily be combined to create a variety of adaptable shelters, classified

in the design catalogue of constructions (Fig. 4). Shelters of type 1 to 3 consisting of a tensile structure or a frame structure can be entirely realised with the components shown in figure 5. Due to their adaptability the shelters can evolve from a basic emergency shelter to a transitional shelter and a semi permanent construction [8]. As a result of the compatibility and their polyvalent design the basic elements can be reused and (re)combined with local materials - stamped mud, wooden boards, etc. - as long as the added elements are designed following the standardization rules. The poorest population unable to afford the reconstruction of a new house can then keep and reuse the basic elements of the shelter kit to (re)build their house.

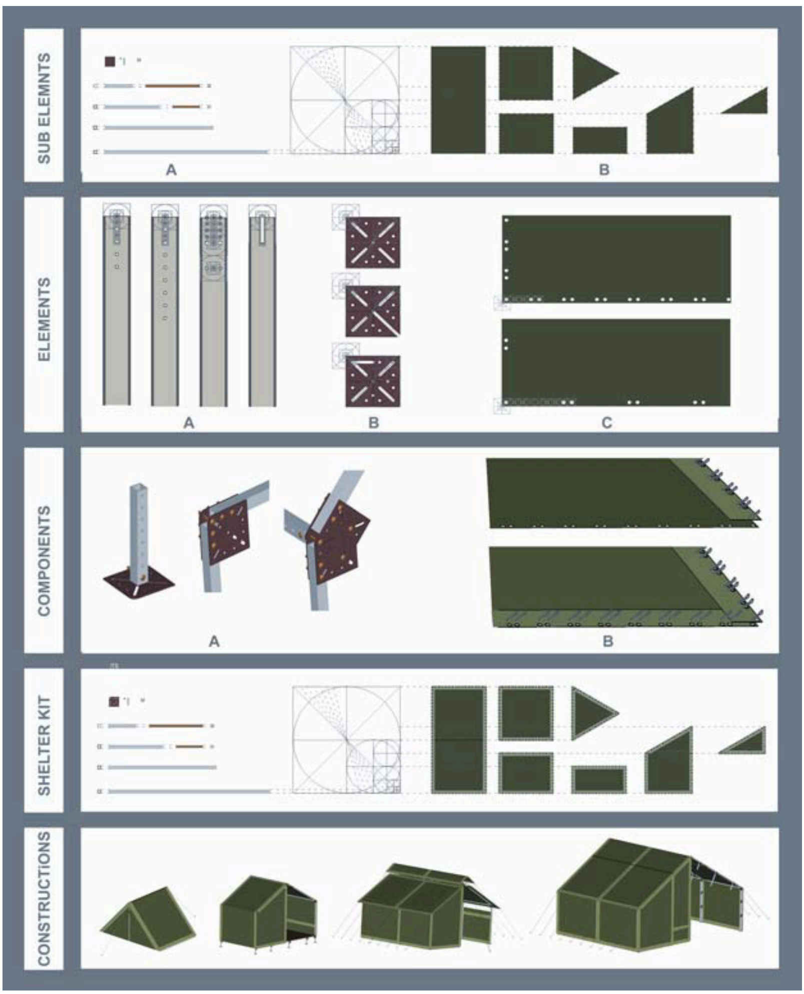


Figure 5: From basic sub element to adaptable construction system.

5 Conclusion

The use of a dynamic and global design strategy enables the development of a global approach for efficient and sustainable shelter intervention after conflict-based and natural disasters. The "4-dimensional design strategy" defines standardisation rules regulating the set up of open design catalogues. Both design tools promote the design of adaptable and polyvalent construction systems and construction kits. These construction systems allow one to design lightweight shelters easily transportable, mountable, transformable, expandable in size and shape, dismountable and adaptable to local climate and cultures. The efficient use and reuse of construction material thus supports efficient shelter relief and a transition from aid to sustainable development.

References

[1] Henrotay, C., *Sustainable Development. A general description and research of its application towards the material support in emergency situations.* Master's Thesis, Vrije Universiteit Brussel, Brussels, Belgium, 2003.

[2] Hall, P., Pfeiffer, U., *Urban future 21. A global agenda for twenty-first century cities*, E&FN Spon, Londen, United Kingdom, 2000.

[3] Brown V. J., *Give Me Shelter. The Global Housing Crisis*, Environmental Health Perspectives, Volume 111, Number 2, February 2003.

[4] European Commission, *Waste Generated and Treated in Europe*, Office for Official Publication of the European Communities, 2003.

[5] Debacker, W., Henrotay, C., de Wilde, W.P. & Hendrickx, H. 2006, *The Hendrickx-Vanwalleghem design strategy*, Proc. Third International Conference on High Performance Structures and Materials, Wessex Institute of Technology, Ostend, Belgium, 2006.

[6] Hendrickx, H. & Vanwalleghem, H., *Syntectuursysteem*, Patent Application, Vrije Universiteit Brussel, Brussels, Belgium.

[7] Gengnagel, C. & Burford, N., *Transformable Structures for Mobile Shelters*, Proceeding of the joint CIB, Tensinet, IASS International Conference on Adaptability in Design and Construction, Eindhoven University of Technology, Eindhoven, The Netherlands, 2006.

[8] Henrotay, C., Mollaert, M. & Hendrickx, H., *Adaptable Construction Systems for Shelter in Emergency Situations*, Proceeding of the joint CIB, Tensinet, IASS International Conference on Adaptability in Design and Construction, Eindhoven University of Technology, Eindhoven, The Netherlands, 2006.

Climate Adaptive Skins: towards the new energy-efficient façade

B. L. H. Hasselaar
Department of Building Technology, Chair of Building Physics, Delft University of Technology, The Netherlands

Abstract

Façades perform many different functions and are made up of many different parts and materials. Mike Davies was the first to describe a façade made up of one layer that was still able to cater to different functions. In an effort to realise this 'polyvalent façade' as described by Davies, multiple features have since been integrated into the façade, each addressing a specific need. Building skins that are considered to be at the forefront of modern façade technology however are basically all variations on the same theme, hardly doing anything but reacting to the current environmental conditions and the situation created by the façade itself. As such, they are ad hoc devices, able to adjust to a specific circumstance at that specific time, but unable to save, store or prepare for another circumstance. The outdoor climate changes through time and season, alternately supplying energy to and drawing energy from the building skin. Most modern façades have no means to buffer between the two, other than trapping air in a double skin façade to use as a warm blanket in winter. Although some buildings utilise thermal mass in floors/ceilings and aquifers to store heat/cold between day/night and seasons, these are technologies that function independently from the building skin, aiding the building installations. Climate Adaptive Skins (CAS) should differ from 'conventional' façades in a way that they are able to adjust their characteristics to and mediate between the changing environments. By doing so they are able to provide a comfortable indoor temperature, lighting level and air quality (parameters influencing energy consumption) without excess use of energy.
Keywords: façade, adaptive, climate control, comfort, low energy, polyvalent.

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060351

1 Introduction

Historically, the first reason why people sought shelter in caves was protection from the elements. Later, shelters were developed that were more elaborate and able to address people's needs better. Since the industrial revolution this process has taken a huge flight as requirements became more stringent (especially for working/office environments) and technologies for more complicated building skins became available.

The annualised cost for the construction of an (office) building is stated to be between 5 and 10% of the total annual cost for an enterprise over 20 years, while employee cost varies between 75 and 92% of total expenditure [2,3]. Judging from the fact that many people are dissatisfied to some extend with their working environment/climate and that part of that dissatisfaction is building related, a better building performance is likely to reduce health complaints of the users, increasing working efficiency. Paying more attention to the performance of the (new) façade is therefore likely to be profitable.

This paper will describe the first steps in the PhD research towards a new energy-efficient adaptable façade concept. The research will focus mainly on climatic conditions similar to those in The Netherlands, as that is where the research is being conducted.

2 Functions of a façade

To provide optimal comfort to users, it is difficult to cater for every physical aspect as one might contradict another: requirements for acoustic quality differ from those for thermal insulation, which in turn differ from the requirements for daylight admittance. However, there are some basic functions that every building skin needs to encompass, such as being waterproof, providing shelter from wind and thermal insulation as well as being structurally sound. Regarding the office environment, a set of additional requirements of the façade can be formulated, to provide the user with an optimal working environment: thermal comfort, indoor air quality, agreeable lighting level and an acoustic comfortable environment.

3 Current situation

Buildings and climate control systems are normally designed separately rather than integrated, with the result that substantial Heating, Ventilation and Air-Conditioning (HVAC) installations are needed to create an agreeable indoor environment. However, this does not automatically result in the most energy efficient design, nor does it mean that users will perceive the indoor climate as comfortable. The latest insights indicate that occupants who have greater access to controls (e.g. those close to a window) report less discomfort than those who have less access (e.g. away from the window) [4], with current research into Adaptive Temperature Limits indicating that buildings that are (partly) naturally ventilated and have a certain level of user control allow higher indoor temperatures [5].

3.1 Definitions

This research focussed on the development of new concepts for energy efficient façades. As such, it is important to define some of the terms that are used, to avoid confusion. Some of the terms defined however, are used in the same context by people in the building industry. To avoid confusion, such terms relevant to this research are also defined:

Façade: the front or face of a building.

Skin: the enclosure of a space, the separator of interior and exterior. The skin of a building can be foundation, façade or roof, depending on the context.

Comfort: is a term literally meaning 'state of physical ease and freedom from pain or constraint' [6]. The conditions that make a person feel comfortable however are not unambiguous: they are dependent of e.g. psychological, and physiological factors and can differ from person to person.

Adaptive: 'become adjusted to new conditions'. This definition immediately introduces a second, related term: adjusted, which has two useful definitions: 'alter slightly as to achieve the desired result', and 'become used to a new situation'. Both meanings are of importance to this research, although adaptive and adjustable have two different annotations: *adaptive* means the ability to adjust and adapt to changing circumstances by itself. *Adjustable*, similar to adapt*able* means the ability to adjust by external interference, such as by human hand.

Responsive: meaning 'responding readily and positively' with respond meaning in this context: 'do something as a reaction' [6]. With respect to façades, the difference between adaptive and responsive is that responsive does not actually mean the adjustment of specific characteristics to the environment, but merely responding to a change in climate by, for example, lowering blinds or opening windows.

Intelligent: literally meaning the 'ability to vary its state or action in response to varying situations and past experience'. This implies the presence of a computer or a central control centre, since past experiences are used to determine the action to be undertaken next. Adjustment of façade components takes place after a signal from a sensor that registers a change of circumstances. In order to function properly, they need to be connected to an electricity source.

Smart, or 'having a quick intelligence' [6] is often used as a synonym to intelligent, but has a different meaning in the view of the author. A smart material does not necessarily need electronics to adjust its characteristics and applies mainly to materials, not as much to façades or components as a whole. Materials that adjust their characteristics under the direct influence of the environment are smart materials, such as photochromic glass that changes colour when exposed to (bright) light.

Passive, meaning accepting or allowing what happens or what others do, without active response or resistance. Also: containing no source of energy or electromotive force. Applied to a façade, this means that passive components do not react mechanically to a change in circumstances.

Active, as apposed to passive, responds to its environment. The term itself has multiple meanings according [6], but in this context meaning working/operative and/or: capable of automatic change in response to input or feedback. The input comes from a computer or other kind of control centre (also see 'intelligent'), and enables the active device to change its orientation, shape and/or other feature that enables it to influence the interaction between indoor and outdoor climate.

3.2 Features of modern façades

Many of the buildings that are considered to be energy efficient, use the building itself, not the façade, as part of their energy strategy. There are few façades that incorporate passive or smart technologies to aid the building installations in creating a comfortable indoor climate, apart from shading technologies through blinds or louvers and operable windows for ventilation. Widely used techniques are the use of thermal mass of the floors in combination with night flushing as a passive means for cooling, boreholes in combination with heat pumps for either heating or cooling, and atria to preheat ventilation air before it is distributed to occupied zones. Although these are all more or less passive means of climate control, employing natural principles to adjust the indoor climate, they are technologies that operate practically independent of the façade.

3.2.1 View openings

Many façades can be divided into two parts: a transparent part used for daylight and view, and a non-transparent part, used for insulation, privacy and structural purposes. Each part has its own characteristics, although often they are not fully separated and share certain building aspects, such as ventilation.

Case study research [7–9] has shown that almost any modern façade that claims to be 'adaptable', 'intelligent' or 'responsive' can be deduced to one of five different types of solar control devices [10], as demonstrated in figure 1.

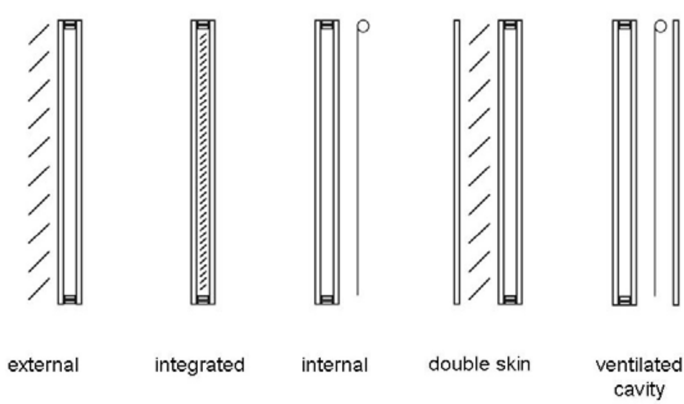


Figure 1: Schematic cross-section of view openings.

External solar control devices reflect and emit the heat absorbed from solar radiation outside the building itself, reducing the need for cooling in summer. The outdoor elements are exposed to the elements resulting in higher costs for maintenance and cleaning. External devices can be fixed (projecting roofs or building sections, awnings, brise-soleils, fixed-angle louver shading) or movable (fabric blinds or screens, slatted or metal louver blinds). Less common are sliding façade units such as panels, screening grids and light-deflecting elements.

Integrated solar control devices are relatively uncommon. Because the devices are contained in an enclosed space, cleaning costs are relatively low. Maintenance costs however can be much higher, especially in cases where the electric motors are also incorporated in the cavity between the panes. An alternative is presented by systems that make use of a magnetic control placed outside the insulating glass.

Internal solar control devices are less effective, as the light passing through glass is transformed into heat by the shading device, which is trapped by the glass and emitted into the room. This effect can to some extend be reduced by drawing off the warm air above these devices. Cleaning and maintaining these devices is considerably easier than exterior and integrated ones, as they can easily be reached. Readily available systems include Venetian blinds and textile materials such as vertical blinds, roller blinds or fabric screens.

Double skin façades have an extra (usually) glass sheet on the outside of a building, posing in front of the actual building façade. The cavity in between the two façades contains some sort of 'intermediate' climate, hence the term 'double skin'. Solar control devices are placed in the cavity between these two skins, which protects them from the influences of weather and air pollution. Heat absorbed by the solar shading is re-radiated and emitted into the intermediate cavity, creating a natural stack effect which causes the air to rise, removing additional heat with the upward flow.

Ventilated cavity façades have an additional single glass sheet on the inside of an interior solar control device. The cavity thus created is mechanically ventilated. Lower pressure in the cavity draws part of the exhaust air from the room into this space where the air warms up, taking most of the heat from the solar control devices, and is then drawn off. The air is extracted on each floor separately, either flowing upwards or downwards in the cavity. For reasons of optimal airflow, horizontal blinds are less suitable. Because temperature differences between the room and the glass surface are small, thermal comfort conditions in the office space near the windows improve.

3.2.2 Combined façades

The classification of façades as described above is mainly concentrated on transparent parts, used for light admittance and view. The non-transparent parts, usually at parapets and ceilings, generally have a very different appearance and function, concealing many of the building services and consisting of very different materials and constructions. As such, these parts can be divided into two variants: cold cavity façades and warm cavity façades [8]:

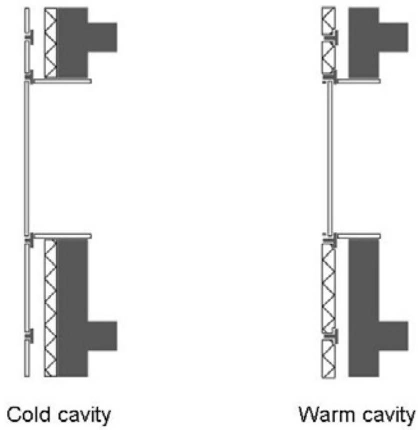


Figure 2: Schematic cross-section of the façade.

Cold cavity façades function as an aesthetic raincoat. The skin gets its strength and rigidity mostly from the underlying construction. The cavity is ventilated with outdoor air; insulation is placed on the outside of the inner leaf. Advantages of this type of facades are the accessibility of the construction and the ease of replacing or altering the façade, giving great flexibility in terms of choice of material for the outer skin. A disadvantage is that the system is relatively unsuitable for building with large façade components.

Warm cavity façades have a cavity separated from outdoor air. The parapet panels contain insulation and are placed at cavity distance from the load bearing structure, making it a complete façade in front of the load bearing façade (see figure 2). This type of façade is suitable for use as component and super component (unit), however, it is also associated with high building costs, limited choice of materials for the non-transparent parts and less accessible construction with inflexibility in case of replacing or altering (parts of) the façade.

3.3 Limitations of current façades

These days, most buildings are insulated to such an extent that, in summer when outdoor temperatures are relatively high, buildings tend to overheat due to the higher sum of internal heat production by sources such as computers, lighting and people themselves and solar gain through windows than, the heat loss through the building skin. Additional insulation is not likely to reduce the building energy consumption but on the contrary causes an increase in energy consumption, as heat inside the building is retained longer; extending the period in which cooling is needed to provide a comfortable indoor climate.

In the Netherlands, the average temperature throughout the year is just under 10°C, well below a comfortable indoor temperature. Even in summer, the average daytime outdoor temperature barely reaches 19 degrees Celsius. Based solely on these numbers one would expect that additional heating is needed throughout the year. However, this does not account for solar radiation and

internal heat production through the use of appliances (e.g. 100 W per computer [11]), or simply the presence of people: one person at rest, standing produces approximately 125 W of metabolic heat [12].

As figure 3 shows, the maximum temperature difference during a 24-hour day is the greatest in summer, averaging up to 11°C, as opposed to about 5°C in winter. This means that the possibility to utilise differences in temperature in the outdoor climate on average is better in summer, which is also when the biggest chance of overheating of the indoor climate occurs. The mean temperature difference between summer and winter is 14,5°C. If it would be possible to store thermal energy for several months, it would be possible to use summer warmth to heat in winter and cold from winter to cool in summer. It is questionable however, whether this is feasible or attainable on façade level.

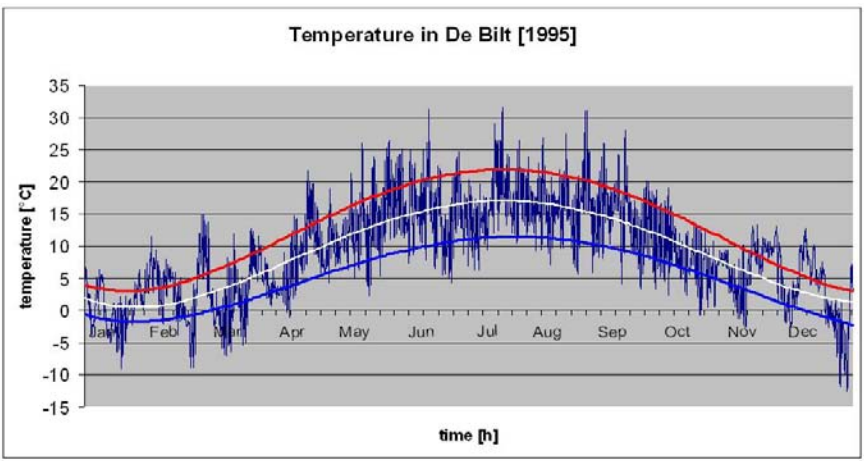


Figure 3: Temperatures during a test reference year.

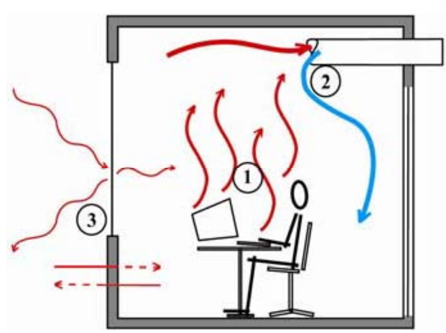


Figure 4: Factors that influence the indoor climate.

If one would be able to isolate the HVAC and the interaction with the outdoor climate, an office environment would function as a net heat producer (see (1) in figure 4): Omitting the influence of the outdoor climate, the building's HVAC

installations are needed to compensate for the internal heat production, taking away excess heat and providing fresh air (2). Basically, these installations produce cool air or water for the indoor environment by drawing heat from it at one location, shedding the heat to the outdoor environment, and transport the cool air or water to another location. In this case, the outdoor environment acts only as a big waste bin for the excess heat. If we would replace the centralised HVAC installations with a small installation for every office room, the devices would be located at each room's façade to be able to shed unwanted heat to the outside. The situation would be slightly more complicated, but essentially the same if we included the influence the outdoor climate has on the indoor climate (3). This influence sometimes calls for additional heating or cooling, or additional lighting. Depending on the climate, the weather and time of year and/or day, the outdoor climate can be used as a source, or should be rejected because it puts too much load onto the indoor environment.

Current façades and HVAC installations are mostly separate entities, influencing each other, but rarely integrated. If a façade would be able to mediate between the indoor and outdoor climate, regulating the exchange of fresh air and heat/cold between the two, and providing additional heating/cooling when desired, the need for additional HVAC installations would be abolished (at least, for the zone within reach of the façade). Although there are some examples of buildings where HVAC installations are located near or in the façade, they are not truly integrated. The integration of HVAC installations within the façade could stimulate integration of different components catering different functions, thus effectively reducing energy consumption needed for the acclimatisation of the indoor environment. By utilising the changes in climate through time and season, the required energy for climate control could potentially be reduced even further.

The principles behind the working of a façade can be explained using terms as described in paragraph 3.1, indicating the amount and type of technology used. The increasing availability of computing power and information through sensors and even the internet makes that many modern buildings utilise intricate, centralised building management systems (BMS) to be optimally able to respond to the outdoor climate. The downside to these technologies is that if (part of) the building management system malfunctions, potentially the whole building climate system is no longer operational, leading to uncomfortable and/or energy consuming situations. Passive technologies are unlikely to fail because they do not actively perform an action but merely influence their environment by their physical characteristics. For that reason however, their application is somewhat limited.

4 Outlook

The intention of this research is to develop principles and concepts for a new type of façades that utilises mostly passive technologies but is still able to provide a comfortable indoor climate throughout the year. By integrating climate controlling elements into the façade, the need for additional, centralised building

services is reduced, reducing the overall consumption of energy for HVAC installations. Because the façade operates using mostly passive technologies, it is less dependent of input (being electricity, thermal energy or data) from the building it is attached to. The independent operation of the façade makes application to both new and renovated buildings easier. The absence of a centralised control centre and many controlling parts reduces maintenance and increases the robustness of the system. Its independence of any external input own enables continuing functionality, even in case of power failure.

An example of the Climate Adaptive Skin (CAS) can be seen in figure 5: an outer sheet of thermotropic glass (1) admits or rejects sunlight, depending on temperature. A valve (2) determines whether ventilation air is pre-heated in the cavity behind the outer sheet (left), or drawn directly from the outside (right), after which the air is brought to temperature by Phase Change Material (PCM) plates (4). A fan (3) pressurises the cavity between (5) and (7), after which the conditioned air is distributed through the micro-porous inner leaf (5). A manually operable window can provide additional ventilation. Exhaust air is drawn from the space above the suspended ceiling.

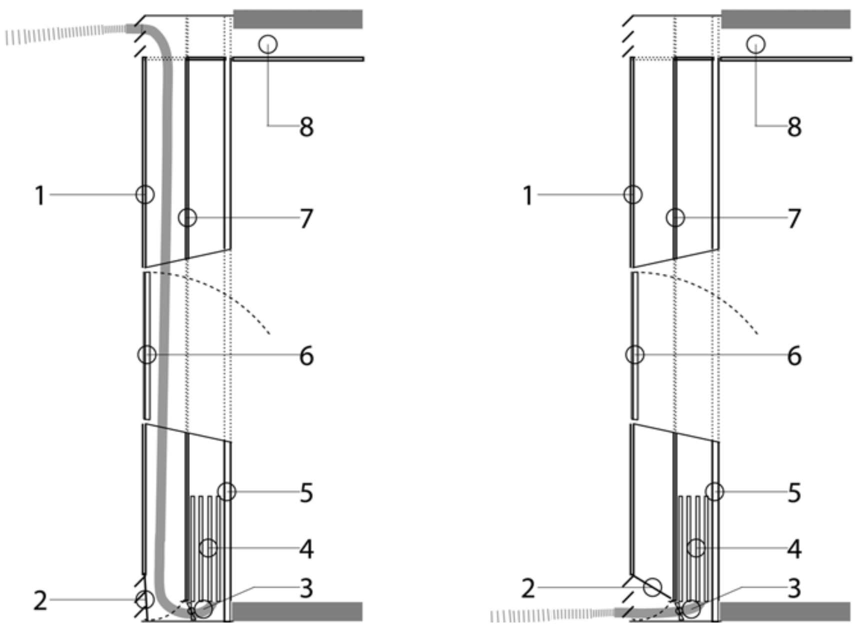


Figure 5: Schematic cross-section of a CAS.

The research it is expected to provide new insights with regard to the functions and functioning of a façade. The design stage, in combination with the production of (a) prototype(s), will deliver a designer's manual which will aid the designer/architect to utilise a new quality in the façade: the ability to react with the changes of the climate, in order to create a comfortable, while energy efficient, indoor environment.

References

[1] Davies, M., A Wall for all seasons + Uses for glass in building, *Riba Journal-Royal Institute of British Architects,* 88 pp. 55-57, 1981

[2] Ree, H.J.v. and Hartjes, Al, Huisvesting in dienst van de medewerker, *BOSS Magazine*, pp. 4-7, 2003

[3] Winch, G.M., Why should we revalue construction, *Revaluing construction 2005: the challenge of change in construction*, Rotterdam, 2005.

[4] Raja, L.A., Nicol, J.F., McCartney, K.J. and Humphreys, M.A., Thermal comfort: use of controls in naturally ventilated buildings, *Energy and Buildings*, 33 pp. 235-244, 2001

[5] Linden, A.C. van der, Boerstra, A.C., Raue, A.K., Kurvers, S.R. and Dear, R.J. de, Adaptive temperature limits: a new guideline in The Netherlands a new approach for the assessment of building performance with respect to thermal indoor climate, *Energy and Buildings*, 38 pp. 8-17, 2006

[6] Oxford Dictionary, Oxford, 2006

[7] Oesterle, E., *Double-skin facades integrated planning,* Prestel, München, 2001

[8] Renckens, J., *Gevels en architectuur; facades in glas en aluminium,* Vereniging Metalen Ramen en Gevelbranche, Nieuwegein, 1996

[9] Wigginton, M. and Harris, J., *Intelligent skins,* Butterworth-Heinemann, Oxford, 2002

[10] Compagno, A., *Intelligente Glasfassaden; Material, Anwendung, Gestaltung,* Artemis, Zürich, 1995

[11] Schalkoort, T.A.J. and Luscuere, P.G. *Klimaatinstallaties; integratie van gebouw en installaties,* TU Delft, Delft, 1996

[12] ISO 8996, *Ergonomics of the thermal environment – Determination of metabolic rate*, 2004

Extant Design: designing things as they are

S. Walker
Faculty of Environmental Design, University of Calgary, Canada

Abstract

A brief overview of two quite different but interrelated ways of addressing sustainability in relation to product design, namely "incremental improvement through design" and "more fundamental considerations of sustainability and product design" is followed by a design exploration that focuses on the second, while also incorporating aspects of the first. The basis of the exploration is described, and supplemented with tangible examples that illustrate the ideas. The focus of the discussion and the design work is on the still-useful objects that are so readily discarded and replaced because they are no longer regarded as useful, or for aesthetic reasons, or because minor technical advancements render these older objects less desirable. The three design explorations included here attempt to offer ways of re-valuing older products, thereby increasing their useful life and, in the process, making some contribution to mitigating the need for and effects of replacement products.
Keywords: sustainability, product design, material culture, novelty, consumerism, moderation.

1 Introduction

A contemporary challenge for design is to re-imagine and re-interpret some of our most fundamental notions of material culture. This is important for two reasons. Firstly, it is critical that we begin addressing sustainable issues in a much more substantial manner than has been evident to date. Secondly, if design is to contribute to human culture in a more meaningful way then it has to move beyond the often shallow, style-based notions of product design that have become so prevalent over the last 50 years.

Here, a discussion of approaches to sustainable design that might be termed 'incremental product improvement' models, are contrasted with approaches that

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060361

take a much more radical approach to material culture. In addition, the latter forms the basis for a design exploration that is described and illustrated.

2 Design and sustainability

The continuous production of new products, with new aesthetic expressions, and accompanied by relentless marketing campaigns, is intended to create a sense of dissatisfaction with one's possessions, leading to the replacement of older products with new, Chapman [1]. Even though the older product may still be perfectly serviceable, it becomes prematurely obsolete because, over time, it grows shabby and outdated in its appearance and because fresh, new styles are made available to us. While the production and consumption of new products satisfies economic objectives, spurs business growth and gives temporary pleasure to the user, there are also a host of well documented problems associated with this conception of material culture that render it fundamentally unsustainable - sustainability, in its basic form, being attendant to the triple bottom-line of environment, ethics and economics, Elkington [2].

Ironically 'sustainable product design', as the name implies, is also often pursued by designing new products that in some manner address sustainable principles – perhaps through the use of more benign materials, or by employing up-to-date, cleaner technologies. In recent years, a host of incremental strategies have been developed to improve the environmental performance and, in some cases, the social ramifications associated with current manufacturing. These include programs such as The Natural Step, Robert [3], Cradle-to-Cradle design, McDonough, W. and M. Braungart [4], Product Life Cycle Assessment, Pré [5], and Factor 10 [6]. All of these offer pragmatic solutions for modifying current practices in order to reduce the negative impacts of product design and production. They are important models for implementing improvements in the existing system, but they do little to challenge the basis of that system. Moreover, these 'incremental product improvement' models result in more objects being manufactured by using more resources and more energy. So this interpretation of sustainable product design, while it may contribute to the improvement of products over time, has some inherent problems associated with it.

These approaches can be seen, in varying degrees, as part of the problem rather than the solution because they bolster, rather than challenge, the current consumerist model of a totally commodified material culture – a model that is manifestly destructive. Modern marketing, of which product design has always been a part, Sparke [7], is based on creating dissatisfaction with what we have, Chapman [8]. It fosters discontentment and tells us, unremittingly, to want more, that we deserve more. It promotes consumerism and desire for the novel, and gives rise to disquiet and yearnings that can never be sated. When we fall victim to this, we can never find happiness; happiness, in many traditions, being synonymous with the notion of human satisfaction, e.g. McCabe [9]. Thus, it could be said that our contemporary market system sells, more than anything else, discontent and unhappiness.

Incremental improvement models do not address this problem in any significant way. They may ameliorate certain negative effects but they do little to find a way out of a production system that is evidently environmentally, socio-culturally and personally damaging.

It is important to acknowledge that there are many new developments in science and technology that do make genuine and valuable contributions. However, the concern in this present discussion is with the plethora of products that are continually restyled, perhaps with the addition of some relatively trivial updates and features to grab our attention. In the context of sustainability, these products represent an immoderate use of resources, they help generate waste and pollution, and, through marketing, they help create dissatisfaction.

3 A changing role for design

If, as I suggested above, the creation of new products is, indeed, part of the problem rather than the solution, then what is left for the designer? Does the designer still have a contribution to make? On the face of it, and within the conventional parameters of product design, it would seem that the answer would be no, or at most, relatively little. However, if we are prepared to broaden our horizons a little, and think of designers not simply as the creators of novel products, but as creative individuals who think about and contribute to the nature, meaning and design of material culture, then they will have an important role to play in re-conceptualizing and redefining our notions of functional objects. Potentially, the creative abilities of the designer can be critical in a world where excessive consumption and waste are leading us down a self-destructive path. In developing this role, design within academia can make a significant and reflective contribution, removed as it is from the pressures of the business environment. As Fuad-Luke [10] has suggested, to develop new roles for contemporary design it is important to separate it, at least temporarily, from economic considerations.

If designers are to address the challenge of sustainability in a more substantive manner, they must question the ways in which they design, the assumptions they make, and the products of design. Some academics and designers are already doing this – by developing new approaches and strategies, and new understandings of the contribution and role of the designer. There are a number of directions that can be seen as much more fundamental explorations of our understanding of material culture, compared to the incremental approaches discussed earlier. For example, Chapman [1] has looked at the psychology of consumerism and considered ways of creating more meaningful relationships with objects; Fuad-Luke [10] has explored the notion of 'slow design', which focuses on personal, socio-cultural and environmental well-being; and Manzini and Jégou, [11] have developed strategies for 'enabling solutions' at the local level, which allow people to become more involved and self-determining in the development of their products and services. Compared to the incremental improvement models, these approaches are more demanding, more difficult to

envision, and less comfortable because they challenge our contemporary conventions and our basic assumptions about the nature of products.

Integral to many of these approaches is the use of design as a method for illustrating alternative notions of functional objects. Designers and design-centred academics can attempt to re-imagine the nature of objects through the practice of design itself. Drawing on the plethora of material that is available on the relationship between material culture, consumerism, and environmental and social effects, designers can explore new design methods and develop alternative approaches to the definition of 'functional object'. There are a number of examples of such designers, including the Droog group based in the Netherlands, Ramakers (ed.) [12], the Campana brothers in Brazil [13], and the Boyms in the United States [14]. All these, in their own distinctive ways, challenge current norms of mass-produced products – conceptually, aesthetically and philosophically.

4 Extant design

The exploration in this paper describes and illustrates, through examples, a direction that leans more to these alternative approaches. In other words, a consideration of our relationship with objects is followed by an engagement in the process of design in order to translate these considerations into tangible, functional objects.

More specifically, it explores an alternative to the novelty-based, voguish approaches to design that are so dominant today, and which tend to maintain design as an arm of advertising to spur consumerism, Woodham [15]. This influential aspect of the design profession may contribute to the continuous growth model of modern economics but, in terms of sustainability, it creates a host of problems. And, as was mentioned earlier, the avid consumerism stimulated by corporate agendas does little to enhance our sense of fulfilment or personal happiness, DeAngelis [16].

Another way of thinking about design, and the products of design, can be based on *acceptance of what is*, rather than continuing the convention of designing the new, the novel and short-lived. The first of the three r's associated with environmental responsibility is *reduce*, which can be applied to consumerism as much as to the use of materials and energy. A reduction in consumerism is logically related to the second of the three r's, i.e. *re-use*. If we reduce consumerism, then we have to re-use what we already have. Reducing and re-using are both more important and should be given higher priority than the third 'r' of *recycling*, yet it is recycling that has received far more attention. This is because is can be more easily accommodated within our current design and production system. Reduction and re-use require a more drastic re-assessment of the nature and norms of contemporary material culture.

The design examples included here also take into consideration a variety of other sustainable principles such as product longevity, localization, and self-determination, and they acknowledge the social and environmental costs of

manufacturing industry, European Commission [17]; Scott [18]. Clearly, these various elements of sustainable design are interlinked and they raise a number of questions about the nature and meaning of design, the place of design, and our understanding of what it means to be a designer in a world straining under the pressures of consumerism.

The objects presented here are examples of ‘academic design’, as distinct from commercial design. They represent an exploration from within the discipline of design itself and are an attempt to articulate ideas about objects. The result is a number of artefacts, each of which includes products that had been discarded. In all cases, new technologies have been added, to enable the still-functional aspects of older products to find renewed usefulness – thus the designs integrate old objects with new, and link contemporary mass-production with the benefits of reduction and re-use. Furthermore, the artefacts are created in ways that are suited to the diversity and restrictions of ‘locale’. Thus, they deal with issues of aesthetic obsolescence, moderation, waste, localization and self-reliance.

These design explorations raise questions about our understandings of products in the context of sustainability. Design is used as a means to re-present familiar and/or unwanted products so they can be re-used and seen anew. This contributes to the pressing debate about sustainability and manufactured objects through an exploration of material culture that stands outside current norms of newness and novelty. The intention is to find a place for design that is more meaningful, more lasting and which can contribute to a culture of *moderation* and *acceptance*.

5 Explorations in Extant Design

The approach taken in this present series of works will, hopefully, begin to illustrate aspects of the changing role of product designers who wish to address sustainability. The focus here is on what we might term ‘the lost products of contemporary society’. These are unwanted objects and older products that, for aesthetic reasons, are no longer regarded as fashionable, but which are not yet old enough to have value as retro or antique objects. These are the products that are so readily cast aside and replaced.

Using the discarded, finding a place for the rejected, and reusing the broken are ways of treasuring, honouring and respecting not only the resources of the earth, but also the time, thought and ingenuity that have already gone into these objects. Exploring how to use them again, beyond their first use, is a way of further justifying and valuing their presence in the world. It may seem unfashionable to say so, but it is a form of appreciation and a way of respecting, even sanctifying, the world, its people and its resources. More than this, re-presenting older, existing products in a new light can allow us to accept them *as they are*, and despite their unfashionable styling and their wear and tear, be content with them, to see them as *enough*.

There are admirable models for these types of explorations, from the ready-mades of Duchamp in the early years of the twentieth century, Ades *et al* [19], to

the more recent work of Rauschenberg in which found objects are incorporated into sculptural pieces. Rauschenberg's work in particular, while not addressing utility, nevertheless salvages and resurrects the discarded, the abused and the lowly objects that are so unconscionably abandoned in today's affluent societies Rose [20].

The first artefact in this present set of design exercises, 'Soundbox' (figure 1), was conceived as a more conventional 'product'. It combines a new MP3 player with recovered computer loudspeakers with an integral amplifier. These objects have been assembled in a very rudimentary manner, by placing them in a re-used cardboard box. While aesthetically crude, the exercise was an attempt to re-use products in a way that could be readily achieved at the local level without special knowledge or tools.

Figure 1: Soundbox. A new personal MP3 player combined with re-used, unfashionable computer speakers with integral amplifier – assembled in a rudimentary manner in a cardboard box.

While this design achieves this objective, it was recognized that, aesthetically, a different approach was needed. In the later explorations, 'Lamplight' (figure 2) and 'Replay' (figure 3) the re-used objects have been separated from their surroundings and used as components within larger compositions.

In these cases, the objects are mounted on specially built, neutral armatures to create 'functional arrangements'. In a sense, this places them on a pedestal, and re-contextualises them so that they may be seen from a somewhat different perspective. The approach is one in which the individual, re-used objects become mere *elements* within a greater whole. Thus, the emphasis is shifted - the focus is no longer on an out-of-date, imperfect old product but, instead, is on the entire composition. In addition, the re-used objects are no longer hidden, as they were in 'Soundbox', nor are they disguised to hide their previous use or their

unfashionable styling. Within the larger compositions, they 'work' as they are and for what they are. Thus, they attain a renewed dignity and a new value by becoming integral components of a contemporary composition.

Figure 2: Lamplight. Re-used bottles, new lampshades, low-energy bulbs on a white armature with shelf.

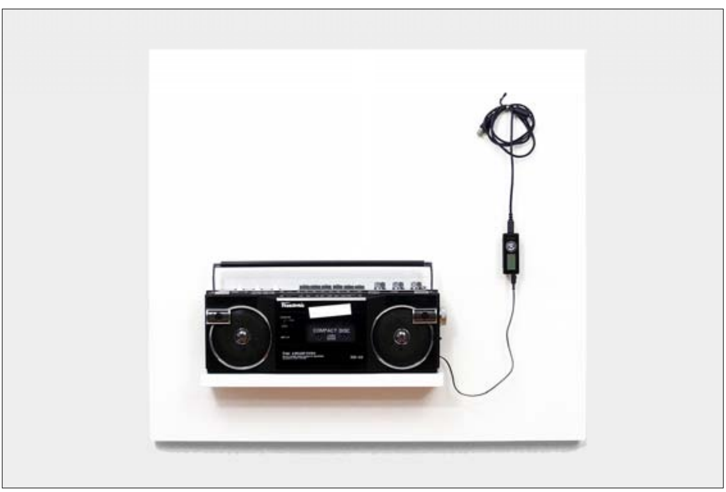


Figure 3: Replay. New personal MP3 player, re-used, 1980s cassette stereo on a white armature with shelf.

These explorations are intended to contribute to the development of a new role for design – a role that is more attuned to the needs of people, society and the environment. These examples are not presented as an end point, but simply

as small steps along a road of rediscovering a new, perhaps more valid and more profound role for design that stands outside the destructive system of unfettered production and waste that is so prevalent today. Moreover, this approach implies a rather different design process than is common today in industrial design; the process here being more akin to collage or assemblage, rather than the *ex nihilo* methods of contemporary practice.

It is pertinent to point out that this approach does not ignore technological development or innovation, instead it attempts to use it where it seems appropriate to do so. For example, in ‘Replay (figure 3), an old stereo cassette player has been slightly modified to allow the use of MP3 technology. In this way the benefits of the MP3, which dispenses with the need for cassette or compact disc production, with their attendant packaging and distribution, are combined with the existing and still serviceable amplifier and speakers of the older product. This is a more moderate solution than discarding the old product and replacing it with a new speaker system designed to go with an MP3 player. The older product is adapted, re-used and, through re-presentation, re-valued.

This example differs in intention from the explorations carried out by Droog designers Wolf, Bader and Oschatz. In their “Bootleg Objects” series, classic designs from the past were updated to use MP3 technology. The chosen products included Dieter Rams’ 1962 Braun Audio, a 1973 Bang & Olufsen system, and the Technics SP1210 record deck from 1980, Williams [21]. All these products are now recognised for their design excellence and therefore already have value; these products were chosen precisely because they are regarded as classics. Hence, they are not typical of the kind of products that end up in landfill. They have passed the transitional period that renders products undesirable and have reached an age and a rarity that makes them, once again, sought after items.

By contrast, the products selected for use in the examples included here have no such design cachet. This is best illustrated in the cassette player that forms part of the MP3 Speaker composition, ‘Replay’ (figure 3). The cassette player was selected precisely because it was no longer wanted, it was no longer regarded as aesthetically pleasing or fashionable, and it had no ‘retro-chic’. The functional composition in which it features is an attempt to:

- Accept objects as they are – with their faded, unfashionable aesthetic and their scars of use.
- Value objects as they are – for the benefits they still offer.
- Respect objects as they are because in doing so we show respect for the thought, ingenuity, time and effort that have gone into their design. We also show respect for the resources and energy that have already been used in their production.
- Moderate our acquisitiveness – by re-representing and revaluing what we already have, rather than simply discarding and replacing older objects with new products.
- Slow the culture of distraction that acquisitiveness engenders. A distracted culture, *ipso facto*, cannot be self-reflective and this, in turn, tends to reinforce the present model of consumerism.

Rather than describing these compositions in more detail, it is perhaps better to allow them to speak for themselves. The language of design is not words, it is not arguments and criteria. It is simply design itself. Words can take us some of the way but the artefacts themselves have to do the rest.

> "Looking is a marvellous thing of which we know but little. Through it we are turned absolutely towards the outside, but when we are most of all so, things happen in us that have waited longingly to be observed; and while they reach completion in us… their significance grows up in the object outside." Rainer Maria Rilke, 1907 [22]

References

[1] Chapman, J., Emotionally Durable Design – Objects, Experience, Empathy, Earthscan, London, p.16, 2005.

[2] Elkington, J., Cannibals with Forks: The Triple Bottom Line of 21st Century Business, New Society Publishers, Gabriola Island, BC, Canada, 1998.

[3] Robert, K. H., The Natural Step Story – Seeding a Quiet Revolution, New Society Publishers, Gabriola Island, BC, Canada, 2002.

[4] McDonough, W. & Braungart, M., Cradle to Cradle – Remaking the Way We Make Things, Douglas & McIntyre, Vancouver, BC, Canada, 2001.

[5] Pré, life cycle tools to improve environmental performance, Pré Consultants, Netherlands, http://www.pre.nl/, accessed November 28 2005.

[6] Factor 10 – an introduction, Factor 10 Institute, Canoules, France, http://www.factor10-institute.org/, accessed November 28 2005.

[7] Sparke, P., An Introduction to Design and Culture in the Twentieth Century, Allen & Unwin, London, pp. 94-98, 1986.

[8] Chapman, J., Emotionally Durable Design, Earthscan, London, p.39, 2005.

[9] McCabe, H., The Good Life, Continuum, London, 2005.

[10] Fuad-Luke, A., SLOW – Slow theory – a paradigm for living sustainably? 2004, accessed at: www.slowdesign.org, accessed November 23rd 2005.

[11] Manzini, E. & Jégou, F., Sustainable Everyday – Scenarios of Everyday Life, Edizioni Ambiente, Milan, 2003, and Manzini, E. & Jégou, F., Album – A Catalogue of Promising Solutions, Edizioni Ambiente, Milan, 2003.

[12] Ramakers, R., (ed.), Simply Droog – 10+1 years of creating innovation and discussion, Droog Publishing, Amsterdam, 2004.

[13] Campana, F. & H., Interview with Fernando + Humberto Campana, 2005, accessed at: http://www.designboom.com/eng/interview/campana.html, accessed November 23 2005.

[14] Boym, C & L. L., Curious Boym – Design Works, Princeton Architectural Press, New York, 2002.

[15] Woodham, J. M., Twentieth-Century Design, Oxford University Press, Oxford, pp. 145, 227-228, 1997.
[16] DeAngelis, T., Consumerism and its discontents, Monitor on Psychology, American Psychological Association, Vol. 35, No. 6 June 2004, http://www.apa.org/monitor/jun04/discontents.html, accessed Nov. 25 2005.
[17] European Commission (2004) EU policy-making - counting the hidden costs, European Commission Environment Research document, August 16 2004, accessed at: http://europa.eu.int/comm/research/environment/newsanddoc/article_1444_en.htm, November 25 2005.
[18] Scott, R. (2003) The high price of 'free' trade, Economic Policy Institute, Washington D.C., EPI Briefing Paper #147, November 17, 2003, accessed at: http://www.epinet.org/content.cfm/briefingpapers_bp147, Nov. 25 2005.
[19] Ades, D., N. Cox & Hopkins, D., Marcel Duchamp, ‘World of Art’ series, Thames & Hudson, London, p. 146, 1999.
[20] Rose, B., Rauschenberg – On and Off the Wall, in “Rauschenberg – On and Off the Wall – Works from the 80’s and 90’s”, Musée d’Art moderne et d’Art contemporain, Nice, France, pp. 47-73, 2005.
[21] Williams. G., Use it again, in Ramakers, R., (ed.), Simply Droog, Droog Publ., Amsterdam, pp. 25-34, 2004.
[22] Rilke, R. M., Looking is a Marvellous Thing – from Rainer Maria Rilke’s letters to his wife Clara, on looking at Cézanne’s pictures in the Salon D’Automne, Paris, October 1907, Hastings Arts Press, Hastings, UK, 1992.

The role of beauty for sustainability: a discussion on responsible consumption, aesthetics attitudes and product design

A. Marchand[1], S. Walker[1] & P. De Coninck[2]
[1]*Faculty of Environmental Design, University of Calgary, Canada*
[2]*École de Design Industriel, Université de Montréal, Canada*

Abstract

This paper explores the relationship between sustainability, and product aesthetic qualities and experience. It presents the result of an ongoing study notably looking at responsible consumption and the resulting visual culture. Initial insights are presented and thoughts that have emerged from this research experience are discussed. The link between eco-ethics, aesthetics, and the intrinsic and extrinsic properties of objects is briefly examined. The paper explores an 'aesthetic attitude' related to the ethical, spiritual and pragmatic facets of sustainability, and proposes the valorisation and celebration of beauty in these terms – as opposed to the surface "aestheticization of everyday life" that we are witnessing in contemporary society.
Keywords: product design, material culture, responsible consumers, ethics, aesthetics, aestheticization, beauty.

1 Introduction

Sustainability has fascinating implications for the aesthetic qualities of built environments, including the functional products that surround us, and for the way we experience them. Indeed, it is expected to affect how products look and are thought about, and to modify our response to a product's visual qualities. The paper provides some examples to illustrate these effects, and starts to explore how the project of sustainability challenges aesthetic conventions and social norms. Ongoing research into the significance of sustainability for product design, visual cultures and products aesthetics experience, is described. This study examines how some individuals promoting sustainability through their

www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060371

consumption habits, and more globally though their lifestyles, experience product aesthetics. Main aesthetic qualities that are meaningful to these identified responsible consumers, and that are derived from eco-ethics concerns are presented. This research leads us to reflections on sustainability, responsible consumption, and the dialogue between ethics and aesthetics. In this regard, the crucial role beauty could play in moving towards a more environmentally and socially viable world is discussed. The paper proposes to valorise and celebrate a type of beauty situated at the meeting of 'being good' and 'looking good'.

2 Sustainable product design

In the field of product development and manufacturing, sustainability involves reconsidering the way objects are thought about, developed, produced, distributed, used, reused, recycled, and disposed. Sustainable production and consumption mean ameliorating and modifying products and, more fundamentally, imagining new ways of living, and new cultural and social models. It is a demanding issue for product design and manufacturing because it is not simply about things, but about how we relate to things. In this context, product designers are invited to propose solutions that are, together with many other characteristics, more resource efficient, socially acceptable, and economically viable. Designers need to be capable of redesigning traditional products in these terms. More than this, they are asked to envision creative solutions where outcomes of traditional given consumption goods are achieved in a different and more sustainable way. This can notably be realized by responding to a particular need, the need for mobility for instance, through product-services-systems (Manzini and Vezzoli [1]). For design disciplines, sustainability has important implications in many areas, including the sphere of aesthetics. In terms of aesthetics, it implies changes in, firstly, the way the material world surrounding us looks and, secondly, in the way we aesthetically experience it.

3 Aesthetics experience and product aesthetics

Discussions about *aesthetics* may be confusing and inconsistent because, in common usage, the term refers to two different concepts (Crilly *et al.* [2]). Firstly, in the context of *product aesthetics*, 'aesthetics' or 'aesthetic qualities' usually relates to what the product presents to the senses. Secondly, in the context of *aesthetics experience*, the term refers to one particular aspect of cognitive response; the perception of how pleasing (or otherwise) is the process of regarding the product. The former is more directly concerned with the specific *intrinsic* qualities of the object, while the latter refers to the experience of these properties, an experience which may also be informed by the object's *extrinsic* properties. Extrinsic properties are those which are external to the artefact and which may notably be the result of various assumptions and *a priori* knowledge. These may be associated with characteristics such as the purpose of the object and its meaning, as well as the type of materials, form, colours, type of

assembly, and, if known, various aspects of the production process and intentions behind the consumer product. For example, in recent years there have been various press revelations about the use of sweatshop labour in developing countries to produce products for Western markets. This extrinsic property of a particular object can, if known by potential purchasers or users, affect their aesthetic experience of it as a thing.

4 Aesthetics and sustainability

As mentioned earlier, sustainability in relationship to aesthetics involves two things: first, the modification of aesthetic qualities of the objects surrounding us, and secondly, changes in the representations of meaning attached to aesthetic qualities of our environment, including objects. It challenges many aesthetics conventions and social norms. For instance, from a sustainable perspective, our attachment to and use of the ubiquitous lawn in residential areas is problematic. Insistence on having a lawn is both highly wasteful and polluting when one considers the excessive amount of drinking water and energy consumed in maintaining one, as well as the common use of pesticides and fertilizers. Replacing lawn area with landscapes composed of native vegetations adapted to local climate would represent a more ecologically sound solution. Native vegetations do not require intensive care and maintenance as a lawn does, and constitute a far more amicable environment for insects and birds. Another example demonstrating how sustainability can challenge aesthetics and social conventions is the wearing of formal business apparel in the typical business setting. The air conditioning of office space has significant negative ecological impacts. Together with more efficient ventilation and air conditioning systems, the establishment of dress codes that are better suited to the climate and the season would be ecologically appropriate. At another level more directly related to the idea of changing the representation of meaning embodied in aesthetic qualities, the example of a varnished apple can be taken. The glossy apple is a widespread representation of the 'ideal' apple. However, in moving towards sustainability, the varnishing of apples could become associated with less environmentally sound practices, consequently modifying the concept of the 'ideal' apple.

5 Ongoing research into material cultures and responsible consumption

In an attempt to explore the potential implications of sustainability for product design, visual cultures and the way we experience product aesthetics, a study among identified responsible consumers has been undertaken. This study examines how some individuals promoting sustainability through their consumption habits, and more largely though their lifestyle choices, experience product aesthetics. Among other things, the research investigates the aesthetic qualities that are meaningful to these people. The study takes note of how ethics inform their aesthetical judgement, and how aesthetic qualities are used as landmarks to evaluate the ethical acceptability of products.

More globally, the aim of this research is to gain insights into what might be termed a "responsible, sustainable material culture" by seeking to understand the choices and perspectives of self-identified responsible consumers with regard to functional products. As understood in this project, material culture refers to the user-product relationship and includes notions such as values, preferences and consumption choices, perception of meaning through object qualities, and aesthetic experience. The underlying premise of material culture is that man-made objects reflect the beliefs of the individuals who made, commissioned, purchased, or used them and, by extension, the beliefs of a particular community or society to which they belong or belonged (St-George [3]). Material culture studies examine aspects of the world of objects, including images, that people make, use and experience. According to Buchli [4], material culture studies refer to the study of the material world, but more exactly, to the study of the materiality of cultural life. Visual culture is here considered to be a part of material culture, more especially in how meaning is created and the relation between beliefs and construction of meaning through visual elements.

Eighteen individual interviews, each of one to three hours duration depending if respondents wanted to meet once of twice, have been conducted among eleven participants. Participants are people who consider themselves responsible consumers and who advocate, in various interpretations and at different levels, the notion of voluntary simplicity. During the semi-directed individual interviews, visual material and objects were used to stimulate the discussions. Images of various products and objects were presented to participants to open up and encourage the discussion of ideas and preconceptions attached to given visual qualities. In addition, visual data were generated by participants to further facilitate the exchanges. Prior to the interviews, respondents were invited to take photographs of objects that are aesthetically pleasing to them. Meetings of existing simple living groups where people primarily discuss the issue of responsible consumption were attended in order to gain further information about products in the context of responsible consumption. Twenty-one meetings were attended, totalling more than 40 hours of participant observation.

Briefly, voluntary simplicity (or simple living), which might be considered as a movement, is an attempt to achieve a better quality of life by minimizing the detrimental impacts of our ways of living on both the human and natural environments (Marchand and Walker [5]). Although this 'philosophy' is far from being a panacea for contemporary consumerism, many authors [6–10] see it as a direct response to consumer culture, and as a movement that holds lessons for a more viable, sustainable society.

Responsible consumption is here considered to be a process through which people make voluntary efforts to both reduce the amount of goods they consume and opt for more environmentally and socially sound solutions. The model proposed by Hansen and Schrader [11] is used here to define responsible, sustainable consumption:

1. *Abstention from consumption*; satisfying certain needs without the purchase and use, and in some cases even questioning the very need for, a given product;

2. *Reducing consumption*;
3. *Using ecologically and/or socially sound product variants*, i.e., choosing product variants that are or are perceived to have greater environmental and social integrity; and
4. *Using ecologically and/or socially sound substitutes*, i.e., switching to products and/or services that are or are perceived to be consistent with sustainable practices

6 Perception of meaning through Objects Visual Qualities

In conducting the study among identified responsible consumers, the considerable influence of ethical concerns on the aesthetic experience of products has been frequently observed. The information these citizens retain about the ecological or ethical integrity of a particular product, as well as the suppositions they make in this regard, was found to have a significant impact on their aesthetic appreciation. In terms of assumptions, aesthetic qualities are used as a basis for evaluating the 'goodness' of the object; to form judgements about the acceptability of the product and the world that lies behind it. Certain aesthetic qualities or typologies are associated with accepted and desirable intentions and larger practices, others with environmentally and socially harmful ones.

For example, their critique of novelty for its own sake and its detrimental effects motivates a certain preference for objects presenting prototypical characteristics when acquiring new and highly mass produced products. Prototypicality here refers to the degree to which an object is representative of a category. Prototypes are usually defined as "[…] central representation of a category as possessing the average or modal value of attributes of that category" (Veryzer and Hutchison [12]). However, differences and diversity of expression are greatly valued as long as they are perceived as being meaningful, of aesthetic value and in-line with sustainable concerns – and not simply as novelty for its own sake or as an abusive marketing ploy.

Also, as they often feel that many contemporary consumer goods are over-equipped, something which they find illustrative of a certain culture of excess, they shared a clear preference for objects with limited features. These are perceived as being longer-lasting as well as easier to maintain and repair. They also proved to be highly sensitive to the quantity of material that composes a product and tend to prefer objects minimizing the use of material.

Respondents highly appreciate products that they can easily understand in terms of constitution and function. Indeed, they esteem intelligible objects where, for example, the mechanism is apparent and/or accessible to the user – in opposition to products where the functional elements are in some way hidden. Being able to understand the functioning of an object gives them the feeling that they have control over the object, and that they can repair it if broken, or at least understand the malfunction. They further report enjoying using objects that

allow them to feel that they are physically involved in achieving the task to be accomplished. For example, we have found that they value objects such as a manual peppermill or a manual coffee percolator.

Finally, participants valued natural and raw materials over composite materials which, as they pointed out, might or might not be recyclable. Preferences for such materials and finishes seem to be in line with a study showing that, in a different but related area, responsible consumers prefer food with strong and authentic tastes, such as dark chocolate over mild chocolate, and non-processed foods (Lauer [13]).

7 Restoring the dialogue between ethics and aesthetics

In *Undoing Aesthetics*, Welsch [14] discusses the relationship between aesthetics and ethics, and describes the term 'aesthetics' as being derived from the Greek term *aisthesis*. Aisthesis has a double meaning: *perception* and *sensation*. On the one hand, as *perception*, it "[…] addresses genuine sensuous qualities such as colours, sounds, tastes, and smells." On the other hand, as *sensation*, it follows an emotive perspective. The former may be understood as being linked to a more 'objective' response while the latter belongs to the more 'subjective' domain involved in the aesthetics judgement and notably comprises the space where the ethical reflection may lie. Welsch severely deplores the fact that perceiving has become autonomous whereas Eaton believes in "[…] a kind of merit in which being good and appearing (looking, sounding, tasting) good are integrated" (Eaton [15]). Welsch condemns the rupture between *perception* and *sensation*, both in academic discourse and day-to-day life, and calls for their reconciliation as a *vital imperative*.

The relationship between these two dimensions is important because their reconciliation allows us to envision a world where aesthetic experience and beauty are synonymous with, if not the source of, justice and vice-versa. Responsible consumption truly seems to constitute a privileged space for restoring this dialogue as it encourages an aesthetic reflection, or an aesthetic attitude, that is informed by environmental and socio-economic responsibilities. While it is clear that correct or truthful information about an object can affect the aesthetic experience of it, it is not clear to what extent an object's appearance may reveal this information [16]. There are obvious limits to how much an object may actually reveal about itself, but to some degree, objects still embody the intentions and purposes of their makers (Buchanan [17]). To a great extent, the essence of a product's aesthetic is a function of the system that produced it, and it follows that the aesthetics of most consumer products are the aesthetics of immoderate waste, and environmentally and socially damaging practices (Walker [18, 19]).

8 Valuing and celebrating beauty against the aestheticization of everyday life

According to Welsch, we are currently experiencing an "aesthetics boom" where "[m]ore and more elements of reality are being aesthetically mantled." Known

as the *aestheticization of everyday life*, this phenomenon "[…] refers to the rapid flow of signs and images which saturate the fabric of everyday life in contemporary society" (Featherstone [20]). As a process of surface aestheticization or styling of the mundane, it lies in "[…] the embellishment and sensualization of everyday objects, environment and experiences" (Hancock [21]). It should be noted that, in this context 'aestheticization' refers to 'styling'.

Outside purely academic circles, the issue of *aestheticization* is also present and discussed. For instance, in a recent article published in *The New York Times Magazine* about product design and business strategy, the following question was asked: "Must even the most mundane household object rise up and join the tyranny of 'good' design?" (Walker [22]). Also, in *Intramuros* [23], a magazine dedicated to designers and design practice, the *aestheticization of everyday life* phenomenon was characterized and discussed under the terms "generalized invasion of aesthetic value" and "propaganda." The main points expressed were that "the democratization of beauty has become an ongoing concern" in which "everything is supposed to 'look good'" and where "everything and everybody has become beautiful." Although the phenomenon is applauded by some for "[…] offering both immense pleasure and rich resources for the construction of identity" (Duncun [24]), its economic purposes and foundations cannot be ignored (Welsch [14]). In turn, one could argue that it serves the interests of multi- and transnational corporations which tend, at the global scale, to reduce diversity of expressions through an unvarying 'corporate aesthetic' or 'style' notwithstanding cultures, people and places.

With regard to what has been described as the "tyranny of 'good' design" and the "generalized invasion of aesthetic value," some critiques can be formulated. In the first place, considering the lucrative aspect of these surface types of 'beauty' concerned with outward appearance and styling, it could be argued that the excessive "democratization of aesthetics" has resulted in an understanding of the concept of beauty as an object of consumption. Supporting a loss of meaning and encouraging today's saturated consumer culture, the excessive quest for 'beauty' that we are actually witnessing may actually undermine our relationship with objects, other people, and the bio-physical world. In fact, as Chabot [25] has said, "[t]he problem is not that beauty is too valorised, the problem rather lies in the growth of impoverished forms of beauty." Postrel [26] has suggested that sensory appeals are becoming ever more prominent in our culture and this is why, according to her, "[t]o maintain a healthy balance between substance and surface, we can no longer simply pretend that surfaces don't matter." Surface really does matter and has legitimate value. The idea here is not to reject our sensory nature, but to look for restoring the dialogue between surface and substance.

However, we believe that the valorisation and celebration of beauty – in the sense of being good and appearing good (looking, sounding, tasting) – has an unsuspected power against superficial styling and its detrimental effects. In fact, an aesthetic experience informed by eco-ethical concerns will not find pleasure through superficial, excessively commercial expressions, but rather only through genuine diversity, creativity and fairness.

9 Conclusions

This paper has explored some aspects of the significance of sustainability for product aesthetic qualities and product aesthetic experiences. The distinction between these two concepts has been made. In the context of product aesthetics, *aesthetic qualities* refers more directly to what the product presents to the senses, while *aesthetic experience* concerns the cognitive response to those properties. This paper suggests that, in moving toward a more sustainable world, both our visual cultures and object visual qualities will be different. In this connection, to explore potential effects of sustainability on visual cultures, the initial results of an ongoing study on responsible consumption and material cultures has been presented. The study establishes the link between eco-ethics concerns and aesthetics appreciation. In addition, the initial results begin to outline some aesthetic qualities that are valorised and appreciated among those people who, through their consumption choices, and more largely in their daily life, work to incorporate some sustainability principles. Based on various assumptions, some typologies are often associated with more environmentally and socially acceptable practices. The study found that the interviewed responsible consumers show a strong preference for objects presenting prototypal characteristics when it is question of new and highly mass produced products, and for natural and raw material over composites. The study also noted a preference for functional products presenting limited features with minimal use of material. Respondents value objects that can be easily understood in terms of their constitution and operation, as well as products that allow them to be physically involved in realising the task they wish to accomplish. Based on some thoughts that have emerged from this research experience, the paper discussed the importance of restoring the dialogue between ethics and aesthetics. To this end, responsible consumption was considered, with some limits, as a privileged space to that allows for this, as it invites aesthetical reflections informed by eco-ethical concerns. Finally, the paper proposed to celebrate and valorise beauty against the surface aestheticization of everyday life, and for sustainability in and of itself.

References

[1] Manzini, E., Vezzoli, C., *Product-Service Systems and Sustainability: Opportunities for Sustainable Solutions*, UNEP: Paris, 2002.

[2] Crilly, N., Moultrie J., Clarkson, J.P., Seeing Things: Consumer Response to the Visual Domain in Product Design, *Design Studies*, 25 (6), pp. 547-577, 2004.

[3] St-George, R.B., *Material Life in America: 1600-1860*, Northeastern UP: Boston, 1988.

[4] Buchli, V., Introduction, *The Material Culture Reader*, ed. V. Buchli, Berg: Oxford and New York, pp. 1-26, 2002.

[5] Marchand, A., Walker, S., Designing Alternatives, Sustainable Consumption, Lifestyles and ‘Responsible Consumers’, *Proc. of the*

SCORE Workshop: Changes to Sustainable Consumption and Production, eds. M. Andersen & A. Tukker, Copenhagen, 2006. Online. [http://www.score-network.org/score]

[6] Shaw, D., Newholm, T., Voluntary Simplicity and the Ethics of Consumption, *Psychology and Marketing*, 19 (2), pp. 167-185, 2002.

[7] Reisch, L.A., Scherhorn, G., Sustainable Consumption, *The Current State of Economic Science*, 2, pp. 657-690, 1999.

[8] De Graaf, J., Wann, D., Naylor, T.H., *Affluenza: The All-Consuming Epidemic*, Berrett-Koehler Publishers: San Francisco, 2001.

[9] Maniates, M., *Confronting Consumption*, eds. T. Princen, M. Maniates, K. Conca, The MIT Press: Cambridge (MA) and London, 2002.

[10] Etzioni, A., Voluntary Simplicity: Psychological Implications, Societal Consequences, *Voluntary Simplicity: Responding to Consumer Culture*, eds. D. Doherty & A. Etzioni, Rowman and Littlefield Publishers: Lanham. pp. 1-25, 2003.

[11] Hansen, U., Schrader, U., A Modern Model of Consumption for a Sustainable Society, *Journal of Consumer Policy*, 20, pp. 443–488, 1997.

[12] Veryzer, R. W., Hutchison, J.W., The Influence of Unity and Prototypicality on Aesthetics Responses to New Product Design, *Journal of Consumer Research*, 24, pp. 374-394, 1998.

[13] Lauer, S., Les 'Alterconsommateurs' Contre-Attaquent: Les Français se Rebellent Contre la Consommation de Masse, *Le Devoir*, June 20th, A-1, 2004.

[14] Welsch, W., *Undoing Aesthetics*, Sage Publications: London, p. 62, 1997.

[15] Eaton, M., *Merit, Aesthetic and Ethical*, Oxford University Press: New York, p. 1, 2001.

[16] Elliott, J., Reconciling Eco-Ethics & Aesthetics in Design, *Design Philosophy Papers*, 2, 2004. Online. [http://www.desphilosophy.com/dpp/home.html]

[17] Buchanan, R., Rhetoric, Humanism, and Design, *Discovering Design*, eds. R. Buchanan & V. Margolin, The University of Chicago Press: Chicago, pp. 23-66, 1995.

[18] Walker, S., Conscientious Objects, *Eternally Yours: Visions on Product Endurance*, ed. E. van Hinte, 010 Publishers: Rotterdam, pp. 164–180, 1997.

[19] Walker, S., Unmasking the Object, *Proc. of the European Academy of Design Conference*, Bremen, 2005. Online. [http://ead06.hfk-bremen.de/conference/tickets.html].

[20] Featherstone, M., *Consumer Culture and Postmodernism*, Sage Publications: London, p. 67, 1991.

[21] Hancock, P. (2002). Aestheticizing the World of Organization: Creating Beautiful Untrue Things, *Tamara: Journal of Critical Postmodern Organization Science*, 1 (1), pp. 91-105, 2002.

[22] Walker, R., The Way We Live Now, *The New York Times Magazine*, February, p. 42, 2004.

[23] Intramuros, Edito, *Intramuros*, 106, p. 13, 2003.

[24] Duncun, P., A Case for an Art Education of Everyday Aesthetic Experience, *Studies in Art Education*, 40 (4), pp. 298-311, p. 295, 1999.
[25] Chabot, M., Honorer le beau, *Revue Notre-Dame*, 100 (6), p. 16-28, p. 26, 2002.
[26] Postrel, V., *The Substance of Style: How the Rise of Aesthetics Value is Remaking Commerce, Culture and Consciousness*, Perennial: New York, 2003.

Structure evolution of spider silk liquid crystalline precursor material

G. De Luca & A. D. Rey
Department of Chemical Engineering,
McGill University, Canada

Abstract

Spiders produce silk fibers with remarkable mechanical properties using an ultra-optimized spinning process. The fluid precursor material used to draw the silk threads is a lyotropic nematic liquid crystal. The mechanical properties of the silk fibers as well as their processability are strongly affected by the complex structural transitions undergone by the nematic liquid crystal precursor along the spinning line. Our work focuses on the particular structure adopted by the nematic precursor in the extrusion duct of the spinning apparatus. This structure is characterized by a succession of well defined point defects located on the axis of the cavity and interacting on each other through elastic mediated forces. The phenomenon described is both important in understanding the process-induced structuring of spider silk fibers and to defect physics.
Keywords: Spider's silk, liquid crystalline spinning, nematic point defects.

1 Introduction

Spiders' ultra-optimized and ecological spinning process produces a fiber with mechanical properties comparable or even superior to the best man-made super-fibers, which use corrosive solvents and cause significant environmental degradation [1, 2]. Hence, there is a great deal of interest in understanding the intricacy of their design and processing routes. Many environment friendly exciting applications are envisaged upon the successful replication of spiders' fibers and complex spinning line [3–5].

Spider silk fibers are known to be spun from a lyotropic nematic liquid crystal emerging from a highly concentrated water-based solution of rod-like molecules or aggregates [6–8]. A quintessential property of this silk precursor is its capacity

doi:10.2495/RAV060381

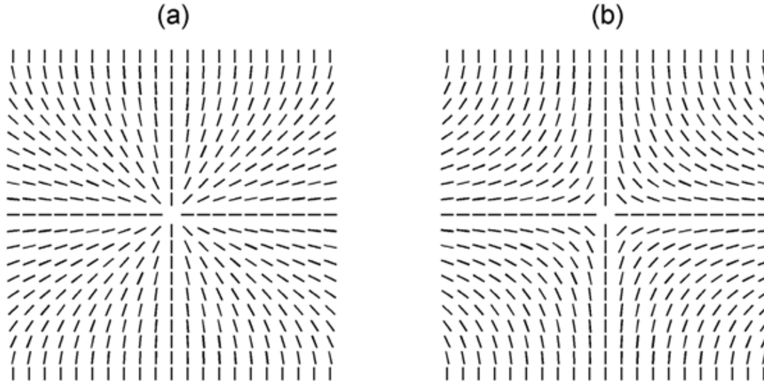


Figure 1: Structure of the radial (a) and hyperbolic (b) point defects found along the spinning duct of spiders in terms of the director field $\mathbf{n}(\mathbf{r})$.

to maintain some degrees of orientational order while still flowing as a liquid. This orientational order is characterized by the tendency that have neighboring rod-like units to align their long axis in parallel along a common direction [9, 10]. This preferred molecular orientation usually varies from subregion to subregion in the mesophase (*i.e.*, intermediate phase) due to elastic effects coupled with geometrical and interfacial constraints [9]. The evolution of orientational order or molecular orientation along the spinning pathway is crucial as it affects the processability of the silk precursor mesophase and determines the microstructural details of the solidified fiber and therefore its remarkable physical properties [6–8].

Spiders' spinning line basically consist in a tail where the silk precursor material is synthesized, a central bag where it is stocked in a very concentrated solution, and an extrusion duct from which the silk fiber is drawn [8, 11]. Observations made by polarized light microscopy in the extrusion duct have revealed the presence of a complex orientation structure known as escaped radial with point defects (ERPD) [8, 12–14]. The point defects, referred as hedgehogs or monopole, are located where the direction of bending distortions changes. At those particular locations, orientational order *melts*. Two types of point defects are alternatively found in the cylindrical cavity of the extrusion duct: the radial and the hyperbolic hedgehogs. Figure 1 shows the characteristic structures of the radial and hyperbolic hedgehogs in term of a director field $\mathbf{n}(\mathbf{r})$ giving the local average preferred orientation of the molecules. The unit directors are drawn arrowhead-free as there is no physical difference between the vector field $\mathbf{n}(\mathbf{r})$ and $-\mathbf{n}(\mathbf{r})$ [9, 10].

Whether this complex molecular structure is an accident of Nature or a necessary ingredient of the spider biospinning process is unknown at present. Nonetheless, one my hypothesized that this configuration with its orientational defects may play an important role in the control of material crystallization along with water pumping, ions exchanges and pH reduction phenomena [8]. Indeed, a premature

crystallization of the silk may indeed cause the permanent blockage of the extrusion system and ultimately lead to the death of the animal [8, 15].

Nematic point defects confined in cylindrical geometries have been first experimentally observed and explained in the early seventies [16–19]. They are typically observed when a nematic mesophase is confined in cylindrical capillary with lateral walls enforcing strong radial anchoring (*i.e.* molecules are forced to orient radially at the surface). Point defects with opposite topological charges are known to annihilate by pairs. This has been experimentally observed [16–21] and theoretically described [20, 22–26]. Results have shown that when two defects are separated by less than a tube diameter they usually attract until they eventually annihilate. As the two defects come closer their speed increases exponentially. At large separating distance the situation is far less clear as some studies support the hypothesis of a total screening of the attraction force [23–25] while others support a repulsion force [22, 27]. Recent experiments have also shown a possible speed anisotropy between the point defects [28] but the role played by elastic anisotropy and back-flow in this phenomenon has not been clearly established yet.

Obviously these phenomena become even further involving when considering not two point defects but rather a whole a array of them. Some studies have in fact *touched* this problem in a statistical manner but no experimental data have yet corroborated them [29]. During their evolution arrays often splits into sub-arrays of few interacting defects with alternating signs. This work aims at describing what can possibly happen inside those sub-arrays. These results should be useful in improving the understanding of arrays of nematic point defects and therefore of their behavior along the spinning duct of spiders, and hence contributing to the on-going efforts to develop systematic technology transfer from Nature to fiber engineering.

2 Modeling

In this section we briefly present the necessary theoretical background to study the dynamics of nematic point defects that is essential to silk bio-spinning.

The continuum dynamic equation describing the structure evolution of a nematic liquid crystal is typically derived from the minimization of a free energy functional depending on some orientational order parameter that characterizes molecular order and macroscopic texture [9]. In the simplest continuum approach, the orientational order of parameter is a unit vector $\mathbf{n}(\mathbf{r})$, called director, giving the average preferred orientation of the molecules at a point $\mathbf{r}$. The energy cost associated with the distortions of the director field is then given by the Frank distortion energy [9]:

$$f_b = \frac{K_1}{2}(\nabla \cdot \mathbf{n})^2 + \frac{K_2}{2}(\mathbf{n} \cdot \nabla \times \mathbf{n})^2 + \frac{K_3}{2}(\mathbf{n} \times \nabla \times \mathbf{n})^2, \tag{1}$$

where K_1, K_2, K_3 are elastic constants for the three modes of orientational distortions occurring in nematic: splay, twist and bend. It is useful and moreover appropriate to adopt the so-called one constant approximation: $K = K_1 = K_2 = K_3$.

Within this approximation, no speed anisotropy can be attributed to elastic effects. The Frank free energy simplifies to:

$$f_b = \frac{K}{2}\nabla\mathbf{n} : (\nabla\mathbf{n})^T. \tag{2}$$

This vectorial approach is generally well suited to study small and continuous deformations of the nematic liquid crystal. However, this approach generally fails in the vicinity of defects where the director field may be discontinuous causing, in turn, the distortion free energy to become infinite. However, this problem can be overcome in a rather straightforward manner by allowing the director to deviate from its unit length constraint and act as an additional order parameter measuring the degree of molecular alignment along itself. In this work we employed a regularized Frank elastic free energy of the form [30, 31]:

$$f_b = K\left[\frac{1}{2}\nabla\mathbf{n} : (\nabla\mathbf{n})^T + \frac{(|\mathbf{n}^2-1)|^2}{4\delta^2}\right], \tag{3}$$

where δ is a penalty parameter related to the size of the defect core. The second term on the right-hand side of Eq. 3 is the penalty function that allows the director to deviate from unity in orientational defects and the distortion energy to be bounded.

The time dependent equation for the rotation of the director is determined by the balance between a viscous and an elastic torque. The latter, which is usually refereed as molecular field, is given by the variational derivative of the Frank elastic free energy. The transient director equation is then:

$$\gamma\frac{\partial\mathbf{n}}{\partial t} = K\left[-\nabla\cdot\nabla\mathbf{n} + \frac{(\mathbf{n}^2-1)\mathbf{n}}{\delta^2}\right], \tag{4}$$

where γ is a constant associated with the rotational viscosity of the director.

In order to reduce the number of parameters and facilitate the analysis of the results we non-dimensionalize the governing equation by introducing the characteristic time and length scales of the problem. Lengths are measure in terms of the capillary radius and therefore: $\bar{\mathbf{r}} = \mathbf{r}/R$. The time scale is determined by the typical relaxation time of the director field and is given by: $\bar{\mathbf{t}} = \mathbf{t}/\tau$ with $\tau = \frac{\gamma R^2}{K}$.

Given that the solutions to our problem have an obvious rotational symmetry around the axis of the cylindrical cavity we consider a two dimensional computational space representing half of a longitudinal cross section and we accordingly employ cylindrical coordinates (r, z). The width L and height R of the computational domain are set to 5 and 1, respectively. Finally, the boundary conditions on the outer wall and end caps are respectively strong radial anchoring and no flux.

3 Results

This section gives some representative results on the collective behavior of subarrays of nematic point defects lying on the axis of a cylindrical capillary and

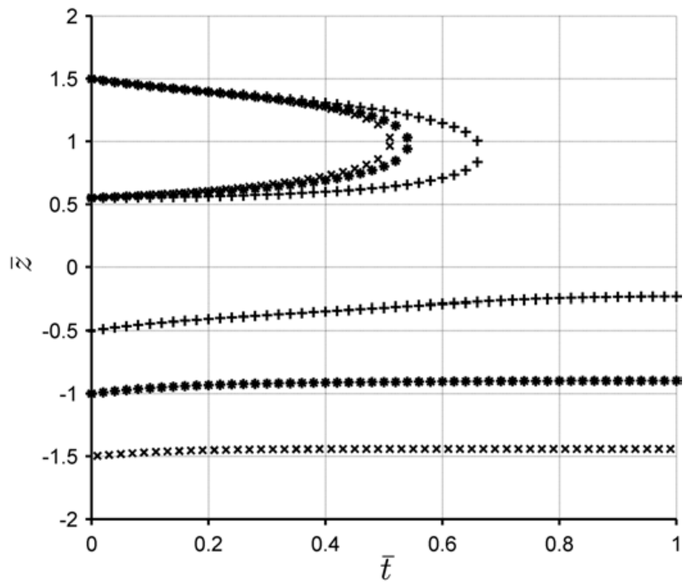


Figure 2: Position of the point defects along the cavity axis as a function of time for three different cases ($d_p = 1.05, 1.55, 2.05$)

interacting between each other. We consider scenarios involving three and then four nematic point defects.

3.1 Interactions between three point defects

In this case we examine the influence of a point defect on the interaction between a neighboring pair. Note that the global charge of the system does not alter the results presented. As we mentioned earlier, when two defects of opposite charges are sufficiently close one another (*i.e.*, when their separating distance is smaller or equal to the tube diameter) they usually attract and finally annihilate living no trace of their previous existence. Furthermore when no back-flow or anisotropic elastic effects are considered, the two defects travel at the same speed and therefore meet at the midpoint between their initial positions. This behavior can be however significantly affected by the presence of an additional defect interacting with the pair.

Figure 2 shows the trajectories of three nematic point defects as a function of time for three different scenarios. In all three cases, two defects, forming a pair, were initially held at the same separating distance $d_i = 0.95$ while a third *perturbing* defect was placed at different distances $d_p = 1.05, 1.55, 2.05$ away from that pair. It can be seen from fig. 2 that as d_p increases, the effect of the *perturbing* defect becomes weaker. For the case $d_p = 2.05$ (the separation tends to the screening length), the pair is annihilating practically at midpoint and roughly

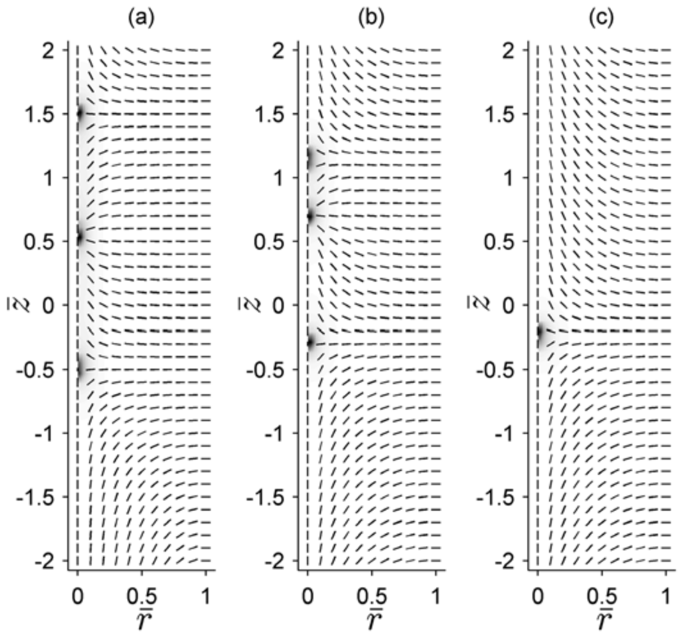


Figure 3: Evolution of structure in the cylindrical capillary as the point defects move along the axis. Segments indicate the orientation of directors while the gray scalar gives the degree of alignment along the directors. Black=no alignment=defect core, white=alignment. $\bar{t} = 0.01(a), 0.6(b)$ and $1(c)$.

unaware of its presence. The reverse is also obviously true for the *perturbing* defect which is just weakly initially attracted by the pair. When $d_p = 1.05$ (the distances between the three defects are comparable), the trajectories become distorted as the system tries to globally reduce the distance separating the defects, including d_p. One can see that the pair does not annihilate at midpoint anymore indicating speed anisotropy between the two defects. Furthermore, the time required for the pair annihilation is found to increase as d_p decreases. Figure 3 illustrates the dynamic structural changes occurring in the case $d_p = 1.05$. In this figure, the small segments give the director field while the gray scale provides and indication of its length thereby providing an alignment scalar order parameter.

3.2 Interactions between four point defects

We now turn to cases where all defects can potentially annihilate by pair and disappear from the system. The system has now two pairs and therefore an additional important length to take into account. We denote by d_j this distance separating the second pair of defects. As for the previous cases of three defects, one pair is held at

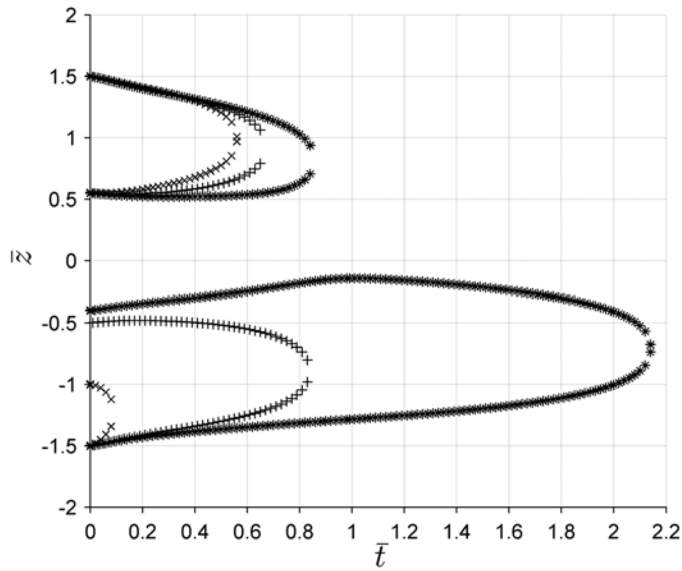


Figure 4: Position of the point defects along the cavity axis as a function of time for three different cases.

constant initial separation. One of the defects of the second pair (the exterior one) is also held at the same initial position while the remaining defect (inner one) is moved at different initial positions thus varying the lengths d_j and d_p at the same time.

Figure 4 shows the trajectories of the four defects as a function of time for different initial positions of one defect. It can be seen that when d_p is large and approaching the screening distance, the two pairs annihilate unaware of each other. The pair whose initial interdefect separating distance is the smallest annihilates the fastest. In the second case $d_i \approx d_j \approx d_p$ and the two pairs annihilate in the same time frame. The trajectories of the defects in each pair are asymmetric as the two exterior defects are traveling faster than their inner counterpart. It is also important to note that the two inner defects do not collapse together despite being separated by the same distance with respect to the exterior defects. This provides evidence that each defect is affected by all the remaining ones and the attraction felt is proportional to the separating distances. In the third and last case, $d_p < d_j$ and the inner defect is more attracted by the other inner defect rather by the closet exterior defect so that three defects tend to go in the same direction. Unfortunately for the *rebelling* defect, it is still slowed down enough by the attraction of the closet exterior defect so that it cannot reach the other pair on time and it is forced to change direction once the first pair has collapsed. This change of direction considerably lengthens the annihilation process of the second pair. As observed from the case of three defects the annihilation time scale increases as d_p decreases. Figure 5 illustrates the dynamic structural changes occurring in the case $d_i \approx d_j \approx d_p$.

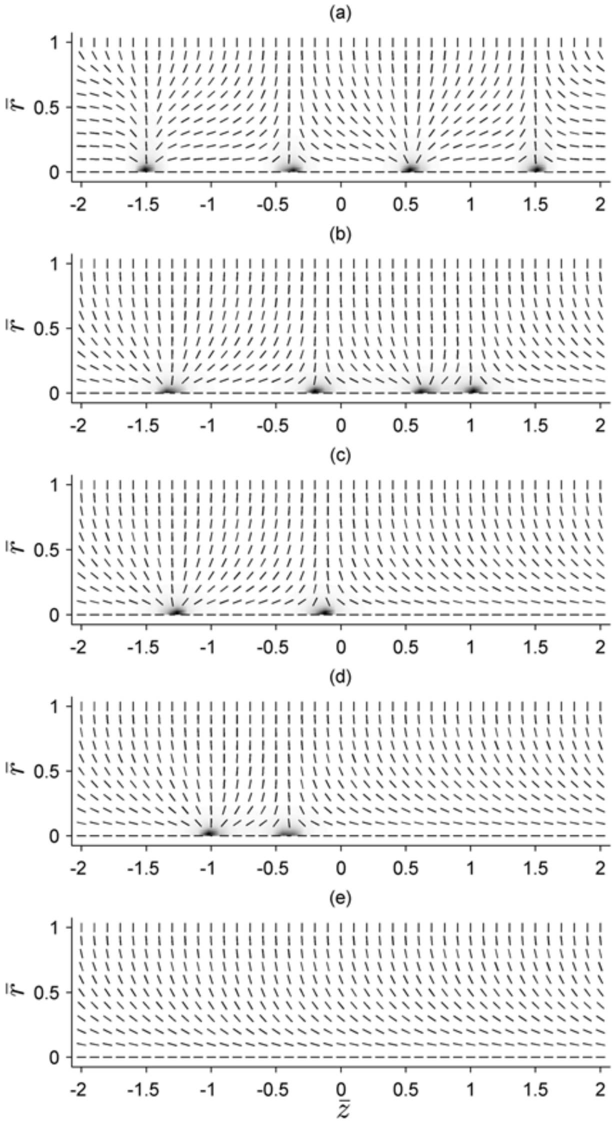


Figure 5: Evolution of structure in the cylindrical capillary as the point defects move along the axis. Black=no alignment=defect core, white=alignment. $\bar{t} = 0.01(a), 0.81(b), 0.99(c), 2(d)$ and $2.2(e)$.

4 Conclusions

We have presented a simple model to study the interactions between nematic point defects lying on the axis of a cylindrical capillary. This work was motivated by

the reported experimental observation of those types of liquid crystal defects and related structures along the spinning extrusion duct of spiders. The cases of sub-arrays of three and four mutually interacting defects were presented. Despite the absence of back-flow and elastic anisotropy effects, strong anisotropy were put in evidence due to the sole effect of collective interactions. As for the much studied case of two interacting point defects, the screening distance after which the defects were unaware of each other and pinned was found to be around one diameter. A phenomenon of direction change that cannot be observed when considering only two defects was also featured. The different simulations have shown that defects are always attracted by their closet complementary neighbor but still affected by all the other defects below the screening distance. The results presented here should be useful in both improving the understanding of point defects in the context of the process-induced structuring of spider silk and to the more universal physics of defects. This work is supported by the Natural Science and Engineering Research Council of Canada (NSERC). G.D. wishes to acknowledge financial support from NSERC through the CGS program.

References

[1] Gosline, J., DeMont, M. & Denny, M., The structure and properties of spider silk. *Endeavour*, **10(1)**, pp. 37–43, 1986.
[2] Ko, F. & Jovicic, J., Modeling of mechanical properties and structural design of spider web. *Biomacromolecules*, **5**, pp. 780–785, 2004.
[3] Atkins, E., Silk's secrets. *Nature*, **424**, p. 1010, 2003.
[4] Jin, H.J. & Kaplan, D., Mechanism of silk processing in insects and spiders. *Nature*, **424**, pp. 1057–1061, 2003.
[5] Turner, J. & Karatzas, C., *Natural Fibers, Plastics and Composites*, Kluwer Academic Publishers, chapter 1, Advanced spider silk fibers by biomimicry, pp. 11–23, 2004.
[6] Kerkam, K., Viney, C., Kaplan, D. & Lombardi, S., Liquid crystallinity of natural silk secretions. *Nature*, **349**, pp. 596–598, 1991.
[7] Viney, C., Huber, A., Dunaway, D., Kerkam, K. & Case, S., *Silk Polymers: Materials Science and Biotechnology*, American Chemical Society, chapter 11, Optical Characterization of Silk Secretions and Fibers, pp. 120–136, 1993.
[8] Vollrath, F. & Knight, D., Liquid crystalline spinning of spider silk. *Nature*, **410**, pp. 541–548, 2001.
[9] de Gennes, P. & Prost, J., *The Physics of Liquid Crystals*. Oxford University Press, 1995.
[10] Collings, P., *Liquid Crystals: Nature's Delicate Phase of Matter*. Princeton University Press, 2001.
[11] Foelix, R., *Biology of spiders*. Oxford University Press, 1996.
[12] Knight, D. & Vollrath, F., Liquid crystals and flow elongation in a spider's silk production line. *Proc R Soc Lond B*, **266**, pp. 519–523, 1999.

[13] Vollrath, F., Strength and structure of spiders' silks. *Reviews in Molecular Biotechnology*, **74**, pp. 67–83, 2000.
[14] Lydon, J., Silk: the original liquid crystalline polymer. *Liquid Crystals Today*, **13(3)**, pp. 1–13, 2004.
[15] Wainwright, S., Biggs, W., Currey, J. & Gosline, J., *Mechanical Design in Organisms*. Princeton University Press, 1982.
[16] Williams, C., Pieranski, P. & Cladis, P.E., Nonsingular s=+1 screw disclination lines in nematics. *Physical Review Letters*, **29(2)**, p. 90, 1972.
[17] Cladis, P. & Kleman, M., Non-singular disclinations of strength s = + 1 in nematics. *Journal de Physique*, **33(5-6)**, p. 591, 1972.
[18] Williams, C., Cladis, P. & Kleman, M., Screw disclinations in nematic samples with cylindrical symmetry. *Molecular Crystals And Liquid Crystals*, **21(3-4)**, pp. 355–373, 1973.
[19] Melzer, D. & Nabarro, F., Cols and noeuds in a nematic liquid crystal with a homeotropic cylindrical boundary. *Philosophical Magazine*, **35(4)**, pp. 907–915, 1977.
[20] Pargellis, A., Turok, N. & Yurke, B., Monopole-antimonopole annihilation in a nematic liquid crystal. *Physical Review Letters*, **67(12)**, pp. 1570–1573, 1991.
[21] Peroli, G., Hillig, G., Saupe, A. & Virga, E., Orientational capillary pressure on a nematic point defect. *Physical Review E*, **58(3)**, pp. 3259–3263, 1998.
[22] Vilfan, I., Vilfan, M. & Zumer, S., Defects structures of nematic liquid crystals in cylindrical cavities. *Physical Review A*, **43(12)**, pp. 6875–6880, 1991.
[23] Peroli, G. & Virga, E., Annihilation of point defects in nematic liquid crystals. *Physical Review E*, **54(5)**, pp. 5235–5241, 1996.
[24] Semenov, A.N., Interaction of point defects in a nematic liquid. *Europhysics Letters*, **46(5)**, pp. 631–636, 1999.
[25] Gartland, E., Sonnet, A. & Virga, E., Elastic forces on nematic point defects. *Continuum Mechanics And Thermodynamics*, **14(3)**, pp. 307–319, 2002.
[26] Bajc, J., Peroli, G., Virga, E. & Zumer, S., Dynamics of nematic point defects in a capillary with tilted boundary conditions. *Liquid Crystals*, **29(2)**, pp. 213–219, 2002.
[27] Holyst, R. & Oswald, P., Annihilation of point defects on a line. *Physical Review E*, **65**, p. 041711, 2002.
[28] Cladis, P. & Brand, H., Hedgehog-antihedgehog pair annihilation to a static soliton. *Physica A-Statistical Mechanics And Its Applications*, **326(3-4)**, pp. 322–332, 2003.
[29] Biscari, P., Peroli, G. & Virga, E., A statistical study for evolving arrays of nematic point defects. *Liquid Crystals*, **26(12)**, pp. 1825–1832, 1999.
[30] Liu, C. & Walkington, N., Approximation of liquid crystal flows. *SIAM J Numer Anal*, **37**, pp. 725–741, 2000.
[31] Yue, P., Feng, J., Liu, C. & J., S., A diffuse-interface method for simulating two-phase flows of complex fluids. *J Fluid Mech*, **515**, pp. 293–317, 2004.

Section 6
Safety

This page intentionally left blank

Risk management of hazardous material transportation

J. Augutis, E. Uspuras & V. Matuzas
Lithuanian Energy Institute, Lithuania

Abstract

In recent years on Lithuanian roads a considerable amount of hazardous materials have been transported, especially oil products. However, there is no common methodology which could assess the risk of such transportation. Lithuania has accepted directives, norms and other acts of law related to risk and hazard assessment and prevention that are valid in the European Union. In relation to this situation, hazard assessment and analysis has even greater significance. The novelty of the work is associated with employing the Markov process to describe a hazard distribution mechanism and to determine a limited hazard distribution in the nodes of networks.
Keywords: Hazard, risk, risk sources, risk distribution, Markov process.

1 Introduction

Hazard identification and assessment are rather complicated tasks, which have received attention in the literature (Adams [1]).

Risk of any activity or process is often defined as a set made of n pairs of frequency of hazardous events and their outcomes. Sometimes these values are multiplied. Hazard measurement is less clearly defined and here such qualitative evaluations as high hazard level, medium hazard level, low hazard level, etc. are used. In certain cases, quantitative expressions are also used.

During transportation of hazardous materials or at the outspread of communicable diseases, etc. hazard is divided, moved from one place to the other, distributed among various structures. In the literature, much attention is devoted to the investigation of various hazard distribution mechanisms. One of the most widely explored among these is the distribution of different pollution

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060391

materials in the atmosphere, water and soil, spread of communicable diseases among people, animals and plants (Lefevre and Picard [2]) and shipping hazardous materials with different means of transportation (Purdy [3]). Latter works provide detailed analyses of distribution mechanisms, speed, process duration, etc. of hazardous materials and diseases.

It is also obvious that together with the improvement of means of transportation and the increase in the quantity and size of the loads, the assessment of hazard distribution becomes more prominent in the systems of transportation. It has to be noted that the majority of the scientific research articles and works on the hazard distribution assessment in the network systems has been made during the last several decades and this topic is still under active investigation.

The main aim of the paper is the analysis of hazard distribution in the network systems. In the context of the analysis, hazard is understood as the amount of hazardous materials, disease concentration, etc. Hazard can be distributed through the channels of various networks and concentrated in the nodes of the networks. Hazard transmission through the channels that connect network nodes can take place in many ways: for example, hazard can be transmitted to a single or to several nodes, as an undivided value or divided into parts. Each node can also have certain protection or immunity against hazard, which blocks its transmission or diminishes it.

The paper has a purpose to present the analysis and mathematical model of hazard distribution in the nodes of the network.

2 Definitions of hazards, transfers and other concepts

As it was already mentioned, hazard in this paper is equalled to such numerical values as the quantity of hazardous materials, the intensity of informational transfer, etc. Hazard will be noted as *H*. We will now define several terms that will be used in the paper:

Hazard source. It is one of the network nodes in which hazard can arise or occur.
Point source of hazard. It is a source of hazard in which hazard occurs only once.
Infinitive source of hazard. It is a source of hazard in which hazard arises periodically, for an infinitive number of times.
Additive hazard. It is a sort of hazard, when hazards in the nodes of the network can be added to or a part of hazard moved to the other nodes. The examples of the additive hazard are: collection of hazardous materials, transport intensity, etc.
Non-additive hazard. It is a sort of hazard when the sum of hazards is equal to the maximum of those hazards: $H_1 + H_2 = \max\{H_1; H_2\}$.
Transfer intensity coefficient between the network nodes. It is a coefficient $q_{i,j}$ that marks the part of the hazard in the node *i* that will be transmitted to the node *j*. It is clear that $\sum_{j=1}^{N} q_{i,j} \leq 1$,

here N - a number of network nodes. The transfer intensity to the node j is $\breve{q}_j = \sum_{\substack{i=1 \\ i \neq j}}^{N} q_{i,j}$ and from the node j is $\widehat{q}_j = \sum_{\substack{k=1 \\ k \neq j}}^{N} q_{j,k}$.

Network. It is system defined as an oriented graph, in which hazard from one node can be transmitted only to one node during a cycle.

Network node immunity. It is a coefficient I_j that marks which part of the hazard is transmitted to the node j $\left(0 \leq I_j \leq 1\right)$. Node immunity can be created by the security systems, ant-virus computer software, etc.

Transfer probability. It is a probability $p_{i,j}$, that during one cycle hazard from the node i will be transmitted to the node j. Transfer probability has the following feature: $\sum_{j=1}^{n} p_{i,j} = 1$.

Hazard transfer cycle. Hazard transfer in the network from one node to the other is regarded as one hazard transfer cycle.

Hazard in network nodes after *k* cycles. Hazard that is accumulated in the node i after k cycles, will be marked as $H_i\left(k\right)$.

Marginal hazard in the network nodes. Marginal hazard in the i node is a steady hazard after an infinitive number of cycles. $H_i = \lim_{n \to \infty} H_i\left(n\right)$.

3 Additive hazard distribution in the network

The distribution of hazard that can be divided or added in the network nodes will be analysed. Two hazard distribution methods will be analysed separately.

In the first case it will be assumed that hazard can be transferred from every node only to one of the possible nodes, while in the second case, let us allow the hazard spreading though the entire network.

3.1 Hazard distribution in Markov chains

Let us suppose that we have a network with N nodes. Hazard from the node i can be transferred only to one node j, which is selected according to transfer probability P_{ij}. Thus, during each cycle, hazard can occur in only one network node. In the paper an assumption will be made that transfer probabilities have Markov properties. Thus, if the hazard that exists in node i after n cycles will be marked as $X(n)$, so

$$\begin{aligned} P_{ij} &= P\left(X\left(n\right) = j \mid X\left(n-1\right) = i\right) = \\ &= P\left(X\left(n\right) = j \mid X\left(1\right) = i_1; X\left(2\right) = i_2; ...; X\left(n-1\right) = i_{n-1}\right) \end{aligned} \tag{1}$$

This way the process *X(n)* will be Markov chain with finite set of the states $\{1;2;\ldots;N\}$. The homogeneous Markov chain should also be discussed since P_{ij} is not dependent on *n*. Let us mark hazard occurrence probability in the *i* node after *n* cycles $\pi_i(n)$. It is clear that $\sum_{i=1}^{N} \pi_i(n) = 1$.

Now it can be returned to the hazard calculation in each node after *n* cycles. Naturally, it is possible to determine only average hazard $\overline{H}_i(n)$ in each node since hazard after *n* steps is a random value. If we would also make an assumption that all the network line flows are equal to 1, the following would be obtained:

$$\overline{H}_i(n) = \frac{1}{n}\sum_{i=1}^{n} H\pi_i(k) = \frac{H}{n}\sum_{k=1}^{n}\pi_i(k) \tag{2}$$

Here H - the hazard that has occurred in one of the network nodes during zero step, i.e., we hold that this node is a point source of the hazard.

From the theory of the Markov chains we know that state probabilities after *n* cycles are described using recursive formulas

$$[\pi_1(1), \pi_2(1), \ldots, \pi_N(1)] = [\pi_1(0), \pi_2(0), \ldots, \pi_N(0)] \cdot [P_{ij}] \tag{3}$$

Or, to put it simpler,

$$\vec{\pi}(1) = \vec{\pi}(0)P \tag{4}$$

here $P = [P_{ij}]$ - transfer probability matrix and $\vec{\pi}(0) = [1, 0, 0, \ldots, 0]$, if we make an assumption that the point source of the hazard is located in the first node.

Then it follows:

$$\vec{\pi}(2) = \vec{\pi}(1)P = (\vec{\pi}(0)P)P = \vec{\pi}(0)P^2,\ i, j = 1,2,\ldots, N \tag{5}$$

Given that $\vec{\pi}(0) = [1, 0, 0, \ldots, 0]$, we receive:

$$\vec{\pi}(n) = \vec{\pi}(n-1)P = \vec{\pi}(0)P^{n-1} = \left[P_{1,1}^{(n-1)}, P_{1,2}^{(n-1)}, \ldots, P_{1,N}^{(n-1)}\right] \tag{6}$$

Thus, we can calculate the average hazard in the node *i* after *n* cycles $\overline{H}_i(n)$ recursively, using the following formula:

$$\overline{H}_i(n) = \frac{H}{n}\sum_{k=1}^{n} P_{1i}^{(k-1)} \tag{7}$$

here $P_{1i}^{(k-1)}$ is the *i* element of the first line of matrix P^{k-1}.

According to the eqn. (2), it is not difficult to prove the theorem of the marginal distribution of the hazard average, when *n* converge to infinity.

If the Markov chain with N states and transfer probability matrix $P = \lfloor P_{ij} \rfloor$ is ergodic, i.e., $\lim_{n\to\infty} \pi_i(n) = \pi_i$, $i = 1,2,\ldots,N$, so marginal hazard average values in all the nodes of the network also exist.

When in the eqn. (2) we reach the limit when n converge to infinity we get:

$$\lim_{n\to\infty} \overline{H}_i(n) = \lim_{n\to\infty} \frac{H}{n} \sum_{k=1}^{n} \pi_i(k) \tag{8}$$

As $\lim_{k\to\infty} \pi_i(k) = \pi_i$, so there is vanishing function $\varepsilon(k)$ which is $\pi_i(k) = \pi_i + \varepsilon_k(i)$, where $\lim_{n\to\infty} \varepsilon_k(i) = 0$, $k = 1,2,\ldots$. Then

$$\begin{aligned} \lim_{n\to\infty} \overline{H}_i(n) &= \lim_{n\to\infty} \frac{H}{n} \sum_{k=1}^{n} (\pi_i + \varepsilon_k(i)) = \\ &= \lim_{n\to\infty} \left(\frac{H}{n} \pi_i \right) n + \lim_{n\to\infty} \frac{H}{n} \sum_{k=1}^{n} \varepsilon_k(i) \end{aligned} \tag{9}$$

Let us select $\varepsilon_{(i)}^{(n)} = \max_{1\le k\le n} \{\varepsilon_1(i); \varepsilon_2(i); \ldots; \varepsilon_n(i)\}$. Then:

$$0 \le \lim_{n\to\infty} \frac{H}{n} \sum_{k=1}^{n} |\varepsilon_k(i)| \le \lim_{n\to\infty} \frac{H}{n} \cdot n\varepsilon^{(n)}(i) = 0 \tag{10}$$

Let $\lim_{n\to\infty} \frac{H}{n} \sum_{k=1}^{n} \varepsilon_k(i) = 0$, and, therefore, $\overline{H}_i = H\pi_i$.

Thus, the marginal average hazard exists in every node, besides, it is equal to the product of the initial H and the marginal node probability.

3.2 The distribution of the additive hazard in the network nodes during the transitional period

In this section, a hazard which is characterized by a value that can be summed or divided as a real number will be analysed. The examples of such hazard are concentrations of significant amounts of hazardous materials, concentration of pollution materials, the amount of water that rises in the reservoir, etc. For calculation of hazard in each network node during the transitional period, systems of equation were made, conditions for marginal hazard existence were specified and for the calculation of marginal hazards in the network systems, systems of equation were formed.

We analyse a network system in which hazard from each node can be transferred to other nodes during one cycle, by dividing hazard $H_i(n)$ of the i node in proportion to the flows q_{ij}, when $j = 1,2,...,N$ and $\sum_{j=1}^{N} q_{ij} \le 1$.

First of all, let us assume that one network node, for example, the first one, is a point source of the additive hazard, in which hazard $H_1(0)$ occurs. Thus, at the zero step we have the following hazard distribution in the nodes:

$$\vec{H}(0) = [H_1(0) \quad 0 \quad ... \quad 0] \tag{11}$$

During the following cycles, hazard modification will occur in each node. From that node hazard will be transferred to other nodes by flows q_{ij}. The total transfer will be:

$$H_i(n)(q_{i1} + q_{i2} ... + q_{i,i-1} + q_{i,i+1} + ... + q_{i,N}) = H_i(n)\hat{q}_i \tag{12}$$

part of hazard. In the node i it will remain

$$H_i(n) - H_i\hat{q}_i = H_i(n)q_{ii} \tag{13}$$

part of hazard. The hazard $\sum_{j=1}^{N} H_{ji} q_{q1}$ will be respectively transferred from other nodes to the node i. Thus, after n cycles, we will have the following hazard in the node i:

$H_i(n+1) = H_1(n)q_{i1} + H_2(n)q_{i2} + ... + H_N(n)q_{iN}$, where $i = 1,2,...,N$.

After defining network transfer matrix $Q = \lfloor q_{ij} \rfloor$, we can write the system of equations in the form of matrix:

$$\vec{H}(n+1) = \vec{H}(n) \cdot Q \tag{14}$$

Thus, we have received hazard distribution in the iterative process. As the process is stationary, i.e. matrix Q is not dependent on the number of cycles n, so irrespective of the initial hazard distribution, this process converges only when all matrix Q own values will be less than one. This is as well the obligatory and sufficient condition for the marginal distribution of the additive hazard in the network systems.

It is easy to ascertain that if the sum of the elements of the lines in square matrix $A = \lfloor a_{ij} \rfloor$ is $\sum_{j=1}^{N} a_{ij} = 1$, where $a_{ij} < 1$, then the sum of the elements of the lines in matrix A^n is also equal to 1, and its elements are $a_{ij}(n) < 1$.

The iterative process of hazard distribution $\vec{H}(n)$ converges when the number of cycles is $n \to \infty$ and it is not dependent on the initial hazard distribution, if all the flows are $0 < q_{ij} < 1$.

It is important to analyse risk distribution after the certain number of iterations.

From the eqn. (14) follows that:

$$\vec{H}(n+1) = \vec{H}(n-1)Q \cdot Q = \vec{H}(n-1)Q^2 = ... = \vec{H}(0)Q^{n+1} \tag{15}$$

Therefore

$$\vec{H}(n+1) = \vec{H}(0)Q^{n+1} \tag{16}$$

This equality allows employing the ideas that are used when proving the ergodic theorems of Markov's chain states.

We shall mark the elements of matrix Q^n this way: $q_{ij}(n)$, and $q_{ij} = q_{ij}(1)$, $i, j = 1,2,...,N$. We will first notice that

$$q_{ij}(n) = \sum_{l=1}^{N} q_{il} q_{lj}(n-1) \geq \min_{1\leq l\leq N} q_{lj}(n-1) \sum_{l=1}^{N} q_{il} = \min_{1\leq l\leq N} q_{lj}(n-1) \tag{17}$$

because $\sum_{l=1}^{N} q_{il} = 1$. This feature is also correct with $q_{ij}(n)$, for which $q_{ij}(n) = \min_{1\leq j\leq N} q_{ij}(n-1)$. Thus,

$$\min_{1\leq i\leq N} q_{ij}(n) \geq \min_{1\leq j\leq N} q_{ij}(n-1) \tag{18}$$

By analogy, it is possible to show that

$$\max_{1\leq i\leq N} q_{ij}(n) \leq \max_{1\leq i\leq N} q_{ij}(n-1) \tag{19}$$

Let us evaluate $q_{ij}(n) - q_{lj}(n)$, for all $i, j, l = 1,2,...,N$. Of course, for any $s < n$:

$$\begin{aligned} q_{ij}(n) - q_{lj}(n) &= \sum_{r=1}^{N} q_{ir}(s) q_{rj}(n-s) - \sum_{r=1}^{N} q_{lr}(s) q_{rj}(n-s) = \\ &= \sum_{r=1}^{N} \left[q_{ir}(s) - q_{rj}(s) \right] q_{rj}(n-s) \end{aligned} \tag{20}$$

Positive differences $q_{ir}(s) - q_{lr}(s)$ will be marked as $\beta_{il}^{(r)}(+)$ and negative ones as $\beta_{il}^{(r)}(-)$. As

$$\sum_{r=1}^{N} q_{ir}\left(s\right) = \sum_{r=1}^{N} q_{lr}\left(s\right) = 1 \tag{21}$$

Thus,

$$0 = \sum_{r=1}^{N} \left[q_{ir}\left(s\right) - q_{lr}\left(s\right)\right] = \sum_{(r)} \beta_{il}^{(r)}\left(+\right) - \sum_{(r)} \beta_{il}^{(r)}\left(-\right) = 0 \tag{22}$$

Let us mark

$$\upsilon_{ij} = \sum_{(r)} \beta_{il}^{(r)}\left(+\right) = \sum_{(r)} \beta_{il}^{(r)}\left(-\right) \tag{23}$$

As all the $q_{ij}\left(s\right) > 0$, so

$$\sum_{(r)} \beta_{il}^{(r)}\left(+\right) < \sum_{r=1}^{N} q_{ir} = 1 \tag{24}$$

Therefore, $0 \le \upsilon_{ij} < 1$. Let us mark $\upsilon = \max_{1 \le i, j \le N} \upsilon_{ij}$. Then $0 \le \upsilon < 1$. It is now clear that:

$$\begin{aligned} \left|q_{ij}^{(n)} - q_{lj}\left(n\right)\right| &= \left|\sum_{r=1}^{N} \beta_{il}^{(r)}\left(+\right) q_{rj}\left(n-s\right) - \sum_{r=1}^{N} \beta_{il}^{(r)}\left(-\right) q_{rj}\left(n-s\right)\right| \le \\ &\le \left|\max_{1\le r\le N} q_{rj}\left(n-s\right) \sum_{r=1}^{N} \beta_{il}^{(r)}\left(+\right) - \min_{1\le r\le N} q_{rj}\left(n-s\right) \sum_{r=1}^{N} \beta_{il}^{(r)}\left(-\right)\right| \le \\ &\le \upsilon \left|\max_{1\le r\le N} q_{rj}\left(n-s\right) - \min_{1\le r\le N} q_{rj}\left(n-s\right)\right| \le \upsilon \max_{1\le i,l\le N} \left|q_{ij}\left(n-s\right) - q_{lj}\left(n-s\right)\right| \end{aligned} \tag{25}$$

For all $i, l = 1,2,\ldots,N$.

Then,

$$\max_{1\le i,l\le N} \left|q_{ij}\left(n\right) - q_{lj}\left(n\right)\right| \le \upsilon \max_{1\le i,l\le N} \left|q_{ij}\left(n\right) - q_{lj}\left(n\right)\right| \tag{26}$$

After using this recursive inequality for $\left]\frac{n}{s}\right[$ of times we get:

$$\max_{1\le i,l\le N} \left|q_{ij}\left(n\right) - q_{lj}\left(n\right)\right| \le \upsilon^{\left]\frac{n}{s}\right[} \max_{1\le i,l\le N} \left|q_{ij}\left(n - \left]\frac{n}{s}\right[\cdot s\right) - q_{lj}\left(n - \left]\frac{n}{s}\right[\cdot s\right)\right| \tag{27}$$

From previous statements $0 < q_{ij}\left(s\right) < 1$, so $\left|q_{ij}\left(s\right) - q_{lj}\left(s\right)\right| \le 1$.

From that we get

$$\max_{1\le i,l\le N} \left|q_{ij}\left(n\right) - q_{lj}\left(n\right)\right| \le \upsilon^{\left]\frac{n}{s}\right[} \tag{28}$$

Then $\lim_{n\to\infty} \max_{1\le i,l\le N} \left| q_{ij}(n) - q_{lj}(n) \right| = 0$.

Let us remember the inequalities of this proof eqn. (18) and eqn. (19). We get that flows $\left\{ \min_{1\le i\le N} q_{ij}(n) \right\}$ and $\left\{ \max_{1\le i\le N} q_{ij}(n) \right\}$ are monotonous and definite ($0 < q_{ij}(n) < 1$), which means that they have limits:

$$q^* = \lim_{n\to\infty} \max_{1\le i\le N} q_{ij}(n) \text{ and } q^{**} = \lim_{n\to\infty} \min_{1\le i\le N} q_{ij}(n).$$

From the eqn. (28) we obtain that there is a limit $q_j = \lim_{n\to\infty} q_{ij} = q_j^* = q_j^{**}$.

It is clear that $1 = \lim_{n\to\infty} \sum_{j=1}^{N} q_{ij}(n) = \sum_{j=1}^{N} q_j$.

Let us return to the eq. (14). When we come to the limit of $n \to \infty$, we mark $\lim_{n\to\infty} Q^{n+1} = \overline{Q}$, $\lim_{n\to\infty} \vec{H}(n+1) = \overline{H}$, and get:

$$\lim_{n\to\infty} \vec{H}(n+1) = \lim_{n\to\infty} \vec{H}(0) Q^{n+1} = H(0)\overline{Q} \tag{29}$$

Or

$$\vec{H} = \vec{H}(0)\overline{Q} = \begin{bmatrix} H & 0 & \dots & 0 \end{bmatrix} \cdot \begin{bmatrix} q_1 & q_2 & \dots & q_N \\ q_1 & q_2 & \dots & q_N \\ \dots & \dots & \dots & \dots \\ q_1 & q_2 & \dots & q_N \end{bmatrix} \tag{30}$$

Thus,

$$\vec{H} = \begin{bmatrix} Hq_1 & Hq_2 & \dots & Hq_N \end{bmatrix} \tag{31}$$

This distribution does not depend on initial conditions.

4 The main results and conclusions

The main aim of the paper is to present the developed hazard distribution mathematical model and its analysis. Here was analysed the mechanism of hazard propagation in network systems in the case of single hazard, evolved in one of the network nodes and in case when hazard arise during each cycle. Few network system cases were analysed. The most important cases are hazard propagation in Markov chains and network systems, where peaks have an immunity or resistance to the hazard characteristic.

References

[1] Adams, J., *Risk*, UCL Press Ltd, London, 1995.

[2] Lefevre, C., Picard, P., An epidemic Model with Fatal Risk, *Mathematical Biosciences,* NY, 117, pp. 127-145, 1993.
[3] Purdy, G., Risk Analysis of the Transportation of Dangerous goods by Road and Rail, *Journal of Hazardous Materials*, 33, pp. 229-259, 1993.

Architecture for humanity: sharing the experience of MERCY Malaysia Core House Project in Banda Aceh, Indonesia

S. S. Zubir, H. Amirrol & N. A. Samah
Faculty of Architecture, Planning & Surveying,
Universiti Teknologi MARA, Malaysia

Abstract

The 26 December 2004 tsunami that devastated many coastal regions of the Indian Ocean has prompted many Government and Non-Government Agencies to look into alternative ways in responding to emergency relief efforts. One way is to adopt traditional methods of construction in order to generate an affordable and sustainable built-environment. The intent of this paper is to delineate the learning experiences from MERCY Malaysia Core Project in providing shelter for the Internally Displaced Persons (IDP) affected by the disaster. This paper focuses on three issues that MERCY took into consideration in formulating a workable praxis to the problems of mass housing in the disaster area. The first was by critically examining the social, cultural and economic factors that were directly connected before the intervention. The second was to establish the most appropriate anti-seismic technology in the construction systems and the third was to generate a design concept that incorporates the above-mentioned issues into a workable solution. Conducting this brief evaluation on the completed batch of the core housing project of MERCY enabled us to rectify future shortcomings and to improve relief efforts succinctly to a particular locality.
Keywords: disaster, humanity, community, social and cultural impact, temporary shelter, anti-seismic, low technology, rehabilitation, reconstruction.

1 Introduction

A sudden change in the life of tsunami survivors in Aceh (Fig. 1) has forced NGOs and particularly architects to come up with ideas that would help reorganize and redirect their lives towards the norm of humanity. Suddenly

www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060401

architecture was all about saving human lives. The appalling condition under which the Internally Displaced Person (IDP) ekes out a meager existence was only too well known. IDPs occupy dilapidated properties that had survived the tsunami. As the absorbing capacity of these areas reached saturation, IDPs were forced to squat on whatever land that was available. As the IDP population increased, sections of them were forced to occupy even the most undesirable places that one could imagine.

The forcible relocation of IDPs into barracks inevitably generated social tension. In isolated instances, NGOs and authorities were compelled to recognize the tenancy rights of IDPs after nearly all means to avoid confrontation, jealousy and social illness have been exhausted [1]. However, these cases are only short-term solutions. Relief works and interventions should seek to reinstate the basic needs of human existence. It was our belief and obligation that a sustainable community planning could provide the framework for the fulfillment of these basic human needs. This belief underpins our approach in providing basic shelter to the victims: mobilizing creative and innovative thinking and technology in safeguarding their future survival on this devastated land.

Figure 1: Banda Aceh and Aceh Besar, before and after tsunami. Courtesy of United Nations Humanitarian Information Center (HIC).

Even as the first relief supplies were being delivered to locations in Aceh, MERCY Malaysia was responding to the emerging long-term needs of the Acehnese people, signing a Service Level Agreement (SLA) letter with the local government to build Core Houses while adopting a whole village in the weeks after the tsunami. The nature of the project will be a redevelopment of a village at Desa Weu Raya, situated one kilometer from the hardest hit coastline of Aceh. The master plan includes a rebuilding proposal to replace 231 units of core houses for the remaining families of the village of Desa Weu Raya, which used to be the homes for more than 3000 lives, now only hosting 600 survivors. The

proposal (Fig. 2) was a module of 52m^2 single storey seismic resistant, semi brick core house with flexibility to be expanded upon future needs, consisting of two bedrooms, a living room and a front and a rear terrace. These terraces may be converted into kitchen, additional bedroom, or to be expanded to the size of the housing unit. Communal toilet facilities were built as requested by the villagers.

2 The approaches

Even before the tsunami, millions of people in the affected areas were living in conditions of poverty unimaginable to most people. In Aceh, the security of lives, possessions, and infrastructure have been threatened by several years of armed conflict. According to the government's own statistics [2], in 2002, nearly 48.5 percent of the population had no access to clean water, 36.2 percent of children under the age of five were undernourished, and 38 percent of the population had no access to health facilities. The poverty rate doubled from 14.7 percent in 1999 to 29.8 percent in 2002. These problems worsened with the recurring yearly disaster that struck Indonesia.

Due to these disturbing statistics, three main issues were carefully considered in the development of a critical praxis around a disaster-hit area. First was to identify ways to approach community needs, social and cultural understanding, and economic factors, which were profoundly connected prior to the interventions. The second issue was to devise anti-seismic systems of constructions. The third issue was to establish design concepts derived from the two mentioned issues, which were then translated into workable solutions. Before rebuilding works could be executed the remaining 600 villagers were relocated temporarily into IDP camps.

Figure 2: Core House plan.

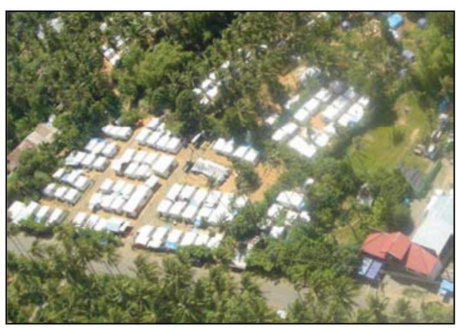


Figure 3: Desa Weu Raya IDP Camp.

A tent village (Fig. 3) to host the families temporarily for six months was constructed while waiting for the Core Houses to be completed. 231 tents were erected on a vacant land donated by one of the tsunami victims. The tent village also included community spaces such as a praying area, mobile clinics, spacious toilet facilities, a kindergarten, an activity area for children and women, office space for the village administration and tents that can be used as shops to help

rebuild the social and economic capacity of the victims. The tents, fabricated of lightweight and durable materials (Fig. 4) were easily assembled (no skill workers required) with minimal logistic requirements. The villagers of Desa Weu Raya managed comfortably in this tent village while waiting for the new houses to be completed (Fig. 5).

Figure 4: Setting up a tent village.

Figure 5: The IDP camp setting.

2.1 Community needs, social and cultural understanding, and economic factors

In general, there are two types of architecture. The first responds to contextual needs and the other driven by market forces. Disaster has taught architects to work like Hassan Fathy [3] who propagated the incorporation of considerations for social-cultural and local components of the built form in the design process. In addition, informal discussions and regular visits to the IDP camp have molded MERCY's understanding of the community needs. The theory of 'small is beautiful' and the concept of self-help approaches to shelter should be accorded greater emphasis.

In general post-disaster settlement must be constructed of permanent structures with the necessary infrastructural facilities and amenities [4]. The World Bank, the United Nations and many other experts have concluded that tsunami-hit areas like Aceh would need billions of dollars of external assistance to finance their basic shelter needs. As this aid is constantly available, but not efficiently distributed, thousands of tsunami survivors are still condemned to live in sub-standard living environments.

The complex relationship between the architect and the community must be established before any interventions can take place. In this project, the feasibility of the interventions, which included availability of materials, work force and construction techniques, had been carefully scrutinized before any final decisions were made, In coordination meetings, agreements were only achieved after both parties had fully understood the reality of the life faced by tsunami victims and the architect's aim. What emerged from these understandings was then put to test by constructing them.

2.2 Anti seismic technology in building construction

In lieu of human losses in recent earthquakes and tsunamis, not much research and development were conducted on seismic resistance of low-cost masonry and

timber buildings [5]. Research centres and universities of the developed and developing countries have carried out extensive amount of earthquake related research and development but these deal mainly with expensive modern buildings.

Due to economic and contextual constraints in post-tsunami Aceh, only masonry and timber houses were the best choice for MERCY Malaysia to construct for the IDPs. This type of construction cannot be replaced quickly and easily with modern earthquake resistant construction systems due to the large dependence on low technology, availability of material and local workforce.

Unreinforced masonry and timber structures generally have a brittle lateral-load-resisting system. These types of buildings have poor seismic performance. Total or partial collapse of these structures has been one of the main causes of the loss of life during past earthquakes [6] (Fig. 6). Some of the important causes of the failure of the walls and the subsequent collapse of roofs of this type of structure are as follows:

1 The low durability and the poor strength of materials (earth/clay) used in the construction of adobe buildings.
2 The inferior quality of the mortar used in the construction of bricks, concrete blocks or masonry buildings.
3 Lack of sufficiently strong connection between the various elements of the structure, such as between walls, partitions and roof. In addition, the brittleness and low strength of the individual elements especially the walls. Often, these structures do not behave as a rigid frame. Under moderate shaking, various elements become disconnected from each other and gradually causing the failure of the whole system.
4 The collapse of the unreinforced load-bearing walls, partitions and spandrels over the openings due to the effects of shear-compression and shear-tension. The failure modes by these effects are very brittle and highly strength degrading. Therefore, under intense tremor, the wall elements crack and subsequently disintegrate the load-bearing walls and partitions, destroying the lateral strength and rigidity of the structure and also the vertical-load carrying capacity of the walls, resulting in the collapse of roofs and supporting structures.
5 Disintegration of the heavy, brittle, non-monolithic and non-rigid timbered roofs or roofs made of unbraced trusses.

Figure 6: Poor seismic performance of domestic buildings in Aceh.

Due to time constraints and shortage of skilled workers, materials, machineries, and capital, MERCY Malaysia's first batch of core houses for the IDPs were constructed without complicated earthquake resistant design. Instead an alternative technique in preventing the collapse of the roof and other structure of houses was used. This technique was developed by MERCY Malaysia in collaboration with other experts. The technique incorporates a simple braced skeleton system (Fig. 7) within the house design that can prevent structure failure and collapse during an earthquake [7]. The function of this skeleton is to tie together various parts of the building. The main design components of this system are:

1 A horizontal in-plane diagonal bracing system to tie together the roof beams so that a monolithic roof with sufficient rigidity is obtained (Fig. 8).
2 An auxiliary system of timber or reinforced concrete columns embedded within the masonry walls so that in case of destruction of the walls, the ductile columns can support the vertical loading.
3 These columns are braced together laterally by means of a vertical system of diagonal bracing in such a way that the lateral instability and subsequent collapse of the structure during earthquake is prevented (Fig. 9).

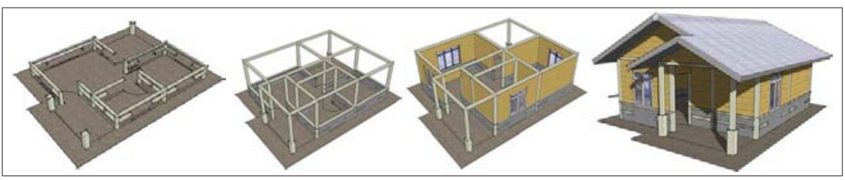


Figure 7: Construction concept and techniques.

In reference to the proposed earthquake protection criteria for the Core Houses, the primary goal is to prevent structural collapse during major earthquakes. From our research and experiments through earthquake simulations (Fig. 10), a structure become less susceptible to collapse under strong earthquake if:

1 The roof and floor of the building remain monolithic with sufficient in-plane rigidity during and after earthquakes;
2 The vertical load-carrying system of the structure can survive the earthquake without impairment to its function;
3 The lateral load-resisting system of that structure during the main earthquake and succeeding after shocks can retain enough residual capacity to resist safely the lateral forces due to lateral instability, wind load and future earthquake lateral loads.

Collaborations between architects, engineers, contractors, experts, humanitarians and local communities have resulted in an idea of tolerance that reduces the risk of construction failure during earthquakes. Without overestimating the omnipotence of new construction method and technology, architects should remain committed to utilizing the progressive potentials of surrounding resources. A limited budget, shortage of materials, machineries, and

skilled workers had driven the architects to undertake creative solutions. The architects are inherently charged with the mission of creating designs that contribute towards the enhancement of the victims' quality of life. The use of technology and techniques should always be constructive to civilization. Technology must be manageable by the locals for their benefit. MERCY Malaysia has collaborated with local and foreign experts from Kyoto University, Japan, Universitas Syiah Kuala, Banda Aceh (UNSYIAH), National Society for Earthquake Technology, Nepal (NSET), Geo Hazards International (GHI) and contractors in the conceptualization and realization of the Core House project.

Figure 8: Roof bracing to obtain sufficient rigidity.

Figure 9: Diagonal bracing to prevent lateral instability.

Figure 10: Shake table earthquake simulation to test the idea of seismic resistant design and construction.

The development of anti-seismic construction methods and techniques has been channeled down to the local community and authorities through workshops, simulations, presentations and discussions. During these stages, MERCY progressively learned to appreciate the contextual components of the area (soil, climate, availability of resources, etc.). It was also important to understand the construction techniques practiced by Aceh builders in order to introduce them to alternative and better techniques. As architecture is gradually absorbed into the knowledge industry, and the role of the architect has shifted from building to

organizing social relations, MERCY Malaysia feels that in any future interventions of the public realms, a better inter-relation between the architects and the end users must be established in order to generate successful aid programmes [8].

3 Implementation concepts

After three months of preparatory work, the proposed Core House scheme for the devastated conurbation started in March 2005 and was successfully completed within six months. This was possible because the programme did not adhere to the static land use requirements and blue prints prepared by the local Shelter Working Group and the Reconstruction and Rehabilitation Body (BRR), which, are so frequently produced but regularly shelved or unexecuted. The deficiencies of the blue print, as well as the constraints imposed by the regional condition were clearly set out in the scheme. The scheme as a whole was an indicative, policy-oriented proposal rather than a definitive housing scheme based on statutory rules, space requirements and other essentially concerns with political decision-making and implementation. The notion of dichotomy between capacity building and social restructuring, which was unrelated to the resource of allocation has been put aside.

With respect to the daily living conditions in Aceh, the major physical deficiencies identified in our studies are the most elemental: adequate and safe shelter and the sub-standard living environments for the foreseeable future. The maximum elevation within the district is only about two meters above mean sea level and the area has little clean water supply. Most of the lands were swept by seawater. The ground spaces in these settlements must be fully and expansively utilized. The absence of trees had caused the area to be susceptible to strong winds and the possibility of another tidal encroachment. There is no proper sewerage system around the remaining land. The vast majority of the people depend on service privies; a small shed containing a seating platform with earthenware bowl underneath for the disposal of excreta.

MERCY Malaysia Core House scheme outlined a desperate position in respect of critical shelter need. The sheer inadequacy and constant deterioration in living condition in the fullest environmental and social senses are central to a broad range of Aceh's current problems, and the plan fully acknowledges our commitment to provide shelter and to improve living conditions in both the short- (IDP Camp settlements) and the long-term (Core House) perspectives (Fig. 11). There are unique strains on working with the community in several parts of the tsunami-hit areas. In some places where a substantial number of affected people we encountered were still holding on to strong principles of traditional ways of lifestyles. As such, MERCY Malaysia and its partners had to find roles in advising and educating them on the best practices of disaster preparedness and safe construction. From previous experiences, we try to ensure that the voices of tsunami-affected people are heard by having:

1 Bi-weekly meetings with villagers, community leaders, religious leaders, authorities and other NGOs, discussing issues of spatial ordering, various

post-disaster management initiatives, and issues related to reconstruction and rehabilitation.

2 On site discussions with each beneficiary (house owner) in determining their own plot of land, house positioning, boundaries, nearest escape routes, and to make sure all beneficiaries are well informed on the designated mitigation plans, safe areas, escape routes, disaster management zones and ways of alerting each other.

3 Effective disaster mitigation that focuses on the key elements of self-help, cooperation and education through activities such as: (a) research projects; (b) training and capacity building; (c) a series of international workshops, and (d) advisory services.

4 Discussions that lead to a development of prevention and mitigation plans. From the plans, we have produced a comprehensive escape routes plan for the village of Desa Weu Raya, designed clear signage and iconographic symbols that lead to the escape routes, and have made clear guidelines on what should be done whenever disaster alert information is being disseminated and made known before, during, and after a disaster.

Figure 11: IDP Camp settlements and the completed Core House.

4 Conclusion

In building houses for a community that live with risks, all plans that have been developed need to be properly studied, discussed, and put to test. All plans applied must be re- analyzed to ensure its effectiveness and workability in responding to a wider global context. Tests have to be done in a practical manner by organizing workshops, forums, simulations, trainings and networking with other inter-related agencies. Public participation is a must to fully understand the needs and in responding to these needs. An overall review on the systems and plans will produce new improvised results, which will help in producing a well-prepared action plan in disaster and other acute cases. The tested idea must be well documented and to be spread to a wider audience of humanitarian workers, experts in the built environment, and the civil society [9].

The intention of this project is not so much a complete ‘otherness’ from other ad-hoc disaster related housing development, but rather a symbiotic merging with its context. Permeable boundaries allow extensions inwards and outwards, and encourage inhabitants to convene, converge, and enjoy the many possibilities offered within the new reconstruction development. This post-tsunami planning substance that connects and separates, is amorphous, fluid, organic, and can contain within itself as well as surround (the other).

Floating within this idea, the Core Houses allow for differences and pronounced simultaneities to grow and mutate but balanced act in the desperate condition of the inhabitants. Their containment allows further expansions. The beneficiaries will find their ideal position within the site by almost naturally gravitating towards desirable adjacencies: family needs, economic conditions, similarities to surrounding typologies, ideal community connections, neighboring activities, or just the most basic shelter provided. It is an amalgam of family togetherness and societies' strength. It is, in essence, architecture for humanity (Fig. 12).

Figure 12: Core House Desa Weu Raya, Lhoknga, Aceh, Indonesia.

References

[1] Morris, Eric, UN Recovery coordinator for Aceh. *'A Place to Stay, A Place to Live', Challenges in Providing Shelter in India, Indonesia and Sri Lanka after the Tsunami,* Oxfam International: London, 2005.

[2] Indonesian Human Development Report 2004, *The Economics of Democracy: Financing Human Development in Indonesia',* BPS-Statistics Indonesia, BAPPENAS and UNDP: Indonesia, 2004.

[3] Fathy, Hassan, *Architecture for the Poor: An Experiment in Rural Egypt,* University of Chicago Press: Chicago, 1973.

[4] Pendick, D., *Build Smart Not Hard,* State University of New York, Courtesy of www.labdna.com: New York, 2004.

[5] Razani R., *Criteria for Seismic Design of Low Cost Housing,* Pergamon Press, The East West Center: New York, 1979.

[6] Goodman, L.J. & Tabujara, E.G., (eds). *Low-Cost Housing Technology, an East-West Perspective,* Pergamon Press, The East West Center: New York, 1979.

[7] Samah, N.A., *Safe Building in Disaster Area (Unpublished Master Thesis),* Department of Engineering, Faculty of Engineering, Universiti Sains Malaysia: Malaysia, 2006.

[8] Mahmood, J., MERCY Malaysia Annual Report 2005, www.mercy.org.my

[9] Amirrol, H. & Samah, N.A., Civil Society Role in Disaster; Disaster Risk Reduction: Learning From Tsunami Experience Training Workshop, MERCY Malaysia: Kuala Lumpur, 2006.

Erosion of forestry land: causes and rehabilitation

T. Ogawa[1], Y. Yamada[2], H. Gotoh[2] & M. Takezawa[2]
[1]*Forest Survey Office, Japan*
[2]*Department of Civil Engineering, College of Science & Technology, Nihon University, Japan*

Abstract

Forests cover 70% of the total land area of Japan. Forest lands within Japan are prone to landslides because weakly resistant geological units are eroded by water flowing down steep slopes that are subjected to annual rainfall amounts that are 2.5-times the global average. The environmental effects of deforestation impact upon atmospheric pollution, wildlife, the hydrological cycle, water resources, soil erosion, and the occurrence of landslides. To mitigate disasters that occur upon forestry land, it is important to forecast landslide development and plan for the provision of remedial measures during disaster rehabilitation. This paper describes the causes of the erosion of forestry land and methods of disaster rehabilitation via a case study of the upper reaches of the Tama River, Japan, which is a national park and an important water resource for the Tokyo Metropolitan area. The causes of erosion of forestry land within the upper reaches of the Tama River are classified as one of the following: shallow landslides related to the loss of under-story vegetation, collapse of steep slopes, damage related to the consumption of vegetation by wildlife, and debris flows that occur during periods of torrential rain. In recent times, heavy rains over eroded forestry land within the upper reaches of the Tama River have produced muddy river water due to the erosion and degradation of mountain slopes. In addition, grazing by Japanese deer has destroyed many trees within the upper reaches of the Tama River, and the torrent bed within this area, previously planted with Japanese horseradish, was lost during a debris flow. In this paper, we describe anti-erosion measures undertaken for disaster rehabilitation of wasted forestry land, including timber-thinning methods and the control of wildlife numbers.
Keywords: erosion of forestry land, deforestation, disaster rehabilitation, landslide.

1 Introduction

It is fair to say that a human life is created by the green of the forest that brings about a mild climate, saves a water resource, and serves the coexistence of animals and plants. The forest is a stable system that can sustain nature. About 70% of the total land area of Japan is forested. Japan is blessed with the most abundant forest resources of any country in the world. The forests of Japan were depleted during the Second World War, but Japanese red cedar (*Cryptomeria japonica*) and Hinoki cypress (*Chamaecyparis obtuse*) were replanted during the post-war period at the demand of the Japanese Government. Plantation forests currently comprise about 40% of the total forests in Japan, but such wood is in poor demand because foreign lumber can get purchased cheaply. The cost of a log of Japanese red cedar or Hinoki cypress is just 33% or less of the price 25 years ago. The rate of self-sufficiency of forests in Japan is less than 20%; consequently, the average age of forestry workers increases and the next generation of forestry owners have moved to the city because forestry is not a desirable career. Hence, the management of forests in Japan has been neglected in recent times, leading to erosion and poor water conservation. In this paper, we discuss the causes of eroded forest land and possible rehabilitation measures.

2 Forest land in Japan

The forested area of Japan has decreased with increasing population and the development of agricultural fields. The cultivated acreage of forest was about 8.62 million ha for about 7,000,000 people in 930 AD, increasing to 54 million ha for 121,000,000 people in 1990: an increase in persons per 1 ha of cultivated acreage from 8.1 to 22.4. Temporal trends in population and cultivated acreage are shown in Table 1 (Iketani [1]).

Table 1: Temporal trends in cultivated acreage and population within Japan.

Year	Cultivated acreage (million ha)	Population (million)	Persons per hectare
930	8.62	7	8.1
1450	9.46	10	10.5
1600	16.35	19.6	12.0
1720	29.70	31	10.4
1874	30.50	34	11.1
1990	54.00	121	22.4

The sizes of areas of protected forest in Japan in 2004 are shown in Table 2. The ‘Other’ category in the table includes shifting-sand prevention forest, windbreak forest, flood-damage prevention forest, tidal wave and salty wind prevention forest, drought-prevention forest, snow-drift prevention forest, fog-inflow prevention forest, snow-avalanche prevention forest, rock-fall prevention forest, fire protection forest, fish-breeding forest, navigation landmark forest,

public health forest, and scenic-site conservation forest. Headwater conservation forest comprises 68.4% of protected forest, while soil run-off prevention forest comprises 21.5%, as shown in Table 2 (JFS [2]).

Table 2: Land areas of different types of protected forest.

Classification	National forest	Non-national forest	Total	Ration
Headwater conservation forest	4,228	3,216	7,444	68.4%
Soil run off prevention forest	935	1,404	2,399	21.5%
Landslide prevention forest	19	37	56	0.5%
Other	458	590	1,048	9.6%

(unit: thousand ha).

Felled forestry land was 19,830 ha; the total area of landslides per 100 ha was 2.38 ha; and the total area of landslides in areas of mature forest (190,328 ha) was 1.17 ha per 100 ha. Differences in landslide development between areas of mature forest and felled areas are shown in Table 3 (JSECE [3]). The total landslide area per 100 ha of felled land is about 2-times that of mature forest areas, as shown in Table 3. Therefore, one of the causes of landslide development is deforestation.

Table 3: Relations between areas of felled forest and landslide development.

Kind	Area (ha)	Landslide (places)	Landslide (ha)	Landslide area (100 ha)
Tree-grown area	190,328	11,286	2,277	1.17
Cut-over land	19,830	2,377	398	2.38
Total	210,158	13,663	2,625	1.25

Trees and plants in general affect the hydrological cycle in a number of significant ways. Therefore, the presence or absence of trees can change the quantity of water upon the land surface, within the soil or groundwater reservoir, and in the atmosphere. Deforestation generally increases the rate of soil erosion by increasing the amount of runoff and reducing the protection of the soil afforded by tree litter. Forestry operations themselves also increase erosion via the development of roads and the use of mechanized equipment. A further cause of landslide development is the grazing of wild animals. In some cases, treeless hills result from the grazing upon growing herbage of wild animals such as deer, monkey, and bear. The surface of the earth is then exposed and washed away during rainfall. Tree roots act to bind soil between the roots and between the

roots and underlying bedrock if the soil is sufficiently shallow. The risk of landslides is therefore increased when trees are removed from steep slopes with shallow soil and when subjected to the grazing of wild animals. The following case study provides an example of the way in which the rehabilitation of eroded forest land is planned and executed in Japan.

3 Case study

The rehabilitation of eroded forest land is planned for the area around the Sakasa River, a branch of the Tama River, as shown in Figure 1 (JT [4]). The study area of the current investigation comprises 131 ha of fast-moving river and hillsides within the 220-ha basin catchment of the Sakasa River. The Sakasa Basin contains 198.07 ha of forest that is designated a protected area with the status of 'protection for headwater conservation'.

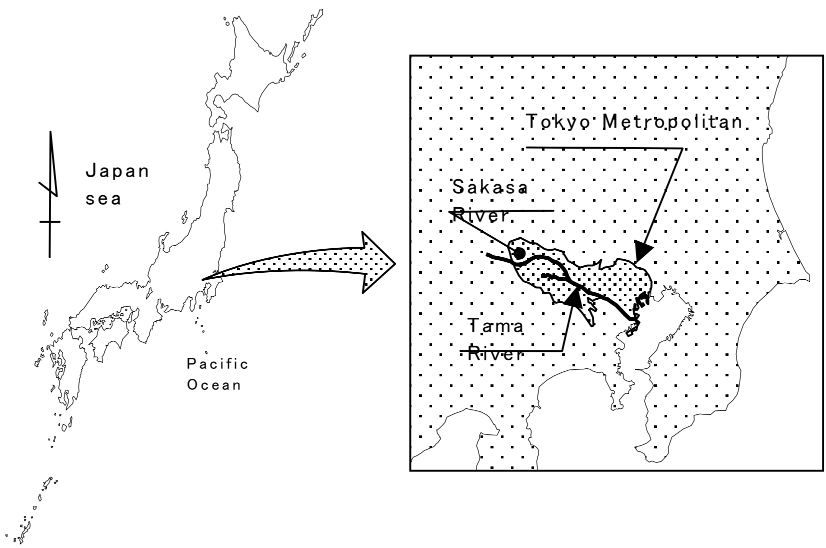


Figure 1: Location map of the study area.

The study area is the mountainous region between Kawanori Mountain, with an altitude of 1353 m, and Honnita Mountain, with an altitude of 2245 m. The upper reaches of the Sakasa River flow to the south, while the middle reaches flow to the west following a 90˚ change in flow direction at an altitude of 790 m related to a geological rift. The topography of the upper Sakasa River formed from crustal movements, and large-scale landslides occur at 3 sites in the upper river. Hillsides in this area are steep, with an average slope of 36˚. Eroded subvertical cliffs occur at the levee foot of mountain streams in particular, and the slopes of the lower parts of such cliffs are in excess of 60˚. The bed slope of the Kawanori River where it joins the Sakasa River is 10˚; this section of the river forms a V-shaped valley with numerous falls of 50–60 m. The average

slope of the riverbed of the upper Sakasa River is 16˚; this section of river contains many falls of 2–10 m, and rocks are exposed along the majority of the rivercourse. There are three main river branches within the Sakasa Basin. Most areas within the river basins are 20–45 ha in size and many bed slopes are in excess of 20˚.

The topography of the Sakasa River consists of six distinct regions, as shown in Table 4 and Figure 2.

Table 4: Characteristics of the topography of the Sakasa River.

Part	Area (ha)	Length (m)	Height ratio	Bed slope (%)	Slope angle (Deg)	Elonga-tion ratio
Lower	35.85	720	160	22	41	
Midstream	10.63	540	90	17	38	
Upper	85.64	1390	460	33	35	0.75
1 stream	19.84	570	290	51	38	0.88
2 stream	22.72	630	330	52	37	0.85
3 stream	45.32	900	360	40	35	0.84
Total	220.00	2650	710	28	37	0.63

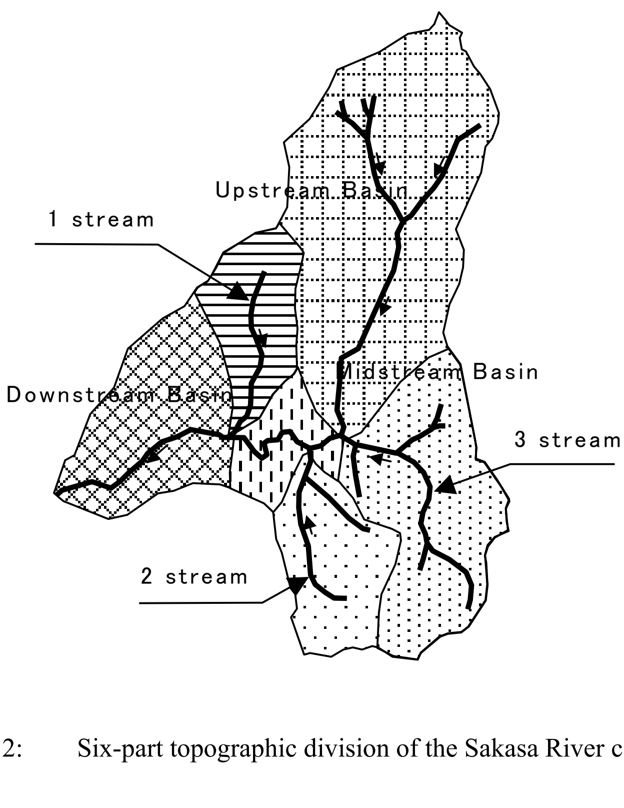


Figure 2: Six-part topographic division of the Sakasa River catchment.

The elongation ratio (E) is given by the equation $E = (2/L)\sqrt{(A/\pi)}$, where L is the river length (m), A is area (ha), and π is the circular constant. The geology of the study area consists of Late Jurassic sandstone underlain by Middle Jurassic shale. The sandstone and shale is overlain by a layer of loam in mountainous areas; the loam is weak and commonly deformed.

Temperature and precipitation data for altitudes of 550 m, 1000 m, and 1363 m are shown in Table 5. The maximum recorded rates of rainfall are 71 mm/hour (1991/8/20), 347 mm/day (2001/9/10), and 634 mm/3 days (2001/9/10).

Cryptomeria japonica and *Chamaecyparis obtuse* planted over the past 30–50 years now covers about 60% of the total forest area across the study site. A classification of the forest area on the basis of tree type is shown in Table 6, while eroded forest land is classified in terms of three types of eroded hillsides and torrent-erosion, as shown in Table 7. Eroded forest land is classified in terms of rainfall, intensity of felling, and damage related to grazing by Japanese deer. In recent years, Japan has been struck by severe typhoons and torrential rain, such that the amount of eroded forest land has increased due to a combination of these rainfall conditions, weak soil, and the grazing of Japanese deer. The main cause of erosion of forest land in the study area is grazing by increasing numbers of Japanese deer that eat the buds and roots of trees. Rehabilitation of the eroded forest land requires the control and management of an appropriate population of Japanese deer. Measures to prevent run-off include hillside works such as the conservation of vegetation within devastated lands and torrent works such as the protection of valleys that contain unstable soil.

Table 5: Temperature and precipitation data for different altitudes.

	Altitude 530 m	Altitude 1000 m	Altitude 1363 m
Average temperature	11.8^0C	9.0^0C	6.8^0C
Volume of temperature Warming	90.1	67.7	51.6
Volume of temperature Coldness	-8.5	-19.7	-30.0
Classification of forest zone	Temperate zone, evergreen and broadleaf forest	Temperate zone and fallen leaf forest	Cold zone and fallen leaf forest
Annual precipitation	1595 mm	1971 mm	2259 mm

Table 8 describes the characteristics of different types of hillside and torrent works in terms of: (1) type of works, (2) effectiveness, (3) durability, (4) value for money, (5) ease of undertaking the work, (6) ease of transportation of materials, (7) protection of the landscape, (8) combats grazing by Japanese deer,

and (9) overall evaluation. The meanings of the symbols used in the figure are as follows: ◎: very good performance; O: good performance; △: poor performance; and ×: very poor performance. Hillside work is classified as either technical hillside work or hillside-seeding work. Table 8 (a) assesses soil-retaining works as part of technical hillside works. Basket-retaining works (A) perform well in both sides of the water-channel work, while log-piling work (C) performs well in soil-retaining works Table 8 (b) assesses water-channel work as part of technical hillside work; the wire net shows the best performance.

Table 6: Classification of forest areas on the basis of tree type.

Kind of tree	Area
Cryptomeria japonica and Chamaecyparis obtuse	60.63 ha (27 %)
Cryptomeria japonica	28.45 ha (13 %)
Chamaecyparis obtuse and Broad leaved trees	33.70 ha (15 %)
Chamaecyparis obtuse	3.30 ha (2 %)
Quercus mongolica var. grosseserrata and Broad leaved trees	49.80 ha (23 %)
Quercus mongolica var. grossesrrata and Larix leptolepis	3.75 ha (2 %)
Quercus mongolica var. grossesrrata	5.57 ha (3 %)
Quercus serrata and Broad leaves trees	16.05 ha (7 %)
Others broad leaves trees	12.05 ha (5 %)
Tsuga sieboldii	4.10 ha (2 %)
Pinus densiflora	0.80 ha (0 %)
Larix leptolepis	1.80 ha (1 %)
Total	220.00 ha (100%)

Table 7: Classification scheme of eroded forest land.

Classification	Form
Eroded hillside (1)	Outflow of surface soil by loss of under-story vegetation
Eroded hillside (2)	Slope failure of steep slope
Eroded hillside (3)	Harm of deer's food
Eroded torrent	Outflow zone of soil by torrent bed gradient

Table 8 (c) provides an assessment of different wattle fence and linear sodding works as part of hillside seeding works. Log wattle fences and log linear sodding provide good performance because thinning lumber can be used.

Table 8 (d) provides an assessment of covering works as part of hillside seeding works. In this case, natural fiber mat provides the best performance.

Table 8 (e) provides an assessment of covering works to protect vegetation from grazing by Japanese deer. Combined works involving a thick layer wire net and SHIKATTO works provides the best results, where the SHIKATTO work provides protection from grazing by deer.

Table 8: (a) Technical hillside works (soil-retaining works); (b) Technical hillside works (water-channel works); (c) Hillside seeding works (wattle fence and linear sodding works); (d) Hillside seeding works (covering works); (e) Hillside seeding works (covering works: ward off Japanese deer); (f) Torrent works.

(a)

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A	○	○	○	○	◎	◎	−	◎
B	○	○	○	△	△	△	−	△
C	◎	△	△	○	○	◎	−	◎
D	◎	△	○	○	○	◎	−	○

A: Basket retain work, B: Wire basket work, C: Log piling work, D: Wood steel wall.

(b)

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
E	○	○	○	△	◎	○	○	○
F	◎	◎	○	○	○	○	◎	◎
G	○	○	△	△	△	○	×	△
H	○	○	△	△	△	△	○	△

E: Sand bag, F: Wire net, G: Sodded channel, H: Corrugated metal pipe.

(c)

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
I	○	△	○	△	△	◎	○	△
J	○	△	○	○	○	◎	○	◎
K	○	○	○	○	○	○	○	○
L	○	○	○	○	○	△	○	○
M	○	△	○	○	○	◎	○	◎
N	○	△	○	○	○	○	○	○

I: Wicker work, J: Log wattle fence, K: Steel linear sodding work, L: Wattle fence with vegetation, M: Log linear sodding work, N: Linear sodding work with vegetation.

(d)

(1)	(2)	(3)	(4)	(5)	(6)	(8)	(9)
O	◎	○	○	△	△	○	△
P	○	△	◎	◎	◎	△	△
Q	◎	○	△	△	△	△	△
R	◎	○	◎	○	○	△	○
S	○	○	◎	○	○	△	○

O: Fagot, P: Straw mat with seed, Q: Vegetation mat with thick top soil, R: Natural fiber mat, S: Chemical fiber mat.

Table 8: Continued.

(e)

(1)	(2)	(3)	(4)	(5)	(6)	(8)	(9)
T	○	○	○	○	○	◎	◎
U	○	○	○	○	○	◎	○
V	◎	○	○	○	○	◎	◎
W	◎	○	○	○	○	◎	◎

T: Wire net together, U: Diamond wire net, V: Thick layer wire net, W: SHIKATTO work.

(f)

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(9)
I	◎	○	○	○	○	◎	◎
II	◎	△	△	○	○	◎	○
III	○	○	△	△	△	○	×
IV	○	○	○	○	△	○	○
V	○	◎	△	△	×	△	×

I : Steel crib structure, II : Wood crib structure, III: Cellular structure, IV: High energy absorber fence, V: Concrete structure.

Table 8 (f) provides an assessment of torrent work for protecting valleys that contain unstable soil. Torrent beds previously planted with Japanese horseradish have been repeatedly lost to debris flows.

Torrent works using trench soils and secondary products that are easily transported are selected for crib dams; the wood crib structure is the structure that is mainly used. The steel crib structure is best in downstream areas and at river junctions.

4 Conclusions

The causes of the erosion of forest lands include intense rainfall events in recent years, deforestation, and increases in wildlife numbers. The effects of global warming are increasing year-by-year: evaporation increases with a warming climate; the average global precipitation increases; soil moisture is likely to decline in many regions; and intense rainstorms are likely to increase in frequency. Deforestation within a basin causes water- and wind-derived erosion, a decline in soil fertility, and the development of landslides. It is known that the roots of Japanese red cedar and Hinoki cypress decay within 7 years of felling. It is therefore unreasonable to expect prolonged resistance to landslide development in areas of felled trees because the roots decay over time. The process of landslide development following the grazing of wild animals involves the removal of surface soil, gullying, and rill formation. The number of Japanese deer in the study area has been increasing rapidly because the area provides a

favorable environment for the deer. The population of Japanese deer in this area has increased by 660% over the past 10 years. The rehabilitation of eroded forest land must involve both forestation measures and measures to control and manage the wildlife population at a suitable level.

The Forest Improvement Plan for the devastated area is to be carried out comprehensively via the Forestation Project and the Protected Forest Improvement Project. The Forest Improvement Plan is concerned with the growing conditions of trees, the existence of under-story vegetation, crown density, the ratio of forms, the ratio of yield, etc. The plan involves the growth of various types of under-story vegetation, the development of an ecologically multi-storied forest with sufficient ground litter, sufficient intensity of illumination under the tree crown, and enhancing the growth rate of trees by periodically carrying out forest improvement works. Such works enhance public services such as water conservation and the prevention of run-off by afforestation, and the best environment is created for wildlife.

Acknowledgement

A part of this research was conducted with the financial support of the Tokyo Metropolitan Government. The authors thank the cooperators of this survey.

References

[1] H. Iketani (1999) Disaster of Debris Flow, Iwanami Series, pp.29.
[2] Japan Forestry Society (2005) Forest Handbook, pp.9.
[3] Japan Society of Erosion Control Engineering (1999) Movable Events of Soil at Slope, Sabo Course Vol. 3, pp.169.
[4] Japan Technology Co. Ltd. (2005) General Survey Report of Forest Conservation Project in the Sakasa River, pp.1~197.

Dual system for management of natural and anthropogenic emergencies and its training

G. A. Sevilla, A. D. Acquesta & G. M. Giráldez
Instituto de Investigaciones Científicas y Técnicas de las Fuerzas Armadas, República Argentina

Abstract

Natural or anthropogenic emergencies change the way in which a community works, generating crisis situations and preventing the community from recovering from the impact by itself. These emergencies may destroy power, communication and drinkable water supply networks, routes of access, engineering works, crops, factories, homes and human lives.

Frequently institutions are forced to respond to those situations with information that is not complete, up-to-date or reliable. Dissimilar organizations must work jointly and use the material and human assets available in an effective way.

Decisions that have a direct effect on the lives of hundreds or even thousands of people are made within short periods of time and with a high degree of uncertainty.

This work aims at approaching the problem in an integral way, presenting the design of an informatics system that enables crisis management as well as training for the members of the organizations involved in the response to a natural or anthropogenic emergency.

We propose an Internet-supported multi-user system with tools that enable the registering, follow-up and monitoring for support requirements, access to an integrated database of available material and human assets, communication and coordination resources among participants, phenomenon mathematic model support for analysis and prognosis, and an external simulation engine to build and carry out emergency management exercises.
Keywords: crisis management, training system, emergency management, natural phenomena, simulation model, risk analysis, risk assessment, risk management, information and communication issues, population protection, control management and protection.

1 Motivation

During the emergency management, several factors appear that might complicate to a greater or lesser extent the tasks of those who must respond to the crisis. The following factors may be considered:

Characteristic of the nature of the phenomenon:

- The phenomenon status can't be known either in real time or to its full extent.
- There is uncertainty about the phenomenon evolution and it is difficult to get accurate forecasts.

Characteristic of the organization:

- The necessary procedures are not up-to-date, familiar or even present.
- There are difficulties to do emergency management exercises.
- The organization is not familiar enough with the information about its own assets – existence, location, status, personnel in charge, involvement and availability.
- The functions and responsibilities of individuals during an emergency are not properly determined.

Characteristic of the communication among organizations:

- Information about the assets of other organizations involved – existence, location, status, personnel in charge, involvement and availability – is unknown.
- Information gathered by other organizations is unknown.
- The communication among organizations is poor, infrequent or nonexistent.
- Plans and actions by other institutions are unknown.

Characteristic of the information quality:

- Different organizations have different information about the same facts.
- Each organization tends to trust in its own means of perception exclusively.
- The means to perceive and capture information are insufficient.
- The information access policies are inadequate.

The aim of the design presented in this work is to approach the difficulties caused by these factors.

2 Management system

The management system must give support during the emergency response phase in an integral way. This can be achieved by unifying the information necessary to make decisions, establishing communication among the groups of different organizations that must interact, and offering joint logistic planning and coordination tools.

2.1 Description

The management system is an IT system that collects all information available for an organization and sets tools to manage, profit from and share that information with other institutions.

The system has a client–server multi-user architecture, uses an Internet connection and is designed to be failure resistant by means of a server ring with data redundancy and load distribution.

By means of a proper profile setting policy and adequate information access permissions, the system offers the users the relevant tools to carry out the tasks pertaining their functions. The operations respective to each specific user responsibility are grouped in separate components.

Different types of information from different sources are integrated with a geographic information system in Situation databases without preventing the access to the data being updated.

2.2 Users

The information goes through the system in the shape of reports (GIS, Message, Form, News, Request entry, Update, etc.).

Each organization has a data base named Situation database, accessible to all its groups. Every time this database is modified, all users connected to it through the Internet will see the information update. Every organization can share and publish information. Note that, for the system purposes, the organization only operates through its groups. The only IT element that belongs to the organization as a whole is its Situation database.

Published information may be viewed by all organizations but can only be modified by the source one. Shared information can be edited by more than one organization. Only the source organization can access the information which has not been published or shared.

The system includes the set of organizations from which several Internet-connected groups with specific responsibilities depend. Seen as a whole, this system should be able to receive: messages, request entries, updates (about assets, requests, phenomena, etc.), forms, forecasts, alerts, and news. It should be able to send: messages, commands, plans, decisions, forecasts, alerts.

Each group of an organization can only access to the Situation database information related to its responsibilities. These groups may have their own database for assessment purposes that enables to run phenomenon evolution models to analyze possible developments of the events.

Each system user belongs to a particular group, each one of them with very different responsibilities, such as: entering support requirements, entering human or material assets, planning involvements, updating requirement status, updating requirement status, writing reports, sending messages, generating forecasts, broadcasting alerts, issuing orders, monitoring phenomena, assessing risks, registering news, generating situation letters, assessing statistics, etc.

The tasks carried out by a group during the emergency management can be distributed among several users with the same profile, which is determined by the group responsibilities. Although all users of the same group shares the same profile, each user may be devoted to different activities or the same activity applied to different regions.

The profiles gather several interface modules called components, highly independent from each other, that enable the group to carry out the activities comprising its responsibility.

The following list proposes a number of components and their offered operations:

Message service: Viewing, sending, receiving, searching and printing messages.
News: Viewing, writing, importing, publishing, searching and printing news.
GIS: Viewing, modifying, importing, exporting, searching, monitoring, sending and receiving geographical table and document information.
Chat: Viewing, sending and receiving instant messages.
Assets: Adding, removing, modifying, searching and consulting information related to the status, location and involvement of the available assets to respond to the emergency (e.g. consumables, personnel, vehicles, facilities, etc.).
Requests: Entering the request, involve assets, issuing orders, monitoring progress, and entering updates.
Users: Creating, deleting, modifying and consulting information about users.
Groups: Creating, deleting, modifying and consulting information about organization groups.
Statistics: Creating, viewing and printing management statistics.

A permission policy should be offered to limit the viewing and modifying of information only to groups who have that responsibility. This allows the different decision-making levels to access more or less summarized information and therefore avoid data excess. On the other hand, this policy allows keeping the layer modification access restricted.

The system should trigger visual or sound alarms when a certain piece of information requires the user's attention. The use of permissions ensures that the information received by a user will be relevant to his/her responsibilities.

2.3 Information organization

Each organization has its own Situation database that consists of the information available to that organization at the beginning of the emergency management. Part of this information might be modified throughout the emergency by means of updates from relevant groups. This data base could include the following initial information: bridges, settlement points, international borders, internal borders, road network, river network, train network, police stations, firefighters, hospitals, schools, evacuation centers, storage centers, contour lines, risk maps, optic and radar satellite images (in normal status); assets: vehicles, equipment, personnel, consumables, mattresses, medicines, blankets, drinking water, requests; phenomenon information: flooded region, land saturation, digital elevation model, optic and radar satellite images, earthquakes, faults, tectonic plates, volcanoes, spills; communications: messages, news, geographical reports, forms.

After determining the validity, opportunity and completeness of the new pieces of information, the profiles with data base modification capacities update the Situation database.

Dynamic information is a set of tabular data and layers that describe the current status of the situation. This information remains valid because of its capacity of change and describes reality aspects evolving throughout the whole management process. This information is used to learn about the current status of the emergency and make decisions accordingly. Dynamic information can be modified, shared and published. By sharing dynamic information, other organizations will have the capacity (and the responsibility) of modifying it; the system is in charge of synchronizing it. By publishing dynamic information, other organizations may see it but can't modify it.

Static information is formed by the set of received and accumulated reports which modified the situation. Since they have a specific date and time of issuing, they are just event notifications, that can't be modified and may quickly lose its validity. When these reports are applied to the Situation database, they change the dynamic information status. Static information can be generated, sent and received. When static information is sent, a report is distributed among the relevant groups and organizations. Static information will not change the dynamic information of the Situation database until it is applied to it.

Only published or shared dynamic information uses automatic synchronization tools to provide other organizations with continuous access to up-to-date information.

2.4 Architecture

The system is developed following a client–server approach. Permanent workstations, like those for monitoring or risk assessment, have a dedicated client application connected to the organization server. Temporary workstations, like update entry or order reception stations, can access the system through a web browser. The system uses a dynamic number of servers. There has to be communication among all servers to maintain information redundancy, synchronize shared or published information layers, ensure system operational capacity in the event of failures and spread processing load among them during the emergency response.

2.5 Cooperation and coordination

Communications go among participants in the shape of reports. This way the system can save a log with every communication, message, decision and update, which will enable a future analysis of the efficiency of procedures or a study of the relevance of modifying them.

Messages are viewed by integrating a standard e-mail server to the system. This way the user can have their mailbox and other components necessary to comply with their functions during management process in a single tool.

Users in charge of media exploitation have a tool to publish articles from different sources (press, radio, TV, digital media, etc.) relevant to the emergency management in the system. These articles will be displayed to the users by means of the news component, which has a digital newspaper structure and allows the user to access news from previous days and search articles by both title and body.

The system accepts both vector and raster geographic information in standard formats. When published geographical information is modified, every user with permission to view that information will receive those updates and the GIS component will activate an alarm and show the updated information.

The forms are pieces of text information entered in a specific format to be able to be automatically processed and then generate statistics that evolve with the reception of each new form. For instance, if a daily report detailing the number of evacuated people classified by gender and age is requested, the user in charge of filling it in will enter the data in a numeric format, and once the form sent, the system will be able to add them automatically and process them accordingly.

The request entries are records of necessities communicated to the operators (by phone, radio, etc.) whose responsibility is to enter them to the system. Once the request has been entered, an alarm will be triggered and the users in charge of analyzing the request will be able to devote assets to act to it. In order to do so, the users will need to have access to the assets future engagements and know their availability status. Since assets information includes information about their personnel in charge, the user will be able to send the orders directly to the individuals in charge of each asset for them to act to the request.

2.6 Phenomena analysis

The participating groups may have the responsibility of maintaining and running mathematic models to analyze the possible evolution of a phenomenon in the near future.

A model is an algorithm that, from the current data base status, processes the received stimuli to generate the following status of that database. These models represent the relation between a phenomenon and certain elements common to every natural or anthropogenic emergency, such as evacuation, rescue or medical assistance requests; damages and operability of facilities and equipments; necessity of mattresses, blankets, medicines, food and drinking water; injured, missing and deceased people; security problems in health centers; evacuation; storage, etc. The models may be used to study the more likely evolution of the situation during emergency management.

The result of this process gives the opportunity to prepare a response for future events. Since experts in different areas have the knowledge about these models, the system offers a generic language that enables the description of a large number of mathematic models that can be loaded and installed in a model server and used taking the Situation data base as its basis. As this data base consists of the best current information the organization has about the emergency, the system should offer the groups the possibility of working on a copy of these data which can be modified without affecting the information accessible by the other groups. This new database owned exclusively by the group is called Study, and all its layers may be copied from the Situation database and modified freely.

Models may be used for the generation of forecasts and alert bulletins to be sent later to other parts of the management process as reports.

The better the aspects involved in a natural or anthropogenic emergency are represented, the more useful the forecasts will be.

3 Training system

An important part of our proposal is to think of the training system as the management system immersed in a simulated environment. This way, the management system is also used during training, offering the following advantages over having two independent systems:

- The user is familiar with the interface before the emergency management process.
- Upgrades in the management system are immediately available for training.
- The management system is constantly tested.
- The training cost and effort are lower.
- The system maintenance cost and effort are lower.
- The simulation tools and the management system are highly independent.
- In the training system, the trainees do not perceive the directors controlling the exercise, bringing a complete immersion in the simulation.

In the training configuration, the simulation server and the exercise director application are added to the servers from each organization and the group applications.

The simulation server is in charge of supporting the event tree, triggering the events and its associated reports, enabling the directors to control and monitor the execution of the exercise and applying changes to the Reality data base, which describes constantly the current status of the simulation.

As in the management system, the profiles enable the customization of the interfaces for each user – director, trainee, observer, etc.
In the director profile, the following components are included:
Exercise management: to create exercises, prepare them and control their execution.
Events: to schedule the event tree ramifications, events, reports and changes.
Tables: to manage users, groups and profiles.

3.1 Users

During the exercise, the trainees represent the groups that take part in a crisis management, preferably with the same responsibilities they would have during a real crisis. The directors control the delivery of reports to the trainees, the simulated time passage and the phenomena evolution.

The director interface must be able to send reports impersonating groups interacting with the trainees during the exercise. The trainees do not perceive any difference between information from other trainees or from the direction, since all data are treated in the same manner.

3.2 Data organization

The Reality database simulates an emergency evolution in a dynamic way. It is the origin of the information in the reports that are sent to each group and update

the Situation database of its organization. Note that, since these reports are not necessarily complete, truthful nor timely, the Situation database of each organization may show different information, as it is the case in a real emergency. Trainees should use the management system capacities to view, modify, share and publish the information of their Situation data bases in order to form collectively a depiction of the situation as close to the simulated reality as possible.

In a real emergency management, each organization has means, external or of their own, to perceive changes and enter updates. In training, these means of perception – inspectors, reconnaissance flights, damage assessment equipments, victims, involved personnel, etc. – are simulated through IT agents called perceptors. The behavior of the perceptor depends on its type and the orders received from the trainees. For instance, an optical satellite perceptor could add cloud clusters and flooded areas to a clear satellite image using the Reality data base data for a time T, every N hours, and send the report to the group that requested those images with a delay of D hours. On the other hand, the inspector perceptors could send updates about the area to which they were assigned with no delay, but with greater spatial limitations. This way, the system enables the training of operators to make decisions with uncertainty, since each organization has a subjective view of the situation.

3.3 Architecture

To carry out training exercises for emergency operators, the management system receives reports from a simulation server. This way, the operation of the management system is not altered by being immersed in a training context and can offer its capacities – geographical representation, document organization, information exchange, logistic planning – in a genuine manner. Director applications offer the possibility of creating and modifying the events of the event tree, impersonating the groups not represented by trainees, monitoring communications and modifying the Reality data base directly or by means of stimuli sent to the model server to keep or change the course of events. Many directors may work on the exercise data at the same time, so an adequate concurrence policy must be implemented.

3.4 Exercise direction

For the preparation of an exercise, an educational goal at which the direction aims should be stated first. Next, a list of participating groups should be included, together with the designation of a trainee or a director as the impersonator for each group. The system maintains an event tree, that is to say, a tree-shaped event structure with multiple courses of events to richly describe the foreseeable development for the exercise. The tree events include changes for the Reality database, reports to be sent to groups and stimuli for mathematic models. The better the aspects involved in a natural or anthropogenic emergency are represented, the more realistic the exercises will be. The evolution of a phenomenon can also be described by simply scheduling consecutive events that change the phenomenon state.

Exercises already prepared may be totally or partially reused, progressively simplifying the process of preparation of new exercises.

The director interface controls the simulated time passage, enabling time jumps, changes in the time passage rate, and stopping and resuming time passage.

While the exercise is being executed, events can still be added, deleted, modified, activated and deactivated to adjust the development to the educational goals of each exercise session.

4 Conclusions

The dual system for management and training reduces the negative effects of the factors exposed before in the following way:

Characteristic of the nature of the phenomenon:

- The use of the system will derive in a greater knowledge of the phenomenon status, since information acquired by any group can be shared immediately, and changes in a greater spatial extension can be known.
- The system increases the possibilities for forecast generation by means of the execution of mathematic models.
- Forecasts published in the system are seen by the participants automatically, without having to make additional operations or queries.
- Characteristic of the organization:
- The system enables the organization to have a greater access to information – existence, location, status, personnel in charge, involvement and availability – about its own assets.
- The system enables a greater access to information – existence, location, status, personnel in charge, involvement and availability – about the assets of the other organizations.
- The practice of crisis management exercises aids in the writing, testing and updating of procedures for real emergencies.
- The use of the system reduces to a great extent the difficulties to do emergency management exercises.
- The system will be a progressive aid in the determination of functions and responsibilities of each user during an emergency, and the statement of adequate information access policies.
- Characteristic of the communication among organizations:
- The system enables the access to information obtained and published by other organizations.
- The system improves the communication among organizations.
- The system provides information about the plans and actions by other organizations to coordinate joint operations.
- Characteristic of the information quality:
- The system enables the organizations to access more information that is more up-to-date.
- The system enables its users to receive updates captured by means of perception owned by different organizations.

- The system increases the information sources available for each organization.
- The system can maintain adequate information access policies by means of profiles and permissions associated to each group.

Dealing with an emergency situation poses a great challenge, and its resolution involves coordination, rationality, assertiveness, efficiency and speed in the decision-making process. To increase the likelihood of success, it is extremely important that an adequate training program be offered that is close to reality, adaptable to multiple organizations and different scales, thus giving the opportunity to teach both the tactical-operational and the strategic levels and without removing the personnel from their workplace. This enormous responsibility is significantly relieved when a dual system with the characteristics mentioned above is available. By the use of the capacities of the proposed design, we are convinced that we will have more and better opportunities to survive a crisis satisfactorily, with the consequent benefits for the society.

Acknowledgements

Translation credits: Romina L. Ferraro and Patricia N. Rosenberg.

References

[1] Borkulo, E. van, Scholten, H.J., Zlatanova, S. and den Brink, A van, 2005. Decision making in response and relief phases, *Geo-information for disaster management.*

[2] Bowers III, F. A. and Prochnow, D. L., 2003. JTLS-JCATS Federation Support for Emergency Response Training, *Proceedings of the 2003 Winter Simulation Conference.*

[3] Mendonça, D. and Fiedrich, F. 2004. "Training for Improvisation in Emergency Management: Opportunities and Limits for Information Technology." *International Journal of Emergency Management*.

[4] Turoff, M., Chumer, M., Van de Walle, B. and Yao X., 2004. “The Design of a Dynamic Emergency Response Management Information System (DERMIS)”, *The Journal of Information Technology Theory and Application (JITTA).*

Section 7
Water resources

This page intentionally left blank

The water crisis in southern Portugal: how did we get there and how should we solve it

L. Nunes[1], J. P. Monteiro[1], M. C. Cunha[2], J. Vieira[2], H. Lucas[3] & L. Ribeiro[4]
[1]*University of Algarve, Portugal*
[2]*University of Coimbra, Portugal*
[3]*Águas do Algarve, S.A., Portugal*
[4]*Technical University of Lisbon, Portugal*

Abstract

Until very recently, the public water supply in the Algarve region was almost entirely supported by groundwater wells. However, in the last years of the 20th Century, the Portuguese government defined a scheme for the public water supply sector entirely based on surface water from large dams, in order to guarantee the public water supply. The efforts to abandon groundwater as a source for public supply started in 1998, after a large investment in new infrastructures and rehabilitation of some existing ones. However, the practical implementation of this water supply scheme showed that an integrated resource management is needed in order to implement a more economical and reliable solution. The present paper describes the historical background and the evolution of water use in recent decades until the present time in the Algarve, and a proposal for restructuring the management of the water supplies based on the development of a decision support system within an integrated water resources management scheme.

1 Introduction

The Algarve region (5400 km^2) is the southernmost province of Portugal, as indicated in Figure 1. The region is characterized by a warm Mediterranean climate. A mean annual precipitation of 653 mm was calculated for the period 1941/42-1973/74 [1]. Unfortunately the precipitation regime is not regular,

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060431

having intermittent periods with short and sharp floods in the winter and a long dry period in the summer. In addition, there may be extreme events such as inter-annual periods of drought.

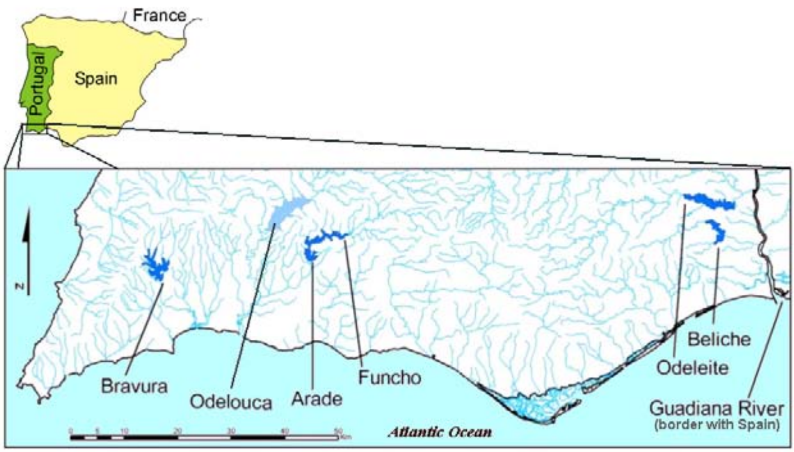


Figure 1: Location of dams in the Algarve region (existing dams in dark blue; dam under construction in light blue).

With such a climate, an integrated management of water resources at the basin scale is needed. With regard to the joint use of groundwater and surface water, Burt [2] remarked: "the inherently random nature of surface water supplies and the natural recharge to an aquifer give groundwater stocks an important role as a contingent supply for times when surface water stocks (*i.e.* dams) are below average. Additionally, optimal intemporal allocation of groundwater used conjunctively with surface water will impute a higher value to the surface water than it would have in an unmanaged basin". Integrated management schemes improve feasibility and reduce systems' vulnerability, and they may also improve the quality of the delivered water and offer important advantages when facing increasing demands, delaying the construction of new and costly infrastructures [3].

At present demand in the region is in the order of 260×10^6 m^3/year, for agriculture (69%), public water supply (27%) and golf activity (10%). This value corresponds to a dramatic rise in the last 40 years, mainly thanks to an increase in the irrigated area (currently about 200 km^2) and tourism (about 10 million visitors per year with an average stay of about 10 days).

Agriculture and golf mostly use groundwater, with much less consumption of surface waters. In the opposite situation, and according to the present policy for the sector adopted by the Portuguese central administration, the public water supply should be guaranteed only by surface waters from large dams (Figure 1 and Table 1). This is a high risk situation for the public water supply system, and its consequences were felt in the last drought that affected Portugal in 2004 and

2005. The dams were almost depleted and restrictions on water use were imposed in all sectors. The economic damage was not greater because groundwater played an important role in helping to cope with the situation.

Table 1: Some characteristics of the large dams in the Algarve region (note: the Arade and Funcho dams are in the same water course).

Dam	Year	Height (m)	Watershed (km^2)	Type	Maximum storage ($\times 10^6 m^3$)
Arade	1956	46	224	embankment	27
Bravura	1958	41	75	concrete	32
Beliche	1986	54	99	embankment	48
Funcho	1993	49	211	concrete	43
Odeleite	1996	65	352	embankment	117
Odelouca	2010(?)	96	385	embankment	135

Table 2: Some characteristics of the 17 aquifer systems identified in the Algarve [4].

Aquifer system	Dominant lithology	Area (km^2)	Mean recharge ($\times 10^6 m^3$/year)
M1-Covões	L, D	22.56	6
M2-Almádena - Odeáxere	L, D	63.49	16-24
M3-Mex. Grande - Portimão	L, D, S	51.71	10
M4-Ferragudo - Albufeira	S, L, M	117.10	8
M5-Querença - Silves	L, D	317.85	53-87
M6-Albufeira - R. Quarteira	L, D	54.55	9
M7-Quarteira	L, D, S	81.19	12
M8-S. Brás de Alportel	L, D	34.42	5-6
M9-Almansil - Medronhal	L, D	23.35	6-7
M10-S. J. Venda - Quelfes	C, S, L, M	113.31	9
M11-C. Cevada - Q. J. Ourém	L, D	5.34	2
M12-Campina de Faro	S, L	86.39	10
M13-Peral - Moncarapacho	L	44.07	10
M14-Malhão	L, D	11.83	3
M15-Luz - Tavira	L, S	27.72	4
M16-S. Bartolomeu	L, D	10.60	3
M17-Monte Gordo	S	9.62	3

Legend: C-clays; D-dolomite; L-limestone; M-marl; S-sand.

Groundwater stocks are by far the biggest reservoir in the region and current knowledge about the hydrogeology of the Algarve allows the identification of 17 aquifer systems at the regional scale [4]. The carbonate rocks (mainly Jurassic and Miocene period) are the support of the most important aquifers, both by their extent and by the volume of stored water (Figure 2 and Table 2). Though ground

waters in the carbonate rocks have a high natural degree of hardness, it has already been noted that their joint use with softer surface waters may have important advantages in treatment requirements [5].

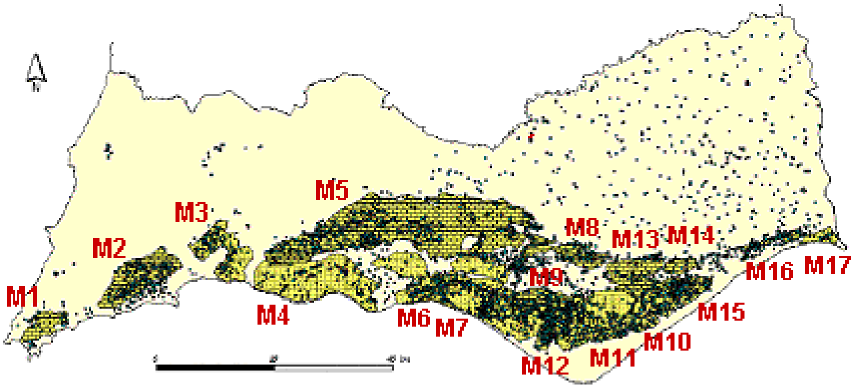


Figure 2: Location of dug wells and boreholes and geometry of the 17 aquifer systems with regional expression in the Algarve.

Recent studies conducted by the company Águas do Algarve S.A. (regional water utility for the public supply) include the use of both surface and ground waters, and the possibility for using desalinated water and treated wastewaters for purposes compatible with its quality The company wants to consolidate the water system to be able to respond to the different possible scenarios, including extreme events like long periods of drought that may periodically happen. Huge investments are required to implement and to operate the different technical solutions. Cunha [6] remarks that planners should use the proper methodological tools to obtain the best response (or a set of good alternative responses) to such problems. The author also argues that decision models can play an important role in dealing with such situation by simultaneously integrating all the relevant aspects (physical, hydrological, technological, financial, etc.). In Section 2 of this paper the authors present a general overview of the changes in the water use during the last 50 years, and in Section 3 they briefly describe the decision support system that is being developed to make the public water supply system more robust and resilient.

2 Evolution of water use in the Algarve during the second half of the 20th century

Until the midway through the 20th century water demand in the Algarve region was entirely supported by mechanical means. Groundwater was exploited through hand-dug wells with large diameter and depths usually less than 30 meters. Surface waters were exploited in small earth dams and water was channelled through sluices. The Bravura and Arade dams (Figure 1 and Table 1),

built in the 1950s, are the first man-made structures supporting the use of surface waters with expression on a regional scale. The water in these two dams was used initially for the irrigation of agricultural areas in two public areas of 23 km^2 and 17 km^2. In the 60s, a period of exponential rise in the water demand began, mainly due to the introduction of many and new irrigated agricultural areas and to the growth in tourism. Rough calculations indicate that by the end of the 20th century demands were ten times higher than in the early 50s.

It was only possible to meet increasing demands thanks to the introduction of drilling technology that allowed the construction of thousands of boreholes in the region. These boreholes, often more than 100 m deep, were implanted without the support of adequate regional planning, that should have been based on the existing knowledge of the hydrogeology of the region. Despite the efforts of the regional office of the National Water Authority, the inventory of the existing boreholes is still far from being exhaustive, although more than 8000 wells and boreholes have already been mapped (Figure 2). At that time, the public water supply was locally and independently dimensioned by each of the 16 municipalities of the Algarve. Only a few of the boreholes operated by the municipalities were established taking in account the characteristics of the aquifers and backed up with pumping tests and technical reports. However, most of the boreholes, in particular those that were built for agriculture, were implemented without any technical or scientific support.

The intensive augmentation of groundwater use during this period was certainly one of the reasons contributing to the large number of scientific contributions related to the characterisation of groundwater resources in the Algarve region. The first contributions to the characterization of groundwater at a regional scale appeared in the early 80s [7–9]. However, due to the lack of technical capacity, those studies did not lead to visible improvements in the efficiency of the regional water management. In practice, the knowledge contained in those works was not used to support current water management. Increasing pressures on the groundwater resources resulted in the exposure of serious problems. High levels of groundwater abstraction were at least partly responsible for high nitrate contamination in irrigated areas [10] and for the displacement of the fresh water-salt water interface in some sectors of coastal aquifers [11].

These situations, together with the relentless increase in water demand, led to the present policy defined by the Portuguese government, which consists of the construction of a multi-municipal system, entirely based on surface waters from large dams to guarantee the public water supply and to replace all the individual municipal systems. The efforts to abandon groundwater as a source for public water supply started in 1998, after a large investment in new infrastructures (Funcho, Odeleite dams; water treatment plants, regional distribution system) and in the rehabilitation of others ones (Beliche and Bravura dams). The Arade dam is used only for irrigation. One dam is missing to complete the system: the Odelouca dam. This dam should already have been built but a series of delays related to the environmental impact assessment procedure have reduced its maximum capacity and postponed its commissioning until 2010, at least. The

Odeleite and Beliche dams support a multi-purpose system. Besides public water supply these dams allowed the creation of a new irrigation area of 81 km^2. This new area, together with the aforementioned area irrigated by the Bravura and Arade dams, allows a total of about 120 km^2 to be irrigated with surface waters.

At a first glance, the immediate benefits associated with the implementation of the multi-municipal system seem evident. From the environmental point of view, it was often claimed that reducing the exploitation of the coastal aquifers, plus with the abandoning of many boreholes used for public supply in urban areas, would lead to the attenuation of the effects of saltwater intrusion. However, as water use for urban supply constitutes only 27% of the total water use, this is not necessarily true, because the effects of these phenomena are strictly conditioned by the pumping schedules at the scale of the aquifer sectors where boreholes were in use. The major benefits were felt in the quality of the water delivered for human consumption and in the control of water quality because all surface water is properly treated for use, and it is far easier to control water quality in a few dams than in a large number of boreholes.

Even without the Odelouca dam, which will be exclusively operated by Águas do Algarve for the public water supply, all needs were satisfied until the last long drought that occurred in 2004 and 2005. In that period all the dams were almost depleted and restrictions on water use were imposed for all economic sectors, with exceptional measures being taken to avoid the total disruption of the public water system. When the public water supply could not be ensured by surface waters "emergency boreholes" were made and the old municipal boreholes were reactivated. Unfortunately, the last solution was not completely successful. In too many cases, the infrastructures (*i.e.* boreholes, pumps and pipes) had been abandoned and were in a bad state of repair. Figure 3 shows some images that confirm the abandonment of infrastructures that represented huge investments in the relatively recent past. One may think that if the Odelouca dam had already been built there would not have been any problems meeting all demands in the last drought. With a maximum storage capacity of 135×10^6 m^3 and mean annual inflows of 121×10^6 m^3, the Odelouca dam will contribute significantly to solving major water scarcity problems and will release the Funcho and Bravura dams for almost exclusive irrigation purposes. However, the measured inflows to the Odelouca dam are between a minimum value of 4×10^6 m^3/year and an extreme value of 309×10^6 m^3/year. The inflows series show that there will be many years when the demand cannot be satisfied. Even with inter-annual regularization due to the construction of the dam, it will be hard to meet the demand in the long term (projected to be 75×10^6 m^3 in 2025) in long periods of drought.

As claimed by Monteiro and Costa [12]: (1) after a period (in the past) in which public water supply was almost entirely supported by groundwater; (2) there is a (present) period characterised by large investments in replacing groundwater by surface waters from dams; (3) but there will be a (future) period where the prevailing hydrological conditions and the conflicts between users will force the administration to define an integrated management scheme to guarantee the public water supply at basin scale.

Figure 3: Abandoned sluices, pumping stations and boreholes (São João da Ourém, March 2006 – Photos by José P. Monteiro).

3 A decision support system for the public water supply system of the Algarve

From the description of the evolution of water use in the Algarve, it can be accepted that the public water supply system in the Algarve must include new water sources under an integrated resources management scheme. New infrastructures have to be sited, designed and operated in order to have a more robust and resilient system. To accomplish those tasks, the development of a decision support system based on an optimisation model is being built in conjunction with regional water supply utility (Águas do Algarve, S.A.).

The decision-support system should help with the following decisions:

- Different sources to be mobilized: surface water, groundwater, desalinated seawater or even treated wastewater;
- The location and the size of water intake and water distribution infrastructures;
- The location, size and technologies for water treatment, wastewater treatment, water desalination infrastructures;
- The operating policy of such infrastructures for different scenarios.

The framework for the development of the decision-support system is presented in Figure 4 and it will be designed to promote the efficient, economic and sustainable use of the different sources. The decision model is being built for cost minimization (the cost functions are being tailored to take into account the Water Framework Directive recommendations [13] on environmental costs, recovery costs, etc.) of interventions, but the decisions taken will have to be sustainable and promote the efficient distribution of water among users. This can be achieved by considering various constraints. The fulfilment of such constraints is evaluated through the use of different simulation models:

- A groundwater simulation model to allow the effects of the pumping policies to be assessed, both in terms of costs and also considering that one is working in a coastal area and saltwater intrusion should be avoided;

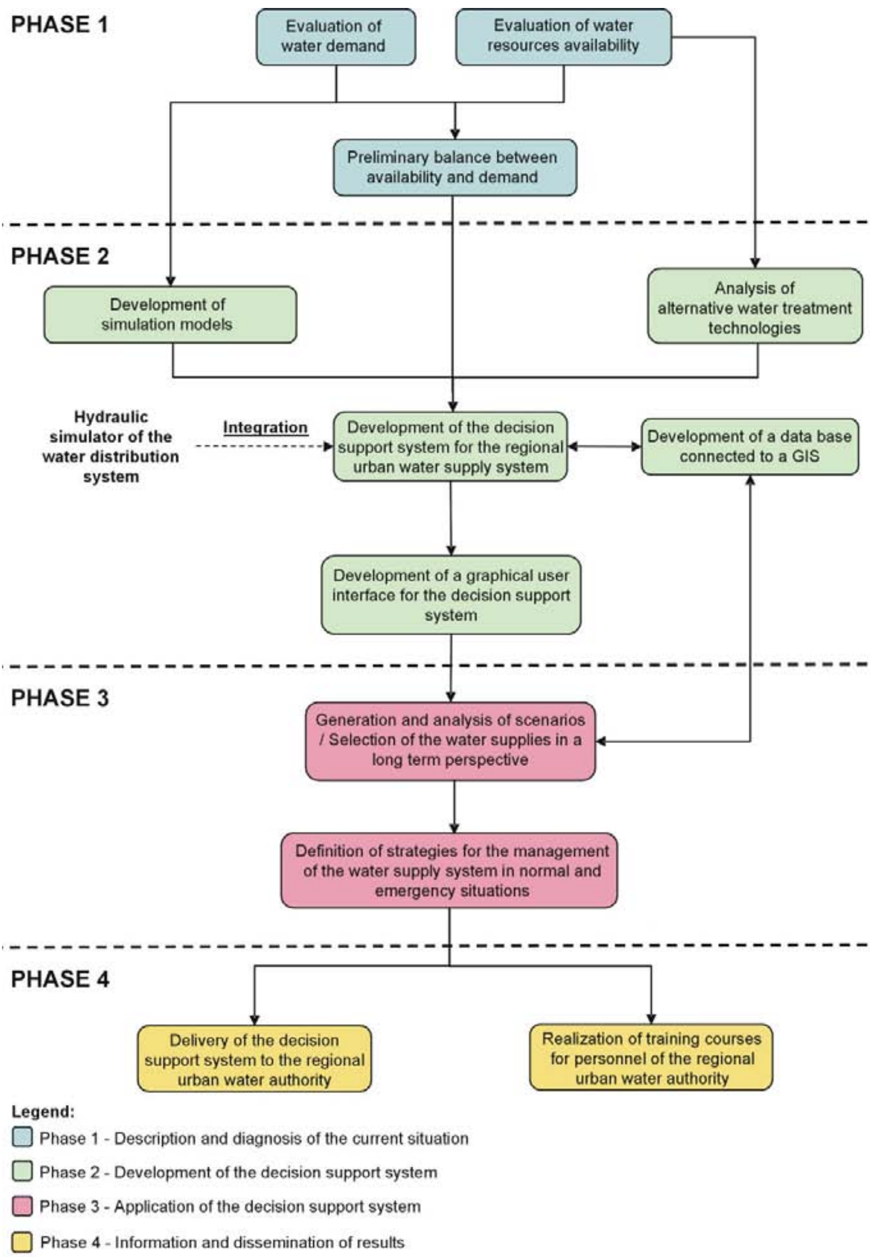


Figure 4: Framework for the development and application of the decision support system.

- Simulation models for dams, to quantify inflows and outflows for different climate scenarios (e.g., due to induced climate changes, and during drought periods);
- Water quality simulation models that will allow the definition of the percentage of each type of water in the final mixture to be supplied, because of the different physical and chemical characteristics of the water coming from surface or groundwater origins;
- Water distribution simulation models to allow assessment of the performance of the distribution system for different scenarios (e.g.: pressure, velocity, etc.).

The integrated model is a non-linear mixed integer one, and a new algorithm based on heuristics optimisation techniques is under development.

4 Conclusions

The analysis of the evolution of the water use in the Algarve during the second half of the 20th century and the casuistic interventions of the central water authority shows the reasons for the water crises faced by the Algarve region in times of drought. History has shown that water supply cannot rely on a single origin, surface or groundwater (or even desalination, or treated wastewater). In this article we have presented a decision-support system that will help to manage the allocation of resources (extractions from each source) by one of the major stakeholders in the region. It is hoped that future (next three to four years) operational results will make the remaining stakeholders and central administration follow the path of integrated water resources management, by including all stakeholders and all available sources, with the least environmental impact, lowest social cost, and greater sustainability.

Acknowledgment

Funding for this work is being provided by Águas do Algarve S.A.

References

[1] Loureiro, J.M. & Nunes, M. F., *Monografia hidrológica do Algarve* [*Hydrologic monograph of the Algarve*], Direcção Geral dos Recursos e Aproveitamentos Hidráulicos: Lisbon, Portugal, 1980.

[2] Burt, O.R., Groundwater management and surface water development for irrigation. *Economic Modeling for Water Policy Evaluation*, ed R.M. Thrall et al., North-Holland: New York, pp. 75-95, 1976.

[3] Pulido-Velázquez, M., *Optimizacíon económica de la gestión del uso conjunto de aguas superficiales y subterráneas en un sistema de recursos hídricos* [*Economic optimization of conjunctive use management of

surface and groudwaters], PhD dissertation, Universidad Politécnica de Valencia, Valencia, Spain, 2003.
[4] Almeida, C., Mendonça, J.L., Jesus, M.R. & Gomes, A.J., Sistemas aquíferos de Portugal Continental [Aquifer systems in Portugal mainland], INAG, Lisbon, Portugal, 2000.
[5] Campinas, M., Lucas, H. & Rosa, M.J., O tratamento conjunto de agues subterrâneas será vantajoso? [Will it be advantageous the conjunctive treatment of ground and surface waters?], *Proc of VI Water Congress*, Oporto, Portugal, 2000.
[6] Cunha, M.C., Water systems planning: the optimization perspective, *Engineering Optimization*, **3**, pp. 255-266, 2003.
[7] Trac, N. Q. (Coordinator), *Evaluation des resources en eaux des systèmes aquifères de l'Algarve*, Technical Report, UNESCO/PNUD, 1981.
[8] Silva, M. O., *Hidrogeologia do Algarve Oriental* [*Hydrogeology of eastearn Algarve*], PhD dissertation, Faculdade de Ciências da Universidade de Lisboa, Lisbon, Portugal, 1984.
[9] Almeida, C., *Hidrogeologia do Algarve Central* [*Hydrogeology of central Algarve*], PhD dissertation, Faculdade de Ciências da Universidade de Lisboa, Lisbon, Portugal, 1986.
[10] Stigter, T., Van Ooijen S.; Post, V., Appelo, T.; Dill, A., Hydrogeological and hydrochemical explanation of the groundwater composition under irrigated land in a Mediterranean environment, Algarve, Portugal *Journal of Hydrology*, **208**, pp. 262-279.
[11] Carreira, P.M., *Mecanismos de salinização dos aquíferos costeiros do Algarve* [*Salination mechanisms of Algarve coastal aquifers*], Laboratório Nacional de Engenharia e Tecnologia Industrial: Sacavém, Portugal, 1991.
[12] Monteiro, J.P. & Costa, M.S., Dams, Groundwater Modelling and Water Management at the Regional Scale in a Coastal Mediterranean Area (The Southern Portugal Region–Algarve), *Larhyss Journal*, **3**, pp. 157-169, 2004.
[13] EC, Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy, *Official Journal of the European Communities*, **L327**, pp. 1-72, 2000.

Water resources management under drought conditions

J. B. Valdés[1,4], J. González[2], J. Cañón-Barriga[3,5] & G. Woodard[4]
[1]*Department of Civil Engineering and Engineering Mechanics, The University of Arizona, Tucson, AZ, USA*
[2]*Department of Civil Engineering, University of Castilla-La Mancha, Avda, Camilo José Cela s/n, Ciudad Real, Spain*
[3]*Department of Hydrology and Water Resources, The University of Arizona, Tucson, AZ, USA*
[4]*Center for Sustainability of Semi-Arid Hydrology and Riparian Areas (SAHRA), The University of Arizona, Tucson, AZ, USA*
[5]*Universidad de Antioquia, Medellín, Colombia*

Abstract

The semi-arid/arid US Southwest and Northwestern Mexico are experiencing severe droughts while at the same time having significant increases in their water demands. This paper evaluates the demand and the efforts to manage water resources under drought conditions. Innovative water resources management techniques need to be implemented to address this scarcity which is aggravated under drought conditions. This paper also presents an application of a new drought index (Gonzalez and Valdes, 2006ab) in characterizing the spatial and temporal variability of droughts and its use in water resources management, particularly multi-reservoirs.
Keywords: droughts, semi-arid regions, water resources management, US Southwest, Northern Mexico.

1 Introduction

Water managers in the Southwestern U.S. face a number of daunting problems. On the water demand side, population growth in Southwest states is the fastest in the nation. Arizona and Nevada, the two most arid states, have been experiencing annual growth rates of 4% and 5.5%, respectively. Changing socio-

demographics, such as decreasing household size, exurbs, and second homes, are increasing per-capita consumptive demand and exacerbating peak seasonal demand.

Weather and climate trends pose another set of challenges. Urban heat island effects in large Southwestern cities mean warmer nights and longer growing seasons for landscape plants. This extends the irrigation season. Climate fluctuations, such as the eight-year drought impacting parts of the Southwest, have directly killed large areas of forests, and created ideal conditions for beetle infestations and wild fires, eradicating additional forested areas. These abrupt, large-scale land cover changes fundamentally alter the partitioning of precipitation between infiltration and runoff. While the long-term impacts of climate change are still uncertain, there is a growing consensus that earlier snow melt will effectively reduce the amount of water that can be stored in the spring in reservoirs for use in summer and fall.

Other economic, legal, and political realities are creating pressures to shift over-allocated supplies among existing and new water demands. These include greater interest in water-based recreation, endangered species protection, settling American Indian water rights claims, and meeting the needs of a resurgent copper mining industry.

The US Southwest and Northern Mexico are populated semi-arid regions that suffer the effects of long droughts with the consequent stress over their water resources. The significant temporal and spatial variability of water resources and related droughts have also a significant impact on water resources availability and management. Integrating this uncertainty into planning and decision-making is a key issue, both for operation optimization and risk management strategies. In this paper a new indicator of drought exceptionality, the Drought Frequency Index (DFI, González and Valdés, [4]) is used in the management of a complex system of reservoirs in the semi-arid Lower Rio Grande/Rio Bravo basin. The approach is compared to alternative operating policies developed, as part of a Decision Support System (DSS) by Gastelum et al. [1].

2 The Conchos River Basin System

2.1 The Conchos River Basin (Mexico)

The Conchos basin is one of six Mexican tributaries to the international Lower Rio Grande/Rio Bravo (LRGRB) basin. It has an area of 64,000 km^2, which represents 14% of the surface area of the LRGRB. The entire basin is shown in Figure 1.

Agriculture in the Conchos Basin accounts for 72% (2,536 Hm^3/yr) of the total water demand. The water demands come from three Irrigation Districts (IDs), (ID005 Delicias, ID090 Rio Florido and ID103 Bajo Conchos), and several Irrigation Units (IUs) located along the river with a potential irrigated area of 108,562 and 14,509 Ha, respectively.

Low water distribution efficiencies are one or the more significant problems of the IDs (~40%) and the IUs (~48%). Commercial, energy and mining water

requirements are 22% and domestic consumption is approximately 6%. The cities of Hidalgo del Parral and Delicias account for 87% of the total population. Runoff from the Conchos is controlled through a series of seven reservoirs (Table 1), with a combined storage capacity of 3,943 Hm^3, most of which is provided by the 2,903 Hm^3 capacity of La Boquilla reservoir. The reservoirs are conjunctively managed to provide irrigation, hydropower and flood control.

Figure 1: Lower Rio Grande/Rio Bravo [12].

Table 1: Reservoir characteristics Conchos Basin.

River	Reservoir	Storage Capacity Hm^3 (MAF)	Uses
Florido	San Gabriel	255 (0.21)	Irrigation District 103 flood control
Florido	Pico de Águila	50 (0.045)	
Conchos	La Boquilla	2903 (2.34)	Irrigation District 005 and Hydropower
Conchos	La Colina	24 (0.195)	
San Pedro	F. Madero	348 (0.28)	
Chuviscar	Chihuahua	26 (0.021)	Municipal, Irrigation, flood control
Conchos	Luis L. Leon	337 (0.29)	Irrigation District 090, flood control

2.2 The 1944 International Water Treaty

In response to the rapid development of the RGRB region in Mexico and the U.S. State of Texas during the early part of the 20th century; to the differing

institutional structures in the two nations and to the frequent conflicts over water due to variability of runoff, an international water treaty was negotiated in 1944 between Mexico and the United States which divided the waters of the RGRB system flowing between Fort Quitman, Texas and the Gulf of Mexico based on shares of runoff of the rivers.

The Treaty was designed to ensure that each nation had access to adequate water in drought years. The Treaty divides the waters of the RGRB equally while simultaneously requires Mexico to deliver an average of 431 Hm^3 per year over a 5-year period (TCPS [10]). One provision of the treaty allows Mexico to deliver less than 431 Hm^3 per year in cases of "extraordinary drought" as long as the deficit is made up in the following 5-year cycle. What constitutes "extraordinary," however, is not explicitly defined in the treaty. Interests on the U.S. side have interpreted it to mean that flow from Mexican tributaries must cease entirely for an extraordinary drought to have occurred (TCPS [10]).

Based on the absence of an accurate definition of "extraordinary drought," the DFI index was used to develop an operating policy for the reservoirs that takes into account the spatial and temporal characteristics of drought in the LRGRB.

3 Methodology

3.1 Drought areal-intensity-frequency characterization

Although the estimation of drought severity at a point gives generally useful information for water management, a more regional indicator usually is required to assess the drought for the entire basin. This regional drought analysis is useful for declaring the drought condition or determining the drought intensity during a particular year (Shin and Salas, [9]) and it is important to address the "extraordinary drought" concept of the 1944 treaty.

One approach to assessing the regional status of a drought is the drought severity-area-frequency curve, which was originally proposed by Henriques and Santos [5] and expanded by Kim et al [6]. This approach regionalized the Palmer Drought Severity Index (PDSI), using geostatistical techniques and applying extreme value analysis to the regional values. The drought intensity (DI) is calculated multiplying the cumulative PDSI in a dry spell by the probability of drought occurrence (Figure 2) for each year. In this way, each drought event is allotted evenly for a particular year (Figure 3).

The drought intensity-areal extent-frequency curve provides useful information which contains drought intensity and area subjected to drought for a given drought return period (Figure 4).

The study by Kim et al. [6], suggests that the Conchos River basin has experienced severe droughts in the last 60 years. Particularly it suffered an extreme drought in the 1990s based on the PDSI indicator. Using the drought intensity-areal extent-frequency curve, shown in Figure 3, the drought in the basin based on the intensity of the return period may be examined. Note that the drought intensity means time average of drought severity in dry spell and represent the drought severity classified by Palmer, which is indicated by a gray

scale between 0 and -6. Finally, in Figure 4, the historical droughts are compared with the drought intensity-areal extent-frequency curve, which was constructed using the recorded data.

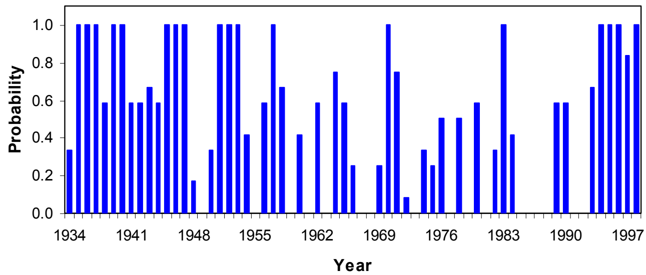
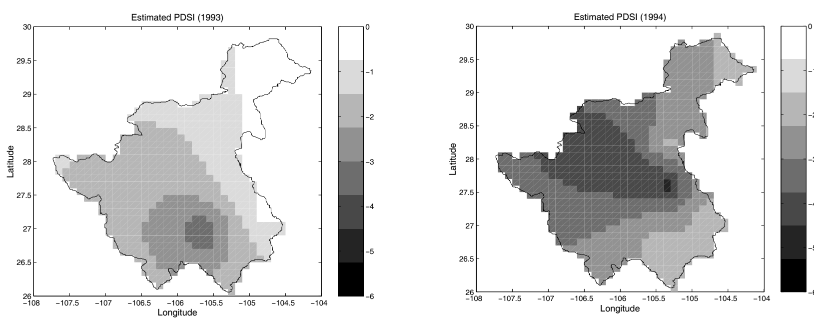


Figure 2: Annual probability of drought occurrence in the Conchos Basin (1934-1998) (from Kim et al, [6]).



(a) (b)

Figure 3: Areal distribution of drought intensity for historical droughts in the Conchos Basin (a) 1993 (b) 1994 (from Kim et al, [6]).

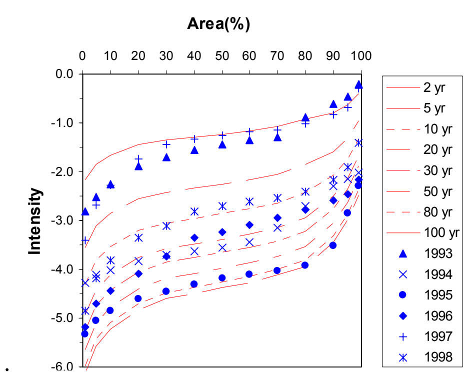


Figure 4: Drought Intensity-Areal Extent-Frequency Curve for the Conchos River basin for the historical droughts occurring in the 1990s (from Kim et al, [6]).

From this analysis, the droughts that occurred in 1993 and 1997 have an associated return period of 2-5 years; however, the droughts that occurred in 1994, 1996, and 1998 have an associated return period of 10-30 years. The drought that occurred in 1995 is the most severe drought the basin has ever experienced in the period of record. The data from after 1990 were not used to construct the drought intensity-areal extent-frequency curve, and the 1990's drought has a higher return period of more than 80 years. The drought intensity is close to a return period of 100 years with an increase in areal extent. In addition to the high return period, considering the lack of the data period, it is the most extreme drought in that more than 70% of the basin experienced a PDSI below –4 (extreme drought condition). Exploding water demands in the basin, as in most of the northern Mexico has contributed to 1990s droughts having severe impacts.

3.2 The Drought Frequency Index (DFI)

The PDSI and the Standardized Precipitation Index (McKee et al., [8]) are frequently used drought indicators, but they have some limitations. González and Valdés [3,4] presented an approach for the stochastic characterization of extreme hydrologic droughts according to their random nature. The approach is based on the characterization of random variable extreme persistent deviations, referred to as the variable's normal variation regime. The characterization is quantified in terms of the mean frequency of recurrence, providing the basis of a new drought index: the Drought Frequency Index (DFI).

The DFI index allows analyzing and evaluating droughts over any random hydro-climatic variable affected by droughts. To complement the approach, a methodology is proposed to analyze the spatial-temporal progress of a drought over a region by generating DFI maps and characterizing droughts from a stochastic point of view, based on their extraordinary persistence and areal extent. In the computation, the DFI algorithm searches and analyzes, for every time step, the period at this step which is the most extreme from the point of view of persistent lower deviation of the random variable. Each of these periods is characterized in term of its mean frequency of recurrence and this value is associated to the corresponding time step. The scale provided by the DFI is general, universal, and attend to the random nature of the phenomenon. Each application may use the DFI scale to define drought state, setting the threshold frequency according to its vulnerability, which will allow homogenizing drought definitions on a single scale.

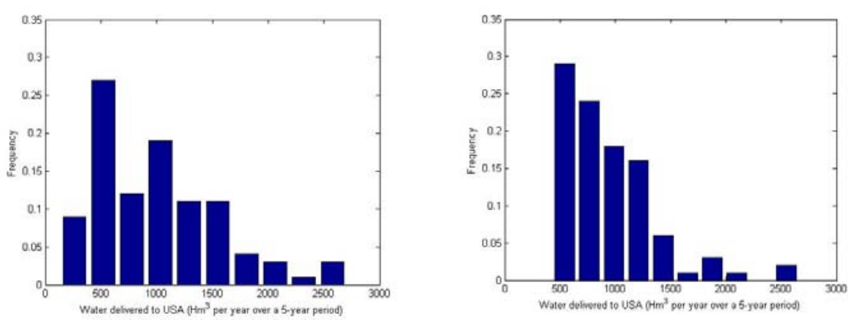
3.3 Operation of multi-reservoir system under drought conditions

A comprehensive review of the most recent application on the use of optimization techniques for reservoir operation is presented in Labadie [7]. Labadie acknowledges that a wide gap exists between operations research in academic or scientific communities and its real-world implementation in existing reservoir systems operations, where more simple and intuitive rules are preferred.

In this section the operation of the multi-reservoir system in the Conchos basin is operated under two alternative operating policies. One uses a linear optimization procedure that tries to minimize long-term deficits without using an explicit definition of drought. The second, an extension of the approach proposed by Gonzalez [2], utilizes the DFI as a criterion for determining releases in a given period. Both models were also compared with the preliminary results of the DSS developed by Gastelum et al. [1] that are described in Stewart et al. (2004). In both optimization models there was a priority in meeting the international treaty obligations during a 5-year cycle. However, the DFI-based operating policy considered the case of an "extraordinary drought" to postpone deliveries in a 5-year cycle to be met in the next 5-year cycle.

The DFI-based operation model has a two-step optimization. First a linear optimization is used to determine an approximate operating policy, which is then used to derive the operating policy utilizing the DFI indicator. For the particular case of the Conchos basin the definition of "extraordinary drought" was for a DFI value of 70 years return period. Below a DFI value of 70 years the deficits are significantly lower. As discussed in Gonzalez [2] using the DFI to define operating policies for multiple reservoirs is more attractive when the demand is more resilient to deficits, i.e., irrigation vs. municipal water supply.

The results were computed at the main irrigation districts (Rio Florido, Delicias and Bajo Conchos). Figure 5 shows the histograms of water deliveries to the US from both optimization models. Although similar, the DFI optimization performs better than the linear optimization rule in terms of the minimum requirements for deliveries to the US. Figure 5b shows that the DFI-based model is able to deliver the US requirements every 5-year period as required in the treaty while at the same time decreasing the number of uncontrolled releases in the system.

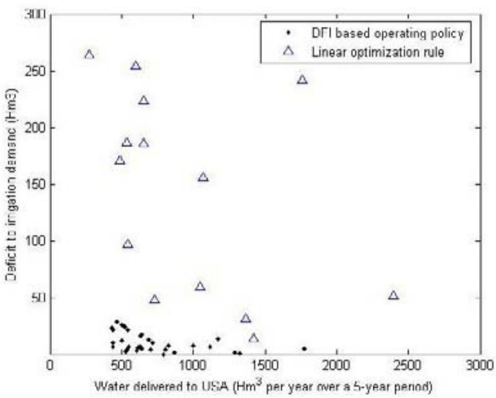


(a) linear optimization (b) DFI-based optimization

Figure 5: Histogram of water deliveries to US for both optimization models.

Figure 6 compares the cumulative deficits in the irrigation districts demands for 5-year periods with the releases to the US for the same periods. This figure shows that the DFI-based operating policy is able to significantly reduce the

magnitude of irrigation demand deficits for all districts, e.g., no irrigation deficit in the DFI-based operating policy is above 50 Hm^3/yr, significantly less than the deficits in the linear optimization policy. In contrast, the number of 5-year periods that experience some deficit with the DFI-based operating policy doubles that of the linear optimization.



	Optimization model	
	DFI	Linear
5-year periods with deficit (%)	34	14
Mean deficit (Hm^3)	1.05	142
Std dev deficit. (Hm^3)	0.8	89

Figure 6: Comparison of deficits to irrigation demands vs. supplies to the US.

4 Final comments

This paper describes a criterion to characterize droughts and its potential use in the management of complex water resource systems. The characterization acknowledges the multiple dimensions that define the severity of a drought (intensity, duration, and areal extent) while at the same time expressing them in a concept easier for water resources managers to understand, i.e., the return period. The temporal and spatial characteristics of a drought are represented by several criteria, e.g. areal-intensity-frequency curves and spatial DFIs. The non-parametric evaluation of the bivariate characteristics of droughts offers more flexibility than the more frequently used extreme value distributions. The use of the DFI as a criterion for the operation policy of reservoirs in the Conchos Basin proves to be advantageous over conventional optimization rules, promising reliable applications to similar single-purpose reservoir systems.

References

[1] Gastelum, J., J.B. Valdés and S. Stewart, A Decision Support System to Improve Water Resources Management on the Conchos Basin, submitted to *System Dynamics Review*, 2006

[2] Gonzalez, J, "Caracterización estocástica de la realización espacio-temporal de eventos hidrológicos extremos de sequías," PhD dissertation, Polytechnic University of Valencia, Valencia, Spain, 2005.

[3] González, J. and J.B. Valdés, "Statistical Process Control Based on Extreme Persistence Detection," submitted to *Journal of Quality Technology*, April 2006.

[4] González, J. and J. B. Valdés, "The new Drought Frequency Index: definition and comparative performance analysis, accepted for publication in *Water Resources Research*, June 2006.

[5] Henriques, A. G. and M. J. J. Santos, "Regional drought distribution model." *Phys. Chem. Earth (B)* 24, No. 1-2: 19-22, 1999.

[6] Kim, T., J.B. Valdés and J. Aparicio, "Frequency and Spatial Characteristics of Droughts in the Conchos River Basin, Mexico," *Water International*, 27(3), 420-430, 2002.

[7] Labadie J, "Optimal Operation of multireservoir systems: state of the art review," *Journal of Water Resources Planning and Management* 130(2):93–111, 2004.

[8] McKee, T.B., Doesken, N.J., Kleist, J., 1993. The Relationship of Drought Frequency and Duration to Time Series. 8th Conference on Applied Climatology, Anaheim, CA, 179-187.

[9] Shin, H.S. and J.D. Salas, "Regional Drought Analysis Based on Neural Networks." *J. Hydrologic Engineering* 5(2), 145–155, 2000.

[10] Texas Center for Policy Studies (TCPS), (2002). The dispute over shared waters of the Rio Grande/Rio Bravo: a primer. Available at http://www.texascenter.org/borderwater

[11] Stewart, S. L.; J. B. Valdés, J. Gastelum, D. Brookshire, J. Aparicio, J. Hidalgo and I. Velazco, "A Decision Support System for Demand Management in the Rio Conchos Basin, Mexico, " in *Hydrology: Science and Practice for the 21st Century Vol. II,* British Hydrological Society, 487-494, 2004.

[12] Schmandt, J. (2002). Bi-national water issues in the Rio Grande/Rio Bravo Basin. Water Policy, 4(2): 137155.

This page intentionally left blank

Growing vulnerability of the arid zones to drought and its impacts

I. Velasco
Hydrology Department, Mexican Institute of Water Technology, Mexico

Abstract

Although drought and aridity are two subtly different natural phenomena, they are not unrelated, and when they happen simultaneously, the implications can be disastrous.

Partly because of the natural climatic march, and partly as a consequence of the human activities, the deserts are increasing and droughts are becoming more persistent and recurrent. The variability and environmental changes are factors that trigger the combination desertification-drought as a vicious circle that leads to potential crisis conditions because of the severe impact on society and its activities.

Given the complexity of both phenomena, and their great coverage and slow but relentless development, it is very difficult to confront them with success, and then neither to avoid them, but we must make, as society members, our best efforts to formulate strategies and improve planning in order to confront its occurrence and to mitigate its impacts.

Nevertheless, the structural (engineering) aspect is decisive for this goal, the non-structural phase is probably of greater importance, and it is directly related with the social participation as a group, beyond the official institutions and government dependences; after all, we are all water users, and then we are all indebted to contribute and to participate in the actions that allow us to overcome these risk conditions in order to avoid expensive crises and catastrophes which injure the humankind.
Keywords: drought, desertification, environment deterioration, water stress.

1 Introduction

Mexico is a country of weather and climate contrasts: although around 2/3 of its territory is classified as arid or semi-arid [1,2], what is reflected in the inhabitant water availability (Figure 1); also, a significant part, mainly on coastal zones is subject to the impacts of the hurricanes.

Both extreme phenomena in connection with the water are potentially of high impact, and paradoxically, frequently they happen simultaneously: while Northern is razed by intense and extensive droughts, South and Southeast suffer the damages because the hurricanes.

The evidences point to that in the last years, these events (droughts and hurricanes) have become more persistent and more intense [6]. Wide regions exist in Mexico, where lack of rain has taken more than 10 continuous years, and also areas that have been whipped by several hurricanes in the same season.

Especially toward North of the Cancer Tropic, the rain it has been smaller than normal recurrently, what has caused severe problems of drought that have rebounded even in binational conflicts, for the inadequacy of water in the shared basins whose International Treaties have not been possible to complete [8], creating so much stress between water users and local authorities, and of diplomatic character too.

This way, the "normal" water demand has not been possible to satisfy fully, causing serious economic and social problems, propitiating that the available resources reach a use level higher than their capacity of natural regeneration, and then making that the risk of desertification grows significantly [9].

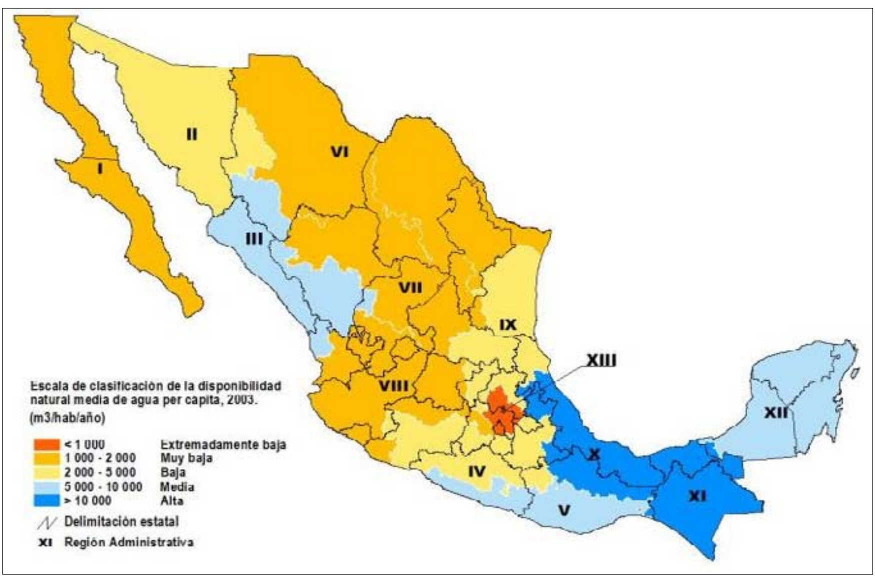


Figure 1: Classification scale of natural water availability, m^3/per capita/year, estimated for Mexico in year 2003 [2].

2 Water availability

According to UNO water standards [5], Mexico is a country with medium water availability (Figures 1 and 3), although this situation shows a trend to decrease, in function of the population increasing. It is considered that the rain annual average in Mexico is about 770 mm/year (Figure 2); nevertheless, its space distribution is very irregular, because while in some areas on the humid tropic it rains more than 3,000 mm/year, some areas in the North hardly reach 100 mm [7].

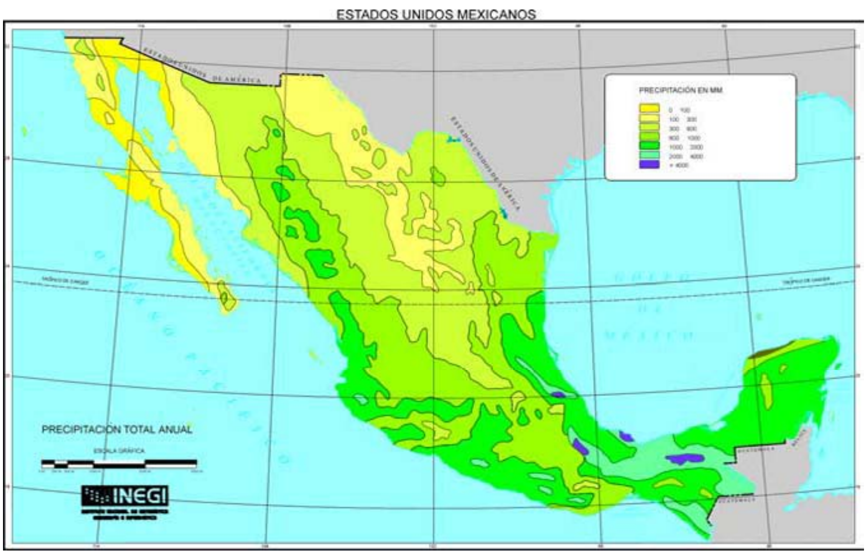


Figure 2: Annual average rainfall in México, in millimeters [1].

Paradoxically, most of the population, the industry and the agriculture are in the less rainy areas (Figure 3), creating with it strong pressures on the water and associate resources.

The population growing, the relative decreasing in the available water, in quality and quantity, and the every time bigger demand required to satisfy the necessities in the diverse uses, outlines serious problems to the environmental management, and increasing the stress in the use of the resource (Figure 4).

One aspect that has more influence in the environmental problems, related with the water requirements and availabilities, and with diverse products as vegetable fibers, wood, firewood, vegetable coal, etc., it is the ravage of the forests and natural grasslands, whose deterioration to quick rate, and with the minimal opportunity of recovery, it is causing that the desertification levels are really alarming [1,5]. In this way, soil, aquifers, surface streams, flora, and wildlife are reducing, and the disappareance risk is very high; that is unforgivable and a big tragedy, because the deep damage to environment and the natural balance.

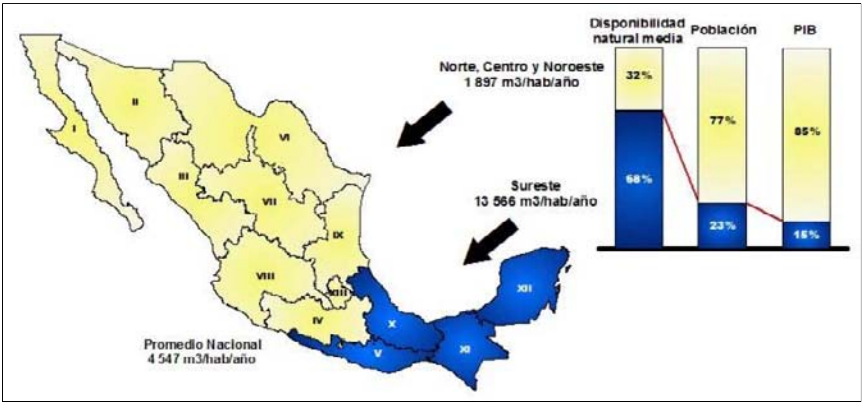


Figure 3: Natural water availability in Mexico, m^3/per capita/year [2].

3 Environmental deterioration

It is considered that among the years 1993 and 2000, 7,700 km^2 of forest lands have been destroyed, 5,100 of which were forests or tropical jungle; the rate of deterioration in this sense it was about 1.1%, one of the highest in the planet [2]. The main causes of this fact were the illegal pruning, the extensive cattle raising, agricultural land cleaning, and the forest fires.

In the last years, although the government efforts have diminished this rate, in general it is still high, and the reforestation has not still surrendered tangible fruits; in the best of the cases, the results are long-time, and on the other hand, to reforest is not only to plant small trees, but taking care of them so that they take root and be integrated to the environment, as well as so they reach the appropriate size.

Without the protection that the trees and the natural grasslands provide to the soil, this is degraded quickly, because the effect of water and wind, converting extensive areas before green and of great biological diversity, in unproductive barrens that finally give way, causing desolation and impoverishment, excluding wide geographical areas and population sectors (it Figures 5), condemning them to emigrate or to continue exploiting the resources until irreversible levels [7].

4 Drought role

This way, the drought is a trigger factor of the process, and according to its recurrence and severity, it becomes a lash of the affected areas, since its impacts are prolonged during years, and the social and economic costs are frequently beyond the reach of the affected people, and the government resources are insufficient, sometimes still to palliate the most elementary necessities [4]. Due to drought complexity, is very common to think that there will not be another drought episode equal or greater than the last, but unfortunately, subsequent events can be worse, if there is no planning to confront it.

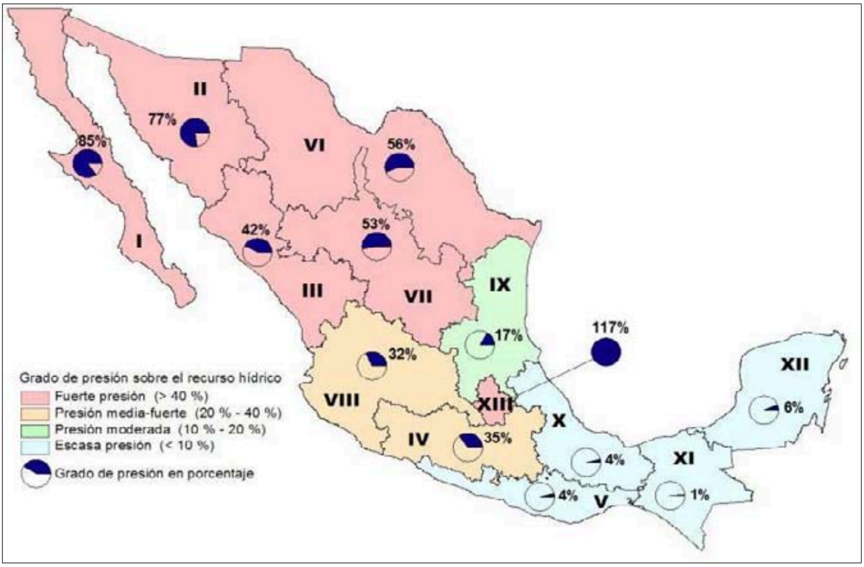


Figure 4: Stress level on available water in Mexico [2].

Although drought doesn't recognize political borders neither development levels, when it is presented, it affects more drastically the underdeveloped countries, in which it is capable to alter the natural and human environment significantly, in order to cause intense famines and human suffering.

These aspects make that, in a country like Mexico, the vulnerability to drought presence is a first-order factor. Irrigated areas, fundamental for the national agricultural development, and basically located in the areas that have been more affected by drought, they face serious problems of water availability, and the conflicts between uses and users they are something common and constant.

The areas that are non-irrigated, besides being more vulnerable, have a bigger risk of deterioration, since the population exploits to the limit the natural resources, and the recovering opportunities are also minimum [3]. In these areas is where the population inhabits are poorer than other population sectors, and where the underdevelopment conditions are accentuated; the problems like migration and marginalization reach levels of catastrophe.

5 Conclusions

The vicious circle of *drought-deterioration-marginalization-poverty-desertificación-abandonment* is difficult to break, and just intense and constant joint efforts of the population and government are able revert the effects.

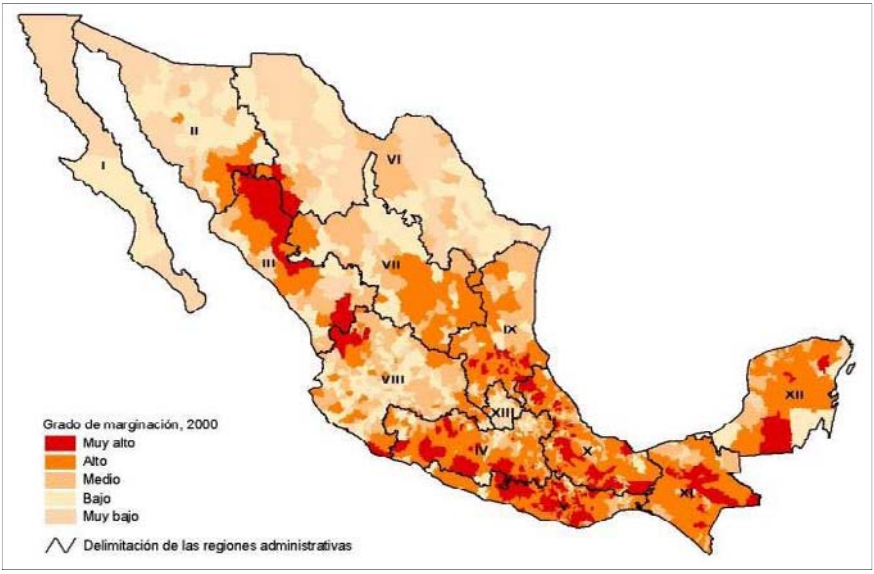


Figure 5: Population marginalization level in Mexico [2].

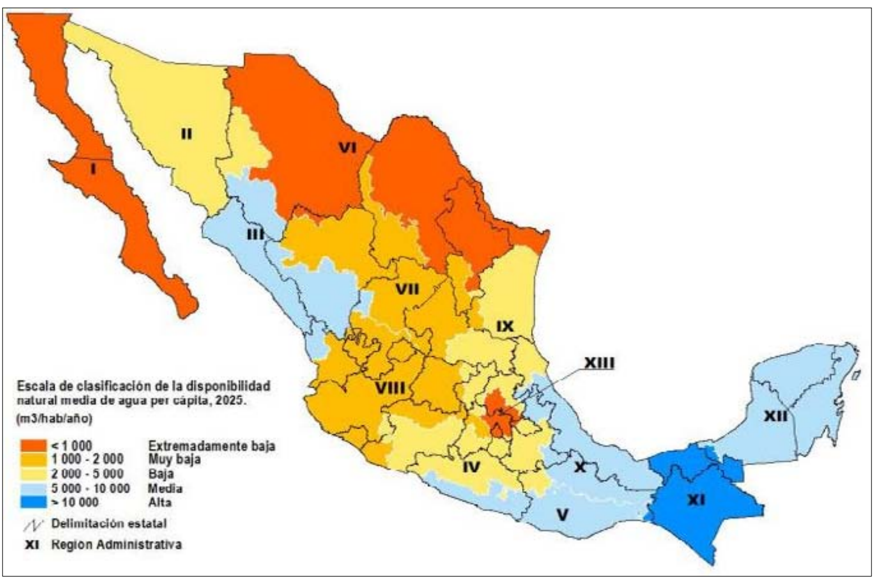


Figure 6: Classification scale of natural water availability, m^3/per capita/year, estimated for Mexico in year 2025 [2].

In connection with drought, the most appropriate form to mitigate the impacts and damages is the prevention, closely tied with the aspect of "water culture" or

“water education”. Only when there are appropriate preventive measures and in the fact that each part of the society and government assumes the list that correspond to each one, it will be possible that the environmental deterioration stops, to reach the human well-being and the sustainability in the use of water and associated resources is possible.

Natural conditions propitiates that Mexico is a low-medium water availability country; its 2/3 territory is located in arid and semiarid zones, where water is naturally scarce (Northern desert fringe).

These regions plow highly vulnerable to natural phenomena like droughts and hurricanes. Mexico is making great efforts to improve water management through diverse mechanisms: laws, strategies, and investments in structural and non-structural aspects, in order to reach the conservation of its natural resources.

Future perspectives are positive if government and society fulfill their respective roles; otherwise, the panorama is not very pleasant (Figure 6).

References

[1] Análisis del cambio de uso del suelo. Mapas del Análisis del cambio de uso del suelo (Soil Use and Change Análisis). INEGI - UNAM, www.ine.gob.mx/dgoece/xid/dgioece/i_usv/

[2] CNA (Comisión Nacional del Agua, Nacional Water Comisión). http://www.cna.gob.mx/eCNA/Espaniol/Estadisticas/Central/Estadisticas_Agua_2004/SWM_2004.htm

[3] Desert Science. www.iydd.org/

[4] Martínez Fernández, Julia. Agua y sostenibilidad: algunas claves desde los sistemas áridos (Water and Sustainability: some Keys from the Arid Systems). *Proc. of the Encuentro por una nueva cultura del agua en América Latina. Fortaleza, Brasil,* pp. 15-23, 2005.

[5] OCDE. *Evaluación del desempeño ambiental Mexico (OECD Environmental Performance Reviews: Mexico)*. SEMARNAT, México, pp. 67-91, 2003.

[6] Postel, S. *Last Oasis*. The Worldwatch Environment Alert Series. W. W. Norton & Company. New York, USA. pp. 45-59, 1992.

[7] Velasco-Molina, H. A. *Las zonas áridas y semiáridas. Sus características y manejo (Characteristics and Management of Arid and Semiarid Zones).* Editorial Limusa, S.A. de C. V. México. pp. 725, 1991.

[8] Velasco, I., F. J. Aparicio. Drought Detection and Evaluation Through a GIS for the Conchos River Basin, Mexico. *Proc. of the Third International Symposium on Transboundary Waters Management.* May, 29- Jun2. Session 2.3.C: Analytical Tools and Frameworks. Ciudad Real, Spain, 2006.

[9] Velasco, I., J. Aparicio, J. B. Valdés, J. Velázquez, T-W. Kim. Evaluación de índices de sequía en las cuencas de afluentes del río Bravo/Grande (Drought Indices Evaluation on the Rio Grande Tributary Watersheds). *Ingeniería Hidráulica en México (Hydraulic Engineering in Mexico)*. Vol. XIX, No. 3, pp. 37-52. Mexico, 2004.

Monitoring land use and land cover changes in Turkmenistan using remote sensing

L. Orlovsky[1], S. Kaplan[1], N. Orlovsky[2], D. Blumberg[3] & E. Mamedov[4]
[1]*Department of Solar Energy and Environmental Physics, Jacob Blaustein Institutes for Desert Research, Ben-Gurion University of the Negev, Sede-Boqer, Israel*
[2]*Department of Drylands Biotechnologies, Jacob Blaustein Institutes for Desert Research, Ben-Gurion University of the Negev, Sede-Boqer, Israel*
[3]*Department of Geography and Environmental Development, Ben-Gurion University of the Negev, Beer-Sheva, Israel*
[4]*National Institute of the Deserts, Flora and Fauna, Ministry of Nature Protection, Ashgabat, Turkmenistan*

Abstract

In Turkmenistan the most prominent cause for desertification is inappropriate land use practices. The natural arid pastures have limited carrying capacity and any changes of the fragile balance can lead to the destruction of this valuable resource. One of the most appropriate tools for monitoring these processes is change detection through remote sensing imagery. Accurate monitoring of changes on the Earth's surface is important to understand the relationship between man and nature and to provide decision makers with relevant information. The information on vegetation change is the most important of these relationships. Vegetation cover is also a useful indicator of the magnitude of land degradation that is easily assessed by multispectral remote sensing. The reduced vegetation cover causes an increase in albedo, which can also be monitored by remote sensing. The combination of these two parameters can give us a better map of the pasture status and its degradation rate. Landsat TM and ETM+ images were processed to maps of land use/land cover changes in northern Turkmenistan. The data were further processed in GIS and revealed the shrinking and the degradation of the pasture area. From the 1970s a total of ~4000km^2 of pasture were transformed into agricultural land, increasing the grazing pressure in the remaining areas. By applying advanced techniques for image based end-member retrieval and spectral mixture analysis a sub-pixel fraction was obtained for each end-member. The fractions of soil and vegetation emphasize the most degraded/rehabilitated sectors of the study area. Our results indicate the reduction of vegetation in specific areas while most of the desert experiences an increase in the vegetation cover. Our current study focuses on combining the spectral mixture analysis products with other degradation criteria such as change detection using albedo and vegetation indices to produce a more detailed assessment and understanding of the processes leading to these changes.
Keywords: Central Asia, pastures, irrigation, secondary salinization.

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060461

1 Introduction

The Asian continent suffers heavily from desertification, in particular severe salinization, waterlogging, vegetation degradation, and soil erosion. Irrigated agriculture and livestock farming are main branches of Turkmen economy. 95% (38.34 million hectares) of the agricultural lands of Turkmenistan are used as around-the-year natural desert pastures. During the last 30 years the total livestock number increased from 4,709 thousand (1970) to 7,792 thousand (1996) and 17 million heads (2005), at the same time population increased more than 2.5 times. Consequently, the anthropogenic pressure on these areas has increased. During the Soviet regime organized delivery of water to the distant pastures existed. Nowadays, due to the financial constraints this approach is not feasible. As a result more than 25% of pastures have not being used because of water lack [1]. The traditional seasonal migrations are ceased. The herds are concentrated around the settlements and near existing watering points all year round (*ibid*). According to different estimations, from 60 to 80% of the pastures are degraded at different rates, but the last systematic survey and assessment of desert and marginal lands was carried out in 1980s, while the most dramatic changes occurred during last 10-15 years. The increase in the cultivated irrigation led to 1) increased demand in irrigation water and water for drainage; 2) further rise of ground water level; 3) secondary soil salinization processes.

Since the natural arid pastures have limited carrying capacity, any factor that changes the fragile balance can lead to the destruction of this valuable resource. Changes in vegetation cover and its productivity in the temporal and spatial aspect is probably the most sensitive sign and early indicator to desertification processes; however, it is not an easy task to monitor vegetation state over large areas.

The vastness of Central Asian natural pastures, financial constraints of these newly independent states and lack of the modern technical means restrict proper monitoring of pasture land. Modern and updated assessment of rangelands is essential since dramatic political and economical changes took place all over central Asia.

Remote sensing has been considered to be an appropriate tool for studying changes in the arid and semi-arid environments due to its multitemporal coverage, synoptic view and digital format which enable monitoring of the Earth surface quickly and repetitively for large areas that are difficult to study by conventional methods [2–4]. However, only few studies have been done in Turkmenistan using remote sensing methods.

Usually the Instantaneous Field Of View (IFOV) of a sensor contains more than one land cover type, especially in arid environment that are commonly a combination of vegetation and soil in different proportions, creating mixed pixels. In order to solve the mixed pixels problem the Spectral Mixture Analysis method has been developed. Most studies involved in un-mixing of land cover elements in arid regions used multispectral unmixing techniques, where vegetation fraction was quantified relative to the soil and rock fractions [5–8].

The aim of this study is to assess land use/land cover changes in north Turkmenistan and reveal trends in these changes using remote sensing methods.

2 Study area

The research was conducted in the Dashoguz velayat (province) in northern Turkmenistan located between 39.5°-42.7°N and 56.4°-61°E. The area of the province is 74,588 km^2, it is populated by 1,196.7 thousands people. The portion of rural population is 68%. Climate is arid with annual average precipitation less than 110 mm and with notable contrast between seasons and years. Most of precipitation occurs in winter and early spring. Average annual temperatures are 12°C - 16°C with high seasonal amplitude up to 34°C [9]. The province has two main physics-geographic regions: a) the northern flat lowland consisting of Sarykamish depression and adjacent ancient alluvial deltaic plain of Amudarya River with spots of takyrs and small sandy massifs; b) Trans-Unguz Karakum Desert with sub-meridianal vegetated sand dunes and “kyrs”– flat-topped ancient ridges consisting of gypsum and carbonates. The irrigated lands of Dashoguz province constitute more than one quarter of the arable lands in Turkmenistan. One third of Turkmen cotton, more than 80% of rice and 50% of corn are cultivated here. Within the study area this paper will focus on the transition zone between irrigated agricultural land of Sarykamish ancient delta and pastures of Transunguz Karakum Desert. This transition zone is strongly influenced by human impact and thus highly dynamic.

3 Methodology

3.1 Selection of images and pre-processing

The series of Landsat TM (1988) and ETM+ (2004) images have been selected for comparison, providing a full coverage of the Dashoguz province. This choice was determined according to the similarity of precipitation amount and distribution in the rainy seasons of 1987-1988 and 2003-2004. The images were taken in/or close to the presumed peak of vegetation season – at the end of March – beginning of April, when both annual and perennial vegetation is present. Prior to performing multi-temporal image data analysis, geometric, atmospheric and radiometric corrections were carried out in order to improve the image quality and accuracy of analysis.

3.2 Spectral Mixture Analysis

The SMA is designed to derive the proportions of different basic land cover components that compose a mixed pixel [4]. We chose the linear model, which assumes that the reflectance of a pixel is a mixture of the reflectance of 'pure' components (end-members) in proportion to the area that they cover:

$$\rho_{i,j,k} = \sum_{m=1,\rho} F_{i,j,m}\rho_{m,k} + e_{i,j,k} \tag{1}$$

where $\rho_{i,j,k}$ *is* the reflectance of the *i, j* pixel in band *k*; $F_{i,j,m}$ is the fraction of the *m* component in that pixel for each of the ρ pure categories; $\rho_{m,k}$ is the reflectance

of the pure cover *m* in the same band *k,* and $e_{i,j,k}$ is the error term for that pixel. The fraction of each pure category covering the pixels of the image ($F_{i,j,m}$) is an unknown value, which is being solved by spectral unmixing of the image. The best mathematical approach is to minimize *e*, using a least-squares solution [10]. The SMA can only be solved if there are less end-members than spectral bands.

According to our results, the SMA provides better estimation of the change processes in the desert comparing to vegetation indices, since both quality and quantity assessment of vegetation and soil proportions are obtained.

The un-mixing procedure was carried out using ENVI 4.1 software. The main end-members of the area are sandy soil and vegetation represented by *Haloxylon sp., Salsola sp.* and *Callygonum* bushes. In the transition zone between the desert and irrigated area, water starts playing a significant role and is included as an end-member for that part of the image solely.

The increase in vegetation fraction in the desert area points on rehabilitation processes, and vice versa. The presence of water in the desert transition area indicates the process of flooding due to the leakage from the canals and flow of drainage water from the irrigated fields.

The vegetation and sand signatures were measured in five wavelengths corresponding to the Landsat bands in the field using CropScan portable radiometer. The water spectral signature was taken from [11], who measured it at six wavelengths, fig. 1.

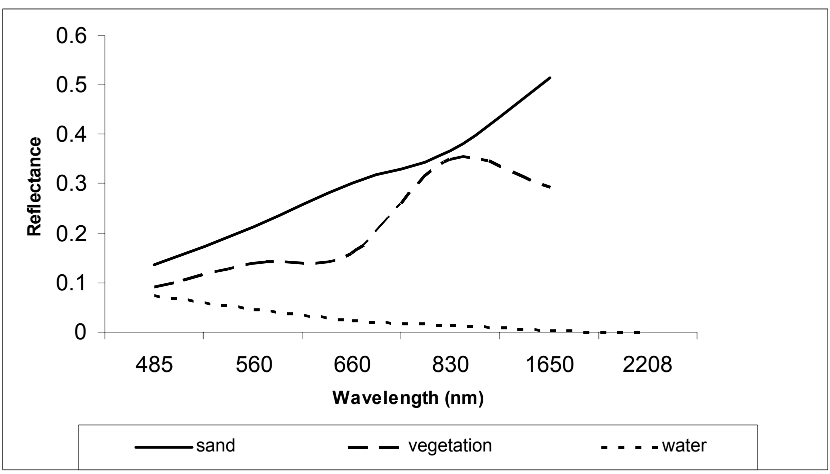


Figure 1: Spectral signatures used for unmixing algorithm.

3.3 Image differencing

Change detection is the process of identifying differences in the state of an object or phenomenon by observing it at different times. Using GIS and image differencing technique NDVI, SAVI and soil and vegetation fraction were calculated. The resulting image indicating the changes occurred in 1988-2004 was produced by subtracting the previous image from the subsequent one.

4 Results

4.1 Land use changes

There are two main types of land use in the study area – irrigated agriculture and natural pastures' use. The irrigated area was calculated from the time series of the Landsat images using GIS 9.0 software, fig. 2. The cultivated area in this region of Turkmenistan almost doubled in the last three decades, while the most rapid increase took place between 1973 and 1987, coinciding with development of large-scale irrigation in Central Asia for growing cotton. After collapse of the USSR in early 1990's the irrigated area keeps growing – again on the account of adjacent pastures – at that time mainly for growing wheat.

4.2 Vegetation cover changes

Change detection analysis of the Normalize Difference Vegetation Index (NDVI) and Soil Adjusted Vegetation Index (SAVI) images showed that areas around watering points and wells and other man made features such as gas pipes experienced degradation while in most pasture areas positive changes took place in 1988 – 2004. Anyhow, such analysis gives only qualitative assessment and shows the trend of changes. In attempt to quantify these changes we applied the SMA, assuming that a pixel of image taken from Trans-Unguz Karakum consists of two end-members – sand and vegetation. In the pixels/areas, which undergoing the negative changes (decrease in NDVI and SAVI values) the unmixed vegetation fraction doesn't exceed 25-30%. It occurs in the areas around watering point, at the tops of sandy ridges, and in the areas of so-called technogenic desertification (near newly constructed canals, water- and gas-pipes). Unmixed values of vegetation fraction in the inter-ridge depressions and lower parts of the ridges' slopes, which were identified by vegetation indices as positively changed, ranges from 45 to 60%.

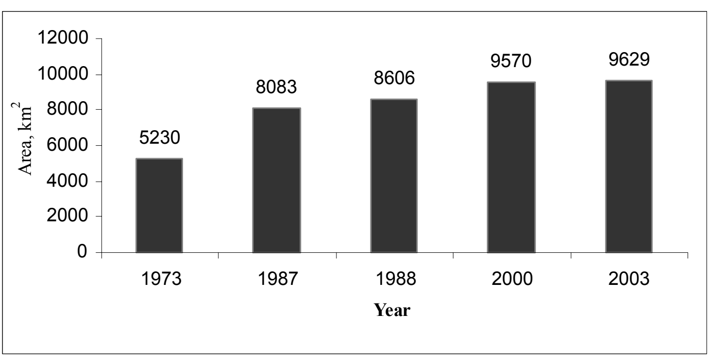


Figure 2: Increase of cultivated area in Dashoguz province of Turkmenistan.

Figure 3 shows the results of the change detection procedure for the vegetation fraction. The trends of the changes revealed by vegetation indices are

confirmed by SMA change detection. Most of the area experiences an increase in vegetation cover by 10%-15%, except desertification plots around the water points and wells, water- and gas-pipes. Usually the radius of desertification plots around the well in the Karakum Desert is 5 or even 10 km [12] - according to daily migration route of sheep. However, the current situation differs from the traditional concept: both images and field survey carried out in spring 2006 reveal the rehabilitation processes at the distance of 800 m and even less from the well. Moreover, further advance from the “epicenter” of desertification plot around Ajikui well (41.079°N, 59.386°E) revealed developing biogenic crusts (“karakharsangs) consisting of moss Tortula desertorum, lichens and cyanobactreia, which serves as indicator of undergrazing –degradation of vegetation due to under-use [13]. The situation is similar near the water- and gas–pipes, although without biogenic crusts.

Our initial assumption/hypothesis was that the remote desert pastures are rehabilitating, while vegetation in the areas adjacent to oases is degrading. Image analysis showed the opposite: areas adjoining irrigated massif are undergoing the rehabilitation. Field survey revealed that in these areas, which were classified as degraded in late 1980s, Haloxylon aphyllum (traditional firewood for the local dwellers) was planted as part of combating desertification measures. Since the gas-pipe has been constructed approximately at the same time and free natural gas is supplied to local population, cutting of trees and shrubs for firewood ceased.

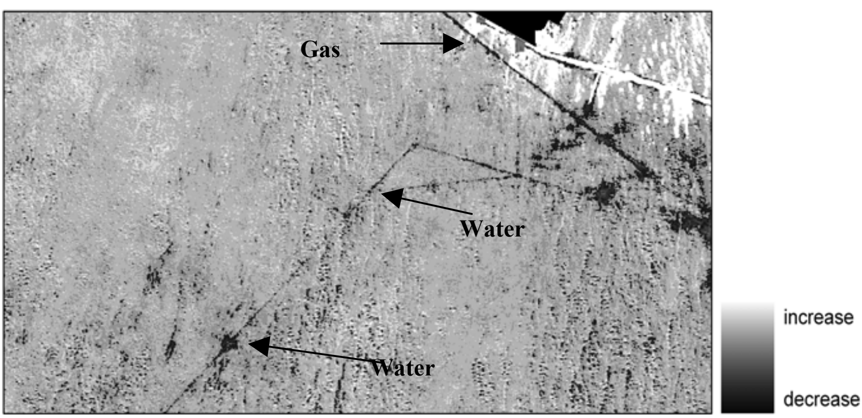


Figure 3: Change of vegetation fraction image. Brighter colors represent rehabilitation of vegetation and darker colors represent degradation.

4.3 Water logging and flooding

The processes of water logging and flooding take place mainly in the boundary area between irrigated fields and desert. They occur as a result of 1) filtration of water from the irrigation canals, which were constructed without waterproof isolation, and 2) discharge of excessive irrigation water and drainage water to the

adjacent desert areas. In spite of increase of vegetation cover and productivity around filtration lakes and waterlogged areas, this process is considered as pasture's degradation, since the available pasture is shrinking, hydrophilic species in many cases are unpalatable, and the migration routes of the livestock is disturbed. The SMA shows the increase of flooded area in Trans-Unguz Karakum from 93 km^2 in 1988 to 268 km^2 in 2004 (including irrigation and drainage canals), fig. 4.

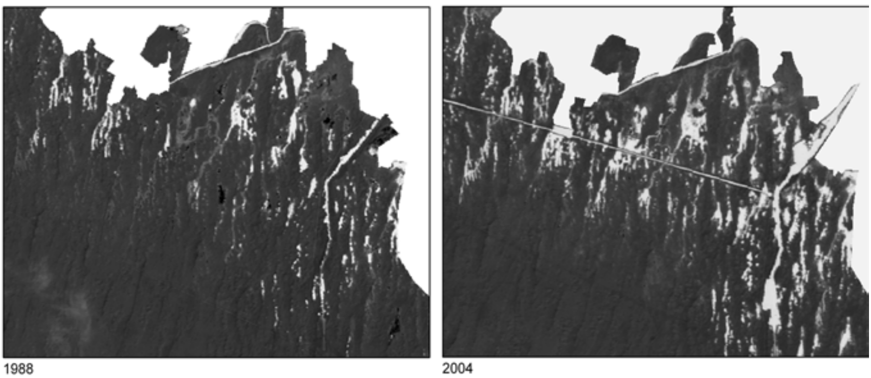


Figure 4: Water component comparison. Brighter colour represents higher water fraction.

4.4 Secondary soil salinization

While the most natural pastures of the Dashoguz province are undergone the rehabilitation processes, the situation in the irrigated areas is just opposite: there are no irrigated fields, which are not salinized. The main reasons of secondary salinization of irrigated lands in Dashoguz are: 1) improper quality of irrigation water; 2) antiquated irrigation techniques; 3) excessive irrigation rates; 4) improper soil washing. Salinity of Amudarya River water (the main source of irrigation water) increased from 0.2-0.4 g per litre at the beginning of 1960s to 2-2.5 g/l at present due to return flow of highly salinized drainage water to the River. Every year from 9 to 32 tons of salts per hectare are brought to the cotton field [14]. Irrigation canals (both main and intra-farming) have been constructed directly in the ground without impervious cover, which leads to the high losses during transportation (up to 25%) and formation of filtration lakes. The large size of the fields and out-of-date irrigation techniques (flooding and furrow irrigation) makes impossible application of biologically required irrigation rates – 8.88 thousand m^3 per hectare. The length of the drainage network is about 10,000 km, or 22.3 running meter per irrigated hectare, which provide about one third of irrigated lands by drainage. The volume of produced drainage water in the province is 2.72 km^3 with average salinity 4.38 g per litre. The main reason of soil salinization is the high level of ground water. Insufficient effectiveness of the drainage network leads to increase of the ground water level – 94% of the irrigated lands are in the area with ground water depth less than 3 m.

5 Conclusions

Vegetation is an important natural resource of Turkmenistan's desert. However, since the 1980's no research was carried out concerning the ongoing changes of vegetation cover in the desert pastures. In spite of the sparse desert vegetation and relatively low spatial resolution Landsat imagery gives the correct results on the current changes occurring in the arid lands of Turkmenistan. Primary results of this research show the potential of Spectral Mixture Analysis for accurate vegetation mapping and land change detection. The vegetation change detection procedure shows that although there is a reduction in the available area for pasture in Dashoguz province of Turkmenistan due to expansion of irrigated fields, most of the pasture area is rehabilitating. The water end-member analysis shows the leaks from the irrigation and drainage canals, which are causing flooding and water logging on the account of pasture area, thus, triggering more degradation. Mismanagement of native pasture may harm natural re-vegetation of dunes by palatable species, which are well adapted to the arid environment and uses as livestock food. By using remotely sensed data this valuable resource is better understood and preserved for an ongoing sustainable development. However, further research is necessary, which will include developing of a complete unmixing algorithm.

Acknowledgement

The authors express their sincere gratitude to U.S. Agency for International Development for funding the project CA22-010, which made this study possible.

References

[1] Babaev, A. G. (ed.). *Desert problems and desertification in Central Asia*. Springer: Berlin, pp. 260, 1999

[2] Singh, A., Digital change detection techniques using remotely sensed data. *International Journal of Remote Sensing,* 10, pp. 989-1003, 1989.

[3] Mas, J. F., Monitoring land cover changes: a comparison of change detection techniques, *International Journal of Remote Sensing* **20,** pp. 139-152, 1999.

[4] Collado, A. D., Chuvieco, E., Camarasa, A., Satellite remote sensing analysis to monitor desertification processes in the crop-rangeland boundary of Argentina, Journal *of Arid Environment* **52,** pp**.** 121, 2002.

[5] Pech R. P., Graetz, R. D., Davis, A.W., Reflectance modelling and derivation of vegetation indices for an Australian semi-arid shrubland, *International journal of remote sensing* **7**(3), pp. 389-403, 1986.

[6] Smith, M. O., Ustin, S. L., Adams, J.B., Gillespie, A.R., Vegetation in deserts: I. A regional measure of abundance from multi-spectral images, *Remote Sensing of Environment* **31,** pp.1-51, 1990.

[7] Shoshany, M., Kutiel, P., Lavee, H., Monitoring temporal vegetation cover changes in Mediterranean and arid ecosystem using a remote

sensing technique: case study of the Judean mountain and the Judean desert, *Journal of Arid Environment,* **32,** pp. 1-13, 1996.
[8] Ustin, S. L., Hart, Q. J., Duan L., Scheer G., Vegetation mapping on hardwood rangeland in California, International *Journal of Remote Sensing*, **17**, pp. 3015-3036, 1996.
[9] Orlovsky, N. S., Climate of Turkmenistan. *Biogeography and Ecology of Turkmenistan*. V. Fet and K. I. Atamuradov (eds.). Kluwer Academic Publishers: Dordrecht / Boston /London, pp. 23-48, 1994.
[10] Shimabukuro, Y. E., & Smith, J. A., The least squares mixing models to generate fraction images derived from remote sensing multispectral data, *IEEE Transactions on Geoscience and Remote Sensing*, **29**, pp. 16-20, 1991.
[11] Small C., The Landsat ETM+ spectral mixing space. *Remote Sensing of Environment* **93**, pp. 1-17, 2004.
[12] Babaev, A.G. (ed). *Productivity of vegetation in Central Karakum Desert under different regime of use*. Nauka: Moscow, pp.255, 1979.
[13] Orlovsky, L., Dourikov, M., & Babaev, A., Temporal dynamics and productivity of biogenic soil crusts in the Central Karakum desert, Turkmenistan, *Journal of Arid Environments* **56(4)**, pp. 579-601, 2004.
[14] Esenov P., Ecological and ameliorative state of irrigation zone of Dashoguz velayat of Turkmenistan, *Problems of desert Development*, **6**, pp. 29-34, 1995.

Trace metals in molluscs from the Beagle Channel (Argentina): a preliminary study

M. E. Conti[1], J. Stripeikis[2], M. Iacobucci[3], D. Cucina[1], G. Cecchetti[3] & M. B. Tudino[2]

[1]Dipartimento di Controllo e Gestione delle Merci e del loro Impatto sull'Ambiente, Facoltà di Economia, Università "La Sapienza", Roma, Italy

[2]INQUIMAE, Departamento de Química Inorgánica, Analítica y Química Física, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Buenos Aires, Argentina

[3]Centro per le Valutazioni Ambientali delle Attività Industriali, Facoltà di Scienze Ambientali, Università degli Studi di Urbino, Urbino, Italy

Abstract

Individuals of *Mytilus chiloensis* (Hupè, 1854) and *Nacella magellanica metalica* (Rochebrune and Mabille, 1885) were evaluated as possible biomonitors of trace metal baseline contamination in the Beagle Channel, Tierra de Fuego (Argentina). The selected species have the necessary prerequisites for use as biomonitors. They are well distributed in all coastal areas of the Beagle Channel and other regional seas, they are easy to identify and available all year round.

The first aim of this preliminary survey is to evaluate the concentrations of Cr, Cu, Pb and Zn present in soft tissues of the selected species in order to have more information on the use of these selected species as possible cosmopolitan biomonitors. Samples were collected in seven stations situated along 170 km of the coast in the tidal zone. The validity of these two species as bioaccumulators was tested.

Significant differences between metal concentrations in molluscs from different stations were tested by ANCOVA on $\log_{10}$ transformed data with body weight as covariate. Multiple comparison tests (MCT) were conducted when significant differences were detected among the stations.

In *M. chiloensis* metal concentrations decreased in the order: Zn > Cu > Pb ≥ Cr while for *Nacella* metal concentrations decreased in the order: Zn > Cu > Cr > Pb showing good bioaccumulation ability. At present, they can be considered as good candidates as trace metals biomonitors for the studied area.

From all the obtained data and statistical analysis (ANCOVA, MCT) the results showed clearly that there is no one site univocally more contaminated (with clearly high levels of metals accumulation in biomonitors) than any other. Thus the possible hypothesis of the Ushuaia Harbour as being the most contaminated site must be reconsidered. Metal concentrations recorded may be used for background levels for intraspecific comparison within the Patagonian seas.

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060471

1 Introduction

The use of biological species in the monitoring of trace metals in coastal areas allows evaluating the contaminants bioavailability or the effects of contaminants on marine organisms. Molluscs are among the organisms most used as bioindicators for trace metal pollution in biomonitoring programmes. Limpets and particularly bivalve molluscs can accumulate and integrate concentrations of several metals in seawater for relatively long intervals. They also assimilate trace metals from their food and from the ingestion of inorganic particulate material [1-4]. The levels of heavy metals accumulated by marine organisms are also a function of many factors such as temperature, salinity, diet, spawning, seasonal variations [5] and can also constitute a potential hazard for humans and other mammals. This is because some molluscs and limpets constitute a seafood group of particular interest.

The main advantage of the biomonitoring approach is a remarkable reduction in time and costs, consumed in frequent analyses of abiotic matrices of water and sediment samples. Bioaccumulators must have necessary requirements: they must be cosmopolitan and available all year round; they must accumulate the pollutant without being killed by the levels it comes in touch with; they must present a high concentration value of the contaminant and must be easy to sample and to preserve. Above all they must have a contaminant concentration that could be easily correlated with the concentration in the surrounding environment [5-7]. Mussels filter the surrounding water and they constantly accumulate metals in their tissues. Besides, they are easy to collect and to identify. *Mytilus chiloensis* (Hupè, 1854) is well distributed in South American seas (i.e. Beagle Channel, Strait of Magallanes, etc.); also a limpet *Nacella magellanica metalica* (Rochebrune & Mabille, 1885) is very abundant and it constitutes a good candidate as a possible biomonitor of trace metals in seawater. It lives on rocky substrata of tidelands at relatively shallow depths and tolerates fairly long periods of time outside of water. It eats algae and vegetable deposits that it scratches from the rocks. Moreover, it is important to underline that this herbivorous gastropod generally takes metals principally from the diet [8]. These molluscs are also relevant from the nutritional point of view. In fact they constitute a popular food in the studied areas.

The aim of the preliminary present study was to have more information on the use of these selected species as possible cosmopolitan biomonitors of Cr, Pb, Cu and Zn present in the Beagle Channel sea ecosystem.

2 Materials and methods

Beagle Channel is a strait in Tierra del Fuego, near the southern tip of South America. The channel has high ecological relevance and is about 240 km in length and between 5 and 14 km wide. It separates Isla Grande de Tierra del Fuego from several smaller islands to the south. The Beagle Channel was named for the British ship *Beagle,* in which Charles Darwin explored the area (1833–34).

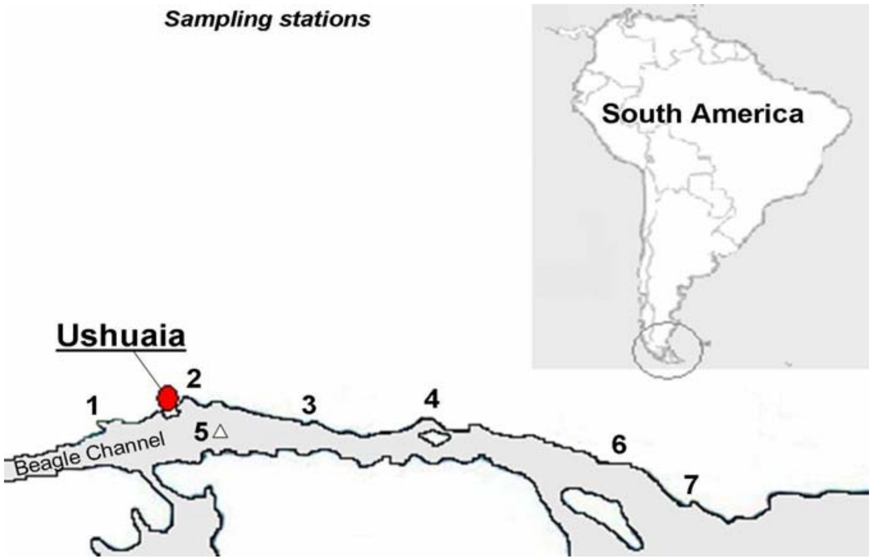


Figure 1: Sampling stations: 1) Lapataia Bay, 2) Ushuaia Harbour, 3) Punta Paranà, 4) Brown Bay, 5) Bridges Island, 6) Este Bay, 7) Punta Moat.

The molluscs have been sampled in February 2005 in 7 stations situated along of 170 km of the coast of the Beagle Channel (figure 1).

Individuals of *Mytilus chiloensis* (Hupè, 1854) (n=202) and *Nacella magellanica metalica* (Rochebrune & Mabille, 1885) (n = 105) have been picked in the tidal zone; then they were immersed (t = 24 hours) in filtered seawater to be purified.

Subsequently, the soft parts were taken out of the shell using tools (hammer and spatula) made exclusively of plastics, so as to prevent metal contamination, and then they were rinsed with deionized MilliQ water, in order to remove every residue of shell. Samples were placed in polyethylene bags, ice deep-frozen and transported to the laboratory.

Each sample of the mollusc previously homogenized (400-800 mg) was treated with 8 ml of 70 % (w/w) nitric acid Suprapur (Merck) and 2 ml of 30 % (w/w) hydrogen peroxide Suprapur (Merck) in PTFE vessels. The microwave digestion system MDS-2000 provided by CEM was used for the mineralization process.

Heavy metals in mollusc species were determined using a Shimadzu 6800 Atomic Absorption Spectrometer coupled to different atomic vapors generators depending on the analytical concentration. A graphite furnace accessory GFA-6000 and an autosampler ASC-6000 was employed for Pb and Cr measurements and an air/acetylene flame was used for Zn and Cu determinations.

The calculation of dry weight (d.w.) on the studied species (10 replicates for each location) was carried out through oven drying at 105 ºC until constant weight was achieved.

Traceability of results was obtained from the analysis of the certified reference material Antartic krill MURST-ISS-A2 (Italian Research Programme in Antartica). The mean recovery percentages (five replicates) were: Cr = 98.1±1.0 %; Cu = 101.1±1.3 %; Pb: 96.5±0.6 % and Zn: 102.1±2.9 %.

The detection limits (LODs) based on three times the standard deviation of the blank (n=11) were: Cr = 0.0002 $mg.L^{-1}$; Cu = 0.020 $mg.L^{-1}$; Pb = 0.001 $mg.L^{-1}$ and Zn = 0.010 $mg.L^{-1}$.

Significant differences between metal concentrations in molluscs from different stations were tested by ANCOVA on log_{10} transformed data with body weight as covariate. Multiple comparison tests were conducted when significant differences were detected among the stations. SPSS 13.0 software was used for statistical analysis.

3 Results

Table 1 shows mean metal concentrations obtained by the analyses of the molluscs. Mean levels in biota are referred to dry weight to reduce variability of measurements determined by habitat, life conditions, pre-treatment and conservation of organisms after sampling [5].

In *M. chiloensis* metal concentrations decreased in the following order: Zn > Cu >Pb ≥ Cr. For *M. chiloensis* the values obtained were in the range of 57.91-111.54 μg g^{-1} dw for Zn; 3.90-7.41 μg g^{-1} dw for Cu; 0.21-0.76 μg g^{-1} dw for Pb and 0.26-0.98 μg g^{-1} dw for Cr. Very few data are available on *M. chiloensis*.

Mean Cu and Zn levels determined in the Beagle Channel (5.78 μg g^{-1} dw and 80.64 μg g^{-1} dw respectively, see table 1) for *M. chiloensis* were higher with respect to those obtained for Strait of Magallanes (Chile) [9]. The comparison with the only other available dated study [10] was not possible because wet/dry weight ratio for *M. chiloensis* is not reported. We think that reporting data referred to dry weight basis is of paramount importance because this procedure can avoid high variability in the obtained data. However, if we compare our bivalves data with other *Mytilus* species from Mediterranean seas [1], in areas with low-medium contamination levels, we observe significantly (p = 0.05) lower mean values for Cr, Cu, Pb and Zn in the Beagle Channel.

For *Nacella* the values obtained were in the range of 62,12-76.15 μg g^{-1} dw for Zn; 6.77-11.87 μg g^{-1} dw for Cu; 0.66-2.48 μg g^{-1} dw for Cr and 0.11-2.13 μg g^{-1} dw for Pb. As to *Nacella* metal concentrations decreased in the order: Zn > Cu > Cr > Pb. Mean Pb levels for *Nacella* were clearly lower than Cr while for *M. chiloensis* the mean accumulation levels of Cr and Pb were very similar.

No data are available in the literature for metal accumulation in *N. magellanica metalica.* However, if we compare our results with data relative to other *Nacella* species collected in Magallanes Strait [9] we observe that our mean values (p = 0.05) for Cu and Zn have higher levels. Comparing also mean accumulation data from this study with other patellid limpet from Mediterranean sea (*P. caerulea*) [1], mean levels of Cu and Zn obtained were significantly lower (p = 0.05) in *Nacella* samples from the B. Channel.

To assess the existence of significant differences in the metal concentrations in molluscs, use of ANCOVA on log_{10} transformed data with body weight as

covariate was taken into account. This approach was chosen because the high variability on individual weights of the collected samples. For *Mytilus*, normality and homoscedasticity were checked, and resulted to be substantially improved by the log transformation with the exception of Pb values.

Table 1: Concentrations of metals (µg g^{-1} dry weight) in the soft tissues of *Mytilus chiloensis* and *Nacella magellanica metalica*, (mean ± s.d.).

Sampling station	Individuals	Cr	Cu	Pb	Zn
Mytilus chiloensis					
1	30	0.28±0.15	6.82±2.49	0.35±0.27	57.91±14.28
2	30	0.27±0.13	5.46±1.74	0.76±0.66	94.14±50.79
3	30	0.26±0.35	4.94±1.95	0.23±0.19	72.51±45.23
4	30	0.27±0.15	5.98±1.75	0.41±0.39	111.54±82.16
5	30	0.43±0.27	7.41±2.00	0.21±0.16	77.64±54.49
6	30	0.50±0.27	5.47±1.49	0.59±0.37	85.67±61.17
7	22	0.98±0.38	3.90±1.62	0.47±0.44	59.36±23.15
Total / Mean	**202**	**0.41±0.33**	**5.78±2.14**	**0.43±0.42**	**80.64±58.83**
Nacella magellanica metalica					
1	15	2.38±0.97	11.87±5.82	1.49±0.48	76.15±11,16
2	15	2.48±1.02	10.04±4.70	2.13±1.03	74.38±12.00
3	15	1.04±0.97	8.11±2.85	0.65±0.25	66.58±8.54
4	15	1.48±0.79	8.20±3.41	0.11±0.02	67.84±7.75
5	15	0.66±0.47	7.48±2.37	0.47±0.25	62.12±8.62
6	15	1.91±1.66	10.32±4.57	0.41±0.33	70.18±13.55
7	15	1.66±0.79	6.77±1.82	0.43±0.21	63.86±6.74
Total / Mean	**105**	**1.66±1.16**	**8.97±4.13**	**0.81±0.82**	**68.73±10.87**

Fresh/dry weight ratio: 3,54 (*Mytilus*) and 4,88 (*Nacella*).

Parallelism of the regression lines between all sites was tested throughout at a *p* level of 0.05. For Cr and Zn (*M. chiloensis*) the condition of parallelism of the regression lines was met (F=0.865, sig.=0.522 for Cr; and F=0.339, sig.=0.915 for Zn), then ANCOVA was performed. Results showed significant differences in Cr and Zn concentrations in *M. chiloensis* between the studied sites (Tables 2-3).

Table 2: ANCOVA test for Cr between sites (*M. chiloensis*).

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	9,224(a)	7	1,318	19,204	,000
Intercept	3,503	1	3,503	51,053	,000
Weight	,503	1	,503	7,330	,007
Site	4,665	6	,778	11,332	,000
Error	13,311	194	,069		
Total	75,548	202			
Corrected Total	22,535	201			

Table 3: ANCOVA test for Zn between sites (*M. chiloensis*).

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	1,337(a)	7	,191	4,211	,000
Intercept	95,387	1	95,387	2102,252	,000
Weight Total	,147	1	,147	3,240	,073
Site	1,261	6	,210	4,630	,000
Error	8,802	194	,045		
Total	694,189	202			
Corrected Total	10,140	201			

For Cu (*M. chiloensis*) the parallelism condition was not met (F=3.405, sig =0.003) and then ANCOVA was not performed. From the valuation of different regression models (logarithmic, exponential, polynomial, etc.) resulted that Cu depends slightly from weight. Then, analysis of variance (ANOVA) was performed. ANOVA results showed that significant differences are present between sites for Cu accumulation in *Mytilus*.

The tests of normality and homoscedasticity for Pb resulted to be not improved by the log transformation. For this reason, a non-parametric Kruskal-Wallis test was conducted. Results showed significant differences for mean Pb concentrations in *M. chiloensis* between the studied sites (Table 4).

Multiple comparison tests (MCT) were conducted for Cr, Cu and Zn (Tukey HSD; Bonferroni; Tamhane; Dunnett T3 and Scheffè) and non-parametric MCT were also conducted for Pb (Games-Howell, Dunnett's C, Tamhane's T2) in order to know which station is more contaminated than another.

Table 4: Kruskal-Wallis test for Pb between sites (*M. chiloensis*).

	Log Pb
Chi-Square	42,214
df	6
Asymp. Sig.	,000

Results for Cr (*M. chiloensis*) showed that site 7 (Punta Moat) has mean values significantly higher (0.98 µg g^{-1} dw; $p = 0.05$) than the all other sites. The sites 5 (Bridges islands) and 6 (Este Bay) have similar Cr mean values and significantly higher than sites 1 (Lapataia Bay), 2 (U. Harbour) and 3 (Punta Paraná) and 4 (Brown Bay) but significantly lower than the site 7.

MCT conducted for Cu in *M. chiloensis* showed that site 7 (P. Moat) has lower mean Cu levels than the others with the exception of site 3 (P. Paraná) that has similar levels of site 7.

MCT conducted for Zn in *M. chiloensis* showed that sites 2 (U. Harbour) and 4 (Brown Bay) have mean Zn values significantly higher than the all other sites. Sites 1 (Lapataia) and 7 (P. Moat) have mean Zn concentrations lower than sites 2 and 4.

MCT conducted for Pb showed that sites 2 (U. Harbour) and 6 (Este Bay) has significantly higher mean Pb concentrations than the others with the exception of site 7 that showed average levels of Pb.

For *Nacella* normality and homoscedasticity were checked and resulted to be clearly improved by the log transformation for all metals. The same ANCOVA model was employed. Parallelism of the regression lines between all sites was tested and the condition of parallelism of the regression lines was met for all metals. Then, it was possible to use ANCOVA for testing differences in mean metal concentrations between sites.

Results of ANCOVA showed that there are no significantly differences between mean Zn and Cu concentrations for *Nacella* between the selected sites. Besides, ANCOVA showed significantly differences between almost a couple of sites for Cr and Pb for *Nacella* species. Then, MCT were conducted for these metals.

MCT conducted for Cr in *Nacella* showed that sites 1 (Lapataia Bay) and 2 (U. Harbour) have the higher Cr levels; the sites 4 (B. Brown), 6 (E. Bay) and 7 (P. Moat) have average Cr levels; the sites 3 (P. Paraná) and 5 (Bridges islands) showed lower Cr levels.

MCT for Pb in *Nacella* showed that site 4 (Brown Bay) has mean Pb concentrations significantly lower than the other sites. The sites 3 (P. Paraná), 5 (Bridges Islands), 6 (Este Bay) and 7 (P. Moat) have similar Pb levels and significantly lower than sites 1 (Lapataia Bay) and 2 (U. Harbor). The sites 1 and 2 presented the higher Pb levels.

In conclusion, from the reported statistical analysis we can observe that there is no one site more univocally contaminated than another for the studied trace metals in the selected biomonitors.

We have also tested the bioaccumulation ability of *M. chiloensis* and *Nacella* by using the *t*-student test (Table 5).

This study confirms that *Nacella* has clearly high bioaccumulation levels of Cr, Cu, Pb and Zn than *Mytilus* (figures 2-5).

Table 5: *t*-student test for all metals. Equal variances not assumed.

Variable	*t*-student	df	Sig. (2-tailed)	Mean Difference
Cr	-10,818	112,701	,000	-1,2525163
Cu	-7,407	133,711	,000	-3,1885070
Pb	-4,488	133,744	,000	-,3811073
Zn	2,976	230,020	,003	11,9064298

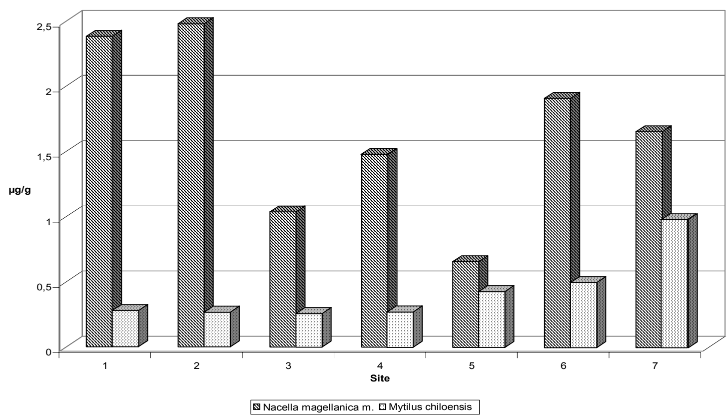


Figure 2: Cr concentrations in *Mytilus* and *Nacella* in the selected sites.

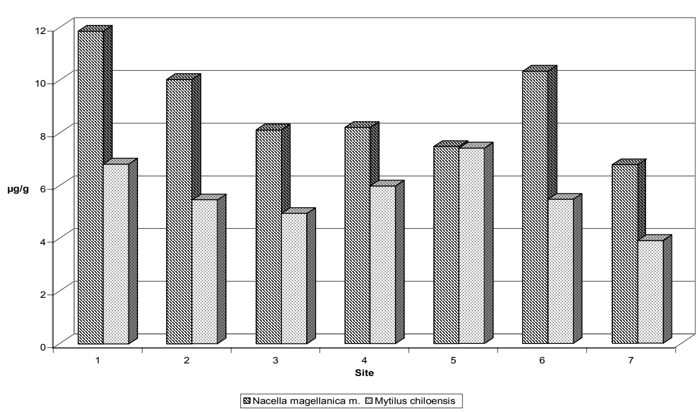


Figure 3: Cu concentrations in *Mytilus* and *Nacella* in the selected sites.

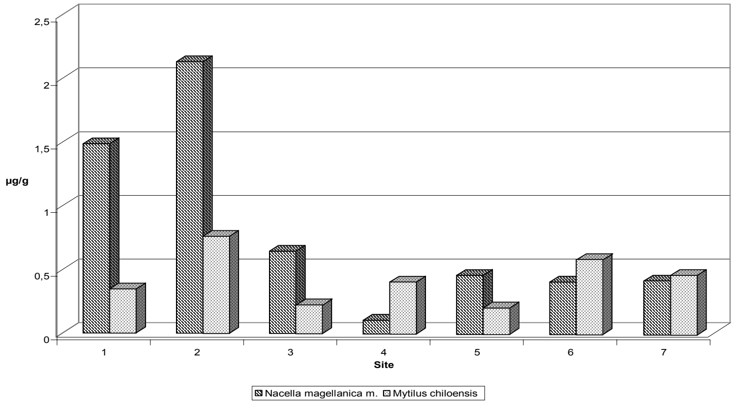


Figure 4: Pb concentrations in *Mytilus* and *Nacella* in the selected sites.

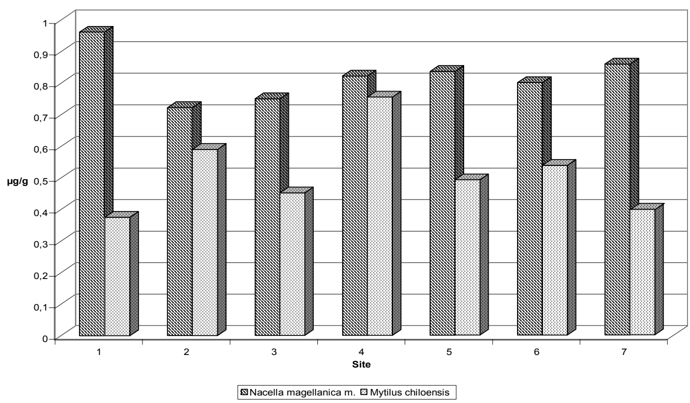


Figure 5: Zn concentrations in *Mytilus* and *Nacella* in the selected sites.

4 Conclusions

The use of the selected species turned out to be very valuable for the study of a coastal area with low levels of contamination. Then, *M. chiloensis* and *N. magellanica metalica* can be considered as potential biomonitors of trace metal pollution in the considered geographical areas. The selected biomonitors showed a very good ability to concentrate trace metals from seawater.

These two mollusc species have the necessary requisites: they are easy to identify and capture, they are sessile and sedentary, they are available all year round, and they are present along all coasts of the Beagle Channel.

Comparing our data with those available for *M. chiloensis* from the Magallanes Strait [9], the Beagle Channel bivalves showed higher Cu and Zn levels. However, this values are significantly lower ($p = 0.05$) than those of other *Mytilus* species from Mediterranean areas [1]. Comparing also our data with

those available for other patellid limpets, our Cu and Zn accumulation levels for *Nacella magellanica metalica* were comparable to those of the Magellan Strait [9] and significantly lower than samples of Mediterranean areas [1].

Statistical analysis for *Mytilus* revealed that, as far as Cr is concerned, Punta Moat station presents higher Cr concentrations than the other sites. On the other hand the same site shows, with some exception, lower levels of Cu. Ushuaia Harbour and Brown Bay have Zn values significantly higher than all the other sites. Pb accumulation values for *Mytilus* showed that U. Harbour and Este Bay sites have, with some exception, higher Pb concentrations than the others.

By using *Nacella*, results showed that there are no significantly differences for Zn and Cu concentrations between the selected sites.

The Cr accumulation levels in *Nacella* were significantly higher in Lapataia Bay and Ushuaia Harbour. Punta Paraná and Bridges islands resulted with a lower Cr levels. The Pb accumulation levels for *Nacella* showed that Brown Bay has mean Pb concentrations significantly lower than the other sites. Lapataia and U. Harbour resulted with the highest Pb levels.

From the reported data and statistical analysis, it is clear that there is no one site univocally more contaminated (with clearly high levels of metals accumulation in biomonitors) than another. Thus, the possible hypothesis of the Harbour as being the most contaminated site must be reconsidered.

The follow-up of this study, extended to other biomonitors and contaminants along time and including metal determination in surrounding seawater, might further confirm these first results.

References

[1] Conti, M.E. & Cecchetti, G., A biomonitoring study: trace metals in algae and molluscs from Tyrrhenian coastal areas. *Environmental Research*, **93**, pp. 99-112, 2003.

[2] Phillips, D. & Segar, D. A., Use of bio-indicators in monitoring conservative contaminants: programme design imperatives. *Marine Pollution Bulletin*, **26**, pp. 593-601, 1986.

[3] Campanella, L., Conti, M.E., Cubadda, F. & Sucapane, C., Trace metals in seagrass, algae and molluscs from an uncontaminated area in the Mediterranean. *Environmental Pollution*, **111**, pp. 117-126, 2001.

[4] Conti, M.E., Iacobucci, M. & Cecchetti, G., The biomonitoring approach as a tool of trace metal assessment in an uncontaminated marine ecosystem: Ustica island (Sicily, Italy), in: *Geo-Environment, Monitoring, Simulation and Remediation of the Geological Environment*; J.F. Martin-Duque, C.A. Brebbia, A.E. Godfrey, J.R. Dìaz de Teràn (Eds.), WIT Press, 335-344, 2004.

[5] Conti, M.E., Il monitoraggio biologico della qualità ambientale, ed. SEAM: Roma, pp. 35-38, 2002.

[6] Conti, M.E. & Cecchetti, G., Biological Monitoring: lichens as bioindicators of air pollution assessment - a review. *Environmental Pollution*, 114, pp. 471-492, 2001.

[7] Conti, M.E. , Iacobucci, M., Mecozzi, M., Cecchetti, G. (2006) Trace metals in soft tissues of two marine gastropod molluscs: *Monodonta turbinata* B. and *Patella caerulea* L. collected in a marine reference ecosystem, in: Environmental problems in coastal regions VI, including oil spill studies; C.A. Brebbia, (Ed.), WITpress, 3-11.
[8] Ahn, I.-Y., Kim, K.-W. & Choi, H.J. A baseline study on metal concentrations in the Antarctic limpet *Nacella concinna* (Gastropoda: Patellidae) on King George Island: variations with sex and body parts, *Marine Pollution Bulletin*, 44, 421-431, 2002.
[9] Astorga España, M.S., Rodríguez Rodríguez, E.M. & Díaz Romero, C., Sodium, K, Ca, Mg, Fe, Cu, and Zn concentrations in molluscs from the Magellan Strait (Chile): their contribution to dietary intake, *International Journal of Food Sciences and Nutrition*, 56, 337-347, 2005.
[10] Astorga España, M.S., Peña Méndez, E.M., Lecaros Palma, O. & García Montelongo, F.J., Heavy Metals in *Mytilus chilensis* from the Strait of Magallanes (Chile), *Marine Pollution Bulletin,* 36, 542-546, 1998.

This page intentionally left blank

Toxic and interactive toxic effects of agrochemical substances and copper on *Vibrio fischeri*

A. Kungolos[1], V. Tsiridis[1], P. Samaras[2] & N. Tsiropoulos[3]
[1]*Department of Planning and Regional Development, University of Thessaly, Volos, Greece*
[2]*Department of Pollution Control Technologies, Technological Educational Institute of West Macedonia, Kozani, Greece*
[3]*Department of Agriculture, Crop Production and Rural Environment, University of Thessaly, Nea Ionia-Volos, Greece*

Abstract

The toxic and interactive toxic effects of two agrochemicals (fosthiazate and metalaxyl-M) and copper were investigated in this study on the photobacterium *Vibrio fischeri*. The toxicities of all tested compounds were generally comparable. The order of toxicity was: copper > fosthiazate > metalaxyl-M. The interactive effects of fosthiazate and metalaxyl-M mixtures were found to be additive for most of the concentration combinations tested. For the binary mixtures of fosthiazate and copper the interactive effect was antagonistic for all concentration combinations tested. Both agrochemicals showed a similar toxic response, while the toxicity of copper increased significantly by the increase of exposure time and a steep increase of copper toxicity was observed with a small increase of its concentration.
Keywords: pesticides, agrochemicals, copper, bioassays, Vibrio fischeri, Microtox, interactive effects.

1 Introduction

The toxicity of various solid and liquid substances to aquatic life forms has started to be the subject of considerable research in the last decade. Quantification of the different compounds based on standard chemical analysis

www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060481

methods indicates the concentration of toxic compounds that released in the environment. Furthermore, the purpose of regulating acts issued by national and international authorities regarding the management and disposal of liquid and solid wastes to the environment is to protect both human health and ecosystem. However, most of these acts are relied on physicochemical analyses for the determination of physical and chemical parameters of waste or of receiving water body. As a result, the assessment of water quality contaminated with chemicals has been based on chemical analyses and the compliance or not of the measured concentrations with limit concentrations imposed by the legislation.

The toxicity tests in addition to chemical analysis can be used for the assessment of the direct effect of the compounds on the terrestrial and aquatic environment. Furthermore, chemical analysis methods do not take into consideration possible synergistic or antagonistic effects of the various compounds on aquatic life forms [1, 2]. Among the anthropogenic substances that are released in the environment, agrochemicals pose a serious risk of toxic effects on the ecosystem. Agrochemicals may be transported from the treated soils in aquatic environment by waterborne runoff or by direct deposition from the atmosphere [3]. Microtox test is among the widely used bioassays for the toxicity determination of various compounds, including agrochemical substances [3–6]. The aim of this study was the evaluation of toxic response of two agrochemicals and copper, when act alone or in combination, using the Microtox test.

2 Materials and methods

The agrochemicals examined for their toxic properties were: fosthiazate and metalaxyl-M using their commercial formulations of Nemathorin and Ridomil, respectively. The toxicity of copper was also evaluated using copper chloride dehydrate, ($CuCl_2{\cdot}2H_2O$) provided by J.T. Baker, Holland. The toxicity of agrochemicals and copper was evaluated using the bioluminescence bacteria *V. fischeri* (Microtox test) that were in freeze-dried form (SDI, USA) and activated prior to use by a reconstitution solution. Since *V. fischeri* is a marine organism, an adjustment of the osmotic pressure of the samples was applied to obtain samples with 2% salinity, using a concentrated salt solution (solution containing 22% NaCl in deionized water). The light emission of the test organisms obtained by their direct contact with the samples was measured using the Microtox 500 analyzer (SDI) within exposure times of 5, 15 and 30 min. The data processing was performed using the MicrotoxOmni software (SDI). The IC_{50} values (the tested compound concentration that caused 50% inhibition on the bioluminescence of *V. fischeri*; expressed as mg/L of fosthiazate, metalaxil-M or Cu) of the tested compounds were evaluated using the 45% basic test [7].

The interactive toxic effects between fosthiazate and copper or metalaxyl-M on *V. fischeri* were also investigated and the evaluation of the results was performed by statistical analysis. The concentrations of the tested compounds used were obtained from the IC_{50} estimation experiments of each compound. The theoretically expected effect of the binary mixtures was evaluated using a simple

mathematical model based on the theory of probabilities [1]. According to this model, if P_1 is the inhibition caused by a certain concentration of chemical A_1 and P_2 the inhibition caused by a certain concentration of chemical A_2, then the theoretically expected additive inhibition P(E), when those concentrations of the two chemicals are applied together, is given by the following equation:

$$P(E) = P_1 + P_2 - P_1 \cdot P_2/100 \tag{1}$$

The null hypotheses were that the observed values were higher or lower than the theoretically predicted values, for synergistic and antagonistic effects, respectively. The result was considered to be antagonistic or synergistic only if the observed effect was significantly lower or higher, respectively, than the theoretically predicted value at the 0.05 level of significance [8].

3 Results and discussion

The toxicities of the single agrochemicals and copper were first examined by applying several dilutions for each one of them and the IC_{50} value was evaluated as an endpoint. The IC_{50} values and the corresponding confidence ranges of the tested compounds, calculated using MicrotoxOmni software, are presented in Table 1, while the bioluminescence inhibition of *V. fischeri* caused by the tested compound concentrations are given in Figure 1 (dose response curves).

Table 1: IC_{50} values and the corresponding confidence ranges for the tested compounds on *V. fischeri*, for exposure time 5 and 30 min.

Compound	5 min IC_{50}, mg/L (C.R)	30 min IC_{50}, mg/L (C.R)
Fosthiazate (Nemathorin)	0.15 (0.13-0.18)	0.20 (0.17-0.25)
Metalaxyl-M (Ridomil)	0.57 (0.33-0.97)	0.88 (0.35-2.21)
Copper ($CuCl_2 \cdot H_2O$)	0.37 (0.32-0.44)	0.18 (0.17-0.19)

C.R.: confidence range

As it is shown in Table 1, the toxicity of both agrochemicals was almost constant within 5 and 30 min exposure time, while the toxicity of copper for 30 min exposure time was significantly higher than that of 5 min exposure time. The toxicities of copper and fosthiazate were generally comparable, while the toxicity of metalaxyl-M was slightly lower than those of the other two tested compounds. The toxicity of copper may be affected from chemical species formed by the hydrolysis reactions of metals and is mainly correlated to the free metal ion concentration [2]. Electrostatic interactions between ions presented in the solutions may cause a slow response of the bacteria. The diffusion of copper species through the cell membrane may be possibly delayed and the action might

not be completed within a short exposure time (i.e. 5 or 15 min). The assessment of IC_{50} values of various pesticides for 5 and 15 min showed that in most cases their action was completed within 15 min of exposure time; however certain pesticides might present a longer time effect, indicating that the type of the chemical might affect the test organism by different ways [4, 9]. Furthermore, the dose response curves of fosthiazate and metalaxyl-M showed a similar pattern, while a steep increase of copper toxicity was observed with small increase of its concentration.

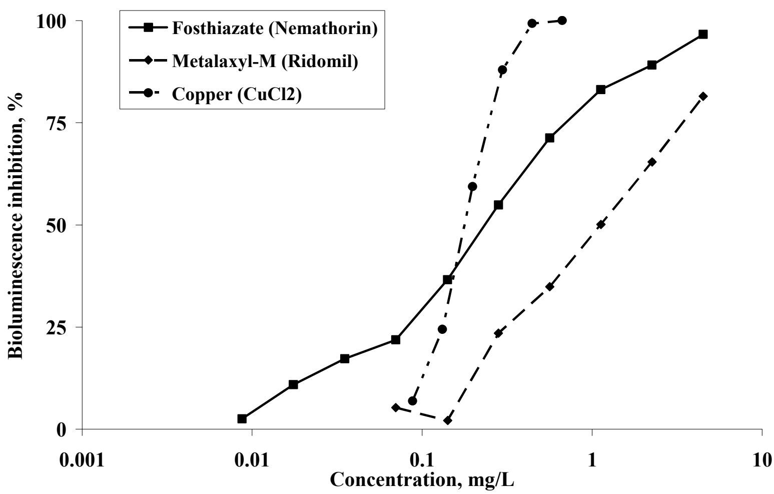


Figure 1: Bioluminescence inhibition of *V. fischeri* caused by the tested concentration of agrochemicals and copper, for exposure time 30 min (dose response curves).

After having tested the effects of the compounds alone on *V. fsicheri*, the interactive effect between fosthiazate and metalaxyl-M or copper were evaluated. The theoretical expected interactive effects, as calculated by Equation 1, and the observed effects for the combine action of fosthiazate and metalaxyl-M, as well as between fosthiazate and copper are illustrated in Figures 2 and 3, respectively.

The interactive effect between fosthiazate and metalaxyl-M was additive in most cases (Figure 2). For the concentration combinations B and G the observed bioluminescence inhibition was significantly lower than the corresponding theoretically expected, indicating an antagonistic action, while for the other concentration combination the effect was additive.

As shown in Figure 3, the observed bioluminescence inhibition of *V. fischeri* for the binary mixtures of fosthiazate and copper was significantly lower than the theoretically expected for all concentration combinations, indicating an antagonistic action. Similar results were found by Kungolos *et al.* [10] for the examination of interactive toxic effect of fosthiazate and copper on the

crustacean *Daphnia magna*. The antagonistic action between fosthiazate and copper could be attributed to the presence of additives contained in commercial formulation of fosthiazate, that may potentially reduce the bioavailability of copper.

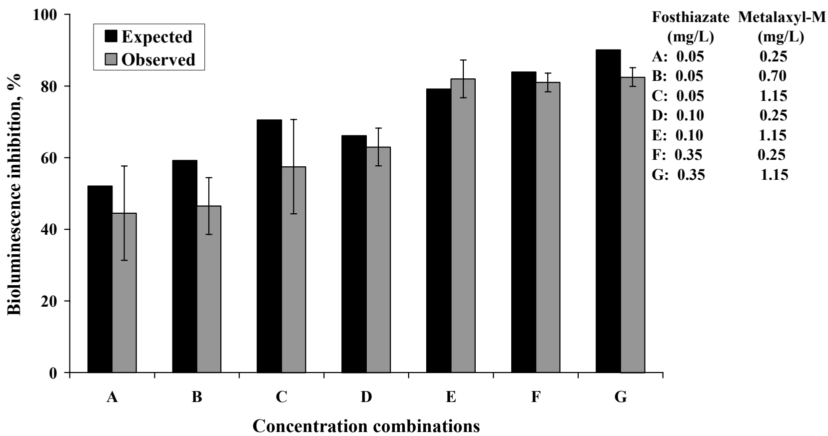


Figure 2: Comparison between theoretically expected and observed inhibitions for the combined effect of fosthiazate and metalaxyl-M on bioluminescence of *V. fischeri*.

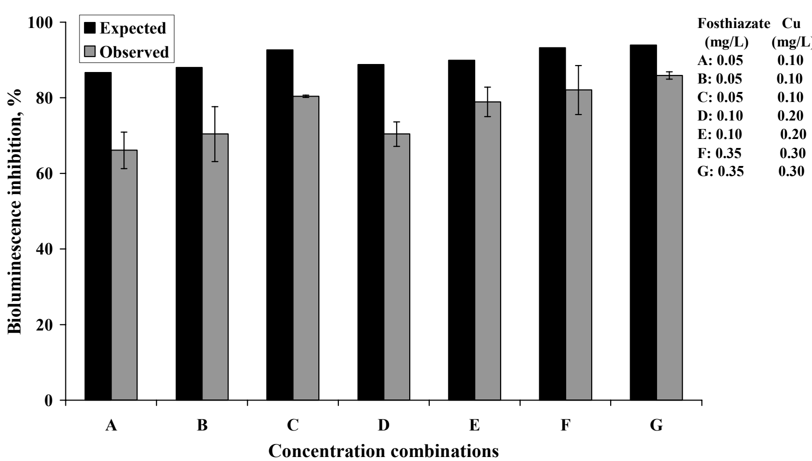


Figure 3: Comparison between theoretically expected and observed inhibitions for the combined effect of fosthiazate and copper on bioluminescence of *V. fischeri*.

4 Conclusions

The interactive effects of fosthiazate and metalaxyl-M mixtures on *V. fischeri* were found to be additive for most of the concentration combinations tested. For the binary mixtures of fosthiazate and copper the interactive effect was antagonistic for all concentration combinations tested. Both agrochemicals showed a similar toxic response; their toxicity was not significantly varied by the exposure time, while their toxicity was not sharply increased by the increase of concentration. On the other hand, the toxicity of copper increased significantly with the increase of exposure time and a steep increase of copper toxicity was observed with small increase of its concentration.

Acknowledgement

The present study was supported by grant of Ministry of Education and EC (programme Pythagoras II, contract number 52319.04).

References

[1] Kungolos A., Samaras P., Kipopoulou A.M., Zoumboulis A. and Sakellaropoulos G.P., Interactive toxic effects of agrochemicals on aquatic organisms. *Water Science and Technology*, **40(1)**, pp. 357-364, 1999.

[2] Tsiridis V., Petala M., Samaras P., Hadjispyrou S., Sakellaropoulos G. and Kungolos A., Interactive toxic effects of heavy metals and humic acids on *Vibrio fischeri. Ecotoxicology and Environmental Safety*, **63(1)**, pp. 158-167, 2006.

[3] Okamura H., Piao M., Aoyama I., Sudo M., Okubo T., and Nakamura M., Algal growth inhibition by river water pollutants in the agricultural area around Lake Biwa, Japan. *Environmental pollution*, **117**, pp. 411-419, 2002.

[4] Kaiser K.L.E. and Palabrica V.S., Photobacterium phosphoreum Toxicity Data Index. *Water Pollution Research Journal of Canada*, **26(3)**, pp. 361-431, 1991.

[5] Jones K.D., Huang W.H., Evaluation of toxicity of the pesticides, chlorpyrifos and arsenic, in the presence of compost humic substances in aqueous systems. *Journal of Hazardous Materials*, **B103**, pp. 93-105, 2003.

[6] Strachan G., Preston S., Maciel H., Porter A.J.R. and Paton G. I., Use of bacterial biosensors to interpret the toxicity and mixture toxicity of herbicides in freshwater. Water Research, **35(14)**, pp. 3490–3495, 2001.

[7] Microbics Corporation, Microtox Manual, *A Toxicity Testing Handbook*, **Vol. 1–5**, Carlsbad, CA, USA, 1992.

[8] Sanders D.H., Eng R.J. and Murph A.F., *Statistics, a fresh approach*. McGraw-Hill Company, New York, 1985.

[9] Ruiz M.J. López-Jaramillo L., Redondo M.J. and Font G., Toxicity Assessment of Pesticides Using the Microtox Test: Application to Environmental Samples' *Bulletin of Environmental Contamination and Toxicology*, **59**, pp. 619-625, 1997.
[10] Kungolos A., Tsiridis V., Nassopoulos H., Samaras P. and Tsiropoulos N. Toxicity assessment of fosthiazate, metalaxyl-M and imidacloprid and their interaction with copper on *Daphnia magna*. *Environmental Toxicology*, editors A. Kungolos, C.A. Brebbia, P. Samaras and V. Popov, WIT press, Southampton, UK, pp. 223-229, 2006.

Kaunas hydropower system management

J. Simaityte[1], J. Augutis[2] & E. Uspuras[2]
[1]*Vytautas Magnus University, Lithuania*
[2]*Lithuanian Energy Institute, Lithuania*

Abstract

The floods occurring in Europe recently confirm that some natural phenomena are unexpected and very difficult to control. Despite this, the proper subjects of infrastructure and scientific institutes try to find common decisions to minimize flood risk levels as much as possible. Currently the European Commission is reviewing legal documents and in the near future EU countries will have to implement flood management directives.

This paper presents a developed mathematical model of flood control which was applied for Kaunas hydropower system. The model consists of two parts: water level calculation algorithms, which were developed using only operational system data, and a risk based water level control model based on an optimization algorithm and integrated risk parameters. The model helps to accept risk based decisions during the flood period and offers alternatives of discharge and water level control near the dam. As a result the flood control characteristics and extreme scenarios are presented.
Keywords: risk based decision making, flood management, optimization model.

1 Introduction

Floods in recent years have brought huge economical losses and put a significant number of human lives in big danger. In 2006, the European Commission will review and negotiate six legislation acts and one of them is a flood management directive. On 18/01/2006 the European Commission proposed a directive on the assessment and management of floods [1]. Its aim is to reduce and manage the risks that floods pose to human health, the environment, infrastructure and property. Two trends point to an increased flood risk and to greater economic damage from floods in Europe. Firstly, the scale and frequency of floods are likely to increase in the future as a result of climate change, inappropriate river

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060491

management and construction in flood risk areas. Second, there has been a marked increase in vulnerability due to the number of people and economic assets located in flood risk zones.

This paper analyses the main problems of the flood control in the biggest Lithuanian hydropower system (HS) which consists of Kaunas hydropower plant (KHP) and dam, and Kruonis hydro pump storage power plant (KHPS). Both plants use water from Kaunas reservoir which is on the biggest Lithuanian river Nemunas. Extreme inflow from the river is the main initiative event which can impose failure of the system and must be considered, in order to keep acceptable the level of the risk and to ensure safety of downstream population, industry and nature.

Flood control problems in HS are multi objective and complex which requires an optimal solution between many parameters such as: flood inflow into reservoir, nominal water level near the Kaunas dam, nominal water level near the KHPS, discharge thorough the dam, acceptable risk level related with dam failure, acceptable water fluctuation in the reservoir and – also from economical point of view – useful reservoir capacity. The present work is concentrated on the situations related to extreme floods which are bigger than the millennium flood. Statistical analysis of historical river Nemunas floods revealed that floods with a peak of about 3000 m^3/s can occur once per thousand years. Such kinds of floods are highly dangerous for the Kaunas dam as the biggest discharge thorough the dam is 3990 m^3/s if all spillways and turbines operate properly. Although there is always the probability that the necessary actions will not be performed or it will be too late to ensure the necessary discharge thorough the dam. One of the purposes of the work is to analyse which scenarios of the flood control near the dam are most dangerous and how to manage the risk level so the economical losses and environmental harm would be as low as possible. The measure of the risk, solving this problem, was assumed as dam overtopping probability, which is the cause of about 50% of dam failures in the world [2].

The main goals of this work are: 1) water level (WL) control model development; 2) risk parameters integration into WL control model; 3) issues of alternatives of risk based strategies for WL control during extreme flood periods.

2 WL control model development

Decisions concerning large structures such as dams are based on different types of criteria, which range from strictly objective to highly subjective. Usually decision making issues used for risk assessment analyses are as follows [2]: a) regulatory type decision process based on standards and criteria; b) normative decision making with explicit or implicit decision rules; c) decision making based on multiple objectives (eg. multiobjective optimization).

The present water level control model was developed under the basis of an optimization algorithm model, as the main task is to create strategies which would find optimal solutions related to the water level near the dam during flood periods. The main input into the model is stochastic – river inflow into the reservoir during flood periods. The next step is real time flood forecasts based on

online data and historical information. The flood forecast results are integrated into the WL control model which uses the results of a discharge model, risk analysis, a decision making model and also evaluates the extra input. In this case input is extra flow from Kruonis HSP, which also depends on Ignalina operation in abnormal conditions and can be assumed to be stochastic.

The general scheme of the WL control model is presented in Fig. 1. The application of this scheme is explained in section 3.

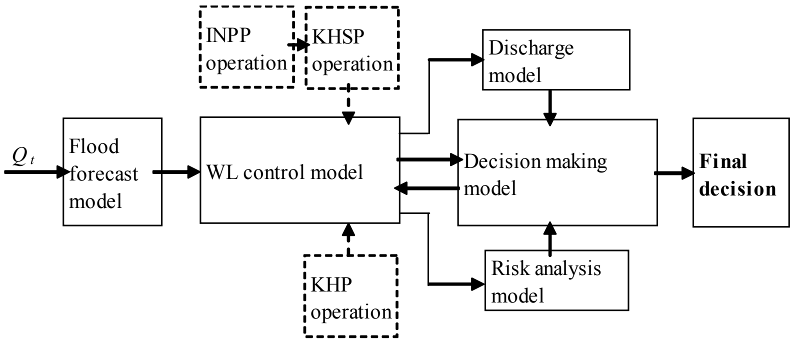


Figure 1: Water level control model general scheme.

2.1 Water level optimization

The formulation of any water recourses optimization problem should incorporate all elements that are significantly relevant to system effectiveness [3]. The framework upon which effectiveness is evaluated consists of three components: declaration of decision or independent variables, and state or dependent variables, at time *t*, a discrete time or continuous objective function, or performance criterion *Z*, to be minimized; and a series of constraints that provide for feasible management strategies. In general terms, a discrete time engineering management problem might be formulated.

Let us construct the optimization model for the water level near the dam problem. As the main task of the problem is to find the best strategies for water level control, than the main equation should refer to it:

$$Y_t = F(Q_t, Q_t^K, D_t^{Tr}, D_t^G, t), \tag{1}$$

where Y_t – water level near the dam at time *t*; Q_t – inflow to the reservoir from the river at time *t*; Q_t^K – inflow to the reservoir from other sources (artificial reservoir) at time *t*; D_t^{Tr} – discharge from reservoir thorough plant turbines at time *t*; D_t^G – discharge from reservoir thorough gates at time *t*.

The function of Y_t at the moment *t* is interpreted as a deterministic function, which has known parameters. But, in fact, the input parameter Q_t and sometimes Q_t^K are stochastic parameters, therefore the evaluation of function Y_t can be not so easy. The problem was solved using a simplified operational model, which let one update function Y_t parameters at every time *t* step using a recursive link and

operating controllable input variable $D = D_t^{Tr} + D_t^{G}$. The stochastic model input is flow into the reservoir $Q = Q_t + Q_t^{K}$. The model scheme is presented in Fig. 2.

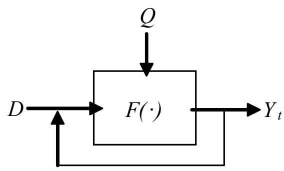


Figure 2: Water level function parameters estimation.

Note R_t as a dam overtopping probability dependent on water level near the dam, which indicates hydropower system risk level. Value R_t can have various expressions, such as dam overtopping or dam failure probability, number of fatalities during dam failure, economical losses, etc. Water level function Y_t is a variable dependent on many contributors, which also are the main elements in the discrete objective function *Z*, so we have the optimization model:

$$Min\, Z = \sum_{t=1}^{T} f(R(Y_t), Y_t, Q_t, Q_t^K, D_t^{Tr}, D_t^G, t) \cdot \tag{2}$$

The constraints of the model in the WL control problem have the following expressions:

$$\begin{aligned} &D_t^T = G_1(Y_{t-1}, t), \\ &D_t^G = G_2(\,Y_{t-1}, t), \cdot \\ &\underline{Y} \le Y_t \le \overline{Y}. \end{aligned} \tag{3}$$

These constraints show that the discharge also depends on water level, which makes the problem more complicated. *G*(·) is the functional relationship between the water level and discharge. When Y_{t-1} is given, than D_t^{Tr} and D_t^G are obtained by using interpolation methods. The inequality in (6) represents the boundaries of the water level fluctuations, which also can be defined as functions. These boundaries are related with the restrictions for WL fluctuations.

As the function *Z* in this case is discrete, using numerical calculations we solve (5) and (6) and under the given uncontrollable values Q_t and Q_t^K we find the optimal values of $R(Y_t), Y_t, D_t^{Tr}$ and D_t^G which are used in the decision making model. If these values agree with risk criteria, than the final decision is formulated for the WL control strategy at a time *t*, *t*+1, … *T*, where *T* is a time interval with sufficient confidence level for Q_t forecast.

2.2 Water level risk based control

The interaction of the Risk analysis model and Decision making model (Fig.1.) generates alternatives according to the flood situation. After evaluating the water

level at the dam, it is necessary to evaluate the risk level which the system reaches at a time t (dam overtopping probability) [4]. Than using the flood forecast model the water level Y_t is evaluated for a few days onrward, $t = 1, 2, \ldots n$, where n is the number of days, hours or another time lag. If after n time intervals WL does not reach the critical water level Y_{cr}, than the decision alternatives for the next time lags are selected using the deterministic WL model (5), which evaluates the necessary discharge thorough the dam and does not directly integrate additional risk parameters.

The risk based decisions are important, when after a selected time interval the situation is critical, and

$$Y_n - Y_{cr} = r\,,\, r > 0. \tag{4}$$

This means that the water level exceeded the critical dam water level and loads on the dam body are unacceptable. The risk level is calculated considering several scenarios that in the time interval [1, n] it is impossible to manage the situation, and the probability of overtopping is evaluated. If the evaluated probability exceeds the criteria, developed for a specific dam, the extra control actions must be activated: the water level must be lowered in advance by the value r, in order to avoid the unacceptable risk level.

It is necessary to mention that the WL control model in this situation not only considers the risk based decision but also at every time step re-evaluates the objective function Z (6). And the last issue of the modelling in this situation is the final decision, which presents the strategies of discharge thorough the dam and keeps an acceptable risk level for the system.

3 Risk WL control model application

3.1 Data analysis and functional relationships

The developed WL control model (section 2) was applied for Kaunas hydropower system. The primary analysis steps were to find functional relationships between the main hydropower system parameters: Y_t, Q_t, Q_t^K, D_t^{Tr} and D_t^G. For this purpose the statistical and operational hydropower system data were used. The dependence of inflow and flood volume can be expressed as

$$V_{t+1} = (Q_t + Q_t^K) \cdot \Delta t\,, \tag{5}$$

where Δt is one day time interval between t and t+1 with unit in s. Flow measurements Q_t and Q_t^K are in m^3/s and food volume V_{t+1} is in m^3.

The operational data showed the relationship between reservoir volume V_t^R and WL near the dam. A non-linear dependence at time moment T was obtained:

$$Y_t = \alpha_1 \log V_t^R + \beta_1\,. \tag{6}$$

where the parameters evaluated were $\alpha_1 = 7.24$ and $\beta_1 = 0.08$.

The other important relations are between the WL, discharge thorough the gates and discharge thorough the turbines. Analysing KHP operational data it was noticed that significant water amount during the flood period used to be discharged thorough the turbines. According to the specifications, the discharge through the turbines depends on the downstream water level and the upstream water level near the dam. A nonlinear dependence of the discharge thorough the turbines and the difference of water levels was determined, when four turbines operate (Fig. 3.):

$$D_t^{Tr} = \alpha_2 H_t^2 + \beta_2 H_t + \gamma_2 , \tag{7}$$

where $H_t = Y_t - Y_t^{low}$, Y_t^{low} – downstream WL of the dam. The parameters evaluated respectively were $\alpha_2 = -8.852$, $\beta_2 = 255.22$, $\gamma_2 = -1135.2$.

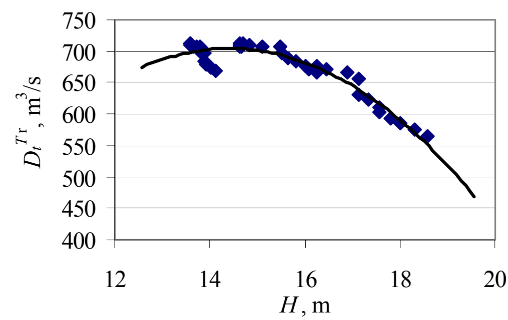


Figure 3: Relationship between turbine discharge s and difference of levels H to the Kaunas dam.

The functional relations of discharge thorough the gates are linear:

$$D_t^G = \alpha_j Y_t + \beta_j . \tag{8}$$

The evaluated coefficients of the equation are presented in Table 1. Coefficients α and β are dependent on c – how much the gate is opened, in meters.

Table 1: The coefficients of linear dependence between WL and discharge.

c	A	β
0.41 m	1.01	37.92
0.84 m	10.49	56.06
1.73 m	13.62	137.35
2.21 m	15.00	209.00
2.69 m	38.37	-1301.10
3.18 m	23.536	326.57
7.00 m	189.39	-7528.00

3.2 Kruonis HPS and Ignalina NPP operation

The Kruonis hydro pump storage power plant was constructed as the main auxiliary power supply system for the peak energy demands. The main function of KHSP is the balancing of electricity generation and consumption as well as the prevention of accidents and their liquidation in the power system. To liquidate the capacity deficit in the event of the disconnection of the Ignalina Nuclear Power Plant (INPP), the Kruonis Pumped Storage Plant generators are automatically launched into operation. KHPS uses other systems energy for water supply from Kaunas reservoir to its artificial reservoir, and when there is a large energy demand it is working on turbine mode and producing electricity. KHSP artificial reservoir volume is 48.78 millions m^3, installed capacity of the pumped storage plant 900 MW (4 units, 225 MW each), maximum discharge at turbine mode 226 m^3/s. Usually, when Ignalina NPP operates in normal regime than 1, 2 or sometimes 3 turbines are working in KHSP.

In risk based decision making analysis the important thing, concerning KHSP and INPP, is the plants operation during a flood period. The most critical period would occur if the WL near the dam already reaches maximum limit, and INPP shutdowns, than KHSP would start to work automatically by total capability and would give extra inflow into Kaunas reservoir. This scenario was included into the model and influenced the most critical decisions made during extreme flood periods.

3.3 Risk analysis model and flood forecast

The risk analysis model [4] was complemented by additional analysis including more parameters: the water level at the dam at every time step and the inflow of the KHPS, as these can influence the system risk level. The risk level value for the present model step was chosen as discharge unavailability through the dam probability, which was calculated at every time step using a developed event and fault tree model. When the flood begins, the algorithm does extra calculations which indicate the discharge unavailability probability dependent on the current situation, i.e. if 0, 1 or 2 gates are already opened and flood forecast results are known, what is the probability that 3, 2, 1 or 0 gates will fail to open. Virtually, in this case, the evaluated probability also means the dam overtopping, if no action will be taken to emerge successfully from the critical situation. For example, if one gate is opened and extreme flow is coming into the reservoir, it is possible to evaluate the necessary amount of water which must be discharged thorough the dam in the further time steps. Let us assume, in this situation one gate must be opened during the following two time steps (two days). So if one gate will fail to open, the water level will increase until the critical and dam overtopping will occur. In this example, dam overtopping probability will be equal 2.0E-04. According to the assumed criteria, extra actions must be taken to lower the evaluated risk level; therefore, in order to have some conservatism, it is recommended to increase the discharge as soon as possible in order to have acceptable risk levels in two days. If the risk level is acceptable, no extra actions are necessary at the moment, only those which are foreseen for ordinary flood control.

Table 2: Probabilities to fail to open the gates at different scenarios.

Number of opened gates	Probability to fail to open the rest or the gates			
	No fail	1	2	3
2	9.999E-01	1.000E-04		
1	9.998E-01	2.000E-04	1.000E-08	
0	9.997E-01	2.999E-04	3.000E-08	1.000E-12

The evaluated probabilities to fail to open the necessary number of gates are presented in Table 2. The calculations are performed for different scenarios dependent on the current situation: that already 2, 1 or 0 gates are open.

This part of the modelling is necessary only for a very narrow spectrum of situations: for millennial floods, for which the peak probabilities are from 1.0E-04 to 1.0E-05. For preliminary analysis these probabilities were chose as risk level criteria. If the flood peak is less than the evaluated millennial flood, then the situations usually are not critical and the dam manages to emerge successfully. If it is bigger, then extra actions will not help to manage the situation.

The results of the risk analysis model are used in a further step – the decision making model, which integrates all available information: actual river flow and its forecast, water level, discharge and also KHSP flow as an unexpected event at a time t.

The principle of the forecast is statistical analysis of the historical spring flood, which exceeds the base level. When the individual flood has its first day flow measurements, the model analyses historical floods, and attributes the special weight coefficients for each flood flows and then the future flood flows are calculated. The forecast model parameters are updated every day, using Bayesian techniques, when new flow data are included into the model. To obtain better results the probabilistic model must be integrated with a deterministic hydrological flood forecast model. The details of the flood forecast model can be found in [5].

3.4 Decision making model

A risk based decision making model was applied to analyse the most critical situations in Kaunas HS related with extreme floods and all characteristics described in the previous sections, where evaluated using extreme flood data.

Let us analyse the situation of spring floods with maximum peak of 1055 m^3/s (1999 year flood). The developed water level model, presented in section 2.1, was applied in order to analyse the possible discharge alternatives. For the same situation applying the risk based decision model the results were found to be the same as applying the developed deterministic WL control model without risk parameters, because the risk levels were not exceeded in all time steps. An example of the results is presented in Figure 4.

The situation can be different when the coming flood peak is bigger than 3000 m^3/s. In this case any failure of the actions can bring overtopping or even

the failure of the dam. For the decision model application the biggest known flood flow data from 1958 was chosen, for which the peak was 3450 m^3/s. Using the developed model the necessary discharge thorough the dam was evaluated every day, performing presented optimization WL algorithms and using a 3 day forecast. At every time step the model proposes the discharge, which must not exceed the acceptable risk level. Fig. 5 presents an example of one of the generated alternatives, which was accepted using the risk based decision model.

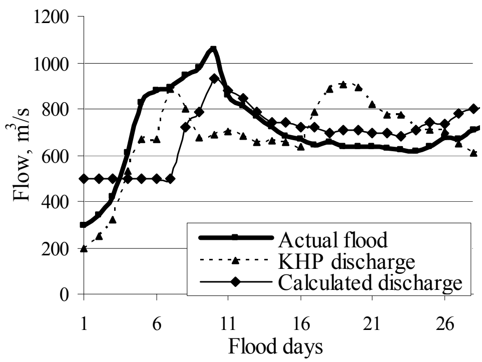


Figure 4: Calculated and operational discharge data for 1999 flood.

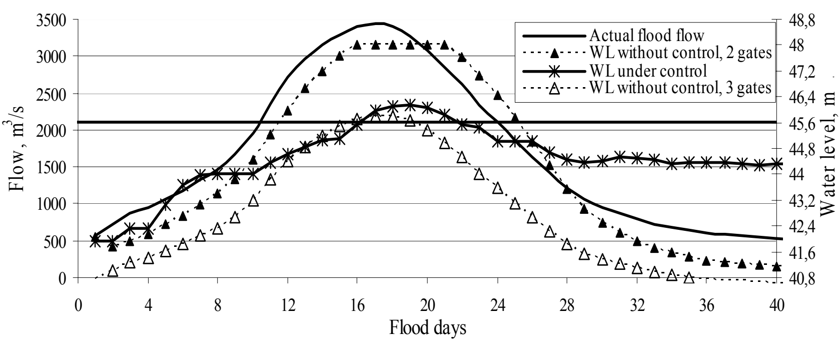


Figure 5: Calculated and operational water level data for 1999 flood.

The solid curve in Fig. 5 shows the flow data of the actual flood. The dashed curve is the water level which would be reached in uncontrollable situation: WL reaches the top of the dam in 14 days when only 2 gates are opened and WL reaches 45.6 m in 15 days when all 3 gates are opened during all the flood period from the very beginning. The scenario with 3 gates open presents the other extremity: if discharge is very big from the very beginning, dam downstream

areas can be flooded and economical loses obtained in power generation. So the task is to find a solution in between these two scenarios. The “WL under control” curve shows the water level which it is possible to achieve if the risk based decision model is used: no overtopping of the dam occurs and the critical water level 45.6 m is exceeded only for 6 days. Also extra conditions were assumed that when the flood passes, the water level near the dam remains nominal (44 m). This was also achieved using the WL control model.

4 Results and conclusions

1. A risk based mathematical model was developed and applied for water level control during extreme flood periods in Kaunas hydropower system.
2. Applying the developed water level control model for floods, for which peaks are less than 3000 m^3/s, the alternatives which were obtained under the risk based decisions are the same as the water level calculated by not applying risk analysis.
3. Risk based decisions for millennial floods decrease dam overtopping probability.
4. Applying the risk based decision model the alternatives of WL control strategies, near the Kaunas dam, are provided.

References

[1] Proposal for a Directive of the European Parliament and of the Council on the assessment and management of floods, SEC(2006) 66, Brussels, 18.01.2006.

[2] Graham L.P., Bartsch M., Risk analysis safety assessment for use at Swedish dams, Proc. of Symposium on Research & Development in the field of Dams, Switzerland, 1995.

[3] Nicklow J.W., Discrete-Time Optimal Control for Water Recourses Engineering and Management, Water International, Volume 25, Number 1. P. 89-95, ISSN 0250-8060.

[4] Simaityte Volskiene J., Augutis J., Uspuras E. Kriauciunene J. Risk Analysis of Kaunas Hydropower System // Proceed. of the 4th on Computer Simulation in Risk Analysis and Hazard Mitigation, Rhodes, Greece, 27-29 September 2004. ISBN 1-85312-736-1. - P. 554-561.

[5] Simaityte Volskiene J., Uspuras E., Augutis J. Flood Forecast Model and Probabilistic Analysis // Proceed. of the 7th Intern. Conf. on Probabilistic Safety Assessment and Management, Berlin, Germany 14-18 June 2004. - Springer-Verlag. ISBN 1-85233-827-X. - 2004. - P. 3079-3085.

[6] Simaityte Volskiene J. Maximal Floods forecast of the Nemunas River (in Lithuanian) // ISSN 0235 – 7208. Energetika, 2005, Nr.1. P. 34-39.

Particle-tracking method applied to transport problems in water bodies

F. C. B. Mascarenhas & A. E. Trento
Computational Hydraulics Laboratory,
Federal University of Rio de Janeiro, UFRJ, Brazil

Abstract

In this work a particle-tracking model was implemented and applied to simulate the instantaneous injection of a dye tracer in the Setubal Lagoon, Santa Fe, Argentina. It focused on the study of the water quality problem generated in point sources of contaminated water through a rainfall drainage system. This problem is difficult to treat with the so-called Eulerian methods. The interpolation for the field velocity, the number of particles used in the simulations, the representation of boundary conditions and the mesh size for the calculation of concentrations are discussed. A field experiment using tracers is described in order to compare the numerical results and the simulations with real data. The diffusion coefficients calculated from the experiment and the hydrodynamic modelling give the necessary inputs for the simulations using the particle-tracking modelling of the measured four clouds. The results support the main advantages of this method in terms of strict mass conservation, good representation of high concentration gradients and point sources and as a consequence, a good representation of the maximum concentration and the general form of the clouds.
Keywords: particle-tracking, Lagrangian, dispersion, Setubal.

1 Introduction

Point sources in water quality problems, like rainfall runoff outlets, are difficult to represent in the so-called Eulerian methods, and the particle-tracking method may represent a good alternative for solving this kind of transport problems. In a pioneer work Chorin [1] employed the particle-tracking method, solving the Navier-Stokes equations, for determining the vorticity field of the two-

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060501

dimensional flow around a circular cylinder with high Reynolds numbers. Since then, the particle-tracking models have been applied to a wide range of engineering problems, for example, cohesive sediment transport (Vinzón and Paiva [2]); simulation of solute transport in coastal waters (Periañez and Elliot [3]). This method is included into the Lagrangian class, where the contaminant is transported as a consequence of the random movement of a large number of discrete particles. The method is free of oscillations as well as of numerical diffusion, and therefore it is also free of negative concentrations and mass loss. These problems frequently appear in the traditional finite differences or finite elements Eulerian methods, when the transport is dominated by advection, or the sources are too small related to the hydrodynamic mesh size, mainly around the sources or where strong concentration gradients take place. Particle tracking methods are particularly suited for this kind of problems (Periañez and Elliot [3]). The mesh for the calculation of the concentration is independent of the mesh used to solve the hydrodynamics, which makes it suitable for the representation of point or linear sources. The main disadvantage of the method lies in the fact that the number of particles (*NP*) used to represent the contaminant should be sufficient to represent the concentrations in the mesh, and then the solution may not be a smooth function, due to the stochastic nature of the dispersion stage. Nevertheless, increasing the *NP* can result in a significant increase of the computational time.

In this work we implemented a particle model for the passive substance transport and applied it to the case of the instantaneous injection of a dye tracer in the Setubal Lagoon, Santa Fe, Argentina. A tracer experiment was accomplished in order to evaluate the characteristic transport parameters of the receiving water body, e.g. the dispersion coefficients. The experiment involves the instantaneous release of a dye tracer and the recovery of the cloud downstream. The particle-tracking method is discussed, pointing out its main advantages and disadvantages, as well as some details of its application. The case study is described, and the modelling results are compared with field data.

2 Random walk particle-tracking method for non-homogeneous turbulence

Mathematically, the transport process can be expressed, in a one-dimensional domain, as follows:

$$x^{n+1} = x^n + \left(U + \frac{dD}{dx} \right) \Delta t + z\sqrt{2\,D\,\Delta t} \tag{1}$$

where x^{n+1} and x^n are the positions at integer time indexes n+1 and n, respectively, z is a random number having normal distribution with zero mean and unit variance, U is the fluid velocity and D is a non-homogeneous diffusion coefficient. The second term on the right hand of equation (1) includes the particle advection $(U\,\Delta t)$, and a “drift” term, as a result of the non-homogeneous dispersion, due to the spatial variation of the dispersion coefficient. The third term in Eq. (1) represents the random displacement of the particles. Hathhorn [4]

has demonstrated the equivalence of results using the random number *z* normally as well as uniformly distributed. If one adopts a uniform distribution, this may imply a considerable computational time saving. Therefore the normally distributed random number *z* in the range (-0.5, 0.5) can be obtained by a commonly computer generated random number *r* in the range (0, 1), with mean 0.5 and standard deviation $\sqrt{1/12}$. In this case, the Eq. (1) must be expressed as:

$$x^{n+1} = x^n + \left(U + \frac{dD}{dx} \right) \Delta t + (2r - 1)\sqrt{6\, D\, \Delta t} \qquad (2)$$

The presence of sources is represented by the injection of a finite *NP*, at a rate corresponding to the simulated substance discharge. The mass of each particle *i* is represented by a mass quantity *MP*i. The particles move at each time interval Δt due to the combined effect of advection and dispersion. When choosing the time and space steps we imposed the condition that the maximum particle path could not be greater than the cell size, thus avoiding the extrapolation of the velocity corresponding to the original position of the particle, beyond its neighbourhood. This is an important condition for non-uniform field velocities.

For a continuous source the mass is injected at each time interval under a certain rate $M/\Delta t$, and for an instantaneous source all the mass *M* is available at $t = 0$. In this work, the concentrations are evaluated by counting the number of particles in each grid cell, multiplying them by the mass of each particle, dividing the total mass by the cell volume and assigning this concentration value to the cell centre. The most important factors to be considered to achieve good simulations of the concentration distribution are the *NP*, the velocity interpolation procedure and the boundary condition prescription.

2.1 Computation of the particle velocities

For the advection of passive substances, the velocity of each particle is considered equal to the stream velocity where the particle is located, obtained by interpolation over the hydrodynamic mesh at each time step. The selected algorithm for the interpolation is a key feature of the model (Narayanan *et al.* [5]). In this work we implemented linear, bilinear and second order interpolations, based on Taylor series expansion. Using Molenkamp's classical experiment, it was observed that a second order interpolation is more accurate than the bilinear scheme, while the linear method produced a completely erratic path of the particles and, as a consequence, unacceptable errors for the advective transport.

2.2 Initial and boundary conditions

The model requires the prescription of initial and boundary conditions. The initial condition is given by the nature of the problem. For our case, a two-dimensional horizontal transport problem where the initial condition is an instantaneous linear source of dye tracer, an *NP* proportional to the mass injected is considered at a single point where the source is located. As regards the open boundaries, the particles are considered lost when they leave the domain. For closed boundaries, the prescribed condition is zero flux through the boundary.

3 Case study: Setubal Lagoon (Argentina)

The Setubal Lagoon is located in the alluvial valley of the Parana River (60°20'W and 60°40'W, 31° 20'S and 31° 37'S), province of Santa Fe, Argentina. It is part of a complex system composed of lagoons and creeks connected to this river. Most of the lagoon bottom is almost flat and strong bathymetric gradients are observed at the outlet reach. Its width ranges from 2000m, with a mean depth of 3m, to 300m and depths of up to 25m. The bottom is composed mainly of silt and clays and, to a less extent, of fine sand. On the right margin (RM) of the lagoon, along 5 km, Santa Fe city (360,000 inhabitants) is situated**.** Along this coast there are storm drains, and mixing of the rainfall water with sewer discharges causes environmental concerns, since the lagoon beaches are used for recreational purposes and a water intake for the city water supply is located downstream. Figure 1 shows the location of Setubal Lagoon and the position of the main storm drainage pipes. Several studies have been done related to the water quality problem of the Setubal Lagoon. Tessi *et al.* [6] described the high bacterial contamination indexes in the waters and sediments at the beaches, which match the maximum values established by the European standards (WHO [7]) and were adopted as a policy by the local Government. These facts motivated the need of studying the contaminant dispersion dynamics.

A dye tracer experiment was performed with this objective, and the tracer clouds were simulated using a 2-DH particle-tracking model. The hydrodynamics of the lagoon was calibrated and used for simulating the flow occurring during the experiment.

3.1 Tracer experiment

At the point indicated in Figure 1, an instantaneous tracer injection was done, using two tracers, Amidorhodamine G (Acid Red 50, colour index 45220) and Uranine (Acid Yellow, colour index 45350). At 11:21 a.m., 1.0 kg of Amidorhodamine G and 1.0 kg of Uranine, dissolved into 70 litres of water, were released from a boat. Based on previous results from the hydrodynamic model (Trento *et al.* [8]), five cross sections (*A, B, C, D* and *E*) were established for measurements, at about 1000, 2700, 4200, 6000 and 8000m from the injection point. At each cross section, the boat moved forward and backward, using DGPS for positioning (with an estimated error of about 5m). For the tracer cloud detection, water was continuously pumped through two serial fluorimeters, one equipped with filters for detection of Amidorhodamine G and the other for Uranine. Considering the low depth of the lagoon (about 3m), the measurements were carried out only at middle depth, expecting a complete mixing in the water column. A vertical profile for each cloud was measured in order to verify this hypothesis.

Water samples (400) were also collected for laboratory analyses. The minimum detection limit of the tracer under laboratory conditions was 0.05 ppb (0.05 mg/m^3), while for the on site detection it was ~0.2 ppb. The results obtained from the laboratory analyses were complemented with the on site

measurements. The samples were collected at 10s intervals at the cross sections *B* and *C* and every 20s at the cross sections *D* and *E*. As the laboratory analyses showed that the cloud at cross section *A* was poorly measured, it was not taken into account. Furthermore, only the tracer Amidorhodamine G was considered, since the tracer Uranine showed significant losses. These losses can be related to the long storage time (approximately 3 months) between the experiment and laboratory analyses. The clouds were thus measured in a Eulerian way, recording the curves as the substance passed through a particular cross section for each cloud. In order to reconstitute the clouds in a Lagrangian way, the flow velocity was then determined as the distance and time between the maximum concentrations of two consecutive clouds. The new coordinates of each cloud point were evaluated according to $X_i^{'} = X_i + u_x\,(t_i - t_0)$ and $Y_i^{'} = Y_i + u_y\,(t_i - t_0)$ where $X_i^{'}$, X_i, and u_x are the new and original coordinates and velocity in the East-West direction; $Y_i^{'}$, Y_i, and u_y are the new and original coordinates and velocity in the North-South direction; t_i is the *i-th* sampling time and t_0 the time corresponding to the occurrence of the maximum concentration in the cloud.

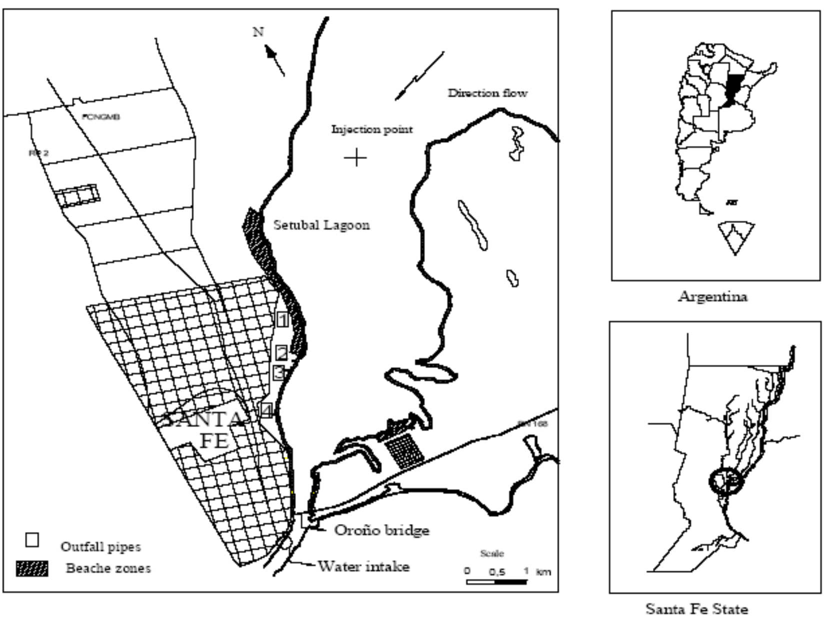


Figure 1: Geographic location of the Setubal Lagoon and the main four storm water pipes.

3.2 Dispersion coefficients from the field data

According to Holley and Jirka [9], for natural flows it is more adequate to adopt a “longitudinal” axis, following approximately the mean longitudinal flow direction, and a “transverse” axis, orthogonal to it, than a geographic coordinate

system (North-South Y axis and East-West X axis). Considering the longitudinal and transverse concentration distributions, if they follow a Gaussian profile, it is possible to adopt the method of the moments (Harleman [10]) to determine the longitudinal and transverse dispersion coefficients. The symmetrical form of the clouds *B, C* and *D* indicated a Gaussian behaviour, thus justifying the use of this method. Once the longitudinal (S_L^2) and transverse (S_T^2) variances were obtained, the dispersion coefficients for each cloud were evaluated using the following relationships:

$$D\varepsilon = \frac{\Delta S_L^2}{2T_d} \tag{3}$$

$$D\eta = \frac{\Delta S_T^2}{2T_d} \tag{4}$$

where ΔS_L^2 and ΔS_T^2 are the differences in the longitudinal and transverse variances between two successive clouds, respectively, and T_d is the effective diffusion time, or time associated to the transit of their maximum concentration. These results are shown in Table 1.

Table 1: Effective diffusion time (T_d), longitudinal (D_ε) and transverse (D_η) dispersion coefficients.

	X (m)	Y (m)	Cmax (ppb)	T_d (s)	Dist (m)	S_L^2 (m^2)	S_T^2 (m^2)	D_ε (m^2s^{-1})	D_η (m^2s^{-1})
Injection	1950	8846	----	----	----	----	----	----	----
Cloud B	1539	6246	25.19	10200	2632	7355	468	0.36	0.023
Cloud C	1595	4798	12.67	15900	4063	24289	542	0.76	0.017
Cloud D	1344	3006	6.64	21060	5871	60472	456	1.43	0.011
Cloud E	848	1098	2.21	26310	7953	----	----	----	----

4 Numerical simulation of the clouds

The hydrodynamic model is two-dimensional, vertically averaged and based on finite differences (Borche Casalas [11]). The modelling domain was defined in order to cover the area of the tracer experiment. It was divided into 3456 square elements of 100m lengths. The adopted time interval according to a stability condition was 15s. The bathymetry was surveyed for the modelling purpose in a mesh that ranged from 25m to 100m. The surface slope, the wind velocity and the discharge for the hydrodynamic conditions occurring during the tracer experiment were measured, resulting in 0.025 cm/km, 6 kmh^{-1} in the NE direction, and 1560 m^{3}s^{-1}, respectively. With these data, the Manning coefficient *n* was adjusted in a value of 0.028 sm$^{-1/3}$. Figure 2 shows the bathymetry and the velocity field computed and used for the transport model. The dispersion coefficients considered were those evaluated from the tracer experiment, shown in Table 1. The coefficient obtained for the reach C-*D* was also considered for the reach *D-E*. The time step for the transport was 30s. The interpolation of the velocity and depth for each particle position was obtained following the bilinear scheme. The boundary condition at the lagoon margins was of zero flux. The

tracer mass, M=1000 g, was represented by 25000 particles of equal mass, while the concentrations of each cloud were calculated over a mesh with $\Delta x = \Delta y = 50$ m. The choice of that particle number is a compromise between the mesh size and the representation of the maximum concentrations observed in each cloud.

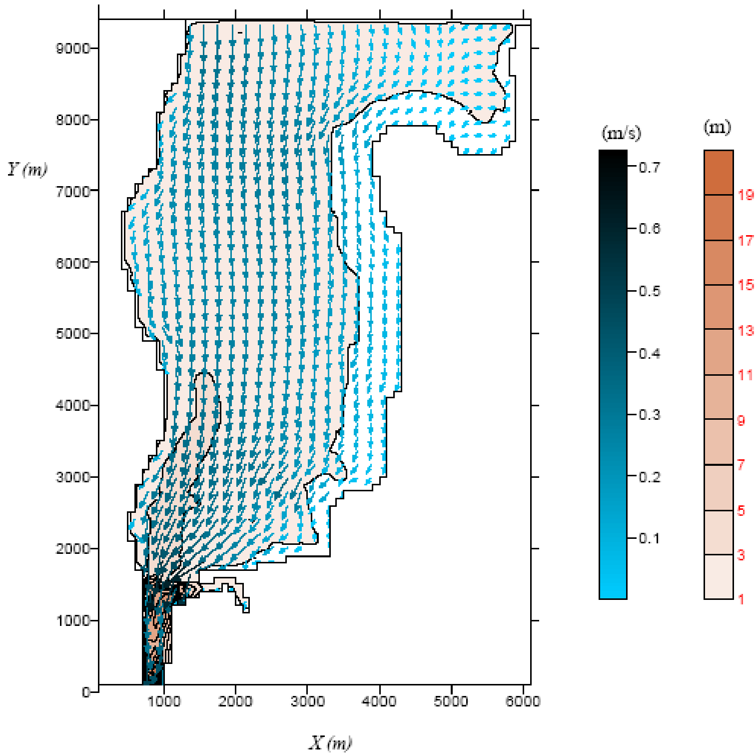


Figure 2: Bathymetry and velocity field obtained with the hydrodynamic model for the flow conditions occurring during the tracer experiment.

The results obtained with the transport model are shown in Table 2, in terms of the coordinates X and Y of the centre of mass of each cloud and their corresponding maximum concentrations. Cloud E has a smaller width than the preceding clouds probably due to the stretching produced when entering in the convergent South part of the lagoon. Considering the variances calculated from the model results for cloud E, starting at the injection point, σ_ε^2 (t)=83065 m^2 and σ_η^2(t)=239 m^2, the dispersion coefficients would be D_ε =1.579 m^{2}s^{-1} and D_η= 0.004 m^{2}s^{-1}. Considering the variances between successive clouds, a negative value D_η, would be obtained for cloud E. We verified a non-linear relationship for variance and time for the transverse direction, indicating that the dispersion process is not "Fickean" at this particular reach of the lagoon. The positions of

the clouds *B, C, D* and *E*, measured and computed, displayed in Table 1 and Table 2, respectively, are shown in Figure 3. The differences in the position of the centre of mass were 32, 37, 27 and 51m. It is important to stress that the measured clouds were actually re-constructed from the passage curves, as explained before. This means that the velocities at all the cloud points were considered uniform and equal to the mean velocity of the whole cloud in the foregoing reach. Table 3 shows a comparison of the maximum concentrations calculated, where we also included all the interpolation schemes (linear, bilinear and second order), which show non-significant differences between them.

Table 2: Results of the numerical simulation: coordinates of the centre of mass and maximum concentrations for clouds *B, C, D* and *E*.

	X (m)	Y (m)	Cmax (ppb)
Injection Point	1950	8846	
Cloud B	1525	6275	24.48
Cloud C	1625	4775	13.20
Cloud D	1325	3025	6.81
Cloud E	816	1058	1.49

In order to compare the model results using the dispersion coefficients obtained experimentally and those calculated according the classical equations by Elder [12] for turbulent flows in channels, the model was implemented using the following expression for longitudinal ($D_\varepsilon = 5.93\ u_* H$) and transverse ($D_\eta = 5.93\ u_* H$) dispersion coefficients, where u_* is the shear velocity ($=\sqrt{\tau_0 / \rho}$), τ_0 is the bottom shear stress and ρ the fluid density. The result of this application is also included in Table 3. The results are similar to those previously obtained. In this case the differential advection is only significant in the flow direction, as the transverse velocity gradients are low, and thus it would match the Elder's hypothesis of turbulent diffusion in the transverse direction.

Table 3: Maximum calculated and measured concentrations (ppb) with the interpolation schemes (linear, bilinear and second order).

Cloud	Velocity Interpolation Scheme			Elder coef. and Bilinear Velocity Interpolation Scheme	Maximum Measured Concentrations
	Linear	Bilinear	Second Order		
B	23.98	24.48	24.37	25.16	25.19
C	14.36	13.20	13.15	17.10	12.67
D	6.83	6.81	6.69	10.65	6.64
E	1.81	1.49	1.44	1.62	2.21

5 Conclusions

The particle-tracking method allows a strict mass conservation and the representation of high concentration gradients without spurious oscillations or numerical diffusion. The method is also flexible for calculating the results in a

grid independent from the hydrodynamic method. This also represents an advantage for the kind of applications of interest, where generally the size of the cloud or plume is smaller than the size of the mesh used for solving the velocity field. Besides considering an adequate size for this post-processing grid, compatible with the scale of the transport problem, special care must be taken in the number of particles selected for the simulation. The relative error of the solution decreases in a quadratic way with the increase of the number of particles, and therefore it is convenient to identify an optimum number of particles to be employed. It was shown by an example that decreasing the number of particles and checking the relative differences between their solutions can achieve this.

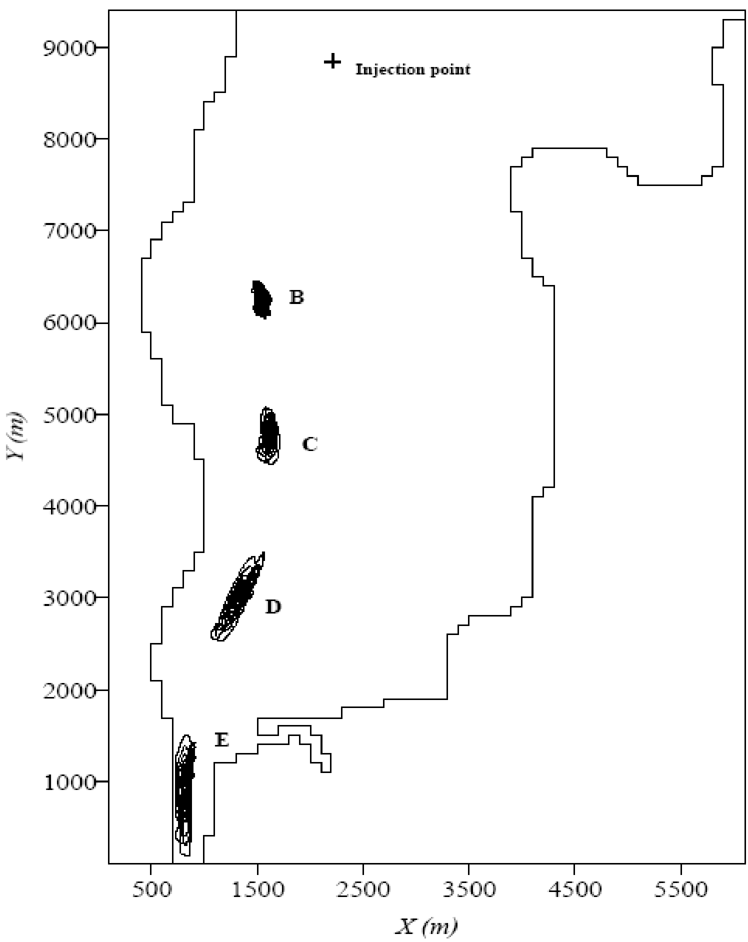


Figure 3: Comparison between the solution of the numerical model (dashed lines), and measured clouds (continuous lines) due to an instantaneous injection.

The method was applied to the case of a point tracer injection held in Setubal lagoon, Santa Fe, Argentina. The flow in this case showed a strong dominance of the longitudinal dispersion related to the transverse one. The model was implemented considering the longitudinal and transverse dispersion coefficients, and from theoretical results. The obtained results, in both cases, represent in an adequate way the concentration of the centre of mass and the general shape of the clouds.

Acknowledgements

Special acknowledgments to Water Resources Management Committee (CT-Hidro / CNPq / FINEP) and to the Universidad Nacional del Litoral, Santa Fe, Argentina, for the funding to perform the surveys.

References

[1] Chorin, A.J., Numerical study of slightly viscous flow, *J. Fluid Mech.*, v. 57 (4), pp. 785-796, 1973.
[2] Vinzón, S. & Paiva A.M., Modelling the Sediment Concentration Profiles at the Amazon Shelf, *Fine Sediment Dynamics in the Marine Env., Proc. In Marine Science*, 5, J. Winterwerp & C. Kranenburg (eds.), Elsevier, 2002.
[3] Periañez R. & Elliot A.A., Particle-tracking method for simulating the dispersion of non-conservative radionuclides in coastal waters, *J. Env. Radiactivity,* v. 58, pp.13-33, 2002.
[4] Hathhorn W., Simplified approach to particle-tracking methods for contaminant transport, *J. Hydr. Eng*., v. 123 (12), pp. 1157-1160, 1997.
[5] Narayanan C. D. Lakehal & Yadigaroglu G., Linear stability analysis of particle-laden mixing layers using Lagrangian particle-tracking, *Powder Technology,* v. 15, pp. 122-130, 2002.
[6] Tessi M. A., Lura M. C., Godoy P. G., Minetti M. L., Rafaghelli de Ramon R. C. & Mir A., Bacterias Indicadoras de Contaminación en las Aguas y Sedimentos de la Cuenca Del Paraná Medio (Zona Santa Fe-Paraná), *Rev. Fac. Ing. Química, Santa Fe*, v. 40, pp. 81-98, 1972 (in Spanish).
[7] World Health Organization, *International Digest of Health Legislation*, v. 27 (4), pp 709-918, 1976.
[8] Trento, A., Vinzón, S., Borche, A., Alvarez, A. & Venturini, V., Modelación Bid. del Transporte de Colif. Fec. en la Laguna Setúbal (Argentina), in *Proc. XVII Cong. Latin. Hidr.*, Guayaquil, Ecuador. v. 3, pp. 353-364, 1996 (in Spanish).
[9] Holley E.R. & Jirka G. H., Mixing in Rivers, *Technical Report E-86-11*, US Army Engineer Waterways Exp. Station, Vicksburg, Mississippi, 1986.
[10] Harleman, D.R.F., *Transport Processes in Environmental Engineering, lecture notes*, Parsons Hydrodynamics Lab., MIT, Cambridge, 1988.

[11] Borche Casalas, A., Modelo Matemático de Correntología do Estuario do Rio Guaiba, *Caderno de Rec. Hídr. No. 12*, IPH/UFRGS, 68 pp, 1985 (in Portuguese).
[12] Elder, J. W., The dispersion of marked fluid in turbulent shear flow, *J. Fluid Mech.*, v. 5 (4), pp. 544-560, 1959.

This page intentionally left blank

Experimental study of an artificial groundwater recharging process

G. Chiaia & G. Ranieri
Water Engineering and Chemistry Department,
1st Engineering Faculty Polytechnic of Bari, Italy

Abstract

Groundwater recharge is gaining ever greater importance in the context of the management of water resources. This practice can improve the quality of aquifers, or at any rate temporarily store resources that would otherwise be lost. The recent stipulation of the “Piano Direttore”, an extract from the Plan for Safeguarding Water Resources (“Piano di Tutela delle Acque”) by the Commissary appointed to deal with the environmental emergency situation in Apulia, has removed the prejudicial clause of illegality introduced by the Law Decree (D. Lgs. 152/99) of allowing rainwater runoff underground. Nevertheless, this type of plant is still rare in the territory, perhaps due to the lack of experimental data supporting the efficacy of the procedure. In the present work the experimental results obtained in an investigation campaign carried out at a recharge site are illustrated and discussed. The recharge site is equipped with suitable devices for continuous monitoring of the rainwater levels, the volume influx to the network, and the variations in the levels of the perched water. Finally, the experimental data are used as the preliminary settings for a mathematical model drawn up to represent the phenomenon.
Keywords: aquifer recharge, experimental investigation, mathematical model.

1 Introduction

The recent national and regional norms regulating the dispersion of rainwater (D.Lgs. 152/99, D. L.vo 258/00 Piano Direttore, an extract of the “Piano di Tutela delle Acque”) pose fairly strict limits on rainwater runoff intercepted by the drainage/sewage system.

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060511

In a region like Apulia, that is practically devoid of superficial bodies of water, the only catchment system responding to these norms is the soil, or the superficial layers of the subsoil.

Besides, the endemic lack of water resources in the Apulian territory means that alternative solutions for the storage of these water volumes, purged of the first rainfall, need to be found to make them suitable for irrigation and, if possible, for production purposes.

Because of the extent of the volumes involved, together with the limited availability of ground, the idea of storing rainwater in storage tanks has had to be excluded.

A possible alternative would be that of using the rainfall as a means of artificially recharging the superficial groundwater, by allowing it to filter through the superficial layers of the subsoil.

This solution has the dual advantage of storing the resource, making it available for future use, and of raising the piezometric level of the aquifer, contributing to combat the saltwater intrusion phenomena due to the widespread, uncontrolled water pumping for irrigation that occurs near the coastline.

Despite the fact that the works necessary to per-form artificial recharge are relatively simple, few experimental data validating such procedures are currently available.

In the present work, the data collected at an experimental recharge plant set up in the countryside of Avetrana, thanks to an agreement between the Polytechnic of Bari and the municipal administration, are analyzed and discussed.

Figure 1: Geographic setting.

2 The experimental apparatus

The built-up centre of Avetrana (figure 1), is located south-west of the city of Taranto, 5.5 km from the Ionian coast, at an altitude of approximately 60 m above sea level. Geologically, it stands on the permeable formation of quaternary calcarenite, as demonstrated by the high number of tufa extraction quarries present on the edges of the town, now al-most all abandoned.

The final storage area of the urban drainage network is a basin inside an abandoned quarry, lying beside the roadside, and adjacent to the terminal tract of the drainage system [3].

The basin (figure 2) is trapeze-shaped and has suitably shaped banks on three sides, while the fourth side corresponds to the vertical wall of the quarry adjacent to the road. The bottom of the basin is that of the quarry and, being made of permeable calcarenite, promotes natural dispersion. The runoff is irregular but generally downhill in a direction away from the road, so that depending on the volumes that flow into the drainage pipes, well de-fined areas with a growing extension are involved.

Figure 2: Area of the recharge plant.

The outflow of the rainwater collector lies in the vertical wall at a level of 51.78 m above sea level, approximately 4 m above the bottom of the basin, so that apart from infiltration losses, a volume of approximately 18.600 m^3 can accumulate before the collector is swallowed up.

At the outflow of the drainage pipe there is a basket-shaped grating, supported by a reinforced concrete structure, that receives the waste water and al-lows it to settle. This bowl has a planimetric area of 10 x 9 m^2 delimited by reinforced concrete sides built up to approximately 4 m above the bottom of the ba-sin. The water passes from the bowl into the basin through a right-angled triangular weir 2.40 m wide at the top, which is situated 50.67 m above sea level. Issue from the bowl is assured by small lateral openings protected by interception devices.

Drainage of the water stored in the bowl occurs not only by infiltration through the bottom but also through a well, which allows water to sink only through the calcarenite fissures in the base, while the water entry is protected by another grating, situated approximately 0.50 m from the bottom of the well.

Inside the quarry a sentry well has been created to measure entry into the perched water, while a pluviograph, placed at a sufficient distance from the flooding area, records the rainwater falling directly into the plant area.

A hydrometrograph measures the water volume entry, placed in the bowl above the weir; another hydrometrograph is placed on the outside, beside one of the walls of the basin, to record the water levels inside the basin.

Finally, there is a ball cock inside the sentry well at the upper end of the aquifer, serving to measure fluctuations in the levels over time.

3 The hydrogeological situation

The final receiver site has all the typical geological characteristics of the Apulian territory.

In general, the mesozoic calcareous crust at the base is covered by a layer of calcarenite rock ("calcareous tufa") of the quaternary age, having an over-all mean thickness of 20 m. This is a massive structure, with rare, irregular subdivisions into great masses.

From the lithological standpoint, the rock is composed of whitish-yellow calcareous granules, tender and porous, with abundant fossils clustered together at irregularly distributed levels inside the body of the rock. The granules vary in size and have variable degrees of hardness and compactness at the different levels.

The absorption trial carried out at the end of perforation of the first sentry well allowed the hydraulic conductivity of the aquifer to be estimated in the order of 10^{-5} m/s.

More details on this topic and on the geological issue are available in [3].

4 Data analysis

The instruments installed at the plant in Avetrana have enabled a high quantity of data to be collected, over several seasons with few interruptions. As in most other experimental plants set up in Mediterranean climatic conditions, the time of year when the most significant events occurred as regards inflow from the urban drainage system was the summer sea-son, when the heaviest rain normally falls in our climes.

Figure 3 shows the water levels trend over time, recorded by the hydrometrograph positioned in the sentry well to control the upper aquifer area, and the volumes passing through the triangular weir de-scribed above. At the representation scale, these volumes are variable and sometimes impetuous.

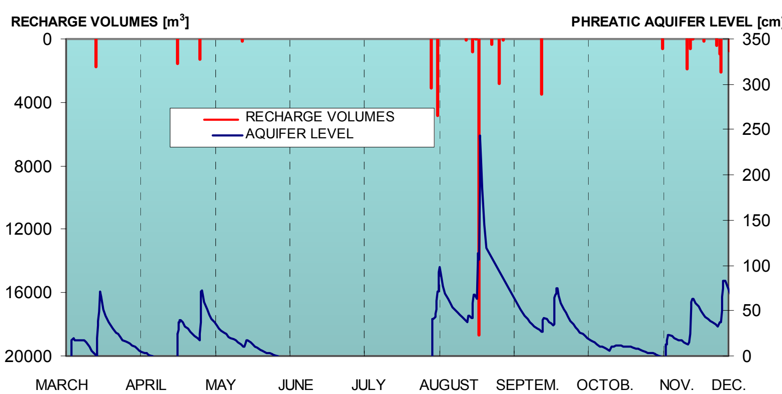


Figure 3: Experimental trend of the overflow volumes and the level of the superficial groundwater.

It can be seen in the figure that at some time intervals, the groundwater level has a value of zero; in fact, this corresponds to the recording of the hydrometrograph ball cock, stuck in the lowest position due to the fall of the water level in the phreatic nappe. According to the controls carried out at the time when the plant was set up, this lowest position can be estimated at approximately 1 m from the bed of the groundwater.

Combined analysis of the trend of the two variables shown in figure 4 shows the close correlation between the volumes overflowing in the bowl and rises in the water levels of the phreatic groundwater.

Figure 4 refers to the event that occurred on the 13th May: the overflow of the bowl was fairly modest, little more than 100 m^3, and did not give rise to a significant filling of the infiltration basin; in fact, the hydrometrograph recording the levels inside the tank indicated a zero value throughout the duration of the event. However, this does not exclude the possibility that some flooding may occur, because this zero value of the instrument corresponds to a geodetic level approximately 15 higher than the minimum level in the tank. It should be noted that the effect of this overflow is felt about 24 hours later: the maximum rise, approximately 6 cm, was reached about 48 hrs after the overflow.

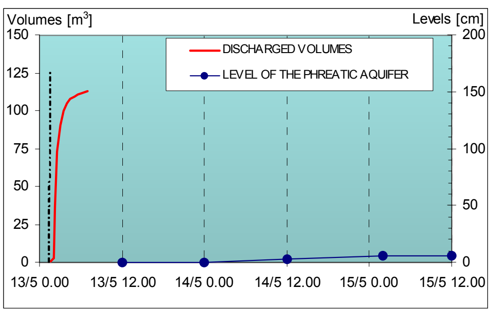


Figure 4: Event on the 12th May.

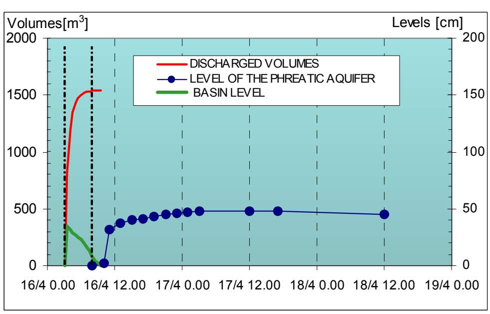


Figure 5: Event on the 16th April.

On the 16th April, the event shown in figure 5 began. The overflow of the bowl, which was greater than in the previous case, caused high filling of the

infiltration basin with a permeable bottom, as demonstrated by the recording of the hydrometrograph, that showed a maximum value of 34 cm (as com-pared to the previous zero), that was certainly less than the quantity flowing into the emergency well.

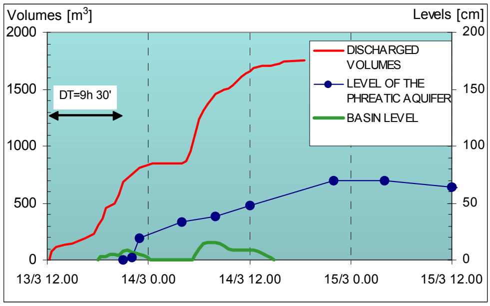


Figure 6: Event on the 13th and 14th March.

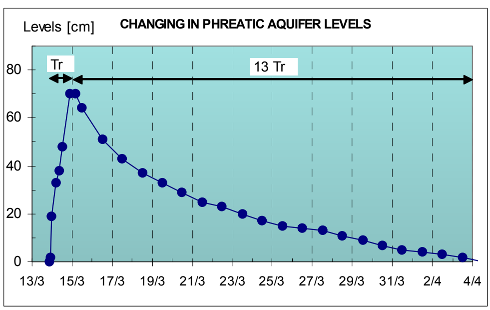


Figure 7: Event on the 13th and 14th March. Complete trend of the variations in the superficial groundwater levels.

The effect of transfer into the groundwater of the overflow volumes started about 7 hours after the beginning of the event, and after a fairly rapid initial rise, the levels reached a maximum increase of 48 cm after about 24 hours.

Comparing the two cases analysed up to now, we can see that the recharge phenomenon depends not only, obviously, on the entry volume but also, bearing in mind the delay before it occurs, on the initial humidity value characterizing the unsaturated state and intensity of the recharge, illustrated to some ex-tent in the diagram showing the steepness of the curve value representing the overflow volume.

Figure 6 shows the event that occurred on the 13th and 14th March, that differed from the previous events because the overflow occurred in two successive time phases. The first phase began on the 13th March at about 12 o'clock and lasted less than 12 hours; it was followed after an interruption of about 6 hours, by a second overflow that reached volumes of 1500 m^3, as compared to the previous volumes of about 700 m^3.

In the infiltration basin, too, the filling occurred in two phases, the first at a maximum height of 9 cm and the second, of 15 cm. It should be noted that the second filling started at virtually the same time as the second overflow of the bowl, as the previous phenomenon had evidently not completely ended during the intermediate phase.

A very similar situation is present as regards the trend of the levels measured in the upper aquifer. In fact, the aquifer began to be affected by the recharge about 10 hours after the beginning of the phenomenon (12 o'clock on the 13th March), while the second delay was only by about three hours. This substantially confirms the fact that the initial conditions (the state of water saturation of the soil) and the hourly intensity of the recharge have a significant effect on the whole phenomenon.

Figure 7 illustrates in detail the trend of the superficial level of the groundwater after the event: the maximum rise, equal to 70 cm, occurred after about 38 hours from the beginning of the overflow of the infiltration basin. The latter time interval is comparable to the duration of the recharge phase. The discharge phase, in the sense of the return to the initial level in the groundwater, lasted about 21 days. The ratio between these two durations is more than 13.

Finally, figure 8 shows the event that caused the greatest rise of the superficial groundwater during the period of observation (178 cm). This is very interesting, as it fully confirms the previous analyses and especially the link between the steepness of the curve representing the overflow volume and the delay before recharge occurs.

It can be seen that the curve indicating the water level inside the infiltration basin shows a horizontal line between the hours of 10 o'clock and 11.30 of the 19th August, at the level of 0.58 m above zero of the measurement tool. This trend can be imputed to the effect of the emergency well, the crown of which is 0.50 higher than the bottom of the infiltration basin.

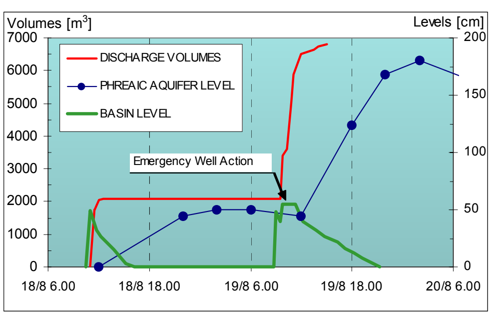


Figure 8: Event on the 18th and 19th August.

5 Modelling the phenomenon

Mathematical modelling of the phenomenon under study must be done by adopting a saturated-unsaturated approach, thus using equations that can describe the filtration process of a water-air blend.

It is possible to simplify this by refraining from making a direct description of the air phase, taking into account its interaction with the liquid phase purely in the retention curve. Empirically built, this shows the correlation between the water charge and the liquid fraction present in the blend, and between these parameters and the real permeability of the aquifer [6].

Taking an ideal vertical two-dimensional domain, the infiltration process can be described by the equation [1]:

$$\left(\Gamma S_S + (1-\Gamma)\frac{\partial \theta_W}{\partial \psi}\right)\frac{\partial \psi}{\partial t} = \frac{\partial}{\partial x}\left(K(\psi)\ \frac{\partial \psi}{\partial x}\right) + \frac{\partial}{\partial z}\left(K(\psi)\ \left(\frac{\partial \psi}{\partial z}+1\right)\right) + Q$$

where

$$\Gamma = \begin{cases} 1 \ \ \textit{Saturated zone} \\ \\ 0 \ \ \textit{Unsaturated zone} \end{cases} \qquad \Psi = \text{hydraulichead} = \begin{cases} \dfrac{\mathrm{p}}{\rho \mathrm{g}} \ \ \textit{Saturated zone} \\ \dfrac{\mathrm{p_c}}{\rho \mathrm{g}} \ \ \textit{Unsaurated zone} \end{cases}$$

$K(\psi)$ = Water conductivity, charge function and hence degree of saturation
θ_W = Water content
S_s = Storage coefficient

The functions:

$$\theta_W = \theta_W(\psi);\ \ K = K(\psi)$$

must be taken as given. The present work adopts the ratios proposed by van Genutchen [7]:

$$\frac{K(\psi)}{K_S} = \frac{\left\{1-(\alpha\psi)^{\beta-1}\left[1+(\alpha\psi)^{\beta}\right]^{\frac{1}{\beta}-1}\right\}^2}{\left[1+(\alpha\psi)^{\beta}\right]^{\frac{\beta-1}{2\beta}}} \tag{1}$$

$$\left(\frac{\theta-\theta_r}{\theta_s-\theta_r}\right) = \left[\frac{1}{1+(\alpha\psi)^{\beta}}\right]^{1-\frac{1}{\beta}} \tag{2}$$

$K(\psi)$= Water conductivity, charge function and hence degree of saturation
θ_W = Water content
S_s = Storage coefficient

where K_S is the water conductivity in saturated conditions, θ_s is the water content when saturated conditions are reached, θ_r is the irreducible water content, α and β are empirical parameters linked to the characteristics of the soil.

The processes described below were carried out using the calculation code SATINSAT [2,5], operating at finite differences. The domain scheme was drawn by a Eulerian grid with variable sized cells. The flow values were calculated on the basis of the elements defined by the cells, while the water charge, degree of

saturation, water conductivity and other scale dimensions were calculated on the basis of the single nodes.

The domain scheme is shown in figure 9. It was drawn with reference to a vertical section of the aquifer, with a homogeneous width, and a depth equal to 9.70 m (the depth of the bed of the phreatic groundwater) passing through the sentry well, the zone behind the grating and so directly involved in the overflow of the latter, and through the hydrometrograph positioned in the bowl.

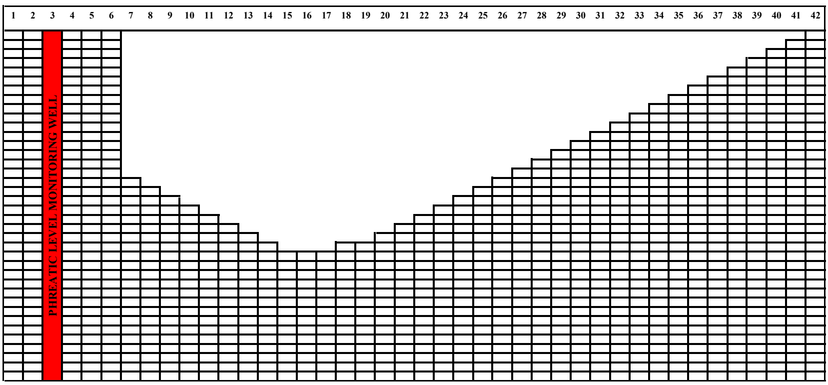


Figure 9: Scheme of the integration domain.

Unlike what is depicted in the figure to make it easier to read, different sized cells were actually used to reduce the number of calculation processes, adopting a much finer grain in the superficial soil zones, where the cells have a much lower height (in the order of cm) than those representing the lower levels. An even rougher grain was used for the saturated soil zone occupied by the phreatic groundwater, represented by the last two rows of cells, each having a height of 1 m. In the same way, the width of the cells was varied according to their site: those involved in the overflow of the bowl (on the right in figure 9) were assigned widths in the order of centimetres, while the other widths are in the order of metres.

6 Settings of the model

The selection of the experimental event to use for the setting of the model was made taking into account the need to reduce as far as possible the uncertainties connected with the initial conditions, especially those having to do with the soil humidity content.

For this reason, among the significant events recorded, the one that occurred on the 30th July was chosen. As can be seen from the graph in figure 3, it was preceded by more than two months without rainfall. This circumstance made it possible, during the flow modelling phase, to consider the initial condition in the zone uninvolved by the groundwater as that of dry soil ($\theta_w = 0$).

The event consisted (figure 10) of a rainfall of 21.2 mm (measured by the pluviograph positioned inside the quarry) lasting about one hour. The volume

overflowing from the tank, that did not give rise to any water head in the infiltration basin, was equal to 310 m^3; the groundwater level rose by about 49 cm.

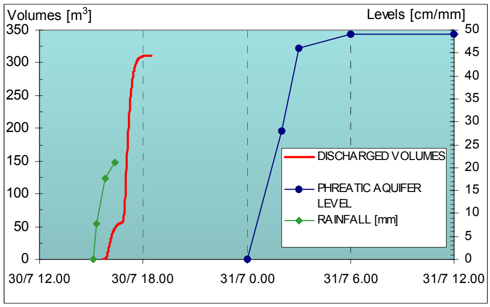


Figure 10: Event on the 18th and 19th August.

The main elements of the physical phenomenon that needs to be reproduced are:

1. time 15.15-16.35 hrs on the 30th July: the recharge basin receives an overall rainfall of 21.20 mm
2. time 16.35 hrs on the 30th July: the overflow from the bowl begins, and ends at time 18.30 hrs (the overall overflow volume is equal to 310 m^3, and involved an area of approximately 1200 m^2)
3. time 17.00 hrs of the 30th July: the ball cock in the infiltration basin records the presence of a water level of 20 cm
4. time 00.00 hrs of the 31st July: the groundwater begins to be affected by the recharge.

When applying and setting the model, the events at points 1, 2 and 3 were used as inputs, and the information at point 4 was used to assess the robustness of the response indicated by the model.

In accordance with the dynamics of the real phenomenon, the simulation was carried out in various steps in order to take account of the variations over time of the border conditions in the integration domain. The condition at the eastern border was an imposed fixed recharge value, at the western border, the free flow of the groundwater, and finally at the bottom, the presence of the impermeable clay limestone layer.

The initial condition was the position of the phreatic line, considered to be 7.70 m below the plane of the countryside.

As mentioned above, the water conductivity of the unsaturated domain was calculated on the basis of the knowledge of its value in saturated conditions, applying the ratios (1) and (2). It was therefore necessary to ascertain the most suitable values for the coefficients α and β.

Indications for the first attempt were taken from the experimental trials carried out by Lakshman and Prasad [4] working with foxhill sandstone, that has a fairly similar behaviour to that of calcarenite.

The next phase of refinement is extremely laborious and, as described below, not entirely satisfactory. The values assigned to α and β were iteratively changed, using a heuristic procedure, within a previously determined physically consistent range, in the attempt to reproduce the behaviour of the system.

Figure 11 shows the final result of simulations, compared with the experimental data. It can be seen that the laborious sensitivity analysis has only yielded a correct reproduction of the shape of the rise of the phreatic groundwater curve, but not the arrival time, which is approximately 5 hours earlier than the experimental time. Similar results were obtained when applying the model with the above settings to other experimental events. The correct reproduction of the shape of the level rise curve continues to be associated with a time lag of the same order as the one described above.

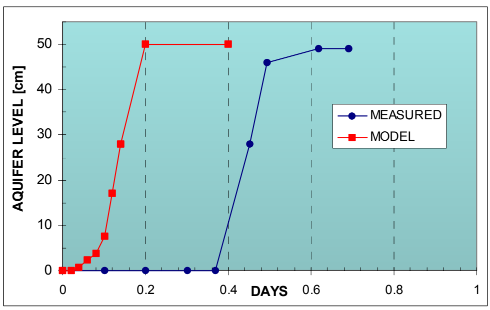


Figure 11: Comparison between the calculated and the measured levels in the groundwater sentry well.

As this difference does not vary according to the event, the proposed calculation model can be considered acceptable, despite the difficulty in transferring the perturbations in the pre-existing movements of the groundwater.

7 Conclusions

This experience of setting up and monitoring an artificial recharge plant of the phreatic groundwater has supplied a large quantity of data. They may be of important interest for the purposes of understanding and carrying out mathematical modelling of processes of infiltration through unsaturated zones.
Comparison of the dynamics of different events has demonstrated that the recharge phenomenon does not only depend on the entry volume but is also conditioned, in view of the delay before it occurs, by the initial value of humidity of the soil characterizing the unsaturated state and by the speed of infiltration.
Analysis of the experimental data collected up to now seems to show that the recharge effects, monitored on the basis of the rise in the free surface of the phreatic groundwater, is exhausted fairly rapidly, in a period in the order of days not weeks.

On the basis of this evidence, it therefore seems that adoption of this practice, at least for relatively small areas of groundwater like the one under study, is not

feasible as an alternative method of storage of water resources, at least on a long term basis. The difficulties encountered in mathematical modelling of the phenomenon, due largely to the lack of published data on the values to be assigned to the van Genutchen parameters, demonstrate the need to carry out further experimental surveys of this kind.

In any case, the laborious sensitivity analysis conducted has made it possible to reproduce correctly in the model, the rise in the curve, experimentally recorded by the sentry well. The considerable delay observed before the water arrives in the groundwater could be attributed not only to incorrect individuation of some hydrogeological parameters, due to the lack of specific experimental analyses carried out on site, but also to the formation of a superficial layer of solid material, carried by the urban drainage system used for the recharge. This material, having different characteristics from the underlying calcarenite mass, could have deposited on the bottom of the basin.

References

[1] Bear, J. (1979) *Hydraulics of Groundwater*, McGraw-Hill, New York.

[2] Castellano L., Masciopinto C., Vurro M., (1993) *Una procedura per lo studio del flusso idrico accoppiando zone sature e non sature: "Satinsat"*, Idrotecnica n. 6

[3] Di Santo A., Maggiore M., Piccinni A.F., Vurro (1992), *Preliminary survey to an experimental research on municipal runoff waters and ground water recharge,* Int. Conf. on "Supplementary irrigation and drought water management"

[4] Lakshman N., Prasad R. (1996), *Field evaluation of unsaturated hydraulic conductivity models and parameters estimation from retention data*, J. of Hydrology, 179, pp. 197-205.

[5] Masciopinto, C., Passarella, G., Vurro, M. and Castellano, L. (1994), *Numerical Simulation for the Evaluation of the Free Surface History in a Porous Media. Comparison between Two Different Approaches*, Adv. in Engin. Soft., vol. 21, pp. 149-157.

[6] Tindall J.A., and Kunkel J.R. (1999), *Unsaturated zone hydrology for Scientists and Engineering*, Prentice-Hall Inc., New Jersey.

[7] Van Genuchten, M.Th. (1980) *A close-form equation for predicting the hydraulic conductivity of unsaturated soils*, Soil Sci. Soc. Am. J., vol. 44, pp. 892-900.

Parametric identification of a karst aquifer

G. Chiaia & G. Ranieri
Water Engineering and Chemistry Department, 1st Engineering Faculty, Polytechnic of Bari, Italy

Abstract

The parametric identification of non-homogeneous aquifers is a highly complex task because it requires the performance of a large number of infield trials, which often yield results that are difficult to interpret or have doubtful reliability. Moreover, the physically continuous nature of the hydrogeological parameters to be identified (typically conductivity) means that the experimental data need to be applied to the territory, a task which is often difficult in view of their great variability.

An alternative approach that has been studied in recent years is parametric identification by means of the solution of the Inverse Problem (I.P.). This method requires a generic hydrogeological knowledge of the aquifer, and the availability of a sufficient number of piezometric survey points on the area where the identification is to be made. The method is potentially able to provide detailed information on the aquifer characteristics, at decidedly competitive costs in terms of money and time as compared with those for the application of the direct method. However, its use is largely limited to the scientific field at present, and few technical applications have been described for the management of territorial scale aquifers.

This is because by its very nature, the inverse problem has the characteristic of an intrinsically ill-posed problem, in the sense that it frequently has no unique, stable solution.

Many researchers have investigated the causes of this poor formulation of the problem, recognizing the fundamental role played by the type of boundary conditions, the accuracy of reconstruction of the piezometric surface and the type and structure of the parametric method adopted.

In the present work, after making a preliminary overview of the theoretical reasons why the I.P. is ill-posed, the above aspects are analysed in the context of a real aquifer with an extension of approximately 100 km^2. After describing the method used to reconstruct the piezometric surface and individuating the boundary conditions, we illustrate the effects on the solution of the I.P. of different types of parametrization of conductivity.

Keywords: inverse solution, territorial scale aquifer, piezometric surface, mathematical model.

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060521

1 Introduction

The solution of the Inverse Problem (I.P.) aims to derive, through direct observation of the values of an easily measured dependent variable (potentiometry), the spatial distribution of the parameter it depends on, i.e. conductivity. It is taken for granted, therefore, that the operator can rely on a sufficient number of piezometric survey points inside the domain under investigation.

The parametric identification procedure is normally performed using indirect methods, applying history matching type criteria to estimate the parameter T(x,y). A problem that arises already at this stage of the procedure is how to model the spatial variability of the parameter being identified. In fact, conductivity is a continuous function of the position and, for the purposes of mathematical modelling of the filtration process, a finite, discrete approximation has to be made.

The process of assigning a discrete value, which in practice allows the unknown quantities of the I.P. to be reduced to a finite number, is known as parameterisation. There are two classic ways of doing this: by zoning or by using geostatistical techniques.

With the former (e.g. [4,7,8]), starting from knowledge of its geology and hydrogeology, the area under study is subdivided into a given number Z of zones, in each of which the parameter T is hypothesized to be constant, although it may vary from zone to zone.

Instead, the geostatistical technique (e.g. [12,20]) considers the parameter to be identified as a random stationary field which undergoes continuous variations in space. A series of nodes is individuated in the area of interest, each of which is associated with a basic local function. A third possible approach to parameterisation is the "hybrid" method proposed by McLaughlin [14] in which, by using a functional analysis method, the unknown value of conductivity is treated as a scale function of the position rather than as a vector.

The vector of the piezometric levels calculated is a function $\mathcal{F}$ of the distribution T' of transmissivity.

After individuating the parameterisation criterion best suited to the case under study, this parameterisation dimension is set at Z (where Z is equal to the number of zones with the zoning technique, or to the number of nodes with the geostatistical approach), the distribution of trial T' is constituted by a vector with Z components. If $\mathbf{h_r}$ is the vector of the identified piezometry, an "indirect" expression of the discrepancy between the real distribution of transmissivity T and the estimated value T', is:

$$Ob(T')=[\mathbf{h_r}- F(\mathbf{T'})]^T * \mathbf{W_p} * [\mathbf{h_r}- F(\mathbf{T'})] \qquad (1)$$

which represents the mean square difference between the calculated levels and the measured levels (the latter being weighted, if necessary, by the matrix $\mathbf{W_p}$). The solution to the Inverse Problem is obtained by minimizing the value (1), that takes on the significance of the objective function (O.F.).

The greatest difficulty in applying this procedure is that of identifying a suitable algorithm that can minimize the O.F. by introducing suitable variations in the components of the vector T'.

This problem can be dealt with by various approaches that differ as regards the optical focus of the mathematical tools required. An extensive list of references on this topic is available in [2,3,18].

The I.P. is often ill posed; that is it hasn't a unique solution or/and the solution is not stable [16].

As shown by several authors [13–15], there is also an intrinsic instability linked to the structure of the equations governing the "direct" problem.

A big role in make the problem well posed is played by:

- o The type of boundary conditions imposed
- o The level (qualitative and quantitative) of knowledge of the piezometric surface
- o The type of parametrization adopted.

2 Experimental processing

The case study was a karst aquifer, on which, the following aspects were analysed at the territorial scale:

- Surveying and processing the piezometric data;
- Individuating the boundary conditions;
- Parameterisation of the unknown conductivity value;

2.1 Description and diagram of the experimental site

The case study area, having an extension of 110 km^2, is the area shown inside the grid in figure 1.

Drinking water needs are partly satisfied by mains water lines supplied by sources outside the Apulian region and partly by pumping of the groundwater at a virtually constant rate of 315 l/s throughout the year, from the artesian wells shown in figure 1 marked as AQP.

The domain is shown in the figure as a diagram with the elements represented on a grid, deliberately built with the vertexes coinciding with points with a known piezometric value, and lying in an orthogonal direction with respect to the coastline.

In view of the large size of the area under study, modelling of the water flow was done according to the porous equivalent approach.
The piezometric survey points were obtained via ten measurement cycles on 44 sentinel artesian wells over a period of two solar years.

Subsequent processing [6] enabled us to individuate 34 points with a high degree of reliability, which were then used to calculate the piezometric surface of the aquifer in the different periods of the year.

2.2 Analysis and processing of the piezometric survey points

As already mentioned above, apart from the drinking water wells, whose position and reserves are known, there are a high number of points for irrigation which pump unknown quantities in a discontinuous manner.

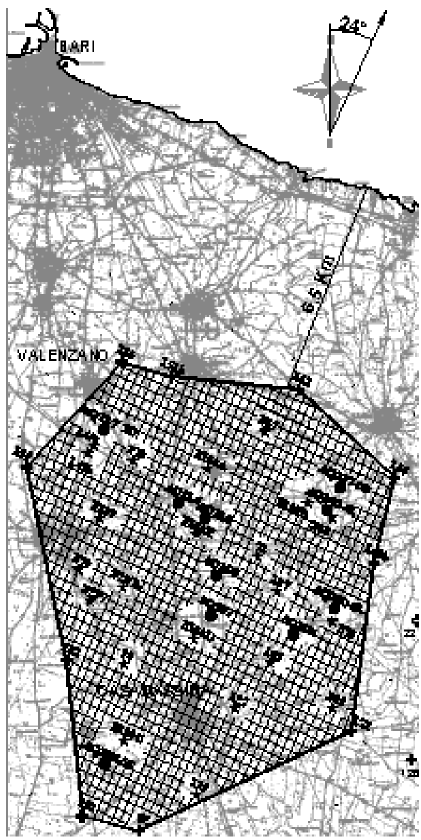


Figure 1.

This constitutes a potential cause of "ill posing" of the inverse problem because, in general, the piezometric points identified are not a unique function of the hydrogeological characteristics of the aquifer, which also depend on other factors which are difficult to measure.

To overcome this shortcoming, the parametric identification procedure was carried out on the basis of the piezometric data obtained in the month of March (figure 2), when the effects on the shape of the piezometric surface of withdrawal for irrigation can be presumed to be over.

As can be seen, the availability of only discrete, discontinuous piezometric survey points often causes ill posing of the inverse problem, due to the incorrect definition of the denominator in (4). It is therefore essential to carry out a series of processes, described in detail in a previous work [6], that yield a unique answer to the calculation of the piezometric surface in different periods of the year.

The above processing yielded the piezometric values of all the nodes in the grid in figure 2.

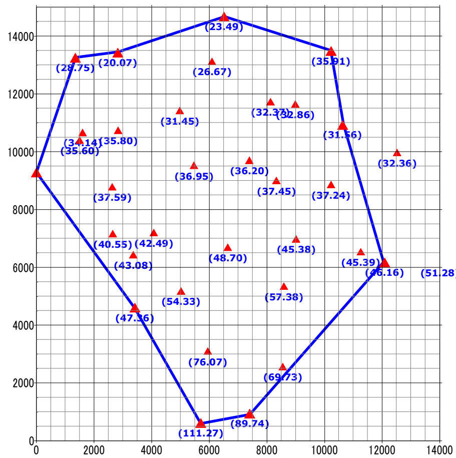


Figure 2.

2.3 Parametrization criteria

The most refined approach to the parametrization of an aquifer is undoubtedly the geostatistical approach, owing to its intrinsic ability to describe complex realities by means of a limited number of parameters (the conductivity values of the nodes).

To use this approach, a well-defined structural model of the parameter to be identified must firstly be identified. In principle, this can be done by following an analogous procedure to the one illustrated above for reconstructing the piezometric surface.

However, it must be noted that the reconstruction of the piezometric structure is relatively simple owing to the physical nature of the piezometric surface, which does not suffer distortion and has minimal variance, so that a fairly complete picture of the spatial variability of the parameter can be gained with a fairly small number of measurements.

The conductivity of a karst aquifer does not have the same characteristics, since it features a notable variability and has a complex, articulated surface. For this reason a much higher number of survey points is required.

Apart from the poor reliability of infield trials, owing to their frequent small scale anisotropy [5], they would be far too costly for use on a real scale aquifer.

Summary indications on the spatial trend of conductivity of the aquifer can be obtained from the results of pumping trials effected at the end of perforations for digging private artesian wells, carried out to determine the admissible withdrawal quantities.

For the area under investigation, the specific charge values (proportional to the local conductivity values) are available of the 115 wells shown in figure 2.

An omnidirectional semivariogram of these values is shown in figure 3.

Despite the limits due to the fact that the data available are not absolutely reliable (having been obtained by the well owners), an examination of the graph,

that shows a nugget type trend which is confirmed even in cases using different lags, demonstrates the poor spatial correlation of the parameter analysed, imputable to the high degree of anisotropy of the aquifer.

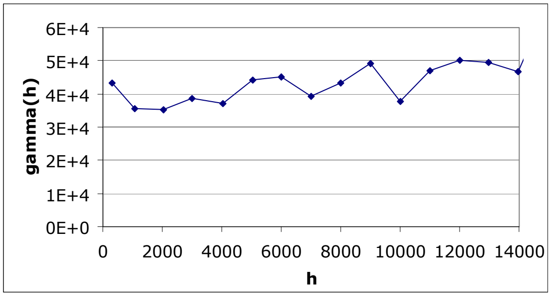


Figure 3: Semivariogram of the specific charges.

This circumstance meant that to assess the aquifer we had to give up the idea of using geostatic type parameterization and to opt for the zoning technique.
It is essential to individuate the number and shape of the zones, or "structure" of the zoning, because this has a strong influence on the correct posing of the inverse problem and the physical plausibility of the solution [14].

In the case of aquifers extending over a wide area, a "direct" approach to individuating the structure of the zoning requires the execution of too high a number of infield trials to be technically and economically feasible. Nevertheless, in absence of the effects of withdrawal for irrigation and local replenishment, and at a suitable distance from drinking water pumping points, the spatial variations in the piezometric surface are largely linked to variations in conductivity.

An "indirect" approach to individuating the zoning type to be adopted can therefore be based on an analysis of the piezometric gradient, following the procedure:

- Calculate the value of the experimental gradient in all the nodes of the integration grid in figure 2
- Individuate the gradient range and subdivisions, according to the criteria illustrated below, at a suitable number of intervals
- Establish the zones by grouping together the elements of the calculation grid with gradient values falling within the same interval.

Figure 4 depicts the qualitative trend of the piezometric gradient in the integration domain.

The zone including the boundary wells 162, 166 and 168b has been excluded because their gradient values are notably lower than the mean, and indeed tend towards zero.

This zone, that probably has special hydrogeological characteristics that for economic reasons are beyond the scope of investigation of this study, has thus been excluded from all the following parametric identification processes.

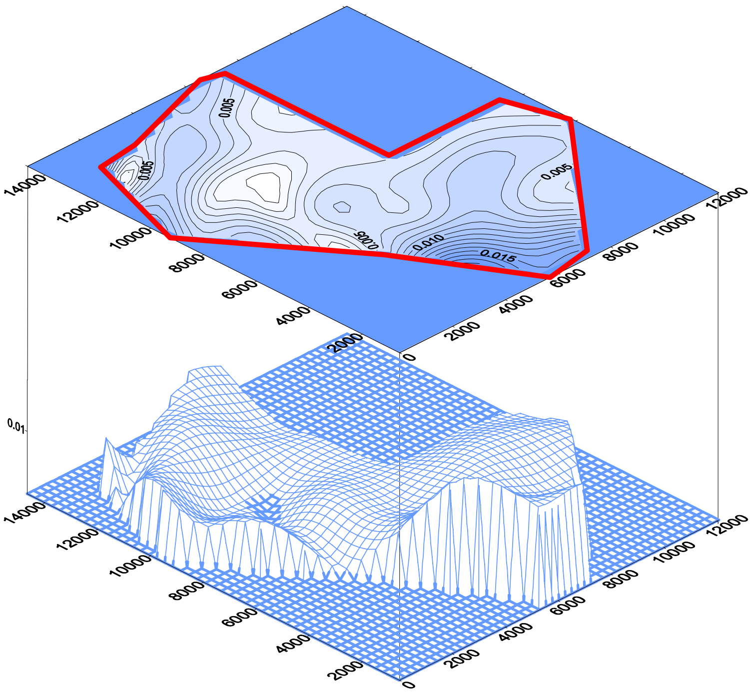


Figure 4: Piezometric gradient of the area under study.

The above procedure yields different zoning parts according to which steps are taken to subdivide the experimental gradient and to the relative criteria adopted to decide this.

With reference to the latter aspect, in the present work three different criteria were used to form the zones:

Zoning type n. 1: ordering the elements by decreasing values of the piezometric gradient and subsequent regrouping to make up zones all having the same number of elements.

Zoning type n. 2: subdividing the range of the logarithm of the piezometric gradient into an equal number of parts and grouping together all elements with gradient values within the same interval (variable elements zoning)

Zoning type n. 3: like 2 but referring to the gradient values not their logarithms.

The parameterization size, i.e. the resulting number of zones, depends on the steps adopted to subdivide the gradient range. In accordance with the indications reported by Yhe et al. [20], increasing order zoning (from 1 to 9), was created, individuating the optimal configuration by means of the sensitivity analysis reported below.

Figure 5 shows, as an example, a comparison of the different zoning types for the same parameterisation size (5 zones).

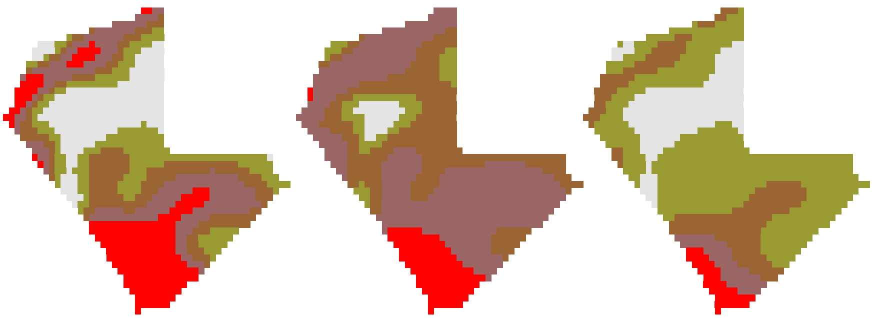


Figure 5: Comparison of zoning types according to different criteria for subdividing the piezometric gradient range. The parameterization size in the case under study is equal to 5.

2.4 Algorithm for the solution of the inverse problem

In the present work, particular attention was paid to the methodology adopted to solve the inverse problem, with the aim of individuating simple, repeatable procedures that could enable technicians responsible for managing aquifers to apply these parametric identification measures to real aquifers.

Processing was done with SGH_D2 software [13], available on the WWW. It is very simple to use and so for technical applications it is preferable to other codes, such as sutra-*1* or *modflow,* which are certainly more refined but are better employed for scientific purposes.

The algorithm used to minimize the objective function is the downhill-simplex, which is particularly efficacious when there are only a small number of parameters to be optimised, as in the present case.

The inverse problem was solved by using the finite elements scheme shown in figure 1, which has a rectangular grid where Dx = 241m and Dy = 305m, and adopting the three different zoning techniques described in the previous section, operating with parameterization sizes ranging from 1 (homogeneous aquifer) to 9.

The software used limits the piezometric control points that can be used to build the lens function to 30. These were selected from the 1295 nodes on the grid depicted in figure 2, according to the criterion of equidistribution.

2.5 Definition of the boundary conditions

The piezometric levels observed in the wells on the boundary of the integration domain (see figure 2) were used to define the Dirichlet boundary conditions.

It was a more complex task to individuate the Neumann conditions, owing to the need to quantify the water flow in and out of the integration domain.

To calculate this it was necessary to discover the hydrological balance in the territorial layer immediately above the study area, shown by the crossing dashed lines in figure 6. This layer is laterally bounded by the subterranean watershed of

the aquifer and in a longitudinal direction, lies parallel to the mean direction of the groundwater flow. Being 20 km wide, it is representative of the mean hydrogeological conditions in the territory.

The aquifer recharging process can be roughly deduced from the following expression:

$$N = P - (E + R + A) \tag{2}$$

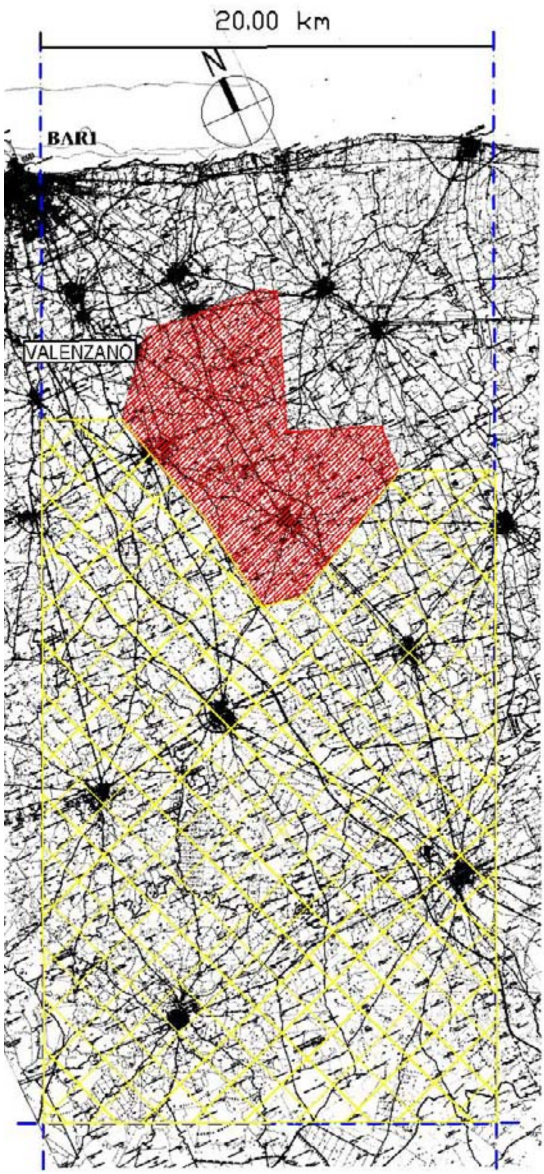


Figure 6: Recharging area of the domain under study.

where: N = recharging of the groundwater; P = rainwater height; E = true evapo-transpiration; R = streaming; A = variation of the water capacity over the terrain. Troisi and Vurro [17] obtained the value of (2) (in mm/year) for each hydrographic-thermometric station of the Murgia basin, calculated by adopting two distinct procedures for assessing the evapo-transpiration terms.

The area under study belongs to the territory served by the stations of Altamura, Masseria Mercadante, Cassano Murge, Casamassima, Turi and Gioia del Colle. The annual recharge load was obtained by reconstructing the Thiessen Polygons for each station, individuating the portions falling within the zone of interest, and multiplying the surface by the relative recharge value. From this sum was then subtracted the total quantity of drinking water withdrawn by the Acquedotto Pugliese from 6 wells situated in the countryside of Acquaviva and Gioia del Colle, which extract a total volume of 75.5 l/s, equal to 2.35 Mm^3/year.

This yielded the net annual average recharge volume in the layer above the integration domain. Divided by the width of 20 km, this made it possible to calculate a specific daily charge of qI [m^3/day*m], entering through the uphill boundary of the domain.

The inflow charge in each of the 50 cells making up the uphill boundary of the integration domain was calculated by multiplying this value by their transverse width, equal to 241 m. The overall mean inflow charge in the domain QI is, of course, equal to q_{IC} x 50.

The annual mean outflow volume from the integration domain (Q_O) was calculated by adding the inflow volume (Q_I*360) to the innate recharge value of the domain (having a surface of 89.17 km^2 all falling within the Thiessen Polygon of Casamassima) and subtracting the drinking water extracted, equal to 180.8l/s = 5.62 Mm^3/year. The discharge q_{oc} from each of the 50 cells on the downhill boundary below the domain is equal to Q_O/(360*50).

A knowledge of the charge crossing the area under study and of the underlying mean gradient allowed us to estimate the equivalent conductivity value of the aquifer as equal to 1670 m^2/day: this was used as the initial value in the parametric identification process.

2.6 Results of the parametric identification

Figure 7 shows the results of the parametric identification made with the three different zoning criteria, varying the parameterization size.

Firstly, it can be seen that the conductivity value obtained with the homogeneous aquifer hypothesis (parameterization size = 1), equal to 1580 m^2/day, is very close to the mean value obtained above.

Referring to zoning 1, it is clear that the different parameterization sizes (at 1,2,…,9 zones) yield conductivity values with a mean increase proportional to the number of zones: this confirms the hypothesis underlying the zoning criterion adopted, which postulates a direct link between conductivity and the piezometric gradient value, which increases with the number of zones.

Again referring to zoning 1, it can be seen that while the parameterization size varies, the conductivity values remain within a physically plausible range (500–6200 m^2/g), compatible with the hydrogeological features of the aquifer.

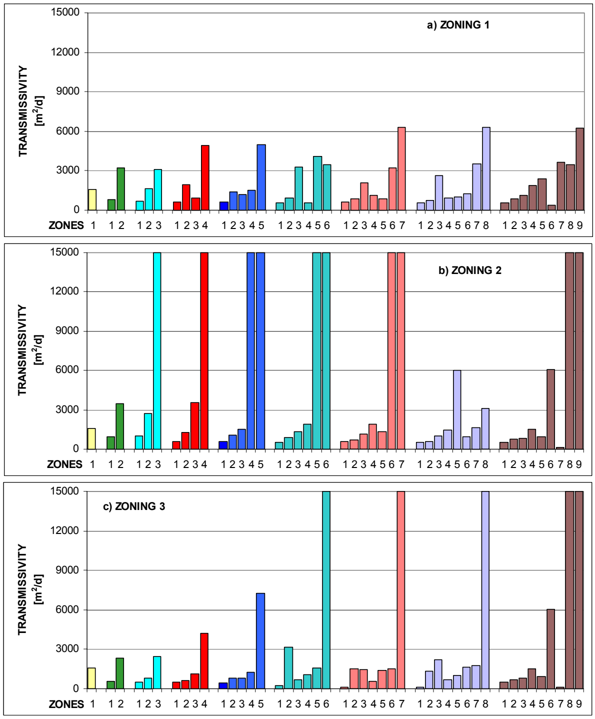


Figure 7: Results of parametric identification with the three zoning criteria at changing parameterization sizes.

Vice-versa, the structure of zoning types 2 and 3 is poorly suited, in the case under study, to modelling the spatial variability of the conductivity parameter: this is demonstrated by the instability of the solution of the inverse problem, which gives rise, in agreement with reports by other authors [18], to values at the upper limit of the variability range of the parameter, set at 15000 m^2/g.

This instability is particularly evident in the case of adoption of zoning type 2, where excessive values are present even at low parameter sizes, while with zoning type 3 excessive values appear when the parameter size is 6 or more.

The zoning geometry that gives rise earliest to instability of the solution to the inverse problem is shown in figure 8. It is evident that in zoning 3, where the instability occurs, there is a small number of elements (33 elements, as compared with 626 elements in zoning 1 and 406 elements in zoning 3) and that in this zoning there is no defined boundary condition of the flows.

Detailed analysis of the other zoning types generating instability of the I.P. solution confirmed that the phenomenon arises in the case of zones with few elements.

The transmissivity value obtained for a given cell in the domain according to the different zoning type and parameterization size is shown in figure 9. It can

be seen that before the above-described instability sets in, the transmissivity values obtained with the different zoning types are fairly similar, demonstrating a certain "robustness" of the algorithm used for the calculation.

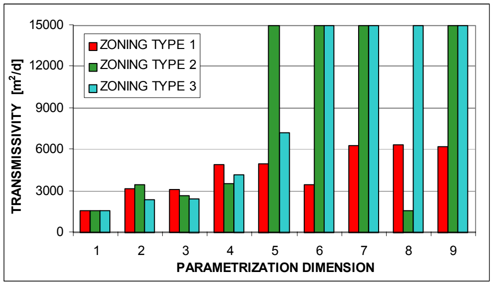


Figure 8: Transmissivity values for the some cell varying parameter dimension.

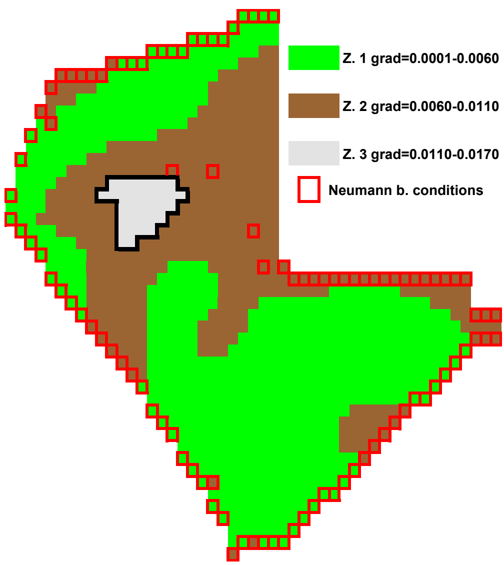


Figure 9: Zoning geometry that gives rise earliest to instability in the I.P. solution.

2.7 Choice of the solution

The results of the processing described in the above section showed that zoning 1 is best able to describe the aquifer, as it can support high parameterization sizes without developing numerical instability.

The solution to the I.P. must therefore be found among the 9 possible configurations depicted in figure 7a, characterized by different parameterization sizes.

In view of the intrinsic ill posing of the I.P. frequently referred in this work, the choice cannot be based on deterministic criteria, but must rely on the concept of optimisation, which yields not "the solution", but "the best among the available solutions".

A classic methodology was proposed by Yhe and Yoon [19], based on comparison of the errors associated with modelling of the aquifer and with the uncertainty of the parameterisation.

The former, as is well known, is the objective function (2); the latter, which affects the stability of the solution, is the norm of the matrix of covariance of conductivity. This can be calculated by the relation:

$$\mathrm{Cov}\left(\mathbf{T}^{*}\right) = \frac{J\left(\mathrm{T}^{*}\right)}{M-L} \cdot \left[J_D^T \cdot J_D\right]^{-1} \tag{3}$$

where T* is obtained by perturbing solution T of the inverse problem with the suitable sized noise [1], $J(\mathrm{T}^*)$ is the value of the objective function, J_D indicates the Jacobean matrix, $J_D^{\,T}$ its transposition, and M and L are the piezometric survey points (30 in the case under study) and the parameterisation size (ranging from 1 to 9), respectively.

The size of the matrix (3) (L x L), and hence the numerical value of its norm, depend on the parameterization size adopted, which hampers comparison between the different configurations.

To circumvent this shortcoming the authors [19] suggest that parameterisation should be adopted, transforming the conductivity vector T* (that in (3) has size L), into a new vector of size N, and N number of elements of the calculation grid.

For this purpose a transformation matrix **G,** is introduced, of size N x L, whose lines are all null elements except one, which is equal to the unit. This matrix associates each of the N elements of the domain with one of the conductivity L deriving from the parameterisation considered. In this way the new vector T_e is obtained, for each element in the domain:

$$\mathbf{T}_e = \mathbf{G}*\mathbf{T}^*$$

So (3) is changed to:

$$\mathrm{Cov}\left(\mathbf{T}_e\right) = \frac{J\left(\mathrm{T}^{*}\right)}{M-L} \cdot G \cdot \left[J_D^T \cdot J_D\right]^{-1} \cdot G^T \tag{3'}$$

The objective function shows a marked decreasing trend for parameterisation sizes between 1 and 4, while subsequent increases up to value 8 do not substantially improve the precision of the modelling but do notably increase the uncertainty of the parameterisation.

Combined analysis of the two curves in figure 10 shows that 5 is the ideal parameterisation size. In fact, at this configuration both the objective function and the norm associated with the matrix of covariance of conductivity have optimal values.

The optimal solution of the inverse problem for the domain under study is shown in figure 11. The mean value of conductivity, that can be calculated as the simple arithmetical mean of the five conductivities thanks to the equal number of elements making up the different zones, is equal to approximately 1900 m^2/day, very close to the previously obtained values. It can be seen that the distribution

of conductivity is coherent with the indications deriving from the shape of the piezometric surface. The presence of a low conductivity zone uphill justifies the high gradient values found in this area (figure 5), while the high value of conductivity T5 explains the presence, in the central zone of the domain, of a layer with very minor gradient values.

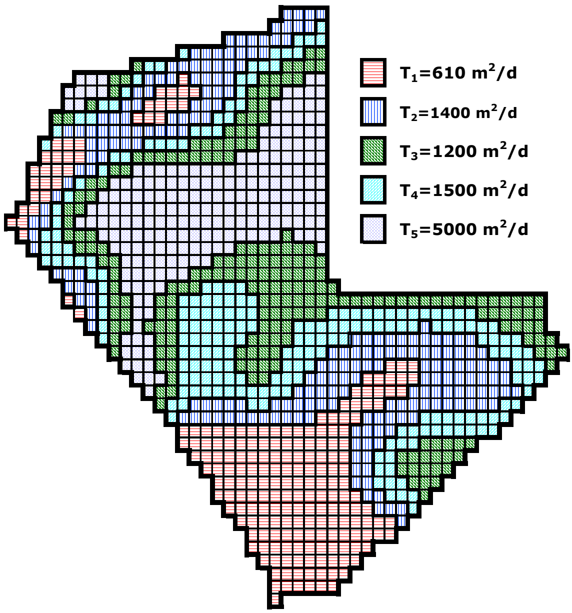


Figure 10: Optimal solution of the Inverse Problem.

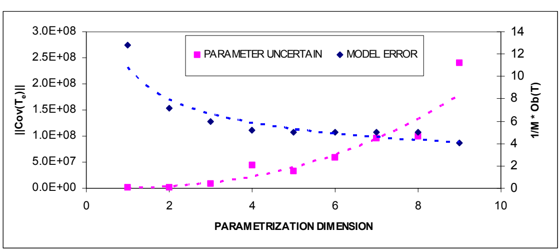


Figure 11: Error of the model and parameterization uncertainty.

3 Conclusions

To apply parametric identification methods based on the solution of the inverse problem to a real aquifer, it is firstly necessary to gain a knowledge of a series of essential elements that are needed to limit the effects of the intrinsic “ill posing” of the I.P.

In particular, theoretical analyses have demonstrated that correct posing of the problem depends very strongly on the level of knowledge of the aquifer piezometry, of the nature of the boundary conditions and of the parameterisation type adopted to assign discrete conductivity values.

In-depth study of these aspects has been made in the present work, focusing on a real aquifer with a planimetric extension in the order of 100 km^2.

The methods used to interpolate the piezometric survey points and to make a preventive hydrogeological assessment of the territory have been analysed in detail. The results of this analysis indicate that zoning is the parameterisation method best suited to the aquifer.

The effects of this method on the solution of the I.P. have been investigated by comparing different zoning techniques, all based on an "a priori" knowledge of the shape of the piezometric surface. Zoning techniques involving the aggregation of a variable number of elements gave rise to an unstable solution of the I.P. and were therefore rejected.

To pose the I.P. correctly, boundary flow conditions need to be imposed, and hence the hydrogeological balance of the area under study had to be identified.

The optimal solution of the I.P. was obtained by means of the sensitivity analysis proposed by Yhe et al. [20].

References

[1] Bard Y., 1974. Nonlinear parameter estimation. Ed. John Wiley – New York.

[2] Butera I., Grella E., Maione U., Tanda M.G., 1988. Applicazione comparativa di tre metodi di risoluzione del problema inverso in un acquifero eterogeneo bidimensionale. *Atti XXVI Convegno di Idraulica e Costruzioni Idrauliche . Catania 9-12 settembre 1998.*

[3] Binley A., 1993. Groundwater model parameter identification. *Basic Lectures of European Training Course on Parameter estimation in groundwater flow and pollution modelling*. Ed. Bios - Cosenza (ITALY).

[4] Carrera J., Neuman S.P., 1986. Estimation of aquifer parameters under transient and steady state conditions. (1-2-3). Water *Resource Research, Vol 22, n. 2 pp. 199-242.*

[5] Chiaia G., Di Santo A., Ranieri M. 1999. Interpretazione di misure sperimentali della falda pugliese. Giornata *Mondiale dell'Acqua – Roma*.

[6] Chiaia G., Ranieri M. 2004. Elaborazione di rilievi piezometrici finalizzate alla identificazione parametrica di un acquifero a scala territoriale. *L'acqua n. 4 pp. 29-40.*

[7] Cooley R.L., 1982. Incorporation of prior information on parameters into non linear regression groundwater flow models. 1, Theory. *Water Resource Research, Vol 18, n. 4 pp. 965-976.*

[8] Cooley R.L., 1983. Incorporation of prior information on parameters into non linear regression groundwater flow models. 2, Applications. *Water Resource Research, Vol 19, n. 3 pp. 662-676.*

[9] Desbarats A.J., Logan C.E., Hinton M.J., Sharpe D.R., 2002 On the kriging of water table elevations using collateral information from digital elevation model. *Journal of Hydrology Vol. 255 pagg. 25-38.*
[10] Gambolati G., Volpi G., 1979 Groundwater Contour Mapping in Venice by Stochastic interpolators 1. Theory. *Water Resource Research, Vol 15, n. 2 pp. 281-290.*
[11] Goovarerts P., 1997 Geostatistics for Natural Resource Evaluation. *Oxford University Press-New York.*
[12] Hoeksema R.J., Kitandis P.K., 1984. An application of the geostatistical approach to the inverse problem in two-dimensional groundwater modeling. *Water Resource Research, Vol 20, n. 7 pp. 1003-1020.*
[13] Karovent T. 2001. Subsurface and groundwater hydrology: basic theory and application of computational methods. *Pubblicazione elettronica*.
[14] McLaughlin D., Townley L.R., 1996. A reassessment of the groundwater inverse problem. *Water Resource Research, Vol 32, n. 5 pp. 1131-1161.*
[15] Sun N.Z., 1994. Inverse problems in groundwater modelling. *Kluwer Academic Publishers - London.*
[16] Tikhonov A.N., Arsenin V.Y., 1977. Solutions of ill-posed problems. *Halsted Press/Wiley – New York.*
[17] Troisi S., Vurro M., 1987. Bilancio idrologico per la valutazione della ricarica naturale: una applicazione ai bacini delle Murge e del Salento – *Idrotecnica, n.2, pp.47-64*.
[18] Yeh W. W-H., 1986. Review of parameter identification procedures in groundwater hydrology: the Inverse Problem. *Water Resource Research, Vol 22, n. 2 pp. 95-108.*
[19] Yeh W. W-H., Yoon Y.S., 1981. Aquifer parameter identification with optimum dimension in parameterisation. *Water Resource Research, Vol 17, n. 3 pp. 664-672.*
[20] Yeh W. W-H., Yoon Y.S., Lee K.S., 1983. Aquifer parameter identification with Kriging and optimum parameterisation *Water Resources Research Vol. 19 n.1 pp. 225-233.*
[21] Yakowitz, Noren, 1976. On the identification of inhomogeneous parameters in dynamic linear partial differential equations. *Journal Math. An. Appli. 53.*

Separation of salt and sweet waters in an area of former mines

P. P. Prochazka[1] & V. Dolezel[2]
[1]Czech Technical University in Prague, Czech Republic
[2]Technical University Pardubice, Czech Republic

Abstract

In old deep abandoned mines, the problem of separation of the layers of a different nature needs to be solved. The problem discussed in this paper relates to mines situated near Prague, capital of the Czech Republic and is concerned with a separation of two kinds of underground water, one being salt (saline) and the other sweet (fresh) water. The layer of the sweet water is positioned at approximately 200 m below the surface, then there was originally a layer of watertight clay (hydrogeological insulator) and from some 300 m downwards, salt water occurs. A problem arose when the shafts, at a depth of 1000 m, have been abandoned due to canceling the mine workings. During mining, pumps were used and no permeability was enabled. Since the mining has stopped, there is no reason to continue pumping and the danger of mixing both kinds of water has occurred. It was decided, in order not to damage the sweet water which is also used for wells and subsoil irrigation, that a stopper should be added into the clay layers to suppress any possibility of mixing both kinds of water. The problem then was oriented to answer several questions. One of the most important that was put forward was on the relationship between permeability coefficients and the thickness of the stopper. Another one was concerned with the optimization of expenses involving the required permeability, expressed in terms of economic indicators, the thickness depending on expenses, etc. As a numerical model has been prepared and it was necessary to feed it by reasonable input data, a scale model from physically equivalent materials was created and tested. The mutual influence of both models, mathematical and experimental, offered a very powerful tool for well-prepared reports on how the stopper should be constructed.
Keywords: scale modeling, boundary element method, leakage problem.

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060531

1 Introduction

The government of the Czech Republic decided that some deep mines would be closed because of their low efficiency. Their location, near the capital Prague underlined the decision, as the air and ground pollution problem should be improved in the capital. The only problem, which occurred before the mines were closed, consisted of the existence of two large layers of different kinds of water. The lower one contained salt and the upper sweet water. The salt water could not be allowed to pollute the upper layer, as plenty of wells in villages in the neighborhood of the mine could be contaminated. Two horizontal clay layers (insulant and semiinsulant) divide both large layers containing salt water and fresh water. Before establishing the deep mines, the vertical shafts were constructed with concrete linings. The salt water was draining away all the time during the existence of the mines. The drainage was designed very carefully to avoid possible infiltration of the salt water to the upper layer. Before the mines abandonment it was necessary to ensure the prevention of contact of both kinds of waters in the shafts. Several alternatives were put forward from which the most efficient appeared to be one which proposed filling up the shafts with spoil soil and the creation of a sealing stopper from soft clay at the horizons between the salt water and fresh water. The material for spoil soil will be delivered from deposits; their distance from the shafts is not too far. Such a material will be disparate; starting from grains of sand to big peaces of slate, slaty clay, and sandstone, reaching to nearly 1 m^3. In order to recover the hydrogeological insulation between both kinds of water, in the depth of 220–300 m, slaty clay will be delivered and an insulating stopper will help keep the split of both kinds of water. A cross-section of the situation is depicted in Fig. 1.

The aim of this paper is to describe the design of the stoppers applied to splitting the contact horizon using coupled equations of effects of streaming water and deforming skeleton. The results from the laboratory are implemented to the numerical model in which optimal material properties of the stopper are sought. The radius of the shafts is 4.45 m.

2 Description of construction of the sealing stoppers

After abandoning the mine and the horizontal seams and access horizontal shafts, only vertical shafts, which ended at the bottom of the horizontal seams, are of our interest. At this stage all technology equipment was removed and partly destroyed. The saline water started to increase its water level and after some time (c. 3 months, which was an equivalent of 300 m increase of the level of the salt water), a spoil material was tipped from the surface into the shaft. It was necessary to equilibrate the levels of the stewed spoil material and that of saline water in such a way that the dry strewed material does not outtop the water level. This stage lasted quite a long time, as the sedimentation of the spoil soil had to be ensured. The sedimentation was enforced by adding coagulators. In the same time the possibility of two or more stoppers were assessed. In the frame of this solution the permeability of stoppers was determined in dependence of their

material properties. The coefficient of filtration (m/s) was gained from laboratory tests in a range of 2.5–4.0 ($*10^{-11}$). Classification of spoil material aimed to filling the shaft and classification of clay sealing material from the standpoint of geomechanical properties comes out from a description of specimens and from results from the laboratory and field (in situ) examinations. With regard to unexpected high groundwater affluent and the breakdown of the electrical power system control room and water-pumping station, water pumping became impossible; this is why backfilling into water is necessary. It is approved that so - called dry rockfill can also invoke some technical difficulties resulting from technological approach.

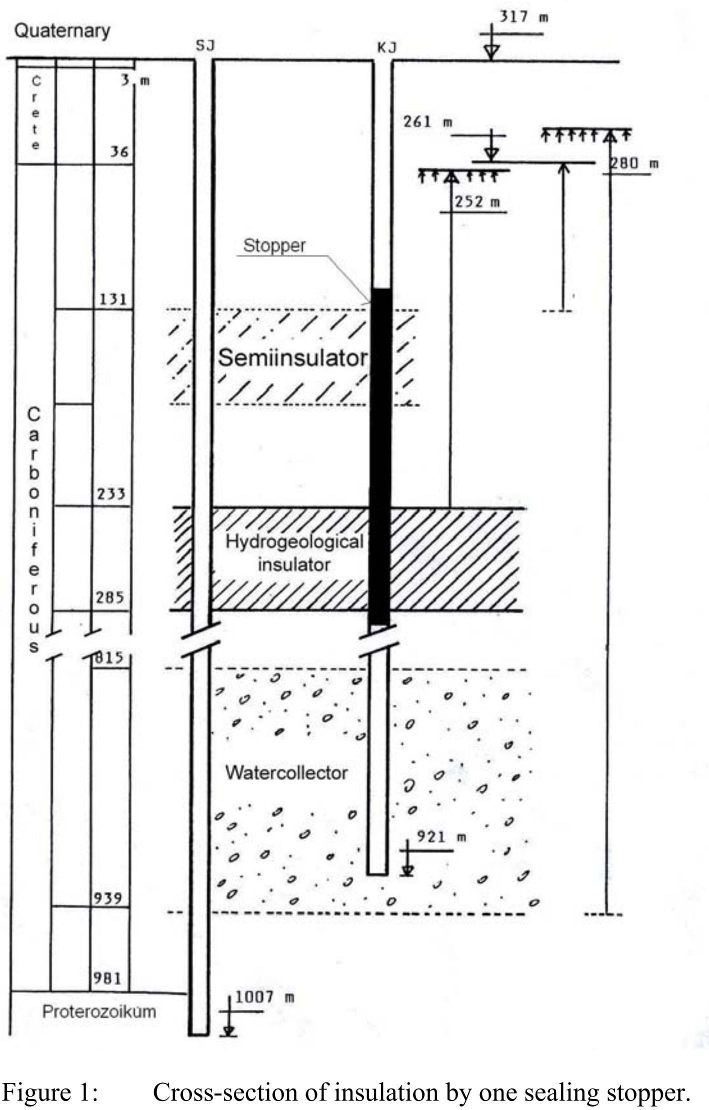


Figure 1: Cross-section of insulation by one sealing stopper.

The principle of the whole problem consists in an assessment of time-dependent behaviour of the backfill, which is found in the subsoil of sealing stoppers, especially concerning its sedimentation and compression. These factors basically influence the function of a sealing stopper.

Providing a decreasing inlet of water to the shafts in dependence of the increasing water level and a rate of strew of 20 m to 40 m a day it is possible to realize such a situation, when backfill strewing will be faster then the water level of the increasing water. Then such a situation can occur that at the height of the dry backfill higher than approximately 20 m due to larger value of angle of internal friction the principal part of the vertical load caused by volume weight is transferred to the lateral sides of the shafts. This phenomenon is dependent on moistening of this material. Correct function of the sealing stopper on such subsoil is very hard and difficult to ensure. If we take into consideration additional inflows of underground water, damage of the sides of shafts, or fluctuation of surface water, it could happen that sinking of the ground of the stoppers appears to be out of control and, consequently, undesirable free space and holes could be created. Too soft stoppers result in a decrease of its required sealing ability. In this case the backfill would be moistened and afterwards it attains the required settlement.

The next problem of identification of the material being necessary for sealing of the stoppers appears to be an assessment of the mechanically physical properties of a bulk spoil material in a high water column (almost to 600 m). By virtue of the following analysis of mechanical and physical properties of the soil and rock in both shafts we may conclude: There is a high degradation of mechanical and deformational characteristics of materials due to the water. Reduction of the deformational characteristics (E – Young's modulus) from about 50 to 60% and mechanical properties from about 30 to 40% results from many performed tests, Tab. 1, where R_d is the compressive strength, R_t is the tensile strength, E_o is the deformation modulus of elasticity, ν is Poisson's number, ϕ is the angle of internal fiction, and R_{sh} is the shear strength. To this it is necessary to count a decrease of the above values owing to transport, incidence and climatic influences, setting material into the shafts, etc. Because of a clash of the backfill material against the water surface and its subsequent spreading out the total decrease of its strength occurs and total destruction of the material appears.

Before placement the relationship between solid and lump material and cohesionless material is estimated at c 1:4. This ratio, after performing the backfill into the water, will further worsen the solid and cohesive material. Then, the relationship between solid and lump material and cohesionless material is changed to c 1:6.

In this technology it is almost impossible to make a selection of the quality of the spoil material. The next problem is to assess the material's properties and mechanical hydraulic behavior of the backfill.

In principle two assumptions exist about the behavior of the backfill below the water:

1. The supporting skeleton of the backfill will have a relatively large porosity (30 to 40 %). Spaces between single solid pieces will successively be filled with fine sediment elements. However, there is a possibility that in the liquid of relatively high density (to 1.3 g/cm) clay contacts are created between single grains of the skeleton, which tighten individual cavities and therefore prevent the next settling.

Table 1: Material properties of the spoil soil.

	middlegr. sandstone		sandy siltstone		silt sandstone		clay siltstone	
	shaft	dump	shaft	dump	shaft	dump	shaft	dump
R_d	46.9	35.2	78.3	15.1	67.6	66.1	15.1	15.1
R_{tp}	2.47		2.98				2.9	
E	7487		10267	2550	18063	6280	9528	
E_o	6603	1200	10209		14816		7366	
ν	0.19		0.26		0.36		0.4	
R_{sh}	4.9		6.7		9.9		7.55	
ϕ	61.6	28	65.2	31.2	60.2		67.9	

2. If the backfill consist mainly of the clay material without solid grains, then single parts of clay will be destructed into single clay elements and subsequent sedimentation can appear. Therefore a new compact clay layer has to be created. The same process will obviously proceed also during filling the clay material of the sealing stopper into the shaft. However, in the same time, the consolidation can be slowed down remarkably owing to the upward pressure and the pore pressures. However, the above-mentioned process can hardly be quantified. In any case it is necessary to support this process with the addition of coagulators.

Except for porosity of virgin material in the range 25–45 % the loose material possesses also great external void fraction (ratio of the volume lacunae between single pieces to general volume of the loose material). In this phase pieces are wedged, spun and crushed until a certain value of the load, the fresh loose material is very permeable. Therefore, it is also very compressible. The speed of the sedimentation is, in this start-up phase, more dependent on changes of the moisture than on time. Only after minimizing of the void fraction, which can take place with depths greater than 50 m, can the process of the classical filtration consolidation with an appropriate reduction of the porosity obviously commence. This process is already markedly time-dependent. The consolidation successively slows at the stable level of the normal load. From results of the shear tests it follows that the values of the shear strength of the clay loose material are impossible to deduce from shear strengths given from entire specimens. It can be stated that the more coherently is overburden removing soil, the less cohesive is its loose material.

Shear tests with loose materials under normal pressures to the 3 MPa brought about piece of knowledge that overstepping the normal tension by about 1–1.5

MPa causes crashing of their structure. This process is in compliance with geostatic pressure in depths 50–75 m. At the same time the angle of internal friction decreases by one half and the growth of the cohesion is doubled.

This phenomenon is observed at the clay loose material with moisture of max. 25–30 % (in dry matter). If the moisture is higher, the described phenomenon shifts to the area of the lower pressure, in some case it disappears. In order to gain this change it is sufficient to moisten the specimens about 5–7 %. Difficulties of the laboratory determination of the shears strength of clay loose material results from the high normal stress. At small stress states the loose material is softened, the total parameters of the shear strength are basically coincident with effective parameters. In the next increasing of the normal stress over certain limit (1-1.5 MPa) pore pressure and air pressure gradually increase. After total decreasing of the void fraction in the loose material the normal stress due to increase of loading is transmitted only to water. Mobilized shear stress in the loose material does not virtually increase. The settlement of the surface of loose material is an external manifestation of various deformation processes that proceed in the entire body. Changes are time-dependent, their size and speed depend on the possibility of the saturation of water. Lumps of the freshly based loose clay material to the dry environ have a solid to rigid consistence. In consequence of the unloading at the excavation of the sealing clay material is in pores of single pieces and the negative pore pressure, because lumps easily absorb the moisture from the neighborhood.

The loose clay material with the initially considerable porosity (40 %, for example) after some time at higher normal pressure (cca 1.5 MPa) turns to become waterproof and airtight.

As far as the deformation processes pass under restricted access of the waters they are caused by crushing of lumps on contacts of the surfaces. After the forthcoming reduction of the void fraction a closure of the air occurs. In this way the agglutinate material is almost waterproof. For such materials, it threads no extreme worsening of strength and deformation characteristics.

On the contrary, the worst properties of the loose material are attained if all space is fulfilled by water. External influences of water are the most dangerous for dry loose material, as the loose material, in the beginning, is permeable. In this case we explain, for instance, of the turn of the loose material to the dump. Gradual plasticization will lead to homogenization of the soil. Mostly, we only obtain the soft till pulpy (pasty) consistence.

Characteristics changed in this way do not allow the next rising of the thickness of the outer dump. At the moment of the placing of next thicknesses all addition of the normal load relates to the pore water with the zero shear strength. The dump becomes unstable as a consequence of active forces under current stagnation of passive forces. The settlement of the soft till mash material can be smaller, providing prevention of the extrusion of the dump to the sides. This case is also possible at the pelted shaft of the deep mine.

Now we try to make an analysis of the progress at which the sealing clay material stopper spill from the big high to the water. The main problem is the behavior of the diffused clay material in the surrounding water. The

gravitational effect dominates during the settlement of grains. The interaction between distinctive grains begins at that time, when forces among single elements dominate over gravitational forces.

3 Numerical description of streaming water

In this section we briefly formulate well known relations with the only target that some known coefficients will have special meaning, and are introduced for tuning with results from scale models. After that, simulation and optimization of the different types of stoppers and their number are considered in numerical models. We start with denotation of two phases composite. The water pressure will be described by function $p = p(\boldsymbol{x},t)$ depending on a position and time. Displacement field in the skeleton is described by function $\boldsymbol{u}(\boldsymbol{x},t) = \{u_1(\boldsymbol{x},t), u_2(\boldsymbol{x},t), u_3(\boldsymbol{x},t)\}$. Time dependent changes of pores are denoted as $\dot{\Theta}$. Total stress is split into effective stress and the influence of water pressure:

$$\sigma_{ij} = \sigma_{ij}^{\text{eff}} - p\delta_{ij} \ .$$

where δ_{ij} is Kronecker's delta indicating that the water pressure does not effect the shear components of stress tensor. The same is valid for strains. If we write the tensor of stresses artificially in terms of vector:

$$\boldsymbol{\sigma} = \{\sigma_{xx}, \sigma_{yy}, \sigma_{zz}, \sigma_{xy}, \sigma_{xz}, \sigma_{yz}\}^{\text{T}} .$$

We also use this notation for strain tensors:

$$\boldsymbol{\varepsilon} = \{\varepsilon_{xx}, \varepsilon_{yy}, \varepsilon_{zz}, \varepsilon_{xy}, \varepsilon_{xz}, \varepsilon_{yz}\}^{\text{T}} .$$

Then

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}^{\text{eff}} - p\boldsymbol{m}$$

where p is scalar and $\boldsymbol{m} = \{1, 1, 1, 0, 0, 0\}^{\text{T}}$. The effective stress can be expressed as:

$$\dot{\boldsymbol{\sigma}}^{\text{eff}} = \boldsymbol{D}(\dot{\boldsymbol{\varepsilon}} + \dot{\boldsymbol{\varepsilon}}^{\text{pl}} - \dot{\boldsymbol{\varepsilon}}_{\text{p}})$$

where dot means time derivative, $\boldsymbol{D}$ is stiffness matrix of elastic media, $\boldsymbol{\varepsilon}$ is elastic stress tensor, $\boldsymbol{\varepsilon}^{\text{pl}}$ is the strain in plastic domain, and $\boldsymbol{\varepsilon}_{\text{p}}$ means the strain tensor caused by water pressure. The time increment of the latter strain (velocity of deformation of the basic material due to pore pressure) can be expressed as:

$$\dot{\varepsilon}_p = -\boldsymbol{m}\left(\frac{\dot{p}}{3K_z}\right),$$

where K_z is the bulk modulus of the basic material. Compressing the liquid using the pore pressure we get after easy algebra the following relation:

$$\dot{\sigma} = \dot{\sigma}^{\,eff} - \dot{p}\boldsymbol{m} = (\dot{\boldsymbol{\varepsilon}} + \dot{\boldsymbol{\varepsilon}}^{\,\mathrm{pl}}) - \alpha \boldsymbol{m}\,\dot{p}$$

where $\alpha = 1 - \frac{1}{3K_z}$, K_z is the coefficient of expansion due to the water. The time increment of changes of pores can be expressed as

$$\dot{\theta} = \alpha\, \boldsymbol{m}^{\mathrm{T}} \dot{\boldsymbol{\varepsilon}} + \left[\frac{n}{K_k} + \frac{1-n}{K_z} + (1 - \alpha)\frac{1}{K_z} \right] \dot{p}$$

where n (porosity) is the increment of the volume of skeleton due to an increment of pore pressure. On the other hand Darcy's law is valid

$$n\,\mathrm{div}\,\mathbf{v} = -\dot{\theta}$$

where $\boldsymbol{v}$ is the vector of relative velocity of the displacement of the liquid with respect to the displacement of the skeleton. According to Darcy's law the relative velocity of streaming liquid is proportional to the gradient of pore pressure:

$$-\frac{c}{K_s}\Delta p + \lambda\,\dot{p} + \alpha \boldsymbol{m}^{\mathrm{T}} \dot{\boldsymbol{\varepsilon}} = 0\,,\ \lambda = \frac{n}{K_k} + \frac{1-n}{K_z} - (1 - \alpha)\frac{1}{K_z},$$

$$\text{for } \alpha \approx 0:\ -\frac{c}{K_s}\Delta p + \alpha \boldsymbol{m}^{\mathrm{T}} \dot{\boldsymbol{\varepsilon}} = 0\,,$$

where K_s is the bulk modulus of elasticity of the skeleton, $c = \frac{K_s k}{\rho_k g}$, ρ_k is the density of the fluid and g is gravitation, k is the coefficient of filtration.

The last equation together with the Navier equations creates a system of partial differential equations, which are solved by finite or boundary elements.

4 Laboratory modeling

The mathematical model serves for solving water pressures in pores and overall (total) stresses in the insulator. Note that the components of stress tensor are virtually impossible to determine in the experimental and "in situ" measurements. On the other hand, no numerical model can be applied without appropriate material and internal parameters. This is why experimental scale models were prepared and the process of construction of the insulator was

simulated in the stands, which are partly glassed basins boxes. The sides of the stands are about 1.5 x 1.5 x 1.5 m^2. Natural materials fill the stands and the water is pushed through the pores of the material.

Classification of the insulator material originates from the description of samples from the site and laboratory tests. With respect to unexpected inflow of ground water and to spare the energy, the continuation or drainage was excluded and the project expected no penetration of saline water to the upper level. Since the problem consists of the time dependent behavior of the insulator, particularly its sedimentation, compression, seepage, change of coefficients of filtration ($k_r = k_\varphi = 0.8\ k_z$; all these coefficients (radial r and hoop φ) are dependent on z-direction, which is axial - vertical - coordinate), and the solution is strongly nonlinear. All these factors essentially influence the insulation function of the water stopper. Under assumption of decreasing inflow of water into the shafts in dependence of increasing level of water and the velocity of construction of the stopper (20 to 40 m a day) it is realistic to imagine a situation, the upper level of the stopper in which "outruns" the increasing water level. If the stopper is finished and the fresh water is let to fang to the shaft, both types of waters should not be in contact, i.e., the upper level of saline water and the lower level of the fresh water may not meet at any stage of construction.

Table 2: Water pressure at different points of the stopper.

$z \backslash r$	0.0	0.6	1.2	1.8	2.4	3.2	3.8	4.45
40	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
37.5	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001
35	0.000	0.000	0.001	0.002	0.002	0.001	0.001	0.002
32.5	0.002	0.002	0.002	0.004	0.004	0.004	0.005	0.005
30	0.004	0.007	0.011	0.013	0.015	0.015	0.017	0.019
27.5	0.014	0.021	0.034	0.046	0.054	0.057	0.057	0.057
25	0.037	0.060	0.107	0.148	0.175	0.189	0.194	0.195
22.5	0.080	0.121	0.205	0.279	0.329	0.354	0.365	0.369
20	0.150	0.223	0.378	0.518	0.609	0.658	0.673	0.675
17.5	0.271	0.406	0.690	0.950	1.122	1.208	1.236	1.239
15	0.483	0.725	1.248	1.733	2.056	2.221	2.266	2.265
12.5	0.835	1.262	2.212	3.127	3.758	4.083	4.159	4.148
10	1.358	2.084	3.759	5.496	6.801	7.518	7.684	7.650
7.5	2.004	3.131	5.877	9.062	11.908	13.837	14.461	14.438
5	2.542	4.050	7.931	13.020	18.778	24.742	28.426	29.187
2.5	2.597	4.205	8.440	14.416	22.530	35.209	58.596	74.068

The scale modeling proves that the correct function of the stopper is very tough and difficult. Moreover, the scale modeling proved again that in the final stage of material characteristics changed principally, e.g., Young's modulus E dropped by 50 to 60 percent and toughness by 30 to 40 percent.

5 Conclusions

In this paper coupled modeling (sort of back analysis) is applied to solution of optimal design of sealing stoppers, which should substitute natural clay insulators dividing two sorts of water. They are built up as recovering of the situation after abandonment of deep mines. The numerical model serves as a tool for decisions on how to design the stoppers in an optimal way, the scale models are tools for delivering necessary data to the numerical model.

We first introduce very important distribution of coefficients of filtration k (m/s) in vertical direction depending on moisture w (%), which was derived from large scale of scale models. The relation is depicted in Fig. 2.

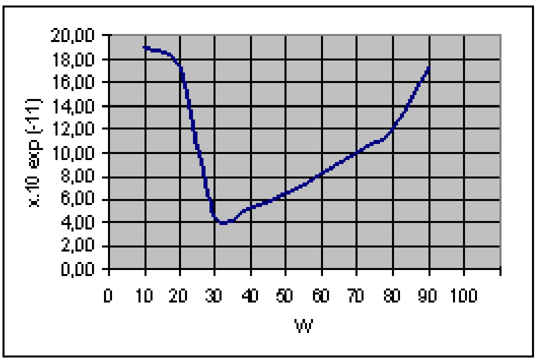


Figure 2: Relation of coefficient of filtration and moisture.

From finite elements distribution of water flow is computed and similarly from Lame's equations stresses are calculated using finite elements again. In certain points strength in the sense of Mohr-Coulomb hypotheses (ideally elasto-plastic law) is attained and the disturbance function desires change in its value at such points (having impact on the coefficients of filtration and others). The resulting pressure for optimal configuration of the stopper appears in Tab. 1, where rows denote the distance from the bottom of stopper in m, r is the radius measured in meters from the shaft axis. It is obviously seen that the highest values of pressure are observed at the outer ring of the stopper and on the lower boundary. Moreover, 40 m from the lower boundary almost no pressure is obtained in the final stage of the construction of the stopper. The distribution of water pressure along the lower boundary shows that the higher pressure is observed at the concrete lining, the lowest at the axis.

Acknowledgments

The first author was financially supported by GA AV ČR, project No. IAA 2119402, the second author appreciates the financial support of GA ČR No. 103/05/0679.

Influence of wind on floating debris distribution in the Balearic Islands

L. Martinez-Ribes[1], G. Basterretxea[1], L. Arqueros[2], A. Jordi[1], T. Estrany[3], J. M. Aguilò[2] & J. Tintoré[1]
[1]*Mediterranean Institute of Advanced Studies, Balearics, Spain*
[2]*Coordination Centre for the Quality Plan for Bathing Waters, Environmental Council, Balearics, Spain*
[3]*Department of Mathematics and Information Technology, Illes Balears University, Spain*

Abstract

In this work we analyse the distribution of floatable debris around the Balearics (Spain) during the summer of 2005 (June-September). Litter abundance and distribution was obtained from the registers of 40 cleaning-ships operating daily around the archipelago. A total of 133.671 Kg. of debris was collected in the coastal waters of Mallorca, and the main component was wood. Different scenarios of litter dispersal are analysed at selected coastal bays depending on wind direction.
Keywords: Balearics, Mediterranean Sea, floating debris, cleaning, wind effect, tourism.

1 Introduction

The economy of the Balearic Islands is based on the tourism sector, and tourism activities are mainly related to the recreational use of the coast. Preserving the quality of the coastal waters is, consequently, a major objective of the local government. As part of the maintenance activities carried out in the coastal environment, 40 cleaning boats sweep the 1,400 km of coastline, collecting floating debris during high tourism season (June to September). This marine litter program requires a strong human effort and an expenditure exceeding 3.5 M€. Cleaning activities are paralleled by a scientific program aimed to detect the origin and distribution of marine debris. It is believed that scientifically based

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060541

knowledge about litter sources and pathways will not only allow the optimization of cleaning efforts but, also, will make possible implementing new palliative strategies and protective policies. Once released into the ocean, floating debris may be concentrated by natural processes along lines of convergent flow, at the core of major current gyres or on beaches and submerged rocky outcrops (Laist [1], Pruter [2]; Galgani *et al.* [3]). As environmental features affecting the distribution of litter and its arrival to beaches, the direction of wind, surface waves and currents (Shaw and Mapes [4]; Pruter [2]; Thiel *et al.* [5]; Abu-Hilal and Al-Najjar [6]) are amongst the most relevant. Vauk and Schrey [7] suggest that floating materials at the surface are primarily influenced by winds and only secondarily by currents.

The removal of floating debris from the sea is not only important because of the visual impact that litter produces on users, but also because of the harmful effects it generates in the marine environment and on the activities developed on the coastal waters. Economic losses from marine litter include the fouling of trawl nets by bottom debris, the blocking of water intake pipes by plastic sheeting and damage to ships and recreational boats following collisions with metal drums or wooden pallets at sea, as well as propeller fouling by floating or semi-submerged nylon ropes and net fragments (Dixon and Dixon [8]). As for the marine ecology, several studies point out the importance that litter has on the life of marine animals such as turtles, and how all kinds of species can be affected by its presence. For instance, the ingestion of plastics in sea turtles in the Mediterranean was studied by Tomás *et al.* [9]), who concluded that more than a 75% of what was found in turtles' stomachs was composed of plastic. The aim of this work is to analyze the characteristics and distribution of floating debris at two selected bays in Mallorca (Spain) and more specifically, the effects of wind in debris distribution along the coast as given by the positions of the cleaning boats.

2 Materials and methods

Herein presented floating debris data was recorded onboard the fleet of cleaning ships operating on a daily basis in the Balearic Islands during summer 2005. Two types of ships are presently used in cleaning tasks; the *Virots* which stay in the nearshore waters (i.e. ¼ mile from shore) and the *Pelicans*, which are able to operate within 4-5 nautical miles from the coast. The GPS position of each unit is visualized and recorded every 3 minutes in the Water Quality Control Office. The routes that the ships follow are selected in terms of sea conditions and litter location information obtained with airborne methods. Collected floating debris is sorted by type of material (i.e. plastics, woods, organic matter, seaweeds and other) and its weight recorded at the end of the day.

As a case study two different bays in Mallorca are selected, Palma and Santa Ponça (see Figure 1). The former is a large bay (~185 km^2) beholding the city of Palma, the largest population (476.000 inhabitants in the metropolitan area) and the main port of the islands. Santa Ponça is a smaller embayment (~32 km^2) located west of Palma bay, mainly dedicated to tourism activities. Both bays are

opened to the south west and present a similar wind regime (see Basterretxea *et al*. [10]). A data set from 9 cleaning boats (6 operating in Palma and 3 in Santa Ponça) containing 73,404 GPS positions has been analyzed. Values corresponding to static ship positions (< 0.5 knots) and steaming data have been disregarded. Density maps have been produced pooling the data in 200x200 m grid cells and standardizing the results of each case by the total number of data. The final data set for each area has been divided attending to the prevailing wind conditions of each day. Meteorological data for the area was obtained from the Spanish National Institute of Meteorology (INM). The wind data was divided in 90 degree sectors based on the mean wind vector direction for the first 12 hours of each day. Direction 1 stands for winds coming from 0˚ (North) to 90˚ (E); direction 2 for winds from 91˚ to 180˚; direction 3 from 181˚ to 270˚ and direction 4 from 271˚ to 359˚.

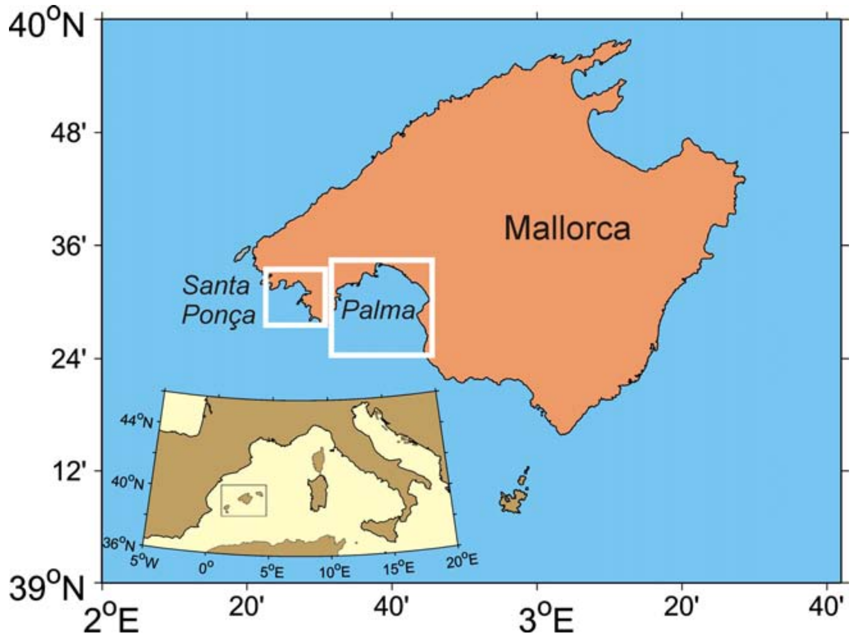


Figure 1: Map of Mallorca indicating the location of Palma and Santa Ponça Bays.

3 Results and discussion

A total amount of 206,758 kg of floating debris was collected during the summer 2005 in the waters of the Balearic Islands, 13% corresponding to the material collected in Palma Bay and 2% in Santa Ponça. Palma is the most polluted area of the archipelago; however, similar amounts of litter were collected in other areas with lower human pressure (i.e. Pollença, in the northern part of the island). Monthly variation of collected litter reveals a similar trend during the

first three months of the survey but, strikingly, large variations are observed in September in the bay of Palma (Figure 2). These variations are attributed to an increase in the inputs of natural wood of terrestrial origin during a storm that took place between September 7th and 9th. Indeed, the highest precipitations of 2005 were recorded during this month (Figure 3). The larger extent of Palma watershed, more than six-fold that of Santa Ponça, may favour the arrival of important amounts of rests of vegetation to this coastal area.

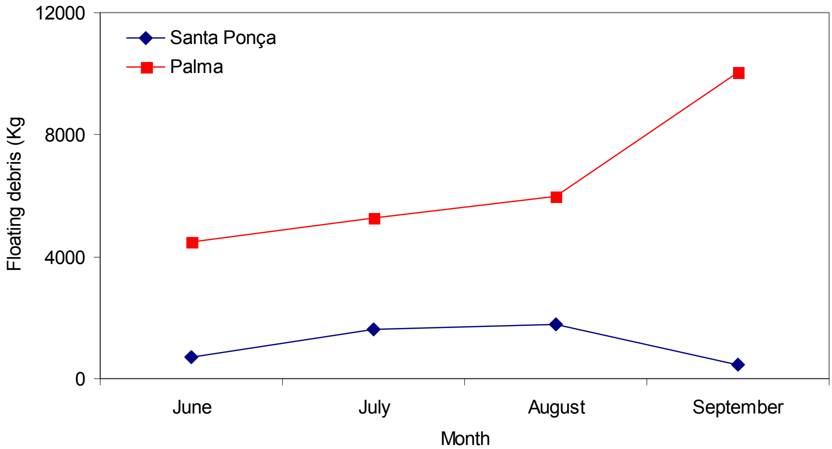


Figure 2: Variation of the amount of floating debris collected in the bays of Palma and Santa Ponça during the 2005 summer cleaning campaign.

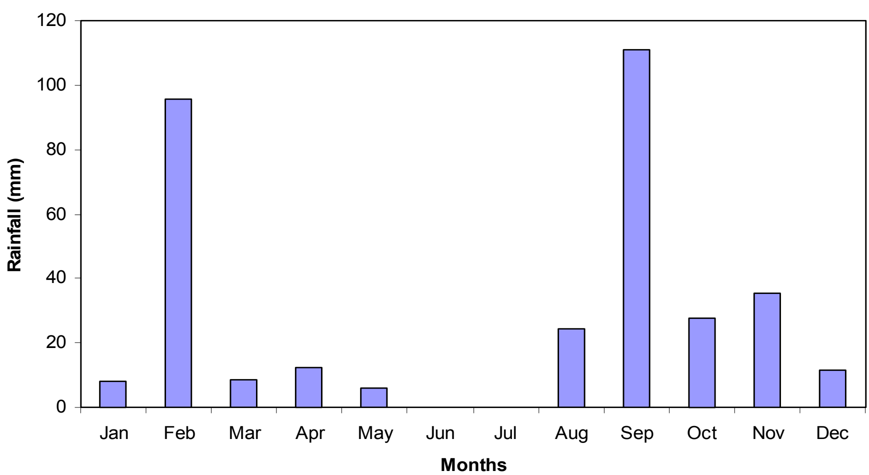


Figure 3: 2005 Monthly values of rainfall at Palma Airport.

The classification of debris into different types of materials shows that wood and its derivates is the most important debris (in weight) in the four months in Palma followed by plastics (Table 1). These two materials account in this bay for between 73 and 86% of the total weight. Similarly, these materials contribute to an average of 66.3±3% of the litter in Santa Ponça, where seaweeds and other materials of heterogeneous nature become more important. The amount of organic matter collected at Santa Ponça is also remarkably higher with the exception of the value in Palma for august, when a large sperm whale rest (5 m long) was collected. In addition, a marine turtle (*Caretta caretta*) was found. It should be noted that although our data referred in terms of weight following the standard procedure of local authorities, when the values are interpreted in terms of collected items plastics become most important. Most of these items (70%) where identified as being of local origin (nearby supermarkets and shops) and presented a low degree of degradation. Wind transport could be an important source of litter to coastal waters, particularly in the case of plastics.

Table 1: Monthly variation in the percentage of floating debris collected at Palma and Santa Ponça (measurements in kg).

	June	*July*	*August*	*September*
PALMA BAY	%	%	%	%
Wood	56.2	42.4	39.7	73
Organic matter	3.7	3	12.3	5.4
Plastic	29.7	33.3	33.1	12.7
Oil	0	1	1.1	0
Algae and seaweeds	5	12	7.3	4.4
Other	5.4	8.2	6.6	4.4
	June	July	August	September
SANTA PONÇA	%	%	%	%
Wood	19.7	30.4	41.5	34.2
Organic matter	9.4	9.8	8.4	7.5
Plastic	47.2	36.2	20.3	35.8
Oil	0	0	0	0
Algae and seaweeds	12.4	11.8	10.3	16.6
Other	11.4	11.8	19.5	5.8

As indicated by the wind rose of Palma Airport, winds blowing from the first and third quadrant are predominant during summer (Figure 4). Wind vectors from SSW and SW are related to the sea breeze in the bay, that rarely exceeds 7 m/s (Werner et al. [11]). Winds from the NNE and ENE are generally milder (<5 m/s) and associated with the land breeze component. However, periodical strong Tramontana events also blow from this direction.

The GPS positions of the cleaning boats corresponding to the above mentioned predominant wind directions have been analyzed for the two study

sites. As shown in the density maps, cleaning activities concentrate in a short portion of the coast, usually in transects running around harbours (Figure 5). Cleaning boats focus their activity in small embayments, where the geomorphology favours the concentration of debris. This is clearly seen in the embayment located to the west of Palma (Cala Nova), and it is particularly remarkable in the west zone of Santa Ponça (Camp de Mar). Moreover, the highest density of boat positions concentrates in the vicinity of Palma harbour, including its inner port. It is in this area where both, the metropolitan sewage and drainage systems converge. Nevertheless, their relative contribution to the total litter amount has not been established, and other sources in the area such as port activities should be considered.

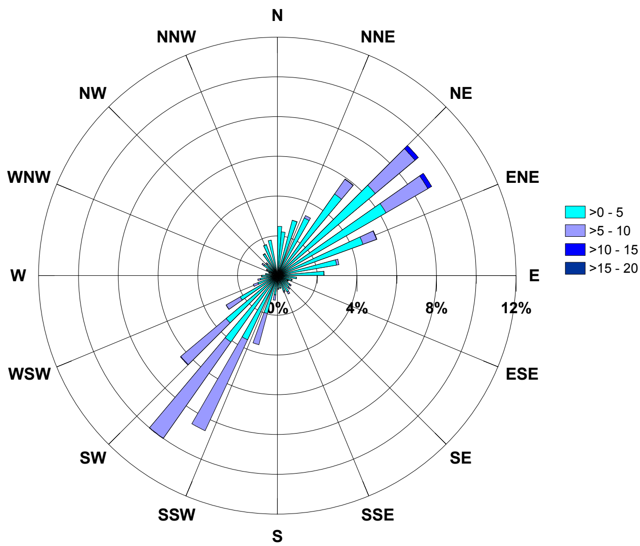


Figure 4: Wind rose for Palma Airport. Data from INM.

Regarding the wind effect, there are not important differences between the analyzed wind directions. Two plausible explanations could explain the low variability observed: (1) debris is collected close to its source and therefore there is no time for wind induced litter dispersal; (2) boats concentrate their cleaning activities based on the skippers experience and only change their routes when other sources of information are provided (i.e. airborne observations, user's complaints). Despite these biases, some relevant features can be depicted from the cleaning routes. During SW winds at Palma, the collection of debris takes place primarily close to the harbour and to the west of the bay (Cala Nova). In the other case (NE winds) debris is more dispersed in the eastern coast of Palma and is gathered farther from the coast. This area corresponds to a long (4.5 km) sandy beach known as El Arenal where, when the winds blow from the sea, cleaning activities concentrate closer to the shoreline and should be basked by

land-based operations. This wind effect is also notable in the Bay of Santa Ponça, where there is a notable spreading under northeasterly wind conditions and a higher cleaning effort at the east of the bay during SW wind conditions.

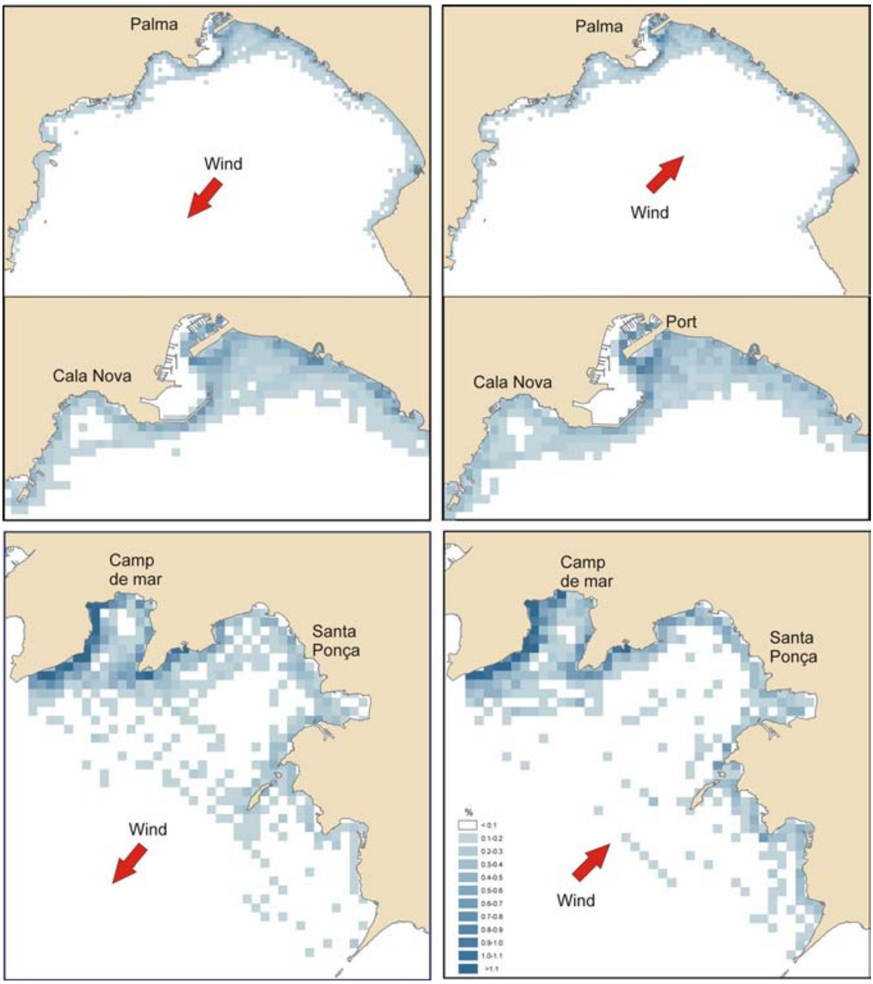


Figure 5: Standardized density maps of cleaning boats positions.

Furthermore, when observing (Figure 6) the quantities of flotsam collected under each of the wind scenarios, in both bays the influence of winds blowing from directions 1 and 3 produces a higher concentration of debris, which is reflected in the amount of it collected. As it was mentioned, the quantities in the Bay of Palma are higher than in the Bay of Santa Ponça. The lowest values were recorded under wind 4 in both bays.

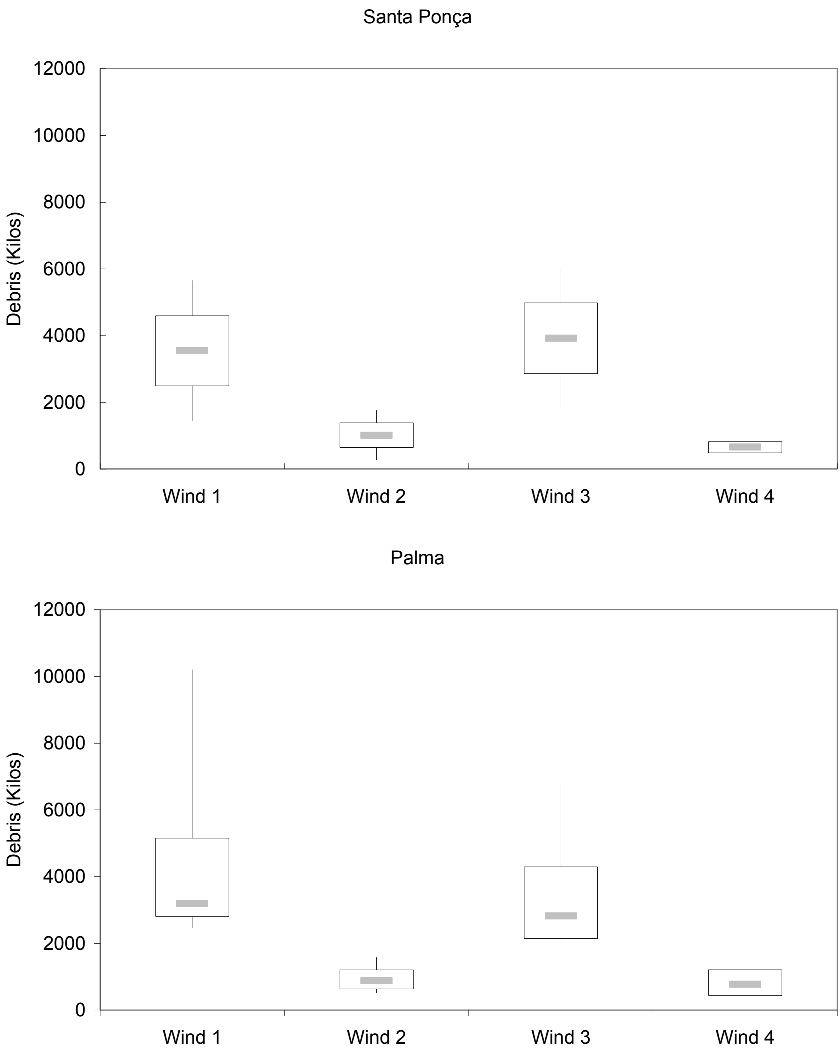


Figure 6: Amount of debris collected under different wind scenarios in the two bays studied, Santa Ponça and Palma. The grey lines represent the median values, and the vertical lines represent the maximum and minimum values.

4 Conclusions

The debris scenario for two bays at Mallorca has been analysed using the positions of cleaning boats. By weight, wood has been the major component of the litter collected, followed by plastics. If counting the number of items, plastics

would be the commonest type of material. Given the high floatability of both kinds of material the wind regime on the island was analysed, showing certain effect on the distribution of the debris. However, other factors are to be thought of, as for instance the geomorphology of the area and the possible sources of the debris. In the study herein presented, local plastics had a high percentage of presence (about 70%), which indicates how important it is to follow good disposal procedures.

Acknowledgements

This work was carried out as part of the program 'Convenio para la evaluación y monitorización de la calidad de las agues costeras de las Islas Baleares' financed by the Conselleria de Medi Ambient of the Balearic Government. We are indebted to Joan Vallespir and Maurici Ruiz for assistance in GIS based data processing.

References

[1] Laist, D. W., Overview of the Biological Effects of Lost and Discarded Plastic Debris in the Marine Environment. Marine Pollution Bulletin, **18**, No. 6B, pp. 319-326, 1987.
[2] Pruter, A. T., Sources, Quantities and Distribution of Persistent Plastics in the Marine Environment. Marine Pollution Bulletin, **18**, No. 6B, pp. 305-310, 1987.
[3] Galgani, F., Leaute, J. P., Moguedet, P., Souplets, A., Verin, Y., Carpentier, A., Goraguer, H., Latrouite, D., Andral, B., Cadiou, Y., Mahe, J. C., Poulard, J. C. & Nerisson, P., Litter on the Sea Floor Along European Coasts. Marine Pollution Bulletin, **40**, No. 6, pp. 516-527, 2000.
[4] Shaw, D. G. & Mapes, G. A., Surface Circulation and the Distribution of Pelagic Tar and Plastic. Marine Pollution Bulletin, **10**, pp. 160-162, 1979.
[5] Thiel, M., Hinojosa, I., Vásquez, N. & Macaya, E., Floating marine debris in coastal waters of the SE-Pacific (Chile). Marine Pollution Bulletin, **46**, pp. 224-231, 2003.
[6] Abu-Hilal, A.H. & Al-Najjar, T., Litter pollution on the Jordanian shores of the Gulf of Aqaba (Red Sea). Marine Environmental Research, **58**, pp. 39-63, 2004.
[7] Vauk, G. J. & Schrey, E., Litter pollution from ships in the German Bight. Marine Pollution Bulletin, **18**, No. 6B, pp. 316-319, 1987.
[8] Dixon, T. J. & Dixon, T. R., Marine Litter Surveillance. Marine Pollution Bulletin, **12**, No. 9, pp. 289-295, 1981.
[9] Tomás, J., Guitart, R., Mateo, R. & Raga, J. A., Marine debris ingestion in loggerhead sea turtles, *Caretta caretta*, from the Western Mediterranean. Marine Pollution Bulletin, **44**, pp. 211-216, 2002.
[10] Basterretxea, G., Garcés, E., Jordi, A., Masó, M & Tintoré, J., Breeze conditions as a favoring mechanisms of *Alexandrium taylori* blooms at a

Mediterranean beach. Estuarine, Coastal and Shelf Science, **62**, pp. 1-12, 2005.
[11] Werner, F. E., Viúdez, A. & Tintoré, J., An explanatory numerical study of the currents off the southern coast of Mallorca including the Cabrera Island complex. Journal Marine Systems, **4**, pp. 45-66, 1993.

Section 7.1
Management of complex systems under extreme conditions, organised by L. Fagherrazzi

This page intentionally left blank

The management of a multi-purpose reservoir

A. Turgeon
Département de mathématiques et de génie industriel,
École Polytechnique de Montréal, Canada

Abstract

This paper addresses the problem of determining the daily operating policy of a reservoir over a one-year period that satisfies conflicting objectives regarding hydropower generation, flood control, navigation, and recreational activities. This optimization problem is stochastic because streamflows are random and can be forecasted only a few days in advance. One of the difficulties involves taking streamflow persistence into account and, more specifically, the fact that there may be long periods of high and low flow. This is important because floods usually occur during long periods of high flow, while water shortages occur during long periods of low flow. The paper shows how to solve the problem with rule curves, Dynamic Programming, and simulation.
Keywords: daily reservoir operation, multi criteria, stochastic optimization, rule curves, multi-lag autoregressive models.

1 Introduction

This paper deals with the problem of determining the optimal daily operating policy of a multi-purpose reservoir over a one-year period. The problem is stochastic because the reservoir inflow is random and cannot be predicted long in advance. The difficulty in solving this problem stems from the fact that there are many competing objectives to meet and the daily inflow in usually correlated with the inflows from the preceding days. The objectives may involve water quality, the preservation of endangered species habitats, recreational activities, flood control, water supply, navigation and hydropower generation (Eschenbach et al. [1]).

The multi-objective reservoir management problem has been the subject of several publications in the past. A list of these publications can be found in the survey papers of Yeh [2], Wurbs [3] and Labadie [4,5]. The paper by Cohon and

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060551

Marks [6] describes and compares the different techniques used for solving multi-objective optimization problems. These techniques are divided into three classes: the generating techniques, techniques that rely on the prior articulation of preferences, and techniques that rely on the progressive articulation of preferences. The generating techniques, like the weighting and constraint methods, were the first multi-objective solution procedures that were developed. They are easy to apply, but are not recommended for problems with more than three objectives (Cohon and Marks [6]). The second class of techniques includes the Goal Programming method by Charnes and Cooper [7] and the Surrogate worth trade off method by Haimes and Hall [8]. The third class consists of iterative and step methods, like the Stem method by Benayoun et al. [9].

Multi-objective optimization has been applied to the management of several rivers, and notably to the Trent River in Canada, the TVA system in U.S., the Svarta River in Sweden, the Iguaçu and San Francisco rivers in Brazil, the Chaliyar River in India, and the Tone River in Japan. It has also been applied to the management of Lake Como in Italy and the Hoover reservoir in U.S. The Trent River basin covers an area of 4400 square miles and contains 48 reservoirs and 14 hydroelectric power plants. The multi-objective operating policy of this huge basin is determined iteratively with a simulation model (Sigvaldason [10]). The daily management of the TVA system, which consists of 46 reservoirs, is done using a decision support system called RiverWare (Eschenbach et al. [1]). This system solves the multi-objective optimization problem with the Goal Programming method. The Svarta River basin, located in south central Sweden, has two reservoirs which are managed to satisfy four competing objectives: hydroelectricity generation, irrigation, water supply, and urban water supply. The problem of allocating the water between the four objectives is solved with a weighting method (Goulter and Castensson [11]). The optimal daily operating policy of the four hydroelectric power plants on the Iguaçu River in Southern Brazil is determined with Dynamic Programming. Two of the four power plants are own by a utility named COPEL and the other two by a utility named ELECTROSUL. Since the objectives of both utilities must be satisfied, the optimization problem has therefore more than one objective and is thus a multi-objective problem. The daily operating policy of the Sobradinho reservoir on the San Francisco River in Brazil was determined with a technique called Sampling Dynamic Programming (Dias et al. [12]). The problem has two objectives: to maximize the production of electricity and minimize the cost of flooding, which can easily be rewritten as a single objective: to maximize the revenue of the generated electricity minus the cost of flooding. The Chaliyar River basin in India has five reservoirs which are managed to satisfy three objectives: drinking water supply, irrigation, and hydropower generation. The operating policy for these reservoirs is determined with a linear multi-objective programming model that maximizes the hydroelectric generation subject to constraints on the drinking water supply and irrigation. The optimal monthly operation of the three reservoirs in the Tone River basin in Japan was determined with Stochastic Dynamic Programming (Wang et al. [13]). The approach consisted in breaking down the optimization problem of three reservoirs into three sub-problems of

one reservoir. This breakdown was made possible by the fact that the three reservoirs are located on three different tributaries of the Tone River. Stochastic Dynamic Programming was applied to each sub-problem separately. The problem has objectives for the firm water supply, flood control and hydropower generation. The two first objectives were converted into constraints so that the sole objective of the optimization problem was to maximize hydropower generation. The main interest of this paper is that it takes the stochasticity of the reservoir inflows into account, which is not the case for the other papers cited above.

The paper is organized as follows. In section 2, the problem of determining the daily reservoir operating policy that satisfies several objectives is formulated. Section 3 shows how the stochastic optimization problem can be solved iteratively with Dynamic Programming and simulation. Section 4 shows how the lag-n autocorrelation of the inflow can be taken into account by the optimization problem. Finally, section 5 shows how the problem can be solved with rules curves and Dynamic Programming.

2 Problem formulation

The problem consists in determining the volume of water to release from the reservoir each day of the year in order to satisfy the following objectives:

a) Maximize the generation of the hydroelectric powerplant fed by the reservoir;
b) Maintain the reservoir and discharge at a high enough level to preserve endangered species habitat and permit navigation and recreational activities;
c) Minimize the risks of flooding.

Let us represent by:

S_t	the reservoir content at the end of day t in hm^3
R_t	the reservoir discharge on day t in hm^3
Q_t	the reservoir inflow on day t in hm^3
D_t	electricity demand on day t in MWh
$G(S_{t-1}, S_t, R_t)$	the generation of the powerplant on day t in MWh. This generation is assumed to be a function of R_t and the water head corresponding to the average reservoir content for day t, which is assumed to be equal to $(S_{t-1} + S_t)/2$

$\Phi\left(S_{365}\right)$ the value of the water stored in the reservoir at the end of the year,

and suppose that:

- Objective b is satisfied on day t when $S_t \geq s_t^{\min}$ and $R_t \geq r_t^{\min}$
- Floods occur when $S_t > s^{\max}$ or $R_t > r^{\max}$

Unlike $s^{\max}$ and $r^{\max}$, the lower bounds $s_t^{\min}$ and $r_t^{\min}$ are functions of t because environmental constraints, navigation constraints and recreational activities vary throughout the year.

The multi-objective reservoir management problem can be formulated mathematically as follows:

$$\text{Maximize } \sum_{t=1}^{365} G\left(S_{t-1}, S_t, R_t\right) + \Phi\left(S_{365}\right) \tag{1}$$

subject to:

$$S_t = S_{t-1} + Q_t - R_t \tag{2}$$

$$r_t^{\min} \leq R_t \leq r^{\max} \tag{3}$$

$$s_t^{\min} \leq S_t \leq s^{\max} \tag{4}$$

$$G\left(S_{t-1}, S_t, R_t\right) \geq D_t \tag{5}$$

This formulation corresponds to the constraint method in the paper by Cohon and Marks [6]. This formulation supposes that inflows $Q_1, Q_2, ..., Q_{365}$ are known at the beginning of the year. Depending of these inflows, there may or may not be a solution to problem (1)-(5). For instance, if the inflows are very high, it might not be possible to always respect the upper bounds in inequalities (3) and (4) and, hence, avoid flooding. If the inflows are very low, the lower bounds in inequalities (3)-(5) might not always be respected. When there is no feasible solution to problem (1)-(5), the bounds must be modified until a feasible solution is found. The problem is to decide which bounds to modify and by how much.

Problem (1)-(5) can be reformulated in a way that all solutions are feasible. This formulation is obtained by replacing inequalities (3)-(5) by a penalty function in the objective function which assigns a cost to the violations of the bounds. This gives the following model:

$$\text{maximize } \sum_{t=1}^{365} \left[G\left(S_{t-1}, S_t, R_t\right) - L_t\left(S_t, R_t\right) \right] + \Phi\left(S_{365}\right) \tag{6}$$

subject to:

$$S_t = S_{t-1} + Q_t - R_t \tag{7}$$

$$S_t, R_t \geq 0 \tag{8}$$

where:

$$\begin{aligned} L_t\left(S_t, R_t\right) = {} & \alpha_1 \cdot \left[\max\left(0, s_t^{\min} - S_t\right)\right]^2 + \alpha_2 \cdot \left[\max\left(0, S_t - s^{\max}\right)\right]^2 \\ & + \alpha_3 \cdot \left[\max\left(0, r_t^{\min} - R_t\right)\right]^2 + \alpha_4 \cdot \left[\max\left(0, R_t - r^{\max}\right)\right]^2 \\ & + \alpha_5 \cdot \left[\max\left(0, D_t - G\left(S_{t-1}, S_t, R_t\right)\right)\right]^2 \end{aligned} \tag{9}$$

The penalty function $L_t\left(S_t, R_t\right)$ is equal to zero when inequalities (3)-(5) are respected. When they are not, the penalty cost is equal to the sum of the square of the violations. The values of parameters α_i, $i = 1, 2, ..., 5$, should be adjusted to reflect the importance of the objectives. If objective c is more important than objectives a and b, the values of α_2 and α_4 should be greater than those of α_1, α_3 and α_5. If satisfying inequality (5) is more important than satisfying objective b, the value of α_5 should be greater than those of α_1 and α_3. The value of α_1 should be greater than that of α_3 if maintaining the level of the reservoir above $s_t^{\min}$ is more important than maintaining the stream flow above $r_t^{\min}$. Similarly, the value of α_2 should be greater than the value of α_4 if the flood damages are greater when $s^{\max}$ is increased than when $r^{\max}$ is. One can naturally use linear functions instead of quadratic ones in (9) to penalize violations. The advantage of using quadratic functions is that it discourages major violations of the bounds.

3 Optimisation and simulation

The optimization models presented in section 2 can solve multi-objective reservoir management problems only when the reservoir inflows are assumed to be known a year in advance. Since inflows are only known a few days in advance in real life, these models cannot be applied to real problems.

The fact that inflows are not known many days in advance does not mean, however, that there exists no information on the inflows that might occur in the future. If we assume that future inflows will resemble those of the past, historical

inflow data can be used to determine the probability distributions of future inflows. Once the probability distributions of the daily inflows are known, the reservoir operating policy can be determined by solving the following problem:

$$\text{Maximize } E\left\{\sum_{t=1}^{365}\left[G\left(S_{t-1},S_t,R_t\right)-L_t\left(S_t,R_t\right)\right]\right\}+\Phi\left(S_{365}\right) \quad (10)$$

subject to constraint (7)-(9). The symbol E in (10) stands for the expected value. This problem can be solved with Stochastic Dynamic Programming (SDP), which consists in solving the following functional equation backward in time:

$$F_t\left(S_{t-1}\right)=\underset{Q_t}{E}\left\{\max_{R_t}\left[G\left(S_{t-1},S_t,R_t\right)-L_t\left(S_t,R_t\right)+F_{t+1}\left(S_t\right)\right]\right\} \quad (11)$$

with $F_{366}\left(S_{365}\right)=\Phi\left(S_{365}\right)$. When function $\Phi\left(S_{365}\right)$ is not known, the following procedure should be followed. Set $\Phi(\ )=0$ and solve equation (11) backward in time for one year. Next set $\Phi(\ )=F_1(\ )$ and solve equation (11) backward again for another year. Repeat the procedure until the marginal values of $\Phi\left(S_{365}\right)$ do not change in two consecutive iterations.

The solution determined by functional equation (11) is a feedback solution and can be denoted by $R_t^{opt}\left(S_{t-1},Q_t\right)$. This solution is optimal for the values of the parameters used in (9). If the values of the parameters are changed, the solution of the optimization problem will probably change. The problem is to find the values of the parameters α_i, $i=1,2,...,5$, that give the best solution, i.e. the solution that best satisfies the objectives. This problem can be solved by simulating the operation of the reservoir over a period of N years with the historical inflow data and the operating policy $R_t^{opt}\left(S_{t-1},Q_t\right)$, $t=1,2,...,365$, determined by (11). The results of the simulation can be used afterwards to determine the probability of flooding, the number of days in N years of simulation that the reservoir level has been lower than $s_t^{\min}$, the reservoir discharge smaller than $r_t^{\min}$ and the generation less than D_t. If the results are not acceptable, the parameter values must be changed and the problem solved again.

4 Inflow model

Functional equation (11) supposes that Q_t, the inflow on day t, is an independent random variable. In reality, Q_t is generally not independent but

correlated to the inflows of the preceding days. This is called persistence. It exists when there are periods of several consecutive days of high flows and low flows in the historical record. Long periods of high flow occur during the spring freshet and when it rains heavily during many consecutive days. Long periods of low flow often occur in the summer when it does not rain for many consecutive days. These long periods of high flow and low flow must absolutely be taken into account by the optimization model because the risks of not satisfying objectives b and c are higher in those periods. The correlation between the inflow of day t and that of days $t-1$, $t-2,...,$ $t-p$ will be taken into account if functional equation (11) is replaced by the following:

$$F_t\left(S_{t-1}, Q_{t-1}, ..., Q_{t-p}\right) = \underset{Q_t \left| Q_{t-1}, ..., Q_{t-p}\right.}{E} \left\{ \max_{R_t} \begin{bmatrix} G\left(S_{t-1}, S_t, R_t\right) - L_t\left(S_t, R_t\right) + \\ F_{t+1}\left(S_t, Q_t, ..., Q_{t+1-p}\right) \end{bmatrix} \right\} \tag{12}$$

Equation (11) uses $\Gamma_t\left(Q_t\right)$, the probability distribution of Q_t, to solve the problem. Equation (12) uses $\Gamma_t\left(Q_t \left| Q_{t-1}, Q_{t-2}, ..., Q_{t-p}\right.\right)$, the conditional probability distribution of Q_t, to solve the same problem. The conditional probabilities can be determined from the historical data or by a mathematical model of the inflows. The most commonly used models are the autoregressive models (AR), the autoregressive-moving-average models (ARMA) and the disaggregation models (Salas et al. [14]).

The biggest problem with solving equation (12) is that the computer time and memory space increase exponentially with the number of state variables. Bellman [15] called this “the curse of dimensionality”. As a result, functional equations with more than four state variables cannot be solved in a reasonable time, which means that (12) cannot be solved when $p > 3$. Since the daily inflows are often correlated with those of the past 5 to 10 days, not all reservoir management problems can be solved with equation (12).

Turgeon [16,17] has shown that the past p inflows in equation (12) can be represented by a single state variable when the inflows can be correctly represented by a AR or ARMA model. Let us suppose that they can be represented by the following AR model:

$$Q_t = \phi_{0,t} + \phi_{1,t} Q_{t-1} + \phi_{2,t} Q_{t-2} + ... + \phi_{p,t} Q_{t-p} + \varepsilon_t \tag{13}$$

where $\phi_{0,t}$, $\phi_{1,t}$,.., $\phi_{p,t}$ are parameters and ε_t is the error term. Now if we set:

$$H_t = \phi_{1,t} Q_{t-1} + \phi_{2,t} Q_{t-2} + ... + \phi_{p,t} Q_{t-p} \tag{14}$$

equation (13) can be rewritten as:

$$Q_t = \phi_{0,t} + H_t + \varepsilon_t \tag{15}$$

According to (15), Q_t is only a function of H_t. Functional equation (12) can therefore be rewritten as:

$$F_t\left(S_t, H_t\right) = \underset{Q_t|H_t}{E}\left\{\max_{R_t}\begin{bmatrix} G\left(S_{t-1}, S_t, R_t\right) - L_t\left(S_t, R_t\right) + \\ \underset{H_{t+1}|H_t, Q_t}{E}\left[F_{t+1}\left(S_{t+1}, H_{t+1}\right)\right] \end{bmatrix}\right\} \tag{16}$$

Turgeon [16] showed that the error committed using (16) instead of (12) to determine the optimal operating policy of the reservoir is very small.

5 Inflow scenarios

Characteristics of the inflows, like the mean, variance, skewness and persistence, are better taken into account when the optimization problem is solved with the actual historical inflow data than with the probability distributions of these data. Turgeon [18,19] showed how to use the historical inflow data to determine warning curves for the reservoir. A warning curve gives the level above or below which, the reservoir must be at each point in time to satisfy a probabilistic constraint. For instance, the lowest warning curve in Figure 1 could represent the trajectory above which the reservoir must be maintained to satisfy objective b with the desired probability. The highest warning curve could correspond to the trajectory below which the reservoir must be kept to satisfy objective c. Objectives b and c would therefore be met if the reservoir level would be maintained inside the corridor between the two warning curves.

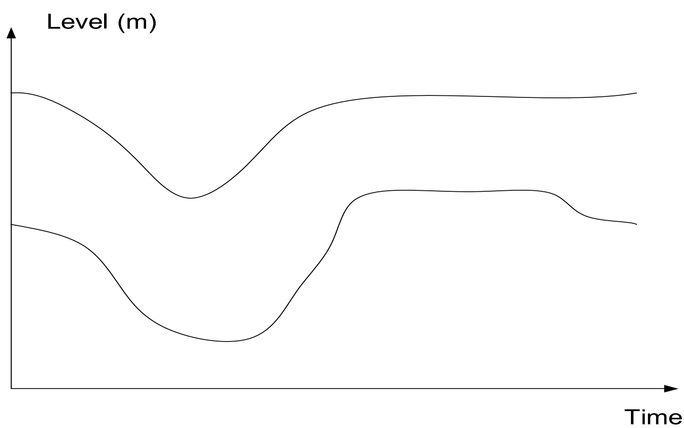


Figure 1: Warning curves.

With the warning curves, the optimization problem becomes:

$$\text{maximize} \sum_{t=1}^{365}\left[G\left(S_{t-1},S_t,R_t\right)-\alpha_5\cdot\Upsilon_t^2\right]+\Phi\left(S_{365}\right) \tag{17}$$

subject to:

$$S_t = S_{t-1} + Q_t - R_t \tag{18}$$

$$G\left(S_{t-1},S_t,R_t\right)+\Upsilon_t \geq D_t \tag{19}$$

$$S_t^{low} \leq S_t \leq S_t^{high} \tag{20}$$

$$R_t, \Upsilon_t \geq 0 \tag{21}$$

where S_t^{low} and S_t^{high} represent the warning levels, and Υ_t a dummy variable whose value is positive when demand, D_t, is not satisfied. If demand is too often not satisfied, the corridor in Figure 1 should be widened by increasing the probabilities of not satisfying objective b and/or c.

The method presented in this section gives good results when the warning curves are determined with a large set of inflow scenarios. Turgeon [18] used 251 synthetic inflow scenarios to build the warning curves. Since the historical data rarely have more than 50 scenarios, synthetic data must be generated before using this method.

6 Conclusion

This paper presents three different mathematical formulations of the multi-objective reservoir management problem. The first, which corresponds to equations (1)-(5), may not give a feasible solution. The second formulation, given by equations (6)-(9), always finds a solution, but this solution is a function of five parameters. The problem is to adjust the five parameters so as to meet the three objectives. The third formulation is the best because it better takes into account the characteristics of the inflows and has only three parameters to adjust: the probabilities of not satisfying objectives *b* and *c*, and the value of α_5.

References

[1] Eschenbach, E.A., Magee, T., Zagona, E., Goranflo, M. & Shane, R., *Goal programming decision support system for multiobjective oeration of*

reservoir systems, Journal of Water Resources Planning and Management, 127(2), 108-120, 2001.
[2] Yeh, W. W-G., *Reservoir management and operation models: a state-of-the-art review,* Water Resources Research, 21(12), 1797-1818, 1985.
[3] Wurbs, R.S., *Reservoir-system simulation and optimization models,* Journal of Water Resources Planning and Management, 119(4), 455-472, 1993.
[4] Labadie, J.W., *Reservoir system optimization models,* Water Resources Update Journal, 107, 83-109, 1998.
[5] Labadie, J.W., *Optimal operation of multireservoir systems: state-of-the-art review,* Journal of Water Resources Planning and Management, 130(2), 93-111, 2004.
[6] Cohon, J.L. & Marks, D.H., *A review and evaluation of multiobjective programing techniques,* 11(2), 208-220, 1975.
[7] Charnes, A. & Cooper, W.W., *Management models and industrial applications of linear programming,* vol. 1, John Wiley, New York, 1961.
[8] Haimes, Y.Y. & Hall, W.A., *Multiobjectives in water resources systems analysis: The surrogate worth trade off method,* Water Resources Research, 10(4), 615-624, 1974.
[9] Banayoun, R.J., de Montgolfier, J., Tergny, J. & Laritchev, O., *Linear programming with multiple objective functions: Step method (Stem),* Mathematical Programming, 1(3), 366-375, 1971.
[10] Sigvaldason, O.T., *A simulation model for operating a multipurpose multireservoir system,* Water Resources Research, 12(2), 263-278, 1976.
[11] Goulter, I.C. & Castensson, R., *Multiobjective allocation of water shortage in the Svarta River, Sweden,* Water Resources Bulletin, 24(4), 761-773, 1988.
[12] Dias, N.L.C., Pereira, M.V.F. & Kelman, J., *Optimization of flood control and power generation requirements in a multi-purpose reservoir,* IFAC Symposium on Planning and Operation of Electric Energy Systems, Rio de Janeiro, 101-104, 1985.
[13] Wang, Y.C., Yoshitani, J. & Fukami, K., *Stochastic multiobjective optimization of reservoirs in parallel,* Hydrological Processes, 19, 3551-3567, 2005.
[14] Salas, J.D., Delleur, J.W., Yevjevich, V. & Lane, W.L., *Applied modeling of hydrologic time series,* Water Resources Publications, Littleton, Colorado, 1985.
[15] Bellman, R., *Dynamic programming,* Princeton University Press, Princeton, N.J., 1957.
[16] Turgeon, A., *Solving a stochastic reservoir management problem with multi-lag autocorrelated inflows,* Water Resources Research, 41(12), W12414, 2005.
[17] Turgeon, A., *Solving reservoir management problems with serially correlated inflows,* River Basin Management III, C.A. Brebbia and J.S. Antunes do Carmo (eds.), WIT Press, Southampton, UK, 247-256, 2005.

[18] Turgeon, A., *Daily operation of reservoir subject to yearly probabilistic constraints,* Journal of Water Resources Planning and Management, 131(5), 342-350, 2005.

[19] Turgeon, A., *Optimal daily operation of reservoirs subject to probabilistic flood constraints,* River Basin Management II, C.A. Brebbia (ed.), WIT Press, Southampton, UK, 153-162, 2003.

A daily hydrological system management model that takes meteorological forecast errors into account

S. Krau[1], M. Latraverse[2], D. Tremblay[2] & A. Turgeon[1]
[1] *École Polytechnique de Montréal, Canada*
[2] *Hydro-Québec, Canada*

Abstract

Meteorological forecasts for the next few days are usually not detailed enough to determine the exact probability distribution of the forecast precipitation, and, hence, the probability distribution of the forecast inflow into the hydrological system. However, the hydrological system manager is responsible for assessing the risk of flooding and violating constraints when operating the reservoirs. This paper deals with the problems involved in modeling forecast errors and of determining an operating policy that takes such errors into account. This is accomplished by first building an inflow scenario tree that takes forecast errors into account and then by solving the reservoir management problem with that tree. This paper describes how the tree was built.
Keywords: meteorological, forecast, hydrological, management, risk, flooding, daily, stochastic, optimisation.

1 Introduction

The management of a hydrological system often implies the pursuit of various divergent objectives. There is a need to optimize power plant generation but also to regulate the flow of rivers, ensure a sufficient supply of drinking water, comply with vacationing-related constraints and ensure the safety of waterside communities as well as that of the structures. Management models most often take these different objectives into account through flexible constraints, i.e. constraints for which violation is tolerated, while minimizing the overall costs incurred by such potential violations. Daily management can then be reduced to a stochastic optimization problem related to a single objective. It consists in

www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060561

determining for the current day the water released from the reservoirs which maximize the hope of future profits, with the latter defined as the difference between the value of the hydroelectric generation and the total cost of violating the constraints.

The management of a hydrological system becomes especially critical during periods of high runoff. Several constraints run the risk of being violated and operating margins become very tight. In these conditions, run-of-river systems are the most vulnerable due to their low storage capacity and the inhabited areas they cross. An adequate representation of the uncertainty of water inflows within a management model thus becomes critical.

The uncertainty of water inflows stems from errors in the calibration of the hydrological model parameters, uncertainty regarding the water conditions in catchment areas, and the inaccuracy of weather forecasts. Weather forecasts are generally based on a deterministic meteorological model that only proposes one precipitation and temperature scenario over a ten-day time frame. Obviously, such forecasts are not perfect and it would be useful to know its uncertainty in order to soundly manage the hydrological system.

We are proposing to formulate the daily management problem using a multistage stochastic mathematical program with recourse within which the uncertainty of the deterministic meteorologist forecast is represented. We will present the perturbation methods used to transform a deterministic meteorological forecast into a probabilistic meteorological forecast and build a water inflow scenario tree.

2 General description of the management problem

The hydrological system management problem has received substantial treatment in scientific literature. Starting with a general mathematical formulation, Labadie [8] shows that a water inflow management model can be written as a non-linear, non-convex and possibly large-scale stochastic optimization problem. The author presents the state of the art for the various approaches used in practice to solve the problem. We have chosen the linear programming approach, one that is commonly used (Loucks [11]) since it can serve to easily model the constraints that temporally link decisions, such as the constraints pertaining to the flow times between reservoirs or those associated with flow variation boundaries. This approach benefits in practice from the sturdiness and power of commercial linear programming software.

A linear management model is obtained by approximating separable non-linear functions from non-separable ones (generation functions) in the vicinity of a point given by non-linear functions in a piecemeal manner. The problem of finding the water release from the current day t which maximizes the hope of future profits is then formulated as a multistage linear stochastic mathematical program with recourse (Birge *et al.* [2]). This program can be reformulated into a deterministic linear mathematical program known as a *deterministic equivalent* (Birge *et al.* [2]). All that is needed is to know the discrete distributions of the daily inflows, and then build the resulting decision tree.

Note that the size of the constraint matrix increases linearly with the number of nodes in the inflow tree. Should the tree become too large in size, modeling compromises and/or the use of decomposition methods (Birge [3]) would then have to be used to solve the program.

3 Characterization of the meteorological forecast error

Uncertainty in meteorological forecasting is predominant compared to the other sources of uncertainty related to a hydrological forecasting model (Krzysztofowicz [7]). Coulibaly *et al.* [4] show that, even for fairly large errors, meteorological forecasting significantly improves the precision of inflow forecasts over a 7-day time frame. Numerous methods have been developed to characterize the uncertainty related to inflow forecasting (Tamea *et al.* [13]; Collischonn *et al.* [4]; Lefevre [10]). Maskey *et al.* [12] proposed an approach used to represent the uncertainty of meteorological forecasting within a deterministic hydrological model. To our knowledge, there is no management model that explicitly integrates uncertainty in meteorological forecasting.

The deterministic meteorological forecasting that we are considering is provided to Hydro-Québec by Environment Canada and takes the form of a set of grids at six-hour intervals over a nine-day time frame, for a total of 36 grids. The first 48 hours are generated by the regional climate model (15-km increments) and are available in six-hour blocks, while days 3 to 9 are generated by the global model (110-km increments) and are available in 12-hour blocks. Hydro-Québec forecasters increase the value over the first 24 hours and merge the regional and global models, while producing 6-hour blocks for the entire period, followed by an interpolation at every 10 km. The 36 proposed precipitation grids are then merged into 9 daily forecasting grids. The daily forecasting grids are set, i.e. they cover the same geographic region. At each of the grid points (spaced 10 km apart) is found a total amount of forecast precipitation along with the daily minimum and maximum forecast temperatures.

Given the current day t and a deterministic meteorological forecast issued over T days, there are T grids of daily meteorological forecasts issued. Let $\hat{\Pi}_{(t+T)|t}$, $\hat{z}_{(t+k)|t}$, $\hat{T}^{\min}_{i,j,(t+k)|t}$, $\hat{T}^{\max}_{i,j,(t+k)|t}$ be respectively the daily forecast grid for day t+T, issued on day t, the amount of forecast precipitation at point (i,j) for day (t+k) issued on day t, the minimum and maximum forecast temperature at point (i,j) for day (t+k) issued on day t. Let (x_i, x_j), I, J, be respectively latitude and longitude coordinates of grid points $(i,j) \in I \times J$, the number of indexes of longitudinal and latitudinal grid points.

We are proposing a characterization of forecasting errors for the amount of forecast precipitation and based on the history of variations between the forecasting grids and the observation grids. The observation grids are built based on an extrapolation of the values recorded at the observation stations located in the hydrological system catchments. Given eqn. (2) that links the amount of forecast precipitation to the precipitation forecast issued on day t for the day (t+k), to grid point (i,j).

$$z_{i,j,(t+k)} = \varepsilon_{i,j,(t+k)|t}(\hat{z}_{i,j,(t+k)}); \forall t; \forall (i,j) \in (I \times J); k = 0..T-1 \quad (1)$$

where:
$z_{i,j,(t+k)}$ is the random variable representing the quantity of precipitation observed at point (i,j), for day $(t+k)$ and $\varepsilon_{i,j,(t+k)|t}(x)$ is the random function representing the uncertainty of a deterministic forecast issued on day t, at point (i,j), for day $(t+k)$. Functions $\varepsilon_{i,j,(t+k)|t}(x)$ are spatially and temporally correlated and the full determination of their parameters is complex, even impossible, to realize given the field's spatial and temporal range.

The forecasting uncertainty is made up of three types of uncertainties: (1) uncertainty regarding the total amount of forecast precipitation over the forecasting time frame, (2) uncertainty regarding the temporal distribution of the precipitation, (3) uncertainty regarding the spatial distribution of the precipitation. We are representing the first two uncertainties by a perturbation (random data transformation) that we will call *temporal perturbation of the forecast precipitation*. We will call the uncertainty related to the spatial distribution of the forecast precipitation over the forecasting time frame *spatial perturbation of the forecast precipitation*. We assume that these two perturbations are independent. They are variations of the perturbations proposed by Latraverse [9] regarding the same forecasting grids over a two-day time frame. The perturbations that we are proposing can be used to transform deterministic meteorological forecasting into probabilistic forecasting in the form of a scenario tree.

4 Temporal perturbation of the forecast precipitation

Temporal perturbation consists in reasoning not on the amount of forecast precipitation at the grid points but on the total amount of precipitation on the grids. Latraverse [9] shows forecasting errors on these quantities are independent from one day to the next. The random variable q_{t+k}, associated with the total amount of precipitation observed on the grid for day ($t+k$), verifies eqn. (3):

$$q_{t+k} = f(\hat{q}_{(t+k)|t});\ k=0..T\text{-}1;\ \forall\ t. \quad (2)$$

with $\hat{q}_{(t+k)|t} = \sum_{i\in I}\sum_{j\in J}\hat{z}_{i,j,(t+k)|t}$, $q_{(t+k)|t} = \sum_{i\in I}\sum_{j\in J} z_{i,j,(t+k)|t}$ for $k=0...T\text{-}1,\ \forall\ t$, and $f(\hat{q}_{(t+k)|t})$ a random variable. We propose to discretize the function $f(x)$ using discrete random variable $\Psi_1^V(\hat{q}_{(t+k)|t})$. Given a history of forecasting grids and observation grids and the following sets (see eqn. (4)):

$$\hat{A}_k = \left\{\hat{q}_{(t+k)|t}; t = \underline{t_h}..\overline{t_h} - k\right\},\ A_k = \left\{q_{(t+k)|t}; t = \underline{t_h}..\overline{t_h} - k\right\};\ k=0..T\text{-}1 \quad (3)$$

with $\overline{t_h}$, $\underline{t_h}$ and K , be respectively the history starting date, the history ending date and the number of days in typical forecasting time frame. Set $\hat{A}_k$ is made up of the amount of forecast precipitation based on a forecast history and associated with kth day. Set A_k is made up of observed amounts of precipitation corresponding to elements of $\hat{A}_k$. Given $\hat{E}_k^v$ respectively E_k^v, the subset of set $\hat{A}_k$ made up of v elements that are closest to $\hat{q}_{(t+k)|t}$, respectively the subset of set A_k made up of the observed elements associated with elements of $\hat{E}_k^v$. Given an arbitrary division of E_k^v (see eqn. (4)) into N_k classes ($k = 0..T-1$):

$$E_k^v = C_k^1 \cup .. \cup C_k^{N_k};\ C_k^{i_k} \cap C_k^{j_k} = \varnothing;\ \forall i,j;\ i_k = 1..N_k;\ j_k = 1..N_k. \tag{4}$$

Let $c_k^{i_k}$ be the centroid of the class $C_k^{i_k}$ for $i_k = 1..N_k$. Random variable $\Psi_1^v(\hat{q}_{t+k})$ for a given t is defined by N_k amounts of precipitation and eqn. (5):

$$\Pr ob\left[\Psi_1^v(\hat{q}_{t+k}) = c_k^{i_k}\right] = P_1(k,i_k) = \frac{\left|C_k^i\right|}{\left|E_k^v\right|};\ i_k = 1..N_k; k = 0..T-1 \tag{5}$$

Given a deterministic meteorological forecast of the total amount of daily precipitations $\left\{\hat{q}_{t|t}, \hat{q}_{(t+1)|t}, .., \hat{q}_{(t+T-1)|t}\right\}$ issued at the start of current day t. Based on the distribution of discrete random variables $(\Psi_1^v(\hat{q}_{t+k}); k = 0..T-1)$, an event tree could be built that represents the scenarios of the amounts of probable precipitations with as many stages as days in the forecasting time frame, the tree becomes enormous in size, even if the forecasting time frame is limited to only about a dozen days. We resort to the use of a scenario tree reduction algorithm (Heitsch and Romisch [6]) as to only keep a limited number of scenarios forming a reduced tree that is statistically close to the original tree. The following notation is used for the remainder of the article: $N_k^{'}$, i_k^- and $d_k^{i_k}$ respectively the number of nodes in the reduced tree for stage k, the predecessor of i_k th node of stage k of the tree ($i_k = 1..N_k^{'}$, for $k = 0..T-1$) and the amount of precipitations associated with node i_k in the reduced tree ($k = 0..T-1$).

5 Spatial perturbation of the forecast precipitation

The uncertainty regarding the spatial distribution of precipitations is important for the characterization of forecasting errors in relation to the hydrological

system management problem. In fact, the risk of flooding or spillage in a hydrological system for a given amount of precipitation differs considerably depending on the distribution of precipitations in the catchment areas that make up the hydrological system under study.

We are assuming, for modeling purposes, that the spatial distribution of the points in the grid for all the deterministic forecasting grids remains invariant; only the geographic location of the forecasting grids changes. The spatial perturbation takes the form of a spatial translation of the deterministic meteorological forecasting grids. The parameters for these translations are calculated based on the forecasting errors on the location of the maximum precipitation point associated with the first forecasting grid. In reality, the forecasting grids cover a larger area than the catchments in the hydrological system, which allows, post-translation, to have full coverage of the hydrological system catchment areas through the translated grids.

The choice of this modeling resides in approximating the following meteorological phenomenon. Low-pressure systems generally follow the same direction. We are attempting in this modeling to capture the fact that the forecast low-pressure system can travel more to the north or more to the south, while keeping the original spatial distribution of the associated precipitation.

Let $(\bar{i},\bar{j})$, $(\hat{x}_{\bar{i}},\hat{y}_{\bar{j}})$, $(x_{\bar{i}},y_{\bar{j}})$ be respectively the indices for the maximum precipitation grid point ($\hat{z}_{\bar{i},\bar{j},t|t} = Max\left\{\hat{z}_{i,j,t|t}; i \in I, j \in J\right\}$) on the first forecasting grid, the forecast longitude and latitude coordinates of grid point $(\bar{i},\bar{j})$ and the random variable corresponding the coordinates observed at the grid point $(\bar{i},\bar{j})$.

The forecasting error equation on the location of the point of maximum precipitation is written $(x_{\bar{i}},y_{\bar{j}}) = g(\hat{x}_{\bar{i}},\hat{y}_{\bar{j}})$, with $g(\hat{x}_{\bar{i}},\hat{y}_{\bar{j}})$ a random variable.

We are proposing to discretize function $g(\hat{x}_{\bar{i}},\hat{y}_{\bar{j}})$ with discrete random variable $\Psi_2(\hat{x}_{\bar{i}},\hat{y}_{\bar{j}})$. The assumption is made that the forecasting errors longitudinally and latitudinally from the maximum point of precipitation are independent and follow normal laws.

Given R ($R \gg 1$) points $(e_{r,1},e_{r,2})$, generated from a forecast location from the maximum precipitation point and normal laws of errors on the latitude and longitude. Given A_l, $l = 1..L$, the different catchments of the hydrological system and $(u_l = (x_l, y_l),\ l = 1..L)$, the point that minimizes the aggregate of the distances between it and all the generated points part of catchment (A_l, $l = 1..L$). Given $u^* = (x^*, y^*)$ the point minimizing the aggregate of the distances between it and all the generated points that are not part of any of the catchments in the hydrological system. Eqns (7) and (8) then define the random variable $\Psi_2(\hat{x}_{\bar{i}},\hat{y}_{\bar{j}})$.

$$Prob\left[\Psi_2\left(\hat{x}_{\bar{i}},\hat{y}_{\bar{j}}\right)=(x_l,y_l)\right]=P_2(l)=\frac{\sum\limits_{r|(e_{r,1},e_{r,2})\in A_l} 1}{R};l=1..L \tag{6}$$

$$Prob\left[\Psi_2\left(\hat{x}_{\bar{i}},\hat{y}_{\bar{j}}\right)=(x^*,y^*)\right]=P_2(L+1)=1-\sum_{l=1}^{L}P_2(l) \tag{7}$$

The spatial perturbation corresponds to $L+1$ movements for all the forecasting grids. Each movement is defined by the translation which correlates the maximum precipitation point associated with the first forecasting grid, with one of the L+1 locations. The forecast maximum and minimum temperatures at the grid points are updated based on their altitudes in relation to those they had prior to the translations. For instance, a grid point associated with a given geographic altitude will see its forecast temperature decrease if, after translation, said grid point has a greater altitude.

6 Example

The *temporal perturbation of the forecast precipitation* allows a tree of probable precipitations to be obtained. The *spatial perturbation of the forecast precipitation* generates series of grids of translated forecasts compared to the grids in the initial deterministic forecast. The composition of the two perturbations allows a scenario tree to be obtained of probabilistic-forecast grids.

The following figure (fig. 1) depicts results obtained by the described perturbation method for a given day of the past. The hydrological system considered in this example is the Gatineau system, a runfall-runoff system in Quebec. We can see on each grid (in black line) the Gatineau hydrological system frontier. At the beginning of the 14th October of 2003, a daily meteorological forecast grid has been issued (top left corner in fig.1). The other grids have been generated by the perturbation method. The intensity of the precipitation is shown by a color gradation.

7 Conclusion

The integration of meteorological forecasting uncertainty in a daily hydrological system management model can substantially improve the safety of hydrological system operation, especially when the system's water conditions are near their limits. The precise characterization of the uncertainty of the forecasting grids is complex and even impossible to achieve given the vast spatial and temporal field that needs to be considered. Aggregations (perturbations) are then needed to decrease the problem's complexity while attempting to capture the variability of future probable meteorological conditions. The management model we are proposing is a piecemeal linear model within which the representation of meteorological forecasting uncertainty takes the form of a tree of inflow scenarios.

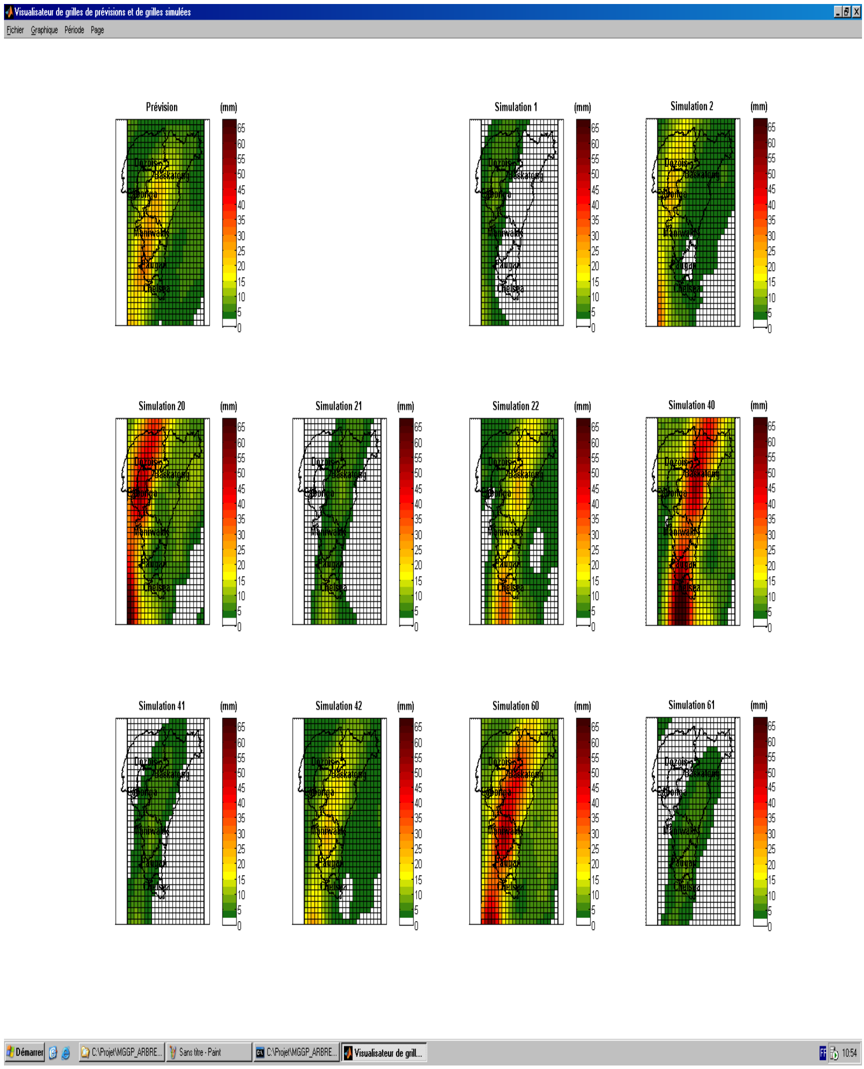


Figure 1: Example of grid perturbation method.

In the operational context, this management model is reset at the start of each day based on deterministic forecasting and must be solved fairly quickly so that decisions regarding reservoir water release are made on time. The transformation of the forecasting grid tree into an inflow scenario tree through the hydrological model is the most costly stage from the standpoint of computation time, which limits the number of meteorological scenarios that can be built for the tree. However, the parameters of the perturbation methods allow the meteorological scenarios that are to be represented in the forecast tree to be "chosen." This allows the risk of flooding and/or spillage associated with extreme meteorological scenarios to be quantified over the forecasting time frame.

In our work, we confined ourselves to characterizing the uncertainty of precipitation forecasts, disregarding inflows resulting from the snowmelt. This therefore limits us to testing the approach in the summer or fall, when temperatures have very little influence on the transformation of precipitation into inflow. The proposed management model that is combined with the Hydrotel hydrological model is in the process of being tested on the Gatineau River hydrological system based on a history of several years of forecasting errors. If the results prove promising, we will integrate the notion of temperature in order to make the model applicable to the spring freshet.

References

[1] Birge, J. R. and Louveaux, F., 1Springer Verlag, New York.1985.

[2] Birge, J. R., Dempster, M., Gassmann, H., Gunn, E., King, A., and Wallace, S., A standard input format for multiperiod stochastic linear programs. Committee on algorithms newsletter, Vol. 17, pp. 1-19.1987.

[3] Birge, J. R., Decomposition and partitioning methods for multi-stage stochastic linear programs. Operations Research, Vol. 33, pp. 989-1007.1985.

[4] Collischonn, W., Hass, R., Andreolli I., Tucci, C. E. M. Forecasting River Uruguay flow using rainfall forecasts from a regional weather-prediction model. Journal of Hydrology.2005.

[5] Coulibaly, P., Impact of meteorological predictions on real-time spring flow forecasting. Hydrological Processes, Vol.17, pp. 3791-3801.2003.

[6] Heitsch, H., Romisch, W., Scenario reduction algorithms in stochastic programming. Computational Optimization and Applications, Vol. 24, pp. 187-206.2003.

[7] Krzysztofowicz, R., Bayesian theory of probabilistic forecasting via deterministic hydrologic model. Water Resource Research, Vol. 35, pp. 2739-2750.1999.

[8] Labadie, J. W., Optimal Operation of Multireservoir Systems: State-of-the-Art Review. Journal of Water Resources Planning and Management, Vol. 130, pp. 93-111.2004.

[9] Latraverse, M., Prévision des apports à court et moyen terme et amélioration de la gestion des systèmes au fil de l'eau. Hydro-Quebec Poduction, Internal Report.2001.

[10] Lefevre, M., Modélisation des erreurs pour le système de prévision PREVIS. Canadian Journal of Civil Engineering, Vol. 31, pp. 892-897.2004.

[11] Loucks, D. P., Multiple-Reservoir Operation in North America. Water Science and Technology, Vol. 1, pp. 711-728.1981.

[12] Maskey, S., Guinot V., Price, R. K., Treatment of precipitation uncertainty in rainfall-runoff modelling: a fuzzy set approach. Advances in Water Resources. 2004.

[13] Tamea, S., Laio, F., Ridolfi, L., Probabilistic nonlinear prediction of river flows. Water Resources Research. Vol. 41, n.9, p. W09421.2005.

[14] Turgeon, A., Solving a stochastic reservoir management problem with multilag autocorrelated inflows. Water Resources Research. Vol. 41, n. 12, p. W12414.2005.

Section 8
Air

This page intentionally left blank

A forecast model to predict the next day's maximum hourly SO_2 in the site of Priolo (Siracusa)

U. Brunelli, V. Piazza & L. Pignato
Dipartimento di Ricerche Energetiche ed Ambientali,
Università di Palermo, Palermo, Italy

Abstract

The purpose of this research is to develop a pure predictive model to forecast the next day's maximum hourly SO_2 in the site of Priolo in two representative monitoring stations. We globally propose a Recurrent Neural Network (RNN) with an inherent dynamic memory to forecast the fluctuations of ground concentration of SO_2 pollutant. The model uses an available time series which was recorded in the industrial site of Priolo, on the east coast of Sicily, Italy for the period between 1st April 1998 and 12th December 2001. The inputs used to train the neural network are three, where two are adimensional variable, which are obtained by the meteorological data of the site and the third is the stability classes of Thomas. This study has some important implications for health warning systems environmental management in places with high pollution concentration.
Keywords: air pollution, neural network, forecasting.

1 Introduction

Air pollution has a negative impact on the environmental and public health when it occurs in the lower atmosphere. In this paper the results obtained by an Artificial Neural Network to forecast the daily maximum SO_2 ground concentration are presented. A neural network is able to treat information that is uncertain and incomplete like the human brain, and it has been utilised to model complex non-linear functional relationships between predictor variables. The

doi:10.2495/RAV060571

pollutant concentration genesis and its dynamics are characterised by highly non-linear processes which are only partially known so then as artificial Neural Network (ANN) appears to be a good approach for pollutant concentration forecasting.

By using neural networks the authors have obtained very interesting results, as shown in Figures 4 and 5. Tasadduq et al. [2] use an MLP neural network trained with the BackPropagation algorithm and a batch learning scheme for the prediction of hourly mean values of ambient temperature 24 h in advance in Jeddah, Saudi Arabia. Full year hourly values of ambient temperature are used to train the network. The performance was evaluated by the mean percentage error between the predicted and measured values for three different years, considering a time window of 50 hours. Elman neural networks have been successfully used in other forecasting applications and time series prediction. Luk et al. [3] evaluated three alternative ANN models for rainfall forecasting: a Multi-Layer Feed-Forward Neural Network, an Elman Neural Network and a Time Delay Neural Network (TDNN). The study eventuates that the three approaches have comparable performance as long as the complexity of the network is variable. In details the Elman network had the simplest structure, but was complex at the same time. Kostela et al. [4] compared the performance of the Multi-Layer Perceptron, FIR and Elman Neural Networks in four different time series analyses: the performance of the Elman neural networks was better or similar to the other neural networks. In the paper, it is also argued that the efficiency of the learning algorithm is more important than the neural model used. In the last year neural networks have became an alternative to conventional methods and they are going to become an important instrument to model the distributions of air pollution (Nagendra and Khare [5]). Viotti et al. [6], use a multi-layer perceptron neural network to forecast short and middle long-term concentrations levels for O_3, NO_x, NO_2, CO. Hooyberghs et al. [7], describe the development of a Multi-Layer Perceptron neural network to forecast the daily average PM10 concentrations in Belgian urban areas one day ahead.

In this paper the authors use a Recurrent Neural Network for prediction of the next day's maximum hourly SO_2. The time series was recorded between April 1st, 1998 and December 31st 2001 and refers to two monitoring stations of air pollution parameters (Melilli, Farodromo) with a lead-time of one hour. These data are kindly provided by CIPA (Consorzio Industriale Protezione Ambiente). The paper is organised as follows. Section 2 gives a theoretical description of the area, section 3 a processing of the data, section 4 the background on the ANN model, and sections 5 and 6 the structures of the Neural Network model used and the obtained experimental results.

2 Site description

The region of Priolo is represented in Figure 1.

We can see that the coastal strip occupied by the industrial settlements extends from S to N laying on a bay, and is limited by land not only on the Eastern side, but also on the Northern one (Capo Izzo with the small town of

Augusta). More importantly, there are relatively high elevations at W-SW (Climiti Mountains, about 400-500 m high), which probably interfere with winds and breezes to generate vertical patterns. As will be seen, they influence the classical partition of the wind directions near the ground, causing an apparent reduction of the frequency of the breezes coming from the sea when compared to those coming from the land. In the natural cavity surrounding the bay there are scattered villages, roads and single houses at different heights above the sea. The pollution sources on the coast can be roughly divided into extensive and fixed; the first ones cover, with a high degree of continuity, the whole costal strip, while the second ones consist mainly of stacks of relevant heights (frequently comprised between 100m and 200 m). This is usually considered a good design, because the maximum ground concentration on a flat plane is inversely proportional to the square of the source height; in this case however, the ground is all but plain and in stable conditions the plumes can impinge on the sides of the mountains at heights which are the same as those of the polluting sources. This results in very high pollutant concentrations in the local area.

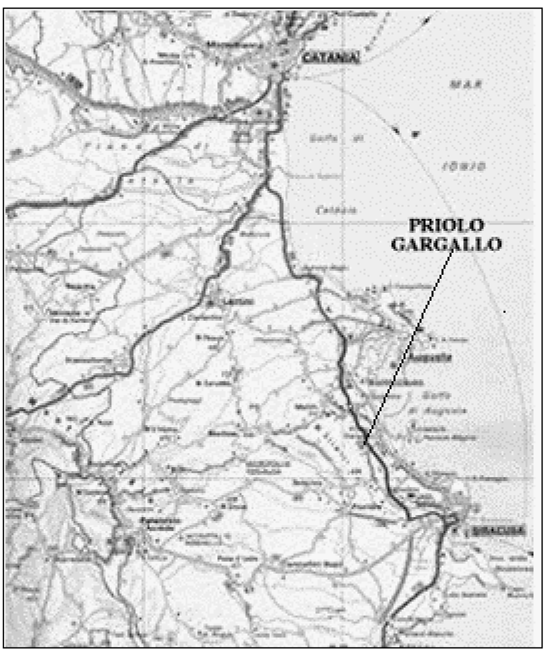


Figure 1: Description of Priolo site.

The facilities of interest are:

- A RASS (Radio Acoustic Sounding System), which is a remote sensing device that allows air temperature measurements of up to 1000 m of height with a spatial resolution of 20-30 m. Its working principle relies on the property of sound waves to propagate with different speeds depending on air temperature: a train of acoustic waves is then emitted along the vertical, letting the latter be

reflected by the travelling wave fronts of the former. The Doppler effect between emitted and reflected radio waves then enables the computation of the sound speed and hence the ambient temperature at the height where the reflection takes place
- A pole of 10m supporting an anemometer located at 10m and two thermometers at 2 and 10m. Both thermometers are shielded to avoid ground radiation and provide an accurate estimation of the temperature gradient at the ground level.
- A SODAR (Sound Detection and Ranging) is able to determine the vertical distribution of wind velocity and direction and to give information on the vertical thermal structure and the turbulence in the first five hundred meters of the atmosphere.
- 11 fixed stations for measuring ground level concentrations of pollutants such as (NHC, CH_4, SO_2, O_3, H_2S, NO, NO_2).

On the Priolo site the monitoring stations are indicated with numbers 1-11; in particular the two monitoring stations of our interest are number 8 for Melilli and number 5 for Farodromo. The CIPA laboratory is indicated by number 12 and it is located on a slightly elevated spot (30m above sea) about 1 Km inland (Figure 2).

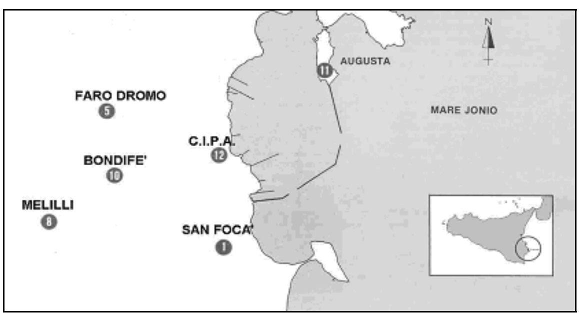


Figure 2: Map of monitoring stations.

3 Data processing

A neural network is a computational model based on data; it is clear that they are the principal component to obtain good results by the neural network.

The phases of their elaboration can be shared in:

1. collection and analysis of the data,
2. pre-processing following by the elaboration of the neural network,
3. post-processing if it is necessary to convert the output values of the neural network in the required format.

For our cases the hourly values used are:

- horizontal wind velocity (m/s)
- vertical wind velocity (m/s)

- base level of the layer of the atmospheric stability (m)
- gradient of the potential temperature (°C/m)
- the difference of the potential temperature of reference (°C)
- wind direction

The data set is recorded from April 1st 1998 to December 31st 2001 and is referred to sulphur dioxide monitoring stations with a lead-time of an hour. These data are submitted to the method of adimensional analysis to obtain same values which are independent of the characteristic of the industrial site which has been considered. By the application of this method the authors have obtained the following two adimensional numbers:

$$Pg = \frac{\gamma_\vartheta}{\Delta_\vartheta} H^3 \tag{1}$$

$$Pz = \frac{V_v}{V_o} \tag{2}$$

where V_v is the vertical wind velocity, V_o is the horizontal wind velocity, γ_θ is the gradient of the potential temperature, Δ_θ is the difference of the potential temperature of reference and H is the base level of the layer of the atmospheric stability. The third input used to train the neural network is the stability classes of Thomas [8]; they are adopted for a simple characterization of the stability of the superficial atmospheric layer. Thomas compared the values of the standard deviation of the vertical wind direction obtained with two different methods:

- measurements of the wind field, at a fixed height, with the SODAR;
- direct measurements of the standard deviation σ_φ, at the same height.

This last parameter was evaluated through the approximate relation:

$$\sigma_\varphi = arctg\left(\frac{\sigma_{vv}}{u}\right) \tag{3}$$

where σ_{vv} is the standard deviation of the vertical wind velocity and u is the wind in the prevailing direction.

In addition, each value was normalized in the range [-1, 1] using the following linear transformation:

$$X' = (X - V_m)/(V_{max} - V_{min}) \tag{4}$$

where X' is the new normalized value, X is the old value, V_{max} is the maximum of the considered data set, V_{min} is the minimum of the considered data set and V_m is the average value of the considered data set.

Finally the input pattern of the neural network is composed by three parameters: the adimensional numbers Pg and Pz and the stability classes of Thomas.

4 Artificial neural network

The Elman neural network (Elman, [1]) is also known as the partial recurrent network or simple recurrent network. In this network, the outputs of the hidden and output layer are allowed to feedback onto itself through a buffer layer, called the context layer. This feedback allows Elman networks to learn, recognize and generate temporal patterns, as well as spatial patterns. Every hidden neuron is connected to only one neuron of the context layer through a constant weight of value one. Hence, the context layer constitutes a kind of copy of the state of the hidden layer, one instant before. The number of context neurons is consequently the same as the number of hidden neurons. Optionally, every neuron of the output layer can be connected to only one neuron of a second context layer through a constant weight of value one (figure 3).

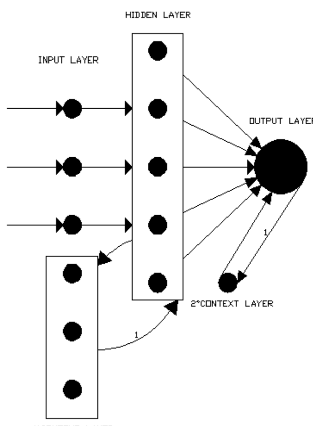


Figure 3: The Elman neural network.

Usually input, output, and context neurons have linear activation functions, while hidden neurons have the sigmoid activation function, but in this case the activation's function of output is the function RM.

$$a_j(t)=\begin{cases}0.85*a_j(t-1)+0.15*net_j(t)*(1-a_j(t-1))\xrightarrow{if} net_j(t)>0\\0.85*a_j(t-1)+0.15*net_j(t)*(1+a_j(t-1))\xrightarrow{if} net_j(t)\leq 0\end{cases}\tag{5}$$

and that for the hidden units is the function TanH:

$$a_j(t) = tanh(net_j(t) + \theta_j) \quad (6)$$

where $a_j(t)$ is the activation of unit j in step t, net_j is the input in unit j in step t and θ_j is the threshold (bias) of unit j. The training algorithm is the Resilient Back Propagation (RProp); it is a local adaptive scheme, performing fast and robust supervised batch learning in neural networks.

5 Network structure

The topology of neural network is a problem that depends on various factors. Whichever the type of ANN model employed, it is important to determine the appropriate network architecture in order to obtain the best results. Several artificial neural network topologies were implemented by changing the number of layers, and the number of the hidden and context units. Neural model simulations were performed using the Stuttgart Neural Network Simulator (SNNS) v. 4.1. The connection weights were initialized to zero-mean random values with adequate upper and lower bounds of (-1, 1). For the ANN the authors use the following values as input: two adimensional variable (Pg, Pz), which are obtained by the meteorological data of the site and the third is the stability classes of Thomas between April 1st, 1998 and December 31st, 2001, to define the optimum ANN structure using the common trial and error method. Different structures of each ANN have been tested with various different hidden nodes. It is found that fifteen hidden nodes are the optimum for both ANN, in this experiment. The primary aim of developing an ANN is to generalise the features of the processed time series. A popular technique to achieve generalisation, avoiding over fitting, is the early stopping method presented by Sarle [10]. According to this method, the generated data set was divided into two subsets: a training set and a test set. The whole training phase was stopped when the lowest error on the training set was reached. With more details, the training set is composed of 1310 values and the test set by 61 values (months of November and December 2001). In the conducted experimental trials, training epochs were set to 150 for each neural network model. To evaluate the model performance, the authors selected two parameters:

- MAE is defined by the following expressions:

$$MAE = \sum_{i=1}^{N} |O_i - P_i| / N \quad (7)$$

where O_i is the observed value at time i, P_i is the predicted value at time i and N is the total number of observations. For a perfect fit, $O_i = P_i$ and MAE = 0. So, the MAE index ranges from 0 to infinity, with 0 corresponding to the ideal condition, and in particular it permits one to compare the appropriateness of the networks used.

Linear Correlation Coefficient (r):

$$r = 1 - (\sum (O_i - P_i)^2 / \sum (P_i - P_m)^2) \tag{8}$$

where P_m is the average value of the observed values. The linear correlation indicates the degree to which two variables are linearly related.

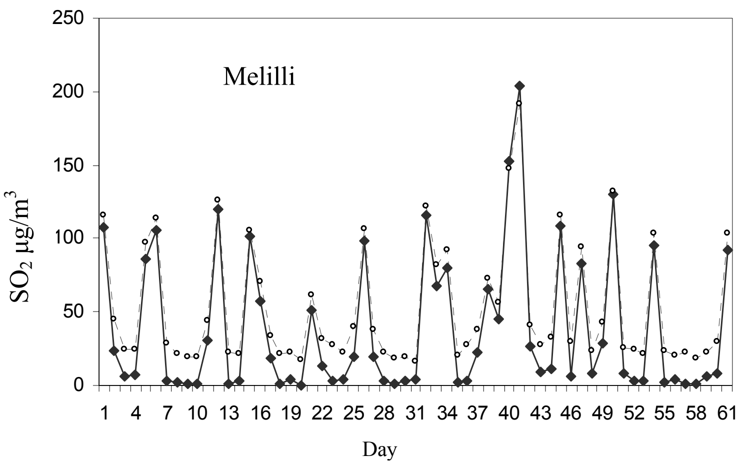


Figure 4: Results with Elman Neural Network, forecasted (white point) and measured SO_2 concentration (black point).

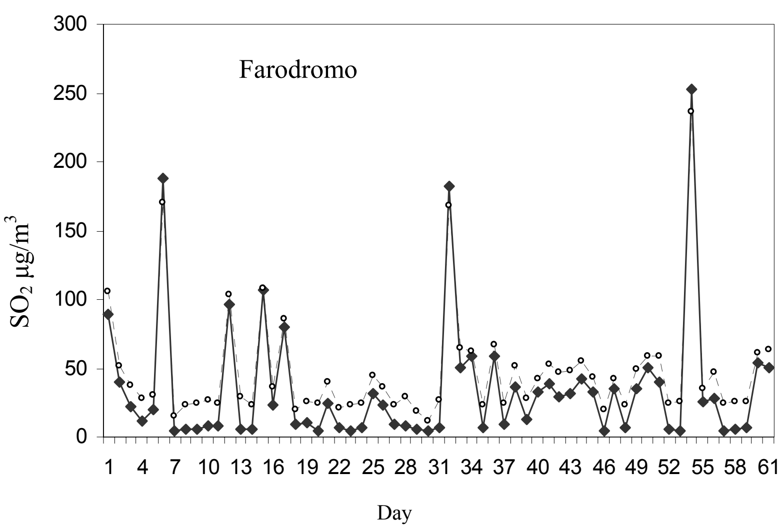


Figure 5: Results with Elman Neural Network, forecasted (white point) and measured SO_2 concentration (black point).

6 Experimental results and conclusions

The forecast of microclimatic phenomena represents a very complex activity with remarkable economic and sociable implications overall about the climatic changes which are happening. These processes, which are not linear at all and for whose cause the genesis and the dynamic are partially known, demand particular models of analysis and simulations to be forecasted with precision; in this situation the model based on artificial neural network represents a good solution for this problem.

The authors have used a Recurrent Neural Network (RNN) to forecast the values of hourly maximum daily sulphur dioxide, and they were tested in two stations of the Priolo region (Melilli and Farodromo). The set of test is constituted by the values of months of November and December 2001 and it is the same for all stations. The peculiarity of the model presented here is that it uses as input patterns two adimensional numbers (Pg, Pz) which are obtained from the meteorological characteristic of the studied place and the stability classes of Thomas. The quality of the developed model was assessed and compared on actual observed SO_2 concentration measurements of the test set and the results are showed in Figures 4 and 5. A statistic index has been used to provide a general indication of the relationship between observed and forecasted values. The selected performance measures included the mean absolute error (MAE) and the linear correlation coefficient (r). The general performance static for the two sites is summarized in Table 1. As we can see, the values of MAE, although being different numerically, are practically identical in terms of relative model performance; in particular, the values of r close to one show that the model gives good results.

Table 1: Statistic indexes.

Station	**r**	**MAE**
Melilli	0,85	16,25
Farodromo	0,84	7,58

Acknowledgements

The authors wish to thank the CIPA laboratory and the director Professor Giuseppe Zerbo for having supplied the experimental data used for this work.

References

[1] Elman, J. L. (1990), “Finding structure in time”. In Cognitive Science, n.14 pp. 179-211.

[2] Imran Tasadduq, Shafiqur Rehman, Khaled Bubshait (2002) “Application of neural networks for the prediction of hourly mean surface temperatures in Saudi Arabia” in Renewable Energy n.25 pp.545–554.

[3] K.C. Luk, J.E. Ball, A. Sharma (2000) "A study of optimal model lag and spatial inputs to artificial neural network for rainfall forecasting" Journal of Hydrology n.227 (2000) pp.56–65.
[4] Koskela T, Lehtokangas M, Saarinen J, Kaski K (1996), Time series prediction with Multilayer Perceptron, FIR and Elman neural networks Proceedings of the World Congress on Neural Networks. INNS Press, pp 491—496.
[5] S.M. Shiva Nagendra and Mukesh Khare. Artificial neural network approach for modelling nitrogen dioxide dispersion from vehicular exhaust emissions. Ecological Modelling, Corrected Proof, Available online 11 July 2005.
[6] P. Viotti, G. Liuti and P. Di Genova. Atmospheric urban pollution: applications of an artificial neural network (ANN) to the city of Perugia. Ecological Modelling, Volume 148, Issue 1, 1 February 2002, Pages 27-46.
[7] Jef Hooyberghs, Clemens Mensink, Gerwin Dumont, Frans Fierens and Olivier Brasseur. A neural network forecast for daily average PM_{10} concentrations in Belgium. Atmospheric Environment, Volume 39, Issue 18, June 2005, Pages 3279-3289.
[8] P. Thomas, "Stability classification by acoustic remote sensing", Atmospheric Research, 20, pp. 165-172, 1986.
[9] SNNS - Stuttgart Neural Network Simulator, url: http://www-ra.informatik.uni-tuebingen.de/SNNS/.
[10] W.S. Sarle, Stopped training and other remedies for over fitting, Proc. of the 27th Symp. On the Interface of Computing Science and Statistic, 352-360, (1995).

Sources of atmospheric pollutants in the North West province of South Africa: a case of the Rustenburg municipality

N. A. Kgabi[1], J. J. Pienaar[2] & M. Kulmala[3]
[1]*Department of Physics, North-West University, Mmabatho, South Africa*
[2]*School of Chemistry, North-West University, Potchefstroom, South Africa*
[3]*Department of Physical Sciences, University of Helsinki, Helsinki, Finland*

Abstract

A number of activities are known to generate substantial quantities of atmospheric pollutants in the form of uncontrolled emissions. Such sources include mineral extraction and stockpiling, landfill sites, materials handling operations and long term construction operations. In this study, the composition of air particulate matter was determined using Scanning Electron Microscopy coupled with Energy Dispersive Spectrometry (SEM/EDS) and the following elements were identified: Si, Fe, Al, Ca, Mg, K, Na, Ti, Cr, C, Cl, S, F, P, V, Pb, Ni and O. Correlation and regression analysis, together with Principal Component Analysis (PCA) were used to determine the sources of atmospheric pollutants; the sources were identified in order of decreasing abundance as: soil dust, industry, biomass burning, and traffic.
Keywords: particulate matter, SEM/EDS, sources, PCA, correlation and regression.

1 Introduction

A better understanding of the chemical constituents as well as the sources of ambient particles is fundamental in bridging the existing knowledge gap between the air quality and its health effects. Exposure to ambient particulate matter (PM) has been associated with a range of adverse health effects including: premature

www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060581

mortality, aggravation of existing respiratory conditions, changes to lung tissues and structures, and altered respiratory defence mechanisms, damages to the immune system, neurological damages, reproductive (i.e. reduced fertility) problems [1].

The major components of PM include sulphate, nitrate, ammonium, and hydrogen ions; trace elements (including toxic and transition metals); organic material; elemental carbon (or soot) and crustal components [2]. PM can also be classified as primary and secondary particles. Primary particles are the direct products from combustion processes while secondary particles are a result of the physical and chemical reactions in the atmosphere [3].

Nriagu [4] observed that metal smelting and fuel combustion are usually the source of non-crustal volatile metals in the atmosphere. The elements Zn and Mn can be used as tracers for smelting sources, K can be used for biomass burning, Cl for sea salt, sulphate for anthropogenic sources, and Al, Si, Ti for soil dust. Zn, excess Mn, Cu and Ni are an unambiguous signal of anthropogenic smelting [5].

The occurrence of Cr with Co or Ni indicates the anthropogenic origin of particulate aerosols while, the co-presence of both Pb and Zn, and moderate correlations (between Pb and V) as observed by Mishra *et al* [6] in Seoul, indicates that the contribution from fuel combustion may be important. The co-presence of high Al, Mg, Si, Ca mixed with K, Fe, Mn, Zn observed by Begum *et al* [7] also indicates contribution of metal smelting source to particulate aerosols.

Different approaches have been proposed to understand the source diversity of aerosol particles, including particle size and chemical composition [8, 9]. The objective of this study was to determine the source contributions to the chemical composition of particulate matter in the Rustenburg municipality of the North West province.

2 Experimental

Sampling of PM was done at Boshoek Primary School (BPS) located at latitude $25^0$30'15"S and longitude $27^0$5'45" E, within the Rustenburg municipality, which is situated at $25^0$39'00"S and $27^0$13'59"E, and is one of the biggest mineral producing districts in South Africa. The site is representative of well-defined environments, exposure situations or source activities like remote areas, urban background, traffic, and industry. The map showing the exact location of BPS is given in Figure 1.

The PM samples were collected onto Teflon-coated borosilicate fibreglass filters using the tapered element oscillating microbalance (TEOM). The ambient sample stream was allowed to passes through the PM_{10} inlet at the flow rate of 16.7 L/min, which was then isokinetically split into a 3L/min sample stream that is sent to the mass transducer and a 13.7L/min exhaust stream.

The scanning electron microscope for environmental samples (ESEM FEI QUANTA 200) coupled with the energy dispersive spectroscope (OXFID ENCA 200 EDS) was used for analysis of filters. The samples were analysed at high

vacuum, with a voltage of 15kV and the working distance of 10mm. The filters were mounted onto sample studs to ensure good electrical connection between the specimen and the microscope stage. The samples were not coated. The filters were scanned several (10) times (1 scan per 10 second) to ensure that the representative portion of the sample is covered. No extraction was performed.

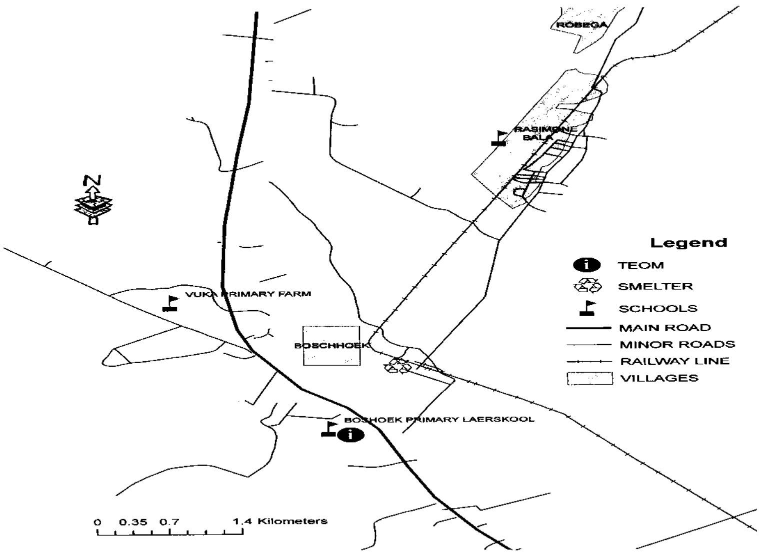


Figure 1: A map showing the Boshoek Primary School (BPS) sampling site in the Rustenburg municipality.

The statistical analysis methods, correlation and regression analysis, together with principal component analysis (PCA) were used to determine the sources of atmospheric pollutants. It is worth noting that a strong correlation between any two elements suggests that the two have a common source whereas the absence of correlation between any two elements may indicate that the two do not have a common source [10].

3 Discussion of results

3.1 PM10 levels

The results obtained from the TEOM levels in the form of monthly concentrations of the inhalable particulate matter at BPS are given in Figure 2. The PM10 levels for the spring season (August, September, and October) are relatively higher than the summer (November and December) levels at BPS. A

peak in PM10 levels was observed during the month of October (spring season), which agrees well with the fact that high wind speeds of 6.4 to 17.1 ms^{-1} reported by the South African Weather Services (SAWS) favour resuspension and transportation of atmospheric pollutants.

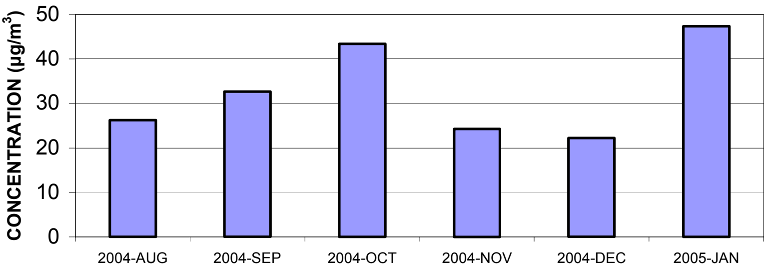


Figure 2: Monthly mass concentrations of PM10 at BPS.

The six-month average concentration at BPS is 33 μgm^{-3}, which is the same as the annual average levels obtained in 1999 at the Goodwood site in Cape Town, South Africa [11]. This agreement in results shows that the same level of pollution can result from urbanisation and industrialisation.

The levels observed from August to October 2004 (26 to 43 μgm^{-3}), and in January 2005 are within the average range observed by Querol *et al* [12], at different countries (Germany, Spain, Sweden, Austria, United Kingdom, and Netherlands) within the European Union, were the PM10 levels (annual mean) ranged from 28 to 42 μgm^{-3} at urban background, and from 37 to 53 μgm^{-3} at kerbside sites.

Characterisation of the fine airborne particulates is becoming increasingly important to regulators and researchers due to their potential impacts on human health, trans-national migration and influence on climate forcing and global warming [13, 14]. In the light of this, the composition of particulate matter and the estimation of source contributions within the Rustenburg area are discussed.

3.2 Composition of PM

Schneider [15], observed that an ordinary particle from traffic or fuel combustion source contain carbon at the core, with toxics and carcinogenic substances attached to their surfaces. This shows the importance of using the composition (elements and their oxides) of pollutants in source determination studies. The composition of PM10 obtained after filter analysis using SEM/EDS is given in Table 1. The levels of oxides of Al, Ca, Si, Na, Cl, S, O, and Fe observed in the samples suggest according to Chong *et al* [16] the presence of aluminosilicates with traces of sulfate or sulfite, and chloride. For most of the samples, the content of oxides of sulphur is the forth highest content.

The S concentration is important because it includes the sulphates and most toxic trace metals exist in the atmosphere in the form of sulphates and/or nitrates.

The relatively high S on the filters is according to Chong *et al* [16], indicative of the presence of sulphur dioxide, sulphuric acids, ammonium sulphate, ammonium bisulphate. The S levels are relatively stable for August, September, November and December. A peak is observed in October and an unusually low value is observed in January. There is a propotionality correlation between S and Pb, and between S and Ni from August to December. The correlation also applies to Cr but from August to November. Further studies need to be conducted to ascertain the cause of these correlations.

Table 1: Concentrations (in μgm^{-3}) of oxides of the elements identified by SEM/EDS analysis of filters.

Sample	PM10	Si	Fe	Al	Ca	Mg	K	Na	Ti	Cr	C	Cl	S	F	V	Ni	Pb	O
Aug 2004	26.2	1.55	0.1	0.58	0.05	1.55	1.55	0.24	0.02	0.03	4.27	-	0.24	4.66	0.01	-	0.03	14.23
Sep 2004	32.62	2.12	0.26	0.88	0.13	0.2	0.13	0.26	0.03	0.1	5.41	0.03	0.29	4.14	-	0.01	0.05	18.56
Oct 2004	43.38	2.47	0.87	0.95	0.26	0.43	0.21	0.48	0.03	0.3	6.68	-	1.69	3.86	-	0.03	0.21	24.99
Nov 2004	24.26	1.77	0.34	0.58	0.12	0.12	0.15	0.32	0.05	0.07	3.2	-	0.27	5.26	-	0.02	0.04	11.94
Dec 2004	22.21	1.18	0.07	0.6	0.09	0.11	0.07	0.14		0.04	4.69	-	0.22	-	0.01	-	0.06	14.93
Jan 2005	36.15	3.15	1.05	1.34	0.14	0.29	0.14	0.11	0.07	0.22	6.04	0.07	0.07	1.59	0.01	-	0.1	21.76
Mean	31.74	2.02	0.46	0.81	0.13	0.46	0.37	0.26	0.03	0.08	5.35	0.01	0.46	3.12	0.01	0.01	0.09	18.49
SD	8.31	0.27	0.18	0.12	0.02	0.19	0.20	0.05	0.01	0.06	0.78	0.01	0.22	0.84	0.00	0.00	0.02	2.53

The oxides of Ni were measured mainly in spring (September and October) and ranged from 0.01 to 0.03 μg.m^{-3}. The levels of V were found to be lowest (0.01 μg.m^{-3}) for the months August, December and January. Pb and Cr were identified throughout the sampling period, and their concentrations were measured as 0.03 to 0.21 μg.m^{-3} for Pb and 0.03 to 0.32 μg.m^{-3} for Cr.

The concentrations of the oxides of crustal elements identified in this study are similar to the levels obtained by Baron and Willeke [17] from four study sites in Atlanta, Rubidoux, Shenandoah National Park, and Meadview between May 1998 and May 1999 in the USA. The Ca concentrations at BPS ranged from 0.05 to 0.14 μgm^{-3} whereas the Ca levels at the four sites in the US ranged from 0.03 to 0.16μgm^{-3}. The 0.1 μgm^{-3} Fe measured at BPS in August 2004 compares with the 0.13 μgm^{-3} in August 1999 in Atlanta. The levels of the oxides of Cr determined ranged from 0.03 to 0.42 μgm^{-3} for this study. This shows that the limit of 1000 ng/m^3 (1 μgm^{-3}) set by NIOSH was not exceeded.

The SEM/EDS technique in this study revealed the presence of atmospheric particles of complex composition including S, Si, Al, Mg, Ca, Pb, Fe, Cr, Ni, V, and Pb among other elements. These particles can be harmful not only to human health, but also to the cultural heritage and the ecosystem as a whole. It is thus advisable for purposes of proper monitoring, to determine the sources of atmospheric pollutants using the elements identified. The world health organization (WHO) [18] also suggests that besides physical aspects such as particle number, size, or surface, the chemical composition of particles is likely to play a crucial role in regard the health implications of particulate matter.

Various health effects of PM, from less serious to very serious ones, are associated according to Sharma and Maloo [19] with its specific chemical and physical components.

3.3 Correlation and regression analysis

The following regression equations (eqns 1 – 2) show a clear relation relationship between the levels of PM10 and the toxic trace metals Cr, and Pb observed during this study.

$$PM10 = 22.6 + 62.7\ Cr \quad (1)$$

$$PM10 = 22.3 + 123\ Pb \quad (2)$$

These relations suggest an increase in the levels of toxic trace metals as the levels of PM10 increases. This imply the possibility of having high levels of trace metals within the PM, thus even when the standards set by the environmental protection agencies are not exceeded, the levels of the cancer-causing agents may still be hazardous to human health.

The relation between carbon and the trace metals is also important since the carbon determined can form part of the elemental and organic carbon. The oxides of carbon may imply the presence of Cr in the form chromium carbonate or even a gas. The carbon content can thus help give an indication of the main trace metals in the area. The r^2 values obtained for trace metals were low, except for C and Cr for which an r^2 value of 0.87 was obtained. The regression equation for C and Cr is given as:

$$C = 3.52 + 12.9\ Cr \quad (3)$$

According to Al-Momani [20], crustal material is the main source of Al. It is therefore logical to use Al as a tracer for crustal material source. The r^2 values obtained from regression analysis ranged from 0.008 to 0.01 for Al and V, 0.008 to 0.32 for Al and Pb, and 0.62 to 0.63 for Al and Cr, 0 and 0.014 for Al and Ni. These rule off the possibility of same source for Al and the trace metals of concern, except for Al and Cr. The high r^2 value for Al and Cr presents the possibility of Cr being from a crustal source. This however, does not rule out anthropogenic activities since they cause resuspension of particulate matter.

Metal smelting and fuel combustion are usually the source of non-crustal volatile metals in the atmosphere [4]. Held *et al* [5] suggests that the elements Zn and Mn can be used as tracers for smelting sources, K for biomass burning, Cl for sea salt, sulphate for anthropogenic sources, and Al, Si, Ti for soil dust. Zn, excess Mn, Cu and Ni are suggested as unambiguous signals of anthropogenic smelting. The correlation between K and Si in this study is $r^2 = 0.58$, and 0.55 between K and Fe, and 0.55 between K and Ca. This suggests that biomass burning is not the major source of the elements identified, since K and Si may be from the same source, and Si and Fe are tracers for soil dust source. The contribution of biomass burning may be observed as a result of some regional source.

In a similar study conducted in an industrial area of Taejon, Korea, Pb was also found to be strongly correlated with crustal elements (Ba, Sb, Co, Fe, etc.) and fuel combustion (V), strong correlations were also noticed by Chow *et al* [21] among Pb, Fe, Zn, and Ti from a study conducted in Los Angeles. The

correlations observed during this study are however, very small $r^2 = 0.008$ for Pb and Ti, $r^2 = 0.006$ for Pb and Fe. This indicates that Pb identified in this study cannot be of crustal origin, but is mainly due to fuel combustion (traffic).

The correlation and regressions discussed in this section suggest soil dust, traffic, and biomass burning as the sources of atmospheric pollutants in the Rustenburg municipality. It is important to note that the regression analysis could not quantify the contribution from different sources of particulate matter thus principal component analysis was used.

3.4 Principal Component Analysis (PCA)

Principal component analysis was also used to apportion the sources of atmospheric pollutants in the Rustenburg municipality. Figure 3 gives the PCA for BPS. Five factors that accounted for 98.4% of the PM were identified. The other 1.6% is labeled as unknown in the figure. This is the part that has eigenvalues less than one and thus could not be apportioned. The first factor, which accounted for 46.1% of the variance contained Si, Fe, Al, Cr, C and O. This is indicative of a soil dust source according to Held *et al* [5]. The occurrence of C in this factor could not be explained since elemental carbon (EC) was suggested by Ito *et al*, Ho *et al* and Salvador *et al* [22–24] as an emission from vehicles. The O is not used as a tracer or signature in this study since it may be linked with carbon or any other element identified thus implying any oxide in the PM10 analysed. The occurrence of Cr is not surprising since it may have been deposited on the earth crust as a result of the long-term ferrochrome mining activities in the area. There may be however, a justification for the co-occurrence of Cr, C and O, since Cr may be present in the form of carbonates.

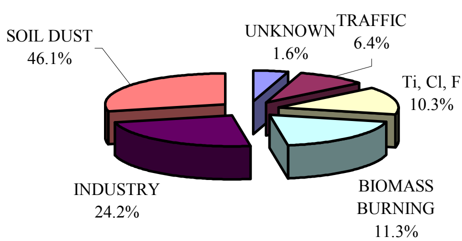


Figure 3: PCA for BPS.

Factor 2, which formed 24.2% and contained Na and Ni, was linked with industrial activities. This is mainly because Ni is linked with smelting, and other anthropogenic activities like oil-fuelled plant emissions [5, 26], while the presence of Na may be due to agricultural activities since it may be found in fertilisers. The third factor accounted for 11.3% of the variance with loadings of Mg and K. The factor was linked to biomass burning [5].

Factor 4, which formed 10.3% of the PM, contained Ti, Cl and F. The source of this factor could not be properly identified, mainly because of the presence of

high loadings of F, which may also be from the filters used. The Cl was identified by Salvador *et al* [24] as an additive to petrol in the form of ethylene dihalide ($C_2H_4Cl_2$), and Ti has been linked to soil dust by Held *et al* [5] and to building industrial activities by Zhu *et al* [25].

The fifth factor (6.4%), which was linked to traffic, contained Ca, S, V and Pb, with high loadings of V. The co-occurrence of Pb and V was also linked to fuel combustion in a study by Mishra *et al* [6]. The source of S was also identified by Watson and Chow [26] as oil-fuelled emissions.

The PCA method suggests the following sources in order of decreasing abundance: soil dust, industry, biomass burning, and traffic.

4 Conclusion

Identification of source contributions to the elemental components of particulate matter within the Rustenburg area was achieved in this study. Both the correlation and regression analysis, and the principal component analysis suggested soil dust, biomass burning and traffic (fuel combustion) as sources of atmospheric pollution. The PCA method proved to be the relevant tool for apportionment of sources atmospheric pollutants, since it yielded quantifiable information about the sources of pollutants within the area. Contributions to the sources of atmospheric pollutants are thus given in order decreasing abundance as soil dust, industry, biomass burning, and traffic.

Acknowledgements

The authors acknowledge Dr L Tiedt of the electron microscopy laboratory for the SEM/EDS analysis of filters. The financial support of the Finnish Environment Institute (SYKE), the National Research Foundation (NRF), and the North-West University is greatly acknowledged.

References

[1] Health Effects Of Wood Smoke, A Report Of Provincial Health Officer Of British Columbia (250/952-0876), 1993.
[2] Review and Assessment of Air Quality in Charnwood, Directorate of Corporate and Environmental Services, October 2000.
[3] Al-Masri, M.S, Al-Kharfan, K., Al-Shamali, K., Speciation of Pb, Cu and Zn determined by sequential extraction for identification of air pollution sources in Syria. *Atmospheric Environment* **40,** pp. 753–761, 2006.
[4] Nriagu, J.O., A global assessment of natural sources of atmospheric trace metals. *Nature* **338**, pp. 47–49, 1989.
[5] Held, G., Gore, B.J., Surridge, A.D., Tosen, G.R., Turner, C.R., Walmsley, R.D. (eds), *Air pollution and its impacts on the South African Highveld*, Environmental Scientific Association, Cleveland, 1996.
[6] Mishra, V.K., Kim, K., Hong, S., Lee, K., Aerosol composition and its sources at the King Sejong Station, Antarctic peninsula. *Atmospheric Environment* **38**, pp. 4069–4084, 2004.

[7] Begum, B.A., Kim, E., Biswas, S.K., Hopke, P.K., Investigation of sources of Atmospheric aerosol at urban and semi-urban areas in Bangladesh. *Atmospheric Environment* **38**, pp. 3025–3038, 2004.
[8] Rasmussen, P.E., Long-range atmospheric transport of trace metals: the need for geoscience perspectives. *Environmental Geology* **33 (2/3)**, pp. 96–108, 1998.
[9] Kim, K.H., Lee, J.H., Jang, M.S., Metals in airborne particulate matters in the Taijon industrial complex area of Korea. *Environmental Pollution* **118**, pp. 41–51, 2002.
[10] Singh, M., Jaques, P.A., Sioutas, C., Size distribution and diurnal characteristics of particle-bound metals in source and receptor sites of the Los Angeles Basin. *Atmospheric Environment* **36**, 1675–1689, 2002.
[11] City of Cape Town, State of Environment Report 1999.
[12] Querol, X., Alastuey, A., Ruiz, C.R., Artinano, B., Hansson, H.C., Harrison, R.M., Buringh, E., Brink, H.M., Lutz, M., Bruckmann, P., Straehl, P., Schneider, J., Speciation and origin of PM10 and PM2.5 in selected European cities. *Atmospheric Environment* **38**, pp. 6547–6555, 2004.
[13] Dockery, D.W., Pope, C.A., Xu, X., Spengler, J.D., Ware, J.H., Fay, M.E., Ferris, B.G., Speizer, F.E., *New England Journal of Medicine* **329**, p. 1753, 1993.
[14] IPCC, The Third Assessment Report of Working Group I of the Intergovernmental Panel on Climate Change: Technical Summary, Lead Authors, Albritton DL (USA), Meira Filho LG (Brazil), Shanghai, 17–20, January 2001.
[15] Schneider CG, Diesel and Health in America: The Lingering Threat; Clean Air Task Force, February 2005.
[16] Chong, N., Sivaramakrishna, K., Wells, M., Jones, K., Characterization of Inhalable particulate matter in ambient air by Scanning Electron Microscopy and Energy-Dispersive X-ray Analysis. *Electronic Journal of Environmental, Agricultural and Food Chemistry*, **1(3),** pp.145 – 1642002.
[17] Baron, P.A., Willeke, K., *Aerosol Measurement: Principles, Techniques, and Applications*, Second Edition, John Wiley and Sons, Inc., 2001.
[18] WHO, Health Aspects of Air Pollution with Particulate Matter, Ozone and Nitrogen Dioxide. World Health Organization, Bonn, 98, 2003.
[19] Sharma, M. & Maloo, S., Assessment of ambient air PM10 and PM2.5 and characterization of PM10 in the city of Kanpur, India. *Atmospheric Environment* **39**, pp. 6015–6026, 2005.
[20] Al-Momani, I.F., Trace elements in atmospheric precipitation at Northern Jordan measured by ICP-MS: acidity and possible sources, *Atmospheric Environment* **37**, pp. 4507–4515, 2003.
[21] Chow, J.C., Watson, J.G., Ashbaugh, L.L., Magliano, K.L., Similarities and differences in PM10 chemical source profiles for geological dust from the San Joaquin Valley, California. *Atmospheric Environment* **37**, pp. 1317–1340, 2003.

[22] Ito, K., Xue, N., Thurston, G., Spatial variation of PM2.5 chemical species and source apportioned mass concentrations in New York City. *Atmospheric Environment* **38**, pp. 5269 – 5282, 2004.
[23] Ho, K.F., Lee, S.C., Chow, J.C., Watson, J.G., Characterization of PM10 And PM2.5 source profiles for fugitive dust in Hong Kong. *Atmospheric Environment* **37**, pp. 1023 – 1032, 2003.
[24] Salvador, P., Artinãno, B., Alonso, D.G., Querol, X., Alastuey, A., Identification and characterisation of sources of PM10 in Madrid (Spain) by statistical methods. *Atmospheric Environment* **38**, PP. 435 – 447, 2004.
[25] Zhu, G., Wang, X., Zhang, R., Source identification of atmospheric aerosols collected at the suburb of Beijing by using PIXE. *Proc. Of the 10TH International Conference on Particle Induced X-ray Emission and its Analytical Applications PIXE 2004*, Portoroz, Slovenia, June 4-8, 2004.
[26] Watson, J.G. & Chow, J.C., Source characterisation of major emission sources in the Imperial and Mexicali Valleys along the US/Mexico border. *The Science of the total environment*, **276**, pp. 1 – 3, August 2001.

Air quality in Buenos Aires Province, Argentina

N. Quaranta[1], M. Caligaris[1], M. Unsen[1], G. Rodríguez[1], H. López[1], C. Giansiracusa[1] & P. Vázquez[2]
[1]*Universidad Tecnológica Nacional, Facultad Regional San Nicolás, Argentina*
[2]*Universidad Nacional de La Plata, Centro CINDECA, Argentina*

Abstract

Studies of different gaseous contaminants and particulate matter were made in several cities of the Buenos Aires Province in Argentina. These cities have noticeable differences in relation to the number of inhabitants, vehicular density, industrial activities, etc. They are La Plata, Bahía Blanca, Mar del Plata, Pergamino and San Nicolás, among other towns.

In each city, continuous monitoring equipment with electrochemical sensor technology was installed, in order to determine the NOx, CO, HC, SO_2 and O_3 concentration.

The particulate matter samples were picked up using high volume equipment and daily concentrations corresponding to total suspended solids (PM) were determined by a gravimetric method. The particles were characterized by optical microscopy, scanning electron microscopy (SEM) and electron diffraction analysis X-ray (EDAX). The results obtained showed a direct relationship between the type and quantity of the present particles and the general characteristics of the towns.

The present study is part of the research project "Study of the Air Quality in Buenos Aires Province", financially supported by the National Agency of Scientific and Technological Promotion, Argentina.
Keywords: air quality, air pollutants, particulate matter.

1 Introduction

Nowadays the increasing impact of human activities on the environment is one of the worries of the international community. These impacts are usually studied

WIT Transactions on Ecology and the Environment, Vol 99,
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060591

locally but it is becoming more frequent to take into account the global effects on human and environmental health.

Science is the key to identify the nature and scale of the impacts and to establish the need for regulatory decisions in the processes. Only if there is a continuous improvement of the knowledge on environmental pollution and its control possibilities will the impacts on the ecosystems be predicted adequately and the effects of the local, regional, national and international changes be evaluated properly.

The “Study of the Air Quality in Buenos Aires Province” project has been designed to carry out the study of the air quality in different cities in Buenos Aires Province, Argentina, in order to build a representation of pollution levels, to determine the influence of the different industries and vehicles contributions, the different geographies, meteorological conditions, etc. and to establish the applicability of different diffusion models that allow to predict future situations.

Buenos Aires is the second largest province in Argentina and congregates the third part of its population.

The most important result to achieve is firstly the characterization of air quality in Buenos Aires Province, secondly the local government awareness of the situation, and of the people of the different areas where the direct measures are taken, and lastly an intense diffusion of the environmental issue locally and provincially together with formation of human resources. All these achievements will be regarded as the beginning of an environmental monitoring and controlling process in Argentina, and the start of collective awareness leading to a modification of individual actions. This would bring about an optimal use of the natural resources and a minimization of contamination.

The objective of the present work is to study the air quality in Buenos Aires Province by means of the contaminants analysis in different points selected for their position and population density, and to establish concentration patterns of the studied contaminants. Not only densely populated areas and their industrial surroundings were studied, other areas with low and no density were also considered. The present situation of air contaminants was determined.

2 Experimental

Measurements were made in different cities of Buenos Aires Province. The pollutant levels established in Argentine norms for ambient air quality follow the values fixed by the Environmental Protection Agency of United States [1].

The determined gases were nitrogen oxides (NO_X), sulphur dioxide (SO_2), carbon monoxide (CO), hydrocarbons (HC), and ozone (O_3). These gases were measured by a monitoring system which included a remote station located in each city and a Central Station which received the data by means of cell phone system located in San Nicolás Regional Faculty. The measurements were performed by electrochemical sensors technology. The complete description of the monitoring system, including the sensors measurement ranges sensitivity to each pollutant was given in a previous work [2].

The presence of particles was also studied, determining concentrations and size distribution and analyzing their physicochemical characteristics. For this purpose, air samples were collected using high volume equipment. The mass of particle samples collected was determined by using a four decimal balance. First the clean dry filter is weighed, then after 24 hours, the filter with the collected sample is dried again and stored in a sterile Petri dish until it is weighed again.

The samples were observed by optical and scanning electron microscopies. The optical observations were made through an Axiotech Zeiss microscope with annexed Philips video camera. The scanning electron microscopy (SEM) analyses were carried out through a Philips 515 scanning electronic microscope with an EDAX X-ray detector.

The contaminants aggressiveness depends mainly on their concentration in the atmosphere. Consequently, the pollution level in a region does not depend solely on the sources but also on the dispersion in the air of the polluting agents. This is directly related with the meteorological conditions of the area, specially winds, rains, and temperature profiles of the surrounding atmosphere, parameters that were also determined in the present study, by using a weather station Davis Monitor II - USA. The data obtained by the National Meteorological Service were also consulted.

In this study the vehicular pollution was specifically analyzed. For this purpose the monitoring equipments were installed in the cities' centre, in order to determine the contaminants concentration. The number of inhabitants and the vehicular density were analyzed in order to infer the corresponding values for cities with similar characteristics, including geographical ones such as the proximity of the city to the sea, the presence of wind channels, etc.

The studied cities were La Plata, Pergamino, Chacabuco, Mar del Plata, Bahía Blanca, Trenque Lauquen, Olavarría, 9 de Julio, Saladillo and San Nicolás.

3 Results and discussion

The first stage of this work involved local studies of air quality in densely populated spots among which are the capital city of the province (La Plata) and other important cities, together with areas which represent zones of medium and low population density. The second stage was the zonal analysis of the results of those local studies. Maps showing the concentration of the different pollutants registered were made.

The evolution of the air quality was performed analyzing the CO concentration as the most representative one. The other primary contaminants analyzed, without taking into account ozone, showed similar evolutions. The ozone concentration depends not only on the amounts of the precursor gases present but on local and seasonal weather conditions.

3.1 Local studies of air quality

The general steps followed for the development of the project in each determined site, within the provincial territory were:

- ✓ Study of the zone to be analyzed.
- ✓ Determination of the data acquisition places.
- ✓ Assembly of the air quality monitoring stations.
- ✓ Collection of meteorological data.
- ✓ Data acquisition and transmission to the Central Station.
- ✓ Data treatment and results analysis.

The study of the zone implies the analysis of the geographical, historical meteorological conditions, characteristics of the running motor vehicles, social behavior of the local people, local industries, etc. The results of these studies allow to determine the most convenient sites for data acquisition and the most representative hours to take into account for the analysis depending on what is to be evaluated, whether the vehicular contributions, the industrial contributions, the seasons differences, etc.

Table 1: Some characteristics of the studied cities.

City	***Inhabitants Number***	***Vehicles Number***	***General characteristics***
Saladillo	29,600	12,409	Agriculture and cattle raising. Moderate winds.
Trenque Lauquen	40,181	17,309	Agriculture and cattle raising. Strong winds.
9 de Julio	45,998	20,039	Agriculture and cattle raising. Moderate winds.
Chacabuco	45,445	20,192	Agriculture and cattle raising. Moderate winds.
San Nicolás	137,867	38,723	Large number of industries: iron steel, energy production, etc. Moderate winds.
Pergamino	99,193	40,863	Agriculture and cattle raising. Moderate winds.
Olavarría	103,952	44,054	Agriculture and cattle raising, cement and mineral industries. Strong winds.
Bahía Blanca	284,776	119,646	Several petrochemical industries. Strong winds – Wind channels.
Mar del Plata	564,056	218,791	Commercial and tourist activities. Strong winds – Wind channels.
La Plata	574,369	285,430	Commercial and administrative activities. Moderate winds.

In Table 1 general characteristics of the analyzed cities are shown. Data in relation to vehicle numbers were obtained from Argentine Department of the Interior, National Direction of the Vehicle Proprietor Register and municipal information. Inhabitant numbers correspond to 2001 National Census.

The wind characteristics in Table 1 must be interpreted as follows:

Moderate classification implies winds from calm to moderate, velocity values less than 40 km/h.

Strong classification implies winds within 40 – 70 km/h.

This classification is taken as a simplification of the known Beaufort Scale. The presence of wind channels corresponds to those cities that have tall buildings and are located near the sea.

In previous works the air quality in San Nicolás [3, 4], and Mar del Plata [5, 6] and the particulate matter in San Nicolás, Vicente López and Coronel Suárez [7] were analyzed in detail. An important influence of different factors on the pollution levels was established in these studies. These factors are: local social and commercial activities, economic situation, relative location with regard to the sea, wind channels formation, vacation periods, seasonal differences, etc.

The commercial activities are an important item to be considered, as regards the standard contamination levels. The CO pollutant level established in Argentine air quality norm is 9 ppm which corresponds to the average of a period of eight hours, taken between 8:00 and 16:00. This period was established considering the situation of the big cities where the commercial and administrative activities are developed without interruption during this time. This situation doesn't occur in towns located in the inner of the province or in small towns, where the activities stop at 13:00 and begin again at 16:00. So, for comparative purposes, in these places the values of the CO concentrations have to be calculated otherwise. In the present work the CO concentration values were calculated taking into account the eight hours of maximum values determined in each analyzed day. In big cities, these two calculation ways give practically the same concentration values.

The most common behaviour observed in the majority of studied cities during working days was the presence of maximum pollution values in two hour bands a day: 12:00 – 13:00 and 20:00 – 21:00. As an example the CO levels determined in Mar del Plata city [6] are shown in Figure 1.

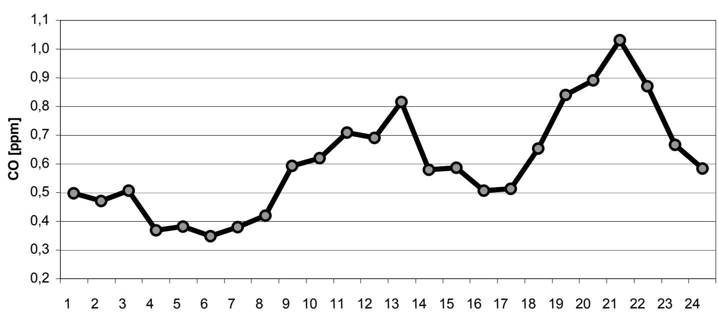


Figure 1: Hourly CO concentrations in Mar del Plata.

The cities which keep daily continue commercial and administrative activities show pollution levels that follow the behaviour presented in Figure 2, with more

stable values during all hours of daylight, which correspond to the studies made in La Plata city. La Plata is the province capital city and it has an important number of Government and Justice offices in the town centre besides the continuing commercial schedule.

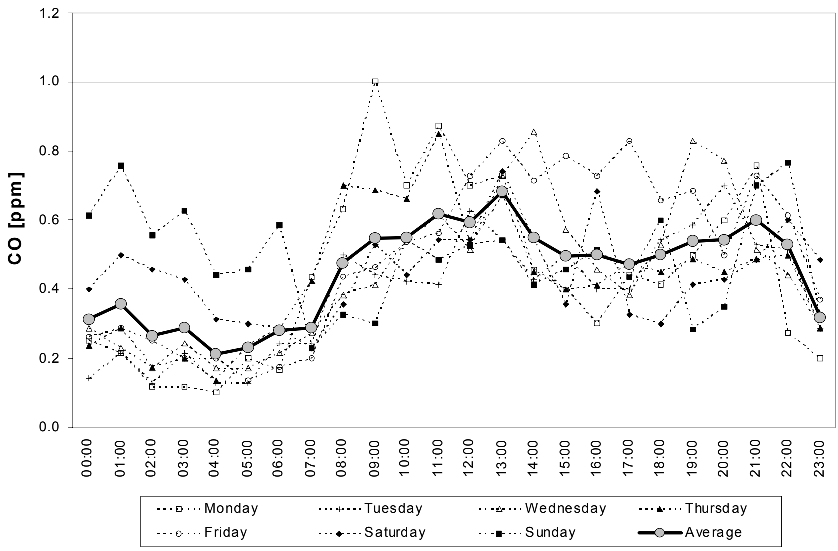


Figure 2: Hourly CO concentrations in La Plata.

The local social activities also produce noticeable differences from the mentioned general behaviour. Pergamino city, for example, shows high CO concentration levels at the last hours of the day practically everyday of the week (Figure 3). This is due to the presence of a bingo hall in the city centre. This is a place where people go to play “bingo” (game in which each player has a card with numbers on, numbers are called out in no particular order and the first player whose numbers are all called out or who has a line of numbers called out wins a money prize). Apart from its own inhabitants Pergamino centre receives during the week numerous visitors from neighboring towns.

The city location is another highly influential factor. For example, the proximity of the cities to the sea in general implies the presence of winds with high velocities and frequencies, and the formation of wind channels in big cities with tall buildings like Mar del Plata. Thus, a comparative study between Mar del Plata and San Nicolás cities is showed in reference [6]. Levels of CO pollutant five times higher in San Nicolás than Mar del Plata were determined, considering similar meteorological conditions, relative orientations between wind direction and streets axes, and similar traffic densities.

All the mentioned factors have to be taken into account in order to avoid the simplification generally observed that considers a direct relationship between pollution levels and vehicular densities. This is rigorously true at a same place with stable meteorological conditions along the analyzed period.

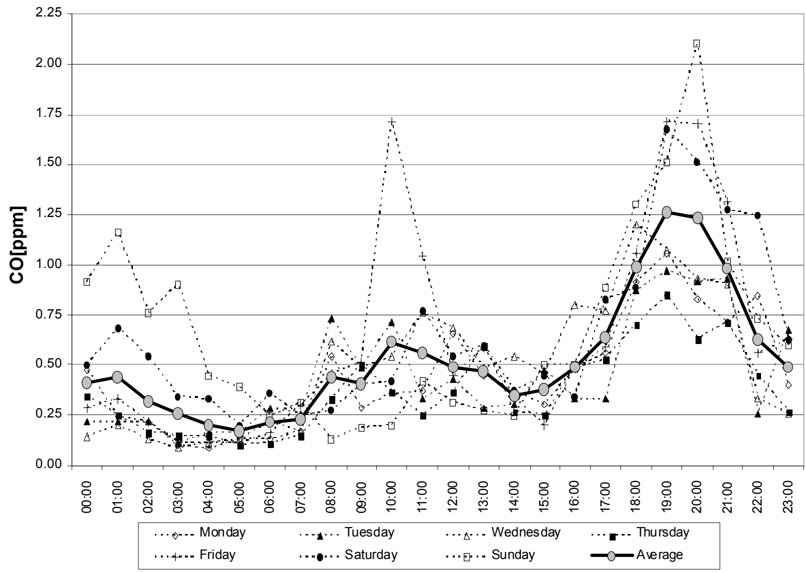


Figure 3: Hourly CO concentrations in Pergamino.

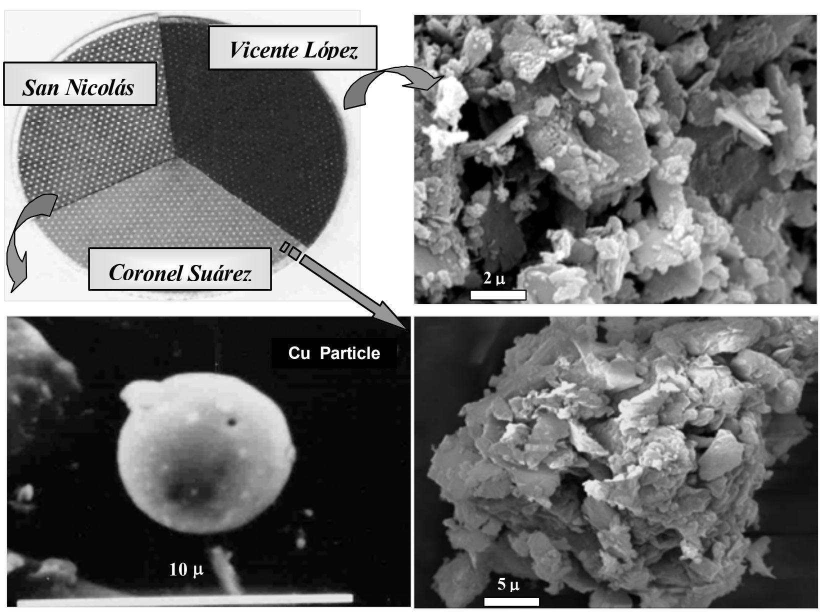


Figure 4: PM characterization in urban, rural and industrial environments.

In relation to the particulate matter analyses a precise detail of the characteristics found in cities with remarkable differences in their main activities is presented in [7]. From it, Figure 4 abstracts the concluding results.

The PM contaminant shows very different compositions depending on its origin environments: urban, industrial and rural.

In Vicente López PM samples (urban) it is possible to observe the presence of small particles nucleating over bigger ones like a mass of agglomerated particles with an important percentage of carbon.

Coronel Suárez PM analyses (rural) present the typical composition found in atmospheric dust, mainly consisting of earth particle in suspension and biological particles like pollen. The chemical compositions are similar to silicoaluminate compounds like clays and feldspars.

In the case of PM analyzed in San Nicolás (industrial) the present particles are the typical ones of vehicular pollution and also those originated from industrial high temperature processes such as metallurgical, steel or metal mechanics industries, characterized by spherical shapes.

3.2 Zonal analysis from the results of local studies

Pollution levels maps were drawn taking into account the contaminants concentrations determined at the sites under analysis.

Figure 5 shows the location of the studied cities within Buenos Aires province territory. The CO concentration values on the figure correspond to the average of the highest values determined in each analyzed site which were repeated six times or more. The other cities on the figure called "inferred sites" were selected taking into account their general characteristics (vehicles number, meteorological conditions, location in relation to the sea, etc.) and comparing them with the analyzed sites, in order to infer the pollution level. Thus, it is possible to infer similar pollution levels as follows:

-*Saladillo:* Baradero, Cañuelas, Cnel. Dorrego, Dolores, Médanos, Pigüé, Salto.
-*Trenque Lauquen:* Bolívar, Gral. Rodríguez, Lincoln, Lobos, 25 de Mayo.
-*9 de Julio and Chacabuco:* Balcarce, Bragado, Coronel Suárez.
-*San Nicolás:* Campana, Zárate.
-*Pergamino:* Escobar, Pilar.
-*Olavarría:* Tandil.
-*Bahía Blanca:* Avellaneda, 3 de Febrero.
-*La Plata:* La Matanza.

The air quality determined by the continuous monitoring stations placed in the different sites of study can be interpreted by means of indexes. In order to reach a better understanding by the population in general an index will be used that classifies the air quality at various levels. In general, the higher the value the poorer the air quality.

One index per contaminant is established according to a concentrations scale that includes the threshold values indicated by regulations. The general index is then determined taking into consideration the individual indexes of each contaminant affected by a factor that considers the danger of such specific contaminant. This quality index will be associated with a pollution level and its influence on health. Thus, pollution levels can be obtained and may be good, moderate, bad, severe or dangerous, representing respectively: almost no

influence, slight effect in the long term, affections to sensitive organisms or general effect in the short or immediate term.

In the United States, the Air Quality Index (AQI) is used for reporting daily air quality [8]. An AQI value is calculated for five major air pollutants: ground-level ozone, particle pollution, carbon monoxide, sulfur dioxide, and nitrogen dioxide. An AQI value of 100 generally corresponds to the national air quality standard for the pollutant, and values below 100 are generally thought of as satisfactory. The highest AQI value for the individual pollutants is the AQI value for that day. AQI is usually showed using colors: green-good, yellow-moderate, orange-unhealthy for sensitive groups, and red-unhealthy for all groups. In France the ATMO index, symbolized by a giraffe, is daily calculated referring to four pollutants: sulfur dioxide, nitrogen dioxide, ozone and particulate matter [9]. This index is based on a 10-point scale: 1 for very good air quality, 10 for very bad air quality. Each pollutant is given one of these numbers and the level of the ATMO index is the highest number.

The pollution levels determined in the different cities in this work can be transformed in indexes if the corresponding values are related with the standard of the Argentine regulation, which in the case of the analyzed gas (CO) is 9 ppm.

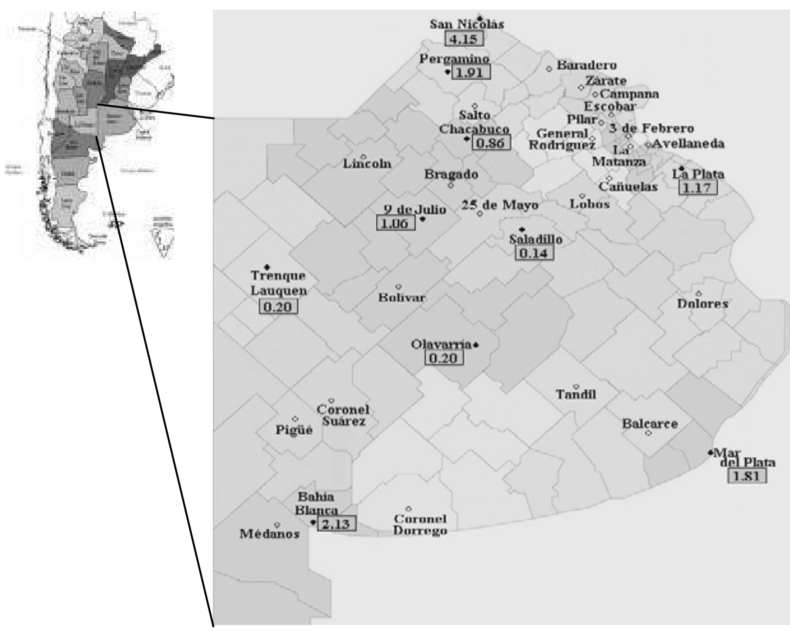


Figure 5: Location of Buenos Aires province cities. •Analyzed sites ○Inferred sites ▢CO concentration in ppm.

4 Conclusions

From the development of this project the influence of different factors on the local pollution levels in several cities of Buenos Aires province was established.

These factors are social and commercial activities, economic situation, relative location with regard to the sea, among others.

The pollutant concentration levels determined in the different analyzed cities resulted lower than the corresponding ones established in the Argentine regulations.

Acknowledgements

The authors wish to thank ANPCyT and CIC for the financial support.

References

[1] Environmental Polity Secretary *Decree 3395/96*, *Law 5965*, modified by *Resolution 242/97*. Argentina.

[2] N. Quaranta and C. Grasselli. Social activities influence on the air quality in San Nicolás City, Argentina. *Air Pollution IX*, G. Latini & C.A. Brebbia Eds., pp. 253–261, 2001.

[3] N. Quaranta, C. Grasselli and G. Merizzi. Air quality: the influence of the economic crisis. *Air Pollution XII*, C.A. Brebbia Ed., pp. 407–414, 2004.

[4] N. Quaranta, C. Grasselli, M. Caligaris and L. Videla. Study of the O_3-NO_x relationship in different periods of the year. *Air Pollution XIII*, C.A. Brebbia Ed., pp. 279–284, 2005.

[5] S. Mendiara, A. Sagedahl, M. García & N. Quaranta. NO_2 measurements at the bus station area in a big tourist city in Argentina. *Air Pollution XIV*, J.W.S. Longhurst & C.A. Brebbia Eds., pp. 251–260, 2006.

[6] N. Quaranta, M. Unsen, M. Caligaris, S. Ringler, S. Mendiara and M. García Air quality in a tourist seashore city during vacations. *Urban Transport XII*, C.A. Brebbia & V. Dolezel Eds., pp. 615–621, 2006.

[7] N. Quaranta, M. Caligaris, M. Unsen and G. Pelozo. Characterization of particulate matter from urban, industrial and rural environments. *The Sustaintable City I*, U. Mander, C. Brebbia & E. Tiezzi Eds, pp. 379–387, 2006.

[8] http://www.epa.gov/airnow

[9] NOR: DEVP0430272A. Arrêté du 22 juillet 2004 relatif aux indices de la qualité de l´air.

Overview of ultrafine particles and human health

J. C. Chow & J. G. Watson
Desert Research Institute, Reno, NV, USA

Abstract

Several hypotheses have been advanced concerning which properties of suspended particulate matter (PM) in the atmosphere cause short-term and long-term health effects. These properties include: 1) PM mass in different size fractions; 2) PM surface area; 3) PM number; 4) transition metals (especially the soluble fraction); 5) acids (especially sulfuric acid); 6) organic compounds; 7) biogenic particles; 8) sulfate and nitrate compounds (typically neutralized in whole or in part by ammonia or sodium); 9) peroxides and other free radicals that accompany and help to form PM; 10) soot (elemental carbon and associated PAH); and 11) correlated cofactors (other pollutants and variations in meteorology). Particle number and surface area are dominated by ultrafine particles (UP), those with aerodynamic diameters less than ~0.1 µm. Although the UP fraction does not contribute large quantities to $PM_{2.5}$ or PM_{10} mass (particles with aerodynamic diameters <2.5 or 10 µm, respectively), it dominates the PM number concentration and most of its surface area. UP are produced by the condensation of hot vapors in fresh combustion emissions. They also form from natural and manmade gases as secondary aerosol by photochemical oxidation of compounds with high vapor pressures to compounds with lower vapor pressures. UP may contain transition metals, organic material, sulfuric acid, and free radicals. Owing to their small size and high mobilities, they diffuse rapidly and may combine with each other, with larger particles, and with nearby deposition surfaces. When UP are inhaled, these surfaces may include those of the upper and lower human respiratory system. Because of their short lifetimes and low mass concentrations, UP are not conveniently measured in source emissions and ambient air. For similar reasons, they are not easily generated for exposure, inhalation, and toxicological studies to determine their potential adverse effects on human health.

Keywords: ultrafine, nanoparticles, PM, health effects.

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060601

1 Introduction

Several hypotheses have been advanced concerning which properties of suspended particulate matter (PM) in the atmosphere cause short-term and long-term health effects [1]. These properties include: 1) PM mass in different size fractions; 2) PM surface area; 3) PM number; 4) transition metals (especially the soluble fraction); 5) acids (especially sulphuric acid); 6) organic compounds; 7) biogenic particles; 8) sulphate and nitrate compounds (typically neutralized in whole or in part by ammonia or sodium); 9) peroxides and other free radicals that accompany and help to form PM; 10) soot (elemental carbon and associated PAH); and 11) correlated cofactors (other pollutants and variations in meteorology).

Particle number and surface area are dominated by ultrafine particles (UP), those with aerodynamic diameters less than ~0.1 μm. Although the UP fraction does not contribute large quantities to $PM_{2.5}$ or PM_{10} mass (particles with aerodynamic diameters <2.5 or 10 μm, respectively), it dominates the PM number concentration and most of its surface area. UP are produced by condensation of hot vapours in fresh combustion emissions. They also form from natural and manmade gases as secondary aerosol by photochemical oxidation of compounds with high vapour pressures to compounds with lower vapour pressures. UP may contain transition metals, organic material, sulphuric acid, and free radicals. Owing to their small size and high mobilities, they diffuse rapidly and may combine with each other, with larger particles, and/or with nearby deposition surfaces. When UP are inhaled, these surfaces may include those of the upper and lower human respiratory system.

Because of their short lifetimes and low mass concentrations, UP are not conveniently measured in source emissions and ambient air. For similar reasons, they are not easily generated for exposure, inhalation, and toxicological studies to determine their potential adverse effects on human health. Much of what is currently known about UP and their effects on health has been discovered over the past decade.

2 Physical and chemical characteristics of ultrafine particles

Figure 1 represents portions of the mass particle size distribution. This is an idealized diagram that accentuates different size modes. Actual particle size and chemical measurements are not so easily classified. Relative magnitudes of the nucleation and Aitken modes are exaggerated, as these small particles dominate particle number but are a miniscule fraction of $PM_{2.5}$ mass. Nucleation (or nanoparticle; <~10 μm) and Aitken (10 to ~100 μm) modes usually overlap, and nanoparticles are often the lower size tail of the Aitken mode. UP includes both nucleation and Aitken modes.

The condensation portion of the accumulation mode forms mostly under dry conditions while the droplet mode is consistent with aqueous-phase reactions in fogs and clouds; more material accumulates within the water droplet that leaves larger particles when the water evaporates. Another interpretation of these

modes for RH>80% is that the water-absorbing (hygroscopic) materials (e.g., sulphate and nitrate) have grown into the droplet mode while the water-repellent (hydrophobic) materials (e.g., soot and some organic carbon) have retained their original sizes. The dotted line indicates the potential for crossover among overlapping distributions; some of the coarse particles are always measured in a $PM_{2.5}$ sample.

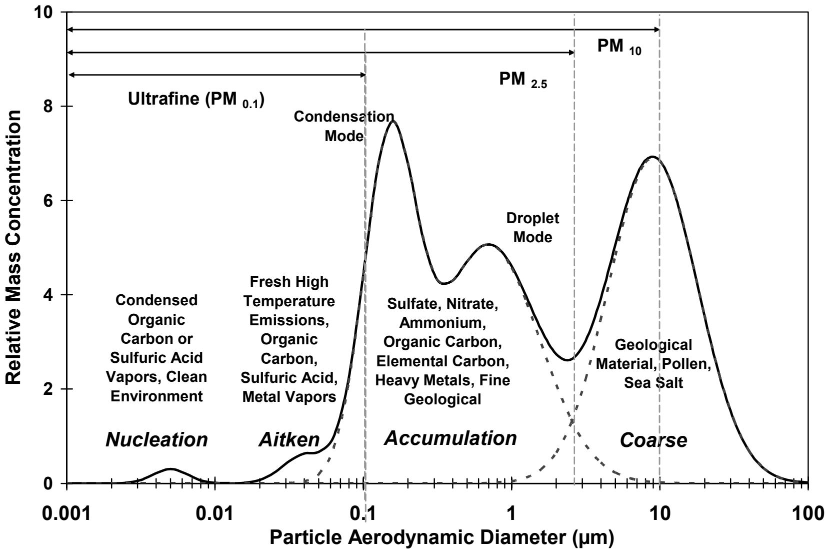


Figure 1: (After Watson, [21]). Illustration of different modes in a typical atmospheric particle size distribution. Nucleation and Aitken modes often overlap. Note that the tail (dotted line) of the accumulation mode penetrates into sizes <0.1 µm, as does the coarse mode into the accumulation mode. Sources and processes that affect each mode are indicated.

Coarse particles are dominated by geological material. They may also contain sea salt near coastal areas as well as pollen and spores, plant parts, and uncontrolled industrial emissions (modern particle removal devices such as precipitators are more efficient for coarse and larger particles than they are for accumulation mode particles). Much of the accumulation mode is occupied by secondary aerosol, particles that form in the atmosphere from directly emitted sulphur dioxide (SO_2), oxides of nitrogen (NO_x), ammonia (NH_3), and some of the volatile organic compounds (VOC, especially aromatics from internal combustion engines and terpenes and sesquiterpenes from vegetation).

Figure 2 compares the particle number and surface area for a 1 $\mu g/m^3$ concentration at different size intervals. Each 0.5 µm diameter particle in the accumulation mode would require one thousand 0.05 µm or one million 0.005 µm diameter particles in the ultrafine mode to account for the same mass

concentration. By the same token, the 0.005 μm particle has 10,000 times the surface area as the 0.5 μm particle. If adverse health effects correspond to the number or contact area of particles rather than their total mass, then UP may have a large effect despite their low mass concentrations. For practical purposes, $PM_{2.5}$ is a good indicator for the accumulation mode while the difference between PM_{10} and $PM_{2.5}$ is a good indicator of the coarse mode. UP is usually within the noise of these commonly measured mass fractions and is typically reported as particle number.

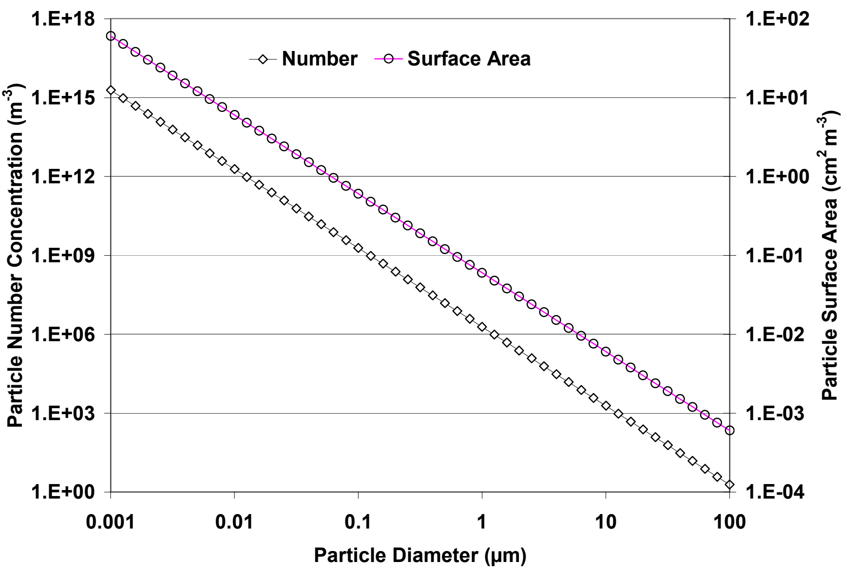


Figure 2: Particle number and surface area corresponding to a 1 μg/m^3 concentration for unit density particles as a function of geometric particle diameter.

UP are both directly emitted by combustion sources and can form in the atmosphere. Coal, petroleum products (residual oil, diesel [including fuel oil, kerosene, and aircraft fuel], and gasoline being the most common), and hydrocarbon gases (mostly methane, but also including propane and butane) are in most common use by internal and external combustion processes. Fossil fuels often include trace amounts of sulphur that can oxidize to sulphuric acid. As organic gases and sulphuric acid cool upon dilution with ambient air, they may condense onto larger particles or nucleate into UP. Black carbon soot is produced by incomplete combustion in oxygen-starved combustion. Some of these may be UP, but they often grow due to condensation and adsorption of vapours or collide with and collect the UP nuclei.

UP are also produced in clean environments. Gaseous hydrocarbons (HC) emitted by natural (e.g., plant life, wildfire) and anthropogenic (e.g., fossil fuel

and biofuel combustion) sources transform into heavier HC and oxygenated semi-volatile organic compounds (SVOC) with lower vapour pressures in the presence of sunlight. These gases can nucleate when their vapour pressure exceeds the saturation vapour pressure at a given temperature. This nucleation usually occurs when PM mass concentrations are low, as the vapours preferentially condense on larger particles when they are available in sufficient quantities. The nucleation mode usually occurs in clean atmospheres but also has been recently observed in urban areas. UP are often found in fresh combustion emissions, but they rapidly combine with each other and larger accumulation-mode particles.

Several competing and complementary processes cause UP number concentrations to decrease, and their sizes to increase, with aging time:

- Nucleation: When the atmosphere is supersaturated with a gas, spontaneous nucleation of small particles with ~1 nm diameter occurs. This dominates mostly in clean environments, as condensation onto existing particles is favoured in more polluted environments.
- Condensation and evaporation: When the ambient vapour pressure is higher than the saturation vapour pressure, condensation occurs and particles are formed or grow to larger sizes. Saturation pressures are larger over very small particles (>5 to 10 nm) than they are over larger particles owing to their curvature (Kelvin effect), so condensation is favoured on larger particles and evaporation is favoured on smaller particles. Evaporation occurs with increasing temperature and with dilution of the gaseous precursors below the saturation vapour pressure. Owing to the Kelvin effect, small particles may evaporate with their vapours condensing on larger particles, thereby leading to growth.
- Coagulation: Particles collide and combine with each other when concentrations area high, thereby decreasing their number and increasing their size.
- Deposition: Particle diffuse and adhere to the ground or other surfaces that they encounter. They can also be intercepted by clouds and fogs with ultimate rainout.

3 Ultrafine particle measurement methods

UP can be measured with respect to number, surface area, particle volume or particle mass. Continuous or semi-continuous direct reading methods include: 1) condensation particle counters, 2) electrostatic classifiers with condensation particle counters; 3) electronic low pressure impactors; and 4) particle mass spectrometers. These instruments can obtain number counts, size distributions, and even some chemical concentrations over averaging periods of seconds to minutes. Other methods acquire samples on substrates that are taken on filters or substrates and returned to the laboratory for mass measurements and other chemical analyses. Electron microscopy quantifies particle sizes and shapes on filter samples that have not necessarily been preceded by a size selective inlet.

4 Ultrafine concentrations and exposures

UP number concentrations and size distributions are highly dynamic. As such, cumulative human exposure will depend on the amount of time that an individual spends in areas with high and low concentrations. The transient nature of UP also has important implications for selection of an indicator that can be monitored to determine safe concentration levels. PM_{10} and $PM_{2.5}$ mass concentrations are currently used as health indicators in the United States because they are practical to measure, show a reasonable spatial homogeneity, and have been epidemiologically related to health effects.

Naturally occurring UP concentrations are measured near natural sources and at global- and continental-scale monitoring sites and from aircraft. Particle number counts are often in the range of 4,000 to 20,000 cm^{-3}, often comparable to the numbers measured in cities. Particle sizes are often lower. UP particle bursts in the nucleation model (3 to 10 nm) have been found frequently in coastal areas and in aircraft measurements above clouds. It is speculated that clouds reflect some of the solar radiation, thereby increasing the photochemical activity and oxidation of trace sulphur compounds that condense to small particles. Number counts of 100,000 cm^{-3} have been reached in these pristine environments under these conditions.

At the other extreme, UP have been measured in the plumes emitted by combustion sources In-plume concentrations are often within the range of 10^6 to 10^9 cm^{-3} and depend on combustion conditions and cooling. Number concentrations change with the sampling method, especially the extent of dilution and cooling. Cooling tends to create new particles by condensation, but dilution with clean air that accompanies dilution often reduces vapour pressures of condensable gases so that they do not nucleate. UP numbers are higher in plumes that have lower $PM_{2.5}$ mass emissions, as UP tend to collect on the larger accumulation mode particles, and condensable vapours favour the larger particles over homogeneous nucleation. For this reason it is difficult, and possibly even meaningless, to obtain UP emission factors for combustion sources. Variability in the emissions testing methods could result in orders of magnitude changes in the number of particles emitted per unit of fuel consumed.

People usually avoid breathing direct emissions from exhaust plumes, but they often are exposed to slightly aged plumes, especially while walking along roadways, waiting for public transportation, or even at homes and workplaces located near major traffic corridors. Much emphasis has been placed on measuring concentrations and determining human exposures at different distances from heavily-trafficked highways UP concentrations decrease more rapidly than larger particles with distance from heavily-traveled roadways.

Much of human exposure takes place indoors, at home, in school, or at the workplace and UP numbers that equal or exceed outdoor concentrations are often found in these environments. Some of the outdoor UP infiltrates inside, but there are a large number of indoor emitters. Vacuum cleaner exhaust [2] contained ~10^4 cm^{-3} of UP that could be reduced by three orders of magnitude by adding a HEPA filter. Fogarty and Nelson [3] report indoor UP numbers approaching

10^5 cm^{-3} near laser printers, photocopiers, vacuum cleaners, cleaning supplies, and tobacco smoke. These far exceed number counts in garages with vehicle exhaust where concentrations are ~20,000 cm^{-3}. Measurements near furnace leaks and chimney cracks were as high as 500,000 cm^{-3}. Rundell [4] found increases from 10,000 cm^{-3} to more than 250,000 cm^{-3} when ice was resurfaced in ice-rinks. Surfacing machines powered by propane yielded slightly higher concentrations than gasoline powered machines.

Normal household concentrations average ~10,000 cm^{-3} [5], with hospital concentrations about half this [6]. Cooking with gas can cause UP numbers to rise two or threefold indoors, but numbers decrease and size increases within minutes after emission [7]. Episodic and long-term measurements have been found at urban-scale sites that range in number from those found at pristine areas to those found at roadside monitors. Typical urban-scale concentrations are in the range of 5000 to 20,000 cm^{-3}, but they are highly variable in space and time. Values vary substantially owing to the different types of measurements applied and the proximity of the monitors to nearby emissions. There is nearly always a morning traffic peak with large numbers and small size distributions during the morning. Sometimes short-duration peaks in number and smaller sizes are observed when an elevated plume mixes to the surface in the afternoon. These events are usually of short duration. Of more recent interest is a banana-shaped afternoon increase in number and size with time. This was previously thought to be a regional-scale phenomenon because urban areas were believed to provide too many existing particles to create many new particles by nucleation. These photochemical mechanisms have been found in several urban areas and appear to provide a mechanism for a more widespread exposure than the transport of plumes from nearby sources.

5 Inhalation and dosage

Suspended particles must be inhaled and deposited in different parts of the human body to cause adverse health effects. Dosage refers to the quantity that is not exhaled and remains in different parts of the body. Figure 3 shows approximately how much of different particle sizes deposit in different parts of the human respiratory system. The extra thoracic region consists of the nasal, oral, pharyngeal and laryngeal airways. The main mechanisms for deposition in this region are impaction and diffusion. Most particles in the coarse particle fraction are removed by impaction in this upper part of the respiratory system. Allergic reactions and sinus ailments result from irritation in this region. The intrathoracic region includes the trachea, bronchial airways and alveoli in the lung. Particles reaching this region can cause lung damage and potentially transfer to the bloodstream for transport to other organs. Most of the UP pass through the upper respiratory tract and deposit in the lower respiratory tract. Deposition efficiencies increase at higher breathing rates, such as those experienced during exercise.

Inhalation models are of varying complexity, with many simplifying assumptions that limit their accuracy and precision, especially for UP. Most include the particle deposition mechanisms of impaction, interception and

diffusion. The complex twists and turns of the upper and lower airways are usually simplified to straight pipes with sharp bends. Constant temperatures and relative humidities are assumed throughout the airway. Condensation, evaporation, hygroscopic properties, and particle agglomeration are omitted.

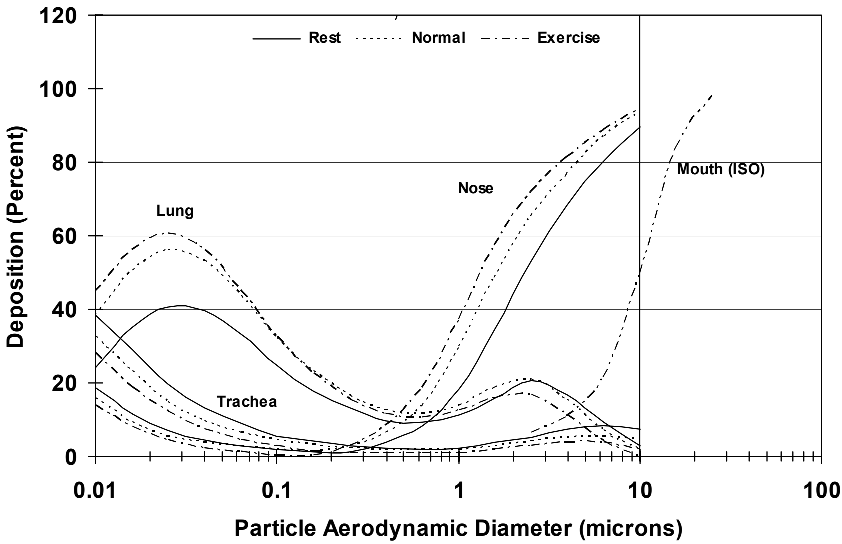


Figure 3: (From Chow, [22]). Ultrafine and larger particle deposition in different portions of the human respiratory tract, based on Phalen et al. [23].

Zhang et al. [8] describe a computational fluid dynamic (CFD) approach to inhalation and deposition modelling that includes many, but not all of these variables. Upper airway and lung bifurcation airways are modelled separately, with the end conditions of the upper airway model serving as initial conditions for the lower airway model. Total dosage is measured by having human subjects inhale particles of known sizes through a mouthpiece with a nose plug, then measuring the concentrations and sizes when the subject exhales. The difference between inhaled and exhaled as a function of size is taken at that which remained within the body. Some deposited material may be dislodged and exhaled in subsequent breathing cycles, but this averages out over multiple cycles. Tests are made with subjects at rest and under respiratory stress, as well as on subjects with different ages and physical characteristics. There is substantial variability among subjects.

6 Toxicology

Air pollution toxicology studies determine the extent to which the human body reacts adversely to physical, chemical, and biological irritations caused by inhaled particles and gases. Toxicological studies are conducted in vivo,

wherein living subjects inhale UP, or in vitro, wherein UP is applied to cell cultures.

Toxicological studies are usually performed by trained physicians and the terminology differs from that used by air quality scientists and engineers. Several commonly-used terms are defined below to facilitate understanding of the published literature.

- Biopersistence [9]: A measure of particle retention in in vivo studies.
- Biodurability [9]: A measure of particle retention in in vitro studies.
- Clearance [10]: Removal of non-soluble particles from the lung via the airways or via the pulmonary interstitium/lymphatic system.
- Cytokine: An intercellular chemical messenger protein released by white blood cells. Cytokines facilitate communication among immune system cells and between immune system cells and the rest of the body.
- Epithelium: The cells or membrane covering the outside of organs.
- Fibroblast: A cell from which connective tissue develops
- Fibrosis: A process producing scar tissue that forms as a reaction to injury or during a healing process.
- Growth factors: Small proteins produced by the human body that enable cells to communicate and effectively coordinate activities between one another. Growth factors in the body affect the individual cells by binding to growth-factor-specific receptors on the cell surface. A specific growth factor may have many cell sources and can use different signal transduction pathways at different times and with different cells. Growth factors are involved in complex feedback loops between the immune, nervous and endocrine systems, and have significant effects on DNA, RNA, protein synthesis, and cell division.
- Lavage: Washing lung surfaces to obtain cells and foreign materials. Usually performed multiple times with a saline solution. Lavage is applied to both live and sacrificed subjects as well as to cell cultures. Lavage contents are analyzed chemically and microscopically.
- Macrophage: A white blood cell that engulfs and digests foreign organisms. Macrophages also stimulate helper T-cells to participate in the immune system response.
- Neutrophil: A type of white blood cell distinguished by a lobed nucleus and granular cytoplasm. It is capable of ingesting and killing bacteria and releases various substances, such as lysozyme (antibacterial enzyme) and oxidizing agents.
- Phagocyte: A cell that is able to engulf and break down foreign particles, cell debris and disease producing micro-organisms in the body. Phagocytes form an important part of the natural defense mechanism in most animals.
- Proteases: Enzymes that cause the decomposition of protein.
- Retained dose [9]: The amount of inhaled particles that remain after clearance mechanisms have reduced the initial dose.

Macrophages, neutrophils, and phagocytes from the lung obtained by lavage are often analyzed with respect to type and number to determine the extent to which a toxicological effect has been observed. Chemicals released by neutrophils and growth factors are often measured as indicators of the functioning of the body's immune system. Very high dosages are given during some tests to determine when the body's defence mechanisms are overwhelmed.

Several hypotheses have been advanced about the toxicological effects of UP and why they might be more intense than those of larger particles [11–13]:

1. UP are more inflammatory than larger particles. The larger UP surface area reacts with macrophages and epithelial cells that prime and activate cells for inflammatory reactions.
2. UP contain more or release more toxic free radicals from their surfaces.
3. UP are not phagocytized by alveolar macrophages as efficiently as larger particles. Much larger UP numbers may spread more widely over the alveolar epithelium, thereby creating a smaller signal to the macrophages than the less numerous, and more spatially distinct, larger particles.
4. UP inhibit phagocytosis, the ability of phagcytes to remove foreign objects.
5. UP translocate from their initial deposition position to other locations. Insoluble UP may translocate from the lungs to other organs through the bloodstream or lypmphatic system.
6. Adverse UP effects are enhanced in the presence of oxidant gases (e.g., ozone).
7. Older people and those with compromised respiratory tracts are more susceptible to UP than to larger particles.

Organic chemicals (e.g., polycyclic aromatic hydrocarbons, PAH) can affect the DNA, causing mutations that may result in cancerous tumours. These are relatively long-term effects that require consistent exposure. The larger number of UP provides a larger probability of interaction with cells and the larger surface area gives more opportunity for irritation and interaction with the DNA. Soluble transition metals can generate hydroxyl radicals which are highly reactive and cause neutrophils to release inflammatory mediators. Plaques may form and cells may rupture as a result of this inflammation. Aveolar machrophages and neutrophils may be activated on ingesting UP, thereby releasing cytokines. These can cause fibrosis in several of the body's organs.

Under good conditions, alveolar macrophages will ingest and clear UP from the deep lung. However, these may become saturated in the brochoalveolar space and adhere or pass through the epithelium. Upon passage, their secretions can become incorporated into the capillary and lymphatic systems, and these secretions may cause fibrosis in other organs. When macrophages are overwhelmed, UP may directly contact, remain on, and even pass through the epithelium. Some of the non-soluble UP may enter the bloodstream and become lodged in other organs such as the heart.

Most toxicological studies have been performed on rodents, notably rats, mice, hamsters, and guinea pigs. Dogs have also been used in limited studies. A few studies have involved a limited number of human subjects. The advantage of using animals is that they can be subjected to much higher, more frequent, and longer UP exposures than would be ethically acceptable or practical for humans. Animals can also be sacrificed after exposure to obtain a wider range of measurements than are possible for human subjects. Animal subjects are also less costly to recruit and maintain. UP and other pollutant concentrations can be delivered intertracheally to animals, thereby bypassing deposition in the upper respiratory tract and providing a better estimate of dosage deep into the lung. The disadvantage of animals is that their inhalation and physiological reaction properties differ substantially form those of humans. The chemical and physical mechanisms may be similar in animals and humans, but the dosage that causes an effect is quite different.

A large number of different UP test aerosols have been generated for toxicological testing. Titanium dioxide (TiO_2) is a solid powder that can be obtained in with sizes in the UP and accumulation modes and can be suspended and dispersed for presentation to the subjects. Elemental carbon (EC) can be reliably produced with electric spark generators. Carbon black is another solid powder that is available in different sizes and has a more complex structure than the EC produced by carbon electrodes. Interest in the effects of metals has led to tests with aluminum, cadmium, cobalt, copper, gallium, nickel, and zinc compounds. Teflon fumes are also used. With the advent of aerosol concentrating inlets described in Section 3, it has become possible to expose subjects to concentrated air pollutants.

A large number of health end-points can be measured. Most of these are determined chemically or by growth cultures after the subjects are sacrificed. Counting of macrophages, neutrophils, and phagoscytes in lavage solutions is most commonly used. There are usually many more macrophages after UP exposure, indicating that the body is attempting to clear the inhaled UP. Alveolar macrophages constitute the main UP clearance mechanism. Substantial and consistent overload is needed before other adverse effects are incurred [14]. This is probably why very large UP concentrations are needed to overwhelm these mechanisms, or they must be already weakened by compromising the test subjects.

7 Epidemiology

Epidemiological studies seek relationships between health end-points such as respiratory illness (morbidity) or death (mortality) and air quality indicators such as $PM_{2.5}$ and PM_{10} mass, gaseous pollutant concentrations, and UP number concentrations. Other end-points are: 1) peak respiratory flow (PEF) as measured with a flow meter is a common end-point, 2) chronic obstructive pulmonary disorder (COPD), 3) frequency and intensity of medication use, and 4) increases in respiratory symptoms such as stuffiness or cough.

A small number of epidemiological studies have been undertaken to relate UP concentrations to adverse affects [11, 15–19].

Cities with epidemiological studies include Koupio and Helsinki Finland, Erfurt Germany, Aberdeen Scotland, Amsterdam Netherlands, and Linz Austria. All of these identified vehicle exhaust as a common source of several pollutants, including UP. $PM_{2.5}$ or PM_{10} mass were also included in most of the analyses, as were soot (usually approximated from British Smoke [BS] or elemental carbon [EC]), oxides of nitrogen (NO, NO_2), sulfur dioxide (SO_2) and carbon monoxide (CO). In some cases, chemical composition was measured on 24-hour average filters using a Harvard Impactor Sampler, but this speciation did not apply to UP. Different regression models were applied to the multivariate ambient concentrations and health end-points to determine the contributions of different pollutants to health.

Most of these studies found consistency with previous epidemiological studies [20] in that health indicators showed more distress as pollution levels increased. No associations of adverse health effects were exclusively attributable to UP, although some of the regression models explained more of the variance in end-points when UP number was included as a variable. Several of the studies noted a good correlation between UP number and $PM_{2.5}$ that would make their isolated effects collinear with each other.

8 Summary and conclusions

UP represent a small fraction of the particle mass, but dominate particle number and surface areas. UP number can be quantified with condensation particle counters. UP size distributions can be measured with scanning electrical mobility analyzers, low pressure impactors, and single particle mass spectrometers. Different instruments give different number counts and size distributions owing to different particle properties and size ranges measured. UP derive from fresh combustion emissions, typically combustion sources, as well as nucleation of oxidized gases. UP have natural as well as manmade origins.

UP are short-lived, on the order of minutes to hours, owing to evaporation and coagulation with other particles. UP last longer in clean remote environments than in polluted urban environments where condensation is preferred on existing larger particles. UP number counts are highly variable with space and time. Background UP concentrations are typically 5000 to 20,000 cm^{-3}, but they can approach 10^5 cm^{-3} during nucleation events. Urban concentrations are typically 10,000 to 50,000, with highest concentrations measured near roadways. Nucleation events have recently been observed in urban environments.

UP deposit less in the upper respiratory tract and more in the deep lungs than do larger particles. Higher numbers increase the probability of contact with lung tissue and transport across the lung epithelium into the interstitium. Toxicological studies, mostly with animal subjects, produce larger measurable health end-points for UP than for equal masses of accumulation mode particles in compromised subjects. However, test aerosols are usually at higher concentrations than ambient and compositions are not typical of those in ambient air. Toxicological studies indicate that UP can be highly oxidizing, thereby

inducing inflammation in macrophages and lung tissues. Very high concentrations can overwhelm the macrophage removal system, especially when subjects are compromise.

Epidemiological studies relating UP to health effects are few. UP is highly correlated with other pollutants and the relative risk it provides is not clearly distinguishable from that associated with large sizes fractions such as $PM_{2.5}$ and PM_{10}. Time lagged regression for a study in Germany showed respiratory effects one day after exposure and cardiovascular effects four days after exposure.

References

[1] Mauderly, J.L., Neas, L.M., & Schlesinger, R.B., PM monitoring needs related to health effects. *Atmospheric Observations: Helping Build the Scientific Basis for Decisions Related to Airborne Particulate Matter*, eds. D.L. Albritton & D.S. Greenbaum, Health Effects Research Institute: Cambridge, MA, pp. 9-14, 1998.

[2] Lioy, P.J., Wainman, T., Zhang, J., & Goldsmith, S., Typical household vacuum cleaners: The collection efficiency and emissions characteristics for fine particles. *J.Air Waste Manage.Assoc.*, **49(2)**, pp. 200-206, 1999.

[3] Fogarty, R. & Nelson, P.A., Tracking ultrafine particles in building investigations. *Indoor Air Quality Handbook*, eds. J.D. Spengler, J.M. Samet, & J.F. McCarthy, McGraw-Hill: New York, p. 50.1-50.18, 2001.

[4] Rundell, K.W., High levels of airborne ultrafine and fine particulate matter in indoor ice arenas. *Inhal.Toxicol.*, **15(3)**, pp. 237-250, 2003.

[5] Wallace, L. & Howard-Reed, C., Continuous monitoring of ultrafine, fine, and coarse particles in a residence for 18 months in 1999-2000. *J.Air Waste Manage.Assoc.*, **52(7)**, pp. 828-844, 2002.

[6] Riesenfeld, E., Chalupa, D., Gibb, F.R., Oberdörster, G., Gelein, R., Morrow, P.E., Utell, M.J., & Frampton, M.W., Ultrafine particle concentrations in a hospital. *Inhal.Toxicol.*, **12(Suppl 2)**, pp. 83-94, 2000.

[7] Dennekamp, M., Howarth, S., Dick, C.A., Cherrie, J.W., Donaldson, K., & Seaton, A., Ultrafine particles and nitrogen oxides generated by gas and electric cooking. *Occup.Environ.Med.*, **58(8)**, pp. 511-516, 2001.

[8] Zhang, Z., Kleinstreuer, C., Kim, C.S., & Cheng, Y.S., Vaporizing microdroplet inhalation, transport, and deposition in a human upper airway model. *Aerosol Sci.Technol.*, **38**pp. 36-49, 2004.

[9] Oberdörster, G., Ferin, J., Soderholm, S.C., Gelein, R., Cox, C., Baggs, R., & Morrow, P.E., Increased pulmonary toxicity of inhaled ultrafine particles: Due to lung overload alone? *Ann.Occup.Hyg.*, **38(Suppl. 1)**, pp. 295-302, 1994.

[10] Ferin, J., Pulmonary retention and clearance of particles. *Toxicol.Lett.*, **72(1-3)**, pp. 121-125, 1994.

[11] Ibald-Mulli, A., Wichmann, H.E., Kreyling, W., & Peters, A., Epidemiological evidence on health effects of ultrafine particles. *J.Aerosol Med.*, **15(2)**, pp. 189-201, 2002.

[12] MacNee, W. & Donaldson, K., How can ultrafine particles be responsible for increased mortality? *Monaldi Arch.Chest Dis.*, **55(2)**, pp. 135-139, 2000.
[13] Oberdörster, G., Toxicology of ultrafine particles: In vivo studies. *Phil.Trans.Roy.Soc.Lond.A*, **358(1775)**, pp. 2719-2740, 2000.
[14] Ferin, J., Oberdörster, G., & Penney, D.P., Pulmonary retention of ultrafine and fine particles in rats. *Am.J.Respir.Cell Mol.Biol.*, **6(5)**, pp. 535-542, 1992.
[15] Pekkanen, J., Peters, A., Hoek, G., Tiittanen, P., Brunekreef, B., de Hartog, J., Heinrich, J., Ibald-Mulli, A., Kreyling, W.G., Lanki, T., Timonen, K.L., & Vanninen, E., Particulate air pollution and risk of ST-segment depression during repeated submaximal exercise tests among subjects with coronary heart disease - The exposure and risk assessment for fine and ultrafine particles in ambient air (ULTRA) study. *Circulation*, **106(8)**, pp. 933-938, 2002.
[16] Peters, A., Wichmann, H.E., Tuch, T., Heinrich, J., & Heyder, J., Respiratory effects are associated with the number of ultra-fine particles. *Am.J.Respir.Crit.Care Med.*, **155(4)**, pp. 1376-1383, 1997.
[17] Spix, C., Tuch, T., & Wichmann, H.E., Daily mortality and fine and ultrafine particulates in Erfurt, East Germany. p. 35, 1996.
[18] von Klot, S., Woelke, G., & Tuch, T., Short-term effects of ultrafine and fine particles on medication use in asthmatic adults. *Am.J.Respir.Crit Care Med.*, **161**p. A310, 2000.
[19] von Klot, S., Wolke, G., Tuch, T., Heinrich, J., Dockery, D.W., Schwartz, J., Kreyling, W.G., Wichmann, H.E., & Peters, A., Increased asthma medication use in association with ambient fine and ultrafine particles. *Eur.Respir.J.*, **20(3)**, pp. 691-702, 2002.
[20] Pope, C.A., III, Dockery, D.W., & Schwartz, J., Review of epidemiological evidence of health effects of particulate air pollution. *Inhal.Toxicol.*, **7(1)**, pp. 1-18, 1995.
[21] Watson, J.G., Visibility: Science and regulation. *J.Air Waste Manage.Assoc.*, **52(6)**, pp. 628-713, 2002.
[22] Chow, J.C., Critical review: Measurement methods to determine compliance with ambient air quality standards for suspended particles. *J.Air Waste Manage.Assoc.*, **45(5)**, pp. 320-382, 1995.
[23] Phalen, R.F., Cuddihy, R.G., Fisher, G.L., Moss, O.R., Schlessinger, R.B., Swift, D.L., & Yeh, H.C., Main features of the proposed NCRP Respiratory Tract Model. *Radiat.Protect.Dosim.*, **38(1/3)**, pp. 179-184, 1991.

Global climate change, air pollution, and women's health

K. Duncan
University of Toronto, Canada

Abstract

Climate change will disturb the Earth's physical systems (e.g. weather patterns) and ecosystems (e.g. disease vector habitats); these disturbances, in turn, will pose direct and indirect risks to human health. Direct risks involve climatic factors that impinge directly on human biology. Indirect risks do not entail direct causal connections between climatic factors and human biology.

The Third Assessment Report (TAR) of the Intergovernmental Panel on Climate Change elucidates the potential human health impacts of global climate change at both a population and regional level.

The impacts on child health, adult health, and the health of the elderly, however, remain largely unexplored. A paucity of research regarding women's health is also extant, despite increasing interest in the issue.

According to the TAR, climate change is projected to affect such key issues as air quality, food yields and nutrition, water-related infectious diseases, and water supply. Exposure to cooking fuels, access to food, distribution of food within the family, and choice of water sources is often determined by gender. Thus, women's contributions may, in some cases, make them more vulnerable than their male counterparts to climate change. Moreover, it is anticipated that health care will significantly help people adapt to climate change. Unfortunately, not everyone has adequate health care. In some countries, fewer than 25 % of women visit health-care professionals.

Climate change is likely to have a strong, positive (worsening) effect on smog and acidic deposition; climate change is likely to have some effect on suspended particulates.

In light of the foregoing, this paper addresses the interrelated and neglected areas of global climate change, air pollution, and women's health.
Keywords: global climate change, air pollution, thermal extremes, women's health.

1 Women's health

It is imperative to address the future of women's health; human health depends considerably on investing in the well-being of women, as their physical condition largely determines the health of their children who are the adults of tomorrow (Duncan [1]; and World Health Organization [2-4]). Moreover, being female or male has a major impact on an individual's health and well-being. For example, the natural course of a disease may be different in women and men, women and men may respond differently to illness, society may respond differently to sick women and men, women and men may have different access to healthcare, may be treated differently by health providers, and may respond differently to treatment (Duncan [5]).

2 Environment and women's health

Since the early 1980's, considerable attention has been devoted to the relationship between the environment and women's health because the survival of women and their families remains closely linked to the health of fisheries, forests, land, and other natural resources. Therefore, exposure to poor environments often leads to the development of disease and ill health. For example, air pollution in Ukraine has been linked to 21 % of all illnesses affecting women and children, and water pollution in Uzbekistan has led to an increase in complications in pregnancy and an increase in birth defects.

The Platform for Action, adopted in Beijing in 1995, called for a greater recognition of women as guardians of the environment, a decrease of environmental risks to women, and enhanced participation of women in the management of environmental resources (Division for the Advancement of Women [6]).

Despite increasing interest in the issue of the environment and women's health, there remains a lack of training in this area, and a dearth of information regarding women and climate change (one of today's most pressing global environmental issues) (Stipp [7]; Houghton et al., [8]; and Climate Change Solutions [9]).

2.1 Climate change and women

For the first twenty years that climate change was debated, gender issues were not even on the agenda—although women (1) are more dependent than men on primary resources (e.g. agriculture and fisheries) that are threatened by changes in climate, and (2) are generally poorer financially than men (Duncan [10]).

Women are, however, particularly vulnerable to climate change. For example, women often shoulder the responsibility for the household water supply and its purification. In Kenya, carrying water may use up to 85 % of a female's daily energy intake (Women's Environment and Development Organization [11]; and World Health Organization [3]).

'She does not know her age, but is probably about twelve years old. She visits the same water hole at 05:30 each day in order to collect water for her family—despite its being muddy and being used by animals. She does not boil the water, as a nearby man tells her that it 'contains no living things'
(World Health Organization [12]).

Women collect water and manage its use in households. They also farm irrigated and rain-fed crops, and they know water's availability, quality, and reliability (Jahnavi [13]).

As a result of the division of labour, women often come in contact with polluted or poor-quality water, and are therefore more vulnerable to water-related diseases than men. These diseases kill between five and twelve million people per year, most of them women and children. And millions more are sickened with diarrhoea, hepatitis, malaria, schistosomiasis, and trachoma that are preventable through access to clean water and health care information (Kumar and Clark [14]; Meade et al., [15]; Women's Environment and Development Organization [11]; and World Health Organization [3]).

In the past ten years diarrhoea killed more children than all those lost to armed conflict since the Second World War. In fact a child dies every 15 seconds from diarrhoea—mainly due to poor sanitation and poor water supply (Duncan [5]).

It is women who bear the main burden of caring for those who are ill. But when it comes to decision-making about climate change, women in most parts of the world remain almost invisible (Duncan [5, 10]; and Roehr [16]).

COP-8, however, at last recognized that women are extremely vulnerable to climate change and that they may bear an unreasonably large share of the adaptation burden. And on March 6^{th} 2003, a Canada-China Workshop on Gender Equality and Climate Change was held in China--a possible first step to ensuring that women become active partners in vulnerability studies, mitigation efforts, and climate change adaptation (Duncan [10]).

3 Climate change

Global average surface temperature increased by 0.6 +/- 0.2^o C over the last century. Rainfall increased in mid- to high-latitudes, and episodes of El Nino increased in frequency, intensity, and persistence since the mid-1970s (Houghton et al., [8, 17]).

Global mean temperature is projected to increase 1.4-5.8^o C over the coming century, a rise greater than any increase experienced by humans during the past 10,000 years, and global sea level is projected to rise 9-88 cm by the year 2100 (Houghton et al., [8]). [It should be noted that preparations for the Fourth Assessment Report by the Intergovernmental Panel on Climate Change are still underway, and that the final reports are due in 2007 (Intergovernmental Panel on Climate Change [18]).]

Susceptibility to climate change differs across countries, populations, regions and sectors (Cohen and Miller [19]). For example, an increase in sea level will

significantly increase the number of people vulnerable to floods. Currently, some 46 million people per year experience flooding due to storm surges; a 50-centimetre sea-level rise would increase the number of vulnerable individuals to 92 million, and a one-metre rise would increase the number to 118 million (McCarthy et al., [20]).

4 Climate change and human health

Climate change will disturb the Earth's physical systems (e.g. weather patterns) and ecosystems (e.g. disease-vector habitats). The disturbances will, in turn, pose direct and indirect risks to human health. Direct risks involve climatic factors, such as extreme thermal events and severe weather that impinge directly on human biology. Indirect risks, including air pollution, decreases in food production, reductions in water quality and quantity, and vector-borne disease, do not entail causal connections between climatic factors and human biology (McMichael and Haines [21]; and Duncan et al., [22]).

Climate change is a growing concern to the World Health Organization because of its potentially serious health consequences, its disproportionate impact on poor countries, and its disproportionate impact on poor and vulnerable groups. For example, a total of 1.3 billion people in the developing world live below the poverty threshold—seventy % of these are women (World Health Organization [2, 23]).

5 Air pollution and human health

Our protective and life-giving atmosphere is no longer in its natural form; that is, humans have altered its chemistry. As early as 1273, King Edward 1 issued a proclamation forbidding the use of impure sea coal, which produced considerable soot and sulphur dioxide when burned. One person was reputedly executed for violating his decree (Mannion [24]).

In spite of the King's mandate, English coal use increased during the fifteenth and sixteenth centuries, and in 1661, the prominent scientist John Evelyn wrote an essay condemning London's filthy air. As industrialization progressed, the smoke problem worsened, and by the 1850's, London was notorious for its dangerous 'pea-soup' fogs. In 1873, one such fog claimed the lives of 700 people, and in 1911, 1150 Londoners succumbed to another deadly haze. More recently, in 1952, a sulphurous London fog killed 4000 people (Hidore [25]; and Ahrens [26]).

Air pollutants are airborne substances, either solids, liquids or gases, that occur in concentrations high enough to threaten the health of people and animals, harm vegetation and structures, or toxify a given environment. Airborne pollutants come from both natural sources and human activities. Examples of natural sources include particulate matter from dust storms, salt from ocean waves, soot from forest fires, and volcanic ash. Human-induced pollution enters the atmosphere from both fixed sources, such as homes, industrial complexes, offices, and power plants, and mobile sources, such as jet aircraft, motor vehicles, and ships (Duncan et al., [22]).

The United States Environmental Protection Agency reported in 1989 on the magnitude of industrial injection of chemicals into the atmosphere. In the United States, a total of 2.7 billion pounds of pollutants, of which 360 million pounds were suspected of being carcinogenic, were released into the atmosphere in 1987 alone. A total of 1600 industrial facilities in 46 states emitted carcinogens into the air, and 125 of these plants released more than 400,000 pounds of chemicals per year (Hidore [25]). In 1997, approximately 107 million people in the United States lived in counties that did not meet air quality standards for at least one regulated pollutant.

Acid precipitation refers to precipitation acidified by atmospheric pollutants, such as sulphur dioxide (SO_2) and nitrogen oxides (NO_x), emitted from smelters and fossil-fuel power stations. These pollutants are converted chemically to sulphuric and nitric acid in the atmosphere. The contaminants may travel long distances before being washed out of the atmosphere in drizzle, freezing rain or snow, hail, or rain (Duncan et al., [22]).

In 1992, EEO provided the following information regarding acid rain in Canada: more than 80 % of Canadians live in areas with 'high acid rain-related pollution levels'. A recent study found that long-term exposure to acid aerosols had a deleterious effect on lung development, growth, and function of children living in 24 communities in Canada and the United States (Dockery et al., [27]). A total of 300,000 Canadian lakes are vulnerable to acid precipitation, 150,000 lakes are being damaged by acid rain, and 14,000 lakes are acidified (Duncan et al., [22]).

Air toxics, including aerosols (fine solid particles and liquid droplets combined), gases, or particulates, may have immediate or long-term adverse effects on human health (e.g. birth defects, cancer, nervous system problems) and the environment (e.g. contamination of food, soil, or water).

One of the most well-known accidental releases of toxics occurred on the night of 2-3, December 1984, when 40 tons of methyl isocyanate (an extremely toxic agent which, in the short term, may cause bronchial pneumonia, bronchitis, pulmonary oedema, and even death) escaped from Union Carbide Corporation's pesticide factory in Bhopal, India, and spread over the surrounding city of nearly 900,000 people. Hundreds of thousands were injured, and thousands died.

Although it is impossible to know the true number of sickened and dead, 7000 shrouds purchased by grieving families in the days following the disaster may give a better estimate of those killed than the approximation of 2000-4000 dead commonly quoted. The long-term health effects are also difficult to evaluate; however, the International Medical Commission on Bhopal estimated that in 1994, upwards of 50,000 people remained partially or totally disabled (Dhara and Dhara [28]).

Air toxics include a wide variety of chemicals, such as polychlorinated biphenyls (PCBs), trace metals, such as arsenic, lead, and mercury, and unburned hydrocarbons, such as benzene, dioxins, furans, formaldehyde, polycyclic aromatic hydrocarbons (PAHs), and trichloroethylene. Air toxics are released by sources such as chemical plants, dry cleaners, printing plants, and motor vehicles.

Particulate matter is a term used to describe the material that can be filtered from the air. Particulate matter includes particles such as ash, dust, fibre, pollen, soot, and other tiny fragments of solid material that are dispersed into the atmosphere by natural sources and human activities. In general, these particulates are produced by human activities, including the burning of diesel fuel by trucks and buses, incineration of garbage, industrial processes (such as agricultural burning, mining operations, and steelmaking), mixing and application of fertilizers and pesticides, road construction, and thermal-power generation plants.

Particulate matter is categorized by size. Fine particulates include the particle sizes PM 2.5 and PM10. PM 2.5 refers to particles at or below a diameter of 2.5 micrometres (um), whereas PM10 refers to particles that are 10 um or less in diameter. Fine particulates are associated with impaired lung function, increased hospitalizations, pulmonary function decrements, pulmonary inflammation, reduced exercise capacity, and even increased mortality from cardiopulmonary disease and lung cancer (Duncan et al., [22]).

Currently, in most large cities in the developing world, airborne particulate levels are five times higher than in developed countries, and more than three million people die each year as a result of air pollution (United Nations Department of Economic and Social Affairs [29]). For example, in Sao Paulo and Rio de Janeiro, 27 million people are exposed to high levels of particulate air pollution, which is thought to cause 4000 deaths annually (Regional Office for Latin America and the Caribbean [30]).

Photochemical smog is produced through complex photochemical reactions when two pollutants, hydrocarbons and nitrogen oxides, react in the presence of strong sunlight, high temperatures (above 18 ^{o}C), and stable air masses. Ground-level ozone is one of the major components of photochemical smog (Duncan et al., [22]; and Ahrens [26]).

More than half of all Canadians live in areas in which ground-level ozone may reach unacceptable levels during the summer months. Peak one-hour concentrations during typical pollution episodes in the Windsor-Quebec City corridor often reach 150 ppb. Windsor exceeds standards for ozone air quality (82 ppb) 30 days per year on average (Duncan et al., [22]).

Possible health effects of smog range from severe, uncommon events (e.g. death) to mild, common effects (e.g. eye, nose, and throat irritation) and asymptomatic changes of unclear clinical significance (e.g. small pulmonary function decrements and pulmonary inflammation). Ozone may pose a particular threat, however, to those who already suffer from respiratory problems, such as asthma, chronic bronchitis, or emphysema; these conditions affect about 7.5 % of the Canadian population (Stieb et al., [31]; and Duncan et al., [22]).

Indoor air pollution is also a problem in much of the world. About 2.5 million women and children die each year from acute respiratory infections due to indoor air pollution. Traditional cook stoves produce carbon monoxide, hydrocarbons, pollutants, and smoke that affect the health of those who tend the stoves. More than two-thirds of deaths are associated with indoor air pollution, which affects mostly women and children (Duncan [10]).

In South Africa, children living in homes with wood stoves are almost five times more likely than others to develop respiratory infections severe enough to require hospitalization. In rural Mexico, coal smoke exposure can increase lung cancer risks by a factor of nine or more, and in India, smoke exposure has been associated with a 50 % increase in stillbirths. (Improved biomass cook stoves, such as the Upesi stove developed in Kenya, not only conserve biomass resources and reduce the time and energy needed for collecting fuel and cooking, but also emit 60 % less smoke.) (United Nations Department of Economic and Social Affairs [29])

Air pollution harms more than 1.1 billion people each year, and kills three million annually. Ninety percent of these deaths occur in developing countries, where air pollution is at its worst; for example, Mexico City exceeded the ozone standard (0.1 parts per million) for 1400 hours over 145 days in 1991 (Duncan [10]).

Chen et al., [32] speculate that air pollution may affect females more than males because females may have a greater deposition of inhaled particles in their lungs, leading to increased health risks. The authors also theorize that females may be more sensitive to airborne pollution since they have fewer red blood cells than males, and thus may be more sensitive to the toxicological influences of air pollutants.

6 Climate change and air pollution

Climate change is likely to have a strong, positive (worsening) effect on smog and acidic deposition; climate change is likely to have some effect on suspended particulates. Health effects range from severe, uncommon events (e.g. death) to mild, common events (e.g. eye, nose and throat irritation). Worsening air quality will therefore further impair the health of women and children who already suffer from indoor air pollution.

7 Climate change, air pollution, thermal extremes, and women

In a warmer world, heat waves are expected to become more frequent and severe (Duncan *et al.,* [22]). A warmer climate and thermal extremes may increase urban air pollution, and will affect local and regional air pollution concentrations (Suzuki [33]).

More specifically, warmer temperatures are expected to be accompanied by an increase in primary and secondary pollutants (McMichael and Haines [21]; and Patz and Kahliq [34]). For example, heat and sunlight are significant variables in the production of smog (Suzuki [33]); smog has the potential to aggravate pre-existing cardiovascular and respiratory conditions (Burnett et al., [35, 36]; and Bransford and Lai [37]), and potentially increase the number of hospitalizations (Duncan et al., [22]).

Particularly vulnerable people include the young, the elderly, the poor, the frail and the ill, and those who live in the top floors of apartment buildings and who lack access to air conditioning, especially in large urban areas. Other vulnerable people are those who take medications that affect their body's thermoregulatory ability (Cohen and Miller [19]).

Significant qualitative and quantitative data suggest that men and women differ in their response to extreme thermal environments. Women sweat less, have a higher working metabolic rate, and have thicker subcutaneous fat that prevents them from cooling themselves as efficiently as men. Women, as a population, are therefore less tolerant of an imposed heat stress; acclimatization, body size, and cardiovascular fitness are important factors in determining tolerance.

In 1984, mean daily temperatures rose from 21.1°C in the preceding week to 28.9°C during a heat wave in New York. Throughout the extreme weather, non-institutionalized elderly, particularly women, were at highest risk of heat-associated death; among those aged 75-84 years, death rates rose 39% for men, and 66% for women; among those over 85 years old, increases were 13% for men and 55% for women (MMWR [38]).

More recently, a heat wave struck France in early August 2003 after warmer than average temperatures in June and July. Total excess mortality from1 August and 20 August was 14 802; in all age groups female mortality was 15-20% higher than male mortality (Duncan [10]).

Gender significantly affects the daily lives of women and men, before, during, and after an extreme event. In most societies, men tend to have greater access to key survival and recovery leaving women more vulnerable to natural events. Particularly vulnerable women include battered women/women at risk of violence, immigrant women/women with language barriers, indigenous women/minority women, isolated women/rural women, poor or low-income women, refugee women and the homeless, senior/frail women, women heading households/single mothers/widows, and women with disabilities (Duncan [10]).

Gender-specific health impacts of extreme weather events include: mental stress as a result of providing emotional care during and after the crisis, and increased violence (Duncan [5]).

Police reports of domestic violence in the seven months following the Mt. St. Helen's eruption (1980) increased by 46%. Following the 1993 Missouri floods, the average state turn-away rate at shelters rose 111%, programs sheltered 400% more flood-impacted women and children than anticipated, and in 1998, a Montreal Police Chief reported that one in four calls received during the 1997 ice storm were from abused women (Duncan [10]).

More recently, Simister and Cooper [39] demonstrated seasonal patterns in workplace events (e.g. 'quitting jobs' and strikes) and violent crime.

As discussed in section six, women are at increased risk during air pollution episodes. Increasing air temperatures (with possible heat-related morbidity, mental stress, and violence) may exacerbate the pollution health impacts on women.

Fortunately, heat-related health impacts can be reduced through behavioural adaptations, such as the use of air conditioners, increased intake of fluids, the development of community-wide heat emergency plans, and improved heat warning systems (Cohen and Miller [19]). But, these measures are often unavailable to women already at risk for air pollution.

8 Adaptation

Regrettably, those with the least resources—that is, the poorest segments of societies--have the least capacity to adapt and are the most vulnerable. Vulnerability is likely to be differentiated by gender.

Adaptation is necessary at all scales to complement climate change mitigation efforts. The ability to adapt and cope with climate change is a function of the ability to spread risk, equity, information, infrastructure, institutions, technology, and wealth.

According to Amartya Sen, 'The voice of women is critically important for the world's future—not just for women's future'. The equal participation of women is absolutely necessary to meet changing climatic conditions.

References

[1] Duncan, K. (Forthcoming). Environment and Health: Corporate Contributions to Our Common Future.
[2] World Health Organization. 2002. The World Health Report 2002. Reducing Risks, Promoting Healthy Life. WHO: Geneva.
[3] World Health Organization. 2004. Maternal Mortality in 2000: Estimates Developed by WHO, UNICEF, UNFPA. WHO.
[4] World Health Organization. 2005. Make Every Mother and Child Count: A Toolkit for Organizers of Activities. WHO/RHR/04.10.
[5] Duncan, K. 2004. Initial links: climate change, water, women, and health. Water and Climate Change: Knowledge for Better Adaptation. Canadian Water Resources Association Conference Proceedings.
[6] Division for the Advancement of Women. Department of Economic and Social Affairs. Nd. Fourth world conference on women. Platform for action. http://www.un.org/womenwatch/daw/beijing/platform/index.html. 29/07/06.
[7] Stipp, D. 1997. At risk: profits and jobs. Fortune December 8, 1997.
[8] Houghton, J., Ding, Y., Griggs, D. et al. (eds). 2001. Climate Change 2001: The Scientific Basis. Contribution of Working Group l to the Third Assessment Report of the Intergovernmental Panel on Climate Change (IPCC). Cambridge University Press: Cambridge.
[9] Climate Change Solutions. nd. Climate change solutions for small to medium-sized enterprises. http://www.climatechnagesolutions.com/emglish/sme/default.htm. 12/07/03.

[10] Duncan, K. 2006. Climate change, health, and women. Climate Change and Health Impacts Atlantic Conference 2006. St. John's, Newfoundland and Labrador, March 23-24, 2006.
[11] Women's Environment and Development Organization. 2004. Gender differences in water use and management. http://www.wedo.org/sus_dev/untapped3.htm. 14/05/04.
[12] World Health Organization. Nd. Water as a human right. http://www.who.int/water_sanitation_health/en/rtw1.pdf. 12/05/04.
[13] Jahnavi, T. 2003. India's village women fight to protect water. http://www.peopleandplanet.net/doc.php?id=1884. 14/05/04.
[14] Kumar, P. and Clark, M. 1987. Clinical Medicine. Bailliere Tindall: London.
[15] Meade, M., Florin, J. and Gesler, W. 2000. Medical Geography. Guildford: New York.
[16] Roehr, U. 2005. Gender and climate change. Why it makes a difference. http://unfccc.int/files/meetings/cop_11/climate_talk_series/application/pdf/cop11_kiosk_roehr-hemmati.pdf. 29/07/06.
[17] Houghton, J., Filho, L., Callender, B. et al. 1995. Climate Change 1995: The Science of Climate Change. Contribution of Working Group l to the Second Assessment Report of the Intergovernmental Panel on Climate Change (IPCC). Cambridge University Press: Cambridge.
[18] Intergovernmental Panel on Climate Change. 2006. Press information note. Preparations for fourth assessment report still underway. http://www.ipcc.ch/press/pr02052006.htm. 30/07/06.
[19] Cohen, S. and Miller, K. 2001. North America. In: McCarthy, J., Canziani, O., Leary, N. et al. Climate Change 2001: Impacts, Adaptation, and Vulnerability. Contribution of Working Group ll to the Third Assessment Report of the Intergovernmental Panel on Climate Change (IPCC). Cambridge University Press: Cambridge.
[20] McCarthy, J., Canziani, O., Leary, N. et al. 2001. Climate Change 2001: Impacts, Adaptation, and Vulnerability. Contribution of Working Group ll to the Third Assessment Report of the Intergovernmental Panel on Climate Change (IPCC). Cambridge University Press: Cambridge.
[21] McMichael, A. and Haines, A. 1997. Global climate change: the potential effects on health. British Medical Journal 315:805-9.
[22] Duncan, K., Guidotti, T., Cheng, W. et al. 1998. Health Sector. In The Canada Country Study: Climate Impacts and Adaptation. Koshida, G. and Avis, W. Environment Canada: Ottawa. pp. 501-590.
[23] World Health Organization. 2003. Climate Change and Human Health - Risks and Responses. WHO: Geneva.
[24] Mannion, A. 1997. Global Environmental Change. Longman Scientific & Technical: Harlow.
[25] Hidore, J. 1996. Global Environmental Change: Its Nature and Impact. Prentice Hall: Upper Saddle River.
[26] Ahrens, C. 2000. Meteorology Today: An Introduction to Weather, Climate, and the Environment. Thomson Learning: Pacific Grove.

[27] Dockery, D., Cunningham, J., and Damokosh, A. 1996. Health effects of acid aerosols on North American children: respiratory symptoms. Environmental Health Perspectives 104 (5): 500-505.
[28] Dhara, V. and Dhara, R. 2002. The Union Carbide disaster in Bhopal: a review of health effects. Environmental Health 57 (5): 391-404.
[29] United Nations Department of Economic and Social Affairs. 2002. Global challenge global opportunity: trends in sustainable development. http://www.johannesburgsummit.org/html/documents/summit_does/criticaltrends_1408.pdf. 19/10/02.
[30] Regional Office for Latin America and the Caribbean. 2000. GEO Report for Latin America and the Caribbean—Environmental Outlook 2000. UNEP/ROLAC: Mexico City.
[31] Stieb, D., Pengelly, L., and Arron, N. 1995. Health effects of air pollution in Canada: expert panel findings for the Canadian Smog Advisory Program. Canadian Respiratory Journal 2: 155-160.
[32] Chen, L., Knutsen, S., Shavlik, D. et al. 2005. The association between fatal coronary heart disease and ambient particulate air pollution: are females at greater risk? Environmental Health Perspectives 113 (12): 1723-9.
[33] Suzuki, D. 2002. Expanding the health care debate. Canadian Medical Association Journal 166(13): 1678 - 9.53.
[34] Patz, J. and Kahliq, M. 2002. Global climate change and health: challenges for future practitioners. Journal of American Medical Association 287(17):2283.
[35] Burnett R., Dales R., Krewski D. et al. 1995. Associations between ambient particulate sulfate and admissions to Ontario hospital for cardiac and respiratory diseases. American Journal of Epidemiology 142(1):15-22.
[36] Burnett R., Brook J., Smith Doiron, M et al. 1999. Effects of particulate and gaseous air pollution on cardiorespiratory hospitalizations. Archives of Environmental Health 54(2):130-9.
[37] Bransford, K., and Lai, J. 2002. Global climate change and air pollution: common origins with common solutions. Journal of American Medical Association 287(17):2285.
[38] MMWR. 1984. Heat-associated mortality—New York City. MMWR July 27, 1984: 33 (29):430-2.
[39] Simister, J. and Cooper, C. 2005. Thermal stress in the USA: effects on violence and employee behaviour. Stress and Health 21 (1): 3-15.

This page intentionally left blank

Visibility and air pollution

J. G. Watson & J. C. Chow
Desert Research Institute, Reno, NV, USA

Abstract

Haze is caused by the scattering and absorption of visibility light by particles and gases. Light is an electromagnetic wave, and just as a plane water wave is deflected by a barrier from its original direction, light waves are scattered when they encounter particles and gas molecules that are approximately the same size as the light's wavelength. The sky is blue because particle-free air also scatters light, but the gas molecules are so small that they scatter the shorter wavelength blue light more than they scatter the longer wavelength red light. Total extinction (b_{ext}), expressed in inverse megameters (Mm^{-1}), is multiplied by the distance between an observer and a target to indicate the amount of light removed from the site path by scattering and absorption by particles and gases, and is the most common measurement used to quantify visibility. Total extinction does not take into consideration light scattered into the sight path that further reduces the target's contrast (ratio of light transmitted from the target to the light transmitted from the horizon). Defining natural conditions is a scientific challenge. Annual average estimates are currently in use, but these will eventually need to be made more event-specific. Wildfires, dust storms, and other natural events will affect visibility on a case-by-case basis. Transport from outside of provincial and national boundaries needs to be considered because this is largely beyond the control of authorities.
Keywords: haze, visibility, PM.

1 Introduction

This paper provides an overview of urban and regional haze and its relationships to other pollutants. It describes the nature and causes of poor visibility, identifies relationships to fine particles and other pollutants, explains regulatory approaches to local and regional haze, and discuses relevance to other regional

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060621

and global pollution problems. It is based on recent reviews of regional haze [1–3] that provide greater detail and references to a larger array of published literature.

2 Causes of haze

Haze is caused by the scattering and absorption of visibility light by particles and gases. Light is an electromagnetic wave; just as a plane water wave is deflected by a barrier from its original direction, light waves are scattered when they encounter particles and gas molecules that are approximately the same size as the light's wavelength. The sky is blue because particle-free air also scatters light, but the gas molecules are so small that they scatter the shorter wavelength blue light more than they scatter the longer wavelength red light.

3 Measures of haze

Total extinction (b_{ext}), expressed in inverse megameters (Mm^{-1}), is multiplied by the distance between an observer and a target to indicate the amount of light removed from the site path by scattering and absorption from particles and gases. b_{ext} is the most common metric used to quantify visibility, but it does not consider light scattered into the sight path that further reduces the target's contrast (ratio of light transmitted from the target to the light transmitted from the horizon). The human eye is also sensitive to the spacing of objects, such as trees and rocky strata, and the observer may perceive texture better even as his or her position becomes more distant with respect to the target.

Visual range, which is approximately $4/b_{ext}$, is often used to express the maximum distance at which a target can be discriminated from its background. The deciview, $dv=10\ln(b_{ext}/10)$, has been adopted for tracking improvements in regional haze in the United States. Twenty-four hour averages of "dry" light scattering are often correlated with $PM_{2.5}$ concentrations.

The United States uses a chemical extinction budget to estimate the deciviews that track regional haze over long periods. Figure 1 shows the network of measurement locations that is used to obtain $PM_{2.5}$ mass and elemental compositions. Chemical extinction (b_{ext}) is estimated as:

$$b_{ext}\ (Mm^{-1}) = \Sigma \text{dry extinction efficiency } (m^2/g) \times \text{humidity modifier} \times \text{species concentration } (\mu g/m^3) = 3f(RH) \times (NH_4)_2SO_4 + 3f(RH) \times NH_4NO_3 + 4 \times \text{Organics} + 10 \times \text{Soot} + 1 \times \text{Soil} + 0.6 \times \text{Coarse Mass} + 10 \times \text{Clear Air Scattering}$$

where:

$f(RH)$ = extinction efficiency increase with RH.

The dry extinction efficiencies depend on particle size distributions and chemical compositions and can span a range of more than a factor of two for different distributions that might occur in the atmosphere. These efficiencies also depend on the complex mixing of particles with different compositions.

The soot fraction does not attain the 10 m^2/g dry scattering in the above formula, but this approximates what is found in nature. This occurs because the absorption efficiency of elemental carbon (or soot) is much higher when it is coated with a layer of sulfate or organic carbon than when it is present as a separate particle. There can be large discrepancies, and large uncertainties, in f(RH) when RH exceeds 90%. In spite of the uncertainties in creating chemical (reconstructed) extinction budgets, they reproduce sight path and point measurements fairly accurately for the 24-hour particle samples taken at U.S. parks and monuments, as shown by Chow et al. [1].

Figure 1: Sampling locations for the IMPROVE (Interagency Monitoring of PROtected Visual Environments) that are located in U.S. National Parks and Wilderness areas. Data from this and other PM networks are available at http://vista.cira.colostate.edu/views/.

The value of the chemical extinction is that it can focus control efforts on the chemical components, and their sources, that are the major causes of the poor visibility. Shenandoah, Great Smoky Mountain, and Acadia national parks in the eastern U.S. have poor visibility caused mostly by sulfate concentrations. Other parks in the western U.S. have better visibility, but it is more evenly distributed among a number of chemical components. The implication is that sulfur dioxide reductions will be the most effective controls in the east, while many sources will need to be targeted to improve western visibility. A large decrease in extinction may not cause a large change in the deciview metric that better represents how people will perceive just noticeable changes in haze. Nearly

30 Mm^{-1} corresponds to a one deciview change when comparing Great Smoky Mountains with Shenandoah, but only a few Mm^{-1} change is perceptible between Bryce Canyon and Denali.

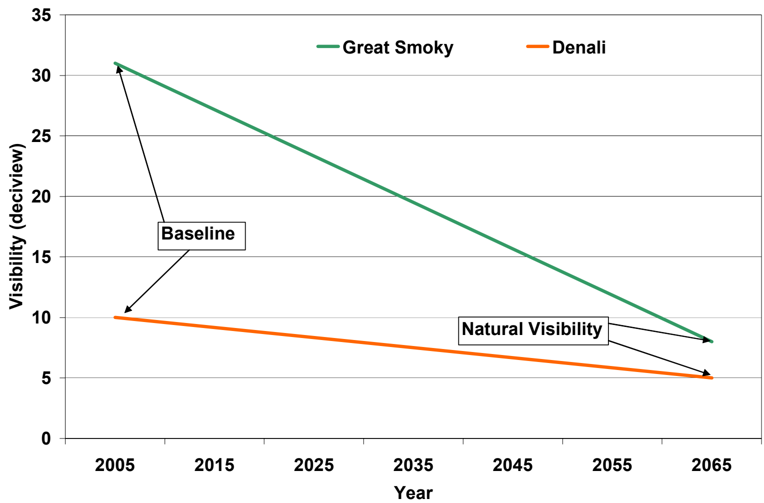


Figure 2: Example of visibility goals for Great Smoky Mountains and Denali National Parks in the U.S. The baseline is the average highest 20% extinction for 2000 through 2004. Natural conditions are derived from annual average estimates of non-anthropogenic carbon, soil, sulfate, and nitrate concentrations. Natural conditions differ among regions owing to different natural sources (e.g., biogenic aerosol formation) and relative humidity.

4 Improving visibility

Reducing regional haze requires emission reductions that cross local, provincial, and international boundaries. The U.S. has established five regional planning organizations with different states as members to track progress toward natural background levels at 156 national parks and wilderness areas with measurements shown in Figure 1. At each of these areas, chemical extinction will be tracked for the next 60 years relative to a baseline for the poorest 20% of the days established by measurements between 2000 and 2004. A linear glide path toward natural visibility conditions, illustrated in Figure 2, will be used to determine progress that will be evaluated at ten year intervals. There are differences in the rate of progress depending on how poor the initial visibility is and what are considered to be natural conditions for an area. Defining natural conditions is a scientific challenge. Annual average estimates are currently in use, but these will eventually need to be made more event-specific. Wildfires,

dust storms, and other natural events will affect visibility on a case-by-case basis. Transport from outside of the U.S. will also need to be considered because this is largely beyond the control of national authorities.

5 Conclusions

Among the most important topics that need further review and evaluation are: (1) practical methods for sampling and analyzing organic and elemental carbon; (2) scientific validity and practical requirements for integrating continuous particle monitoring technology into ongoing networks; (3) North American and global emissions from natural and anthropogenic sources; (4) satellite technology for tracking haze, emission sources, and pollution levels; (5) non-road emission source identification and estimation methods; (6) chemical markers for natural and anthropogenic sources; (7) air quality trend detection and tracking methods; and (8) integration of source and receptor models. Urban and regional haze is an important indicator of air pollution in many cities throughout the world. Quality of life and enjoyment of majestic vistas will improve only when serious emission reductions are undertaken to improve visibility. These have the added benefit of improving public health and reducing property damage due to excessive air pollution.

References

[1] Chow, J.C., Bachmann, J.D., Wierman, S.S.G., Mathai, C.V., Malm, W.C., White, W.H., Mueller, P.K., Kumar, N., & Watson, J.G., 2002 Critical review discussion - Visibility: Science and regulation. *J.Air Waste Manage.Assoc.*, **52(9)**, pp. 973-999, 2002.

[2] Watson, J.G. & Chow, J.C., Clear sky visibility as a challenge for society. *Annu.Rev.Energy Environ.*, **19**pp. 241-266, 1994.

[3] Watson, J.G., Visibility: Science and regulation. *J.Air Waste Manage.Assoc.*, **52(6)**, pp. 628-713, 2002.

This page intentionally left blank

The new millennium Ethics of housing technology

S. Mahmoud Issa & S. Ezzeldeen
Department of Architectural Engineering, Department of Basic Sciences, Menofia University, Egypt

Abstract

Before the turn of the century, information technology was spreading worldwide knowledge regarding global environmental changes and their expected implications on the health of the globe's inhabitants. New approaches to deal with these hazards were initiated, including the technological approach in which *Ethics* stresses human values, human responsibilities and the internal human aspects of our global community.

From an ethical point of view, human conduct and the actions of professionals, especially in the case of ozone depletion, should lead to the revival of nature, the prevention of pollution and the application of sustainability as far as development and production are concerned. This conforms to the principles of the Engineers' code of Ethics as set by ABET. As Housing Technology provides the indoor environment that accommodates inhabitants, this paper will investigate, through an analytical approach, the Ethics that should be applied in providing peoples' needs, given the environmental changes and sustainability concepts in development, production and housing.

The study will include: housing technology functions, indoor housing pollution, sick building syndrome, the role of technology in protecting human rights and the code of practice for professionals and the ethical principles it formulates.

Keywords: global environmental changes, housing technology, indoor pollution, sick building syndrome, ethics, Engineers' code of Ethics.

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060631

1 Introduction

Last century's technology advancement, known as Information Technology, paved the way for closer relations and faster communications between the corners of the globe.

By spreading the news that the "World Meteorological Organization observed a 45 % depletion of the ozone layer over one-third of the northern hemisphere, from Greenland to western Siberia, for several days during the winter of 1995-1996", it became clear that humanity has a pressing need to face the limited nature of its planet and should address the issues related to coping with the impact of environmental hazards to human health. Keeping pressure on the system within manageable limits necessitates various approaches to the problem that should have for the basis for an ethical directive and a distribution of responsibilities for the causes and effects of all the emissions that lead to environmental changes and ozone depletion.

2 Environmental hazards: causes and effects

There has been an increased recognition that environmental hazards to human health include climate change, stratospheric ozone depletion, loss of biodiversity as well as changes in hydrological systems and supplies of freshwater, land degradation and stress on food-producing systems [1]. Human activities, such as the burning of fossil fuels, resulted in the altering of the chemical composition of the atmosphere through the build-up of greenhouse gases and the release of atmospheric pollutants (such as chlorofluorocarbons), which deplete the lower ozone layer and increases human exposure to ultraviolet radiation, causing skin cancer and cataracts. Appreciation of this scale and of the type of influence on human health requires a new perspective, which focuses on ecosystems and on the recognition that the foundations of long-term good health in populations rely in great part on the continued stability and function of the biosphere's life-supporting systems. It also brings an appreciation of the complexity of the systems upon which we depend**.** And that led to the formulation of the "Montreal Protocol, which acts to reduce emissions of pollutants that weaken the ozone layer. Although this international agreement is proving highly effective in reducing risks in the long term, Ultra Violet radiation remains a health hazard" [1].

2.1 Environmental pollutants

Ozone is composed of three Oxygen atoms, and is unstable [2]. The formation of Ozone has been taking place for many millions of years and its stability is attributed to Nitrogen compounds in the atmosphere. Ozone forms in the atmosphere by the action of sunlight on oxygen. Ozone higher up in the atmosphere protects life on Earth [2]. Ozone lower in the atmosphere (troposphere) can be toxic and a pollutant. Chlorofluorocarbons (CFCs) were developed in the 1930s. Their primary uses were for industrial, commercial, and

household appliances. CFCs are non-toxic, non-flammable, and non-reactive with other compounds, and were first invented to replace dangerous sulphur dioxide and ammonia refrigerants. Most CFCs do not dissolve in water and therefore reach the atmosphere [2]. When released into the atmosphere, CFCs rise and are broken down by sunlight. Chlorine molecules interact with Ozone and destroy it.

2.1.1 Indoor pollutants

Studies have shown that the levels of some pollutants can reach high levels indoor, which was not previously known. Well-sealed buildings present protection (to some extent) from air pollution, depending on the type of pollutant. The following lists, some of the pollutants relate to air tight, sealed buildings and which are harmful to human beings:

Asbestos: It is now known that asbestos is a dangerous air pollutant. Asbestos is a fiber material, which has fine hair particles suspended in the air that cause lot of health problems if inhaled into to the lungs. Exposure to this substance for long periods of time results in a chronic disease known as asbestosis. This happens to people living around asbestos mines and near factories using this substance. Building materials containing friable (that can be crushed by hand) asbestos have a significant effect on pollution. Asbestos uses include fire-resistant materials as well as heat and sound insulation. It could also be used as a reinforcement material in cement pipes and in floors and ceiling products.

Formaldehyde: Formaldehyde exists in the compressed wood used in furniture and on walls, in medium density fiberboard and in insulating foam manufactured from urea-formaldehyde. The latter is chemically unstable, and can release free formaldehyde from the volatile formaldehyde that did not react and exists in the resin, and from the hydrolysis of the resins itself. The formaldehyde that does not react is responsible for the high level of formaldehyde indoors.

Volatile Organic compounds (VOCs): This is a category that contains a very large number of easily evaporated organic compounds. It includes hazardous air pollutants. To varying degrees, VOCs contribute to the formation of ground level photochemical ozone.

There are many sources of VOCs. Motor vehicle exhausts are a major and pervasive source, in addition to combustion engines that burn hydrocarbons inefficiently. There are also small sources of VOCs, including wood-paintings and organic solvent-based paints, freshly painted houses (chlordane and asbestos) [3], and mercury, chlordane and asbestos resulting from paintings and the use of acoustical tiles. Additionally, there are some building materials like gypsum, and places such as auto maintenance shops and bakeries, which emit large amounts of ethanol, contributing to smog (visible air pollution) formation [4]. Natural VOC emissions are not a problem by themselves but must be figured

into a strategy to reduce urban ozone formation [5]. VOCs contribute to health problems when sensitive individuals may react to those emissions with attacks of asthma, or other respiratory problems.

Measures to reduce those emissions in building materials implies the use of alternative paints including natural, hypoallergenic and low-biocide paints that have been developed to reduce those emissions [4]. The process of ventilation and storing gypsum boards separately could be applied to prevent absorption of these compounds by other surfaces or materials in the building's interior, as well as the use of recycled gypsum that would be free of volative organic compounds [4].

Radon: Radon is a radioactive material. In general, radioactive pollution results from natural sources or from human activity. Natural sources of radiation are radioactive gases that come out of the soil and rocks. Radiation caused by human activity arises when using weapons and in nuclear reactors used for the generation of energy and from the disposal of waste from these reactors.

The major source of radon comes from the soil, which constitutes 90% of the radon level in buildings. Other sources come from well water, natural gas and building materials.

3 Housing technology: functions, concerns and impacts

Acquiring adequate housing is a major issue for many individuals. The problems encountered to find a place to live that meets one's needs whilst still being affordable and in the preferred location vary according to one's ability and physical condition. This applies in the first place to elderly and disabled people whose additional problems include lack of accessibility and sometimes discriminatory attitudes in the place they have chosen to live.

Through technology and advances in design, housing needs should be satisfied without presenting risks or hazards to the users. Knowledge of innovative ways to build houses that can accommodate residents and visitors with disabilities has been available to professionals with the help of Assistive Technology i.e. as defined in the US "the broad range of devices, services, strategies and practices that are designed to increase the functional competencies of persons with disabilities" [6]. These include, but are not limited to, environmental control systems or "devices designed to permit the user to exercise greater control over devices and functions within their immediate vicinity" [6]. The control systems that the technology provides include: fire alarms, automatic control of entrances and of lighting at night in addition to the systems alarm for burglars at night and the simulation presence for maximal safety and security. That, in addition to processing or setting comfortable temperature through a system of "Automatic regulation of lighting and heating" [7]. The system also includes the use of Remote Control for service functions in order to achieve maximal comfort [7].

Advances in housing technology have led to refined methods of housing construction, which incorporate the advantages of materials testing with the

process of manufacturing the relevant parts in innovative systems. Newly-tested materials could provide structural superiority, improved thermal dynamics, no rotting and protection from pest problems. While conventional building methods and materials have registered incremental changes across time (the major choice for providing the close to nature internal environment) there have been certain issues which the housing industry, the second largest industry worldwide after the food industry, still considers of greater importance than adapting new techniques or using different innovative materials. These include, amongst others, the internal indoor environment and its effect on people's health and on their ability to perform their activities, given the use of highly effective insulation materials, as well as the sustainability of these products for future generations. While innovations and improvements have brought tremendous changes in housing quality, the use of building materials preventing toxic or harmful effects on its users is still a major concern for the housing industry.

"The selection of building materials greatly impacts the sustainability of a project. By choosing building materials wisely, such as considering the complete life cycle of the materials, a designer and/or builder can reduce the impact of the project on the environment. Careful material selection can minimize the depletion of resources, including raw materials, such as wood and metals, as well as energy and water used in the manufacturing process. It also can allow for efficient reuse or recycling of materials and building components if a building or facility is to be deconstructed or demolished. Finally, building materials choices may alleviate environmental impacts created by the manufacturing process, as well as indoor air quality in the final structure" [8].

"Analyzing the environmental impacts of building materials can be a difficult process". A life-cycle analysis, which accounts for the impacts of resource extraction through manufacturing, use, and disposal, involves extensive data collection and analysis and explains the impact of the energy used to do these needed processes. Some of this analysis has already been conducted and results are available through various lists and databases of environmentally preferable building products [8].

3.1 Environmentally preferable materials

The federal government of the USA defines environmentally preferable products (EPP) as goods that have a lesser or reduced effect on human health and the environment when compared to competing products that serve the same purpose. Environmentally preferable attributes include reduced toxicity, the use of recycled materials, and increased energy efficiency.

Building for Environment and Economic Sustainability (BEES) developed BEES software that measures the environmental performance of building products using an environmental life cycle approach. Twelve environmental impacts are now included in each product's environmental performance score: global warming, acidification, fossil fuel depletion, indoor air quality, habitat alteration, ozone depletion, water intake, criteria air pollutants, smog, ecological toxicity, and human health [9].

All the finishes and materials considered to be "Environmentally Preferable" have been formulated and manufactured with children, the elderly, and the chemically sensitive in mind, and were recommended for people with allergies, with sick building syndrome or multiple chemical sensitivity.

3.1.1 Multiple Chemical Sensitivity (MCS)

"Multiple Chemical Sensitivity (MCS) is the name given to the broad issue of reactions to specific or cumulative chemicals in the environment. One of the difficulties in classifying MCS as an illness has been the complex nature of chemicals in the environment and the interaction effects with and within the human body. The length of exposure, the concentration of the chemical(s), and the individual's threshold of resistance are also factors complicating a simple definition. In the relatively few but growing documented cases of severe reactions to chemicals, there seems to be no single stimuli or predictor of reactions" [10].

Table 1: Peoples' chemical sensitivity.

Chemical	Where found	Reactions
Formaldehyde	carpet, plywood, particle board insulation, adhesives	sore throat headaches
Pesticides	bug sprays, lawn chemicals, household cleaners	skin/eye/respiratory digestive tract/nervous system
Nitrogen dioxide	Unvented gas stoves/heaters Carbon monoxide gas appliances	respiratory problems headaches/dizziness
Solvents	household cleansers, paints strippers, gasoline	respiratory problems fatigue/dizziness
Latex	paints, gloves, caulking	allergic reactions
Dyes (especially dark blue)	clothing, curtain, tablecloths, napkins, rags, furniture	skin reaction

The previous table summarizes some chemicals to which people are most likely to have sensitivity.

3.1.2 Sick Building Syndrome (SBS)

Sick Building Syndrome "is a combination of ailments (a syndrome) associated with an individual's place of work (typically, but not always, an office building), although there have also been instances of SBS in residential buildings. A 1984 World Health Organization report into the syndrome suggested up to 30% of new and remodeled buildings worldwide may be linked to symptoms of SBS" [11].

Building occupants complain of symptoms such as: a) Headaches b) Eye, nose, or throat irritation c) Dry cough; dry or itchy skin d) Dizziness and nausea e) Difficulty concentrating f) Fatigue j) Sensitivity to odors.

For SBS to exist, these symptoms must disappear soon after the occupants go outside. The contributing factors often relate to the design of the built environment, and may include combinations of some or all of the following: a) indoor air pollution; b) artificial fragrance, such as dryer sheets c) limited access to natural sunlight; d) poor ventilation or heating; e) bad acoustics; f) poorly designed furniture; g) poor ergonomics; h) chemical contamination; i) biological contamination.

The solution to this problem lies in the following activities: pollutant source removal or modification: maintenance of "Heat Ventilation and Air conditioning HVAC" systems, replacement of water-stained ceiling tiles and carpeting, institution of smoking restrictions, venting containment source emissions to the outdoor storage and use of paints, adhesives, solvents, and pesticides in well ventilated areas, and use of these pollutant sources during periods of non-occupancy.

4 Indoor pollution abatement measures

The following measures are undertaken to reduce the impact of air pollutants inside buildings [12].

4.1 Choosing the appropriate building specification

4.1.1 Exclude the source of pollution

In newly-built houses or renovated ones the choice of building materials should exclude materials emitting pollutants, such as formaldehyde and medium density fibreboard, and use alternatives such as solid wood and decorative gypsum board. Covering floors with composite materials should be avoided and instead ceramic or wood should be used [13]. Air pollution by aliphatic and aromatic hydrocarbons and also by chlorinated hydrocarbons can be prevented by selecting the right paint finishes and furnishings. An example is the water-based paint that emits lower levels of volatile organic compounds than oil-based paints and varnishes.

4.1.2 Removal of the polluting source

This is an effective way to the abatement of formaldehyde. Asbestos is removed when it causes a pollution problem because it is used in some developing countries, although it is now forbidden due to the dangerous impact it has on health.

4.1.3 Choosing the building location

Air pollution due to radon can be avoided indoors by choosing the location of the building. The single most important factor, which enables radon to enter a building, is through the soil. For example, sand soil of high permeability constitutes a good medium for radon transfer, in contrast to muddy soil, which decreases radon penetration.

4.1.4 Climate control

Climate control can be used to decrease indoor pollutant levels. For example the levels of formaldehyde indoors are influenced greatly by the temperature and the relative humidity. This is done by decreasing the source of emission. Taking the science of chemical kinetics into consideration, we can find that temperature is the single most important factor in emitting volatile organic compounds from several construction materials and from some paints.

4.1.5 Source treatment

Modifications can be made to decrease the emission that causes the pollution. For example, materials containing asbestos can be sprayed by an appropriate material to achieve this goal. The cost of this operation is about 50% of the cost of removing asbestos.

4.1.6 Ventilation

Ventilation is used to decrease indoor pollution. This decreases the results from dilution and from the removal of pollutants. Methods involved include infiltration and exfiltration and natural ventilation, in addition to forced or mechanical ventilation. The first two methods constitute the methods for exchange of air in buildings. Forced ventilation and to some extent infiltration and exfiltration constitute the major methods by which air is exchanged in large offices and multi-storey buildings.

4.1.7 Air purification

Air purification or air cleaning is a procedure similar to ventilation since both involve decreasing the air pollutants. Air dust purification systems are more efficient than those used for gaseous pollutants. The principles of collecting dust are applied to a variety of particles regardless of their source, which is different to gaseous pollutants. Common techniques for collecting dust in the air are filtration and electrostatic precipitation.

5 Codes of engineering ethics

As is the case for other professions, engineers of different specializations attempted to formulate rules or standards of conduct in the form of codes of ethics [14]. Engineers in the USA accomplished this code, while others did not have the same opportunity and the situation has been left to one's conscience and to regulations set by the authorities. These codes protect the public and preserve the dignity of the profession. There is no single code of ethics for all engineering specializations [15]. However, they all agree upon certain issues as to what is known as ethical behavior. Therefore, there is a similarity among the codes formulated. One of the most recognized is the code of ethics published by the Accreditation Board for Engineering and Technology (ABET) [16]. The code consists of four fundamental principles and seven fundamental canons or authoritative rules. The fundamental principles state that engineers "uphold and advance the integrity, honour and dignity of the engineering profession by the

following: using their knowledge and skill for the enhancement of human welfare, being honest and impartial and serving with fidelity the public, their employers and clients, striving to increase the competence and prestige of the engineering profession, and supporting the professional and technical societies of their disciplines". These principles and canons constitute the basic rules of professional behavior, which can be applied to different scientific specialties such as the codes of ethics of chemists [17].

Since undertaking ethical decisions is sometimes difficult, more detailed guidelines have been prepared in order to be used with the fundamental canons of ethics.

6 Conclusion

The new millennium has been the beginning of a new era marked with new approaches for dealing with all the environmental problems that characterized the last century and resulted in damage to our planet from as a result of human activities.

While Information Technology has brought the awareness of the threats to human health to our attention, housing technology that provides the environment accommodating its users should fulfill its needs without endangering their health or depriving them from their rights to safety and security. The choice of building materials should be based on careful investigation of their properties, efficiency for reuse and impact and effects on people and the surrounding environment.

The role of housing technology's in allowing disabled and elderly people to benefit from the available assistive technology should be enforced as this allows the members of this group to reside in their own homes and participate in a social life.

As ethics is a major factor affecting professionals' conduct, which in turn reflects on other society members, there is a need for it to be enforced through the following recommendations:

1) Each profession should have a code of ethics to abide by in every country.
2) People all over the world should be encouraged to exchange their views and experiences in the field of ethics.
3) Examples of engineers who made good contributions to their own profession and followed the code of ethics that was formulated or even the common rules of ethics, if it has not been formulated, should be made known to other members of the profession.
4) Knowledge should be made available to everyone.

The following recommendations are presented for the housing sector:

1) The professional level of architects and engineers, such as chemical and civil engineers, should be raised, through awards for the best performance and ethical conduct.
2) Architects and engineers (such as the chemical and civil) should be encouraged to be aware of the latest innovations in the field and to experiment for new safer materials.

References

[1] WHO, Global environmental Change, http://www.who.int/globalchange/en/ 2006
[2] Michigan University, Ozone depletion http://www-personal.umich.edu/ 2006.
[3] Hewehy, M., "Air Quality in Work Environment (Case Study) in Office Buildings in Egypt", CDC 28th Annual Seminar, Cairo, December 1999.
[4] AIA - American Institute of Architects (2001), "Environment Resource Guide", at www.sustainableabc.com/sustaindev.html. 2004.
[5] Hill, M., Understanding Environmental Pollution" Cambridge Press UK.
[6] Allen, B. and Dillon, B., Environmental Control and Field Bus systems: A Study of field bus systems and their potential environmental control application for people with disabilities, Central Remedial Clinic, Dublin Ireland.
[7] Dewsbury, G., "Smart Thinking: Some thoughts on Designing a Smart House" available on line from www.gdewsbary.uykideas.com//smarthouse.html 2003
[8] *Navy WBDG "Green Product" Resource Page,* http://p2library.nfesc.navy.mil/P2_Opportunity_Handbook/13_7.html
[9] Minnesota Office of Environmental Assistance, Sustainable Bldg Products http://www.buildinggreen.com/auth/article.cfm?fileName=040101a.xml
[10] Ohio State Univ. Fact Sheet, //www. ohioline.osu.edu/cd-fact/0192.html
[11] Godish, T., Air Quality, 3rd edition, Lewis Publications, 1997.
[12] Wikipedia the free Encyclopedia, June 2006.
[13] ElHaggar, M., Elkady, M., Ezzeldeen, S., The Complete Guide in Air Pollution and control Technologies, Dar Elfekr Elarabi, 2003 (in Arabic).
[14] www.masterscieng.sunyab.edu/ethics/index.html
[15] www.masterscieng.sunyab.edu/ethics/abet.html
[16] Paul Wright Introduction to Engineering, 2nd Ed., John Wiley & Sons 1994.
[17] www.masterscieng.sunyab.edu/ethics/chemi.html

Impact of greenhouse gas stabilization initiatives on the Croatian petroleum industry

L. Maurovic
University of Zagreb,
Faculty of Mining, Geology and Petroleum Engineering,
Department of Petroleum Engineering, Croatia

Abstract

Even though the impact of greenhouse gas (GHG) emissions on global climate began at the end of 19th century it took more than half a century for scientists to start addressing this problem. Since the impact was global, variable and long term, for a long time there was no initiative of countries or emission intensive industries to control and accept responsibility for GHG emissions. Today the greatest part of the responsibility for adverse environmental impact due to GHG emissions lies with industry and the energy sector. The petroleum industry is one of the GHG emission intensive industries. Beside direct CO_2 and CH_4 emissions with oil and gas production, the petroleum industry is also characterized with high-energy intensity and final products with high emission potentials. Besides facing the same problems as other petroleum industries due to new high ecological standards, the Croatian petroleum industry is also faced with a lack of domestic legal regulative addressing stabilization of GHG emissions and, at the same time, increasing international directives for adoption of international environmental regulations. Croatia has signed and ratified several international agreements considering GHG emission stabilization. The UNFCCC was ratified in 1996 and in 1999 Croatia signed the Kyoto Protocol by which it has committed to reduce GHG emissions by 5% from 1990 levels. Since the admission to the EU is one of main political and economical goals of Croatia the ratification of the Kyoto Protocol in Croatia is not questionable. Due to the wide implementation of GHG emissions limitation laws there is a necessity to incorporate climate change factors into corporative strategies of petroleum companies in order to reduce corporate risks and to assure long-term competitive advantages.

Keywords: climate change, greenhouse emission, Croatian petroleum industry.

1 Introduction

With the beginning of the industrial revolution at the end of 19th century and growth of consumption of all kinds of fuels the negative environmental impact of humankind rapidly started to increase. Even though the impact of greenhouse gases on global climate system has begun during the end of the 19th century the scientific addressing of this problem didn't begin until second half of 20th century. Since the impact was global, variable and long-term for a long time nor countries nor companies took any initiative to control and to accept the responsibility for greenhouse gases emission.

Nowadays the highest share of the responsibility for negative environmental impact, as on local so on regional and global scale, is upon industry and energy sector. Environmental impacts of energy sector, if analyzing from the energy producers to final energy consumers, are various. Most certainly, presently the most important and the most analyzed energy sector environmental impact is the problem of greenhouse gases emission with consequential air pollution, especially in urban areas, acid rain problem, high ozone concentrations and most of all, global CO_2 problem. Rest of the impacts is usually local and therefore could be solved, prevented or put to acceptable levels with improvement of technical and technological performances. The emissions from energy sector are mostly due to fossil fuel combustion and they include carbon dioxide (CO_2), carbon monoxide (CO), sulphur dioxide (SO_2), nitro oxides (NOx) and solid compounds like soot. Beside their negative impact on health of living creatures, SO_2 and NOx are also known as "acid" gases because the products of some chemical reaction of SO_2 and NOx are characterised with high pH values. These products are sedimented from atmosphere as wet (acid rains) and dry deposits. Beside that, nitro oxides with some volatile organic compounds (VOC) are part of ozone (O_3) forming processes by whose concentrations the greenhouse effect is even more intensified. The greenhouse gas mostly responsible for the mentioned greenhouse effect is carbon dioxide which concentrations have increased by 30% from the beginning of the 19th century and currently amounts over 370 parts per million (ppm). Getting out to the atmosphere, mostly by fossil fuel combustion, carbon dioxide is present there for almost 120 years. Due to present predictions, if there are no appropriate steps towards stabilization and decrease of global CO_2 emissions, worlds CO_2 concentrations will, in next 50 years, climb up to warning and disturbing level of 550 ppm [1].

With the development of global ecological consciousness during the 1970s of the 20th century, after the recognition of global warming problem and global climate change as its consequence, worlds leading economies started to take more resolute steps towards decrease of greenhouse gases. In 1988 the United Nations Organization formed the International Panel on Climate Change. After that, in 1992, followed the Rio Earth Summit, the international summit in Rio de Janeiro, that represented first public acknowledgment of the need for greenhouse gases emissions reduction and climate change mitigation. The result of the summit was the United Nations Framework Convention on Climate Change. The United Nations Framework Convention on Climate Change (UNFCCC), ratified

by over 180 nations, came into force in 1994 with long-term goal to achieve "stabilization of greenhouse gas concentrations in the atmosphere at a level that would prevent dangerous anthropogenic interference with the climate system" [2]. The first step towards achieving this goal was set of national emission reduction targets prescribed by the Kyoto Protocol. Kyoto Protocol came into force on February 16th 2005 and became a legal obligation for 128 world countries, which have signed and ratified it. Due to reduction of global costs of achievement of proposed emission reductions there were three flexible mechanisms presented by Kyoto Protocol respectively Joint Implementation (JI), Clean Development Mechanisms (CDM) and Emission Trade (ET) [3].

2 Greenhouse gases emissions in Croatia

In 1996 the Republic of Croatia ratified and became fully responsible member of UNFCCC. By ratifying UNFCCC Croatia committed itself to systematic recording of greenhouse gases emissions of which the IPCC is informed in annual greenhouse gas emissions report, and has also accepted the obligation to keep its greenhouse gases emission under 1990s levels. Beside UNFCCC, measuring and recording of greenhouse gases emissions in Croatia is also obligatory due to national Air Protection Law, which came into force in 1990. Beside UNFCCC Croatia has also signed several other, for national energetics important, international settlements. One of them is the Convention on Long – Range Transboundary Air Pollution (LRTAP) of the United Nations Economic Commission for Europe (UNECE), which resulted with several protocols. That is how Croatia accepted the Protocol on Further Reduction of Sulphur emissions, according to which Croatia has to reduce its sulphur emissions by 22% until 2010 in relation to 1980s levels. That means that the levels of SO_2 emissions should be kept under 117 kilotonnes. In 1999 under LRTAP Convention came out also The Protocol to abate acidification, eutrophication and ground – level ozone (MPME Protocol), which Croatia still didn't ratified. By MPEM Protocol, beside SO_2, the limitations of NOx, Non-methane volatile organic compounds (NMVOC) and amonium (NH3) are also set. By its ratification, which will probably be one of the conditions when joining the European Union, Croatia will be obligated to keep its NOx emissions on 1990s levels, to reduce its SO_2 emissions by 61%, NMVOC by 14% and its NH_3 emissions by 19% until 2010 compared to the same emissions at 1990s levels.

Croatia also signed and ratified Wienna Convention on ozone layer protection, Montreal Protocol on ozone polluting compounds, Heavy metal Protocol and Protocol on organic polluters.

In March of 1999 Croatia has signed the Kyoto Protocol by which it is committed to reduce its greenhouse gases emission by 5% in period from 2008 to 2012 compared to base year 1990. By choosing 1990 for a base year Croatian Ministry of Foreign Affairs has chosen a bad term for Croatia since the emissions were extremely low in 1990 first of all due to the war in the country, but also because of electricity import and high share of renewable sources (hydropower) in domestic electricity production which is one of the most

emission intensive sectors. There is also a problem with energy capacities, which were in possession of the Republic of Croatia in 1999 but were situated on territory of other republics of former Yugoslavia. These capacities were taken into consideration when setting the 1990's emission levels for Croatia. Today they are no longer in Croatian possession. Croatia still didn't ratify the Kyoto Protocol, but since joining to European Union is one of the main goals of Croatia, its ratification is not questionable. Croatia appealed to the UNFCCC for higher emission quotas in the base year. The negotiations are still running [4]. The CO_2-eq emissions from energy industries in Croatia are shown in figure 1 [5].

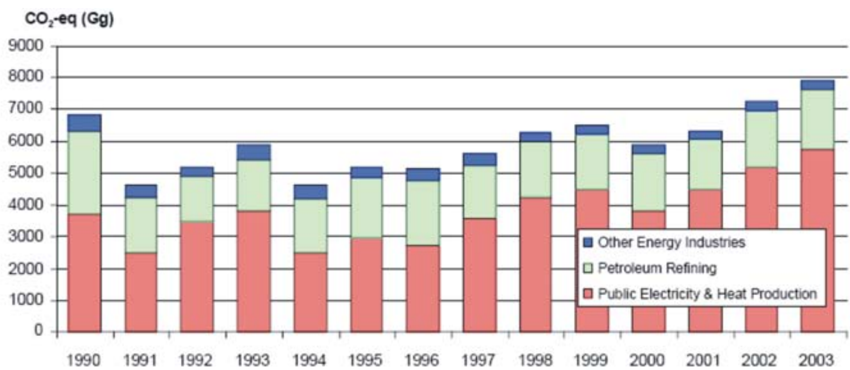


Figure 1: The CO_2-eq emissions from energy industries in Croatia.

Calculations of SO_2 and NOx emissions in Croatia are done by EMEP/ CORINAIR methodology, while calculation of CO_2 emissions is carried out by using IPCC methodology. The combustion emissions have the highest share in overall CO_2, SO_2 and NOx emissions in Croatia (figure 2) [6]. The SO_2 emissions are mainly due to fuel combustion in industrial energy plant and thermal power plants. The decreased levels of SO_2 emissions in last five years are result of emission reductions in power sector due to more favourable structure and better quality of used fuels. The NOx emission is largely the result of fuel combustion in road transport, and depends on type and age of vehicles and technological solution for emission reduction. In 2004 there was a decrease in total CO_2 emissions mainly because favourable hydrometerological conditions and consequently more intensive hydro power plants performance and reduced generation form thermal power plants, but the emissions were still 1% above the commitment defined by the Kyoto Protocol. In addition it is expected that the energy demand in Croatia will rise in next few years so the Kyoto Protocol target will be easily exceeded. The expected difficulties in meeting the Kyoto Protocol targets are, as already mentioned, caused by very low CO_2 emissions of the base year, so Croatia will have to make additional actions in form of subsidies, regulations, laws and environmental policy in general to meet the obligated greenhouse gases emissions targets [6].

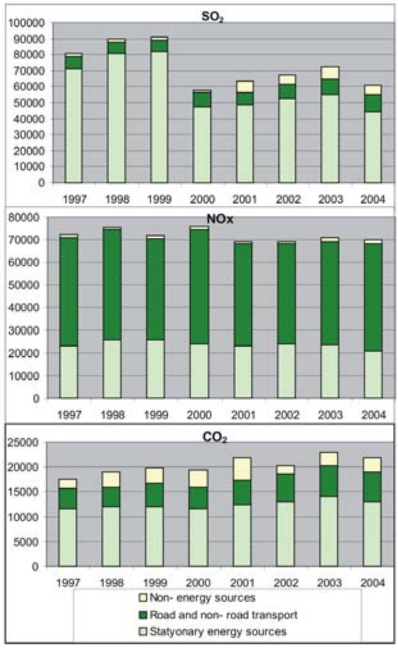


Figure 2: Emission trends in Croatia - SO_2, NOx and CO_2.

3 Croatian petroleum industry and greenhouse gases emissions

Reviewing presently used resources for energy production it is obvious that fossil fuels still have the highest share. Oil, natural gas and coal represent around 90% of commercial energy used worldwide. According to current predictions, considering today's rate of energy consumption and technical and technological development, fossil fuel reserves are expected to last for at least one more century.

Petroleum industry is highly emission intensive. Not only that there is a direct carbon dioxide (CO_2) and methane (CH_4) emissions during oil and natural gas production, but petroleum industry is also characterized with high-energy intensity and with final products of high emission potentials. The fugitive emissions of methane from oil and gas activities in Croatia are shown in figure 2 [5]. Fossil fuels are still dominant sources of energy supply, so to meet the energy demand petroleum industry today is called upon not just to provide an effective management of oil and gas reserves, but also to manage to do that in safe and efficient manner with as low as practically possible ecological footprint.

Since petroleum industry, regarding its business activities, which are considered to be environmentally intensive, takes up special public attention great emphasis has been put on quality of environmental management system based on effective management with systematic approaches and adequate

mechanisms, which have to ensure compliance with present, as national as the global, environmental policies and regulations. Principles of environmental protection were initiated and became embedded in petroleum business for more than half of a century, but it is only few decades ago that environmental protection started to be seen as an integral part of sustainable development approach. Task of petroleum company's environmental management system is not only to provide improvement, necessary to cope with new environmental standards, technological development and to meet public expectations, but also to control and supervise that policy statements, guidelines, programs and field procedures are properly conducted. Environmental management techniques have successfully integrated into all segments of upstream operational procedures and activities, from exploration work, through drilling operations, field development, production of oil or natural gas and ending by abandonment. With effective environmental management system it is ensured that during all of these activities preventive environmental and safety measures are taken even before starting operations. Some consider the high environmental standards and regulations and its implementation to be the main principle that underpins petroleum industry's business and decision - making process.

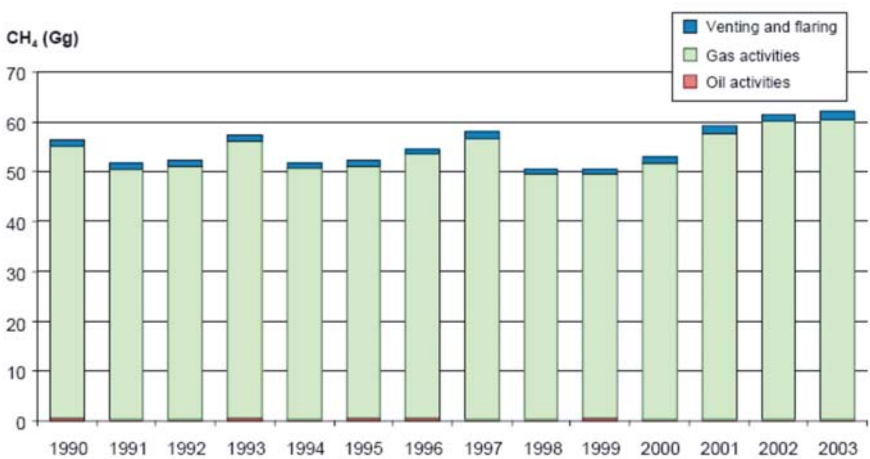


Figure 3: The fugitive emissions of methane from oil and gas activities.

When talking about environmental protection in petroleum business there are several issues that should be addressed such as biodiversity, climate change, oil spill prevention, waste and water management. To support the needs of ever increasing world development, energy demand will, in next few decades, still strongly depend on oil and gas resources. Consequently, in order to fulfil the worldwide increasing demands, petroleum companies have to develop and adopt new technologies to increase recovery and to spread their upstream activities to unexploited areas.

The biggest player in Croatian petroleum industry is INA. INA is medium-sized, vertically integrated petroleum company specialized in oil and natural gas exploration and production, refining and marketing of petroleum products. It owns all of the country's oil and natural gas reserves and refining capacities and it is also responsible for natural gas imports to Croatia. The rest of the petroleum business activities are conducted by independent companies operating under umbrella of INA Group. INA's subsidiaries are engaged in LPG business, natural gas transportation and providing integrated oilfield services. Until 2000, when it was partially privatized, INA was fully state owned company. At the moment, MOL, the Hungarian petroleum company, as INA's strategic partner, owns 25% + 1 share of INA's equities. The rest is in possession of the Croatian government.

INA operates on 43 onshore oil fields, 11 onshore gas fields, 2 offshore gas fields (Adriatic sea), and owns and operates two fuels refineries located in Rijeka and Sisak. Considering greenhouse gases emissions INA is one of the companies that will be affected with the implementation of greenhouse gases stabilization regulations. Among INA's capacities the biggest emitters are gas-condensate field Molve and Rijeka and Sisak refinery. The fugitive emissions of ozone precursors and SO_2 from oil refining in Croatia are shown in table 1 [5].

Table 1: The fugitive emissions of ozone precursors and SO_2 from oil refining in Croatia.

	Emissions (Gg)					
	1990	1995	2000	2001	2002	2003
CO	0.62	0.49	0.47	0.44	0.44	0.44
NOx	0.41	0.33	0.32	0.29	0.30	0.29
NMVOC	4.25	3.37	3.26	3.04	3.05	3.02
SO_2	6.38	5.06	4.90	4.57	4.58	4.53

The Rijeka refinery is medium sized refinery located on the Adriatic coast with access to deep-sea port and JANAF pipeline system. The refinery typically runs 3–3.5 million tonnes per year of crude oil, producing a range of petroleum products for the domestic and export markets.

The Sisak refinery is located 50 km from Zagreb, the capital of Croatia and also the biggest consumption area in Croatia. The refinery runs 2–2.2 million tonnes per year of crude oil produced by INA plus Russian crude oil imported via connections to Druzba 1 and Druzba 2 pipelines.

The gas-condensate field Molve is situated in north-western part of Croatia in most onshore gas productive part of Croatia. It has extremely unfavourable reservoir conditions with pressures up to 1100 bar (16 100 PSI), temperatures up to 235°C, 25% CO_2, 3% SO_2 and significant values of mercury in produced natural gas and condensate [7]. Yearly emissions on gas- condensate field Molve are shown in table 2.

Croatian petroleum industry will be affected by greenhouse gases emission stabilization in country on several levels. Since, even with additional measures,

Croatia won't be able to fulfil its obligations towards Kyoto Protocol, and since it is not likely that EU Commission will approve increase in emission quotas for the base year, Croatia will have to find the solution for its problem of greenhouse gases emissions in Kyoto Protocol flexible mechanisms. Considering present economy status, and since it is one of the Annex I parties, it is most likely that in few years Croatia will take part in Joined Implementation projects as a host country. Through these projects Croatia will gain needed foreign capital, efficient technologies and finally, decrease in greenhouse gases emissions.

Table 2: Emissions on gas-condensate field Molve.

	Emissions (10^6g)					
	1990	1995	2000	2001	2002	2003
CO_2	233.6793	391.5306	310.0001	348.6376	350.6254	353.2544
H_2S	0,149	0.103	0.058	0.065	0.028	0.007065

Croatia's participation in EU Trading System will probably pose costs of €60 million per year. This will also mean that the EU emission trading directives should be interpolated into domestic laws and that Croatia will have to develop national laws for national emission trading system regulation. That is expensive and long-term process that has partially started with the energy and environmental law reforms, but finalization of this process is not in sight in the near future. Another problem that rises up is the development of effective and qualitative monitoring system for continuous monitoring of stationary and mobile emission sources. Only few of Croatian companies have installed systems for continuous emission monitoring according to EU standards, which development and installation will bring additional financial investments to the companies.

The development of national emission trading system in Croatia is questionable for several reasons. One of the reasons is emission allowances allocation that will be done by the government under National Allocation Plan. Making of National Allocation Plan is one of the most important and one of the most controversial as political as economical issues. By making it the government will be directly affecting companies business and their market competitiveness. By gaining a small number of emission allowances the company is forced to make additional investments to achieve bigger emission reductions in order to keep its emissions within obligated levels. Another problem with the development of national emission trading system in Croatia is the fact that many of the potential emission allowance sellers will have to achieve significant emission reductions (in order to create tradable quotas) before the actual emission allowance buyers even exist, That will lead to less number of market transactions than expected. And finally, there is also a question whether the domestic companies are financially strong enough to enter international emission markets.

Croatian petroleum industry is taking steps today to limit greenhouse gas emissions from their own operations through the development and implementation of cleaner and energy efficient technologies. INA's activities to

reduce its climate change impact include investment in cogeneration facilities, research and development of renewable and environment friendly refinery products, reduction and, where possible, elimination of venting and flaring and capture and sequestration of CO_2. INA is also putting a lot of effort to motivate their customers to reduce their own emissions through increasing energy efficiency and use of environment friendly products through several marketing projects. There is also the necessity for INA to start to incorporate climate change factors into its corporative strategy in order to minor corporate risks and to assure long-term competitive advantages.

4 Conclusion

Today the greatest part of the responsibility for adverse environmental impact due to GHG emissions, as on local so on regional and global scale, lies on industry and energy sector. Petroleum industry is one of GHG emission intensive industries. Beside direct CO_2 and CH_4 emissions with oil and gas production, petroleum industry is also characterized with high-energy intensity and final products with high emission potentials.

In March of 1999 Croatia has signed the Kyoto Protocol by which it is committed to reduce its greenhouse gases emission by 5% in period from 2008 to 2012 compared to base year 1990. By choosing 1990 for a base year Croatian Foreign Ministry has chosen a bad term for Croatia since the emissions were extremely low in 1990. Croatia still didn't ratify the Kyoto Protocol, but since joining to European Union is one of the main goals of Croatia, its ratification is not questionable. It is expected that energy demand in Croatia will rise in next few years so the Kyoto Protocol target will be easily exceeded.

Besides facing the same problems as other petroleum industries due to new high ecological standards, Croatian petroleum industry is also faced with lack of domestic legal regulative addressing stabilization of GHG emissions and, at the same time, increasing international directives for adoption of international environmental regulations. Only few of Croatian companies have installed systems for continuous emission monitoring according to EU standards, which development and installation will bring additional financial investments to the petroleum company. There is also the necessity for INA to start to incorporate climate change factors into its corporative strategy in order to minor corporate risks and to assure long-term competitive advantages.

References

[1] http://cdiac.esd.orln.gov
[2] United Nations Framework Convention on Climate Change, http://unfccc.int
[3] The Kyoto Protocol, http://unfccc.int
[4] www.mzopu.hr

[5] ENRG- Energy Research and Environmental Protection Institute, *Croatia's National Inventory Report for the period 1990 – 2003*, Zagreb, 2005
[6] Republic of Croatia, Ministry of Economy, Labour and Entrepreneurship, *Energy in Croatia 2004*, Annual Energy Report, Zagreb, 2006.
[7] www.ina.hr
[8] Energy Institute "Hrvoje Pozar", *Energy Sector Development Strategy*, NN (38/02), Zagreb, 2002.

Prediction of road traffic noise attenuation due to distance, ground absorption and gradient

H. N. Rajakumara[1] & R. M. Mahalinge Gowda[2]
[1]*Department of Civil Engineering, Adhiyamaan College of Engineering, Tamilnadu State, India*
[2]*P.E.S. College of Engineering, Karnataka State, India*

Abstract

The propagation of road traffic noise is also influenced by the ground absorption, distance of source from receiver and road gradient. Hence it is necessary to apply corrections to the measured traffic noise levels. In this study a simple correction model has been developed for ground absorption, the distance and road gradient based on the field data. The method of calculating the ground absorption consists of identifying the locations which are totally or even partially absorbent in nature like backyards with gardens, paddy fields, tree density on either side of the roads and a mixture of absorbent and non-absorbent areas. Field measurements are taken from different locations with a varying percentage of soft ground cover. The correction value for ground absorption is estimated from the difference between the calculated noise levels and measured noise levels at varying percentages of soft ground cover. The correction value for distance is estimated from the difference between the calculated noise levels and measured noise levels at different horizontal and vertical distances over the hard ground surfaces. Similarly, the correction value for road gradient is calculated from the difference between the calculated noise levels and measured noise levels at different road gradients.
Keywords: ground absorption, distance, road gradient.

1 Introduction

The estimation procedure for noise from road traffic can be usually divided into two parts. The first part is concerned with the evaluation of the source noise level in terms of traffic parameters and the second with the evaluation of the effect of

WIT Transactions on Ecology and the Environment, Vol 99,
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060651

the surrounding on the propagation of the noise. If the ground surface between the edge of the near side carriage way of the road segment and the reception point is totally or even partially absorbent in nature (e.g. grass land, cultivated fields or plantations) an additional correction for ground cover, often referred to as ground absorption correction, must be taken into account [1]. The correction factor is progressive with distance and has a significant impact on reception points nearest to the ground. In open areas the measured traffic noise level in the study area is also influenced by distance between the source and reception point. It can also often produce changes in received noise level of a magnitude similar to those expected from difference in ground cover as mentioned above [2]. For ideal point sources it is inversely proportional to the distance squared, while intensity due to line source is inversely proportional to the distance from the source [3]. Therefore, it is necessary to apply necessary correction to the measured noise levels for distances greater than the reference distance from the centre line of roadway. This correction is often referred to as distance correction. Noise emitted by traffic flow is also influenced by the gradient of the road segment. Generally, gradient cause an increase in traffic noise for upward flows and a decrease in noise for downhill flows [4]. Therefore, a correction should be applied to the predicted values of noise levels when gradient exists.

It is obviously desirable to develop some method which is more sensitive to the distribution of ground cover, distance and gradient, but it is equally important to retain the simplicity of application of the prediction method, in view of the large number of variables which can be introduced, to maintain its usefulness. In this paper a sound attenuation model is developed for distance, ground absorption and gradient of roadway. These models are incorporated in a road traffic simulation model to determine the L_{eq} levels from a steam of vehicles.

2 Prediction formula

2.1 Correction model for distance (ΔL_D)

The basic concept followed in this study to develop the distance correction model is given as follows.

$$\Delta L_D = \alpha * 10 * \log (Do/D), \tag{1}$$

where

α is a site parameter whose value depends on site conditions.

D is the equivalent distance from the road segment to the reception point, m.

$= \sqrt{(D_N \times D_F)}$

D_N is the perpendicular distance from the receiver to the centerline of the nearest lane, m

D_F is the perpendicular distance from the receiver to the centerline of the far lane, m.

D_0 = Reference distance = 7.5 m

The correction value for distance is estimated from a linear equation of the difference between the calculated noise levels and measured noise levels at

distances of 10 m, 15 m, 20 m, 30 m, 40 m, 50 m and 60 m from centre line of roadway at the receiver height of 1.2 m, 3 m, 5 m, 7 m, 10 m and 15 m from the local ground surface. Field measurements are taken from 12 locations in outer ring roads of Bangalore city with varying horizontal and vertical distances. The locations chosen are flat ground with no shielding effect and ground absorption. The relation between the attenuation value of traffic noise and distance at every receiver height is shown in Figure 1. It may be observed that the horizontal distance has more influence on noise attenuation values when the distance is less than 60 m and it is almost remains uniform when the horizontal distance exceeds 60 m. Further, the height of receiver is also has significant impact on the noise attenuation values (see Figure 2). However, the attenuation values are remains intact with the horizontal distance when the receiver height reaches 15 m.

Based on the attenuation values of traffic noise as a function of horizontal distance and receiver height, the value of site parameter (α) in equation (1) is calculated as follows.

$$\alpha = \begin{cases} 1.0 : if\ H_r \leq 3m \\ 0.9*(1-H_r/15) : if\ 3m < H_r < 15m \\ 0,\ Otherwise \end{cases} \tag{2}$$

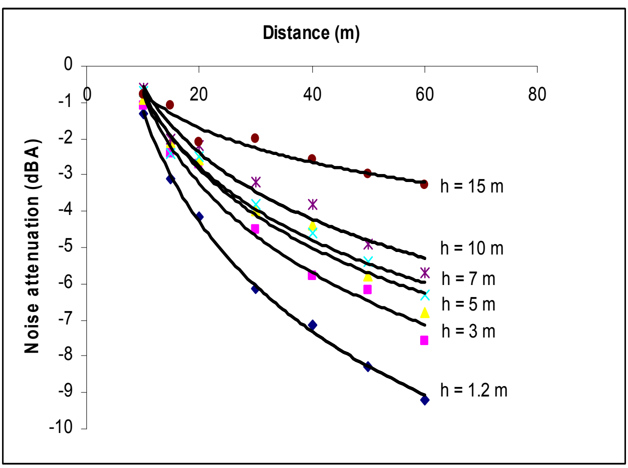


Figure 1: Attenuation of traffic noise as a function of horizontal distance and receiver height when there is no soft ground cover.

2.2 Correction model for ground absorption (ΔL_{GA})

The method of calculating the ground absorption consists of identifying the locations with totally or even partially absorbent in nature, (e.g. grass land, cultivated fields or plantations). In this study, field measurements are taken from 17 locations from in and around Bangalore city with soft ground cover (P) of 25%, 50% and 75%. The locations chosen are backyards with gardens, paddy

fields, tree density on either side of the roads and mixture of absorbent (grass cover) and non-absorbent areas. The site locations are classified into hard or soft based on the criteria given in FHWA table [1]. All measurements are carried out during normal weather conditions and the effects of wind speed and direction, and temperature stratification are neglected.

The basic concept followed in this study to develop the absorption correction model is given as follows.

$$\Delta L_{GA} = \beta * 10 * \log (Do/D), \tag{3}$$

where β is a site parameter whose value depends on site conditions.

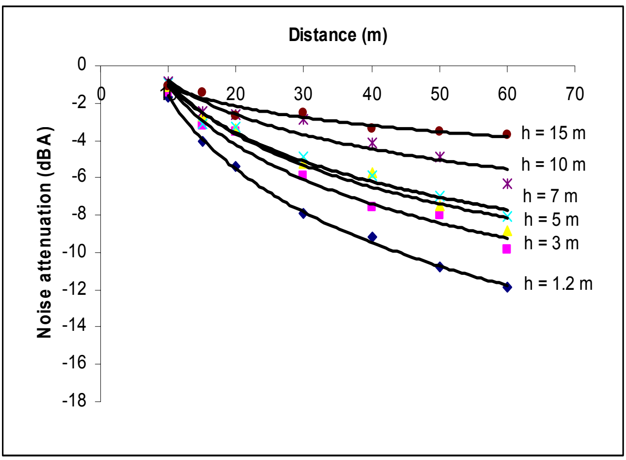


Figure 2: Attenuation of traffic noise as a function of horizontal and receiver height when P (percentage of soft ground cover) = 25%.

The correction value for ground absorption is estimated from the difference between the calculated noise levels and measured noise levels when soft ground cover is equal to 25%, 50% and 75% at the receiver height of 1.2 m, 3 m, 5 m, 7 m, 10 m, 15 m and 18 m, and the horizontal distances of 10 m, 15 m, 20 m, 30 m, 40 m, 50 m and 60 m from centre line of the roadway. The height of source for all types of vehicles is taken as zero. The relation between attenuation value of traffic noise and the percentage of soft ground cover as function of horizontal distance and receiver height is shown in Figure 2 to 4. According to field measurements, it may be observed that the attenuation of traffic noise due to ground absorption is more when the soft ground cover exceeds 50%. Further, the ground absorption is also progressive with horizontal distance and receiver height. However, it remains intact when, H_r (receiver height), reaches 15 m.

Based on results of the study the value of site parameter (β) in the equation (3) is calculated as follows.

$$\beta = \begin{cases} 1.0 * (1 + P/100) : \mathit{if}\ H_r \le 3m \\ 1.2 * (1 + P/100) * (1 - H_r/15) : \mathit{if}\ 3m < H_r < 15m \\ 0,\ Otherwise \end{cases} \qquad (4)$$

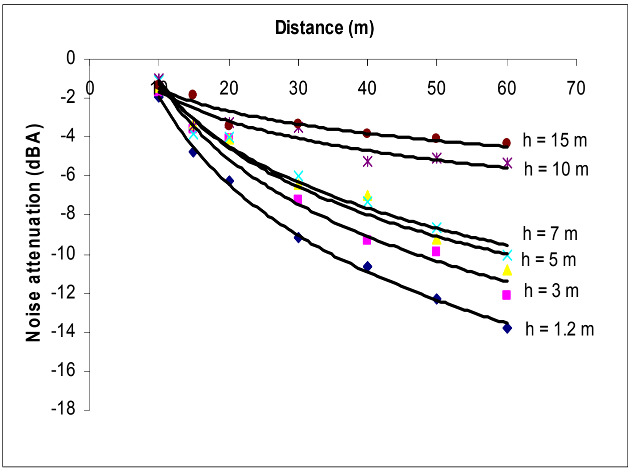


Figure 3: Attenuation of traffic noise as a function of horizontal and receiver height when P (percentage of soft ground cover) = 50%.

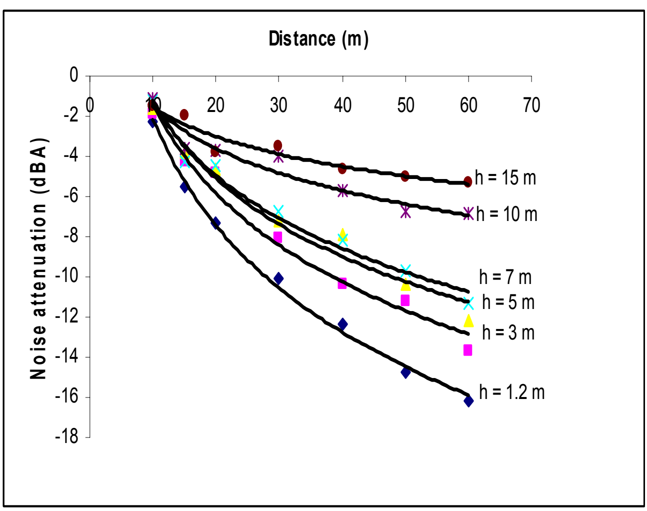


Figure 4: Attenuation of traffic noise as a function of horizontal and receiver height when P (percentage of soft ground cover) = 75%.

2.3 Correction model for road gradient

Field measurements are performed at 14 locations having different gradients ranging from 1.5 – 10.5% on urban road network of Bangalore city, State Highway no.17 and National Highway no.7. The gradient of road surface is measured using Cylone Ghat Tracer. The field study indicated that the gradient above 5% has significant impact on noise attenuation values.

Based on results of the study and considering the road gradient in terms of percentage gradient, G, the increase in noise due to upward gradients can be expresses as 0.40 G, 0.38 G, 0.63 G, 0.45 G, 0.42 G and 0.72 G for Two Wheelers, Car/Jeep/Van, Bus, LCV/Minibus, Autorickshaw and Trucks respectively. Similarly, 0.36 G, 0.32 G, 0.53 G, 0.40 G, 0.38 G and 0.65 G corresponding to decreased noise levels for downhill flow for Two Wheelers, Car/Jeep/Van, Bus, LCV/Minibus, Autorickshaw and Trucks respectively. The necessary traffic noise correction model for roadway segments with gradient can takes the following form:

For Uphill flow:

$$\Delta L_{UG} = \frac{(0.72Q_{Truck} + 0.63Q_{Bus} + 0.45Q_{Lcv/MB} + 0.42\,Q_{Auto} + 0.40\,Q_{Tw} + 0.38Q_{Car/Jeep/Van})*G}{(Q_{Truck} + Q_{Bus} + Q_{LCV/MB} + Q_{Auto} + Q_{TW} + Q_{Car/Jeep/Van})} \tag{5}$$

For Downhill flow:

$$\Delta L_{DG} = \frac{-\,(0.65Q_{Truck} + 0.53Q_{Bus} + 0.40Q_{Lcv/MB} + 0.38\,Q_{Auto} + 0.36\,Q_{Tw} + 0.32Q_{Car/Jeep/Van})*G}{(Q_{Truck} + Q_{Bus} + Q_{LCV/MB} + Q_{Auto} + Q_{TW} + Q_{Car/Jeep/Van})} \tag{6}$$

where
ΔL_{UG} = Correction for Up gradient in dBA.
ΔL_{DG} = Correction for Down gradient in dBA.

3 Conclusions

The outdoor propagation of road traffic noise is influenced by many factors. The important ones are ground absorption, the distance between source and receiver and road gradient. In this study a simple attenuation model is developed for ground absorption, the distance between source and receiver and road gradient. The results of the study indicated that the ground absorption is progressive with distance when the percentage of soft ground cover exceeds 50%. Similarly, it may be observed that the horizontal distance has more influence on noise attenuation values when the distance is less than 60 m from the source and it is almost remains uniform when it exceeds 60 m. Further, the height of receiver is also has significant impact on the noise attenuation values. However, the

attenuation values are remains intact with the horizontal distance when the receiver height reaches 15 m. The results of road gradient effect indicated that the gradient above 5% has significant impact on noise attenuation values.

Acknowledgements

The first author is gratefully acknowledge the financial assistance provided by All India Council for Technical Education, New Delhi, under the scheme Career Award for Young Teachers and the authorities of Adhiyamaan College of Engineering, Hosur, T.N. and P.E.S.College of Engineering, Mandya, Karnataka, for providing necessary facilities and support.

References

[1] Jain, S.S., Parida, M & Bhattacharya, C.C., Development of Comprehensive Highway Noise Prediction Model for Indian Conditions. Journal of Indian Road Congress, 62 (3), pp. 453A - 488a, 2001.
[2] Nirjar, R.S, Jain, S.S., Parida, M and Katiyar, V.S., Study of Transport Related Noise Pollution in Delhi. Journal of Institution of Engineers (I), 84, p.6-15, 2003.
[3] Hother Sall, D.C and Chandler Wilde, S.N., Prediction of the Attenuation of Road Traffic Noise with Distance. Journal of Sound and Vibration, 115 (3), pp.459 – 72, 1987.
[4] Bengang, Li., Shu – Tao & Dawson, W., A GIS based Road Traffic Noise Prediction Model. Applied Acoustics, 63(6), pp.679 – 91, 2002
[5] Harris C, M., Handbook of Noise Control; McGraw Hill, USA, 1979.
[6] Yamamoto and Yamashita, M.A., Simple Model for Estimating Excess Attenuation of Road Traffic Noise. Journal of the Acoustical Society of Japan, 8 (1), pp.13 -32, 1987.
[7] LF Cohn and G R Meroy, Environmental Analysis of Transportation Systems, John Wiley & Sons, New York.

This page intentionally left blank

Section 9
Soil

This page intentionally left blank

Arsenic pollution in the southwest of Tuscany: monitoring of Cornia catchment basin

F. Rossi[1], A. Donati[1], M. Rustici[2], B. Rugani[1] & E. Tiezzi[1]
[1]*Department of Chemical and Biosystems Sciences, University of Siena, Italy*
[2]*Department of Chemistry, University of Sassari, Italy*

Abstract

The territory of Colline Metallifere, in SW Tuscany, is characterized by the presence of strong arsenic anomalies. Some hypotheses, formulated in the last 20 years, based on geological and mineralogical factors have failed to explain the peculiar distribution of this toxic element in soil, fluvial sediments and ground water. Our research group has been studying for four years the problem of arsenic pollution in this district to investigate the origin and the mechanism of As diffusion in the environment. In particular we started a comparative study based on the extensive sampling of the stream sediments of the main waterways of Colline Metallifere (Pecora, Bruna and Cornia and their tributaries). In this work we focused our attention on the stream sediments and the soil of the Cornia basin. The comparative analysis of Cornia and the other rivers of the area, together with the investigation of historical series of data, provided important information about the origin of contamination of the territory and the impact of ancient mining settlements on this phenomenon.
Keywords: arsenic pollution, Tuscany, geochemical investigation, elemental analysis, anthropogenic activity.

1 Introduction

Arsenic is a naturally occurring element, with an average abundance that range from 1 to 10 mg/kg (dried soil) in the earth crust ranking as the 20th abundant element [1].

www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060661

In a growing number of cases the concentration of this metalloid has been found much higher than average all around the world, and most of the time these anomalies were due to anthropogenic activities. The presence of As in the environment at high levels combined with its well known toxicity, has generated a large public concern and a widespread interest in the scientific community and government agencies for health and environment. In order to understand the diffusion routes of Arsenic in the environment and its biogeochemical cycle, a large number of publications on its chemistry in soil and ground water have been produced [2].

The present paper is a part of a wider project that has been started in order to understand the origin of the strong Arsenic contamination in the Scarlino plain in the southwest of Tuscany [3-6]. In this area, a partially filled wetland, close to the sea, a power plant for sulfuric acid production from pyrite roasting was installed in 1962. The plant used the ores found in the mine-field (Colline Metallifere), far about 15 km west from this site. In the last 20 years it has been clear that a large area around the industrial site was strongly contaminated (up to 1290 ppm in the soil and 2700 ppb in the ground water) by Arsenic compounds [7].

However, several chemists and geologists negate anthropogenic origins for these anomalies and formulated another alternative hypothesis which is related with the natural accumulation of the element via fluvial sediments and via underground hydrothermal water fluxes transported from mineralized area [8].

In previous papers [5,6] we described the results of an extensive sampling of fluvial sediments collected from two very similar rivers and their tributaries, which flow almost parallel from the "Colline Metallifere" towards the coastal wetlands: the Pecora River and the Bruna River. From a comparative study of these two environments we were able to establish the extent of Arsenic natural background. Moreover we accumulated strong evidences of the importance of ancient mining and smelter sites in the formation of the environmental geochemistry as it is now.

In this paper we present the analysis of stream sediments and soil samples collected along the Cornia catchment basin, the third most important river of the area. Cornia shares similar geo-morphologic characteristics with the other two rivers, in particular it springs and streams in the same area and along its course it passes through some sites interesting from a geochemical point of view like spas or Tin mines. Moreover all three rivers end in geographically similar areas, even if Scarlino plain (Pecora's mouth), as mentioned above, is highly industrialized conversely to Diaccia Botrona (Bruna's mouth), which is a natural park and Bottegone marsh (Cornia's mouth), which is partly a WWF oasis but it also hosts a thermoelectric power plant.

The territory of Scarlino, considered together with the large area of the "Colline Metallifere", and other uncontaminated coastal wetlands which have same origin, presents indeed some unique features that could help to understand the general mechanisms of Arsenic diffusion. A deeper understanding of Arsenic mobility in particular in wetlands environment and the relationship with its bioavailability, is one of the recent most intriguing topics of scientific debate.

2 Materials and methods

In our campaigns we collected more than 200 samples including stream sediments, soils and waters. In this work we focus our attention on the analysis of 71 samples of stream sediment and 53 samples of soil, collected in five sampling cycles during a period spanning from March to October 2004; for every sample [As], [Cd], [Cr], [Pb] and [Mn] were recorded. The complete location of the samples is shown in fig. 1.

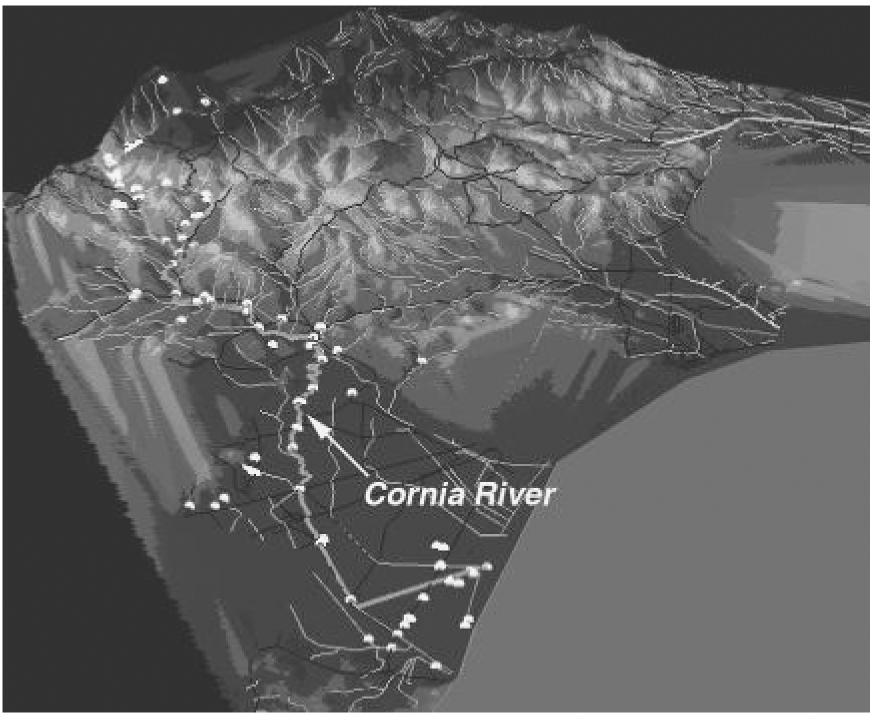


Figure 1: Study area. Dots indicate sampling points.

Samples of sediment were collected along the course of the Cornia River together with samples coming from its tributaries.

Dried stream sediments and soils were treated as follows: the fraction of 150 μm of the dried matter was homogenized and stored in sample stocks of about 100 g. Sub-samples of about 150 mg were subsequently weighted and digested with a mix of fluoridric, percloric, nitric and hydrochloric acid in a microwave oven. Pollutant concentrations were determined by "graphite furnace – atomic absorption spectroscopy" (GF-AAS). In the 30% of the sample stocks, three sub-samples were extracted and in the 10% of the sample stocks, six sub-samples were extracted. For these sample stocks pollutant concentrations were given as the average of the concentration of the sub-samples. The difference of the

measured values from the average is less than 10% in all the considered cases. The following standards were used for the instrument calibration: 1000 mg/l of arsenic acid (As), 1000 mg/l of cadmium nitrate (Cd), 1000 mg/l of chromium nitrate, 1000 mg/l of lead nitrate (Pb), 1000 mg/l of manganese nitrate (Mn), all dissolved in a 0,5 M solution of nitric acid. All the acids used were of ultra-pure grade (Merk), all other chemicals were of analytical grade (Merk). Data are reported in ppm, which correspond to mg of an element in 1 kg of dried matter.

3 Results

In this section results of the analysis will be presented for every element. Only anomalous or particularly interesting values will be reported and commented for every element.

3.1 Cadmium and manganese

Cadmium concentration was very low in all samples analyzed; in most of the cases it was under the instrument detection limits, and in the other cases it was under the law limits, both in stream sediments and top soils. About manganese Italian law does not provide upper limits for its presence in water, soil or stream sediment. Its concentration was about 800 ppm along the entire course of the river and in the terrains nearby, except for few spots in correspondence of the "Campigliese" mining zone, in which we found peaks of 1400-1600 ppm.

3.2 Lead

Italian law provides a lead concentration limit of 100 ppm in stream sediments and 1000 ppm in soils (for commercial and industrial purposes). About stream sediments only one anomalous value was found in "Botro Colombo" stream, a tributary of Cornia River. This high value (791.1 ppm) was probably due to the fact that "Botro Colombo" drains the archeominerary zone of Montioni. In all other samples a mean [Pb] of 30 ppm was found.

Top soil analysis revealed high tenors of lead in the zone of "San Silvestro" Archeominerary Park, in particular [Pb] ranging from about 1200 ppm to about 6500 ppm. Data recorded for lead concentration stand in line with the mean values of other archeominerary areas or industrial sites present in "Colline metallifere". Figure 2 shows trends of [Pb] along the river course.

3.3 Chromium

[Cr] tenors in Cornia (and tributaries) stream sediments are generally low. Exceptions to this behaviour were found in correspondence of hydrothermal springs and former mining districts. In particular, [Cr] ranges from 1560 to 4584 ppm in "Fosso Corniaccia" stream, which drains Venturina hydrothermal zone.

Regarding top soils, we found a distribution similar to stream sediments: in Venturina soils and near former tin mine "Sales" [Cr] ranges from 1400 to 1700ppm.

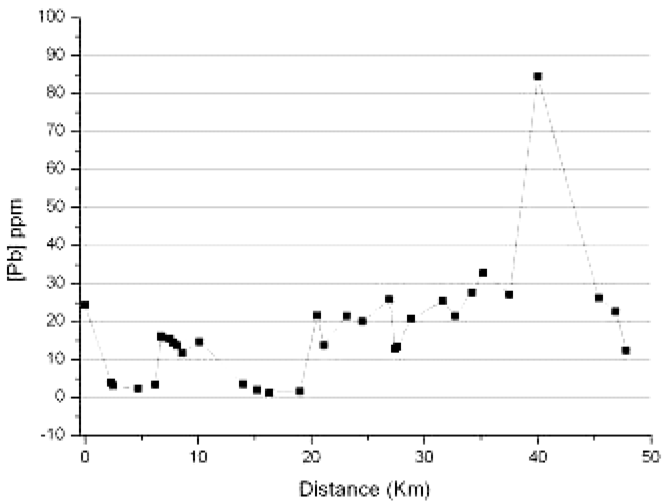


Figure 2: [Pb] trend in stream sediments of Cornia River as a function of the distance from the spring. [Pb] tenor is low along the whole course of the river.

High [Cr] values in "Venturina" area are not explicable only in terms of hydrothermal or mining influences and need further investigations. Figure 3 shows the location of the anomalies, in the inset it is reported [Cr] trend along the main course of Cornia. It is evident from the figure that contamination of tributaries does not influence the pollution level of the main river.

3.4 Arsenic

Figures 4 and 5 show results for [As] analysis in upper and lower Cornia valley. The arsenic content is showed graphically with vertical bars for every sampling point. From figures it is evident that Cornia has generally low [As] tenors along the main course, but localized anomalies were found in its tributaries and in the soils nearby. Major anomalies are present in the upper valley, where [As] reaches 3327 ppm in "Mulino Nuovo" stream and 2047 ppm in the soils near "San Pompeo" power plant. These extremely high values are explicable in terms of hydrothermal water fluxes: "Mulino Nuovo" stream receives water from a hydrothermal spring, which is part of the Hellenistic-roman "Sasso Pisano" archeological complex; soils in the zone of San Pompeo are highly contaminated by the presence of geothermal power plant. It is interesting to note that anomalies are punctual and circumscribed to the pollutant sources. In fact stream sediments samples taken afterwards pollution sources present low level of [As] and generally, as in the case of chromium, [As] in Cornia river main stream is not influenced by the presence of punctual pollution sources.

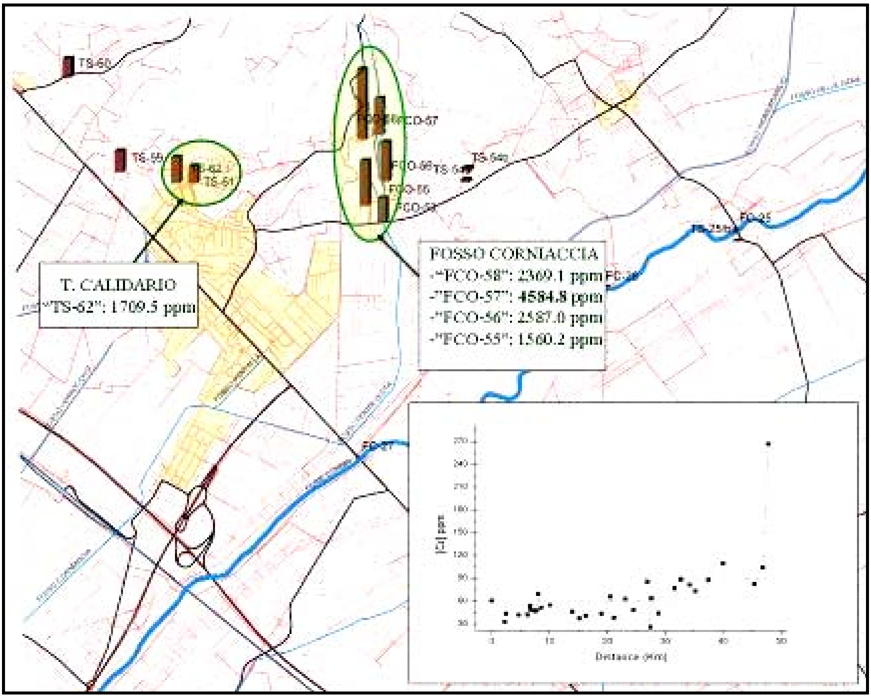


Figure 3: Figure shows the locations of the strong Cr anomalies. The inset reports [Cr] trend along the Cornia River course.

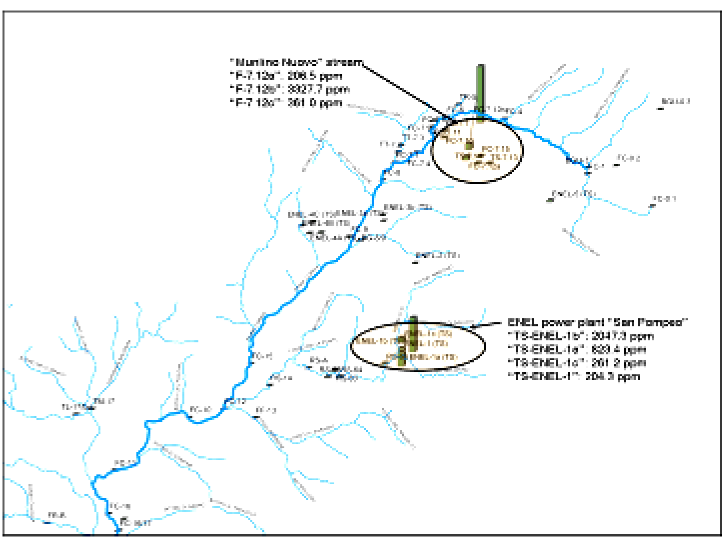


Figure 4: Upper Cornia valley. Circlets highlight anomalies area.

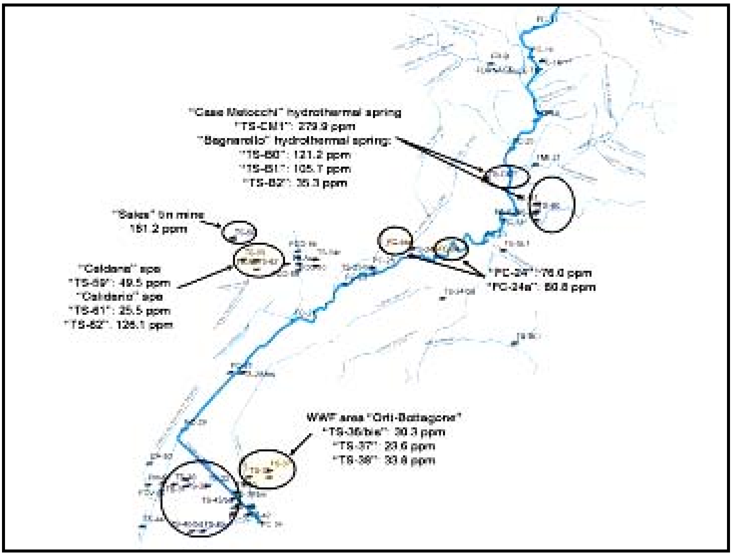


Figure 5: Lower Cornia valley. Circlets highlight anomalies area.

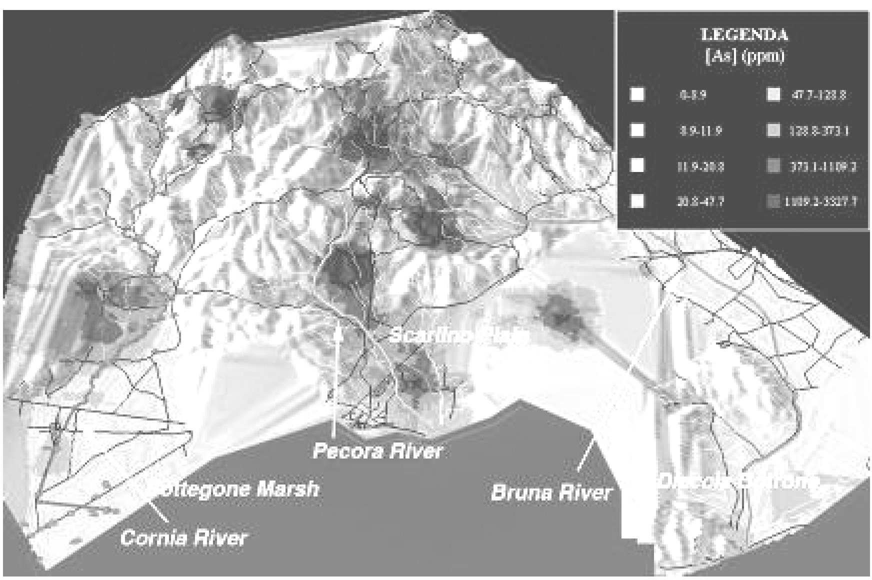


Figure 6: GIS map of the Colline Metallifere area showing concentrations of arsenic in soil and sediments. Darker areas indicate high As concentrations.

4 Conclusions

Data collected over the last years, including data presented in this paper, allowed us to build a geochemical map for Arsenic dispersion in Colline Metallifere area (Fig. 6). The arsenic grid was generated by interpolation of more than 10000 analytical data by an ordinary kriging method.

It is also interesting to compare [As] trend for Bruna and Pecora rivers with results for Cornia. From Figure 7 it is evident that the behaviour of [As] is completely different for the three rivers. Pecora and Bruna shows opposite trends and the highest [As] values are located in correspondence with ancient and modern anthropogenic activities [5,6]. On the Contrary, Cornia River, which flows in a less industrialized area, always presents low arsenic levels.

Arsenic dispersion map, jointly with results for other elements, puts in evidence that natural pollutants levels are significantly lower than the values recorded in the human exploited sites. Moreover wetlands in which the Bruna and Pecora rivers end, did not present Arsenic and other metal anomalies as in the case of the highly industrialized Scarlino plain. Finally, we want to point out that hydrothermal fluxes actually cause high intensity pollution (see the case of Cr and As), but polluted zones are mostly punctual and limited to and it doesn't seem that hydrothermal fluxes are the main vehicles of pollutants dispersion in the area.

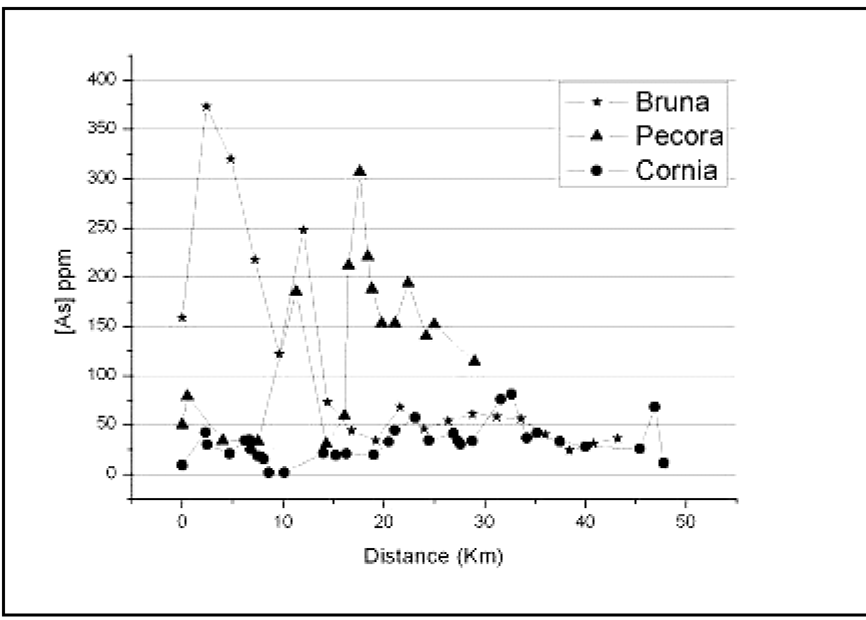


Figure 7: Comparison between [As] trends along the courses of the three main rivers of Colline Metallifere area.

References

[1] Adriano D.C., *Trace Elements in the Terrestrial Environment*. Springer-Verlag, New York: 1986.

[2] a) Cullen W.R. and Reimer K. J. Arsenic speciation in the environment. *Chem. Rev.* **89**, 713-764 (1989); b) Ahman D., Roberts A.L., Krumholz L.R., Morel, F.M.N., *Nature*, **371**, 750, (1994). c) Oremland, S.R., Stolz, J.F. The ecology of Arsenic. *Science*, **300**, 939-944, 2003.

[3] Donati A., Tiezzi E., Marchettini N. *On the Arsenic anomaly in the Scarlino county*, technical report for the remediation project. 2002.

[4] Picchi, M.P., Fugaro L., Donati A. in Brownfield sites: Assessment, Rehabilitation & Development, WIT Press, pp. 121-130 2002.

[5] Donati A., Protano G., Riccobono F., Dallai L., Francovich R., Rossi F. and Tiezzi E., *Influence of ancient mining settlements on arsenic pollution in the southwest of Tuscany*. In Brownfield Sites II, WIT press, pp 183-191, 2004.

[6] Donati A., Pulselli F.M., Riccobono F., Dallai L., Francovich R. and Tiezzi E., *Ann. Chim.*, **95(3-4)**, pp. 161-166, 2005.

[7] a) AQUATER s.p.a. report – *Study for the building of a repository for hematite cinders in the Scarlino Plain*. 1985; b) ARPAT - Regional Agency for Protection of the Environment - report: *Anomalies of Arsenic level in the Scarlino plain: a preliminary study*. 2001.

[8] Costagliela, P., Benvenuti, M., Benvenuti, M.G., Innocenti, A., Mascaro, L., Paolieri, M., Rossato, L. and Tanelli, G., *Arsenic Distribution in the Quaternary Sediments of the Median Valley of the Pecora Stream (Grosseto, Italy)*, In Brownfield Sites II, WIT press, pp 201 - 209, 2004.

Soil contamination and land subsidence raise concern in the Venice watershed, Italy

L. Carbognin[1], G. Gambolati[2], M. Putti[2], F. Rizzetto[1], P. Teatini[2] & L. Tosi[1]
[1]*Institute of Marine Sciences, National Research Council, Venice, Italy*
[2]*Department of Mathematical Methods and Models or Scientific Applications, University of Padua, Italy*

Abstract

The southern catchment of the Venice watershed is threatened by shallow aquifer salinization and anthropogenic land subsidence due primarily to the microbial oxidation of organic soils that outcrop in the coastal farmland reclaimed from the Adriatic Sea over the last century. Recent hydrogeological and geophysical surveys provide documentary evidence that saltwater intrusion may extend inshore up to 20 km away from the Adriatic coastline with the contaminant plume from near ground surface down to 100 m depth in some areas. The actual salt distribution is the outcome of a number of factors, including the ground elevation markedly below the mean sea level (down to -4 m locally), the seawater encroachment along the final 10-15 km of the regional watercourses (Brenta, Bacchiglione, Adige, Gorzone), and the drainage practices implemented in the reclaimed area. The fresh-salt water interface is generally between 2 and 30 m deep and exhibits a pronounced seasonal variation. At the same time an ongoing settlement due to peat oxidation promoted by farming activities is observed in most of the area south of the Venice Lagoon that was reclaimed from 1892 to 1967 and is rich in organic matter. Overall land subsidence over the last 70 years ranges between 1.5-2 m and is still in progress at a rate of 1.5-2 cm/y. As a major result a large fraction of the reclaimed land lies below the mean sea level with an increasing exposure to flooding during severe winter storms and saltwater intrusion from the Adriatic Sea, the nearby Venice Lagoon, and the river beds that locally lie above the surrounding ground surface. To mitigate both hazards the implementation of a drainage strategy of the reclaimed area intended to maintain the water table as high as possible would be required. This could decelerate the peat oxidation (hence the related land subsidence) and oppose the inland subsurface salt convection and dispersion. Moreover the design of mobile gates at a few river mouths (e.g. the Brenta river) could create an effective barrier against the seawater migration upstream the watercourses in the hottest and driest summers.
Keywords: saltwater intrusion, land subsidence, peat oxidation, soil salinization.

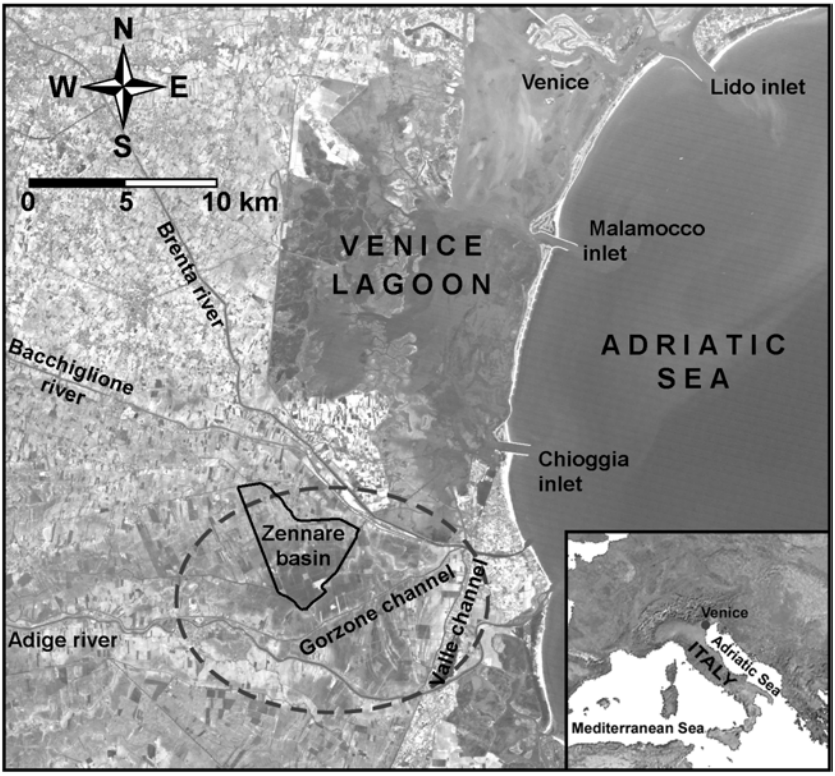


Figure 1: Map of the southern catchment of the Venice Lagoon with the location of the Zennare basin. The area characterized by outcropping organic soils (dashed line) and the major watercourses are shown.

1 Introduction

The Venice watershed is located in the eastern part of the Po river plain, Italy (Figure 1), and includes a very precarious coastal environment subject to both natural and anthropogenic changes. A significant and economically important fraction of the watershed is at present below the mean sea level while the reduced sediment supply to the Adriatic is impacting appreciably on the coastal morphodynamics. Venice herself is the crossroad of a number of critical problems including anthropogenic land subsidence, sea level rise, periodic flooding during severe winter storms, and lagoon water contamination of urban, industrial and agricultural origin.

Land settlement in the watershed is contributed by three different sources [1]: deep regional tectonics and natural compaction of Quaternary sediments, geomechanical compaction caused by subsurface water pumping, and biochemical oxidation of outcropping soils rich in organic matter promoted by

farming activity (e.g., ploughing). While geomechanical subsidence in Venice seems to have come to a halt because of the well shutdown [2], natural subsidence still continues at a rate of 0.5-0.7 mm/y [3] and geochemical subsidence in the southernmost part of the area is currently under way at an alarming rate of 1.5-2 cm/y [4,5]. At the same time the sea level is expected to rise over the present century by more than the amount (≅11 cm) experienced over the last 100 years [6]. The combined effect of sea level rise and land subsidence has enhanced the intrusion of salt water and the related soil salinization in the easternmost part of the Venice watershed.

Starting from the 70's it became evident that saline intrusion affected a large fraction of the Venice coastland. In the late 90's the problem was addressed in much detail in relation to the multiaquifer system (50-350 m) pumping that was the major cause of the Venice geomechanical subsidence in the 60's and early 70's. A risk analysis was performed with the aid of numerical models for several groundwater pumping scenarios pointing to a negligible danger for the deepest aquifer units [7]. In the hydrogeologic context of the Venice coastland, the largest risk of saltwater contamination is concerned with the upper aquifers, especially in the southernmost area, because of the geomorphological setting of the coastal plain south of the Venice Lagoon [8]. Because of the serious environmental and socio-economic impacts induced by groundwater salinization, a project was undertaken in 2000 to identify area and depth involved by the presence of salinized soil that can enhance the risk of soil desertification and compromise the agricultural practices [9].

A recent research topic is concerned with the geochemical subsidence caused in the southernmost part of the Venice watershed by mass loss due to oxidation of the organic soil fraction in response to drainage for farming. Consistent with rates from 1 up to 10 cm/yr as reported worldwide from similar processes [10], peatland settlement of the area in question amounts to 1.5-2 m over the last 70 years with a current rate of 1.5-2 cm/y, and more subsidence expected to occur in the future if the drainage policy is kept as it is now [4,5].

In the present communication soil contamination by saline water and geochemical land subsidence due to histosol oxidation in the Venice watershed are first reviewed with a glance at some of the most undesired consequences. Remediation strategies to alleviate or delay the processes are also outlined.

2 Soil contamination by saline water

Soil contamination by saline water occurs between the southern margin of the Venice Lagoon and the Adige River (Figure 1). The saltwater intrusion is connected to the geomorphological features of this area, i.e. the presence of sandy paleo-channels crossing the farmland with main direction from inland to the lagoon boundary and land elevation well below the mean sea level [8].

A research project has recently been started to map the extent of the occurrence and evaluate its criticality in relation to the economical activities of the area, i.e. mainly agriculture and horticulture. Saltwater contamination has been detected using hydrogeological and geophysical surveys and a new

monitoring network of surface water and shallow groundwater. Namely, 700 existing stratigraphies and penetrometric tests have been analyzed and integrated with 25 new continuously sampled boreholes drilled to a depth ranging from 15 to 100 m, and a network of about 100 monitoring wells has been established to measure the groundwater conductivity. A number of vertical electric soundings (VES) have been processed with a few electrical resistivity tomographies (ERT). Moreover, salinity measurements have been performed on surface water using 400 sampling sites to monitor the connection between drainage network and phreatic aquifer, and the tide encroachment in the river mouth [9].

With reference to the tolerance limits used in the agriculture of the study area three classes of water quality were identified:

- salty, if the water electrical conductivity ρ exceeds 5000 μS/cm and the soil electrical resistivity λ is less than 4.5 ohm·m, i.e. the seawater salinity (about 35 g/l);
- brackish, if ρ ranges between 2000 and 5000 μS/cm and λ between 4.5 and 7 ohm·m with salt concentration higher than 1 g/l and water unsuitable for irrigation purposes;
- fresh, if ρ is less than 2000 μS/cm and is λ larger than 8 ohm·m.

Figures 2 and 3 show the depth from the ground surface of the top and bottom of the saltwater contamination plume. The top depth varies with the different hydraulic and weather conditions with the highest fluctuations occurring in the 0-10 m depth interval. Six depth classes equally distributed between 0 to 30 m are used (Figure 2). The areas belonging to the first class (0-5 m) are most at risk of desertification as the saltwater contamination directly involves the plant and crop root zone. High probability of land degradation characterizes the areas in the second class (5-10 m) as well because of the seasonal climate oscillations that can induce a transient rise of the fresh-saltwater interface. The farmland within the other classes, although somewhat sensitive to seasonal interface oscillations, is much safer in relation to saltwater contamination.

Though of less importance for the soil agricultural use, the bottom of the saltwater plume may also supplies interesting information (Figure 3), e.g. it shows the transition from the aquifer to the aquitard units. The latter act as a barrier to the downward migration of salt, thus preserving the underlying freshwater aquifers. Note that the base of the saltwater plume deepens southward, from a shallow depth of 15-30 m down to 60-75 m, and locally 100 m. The morphology of the bottom interface depends mainly on the areal distribution of the clayey layers, their permeability, as well as on thickness and continuity. Due to the evolution history of the lagoon area, wide spatial continuous impermeable layers have rarely been identified down to depth of investigation, because of the complex geologic features of this zone [8].

The time and space behavior of salt contamination is influenced by other anthropogenic factors, such as the activity of several pumping stations used to keep drained the area, groundwater withdrawals, irrigation and freshwater releases during the summer dry months. Another forcing factor that plays a primary role is the sea and lagoon tide dynamics that is responsible for

significant saltwater encroachment into the river mouths and channels. To prevent seawater migration upstream from the river mouths during periods of low river discharge, a preliminary project for the construction of mobile gates at the mouth of the Brenta River was planned in 2004 [9]. This work which is estimated in about 15 million Euros, would generate a freshwater volume of more than 3×10^6 m^3 in the terminal portion of the Brenta and Bacchiglione rivers and in the Gorzone and Valle channels (Figure 1), presently invaded by seawater.

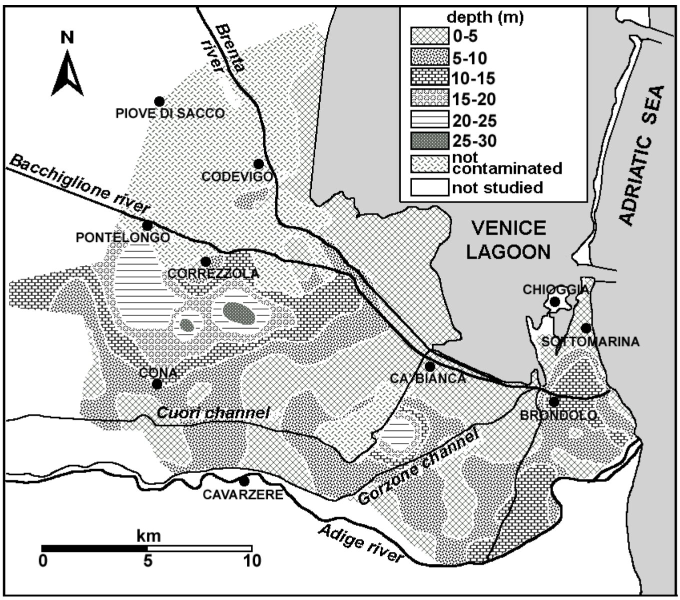


Figure 2: Depth (m below the ground surface) of the top of the saltwater plume.

3 Subsidence due to peatland oxidation

Large areas south of the Venice Lagoon are characterized by outcropping peat (Figure 1). They were reclaimed from 1892 to 1967 and at present lie almost entirely below mean sea level. The actual depth to the water table is regulated by a complex network of ditches and pumping stations that discharge the drainage water into the Venice Lagoon or the Adriatic Sea. Flooding from the sea and the lagoon is prevented in normal conditions by levees.

When drained the soil organic fraction, generally exceeding 50% of the total volume, is subject to oxidation with CO_2 release to the atmosphere, loss of sediment mass, and related land subsidence. This is demonstrated in Figure 4 with the aid of a few examples showing how some important infrastructures look like at present. The evident protrusion above the water level in the drainage network, or even above the ground surface, is a common feature of all hydraulic installations that were built at the time of the basin reclamation. The chemical

reaction underlying the release of CO_2 is primarily controlled by soil temperature and water content [10], and can be accelerated by agricultural practices, e.g. ploughing that bring to the surface deep fibrous unmineralized peat [11].

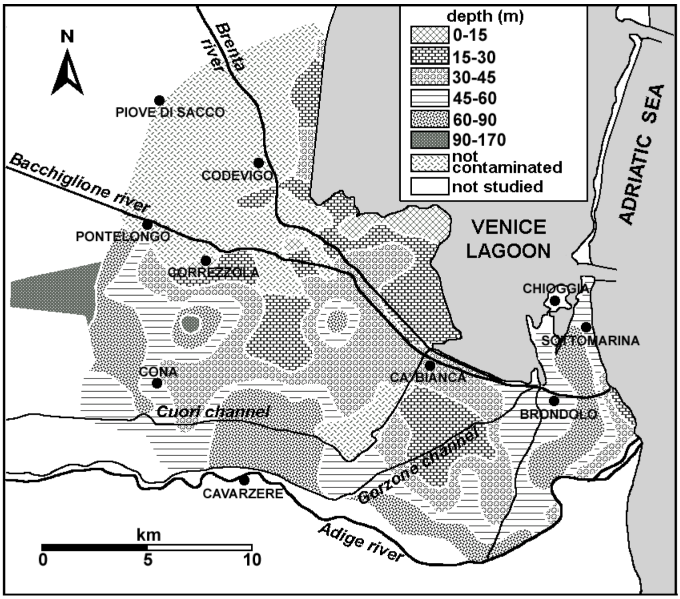


Figure 3: Depth (m below the ground surface) of the bottom of the saltwater plume.

A research project called VOSS (Venice Organic Soil Subsidence) funded by Co.Ri.La was undertaken with the aim at characterizing the peat composition, defining the extent of the settling area, understanding the basic processes, and finally developing a predictive tool to help plan the most appropriate management strategies in relation to the dominant agricultural practices and the maintenance of an efficient drainage network able to safely protect the area from exceptional floods. In the VOSS mainframe, a closed catchment south of the Venice Lagoon, the Zennare basin (Figure 1), has been selected to implement the project.

The site has been instrumented with the purpose of measuring the ongoing land settlement, collecting the main hydrologic variables that control the peat oxidation (i.e. depth to water table, rainfall, soil temperature and moisture, capillary pressure) and quantifying the CO_2 fluxes to the atmosphere with a NSS (Non Steady State) chamber [12]. Currently the thickness of the residual surface peat layer averages 1 m. Prevailing sand and silty sand occur in the northernmost part of the basin with traces of old paleorivers still visible throughout the catchment. The Zennare basin lies almost entirely below mean sea level, except for a small northern corner where mineral soils are dominant. On the basis of an aerial photographic survey performed in 1983 (and never re-performed again),

elevation ranges between -2 and -4 m above mean sea level. Ground surface displacements have been measured and monitored using an ad hoc extensometer. The area extent of peatland has been investigated by satellite data. Several images from IKONOS, ASTER, and LANDSAT-7ETM+ satellites, which combine high geometric (1 m^2 for IKONOS) and high spectral (6 bands for LANSAT and 14 for ASTER) resolution, have been analysed and calibrated against available geomorphologic maps and a large dataset of peat spectral signatures collected in situ with a portable spectrometer. The peat areas detected in this way compare nicely with an 1833 map of the local marshes drawn by government officials of the Lombardo-Veneto kingdom [7].

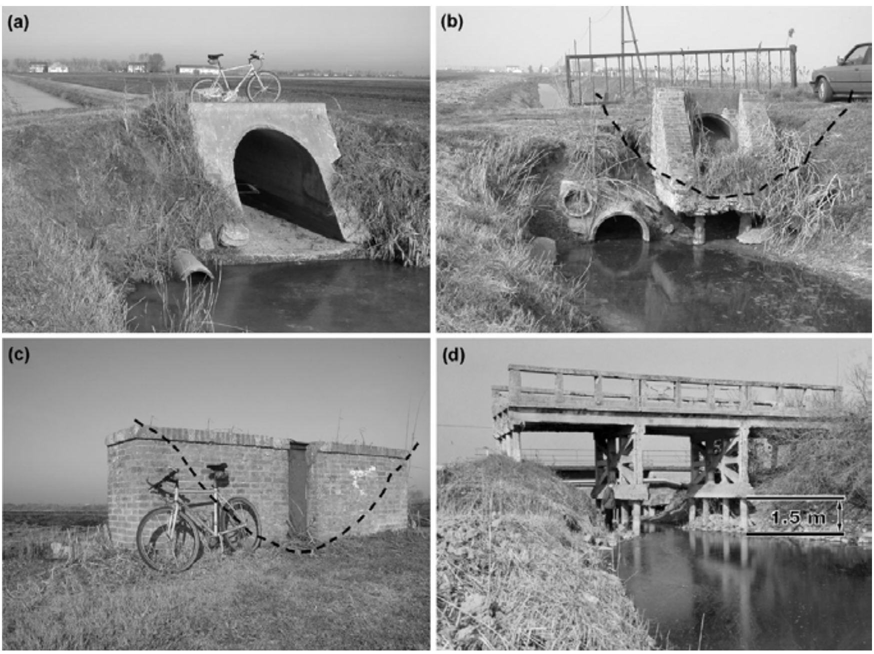


Figure 4: Evidence of the anthropogenic land subsidence in the reclaimed area. (a) A bridge has been turned into a useless structure: the left drainpipe helps convey the water of the channel originally flowing through the protruding infrastructure. (b) An old masonry culvert presently above the water level and substituted by two lower concrete drainpipes, the higher of which already unusable. A qualitative position of the ditch section in the original configuration is sketched. (c) The protrusion of a sluice wall above the bed of an old disappeared channel. (d) An old bridge hanging over the canal bank that settled by 1.5 m.

Although ad hoc campaigns to measure land subsidence in the Zennare basin were never performed before 2000, a number of independent factors point to an average long-term subsidence rate of 2-3 cm/y. Land subsidence estimated from

1964 to 1983 by the reclamation authority with the aid of elevation maps suggest that locally a maximum settlement approaching 1 m was measured near the northwestern and the southwestern Zennare boundaries. The Zennare pumping station provides indirect evidence of the surface sinking experienced in the area. The shaft level of the station that drains off the surplus water from the basin has been lowered from the original 1930 elevation by approximately 1.5 m [5]. Comparing the 1983 DEM with the outcome of a DGPS (Differential Global Positioning System) survey carried out in 2002 shows that some points have settled at a rate of 4-5 cm/y, consistent with the area peat distribution detected from satellites. On the average a 1.5-2 m of subsidence has affected the basin over the last 70 years with a present trend equal to 1.5-2 cm/y. Figure 5 gives the local vertical land displacement measured by the extensometer from January 2002 to December 2003. Note the highly fluctuating and irregular behaviour pointing to the superposition of different processes, including the important elastic peat response to rainfall (mire breathing) and soil freezing. However, a general trend is clearly distinguishable and indicates 2 cm over the two year observation period. The CO_2 efflux measured with the NSS chamber ranges between 0.2 and 0.7 mgr $m^{-2}s^{-1}$ corresponding to subsidence rates that are in good agreement with the direct displacement records.

Finally, a prediction based on the model by Stephen et al's [10] indicates that if the 2003 temperatures are projected into the future, the presently existing peat layer is bound to disappear completely in about 65 years for a constant 60 cm depth to the water table with an additional 75-100 cm of subsidence expected over the next half century. By distinction if a 20 cm water table depth is constantly maintained, 200 years would be required to totally oxidize the land peat. To become an effective management strategy, a shallower phreatic surface calls for innovative agricultural practices along with an accurate and timely control of the drainage system and the pumping stations, possibly with the aid of forecasting models.

4 Conclusion

Water and land management over the last decades has triggered two undesired occurrences in the southern Venice watershed: saltwater contamination of the phreatic and shallowest confined aquifers and geochemical land subsidence. The combined effect of both processes is producing an alarming social and environmental impact on the coastland addressed by the present study, also in relation to the expected global climate change.

Although the area is not experiencing everywhere the saline contamination, a very serious situation has been brought to light in a large portion of the coastal farmland. The fresh/saltwater interface is very close to the ground surface, between 0 to 10 m, with the thickness of the upper freshwater lenses often on the order of 1 m, depending on drainage and weather conditions. The saltwater contamination plume extends up to 20 km inshore and is enhanced by land elevation, generally below the mean sea level, the geomorphological structures of the coastal plain, and the seawater encroachment along the hanging last

portion of the watercourses. A cumulative land settlement of more than 1 m over the past 70 years has been observed in many places. The oxidation of the soil organic matter accounts for a present subsidence rate of 1.5-2.0 cm/y to which other less relevant factors, e.g. neotectonics, natural consolidation of the sedimentary column and groundwater pumping, superpose.

Mitigation of both hazards is of paramount importance. A number of strategies are under investigation, i.e. 1- the careful implementation of an ad hoc drainage management to maintain the water table as high as possible; 2- the control of seawater upstream migration from the river mouths by mobile gates; 3- the use of conservative tilling as a substitute to ploughing to decrease the exposure of new unmineralized peat to atmosphere; and 4- the introduction of cover crops that may partially offset the loss of organic material thus yielding a reduction of the present anthropogenic land subsidence rate.

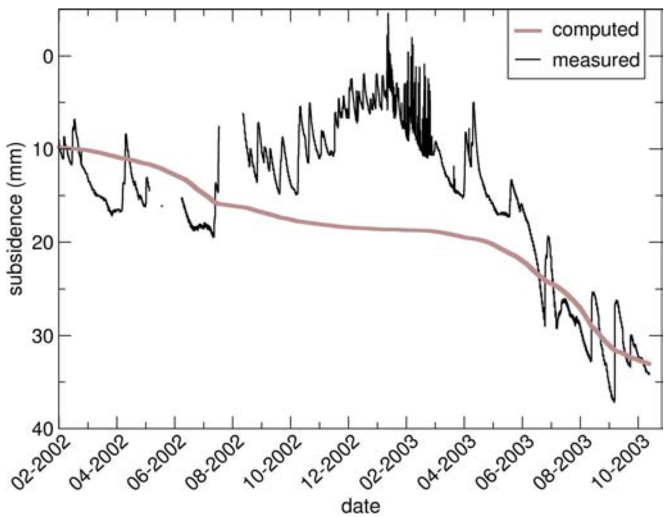


Figure 5: Comparison between computed and measured land subsidence since February 2002 after ice melting.

Acknowledgments

The work has been supported by Co.Ri.La., Subprojects 3.1 and 3.10, and ISES Project.

References

[1] Carbognin, L., Teatini, P. & Tosi, L., Land Subsidence in the Venetian area: known and recent aspects. *Giornale di Geologia Applicata*, **1**, pp. 5–11, 2005.

[2] Brambati, A., Carbognin, L., Quaia, T., Teatini P. & Tosi, L., The Lagoon of Venice: geological setting, evolution and land subsidence. *Episodes*, **26(3)**, pp. 264–268, 2003.

[3] Bortolami, G., Carbognin, L. & Gatto, P., The natural subsidence in the Lagoon of Venice, Italy. *Land Subsidence*, eds. A.I. Johnson et al., IAHS Publ. no. 151, pp. 777–785, 1985.

[4] Gambolati G. et al., Peatland oxidation enhances subsidence in the Venice watershed. *EOS-Trans. Amer. Geoghys. Union*, **86(23)**, pp. 217–224, 2005.

[5] Gambolati, G., Putti, M., Teatini, P. & Gasparetto Stori, G., Subsidence due to peat oxidation and impact on drainage infrastructures in a farmland catchment south of the Venice Lagoon. *Environ. Geology*, **49(6)**, pp. 814–820, 2006.

[6] Carbognin, L., Teatini, P. & Tosi, L., Relative land subsidence in the lagoon of Venice, Italy, at the beginning of the new millennium. *J. Marine Systems*, **51(1-4)**, pp. 345–353, 2004.

[7] Bixio, A.C., Putti, M., Tosi, L., Carbognin, L. & Gambolati, G., Finite element modeling of salt water intrusion in the Venice aquifer system. *Computational Methods in Surface and Ground Water Transport*, eds. V.N. Burganos et al., vol. 2, Computational Mechanics and Elsevier Applied Sciences: Suthampton, UK, pp. 193–200, 1998.

[8] Rizzetto, F., Tosi, L., Carbognin, L., Bonardi, M. & Teatini, P., Geomorphological setting and related hydrogeological implications of the coastal plain south of the Venice Lagoon (Italy). *Hydrology of the Mediterranean and Semiarid Regions*, eds. E. Servat et al., IAHS Publ. n. 278: Wallingford, UK., pp. 463–470, 2003.

[9] Carbognin, L., Rizzetto, F., Tosi, L., Teatini, P. & Gasparetto-Stori, G., L'intrusione salina nel comprensorio lagunare veneziano. Il bacino meridionale. *Giornale di Geologia Applicata*, **2**, pp. 119–124, 2005.

[10] Stephens, J. C., Allen Jr., L. A. & Chen, E., Organic soil subsidence. *Man-Induced Land Subsidence*, ed. T. Holzer, Rev. Eng. Geol., Geological Society of America: Boulder, Colorado, **VI**, pp. 107–122, 1984.

[11] Andriesse, J.P., *Nature and Management of Tropical Peat Soils*, FAO Soils Bulletin 59, 1988.

[12] Hutchinson, G.L. & Rochette P., Non-flow-through steady-state chambers for measuring soil respiration: Numerical evaluation of their performance. *Soil Sci. Soc. Am. J.*, **67**, pp. 166–180, 2003.

Elemental composition of PM_{10} and $PM_{2.5}$ in ambient air downwind of agricultural operations in California's San Joaquin Valley

O. F. Carvacho, L. L. Ashbaugh & R. G. Flocchini
Crocker Nuclear Laboratory, University of California, Davis, CA, USA

Abstract

Fugitive dust emissions from soil are thought to constitute a large fraction of the PM_{10} and $PM_{2.5}$ inventory in California's San Joaquin Valley (SJV) and other western air basins, especially during dry periods. The major sources of these emissions are paved and unpaved roads, construction sites, windblown dust, and agricultural activities. Furthermore, PM_{10} and $PM_{2.5}$ are considered to be among the most harmful of all air pollutants. When inhaled these particles evade natural defenses of the respiratory system and lodge deep in the lungs causing serious health problems. Some heavy metals in small particles have the tendency to donate electrons and to form basic oxides. Biologically, many metals are essential to living systems and are involved in a variety of cellular, physiological, and structural functions. But at high doses, many metals become toxic. The route of exposure may affect the dose and the site where the metal concentrates, and thus the observed toxic effect.

In California's San Joaquin Valley, agricultural operations are highly complex and potentially significant sources of PM_{10} and $PM_{2.5}$, especially during late summer and fall. A series of experiments was conducted to measure PM_{10} and $PM_{2.5}$ emissions with traditional array sampling from agricultural operations in San Joaquin Valley. The elemental analysis of PM_{10} and $PM_{2.5}$ collected in the field samples was conducted using Proton Induced X-Ray Emissions (PIXE), Proton Elastic Scattering Analysis (PESA) and X-Ray Fluorescence (XRF) analytical techniques available in our laboratory.

The composition of PM_{10} dust collected downwind of agricultural operations is different from the composition of the $PM_{2.5}$ dust collected at the same time. The smaller particles are enriched in sulphur and in heavy metals.

Keywords: PM_{10}, $PM_{2.5}$, fugitive dust, agriculture, particle speciation, elemental composition.

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060681

1 Introduction

California's San Joaquin Valley is one of the most productive agricultural regions in the United States. The dominance of fugitive dust from mobile and agricultural sources in the fall has lead to the hypothesis that agricultural operations may contribute significantly to the exceedance of PM_{10} concentrations in the valley.

Carvacho, et al. [1, 2] documented a strong relationship between soil texture as measured by the amount of sand, silt, and clay in the soil and the amount of PM_{10} and $PM_{2.5}$ that could be generated from it (the PM_{10} or $PM_{2.5}$ Index). They also showed that the $PM_{2.5}$ that could be generated from a soil was approximately 10% of the PM_{10} that could be generated from the same soil. Carvacho, et al. [3] examined the composition of PM_{10} and $PM_{2.5}$ from resuspended soil samples collected in the San Joaquin Valley. Finally, Carvacho et al. [4] documented the presence of heavy metals in $PM_{2.5}$ that may be hazardous to human health.

In this study, we document the elemental composition of the ambient PM_{10} and $PM_{2.5}$ dust collected downwind of agricultural operations on a variety of soil textures, and the elemental enrichment of metals and other elements in $PM_{2.5}$ versus PM_{10} dust.

2 Materials and methods

All samples were collected on a single farm near Stratford, CA between July 26 and September 11, 1999. All measurements were made under actual field conditions. A combination of upwind/downwind source isolation and vertical profiling were used to quantify PM_{10} and $PM_{2.5}$ concentrations, as described in Holmén et al. [5], and shown in Figure 1. We collected ambient PM_{10} and $PM_{2.5}$ on Teflon filters using IMPROVE samplers [6, 7] for gravimetric and elemental analyses.

The aerosol mass concentrations were calculated using the gravimetric method. The elemental composition (22 elements) was determined using three analytical methods: PIXE (Proton Induced X-ray Emissions) for elements with atomic mass less than Fe, XRF (X-Ray Fluorescence) for Fe and above, and PESA (Proton Elastic Scattering Analysis) for hydrogen. Further details of these techniques are described elsewhere [7–9]. PIXE and PESA were conducted using 4.5 MeV protons produced by the 76" cyclotron at the Crocker Nuclear Laboratory of the University of California in Davis. XRF analysis used a General Electric grounded anode diffraction type X-ray tube with molybdenum anode.

2.1 Composite variables

A SOIL parameter was calculated using the IMPROVE formula by adding the concentrations of five major soil elements in their typical oxide form [10] as shown in eqn (1).

$$SOIL = 2.20*[Al] + 2.49*[Si] + 1.63*[Ca] + 2.42*[Fe] + 1.94*[Ti] \quad (1)$$

The hydrogen concentration is useful as an estimate of organic mass. Sulphur was used to calculate the sulphate aerosol component, which is assumed to be ammonium sulphate. Organic mass and sulphate were calculated following the IMPROVE formulas [10, 11], shown in eqns (2) and (3).

$$Organic(byH) = 13.75 * ([H] - 0.25 * [S]) \tag{2}$$

$$SO_4^{=} = 4.125 * [S] \tag{3}$$

Figure 1: Photograph of downwind sampling array and meteorological measurement tower in a land preparation field study.

3 Results and discussion

Figure 2 shows the mass concentration and fractional composition of the PM_{10} dust collected from ambient samples downwind of agricultural operations. For all the soil types examined, mineral soil (i.e. the SOIL parameter) accounts for 77% to 87% of the PM_{10} mass in downwind ambient samples. Organic matter comprises 12% to 22% of the PM_{10} mass for all soil types, with sulphate, metals, and other elements accounting for 1% or less.

Figure 3 shows the mass concentration and fractional composition of the $PM_{2.5}$ dust collected from ambient samples downwind of agricultural operations. The $PM_{2.5}$ mass is 6% to 7% of the PM_{10} mass for all soil types except loam, where it is 12% of the PM_{10} mass. Mineral soil accounts for 37% to 66% of the $PM_{2.5}$ mass for all soil types. Organic matter accounts for 32% to 41% of the $PM_{2.5}$ mass, and sulphate is 2% to 21%. Metals and other elements account for 3% or less of $PM_{2.5}$ mass.

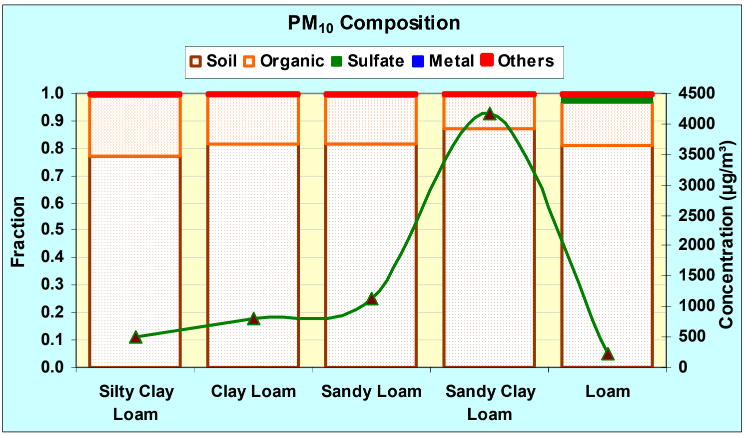


Figure 2: Concentration and composition of PM_{10} soil dust from ambient samples.

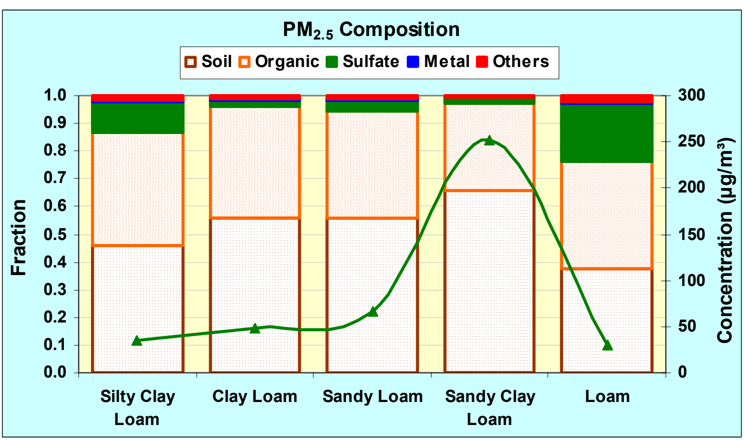


Figure 3: Concentration and composition of $PM_{2.5}$ soil dust from ambient samples.

In general, the elemental composition of the ambient samples collected during agricultural operations suggests the dust is mostly local soil. Exceptions

are for sulphate, metals, and organic mass. This is possibly because the ambient samples contain sulphate and metals from San Joaquin Valley air that are not present, or are present in very small amounts, in the underlying soil. The contribution from ambient background particulate matter was subtracted from the downwind samples, though, so the contribution from agricultural equipment may be a major source of these components.

We calculated enrichment factors (Table 1) relative to silicon using eqn (4) for elements in $PM_{2.5}$ compared to elements in PM_{10} samples.

$$EF = \frac{\left([E_{2.5}]/[Si_{2.5}]\right)}{\left([E_{10}]/[Si_{10}]\right)} \tag{4}$$

In eqn (4) EF is the enrichment factor, $[E_x]$ is the elemental concentration in the $PM_{2.5}$ or PM_{10} fraction, and $[Si_x]$ is the silicon concentration in the $PM_{2.5}$ or PM_{10} fraction.

Table 1: Enrichment factors for elements in $PM_{2.5}$ versus PM_{10} relative to silicon.

Element	Silty Clay Loam	Clay Loam	Sandy Loam	Sandy Clay Loam	Loam	Average all soils
H	4.1	3.8	3.7	3.6	7.1	4.5
Al	1.1	1.1	1.0	0.8	0.8	1.0
Si	1.0	1.0	1.0	1.0	1.0	1.0
Ca	2.4	1.7	2.1	1.7	1.8	2.0
Ti	0.5	1.7	1.5	1.6	1.5	1.3
Fe	1.3	1.3	1.2	1.4	1.4	1.3
S	36.5	11.5	13.1	13.0	17.6	18.4
V	40.0	16.3	4.0	5.7	8.1	14.8
Cr	8.9	12.4	18.3	4.3	11.2	11.0
Mn	5.1	3.8	2.4	2.1	2.0	3.1
Ni	9.3	8.5	7.7	6.5	7.2	7.8
Cu	46.1	21.8	43.9	14.3	12.4	27.7
Zn	15.4	8.0	10.7	9.4	3.5	9.4
Ga	16.2	11.6	13.2	9.6	12.1	12.6
Hg	17.5	11.8	15.1	9.8	7.8	12.4
As	6.9	7.7	4.9	3.8	13.1	7.3
Pb	17.4	16.4	24.6	11.7	4.4	14.9
Se	17.6	13.7	16.1	11.3	11.4	14.0
Br	17.9	15.4	28.7	5.3	17.4	16.9
Rb	2.0	2.0	1.7	1.3	2.7	1.9
Sr	1.6	1.9	1.4	1.2	1.3	1.5

Table 1 shows the enrichment factors for all five soils and for the average of all soils. The enrichment factors are also shown graphically in Figure 4. Four of the five soil elements (Al, Si, Ti, and Fe) are very similar in the $PM_{2.5}$ and PM_{10} fractions for all soil types. Calcium is enriched in the $PM_{2.5}$ fraction by a factor

of two. Rubidium and strontium are also slightly enriched (by less than a factor of two) in the $PM_{2.5}$ fraction. Hydrogen, manganese, nickel, and arsenic are enriched by factors of 3-8. Other metals, including V, Cr, Ga, Hg, Pb, Se, and Br are enriched by factors of 11 to 17 on average, with a wide variation for individual soil types. Sulphur is enriched by a factor of 18 on average with a range of 11.5 to 36.5. Finally, copper is enriched by a factor of nearly 28 on average, with a range of 12 to 46 for different soil types. It's not clear whether the source of these metals is the underlying soil, background $PM_{2.5}$ concentrations in the San Joaquin Valley atmosphere, or the emissions of the agricultural equipment used in the operations being sampled. The background concentrations were subtracted from these measurements, though, so it's unlikely that ambient levels are the source.

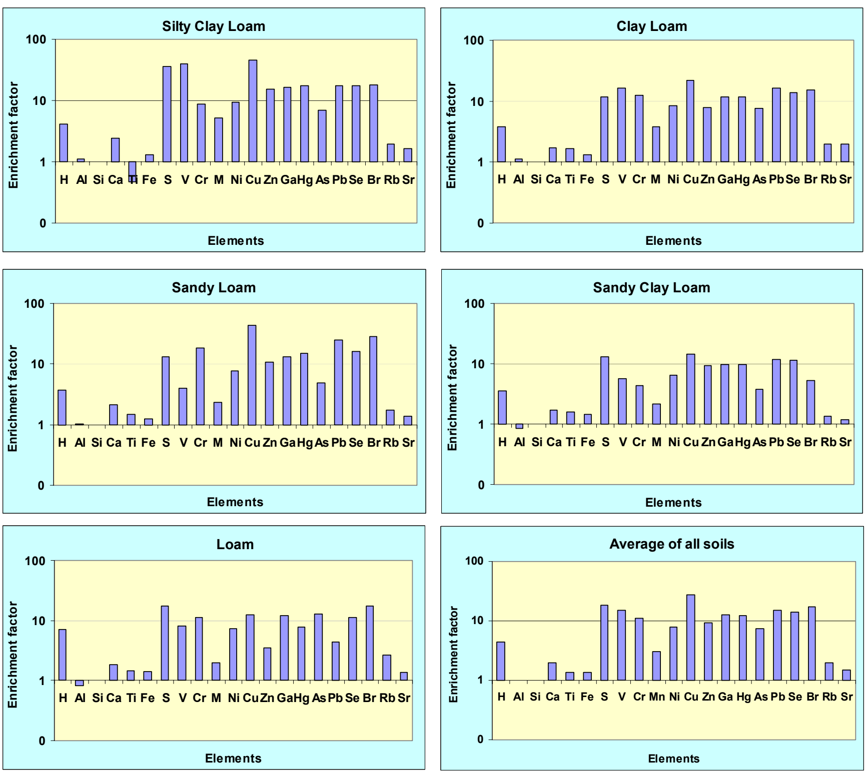


Figure 4: Enrichment factors, relative to silicon, of elements in $PM_{2.5}$ versus PM_{10}.

4 Conclusions

The composition of PM_{10} soil dust collected downwind of agricultural operations is primarily underlying mineral soil with a small fraction representing organic

matter. The PM_{10} fraction is approximately 80% mineral soil (as calculated by the IMPROVE equation) and 20% organic matter, with only minor amounts of sulphate and metals.

The composition of $PM_{2.5}$ soil dust collected downwind of agricultural operations is enriched in sulphur and heavy metals as compared to the PM_{10} fraction. The $PM_{2.5}$ mass concentration is 6-7% of the PM_{10} concentration for operations on most soils, but is higher (12%) for operations on loamy soil. The $PM_{2.5}$ composition is more variable than the PM_{10} composition, with 37-66% mineral soil, 32-41% organic matter, up to 20% sulphate and less than 1% metals.

Sulphate and metals are enriched in the $PM_{2.5}$ fraction compared to the PM_{10} fraction by factors that vary depending on the soil type and test conditions. Calcium is enriched by approximately a factor of two, while rubidium and strontium are enriched by less than a factor of two. Other metals, particularly V, Cr, Cu, Zn, Ga, Hg, Pb, Se, and Br are enriched by about an order of magnitude in the $PM_{2.5}$ fraction relative to the PM_{10} fraction. Sulphur is enriched by a factor or 11-36 (18.4 on average).

References

[1] Carvacho, O. F., L. L. Ashbaugh, Brown, M. S., & Flocchini, R. G. Relationship between San Joaquin Valley soil texture and PM_{10} emission potential using the UC Davis Dust Resuspension Test Chamber. *Transactions of the ASAE* 44(6): 1603-1608, 2001.

[2] Carvacho, O. F., Ashbaugh, L. L., Brown, M. S., & Flocchini, R. G. Measurement of $PM_{2.5}$ emission potential from soil using the UC Davis resuspension test chamber. *Geomorphology*, 59(1-4), 75-80, 2004.

[3] Omar F. Carvacho, Ashbaugh, L. L., Brown, M. S., & Flocchini, R. G. Elemental Composition of PM_{10} and $PM_{2.5}$ from Resuspended Soil in California's San Joaquin Valley. *The First International Conference on Environmental Science and Technology*, Sponsored by American Academy of Science, Fairmont Hotel, New Orleans Louisiana, USA January 23 – 26, 2005.

[4] Carvacho, O. F., Ashbaugh, L. L., & Flocchini, R. G. $PM_{2.5}$ Metals in Ambient Air Downwind of Agricultural Operations in California's San Joaquin Valley. *International Society for Aerosol in Medicine. 15th International Congress*. Perth, Western Australia March 14 – 18, 2005. *Journal of Aerosol Medicine*, Vol. 18, Number 1, pages 100-101, 2005.

[5] Holmén, B. A., T. A. James, L. L. Ashbaugh, & R. G. Flocchini. Lidar-assisted measurement of PM_{10} emissions from agricultural tilling in California's San Joaquin Valley - Part II: emission factors. *Atmospheric Environment*, 35, 3265-3277, 2000.

[6] Eldred, R. A., Cahill, T. A., Pitchford, M., & Malm, W. C. IMPROVE – A new remote area particulate monitoring system for visibility studies. *Proceedings of the 81st Annual Meeting of the Air Pollution Control Association*, Pittsburgh, PA, 1988.

[7] Cahill, T. A., Eldred, R. A., Shadoan, D., Feeney, P. J., Kusco, B. H., Matsuda, Y. Complete Elemental Analysis of Aerosol - PIXE, FAST, LIPM and Mass. *Nuclear Instruments and Methods in Physics Research*, B3, 291-295. 1984.

[8] Morales, J., Dinator, M., Romo-Kroger, C. Caracterización del material particulado atmosférico en suspensión en Santiago de Chile. Parte VII: determinación elemental mediante PIXE. *Nucleotécnica* 15, 47 – 50. 1995.

[9] Rojas, M. C., Artaxo, P., Van Grieken R. Aerosol in Santiago de Chile: a study using receptor modeling with X-Ray fluorescence and single particle analysis. Atmospheric Environment B 24, 227 – 241, 1990.

[10] Eldred, R. A., Cahill, T. A., Feeney, P. J., Particulate monitoring at United-States National-Parks Using PIXE. *Nuclear Instruments & Methods in Physics Research Section B-Beam Interactions with Materials and Atoms*, 22, 289-295, 1987.

[11] Eldred, R. A., Cahill, T. A., Wilkinson, K. L., Feeney, P. J., Particulate Characterization at Remote Sites Across the US: First year results of the NPS/IMPROVE Network. *Air Waste Management Assoc. 82nd Annual Meeting & Exhibition*, Anaheim, California June 25 – 30. Paper 89-151.3, 1989.

Assessment of slopes endangered by groundwater

P. P. Prochazka[1] & J. Trckova[2]
[1]*Czech Technical University Prague, Czech Republic*
[2]*Academy of Sciences, Prague, Czech Republic*

Abstract

In many cases of deposits like tailings from open pit mines, the environment is damaged by the long extension of the slopes. Steeper slopes have a better influence on the landscape. On the other hand, the basic problem is caused by ground water, which a posteriori, after completing the construction of these structures, seeps through the foundation of the slopes. It is very important that no reclaim helps to solve this problem. If the deposits are founded in the wrong way, i.e. in the foundations where no disturbances of capillary water are applied, new methods have to be employed in order to stabilize the slopes. Before this, studies concerning the influence of ground water on the slope stability of the deposits have to be carried out. Natural slopes cannot practically be assessed, as the expanses involved in such tests are enormous. This is why scale models are prepared, together with numerical analysis based on coupled modeling. The influence of water is simulated by free internal parameters, which describe the extent of the hydration. On the other hand, geotechnical parameters change according to rules determined in the laboratory. Such a complex discussion can then be a starting point for several improvements of the slopes. The coupled modeling (scale and numerical) is applied to prepare a parametric study of tailing dams in Northern Bohemia, where large problems occurred because of the moving of current slopes against existing villages in this region. Scale models have been created in compliance with real situations and the results from them implemented in mathematical models. A parametric study was carried out using the numerical models.
Keywords: tailing dams, Desai's distinct state model, scale modeling, numerical modeling, coupled modeling.

1 Introduction

In Northern Bohemia a problem occurs as a result of poor foundations of tailing dams from open pit mines. They move towards existing villages, in some cases at a rate of 1cm a day. The villages are threatened by the moving mass of slopes of very high ridge. The solution to this situation is based on mathematical modeling, which has to be adjusted by comparing results from scale modeling and numerical results. Here, Desai's model [1] of the watering of underground mass is used together with eigenparameters, serving as the design parameters for an optimization problem (actually, smart back analysis is introduced). In [2], the optimal material properties of slopes are sought using similar techniques. An approach considering both scale and numerical modeling is sometimes called coupled modeling. This is widely used in some preceding papers by the authors of this paper, [3-5]. Deposits and tailing dams are also solved in these papers for different real-life cases, but in these papers the influence of water is not considered, which is obviously the most dangerous phenomenon. Now, we have a large scale of tests available, which also covers the response of the slope material to the effect of wetting the slope mass. The aim of this paper is to show an approach, which leads to a very good numerical model for being able to assess different defence technologies against moving deposits.

2 Deformation of the deposit due to underflooding

Deformation of deposits caused by water levels that increase to reach the toe of the deposit slope have been observed on scale models of physically equivalent materials. The models were created as a dam; 10m high, 15m wide and with a slope of 30°. The models were prepared in a modeling stand, which is a basin $800 \times 230 \times 600\text{mm}^3$ (length × width × height), created in a stiff steel frame, in which waterproofed Plexiglas walls of a thickness of 12mm are fastened. The scale of the dimensions of the real:model slope are 1:100.

Measurement points are placed on the surface of the model and in the longitudinal axis, (fig. 1). At measurement points in each stage of the level of groundwater, movement at these points was identified using theodolit, which assesses vertical and horizontal angles α_i and β_i. Knowing the distance d of the axis of theodolit from the measurement points, and with known changes of vertical and horizontal angles in each consecutive stage of water level, the vectors of movements v_i at distinct measurement points are calculated as:

$$a_i = d.\mathrm{tg}\,(\alpha_i - \alpha_{i-1}), \qquad b_i = d.\mathrm{tg}\,(\beta_i - \beta_{i-1}), \qquad v_i = \sqrt{a_i^2 + b_i^2}\ .$$

Water was fed into the model through openings at the base of the stand through hoses connected to a water reservoir. From this reservoir the water was added to the stand according to the requirements of the simulated process. After each increase in water level, photos of the situation in the stands and measurements at measurement points documented the status of the sample.

2.1 Physical model 2023

The first of the models was created from a mixture of very fine balotine, 89,98%, ferrosilicia 10%, and a fat A00 0,02%. This material in the selected scale reasonably simulates stiffened sand soils, which are predisposed to low swelling.

The physical and mechanical properties of the modeled material in a naturally moistened state are defined as:

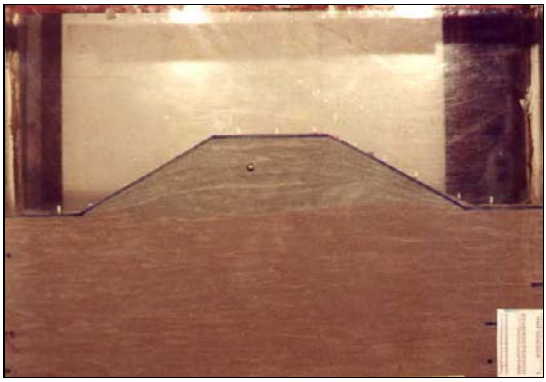
Volume weight ρ	1.93 g/cm^3
Tensile strength σ_c	6.9 kPa
Cohesion (triaxial) c	5.9 kPa
Angle of int. friction φ	15°
Young's modulus E	2.1 MPa
Oedometric modulus of deformation E_u	0.5 MPa
Coefficient of compressibility C	54.6
Poisson's number ν	0.33

Material with an extremely low volume swelling of 1.3% was prepared. After completing the scale models, the water was allowed into the stands. In figs. 2a and 2b, the states before wetting and after the final stage of wetting are depicted. In Table 1, the vectors of displacements at the measurement points are introduced for entire flow of the test. Changes of horizontal x and vertical y displacements, together with the resultant vector d are there shown. The minus sign denotes right and downwards. Values are given in mm.

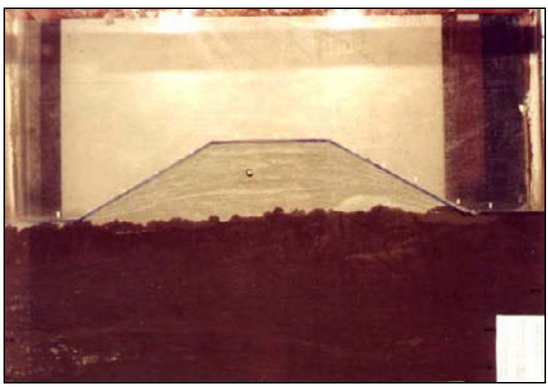
Figure 1: Measurement points.

The test was divided into several stages depending on the height of the water level in the stand. In the first stage, no water takes place in the sample. The displacements in the second stage also involve movements due to the imprecise settlement of the equivalent material around the drainage hose, the water was let into the stand, which is why the values are not introduced in tables. At the third stage, when the water level reached 9.5 cm and the fourth stage (water level at 21cm above the bottom of the stand), the movements were relatively small. The

decisive displacements were registered in the stages when the water started to approach the toe of the slope and seepage into the slope began to take place (stages 5 - 6).



a)



b)

Figure 2: Photo of model 2023 before water was let in, a), and after end of the test, b).

After completing the test, selected material coefficients were tested in the laboratory to obtain information on how material properties changed from before and after water saturation of the soil and checked with the results of the numerical model proposed.

2.2 Physical model 2024

The material for the second test was created from slightly different components. The physical and mechanical properties of the modeled material in a naturally moistened state are defined as follows:

Volume weight ρ	1.93 g/cm^3
Tensile strength σ_c	8.8 kPa
Cohesion (triaxial) c	7.9 kPa

Table 1: Horizontal x, vertical y, and resultant displacements d.

Point	Stage 1			Stage 4			Stage 5 – 6		
	x	z	d	x	z	d	x	z	d
1	0,01	-0,22	0,22	0,06	0,01	0,06	0,17	0,18	0,24
2	-0,06	-0,31	0,32	0,06	0,01	0,07	0,45	0,20	0,49
3	-0,10	-0,36	0,37	0,14	-0,15	0,21	0,40	0,02	0,40
4	-0,13	-0,35	0,38	0,11	-0,13	0,17	0,29	-0,11	0,31
5	-0,13	-0,36	0,38	0,04	-0,16	0,16	0,21	-0,32	0,39
6	-0,17	-0,37	0,41	0,14	-0,11	0,17	0,19	-0,23	0,30
7	-0,10	-0,39	0,40	0,01	0,03	0,03	0,01	-0,04	0,04
8	-0,19	-0,43	0,47	-0,20	-0,02	0,20	-0,14	-0,19	0,23
9	-0,13	-0,38	0,41	-0,12	-0,20	0,23	-0,08	-0,30	0,31
10	-0,13	-0,38	0,40	-0,04	-0,26	0,26	-0,22	-0,06	0,23
11	-0,06	-0,40	0,41	-0,05	-0,13	0,14	-0,33	0,01	0,33
12	-0,03	-0,45	0,45	-0,12	0,12	0,17	-0,44	0,06	0,44
13	-0,09	-0,39	0,40	-0,11	0,11	0,16	-0,19	0,17	0,25

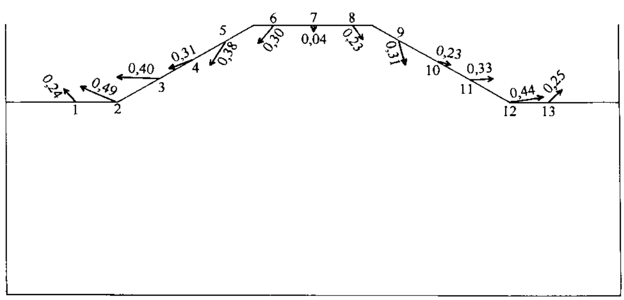


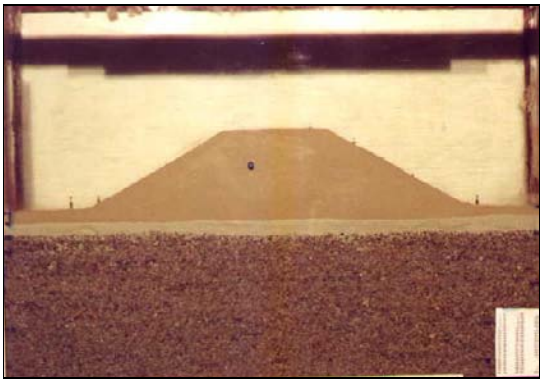
Figure 3: Model 2023 – vectors of displacements at measurement points (100x amplified).

Angle of int. friction φ	10°
Young's modulus E	1.8 MPa
Oedometric modulus of deformation E_u	0.5 MPa
Coefficient of compressibility C	41.3
Poisson's number ν	0.25

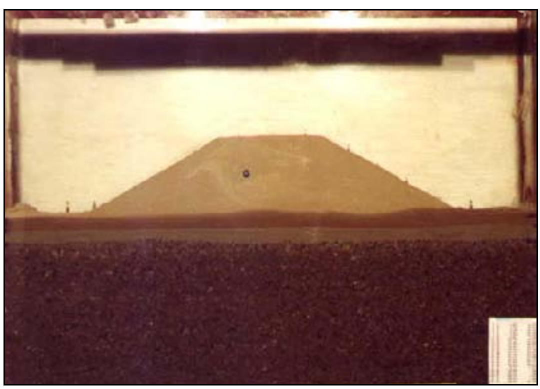
Rocky material of the bedrock was simulated by silica sand with a prevailing grain size of 1-4mm, which was coated by 20 mm thick layer of silica sand with granularity of 0.1 – 0.25mm. This part of the model has virtually no part in

changes to the massif due to underflooding. The body of the deposit and the bedrock to a depth of 20mm below the toe of the slope were created from a mixture of fine silica sand (80%) and bentonite (20%). A volume swelling of 49.6% (very high swelling) is attained after a couple of tests with different equivalent material. For completeness, the dry slope and the slope with the saturated body are seen in photos figs. 4a and 4b.

In Table 2, displacements at distinct measurement points are introduced in a similar way, as in Table 1.



a)



b)

Figure 4: Photo of model 2024 before water was let in, a), and after end of the test, b).

The development of moving points can be described as: When the water level reached the toe of the slope, the process of swelling was triggered. After letting in another appropriate amount of water, the water level increased, but very slowly. The inlet was stopped and it took two days when the water level was horizontal. Letting the next amount of water into the stand, the permeability of the material had decreased due to the swelling of the equivalent materials. It took some 36 hours until the water level was horizontal again.

In the final stage, the water rose up through the material due to capillarity. At this moment, the test was terminated. Vectors of displacements are depicted in fig. 5.

Comparing figs. 3 and 5, in which vectors of displacements are depicted for two models 2023 and 2024 with extreme swelling properties, quite different behavior is shown. In the case of slope from material with virtually no swelling properties, the largest displacements are observed at the toe of the slope, up to about 50mm in reality. Due to the strong swelling properties of the equivalent material of model 2024 and capillarity conditions, displacement upwards was ten times greater than before (fig. 5).

Table 2: Horizontal x, vertical y, and resultant displacements d.

Point	Stage 1			Stage 4			Stage 5		
	x	z	d	x	z	d	x	z	d
1	-0,01	0	0,01	-0,62	3,26	3,31	0,10	0,16	0,19
2	0,06	0,04	0,07	-0,62	1,99	2,08	0,16	0,59	0,61
3	-0,02	-0,01	0,03	-0,19	0,79	0,82	-0,04	0,48	0,48
4	-0,06	0	0,06	-0,04	0,58	0,58	-0,08	0,35	0,36
5	0,02	-0,04	0,05	0,01	0,45	0,45	-0,08	0,36	0,36
6	0,01	0	0,01	0,02	0,43	0,43	0,03	0,31	0,31
7	0,01	0,05	0,05	0,01	0,43	0,43	0	0,44	0,44
8	0,02	0,04	0,04	-0,04	0,38	0,39	0,05	0,41	0,42
9	0	-0,01	0,01	0,05	0,41	0,42	0,02	0,40	0,41
10	0,04	-0,04	0,05	0,04	0,39	0,40	0,06	0,48	0,49
11	0,06	0,02	0,07	0,22	0,63	0,67	0,04	0,53	0,53
12	0,02	0,02	0,03	0,46	1,64	1,70	-0,12	0,49	0,51
13	0	-0,06	0,06	0,38	3,36	3,38	-0,17	0,29	0,33

4 Numerical model

First introduce certain denotations and assumptions. Express displacements $u_i(\xi)$ at an arbitrary point ξ of the domain by virtue of superposition of displacements $\mathbf{u}^{ext}(\xi)$ at ξ due to external loading applied to a linearly elastic medium, or other materials, and a linear hull of influences of, say, the eigenstresses λ, and plastic strains ε^{pl} at other points $\mathbf{x}$. Since we assume that at each point, two values of displacements are necessary in 2D and three values in 3D, the relation stresses $\mathbf{u}^k$ at the points B_k, $k = 1,...,m$, and the eigenstresses and plastic stresses λ^l, $l = 1,...,n$ and $(\sigma^{pl})^l$, $l = 1,..., m$ at A_l becomes (to simplify the expressions, the vector notation for stress and strain tensors is used), cf. (1):

$$(u_i)^k = (u^{\text{ext}}{}_i)^k + \sum_{j=1}^{6} \sum_{l=1}^{m} (T^{\sigma}{}_{ij})^{kl}, (\sigma^{\text{rel}}{}_j)^l + \sum_{j=1}^{6} \sum_{l=1}^{m} (R^{\sigma}{}_{ij})^{kl} (\lambda_j)^l,$$
$$i = 1,...,6,\ k = 1,...,m,$$

or

$$(u_i)^k = (S^{\sigma}{}_i)^k + \sum_{j=1}^{6} \sum_{l=1}^{n} (R^{\sigma}{}_{ij})^{kl} (\lambda_j)^l,\ \ i = 1,...,6,\ k = 1,...,m,$$

where $(S^{\sigma}{}_i)^k$ expresses the current state of the overall displacements involving nonlinear changes in the material.

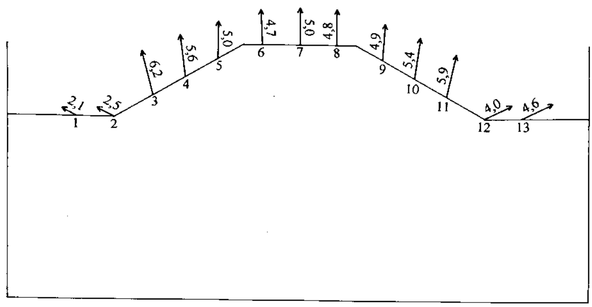


Figure 5: Model 2024 – vectors of displacement at measurement points (amplified ten times).

On the other hand, measured displacements $(u_i^{\text{meas}})^k$ are available in a discrete set of points. A natural requirement is that the values of measured and computed values be as close as possible. This leads us to the optimization of an “error functional”

$$I[(\lambda_j)^l] = \sum_{i=1}^{6} \sum_{k=1}^{m} [(u_i)^k - (u_i^{\text{meas}})^k]^2 \ \rightarrow \ \text{minimum}$$

Differentiating I by $(\lambda_\alpha)^\beta$ yields a linear system of equations for $(\lambda_j)^l$:

$$\sum_{j=1}^{6} \sum_{l=1}^{n} (A_{\alpha j})^{\beta l} (\lambda_j)^l = Y_\alpha^{\beta},\ \ \alpha = 1,...,6,\ \beta = 1,...,\text{m},$$

where

$$(A_{\alpha j})^{\beta l} = \sum_{i=1}^{6} \sum_{k=1}^{m} (R_{ij})^{kl} (R_{i\alpha})^{k\beta},$$

$$Y_\alpha^{\ \beta} = -\sum_{i=1}^{6} \sum_{k=1}^{m} (S_i)^k - (u^{\text{meas}}{}_i)^k + \sum_{j=1}^{6} \sum_{l=1}^{m} (R_{ij})^{kl}(\lambda_j)^l (R_{i\alpha})^{k\beta}$$

For more details of this procedure see [3-5].

5 Conclusions

In this paper, a procedure for determining mechanical behavior of earth moistened deposits from open pit mines is suggested. Information from scale models is taken into account for feeding the input data (material model) to the mathematical formulation of the physical law. Numerical codes are then tuned by an optimization problem to get as much good agreement from the results from scale and numerical models. In contrast to previous papers by the same authors, the influence of water is also considered in this study.

Acknowledgment

This paper was prepared with the financial support of GA AV ČR, project No. IAA 2119402.

References

[1] Desai Ch., Ma Y.: Modelling of joints and interfaces using the disturbed state concept - Int. J. Numer. & Anal. Methods in Geomechanics, Vol. 16, (1995) 623-653.

[2] Trčková J., Procházka P.: Optimal material properties of slopes using coupled modelling - Proc. GEOTECH–YEAR 2000, Developments in Geotechnical Engineering, Asian Inst. of Technology, ed. A.S. Balasubramaniam et al., Bangkok, Thailand, (2000) 439-448.

[3] Trčková J., Procházka P.: Application of coupled modelling to slope stability assessment – Proc. Computational Methods and Experimental Measurements X, WIT Press, Alicante, (2001) 447-456.

[4] Procházka P., Trčková J.: Material properties of tailings of open-pit mines using coupled modeling. Proc. 5th European Conf. on Numerical methods in geotechnical engineering NUMGE 2002, Mestat (ed.), Presses de lÉNPC/LCPC, Paris, (2002a) 273-278.

[5] Procházka P., Trčková J. Material properties of tailings of open-pit mines using coupled modeling. Environmental Studies IX, series: Environmental Studies Volume 7, (2002b) 361-370.

This page intentionally left blank

Radioactive contaminated soil removal from the sites of the former Azgir nuclear test site

D. G. Gilmanov & E. Z. Akhmetov
National Nuclear Centre of the Republic of Kazakhstan (NNC RK), Institute of Nuclear Physics NNC RK, Kazakhstan

Abstract

The research and engineering operations performed to clean and remove radioactive contamination from the Azgir test site improved levels of radiation in and around the site. After the operations, the possibility of contamination of personnel and the local population from radiation doses in the soil that exceed maximal permissible values for radiation were eliminated.
Keywords: Azgir nuclear test site, underground nuclear explosions, radionuclides, underground cavities from nuclear explosion, radioactive wastes.

1 Introduction

At the present time there are problems as a result of controlled peaceful nuclear explosions in Kazakhstan, which include clean-up activities in the contaminated area where the nuclear explosions were conducted. A considerable amount of time has passed since the explosions were carried out. This report addresses problems such as the removal of contaminated soil from the former AZGIR test site in the west of Kazakhstan. The pilot research was carried out from 1964 to 1979 by means of underground nuclear explosions in the rock salt massifs of an underground cavity used for multifunctional storage near to Azgir in Atyrau oblast (250 km from Caspian Sea and 70 km from the Volga River) at the salt dome of East and West Azgir. Among 39 peaceful underground explosions conducted in Kazakhstan (Table 1, Figure 1) [1], 17 explosions (22 blasting charges with energy release in trinitrotoluol equivalent from 0.01–103kt) were carried out at a depth of 160–1500m at 10 sites of the Azgir test site (Table 2). The total activity of underground cavities is about 10^6 Ci [2,3].

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060701

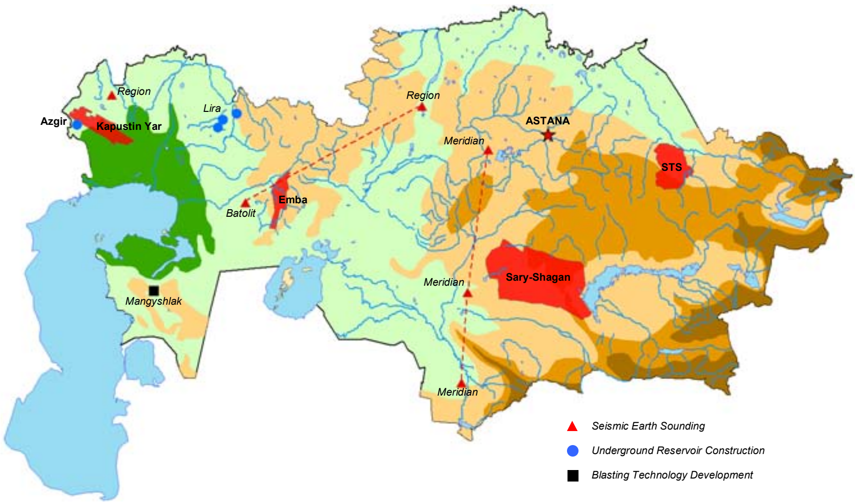


Figure 1: Peaceful nuclear explosions in Kazakhstan territory.

Table 1: Chronology of underground nuclear explosions at the Bolshoi Azgir salt dome.

Site	Explosion date	Depth of charge blasting, m	Explosion yield, trinitrotoluol equivalent, kt
A-I	April 22, 1966	160	1.1
A-II	July 1, 1968	600	27
A-III	December 22, 1971	1000	up to 64
A-IV	July 29, 1976	1000	up to 60
A-V	September 30, 1977	1500	up to 10
A-VII	October 17, 1978	1000	0.001 -20; 20-150
A-IX	December 18, 1978	900	103
A-VIII	January 17, 1979	1000	0.001 -20; 20-150
A-XI	July 14, 1979	1000	20
A-X	October 24, 1979	1000	0.001 -20; 20-150

As a result of the explosions conducted in the rock salt massif, 9 relatively resistant cavities of different volumes (from 10,000–240,000 m^3) were formed over a period of 3 decades (Figures 2 and 3). All the explosions were projected as having been fully confined i.e. without radioactive products reaching the surface. However, this was not achieved, as shown by the data of radioactive discharge duration (having become the reason for the test site contamination) from cavities over a period of some days up to months and of its total activities up to 10^{18} Bq. Site contamination also occurred as a result of some technological operations related to the opening of the cavity by means of additional boreholes, as well as radioactive gases with activity up to 10^7 Ci [4] that were injected into the atmosphere. Radioactive spots at the site were the source of secondary

radioactive contamination of the environment (radionuclide transfer by way of wind and atmospheric fallout) and radiation risk to the public. Therefore the problem of the clearance of contaminated land was considered as current in terms of sanitary and ecological aspects.

Table 2: Peaceful nuclear explosions in Kazakhstan.

Date, years	Place (oblast)	Condition	Amount of explosions
1966-1979	Guryev	Construction of underground tanks of large volume in rock salt massif (A-1 – A-XI)	17
1969-1970	Mangyshlak	Development of commercial explosion technology. T-1, T-2, T-3 boreholes	3
1965-1974	Semipalatinsk Test Site	Experimental works on NET development in behalf of the national economy	7
1972	Ural	Depth seismic “Region” sounding, P-5 borehole	1
1972	Kustanai	Depth seismic “Region” sounding, P-5 borehole	1
1973	Turgai	Depth seismic “Meridian” sounding, MN-1 borehole	1
1973	Shymkent	Depth seismic “Meridian” sounding, MN-2, MN-3 borehole	2
1983-1984	Ural	Construction of underground cavities “Lira” - 1÷6	6
1987	Aktyubinsk	Development of nuclear explosion technology “Batolit”, BT-2 borehole	1
TOTAL			**39**

Research and practical measures regarding the removal of radioactive contamination at the test site were carried out according to the following activities: mapping of the contaminated area, determination of radionuclide concentration in the soil, location and arrangement of radioactive soil disposal, cleansing of the area from radioactive contamination, transfer and burying of contaminated soil, dosimetry and determination of soil specific activity at the contaminated site, refilling of excavation with pure soil, estimation of mass and total activity of removed soil.

Radionuclide concentrations, especially Cs-137, as the concentration of other radionuclides were comparatively low (at the level of natural background and global fallout), allowed to subsume excavated soil (with maximal specific activity up to 9 000 Bq/kg) to the category of low-level radioactive wastes with established intervention level -10^4 Bq/kg and more by Radiation Safety Standards in Kazakhstan. Under this approach specified soil just had to be removed to a nearby ground disposal location at a depth of 10 m.

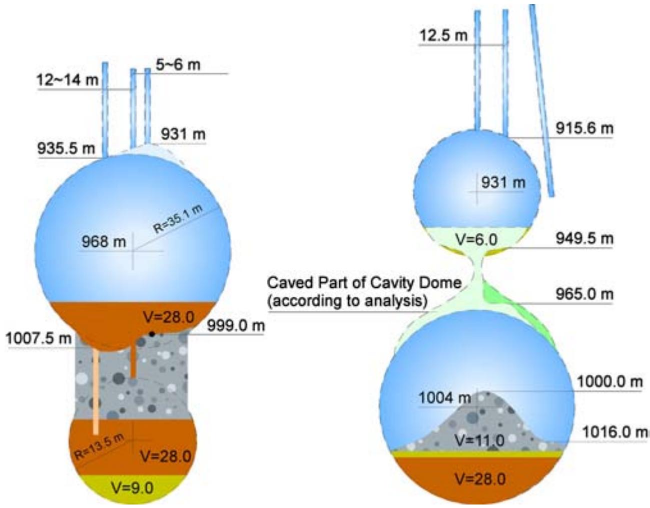


Figure 2: Cavity scheme A-VII and A-VIII.

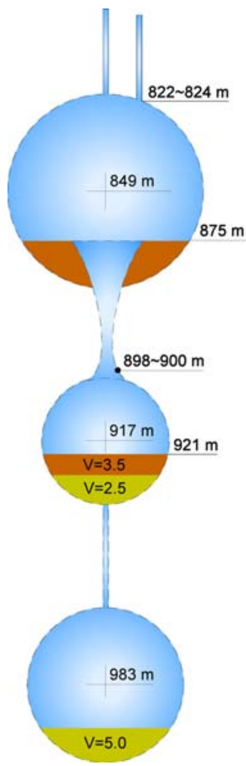


Figure 3: Cavity scheme at site A-XI.

2 Natural characteristics of AZGIR test site range

Geographical features of the region are its relative closeness to the Volga River and the Caspian Sea. The described area is situated in the north-west of the Caspian Sea lowlands. The test site is placed in semi-desert area on a sandy edge in a large basin with a size of approximately 15km × 20km. In geomorphological relation, the region is presented by a series of sand massifs divided by loamy plains. The basin bottom, where the test site is located, consists of clay and heavy loams alternating in some places by low-powered deposits of sand. Atmospheric fallout is accumulated in low relief, these reservoirs are completely dry from June to September. Ground waters underlie depending upon the season, the relief and geological structure at a depth of 0.1–16m from the surface. Soil cover of the territory is brown desert and salt soil. There is a definite differentiation on the genetic horizons: humus, carbonate, salt are characteristic of a brown soil profile. The thickness of the soil layer at the described area is insignificant (2-6 cm). The region's climate is strongly continental, arid; it is formed under the predominant influence of arctic, Iranian and Turan air mass. The average annual fallout is more than 200mm. The average annual air temperature is 7.3°C. North (Summer) and East (Winter) parts steady wind at a speed of 15–20m/s and more are characterized for the region.

3 Characteristic of radioactive contaminated areas at the site

Site radiation control was carried out by a detailed method through ground survey on a grid 10m × 10m (orthogonally related directions with distances in 10m between them). At the surface of A-I, A-II, A-III, A-V and A-X sites radioactive contaminated 0.4–0.6 m^2 areas (spots) are defined. These are the potential source of the second environment pollution. More contaminated sites are A-II and A-X, as by site quantity so by value of exposure (equivalent) dose rate (EDR), Figure 4. A-I, A-II, A-III, A-V and A-X sites in total amount of 200 spots are related to removal.

Soil samples selected from these spots were analyzed using alpha, beta and gamma-spectrometry methods. The following artificial radionuclides with specific activity: Am-241 up to 5.0 Bq/kg, Cs-137 up to 9 000 Bq/kg, Sr-90 up to 17 Bq/kg, Pu-239+240 up to 2.0 Bq/kg were identified; under the Radiation Safety Standards of Kazakhstan, these soils are not related to the radioactive waste category but to disposal at general industrial waste storage. In addition, Cs-137 content soil can determine the overrunning of the maximum permissible public dose because its specific activity is close to the criterion for radioactive waste (10^4 Bq/kg).

For the study of Cs-137 with soil particles degree of association the experiments on its leachability were carried out. Experiment results showed that no less than 95% Cs-137 was in linked form, which is not leachable with ammonium-acetated buffered solution [5].

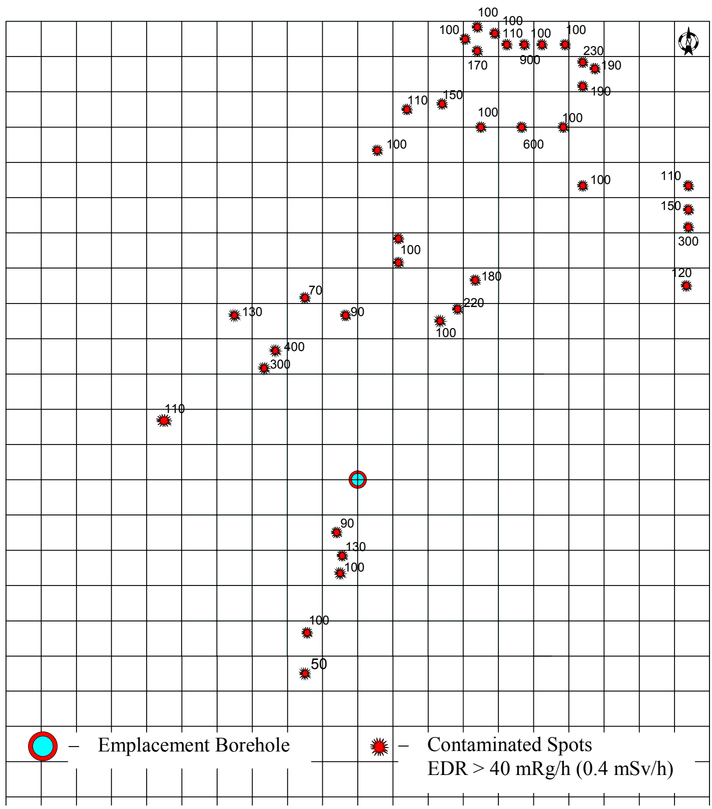


Figure 4: Radiation background at A-X site.

Even under infiltration waters, entry into the storage Cs-137 will not be leached by ground waters that exclude its further migration in solely unfavorable conditions.

Spots scheduled for elimination were located at topsoil. The total spot area does not exceed 90m^2. These are formed during the process of cavity opening, under radioactive ground recovering, sampling of salt solution and different technological operations execution. The organization operating the test site (Russia) temporarily placed all the mentioned contaminations in a trench within the site and buried it in cavities A-III and A-X [6]. Analyses of layer-by-layer samples from local spots showed that basic activity was focused in the upper 10–15cm layer. Besides, maximal concentrations are defined in soil layer 5–10 cm, in comparison with 0–5 cm layer, from which the part of radionuclides was brought into the environment by the wind as well as the part that was washed by rain and melt waters. However, it is not possible to exclude the presence of radioactivity in more deep layers as excavation and re-disposal from the trench into cavity were carried out without proper quality controls. Consequently, soil cover was technologically damaged.

4 Elimination of hot spots

The sample of radioactive soil was drawn by 5 cm layers at the bottom of the excavation at a level of 30 micro-roentgens per hour before the moment of fixing. In cases when EDR values were neither decreasing nor increasing, the soil sample ceased at a depth where the backfilling with clean soil with the 10cm half-value layer made it possible to obtain EDR values in 30 micro-roentgens per hour on the daylight surface (Figure 5).

Figure 5: Dosimetry of pits.

Selected contaminated soil was intensively intermixed in special pans and than placed in a plastic sack. EDR was determined at the surface of each sack. Composited samples with high EDR were selected from the sack in order to determine the specific activity of Cs-137. Sacks were marked and data on the site index, position coordinates of soil sorting, EDR at the surface of sacks, volume and weight of soil from each radioactive spot was entered into the special log. Soil samples were transferred to the Institute of Nuclear Physics (Almaty city) where Cs-137 content was determined under laboratory conditions by meands of gamma-ray spectrometry.

Sacks with contaminated soil were tightly placed in a burial trench of the following size: (width × depth × length): 1m × 3m × 6m. After placing all sacks, the exposure dose rate was measured on the ground level under the open storage and also at the surface of all packed sacks. The trench was backfilled with clean soil which was excavated during the preparation of the storage, the thickness of the layer was about 1m so that the Exposure Dose Rate (EDR) would not exceed 30 micro-roentgens per hour on the daylight surface. The trench coating was tamped and did not exceed 0.5m from the daylight surface.

Reference pegs were located in storage nooks for switching of these points in plan of regular monitoring. After samples of radioactive soil and its disposal, newly formed excavations were filled with clean soil that was excavated from the trench and imported additionally. The backfilled soil did not exceed the relief of the site within limits of 10cm. The results of EDR measurement showed that

the radiation situation in sites which had spots of radioactive contamination had significantly improved after taking some action on the decontamination of areas of sites. Values of EDR amounted to 900 micro-roentgens per hour before room works and they were between 20 to 30 micro-roentgens per hour after room works (Table 3).

Figure 6: Depositing of radioactive soil in trench storage.

Table 3: Dosimetry in A-X site after cleansing, EDR *(mR/h)*.

before elimination		after elimination		before elimination		after elimination		before elimination		after elimination	
April		July	August	April		July	August	April		July	August
No of points	EDR	EDR	EDR	No of points	EDR	EDR	EDR	No of points	EDR	EDR	EDR
1	100	25	25	9	400	30	30	17	150	31	30
2	270	30	30	10	110	30	30	18	150	25	25
3	280	25	30	11	100	31	30	19	120	30	26
4	110	25	27	12	150	30	30	20	900	23	25
5	170	30	30	13	130	32	32	21	100	27	24
6	190	30	32	14	200	33	32	22	110	30	28
7	110	30	30	15	170	30	31	23	100	23	25
8	160	30	32	16	150	29	28	24	110	24	24

5 Conclusion

The total weight of the excavated radioactive soil was 5060kg. The average concentration and total activity of Cs-137 in the removed soil was 5000Bq/kg and 3×10^{7}Bq. The research carried out and the engineering works on decontamination and the removal of radioactive contamination from the area of the Azgir test site made it possible to for the radiological situation at the test site and in surrounding areas to be improved. After performing the abandonment

operations, danger to personnel and population from radiation doses which were previously over the safe rating values for radioactive soil contamination were excluded.

References

[1] S.A. Berezin, D.G. Gilmanov, Zh. R. Zhotabaev, K.K. Kadyrzhanov and others. NNC RK Bulletin, 3rd edition, 2001.

[2] A.S. Krivokhatsky and others. The main characteristics of radiation situation after completion of underground nuclear explosions in the interests of national economy in the salt- dome deposit land of Bolshoi Azgir (Kazakhstan). Preprint RI-223 SPA "Radium Institute", CSRIatominform, 1992.

[3] E.Z. Akhmetov, D.G. Gilmanov and others. AZGIR testing site. Preprint 1-96, Institute of Nuclear physics, Almaty, 1996.

[4] A.S. Krivohatsky and others. Radionuclide effects due to peaceful underground nuclear explosions in the salt deposit land of Bolshoi Azgir. CSRIatominform Bulletin, M, 1993, No 9.

[5] E.Z. Akhmetov et al. The granulometrical fractions of soil, distribution of Cesium-137 on its and the fastening of soil surface contaminated by radionuclides at locations of the "Azgir test site". International workshop. Effect of ionizing radiation on ecological situation of countries from Caucasian region and Caspian Sea basin. Azerbaijan, Baku, 2002, p. 92.

[6] Analysis and phased improvement of radiation situation near the object "Galit". Report of SPA "Radium Institute" .S-P, 1994.

[7] Suitability evaluation of places and facilities in Degelen and Azgir in the Republic of Kazakhstan as location for radioactive waste burial. Report No D -1/96, NNC RK fund, 1997.

[8] D.G. Gilmanov, E.Z. Akhmetov, J. Ziegenhagen and others. West Azgir Salt Dome as Massif for Locating Radioactive Waste Burial. NNC RK Bulletin, 1st edition, 2005.

Accidental mercury spill in the Andes: forensic neuropsychological evaluation six years later

F. Gonzalez & A. Saldivar
Rancho Los Amigos National Rehabilitation Center, USA

Abstract

On June 2, 2000, 330 pounds of elemental mercury from a local gold mine was spilled over a 45-kilometer stretch of the region's main road. The affected road represents the only route between towns and receives heavy foot traffic. People from the surrounding towns, believing the mercury to be valuable, gathered and carried mercury using any means available. They took mercury into their homes and in many cases handled mercury extensively, even going so far as to try to clean it. Adults and children of all ages were exposed. Many were hospitalized and chelated; some were not. Levels of exposure and symptoms varied, and many victims did not receive treatment or care as local resources were quickly overwhelmed. This paper outlines an approach to a forensic assessment model for a culturally diverse and challenging population. The literature on the neuropsychological effects of mercury on children and adults will be briefly discussed. Recommendations from the World Health Organization for assessment of neurotoxicity and their implementation in the current project will be reviewed. Two pediatric cases are presented from a medical and neuropsychological perspective. Discussions regarding the effects of the spill on the local educational system and community as a whole will be presented.
Keywords: neuropsychology, culture, Spanish, mercury exposure, toxic exposure, forensic, neurotoxicity, chelation, penicillamine, child, pediatric.

1 Introduction

Mercury (Hg) was known to ancient Hindus and Chinese before 2000 B.C. The Greeks used it in ointments, the Romans in cosmetics, and it has been found in 3500 year old Egyptian tombs. It is often a by-product of gold mining and frequently used in thermometers, barometers and pressure sensing devices,

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060711

batteries, lamps, dental amalgam, pharmaceuticals, religious practices, etc. Mercury easily forms alloys with other metals and is frequently used to extract gold from its ores. One of the world's largest goldmines is located in the Andes in a town called Yanacocha, located in the northern part of Peru. In June 2000, while transporting mercury that was a by-product of the gold mining process, 330 pounds of elemental mercury was spilled over a 45-kilometer stretch of the region's main road. The affected road is the only route between towns and is travelled heavily by foot by the local inhabitants, most of whom cannot afford cars. When people in the area became aware that there were glittering silver droplets on the road, they began to gather the mercury thinking it could be valuable. They collected the mercury using bottles, cans, cups, brooms, pockets, even their hands. People took the mercury into their homes and in many cases handled it extensively, with some individuals even trying to clean it. Over 1000 adults and children were exposed. When the local authorities became aware of the spill, attempts were made at cleaning the contaminated areas and treating the exposed individuals. A protracted legal battle ensued that is currently ongoing and numerous medical studies were initiated to assess the effects of the mercury exposure on the affected population. Given the cultural and geographic issues, and limited studies, there were few existing protocols or specific guidelines addressing appropriate neuropsychological evaluation of mercurial exposure. The following paper will discuss the development, application, and results of a forensically oriented neuropsychological assessment of a sample of the exposed individuals.

2 Deficits commonly found with mercury exposure

Elemental/metallic mercury (Hg0) is a silvery metal with high vapor pressure that can exist either as a liquid or vapor form in ambient conditions. Elemental mercury is absorbed through the lungs and crosses the blood-brain barrier, the placenta, and other tissues prior to oxidation. Oxidation changes the elemental mercury into inorganic mercury, which cannot cross the blood-brain barrier and is therefore trapped by oxidation within the brain [1]. Rohlings and Demakis [2] report that exposure to elemental mercury can result in tremors, psychological problems (e.g., erethism manifested as anxiety, timidity, shyness, and mood lability), and higher order neuropsychological impairments (including deficits in attention, intellectual disturbances, visuospatial abilities, cognitive abstraction, reaction time, visual memory, and decreased cognitive efficiency).

Given the nature and history of research into mercury exposure, it has become an excellent experimental model for studying the developmental neurotoxicity of heavy metals. Previous research as cited by Deitch [3] noted substantial deficits in visual perceptual constructional skill, nonverbal memory and abstract reasoning related to exposure to Hg vapors in a home by children. Though follow-up testing at one year revealed some improvement, deficits remained. Interestingly, similar levels of exposure to the parents resulted in no findings, leading the authors to conclude that the children were more sensitive to the exposure and the deficits were most likely caused by demyelination in the

peripheral nervous system, since mercury is known to cause demyelination in the nervous system. Kim Deitch [3], notes, "…children are almost always highly sensitive to environmental contaminants, frequently expressing signs and symptoms of toxicity at doses that fail to produce any outward indications of neurological disease in adults" (p.211).

3 Assessment: principles and approaches

In 2001, the World Health Organization (WHO) developed the Neurotoxicity Risk Assessment for Human Health, [4], a report that contained the collective views of an international group of experts on how to assess individuals exposed to toxic chemicals. The report indicates that, "the assessment of potential neurotoxicity in individuals begins with an individual patient in order to establish a differential diagnosis of neurotoxic disease and to rule-out other etiologies." [4]. The evaluation includes a detailed medical history and examination, and depending on the symptoms and type of exposure, may require other assessment procedures such as neuropsychological evaluation.

Neuropsychological testing has long been used as a method of documenting neurological insult as would be noted in toxic exposure, with numerous researchers and authors describing the methodology and areas that should be assessed [5,6]. The WHO report describes the general domains to be examined, including general intellect, attention, executive function, memory, verbal ability and language, visuospatial and visuo-motor ability, as well as mood and personality. Guidelines for conducting epidemiological studies that are more cost effective and allow for the examination of a larger number of individuals have also been published, however, these approaches often use screening measures that are less sensitive and may under-appreciate pathology or miss sub-clinical syndromes. The advantages of a neuropsychological neurotoxicology assessment are outlined by Hartman [5]. These include: 1) concurrent and predictive validity, 2) safe and non-invasive with no risk to the patient, 3) comprehensive and flexible and can assess a wide variety of cortical and subcortical functions, 4) tests are objective and replicable, 5) relatively inexpensive compared to medical and neurological workups, 6) testing is portable and can fit in one or two suitcases, 7) information is usually complimentary to conventional medical screens, and 8) because of its sensitivity the neuropsychological method can provide early warnings of potential brain damage.

3.1 Approach to neuropsychological testing of rural Peruvian population

For this project, the examiners reviewed the available literature in English and Spanish and consulted with other professionals on the availability of appropriate norms and instruments that were developed for specific use with a Peruvian population. Unfortunately, available measures were primarily used with the urban dwelling Peruvian individuals. Even within Peru, psychological testing of rural people such as those living in the Andes is rare. It is also remarkable that available measures commonly used by the local professionals were tests

developed in the United States (WISC-R [7], WPPSI [8]), translated into Spanish but not re-normed on the Peruvian population. Given the questionable fairness of the test and availability of newer technologies and tests, the examiners developed adult and child protocols that were modelled after the WHO multi-center longitudinal study that explored the nature and prevalence of HIV associated neurological disorders in persons living in various geographic and social contexts [9]. This line of research developed measures that were more culture fair and with sensitivity that holds across diverse cultures and consistent with the International Test Commissions Guidelines for Test Use [10].

Demographic trends have shown Hispanics to be a growing segment of the population in the United States, driving efforts to create neuropsychological measures that can be used with Spanish speakers from various countries, taking into account education, age, and cultural variables [11–14]. Of these measures examiners opted to use the NeSBHIS [11] and the Bateria III [13], as well as other tests of motor functioning for examination of adults. For the children, WISC-IV [13] and Bateria-III, along with other measures listed below were chosen.

3.1.1 Teacher interviews

The protocol was piloted in 2005 and teachers within the local affected schools were interviewed at various grade levels to further examine the appropriateness of the battery and review other possible sources of information. In their system, teachers start with a set of students in the first grade and follow them through the sixth grade, such that the teacher would have the same students for six consecutive years before returning to teach a new group of first graders. All the teachers had extensive teaching experience and were present during the exposure. Three teachers were interviewed: Mrs. P (1st grade teacher), Mr. Q (3rd grade teacher), and Mrs. R (4th grade teacher).

3.1.1.1 Teacher interviews #1 Mrs. P reported a sudden change in behaviour and performance among her students in the days following the mercury spill. On the day of the spill and for several days following many of the children brought mercury into the classroom or routinely played with it. From her recall, all the children in her class had direct contact with the mercury.

Mrs. P described having numerous students with visible rashes, complaints of headaches, and notable changes in their behaviour and attention. The children became very inattentive and had difficulty concentrating. They became more emotional and frequently cried or became irritable. Most of the children developed poor appetites and appeared depressed. Academically, they started having difficulty retaining information regardless of what approach to teaching was employed or how much their studying time was increased. Many of them were also noted to have difficulty with language during this period.

Mrs. P explained that where the spill occurred in June, the academic year begins in April and ends in December. In the year of the mercury spill, only 20 of her 37 students completed 1st grade requirements; over 45% of the class had to repeat the 1st grade, whereas in a normal year she typically holds back

approximately 10% of her students. She reported that this phenomenon was not isolated to her classroom, but that the failure rate school-wide also jumped that year with approximately 30% of the school's students not graduating to the next grade.

3.1.1.2 Teacher interviews #2 Mr. Q reported that following the mercury spill the children started to develop numerous physical problems, including rashes, headaches, vomiting, general weakness, and fatigue. Children began to miss more days of school than usual and academically appeared to become very lazy. He noted problems related to poor attention and memory, which appeared to affect their abilities in all academic pursuits. A general drop was noted in their language skills, general communication, and ability to do arithmetic. Out of a class of 26 students, Mr. Q failed eight. Mr. Q also noted that in the years to follow many of the children's problems continued. They did not seem able to recover from the exposure. In many cases, the children were pulled out of school by their parents because of their inability to perform or at the children's request that they be taken out of school. He said that at the time of the spill 350 students were enrolled at his school, but within a few years, enrollment dropped to 227.

3.1.1.3 Teacher Interviews #3 Mrs. R reported that in the year of the mercury spill she failed nine of her 22 students (over 40%). She reported that many of the children described numerous physical complaints in the time immediately following the mercury exposure, including rashes, headaches, weakness, and fatigue. As the children had already developed many academic abilities and skills, they were initially able to continue or maintain what they had already learned. Academically, problems were noted starting in October and November as they got further into the academic year and were challenged with new material and development of novel skills. All the students demonstrated lower achievement and difficulty progressing.

4 Case presentations

Two cases are presented: Andres and Anita. They were evaluated using the following neuropsychological assessment protocol: Clinical Interview, Child Neuropsychological History Questionnaire (a structured clinical interview), Wechsler Intelligence Scales of Intelligence-IV Spanish (WISC-IV) (selected subtest of intellectual and cognitive ability), Bateria III (selected subtests of a test of achievement and aptitude), Rey Complex Figure Test (visuoperceptual and visual memory test requiring copying a complex drawing, then recalling the drawing from memory at later points), Children's Color Trails Test 1 and 2 (tests of simple sustained attention, ability to shift set, and divided attention), Beery-Buktencia Developmental Test of Visual-Motor Integration, 5th Edition - Short Form (VMI) (a measure of visual perception and graphomotor construction), Beck Youth Inventory-II (self-report measure of depression and anxiety), Grip Strength (test of gross motor hand strength), Grooved Pegboard (test of complex fine motor ability), and Finger Tapping Test (test of repetitive fine motor speed).

The preceding measures were chosen to reduce the effects of cultural and language factors, and were given by a bilingual (Spanish/English) and bicultural (Mexican-American) examiner in the presence of a Peruvian neurologist.

4.1 Andres's exposure

On June 2, 2000 Andres was 2 years and 8 months old when he came in contact with elemental mercury while playing in the street near his home and later within his home. His mother reported that she and the family collected approximately 15 kilos of mercury from the roads near their home using spoons and cartons, and believing the mercury to be valuable, they brought it home. In the home they placed it on a table to separate the dirt from the mercury. Andres's mother explained that there was a lot of dirt mixed in with the mercury and that she (with the help of 14 family members) sifted out the dirt. Andres was hospitalized in June of 2000 for the effects of mercury exposure, during which time he received chelation therapy. A toxicological blood analysis done on June 12, 2000 showed a mercury level of 69.94 ug/dl. Prior to the exposure, he was described as a verbal, outgoing, energetic, animated and engaging child who talked about wanting to become an attorney. After the exposure, his mother described him as having significant problems and difficulties, that included changes in his behaviour, personality, increased irritability, and being withdrawn.

4.1.1 Andres's neuropsychological evaluation

Andres's developmental and family history was unremarkable. He was born full term in a normal delivery, had reached developmental milestones at expected age levels, and was described as an academically average student. No systematic diseases, neurological disorders, psychiatric problems, or learning disorders were reported for Andres or his immediate and extended family.

At time of evaluation, Andres was 7 years 11 months. Andres' mother described him as having recurrent stomach problems, pain in his joints (especially his knees), and difficulties in school. In a structured interview Andres' reported problems were separated into the following three categories: 1) cognitive, 2) physical, and 3) emotional. Cognitive problems included: difficulty with novel problem solving, following directions, switching tasks, problems with decision-making, learning new tasks, poor concentration and distractibility, word-finding difficulty, poor articulation ("more like a 5 year-old"), memory problems, forgetting where he puts things, poor writing and fine motor control, easily frustrated, hypersensitive to light and sound, and very fidgety. Physical problems included: difficulty with daily headaches, urinating in his pants and bed on a daily basis, tremors, difficulty with balance, skin rashes, bone pain, and dizzy spells. Emotional problems included: sadness, poor appetite, anxiety, nervousness, difficulty falling asleep, fearful, shy and withdrawn, overly emotional, crying easily, and easily losing his temper. He is currently failing in school and has asked his mother to drop out of school.

On testing, refer to table 1, Andres appears to be in the average range of intellectual functioning. Significant difficulty in the area of nonverbal reasoning was noted, with performance in this domain being variable and ranging from

borderline to average. On a visual memory task he demonstrated impaired recall and low average recognition, indicative of a retrieval deficit. Andres demonstrated variable mental control, with impairments in passive and divided attention. Language skills were depressed, with verbal fluency (i.e., rapid access to verbal information) in the impaired range, well below what would be expected. Verbal recall at the word and paragraph level, visuospatial and visuoconstructional abilities, and fine motor abilities were intact. His psychological profile evidenced symptoms related to depression.

4.2 Anita's exposure

At the time of the mercury spill Anita was 1 year and 1 month old. She was exposed to mercury while playing in the street near her home, within her home, and through nursing. Mrs. M., Anita's mother, reported that she and her family collected mercury from the roads near their home. Anita and her mother played with the mercury, put it on their bodies and stored the mercury in their home. When the mining company became aware of the mercury spill the company set up locations to collect the mercury and reward the individuals who returned it. When Mrs. M. attempted to return the mercury she was told that it had too much dirt and to return after she had cleaned it, and was instructed to boil the dirt out. Mrs. M. boiled the mercury over an open wood fire in her home where ventilation was limited; the walls of her home are constructed of the local dirt with dirt floors. Anita developed a rash, problems breathing, head aches, change in personality (irritability), problems sleeping and had blood in her feces. She was hospitalized in June of 2000 for the effects of mercury exposure and received chelation therapy and was classified as "red code" with risk of organ damage because of her young age. Her body was still developing and was in a critical period. In February 2002 she showed low creatinine depuration denoting renal compromise. In August 2004 Anita continued having breathing problems with wheezing, chronic headaches and chronic dermatitis. Notes by local medical providers noted her to have normal intelligence but with deficits in visual organization and concentration.

4.2.1 Anita's neuropsychological evaluation

Anita's developmental, medical and family history was unremarkable. At time of evaluation she was 6 years old. Her teacher described her as having significant difficulty in reading, writing and academically falling further behind than her peers and expected that over time the gap would widen. She took 3 to 4 hours per day to complete her homework and required constant supervision and tutoring. Anita was notably reserved, less social than her peers, and presented as depressed. She was quick to react and was often aggressive.

On interview, Mrs. M reported that Anita was very active and distractible. In the period following the exposure Anita was described as very withdrawn and difficult to sooth, but over the years received extensive intervention and treatment, through schools, tutors, doctors, and therapists. A level of care and attention that is not typical for most of the individuals exposed. Cognitive problems endorsed: difficulties with novel problem solving, following directions,

switching tasks, learning new tasks, poor concentration and distractibility, easily losing train of thought, and memory problems. Physical symptoms endorsed: occasional headaches, urinates in her pants and bed, skin rashes, bone pain (in her joints which often makes her cry), and very fidgety with difficulty standing still. Emotional problems reported: poor appetite, difficulty with sleep initiation, periodic nightmares, depression, anxiety, nervousness, shy and withdrawn, emotional, crying easily, easily loses temper, aggressive and easily frustrated.

Table 1: Interpretation of neuropsychological testing.

INTELLECTUAL	**Anita**	**Andres**
Raven’s Colored Progressive Matrices	Average	Average
Digit Span-WISC-IV	Impaired	High Average
Coding-WISC-IV	Borderline	Average
Block Design -WISC-IV	Average	High Average
Matrix Reasoning-WISC-IV	Average	Impaired
Symbol Search-WISC-IV	Average	Low Average
Processing Speed Index	Low Average	Average
LANGUAGE		
Picture Vocabulary-Bateria-III	Low Average	Impaired
Verbal Fluency-Bateria-III	Average	Average
Oral Expression-Bateria III	Low average	Low Average
MENTAL CONTROL		
Color Trails 1	Impaired	Borderline
Color Trails 2	Impaired	Average
Digit Span-WISC-IV	Impaired	Average
Coding-WISC-IV	Borderline	Borderline
VISUOSPATIAL		
Raven's Progressive Matrices	Average	Average
Rey Complex Figure Copy	Discontinued	Average
Block Design-WISC-IV	Average	High Average
Bender-Gestalt (Error)	Impaired	Intact
MEMORY-VISUAL		
Rey Complex Figure Immediate	Discontinued	Impaired
Rey Complex Figure Delayed	Discontinued	Impaired
Rey Complex Figure Recognition	Discontinued	Low Average
Design Recall-Bateria-III	Low Average	Average
MEMORY-VERBAL		
Story Recall -Bateria-III (3)	Low Average	Impaired
Story Recall (Delayed)-Bateria-III (12)	Impaired	Impaired
Word Memory-Bateria III	Average	Low average

On testing, Anita presented with average intellectual ability. Language skills were in the low average to average range. Mental control was generally impaired and reflecting difficulty with attention. She did well on non-motor visuospatial tasks. When tasks required the integration of fine motor and visual spatial abilities she demonstrated severe deficits, not being able to complete one task. On a similar task not requiring fine motor ability, she performed in the low average range. Immediate short-term verbal memory was grossly intact, while delayed verbal memory was impaired.

5 Discussion of neuropsychological test results

Testing provided meaningful data, with the selected tests demonstrating appropriateness in terms of language (appropriate Spanish) and findings consistent with the literature and collateral information. In reviewing Andres's and Anita's performance, it is interesting to note the scatter or variability in scores within the various domains assessed (mental control, intelligence, language, visual spatial and memory). Their scores indicate potential for average to high average performance in all domains; however, on specific tests measuring different aspects of those same domains, performance is impaired. Examination of the commonalities suggest possible executive dysfunction, with what appears to be planning, organizational and monitoring difficulty, as well as deficits or processing problems in the area of information retrieval, all of which would significantly impair their ability to function, especially in school. It is well within the realm of medical probability that current difficulties reflect the effect of mercury on Andres' neurodevelopment, given that his exposure to mercury occurred at a time of rapid neurodevelopment. The noted difficulties are consistent with collateral information and reflective of deficits that will have real world implications. They suggest children having difficulty functioning within their school and home environments, despite, in Anita's case, a heroic level of intervention. Given the time elapsed since their exposure and the limited and questionable improvement it is highly likely that both will have persistent impairments and disabilities for life. Will they continue to fall behind, maintain or improve? Only further research can answer that question. At a societal level, the implications of the mercury spill are immeasurable, especially if one considers the number of children affected, the rate of school drop out, and the loss of intellectual ability in a community with limited resources for remediation. These children are growing up with valid psychological fears for their future and well being.

References

[1] ATSDR (1999) Toxicological profile for mercury (update). Atlanta, GA, US Department of Health and Human Services, Public Health Services, Agency for Toxic Substances and Disease registry, March.

[2] Rohling ML & Demakis GJ (2006) A meta-analysis of the neuropsychological effects of occupational exposure to mercury. The Clinical Neuropsychologist, 20, pp 108-132.
[3] Deitch K (2000) Environmental toxicants and psychological development. In Yeates, Ris & Taylor eds. Pediatric Neuropsychology: Research, Theory, and Practice. Guilford Press, New York
[4] WHO (2001) Environmental Health Criteria 223 Neurotoxicity Risk Assessment for Human Health: Principles and approaches, Geneva.
[5] Hartman DE (1995) Neuropsychological toxicology: Identification and assessment of human neurotoxic syndromes, 1st ed. New York, Pergamon Press.
[6] White RF & Proctor Sp (1995) Clinico-neuropsychological assessment methods in behavioural Neurotoxicology. In: Chang LW & Slikker W. ed Neurotoxicology: Approaches and methods. New York, Academic Press, pp 711-726.
[7] Weschler D (1976) Weschler Intelligence Scale for Children-Revised. New York: Psychological Corporation.
[8] Wechsler D (1967) Wechsler Preschool and Primary Scale of Intelligence -Revised. New York: Psychological Corporation.
[9] Maj M, D'Elia, L, Jannsen R, Zaudig, M, Uchiyama, C, Starace F, Galdererisi S & Chervinsky, A (1993) Evaluation of three new neuropsychological test designed to minimize cultural bias in the assessment of HIV-1 Seropositive persons: A WHO Study. Archives of Clinical Neuropsychology, 8 123-135.
[10] International Test Commission (2000) International Guidelines for Test Use, Stockholm.
[11] Ponton MO, Satz P, Herrera L, Ortiz, F, Furst, C & Namerow, N (1996) Normative Data stratified by age and education for the neuropsychological screening battery for Hispanics (NeSBHIS): Initial report. Journal of the International Neuropsychological Society 2 (2), 96-104.
[12] Artiola IF, Hermosilla D, Heaton RK & Pardee RE (1999) Manual de normas y procedimientos para la Bateria Neuropsicologic en Espanol, Tucson, m Press.
[13] Woodcock RW, Muñoz-Sandoval AM, McGrew, KS & Nancy M (2005) Bateria III Woodcock-Munoz, Riverside Publishing.
[14] Wechsler D (2004) Wechsler Intelligence Scale for Children-IV Spanish. San Antonio: Psychological Corporation.

Predicting the intensity of wind-blown removal of dust and sand in the Turkmenistan desert

V. Kostiukovsky[1] & A. Arnageldyev[2]
[1] *Albert Katz Department of Dryland Biotechnologies, Jacob Blaustein Institutes for Desert Research, Ben-Gurion University of the Negev, Israel*
[2] *Turkmenian Institute of Transport, Ashgabat, Turkmenistan*

Abstract

Wind plays a major part in the dynamics of solid non-organic material in the deserts of Turkmenistan. The main direction of movement of sand and dust in this region, due to atmospheric transference, is from north to south.

Data from the literature and the results of field investigations testify that the majority of wind-blown material comes from desert dunes without vegetation (combined area 37,000 km^2) - more than 29 billion tons per year. Material blown from sandy desert terrain anchored by vegetation (206,000 km^2) amounts to 16.5 billion ton/year. Nine hundred and sixty million ton/year comes from the saline plains (20,000 km^2), where the surface is crumbly due to intensive geo-chemical processes, and 352 million ton/year from stony rubble surfaces (45 km^2). Consequently, nearly 50 billion ton/year of sand and dust are implicated in wind-blown solid particle removal in the deserts of Turkmenistan.

The influence of human activity in the deserts during the 20th century came close to being catastrophic. More than 60% of the territory of Turkmenistan (200,000 km^2) is currently undergoing desertification. More and more dust and sand is blown from overgrazed pastures, roads, arable fields, settlements, etc. The major sources of the blown sand, dust and salts are the exposed beds of the Aral Sea and other water bodies that are drying up after extraction of their water for irrigation (nearly 45,000 km^2).

Natural warming of the climate is intensified by the hotbed effect, in which increased human activity has led to an increase in average annual air temperatures by 6-9 °C and an abatement of precipitation by 50%. The result may be expansion of desertification to 90% of the deserts of Turkmenistan, and an increase in the quantity of wind-blown sand and dust to 65 billion ton/year by 2050.
Keywords: wind-blown transportation, soil types, Turkmenistan Desert, desertification, solid substrate.

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060721

1 Introduction

Wind-blown transportation of dust and sand plays a major role in the movement of solid non-organic material in the desert zone, including the deserts of Central Asia and particularly in the extra-arid desert of Turkmenistan (where evaporation capability exceeds atmospheric precipitation by 100:1). The intensity and spread of desertification due to both natural causes and increasing human pressures (Alternative Strategy [1]; Mabbutt and Wilson [2]; UNCCD [3]) contribute to such wind-blowing activity.

The direction of movement of blown material depends on the atmospheric circulation. The major movement of air masses in the desert of Turkmenistan is from the northern part of the horizon (N, NW, NE) to the southern part. Consequently, most blown material travels in a southerly direction, which is indicated by the orientation of the sandy relief forms. By determining the quantity of material involved in the wind-blown process as a result of different environmental conditions, we can calculate a future change in this quantity, related to natural changes in climate and human influence.

2 Materials and methods

The quantity of the blown and transported dust and sand by wind has been investigated in the literature (e.g. Thomas [4]; Fedorovich [5]; Arnageldyev [6]; Bagnold [7]), using data from the hydro-meteorological service and from field investigations. The transportation of blown material in the field was defined using standard dust- and sand-catchers, catch pits, scotch-tapes etc. The change in soil surface beneath the wind-blown process was measured by continuous monitoring using geodesical instruments along fixed profiles in different soil types (sandy, loamy, stony, salty, etc.), and by continuous monitoring of metal and wooden rulers, inserted into the same substrates. Additional methods of investigation included measuring the extent of wind-blowing of desert plants, roads, buildings, archaeological remains, etc. Areas with different surface types were defined using topographical, soil and geomorphological maps and space information from various years.

3 Results

Five types of soil surface were defined by their resistance to wind erosion: bare sand with no vegetation, sandy soil anchored by vegetation, loamy soil, stony-rubble soil and saline plains (Fig. 1). The major source of blown dust and sand in the area is the sandy dunes with no vegetation. Some of this is natural in origin, i.e. places with minimal precipitation and very strong winds, at the seashore, etc. These types of soils are found on the shores of the Caspian Sea, mainly near Cheleken peninsula, in the narrow windy passes between the Large and Small Balkhan Mountains, along the Amudarya River, in part of the Amudarya valley, near the border of the mountains and plain, and in some places in the central part of the Karakum desert.

Apart from the above natural sources of sand and dust, many desert dunes have been formed under continuous human pressure from the prehistoric period (Neolithic remains, visible even today) until the present (oases, wells, settlements, industrial and mining regions, railways, highways, etc.). It is very difficult to identify dust and sand blown from desert dunes due to the constant exchange of particles in the wind. A detailed analysis of topographical maps allowed us to determine the layer of soil blown up from dune sites as 10-50 cm per year. The area of sand without vegetation in the desert of Turkmenistan covers 37,000 km^2, yielding almost 29,600 billion tons of solid material per year (Table 1).

Table 1: The quantity of solid non-organic material blown up in the desert of Turkmenistan.

Soil surface type	Area (1000 km^2)	Average thickness of layer of blown material (mm/ year)	Total quantity of wind-blown material (billion tons/year)
Sands without vegetation	37	100-500	29,600
Sands anchored by vegetation	206	3-10	16,480
Loamy soil	45.5	1-5	288
Stony-rubble soil	44.3	1-7	352
Saline plains	20	10-50	960
Total	**352.8**	**-**	**47,680**

The layer of blown soil in the sandy desert anchored by vegetation is less than the layer of blown soil in the bare sands, though the area covered by anchored sands is five times greater. The measurements show that the change in surface thickness in sands covered by grass and shrubs is not more than 1 cm per year. Most of the dust and sand carried by the wind falls in wind shade of plants and is anchored by stems, roots and other products of plant activity. The vegetation-covered areas in Turkmenistan cover the largest part of the desert (more than 200,000 km^2), while the quantity of substance blown up from there is less than from the bare sands (16,480 billion tons per year). Dust is the predominant substance blown from the anchored sands, in contrast to the bare sands where, over thousands of years of dune shifting, much of the dust has been carried off by the wind.

The dust is mainly material blown from loamy surfaces. This type of desert soil is mainly found in the foot-hills of the Kopet-Dag and Balkhan Mountains, in the Tedjen and Murgab River deltas, which do not reach the sea, and in the Atrek River, which reaches the Caspian Sea only in wet winter/spring periods. Many loamy surfaces ("takyrs") are typical of the drops found amongst dunes and ridges in the sandy desert (Fig. 1).

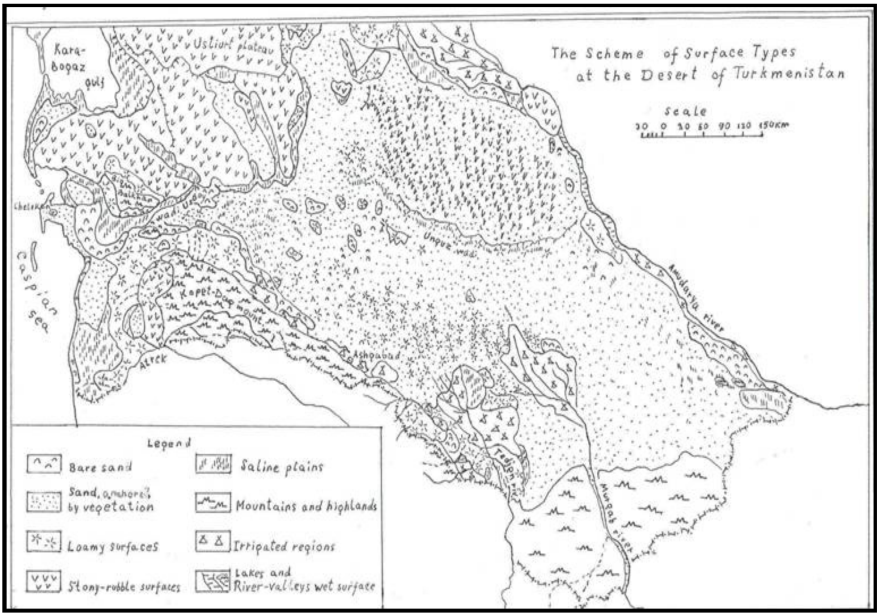


Figure 1: Scheme of surface types in the desert of Turkmenistan.

The loamy soil possesses great adhesive strength and is resistant to being blown away by the wind. Due to the admixture of sandy, dusty and alevrite particles with clayey and loamy ones, and human and animal activity which destroys the surface, the wind blows up a layer of loam between < 1 mm and 5 mm thick per year. As a result, only 288 billion tons per year are blown from 45,500 km^2 of loamy surfaces.

Stony-rubble desert occupies mainly the north-western part of Turkmenistan – from the Kara-Bogaz gulf to Sarykamysh Lake and Wadi Uzboy. The total area is nearly 44,300 km^2. Continued intensive destruction of stony highlands by wind, water, temperature contrasts, and plant, animal and human activity constantly produces new quantities of sand, dust and rubble which involved in the process of wind erosion. Thus from an area of 44,300 km^2 of stony-rubble desert, the wind blows up more than 352 billion tons of dust, sand and rubble per year (Table 1).

One of the most active sources of blown material is saline soil. Large quantities of such soil occur near the south-eastern shores of the Caspian Sea, in the foot-hills of the Ustiurt plateau, in the dry wadis Uzboy and Unguz, in the Murgab and Tedjen River deltas and in the isolated drops between sandy ridges. Saline plains, formed under human influence, are typical of areas containing irrigated fields, canals, water bodies etc. These large areas of saline plains have been exposed during the drying up of the Aral Sea. The soils in the saline plains are chemically active, due to the large quantity of salts they contain which break up the surface. Particles of broken up sand, dust and loam are blown up from the salty plains by the wind. The total area covered by saline plains is less than the

other soil surface types (nearly 20,000 km^2), but the quantity of material blown up from these soils is more than that blown up from the loamy and stony soils together – 960 billion tons per year (Table 1).

We can see that the total quantity of solid non-organic material blown up in the desert of Turkmenistan is between 45 and 50 x 10^9 tons per year. Most of the blown material is carried in the lower layers of wind flow, and is dropped at a distance of several centimetres to hundreds of meters from its origin. Thus the material moves with the wind within the borders of the Turkmenistan desert. At the same time however, particles with a diameter of less than 0.25 mm can fly with the wind upward to an altitude of 4-10 km and be carried hundreds and thousands of kilometres. Dust particles from the Sahara and Sahel Deserts have been located in the island of Barbados (Kovda [8]). The reddish dust from Sahara that fell during the years 1970-1980 reached Ashgabat and formed a layer up to 10 cm thick in the streets.

The average content of small particles which are present in wind flows over all surface types is 0.5-1% (our measurements and data from the literature). This allows us to put the quantity of solid material transported out of the Turkmenistan desert at 220-250 billion tons per year.

The development of future wind erosion will depend on several global and local factors:

1. Global warming due to the hotbed effect, which will increase aridity in desert zones.
2. Increasing human activity, including upsetting of the natural equilibrium.

All of these factors are causes of desertification, the scale of which is now universal. In the desert of Turkmenistan the consequences of desertification will be an enlargement of the area occupied by sandy dunes without vegetation, destruction of soils by cattle, railways, vehicles, industry and construction, etc.; increase in soil salinisation from irrigation, construction of canals, water bodies, etc.; reduction in river flow rates and drying up of lakes.

Bare sands will be formed in the sandy desert – around roads, settlements, industrial sites, areas of cattle concentration, at the dry beds of the Aral Sea and many lakes and water bodes in its vicinity. Accurate prediction of such a process is not possible, because many of the causes develop very slowly, but the current trend enables us to quantify the enlargement of the bare sand area at 350-500 km^2 annually, and blowing up of sand and dust at between 3000 and 5000 billion tons per year (Table 2).

As the area of bare sand enlarges, it takes place of the anchored sands, whose area decreases correspondingly. As a result, the removal of sand and dust from the anchored soils falls every year by 150-300 billion tons.

The area covered by saline soils grows very rapidly under intensified irrigation, increased water use for industry and domestic purposes and the disposal of wastewater into the desert. Nearly 30,000 km^2 of saline plains appeared outside the borders of Turkmenistan in the dry bed of the Aral Sea, and several thousand sq. km appeared inside Turkmenistan around the Aral region. Thus the area covered by saline plains is increasing rapidly and this trend may be

set to continue for quite a while. Moreover, not only are dust and sand blown up into the atmosphere from these saline areas, but also billions of tons of salts.

Table 2: The change in quantity of solid non-organic material blown up in the desert of Turkmenistan under conditions of desertification.

Soil surface type	Area in year 2000 (1000 km^2)	Quantity of material blown up from 1000 km^2 (billion tons/year)	Change in area, (1000 km^2/ year)	Quantity of material blown up (billion tons/year)		
				Year 2000	Year 2010	Year 2020
Sands without vegetation	37	800	+0.3 - 0.5	29,600	32,800	33,600
Sands anchored by vegetation	206	80	-0.3 - 0.5	16,480	16,240	16,080
Loamy soil	45.5	6.3	-	288	409.5	500.5
Stony-rubble soil	44.3	8	-	352	443	531.6
Saline plains	20	48	+0.2 - 0.4	960	1104	1152
Total	**352.8**	-	-	**47,680**	**50,194**	**51,864**

The change in areas covered by loamy and stony-rubble deserts is minor (some settlements, industrial sites, irrigated fields, water bodes), but they are exposed to pressure from development of large roads, railways and off-road vehicles, and the soils are destroyed during construction of pipelines, powerlines, etc. This disturbance in the surface consequently leads to an increase in the layer that is exposed to the wind in industrial regions - up to 10-15 cm, and wind-blowing of this layer increases from 6-8 billion tons per year per 1000 km^2 to 8-12 billion ton per year.

Under such desertification processes the total quantity of wind-blown sand and dust in the desert of Turkmenistan, now ranging from 45-50 x 10^9 ton/year, may reach 47-53 x 10^9 ton by the year 2010 and 49-54 x 10^9 by the year 2020. Strategies for regulating wind erosion and importation of dust and sand to the atmosphere are being jointly worked out by international programs for combating desertification (Alternative Strategy [1] and others). They include regulation of grazing, ceasing the use of off-road vehicles, promotion of progressive methods of water use and soil cultivation, reduction in the quantity of wastewater, industrial and domestic effluents. Afforestation of destroyed lands and preservation of natural and man-made vegetation as a means to combating wind-erosion are especially important.

4 Conclusion

Wind-blown erosion plays a major role in the arid regions of the world, including the extra-arid desert of Turkmenistan. The largest quantities of sand and dust (296,000 billion tons per year) are blown from the desert dunes with no vegetation, and the smallest quantities (288 billion tons per year) from loamy surfaces, due to the great strength of adhesion between shallow particles in the soil.

All soil surface types present in the Turkmenistan Desert together supply the atmosphere with 45-50 x 10^9 tons of dust and sand per year. Most of this falls inside the desert of Turkmenistan, and only 220-240 billion tons of particles with diameter less than 0.25 mm are carried out to other regions annually in the upper layers of atmosphere.

The process of desertification under climatic changes and intense anthropogenic pressures leads to more active wind erosion and enlargement of areas covered by bare sands and saline soils. The result of this process will be an increase in wind-blown removal of dust and sand to 47-53 x 10^9 ton by the year 2010 and to 49-54 x 10^9 ton by the year 2020. Thus we may expect an increase in the quantity of solid non-organic material present in the lower layers of the atmosphere in Turkmenistan.

Implementation of rational measures for combating desertification will lead to a mitigation of the negative consequences of wind erosion.

References

[1] Alternative Strategy for Desert Development and Management. *Proc. Int. Conf. Sacramento, California,* 1977, V.1-4, New York, e.a. I-X, 1982.

[2] Mabbutt, J.A. & Wilson, A.W., (eds). *Social and Environmental Aspects of Desertification*. International Geographical Union / The United Nations University, 40 p., 1980.

[3] UNCCD, *Desertification. An Overview* (A/Conf/74/1). Background Document, UN Conference on Desertification, Nairobi, Kenya, 1977.

[4] Thomas, D.S.G. Arid geomorphology. *Progr. Phys. Geogr*. **12(4)**, pp. 595-606, 1988.

[5] Fedorovich, B.A., *Dynamics and Statute of Relief-Forming in Deserts*. Nauka: Moscow, 236 p., 1983.

[6] Arnageldyev, A., *The Sands of Central Part of Karakum Desert, their Mobility and Struggle Against it*. Ylum: Ashgabat, 120 p. (in Russian), 1979.

[7] Bagnold, R.A. *The Physics of Blown Sands and Desert Dunes*. Methuen: London, 265 p., 1954.

[8] Kovda, V.A., The Aridisation of Dryland and Struggle against Drought. Nauka: Moscow, 272 p., 1977.

This page intentionally left blank

Section 10
Energy

Sustainability concern of housing: emergy storage and flow assessment

R. M. Pulselli, F. M. Pulselli, N. Marchettini & S. Bastianoni
Department of Chemical and Biosystems Sciences, University of Siena, Italy

Abstract

A city is an organized open structure made of assembled materials and buildings that constantly interface with changeable contextual agents such as clime, weather, solar irradiation and human beings. Urban systems feed on energy inflows in order to achieve an organization (e.g. society, economy, architecture) that is maintained in time. The interaction of different inputs from the environment generates the building as a built storage in which energy and materials have been stocked. Energy and materials inflows are required to maintain and to use the building in time (for instance, electricity, water and gas are needed for building use). These interaction processes between buildings and the external environment are the focus of this study.

Is it possible to measure these processes to evaluate sustainability of urban systems? How can the impact due to resource exploitation of housing on local sustainability be measured? Can we evaluate the environmental effects of urban strategies and structural plans?

An environmental accounting method, namely emergy analysis (spelled with an 'm'), was applied to an urban area considering the main activities of an entire human settlement and a detailed analysis was focussed on housing: the general environmental performances of buildings in terms of resource exploitation were evaluated considering their construction, maintenance and use.

As a case study, an emergy analysis of the municipality of Ravenna (north-eastern Italy) is presented with a special focus on housing and on the trend of growth of the building industry.
Keywords: emergy analysis, housing, buildings, neighbourhoods, urban systems, dissipative structures.

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060731

1 A thermodynamic approach to urban systems

An urban system can be conceived as an ecosystem in which there is a continuous interaction between a community of organisms (mainly humans) and the physical environment, whether it was natural or man-made, as once Eugene Odum [1] affirmed: "[an ecosystem] is a unit of biological organization made up of all of the organisms in a given area interacting with the physical environment". In urban regions, many interacting living agents coexist and a physical structure, made of lands, nature, buildings, infrastructures, technologies and other settings, is combined with a social community. This conjunction of non-living things with living agents let cities belong the category of ecosystems.

According to Francis Evans [2]: "in its fundamental aspects, an ecosystem involves the circulation, transformation, and accumulation of energy and matter through the medium of living things and their activities". Similarly, urban systems require inflows of energy and materials for self-maintenance.

Urban regions therefore 1) are extremely energy intensive; 2) require large inputs of energy and materials from the external environment 3) produce copious amounts of waste; 4) are human-dominated systems.

Eugene Odum [3] has further argued that an integral part of the ecosystem concept is a model of an open, thermodynamic non-equilibrium system, with the emphasis on the external environment. In urban regions, an organization (e.g. society, economy, architecture) is achieved and maintained over time. In other words, according to non-equilibrium thermodynamics, urban regions feed on different kinds of resources, whether locally available or imported from outside, that keep them in a steady state, that is a state far from thermodynamic equilibrium [4].

Ilya Prigogine introduced the concept of *dissipative structure* that can be considered to describe the general behaviour of human systems in urban regions, with their population, activities and settings. According to Prigogine, dissipative structures are defined as thermodynamic non equilibrium systems open to both energy and matter that self-organize towards high levels of complexity and organization (Prigogine and Stengers [5]). Therefore, *dissipative structures* are open living systems, far from thermodynamic equilibrium, able to self-maintain in a steady state (dynamicity, diversity, life) at high levels of organization; they constantly exchange energy and matter with the external environment, structuring themselves and evolving on the basis of these interactions.

An urban system, like a dissipative structure, absorbs high quality fluxes of energy and materials from the outside to self-organize; in terms of entropy, this means that it tends towards a state of minimum entropy (Tiezzi [6, 7]).

2 Introduction to the emergy analysis

Emergy analysis (spelled with an "m") is an environmental accounting method that develops an energy systems language for the thermodynamics of open systems (Odum and Odum [8]; Odum [9]). When applied to a building, it is processed to quantify all the environmental resources used for building manufacturing, maintenance and use.

For definition, emergy is the available solar energy previously used up, directly and indirectly, to make a service or product (Odum [9–11]). The emergy evaluation assigns a value to products and services by converting them into equivalents of one form of energy, the solar energy, that is used as the common denominator through which different types of resources, either energy or matter, can be measured and compared to each other. The unit for emergy is the solar emergy joule (*sej*).

The emergy of different products is assessed by multiplying mass quantities (kg) or energy quantities (Joule) by a transformation coefficient, namely transformity or specific emergy. Transformity is the solar emergy required, directly or indirectly, to make one Joule or kilogram of a product or service. Every time a process is evaluated, previously calculated transformities are used as a practical way of determining the emergy (sej) of commonly used products or services.

By definition, the solar emergy B_k of the flow k coming from a given process, for example housing, including the processes of building manufacturing, maintenance and use, is:

$$B_k = \Sigma_i \, Tr_i \, E_i \qquad i = 1,\ldots,n \tag{1}$$

where E_i is the actual energy content of the i-th independent input flow to the process, (e.g. materials, human work, solar irradiation, etc.) and Tr_i is the solar transformity of the i-th input flow.

In this paper an emergy analysis of buildings is presented with special reference to a published work (Pulselli *et al.* [12]) in which an emergy analysis was applied to a specific case study. Results are here expressed in a more general form and then applied to an entire urban region. The emergy of housing presented here refers to the municipality of Ravenna, in north-eastern Italy. An emergy analysis of the urban system of Ravenna as a whole was also published in Pulselli *et al.* [13] and used as a basic reference for comparing outcomes.

3 Emergy analysis of buildings

Referring to Pulselli *et al.* [12], an emergy analysis of housing is here developed in order to give a comprehensive evaluation of a traditional contemporary building. Outcomes refer to a traditional building block (usually in south Europe) with a reinforced concrete frame and brick walls. Results were processed in terms of emergy per unit of built volume (sej/m^3) as a sort of specific emergy of buildings.

In Figure 1 an *energy system diagram* of a building is shown with inflows of energy and materials. In the diagram: the building is shown as a built stock (symbol of storage) that, once manufactured, is maintained in time (ordinary maintenance). More in detail, the analysis was based on three distinguished processes: 1) building manufacturing; 2) building use; 3) building maintenance.

The interaction of different inputs, such as soil, water, energy, machinery, human work, materials, transport and other services (energy and materials flows)

generates the building as a built storage in which energy and materials have been stocked. Also a flow of energy and materials is used for the ordinary maintenance of the building in time. In the analysis this flow is assumed to be constant for 50 years that is the likely building lifetime. Building use is then shown by a rectangle. In this phase, inflows of energy for cooling and heating, electricity, gas and water are constantly needed.

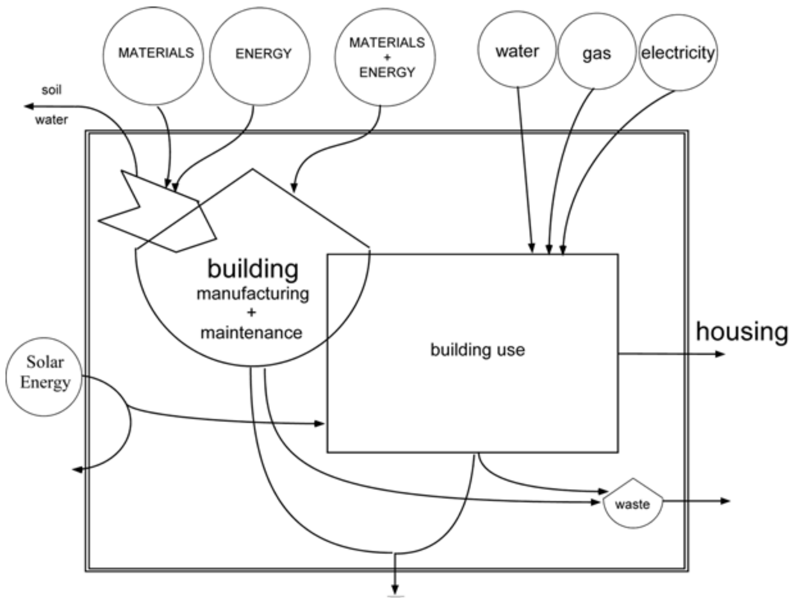


Figure 1: Energy system diagram of a building.

All the inputs to the process are then assessed by collecting and processing data as shown in Table 1. Every raw data (mass quantities) in the *building metric computation* was allocated to a unit of built volume (m^3) and expressed in terms of equivalent solar emergy (sej) through the transformity. Emergy flows refer to the quantity of materials used per m^3 of building and human work, soil erosion and solar irradiation for building manufacturing. Emergy of manufacturing per built m^3 represents the amount of environmental resources stoked in the building as a storage.

The emergy for building maintenance was assessed considering the quantity of energy and materials that are needed to maintain buildings, that is to maintain the emergy stocked in the built storage constant in time, contrasting the entropic degradation. Maintenance is therefore an annual inflow to the building. An emergy assessment of building maintenance was presented considering a 50 years average lifetime of buildings.

Results highlight the ‘environmental cost’ in terms of sej of building materials that is about 65% of the entire process of building manufacturing and maintenance (50 years). Transformities used are available in literature and were

accounted for each material considering the production processes from the origin to the provision (i.e. from the extraction of raw materials, to the final product). The emergy of building materials in a built m^3 are therefore expressed through emergy as a combination of used quantity (kg/m^3) and environmental cost (sej/kg). Results highlight the high impact due to concrete in contemporary architecture.

Table 1: Emergy of building manufacturing and maintenance: sej per built m^3.

RAW DATA	**Composition (kg/m^3)**	**Specific Emergy (sej/kg)**	**Emergy (sej)**	%
Concrete	306.81	1.79E+12	5.49E+14	32.48%
Brick and tiles	76.92	3.68E+12	2.83E+14	16.75%
Mortar	21.24	3.31E+12	7.03E+13	4.16%
Steel	7.72	6.97E+12	5.39E+13	3.19%
Stony materials	18.53	2.44E+12	4.52E+13	2.68%
Plaster	11.50	3.29E+12	3.79E+13	2.24%
Paint	1.14	2.55E+13	2.91E+13	1.72%
Copper	0.09	1.04E+14	9.20E+12	0.54%
Polystyrene	1.03	8.85E+12	9.07E+12	0.54%
PVC	0.58	9.86E+12	5.71E+12	0.34%
Aluminium	0.15	2.13E+13	3.20E+12	0.19%
Wood	0.49	2.40E+12	1.17E+12	0.07%
Glass	0.01	2.16E+12	1.69E+10	0.001%
Human work (in jouls)	1.76E+06	1.24E+07	2.18E+13	1.29%
Land use (in joules)	3.82E+07	1.24E+05	4.74E+12	0.28%
Solar radiation (in jouls)	1.50E+09	1.00E+00	1.50E+09	0.0001%
Emergy manufacturing per built m^3			**1.12E+15**	**66.45%**
Maintenance (year)		set of items	1.13E+13	
Maintenance (50 years)		set of items	**5.67E+14**	**33.55%**
Total Emergy stock per built m^3			**1.69E+15**	**100%**

In Table 2, consumptions of electricity, gas and water are presented per year (as an annual inflow) and during the entire building lifetime (assumed around 50 years). These emergy flows are due to building use.

Results were obtained for the processes of building manufacturing, maintenance and use; they show the following values:

- The emergy of building manufacturing is $1.12x10^{15}$ sej per m^3. It represents the investment in terms of natural capital to provide the building. This amount of emergy could be conceived as emergy stocked in the building. As shown in the energy system diagram buildings are

emergy storages. This emergy investment is made once in the entire building lifetime.

- Emergy of building maintenance is 1.13 x 10^{13} sej/yr per m^3. This flow is spent to maintain the building emergy stock. This is an annual emergy flow.
- The total emergy flow due to building use is 6.71x10^{12} sej/yr per m^3. This is an annual emergy flow.

Table 2: Emergy of building use: specific emergy per built m^3.

RESOURCE USE	Quant./m^3	Unit	Transformity (Sej/unit)	Emergy (Sej/yr)	Emergy (Sej) per 50 years
Electric Energy	3.09E+07	J/yr	2.07E+05	**6.40E+12**	3.20E+14
Gas Heating	8.22E+04	J/yr	6.72E+04	**5.52E+09**	2.76E+11
Water Supply	1.58E+02	kg/yr	1.95E+09	**3.09E+11**	1.54E+13
Total emergy for building use per built m^3				**6.71E+12**	3.35E+14

4 Emergy analysis of urban systems: the case of Ravenna

An emergy analysis was applied to the urban system of Ravenna as a whole and published in Pulselli *et al.* [13]. When applied to a region, many inputs to a multiplicity of processes are taken into account in order to give a measure of different activities and processes that take place in a local area. In the case of Ravenna, we assessed energy and materials inflows relative to physical aspects, energy consumption, materials use, fuel combustion, agriculture and industry, as shown in Table 3. Data and results refer to the year 2003.

The area under study is about 653 km^2 with a population of 141,800 persons (equivalent to a density of 217 persons/km^2). It is a highly industrialized area and relevant amounts of natural gas are also used for a local thermoelectricity production (electricity is also exported to the outside). This industrial vocation is thus highlighted by high emergy values of industry (1.44x10^{22} sej). Also a high value of emergy is due to natural gas (9.30x10^{20} sej) that is used for thermoelectricity production, besides industry and housing. For this reason, electricity consumption was not accounted in order to avoid a double counting. Solar irradiation and wind were accounted in the emergy of rain because they are co-products of solar energy.

The total used emergy of the municipality of Ravenna, in 2003, was 2.80x10^{22} sej.

5 Emergy analysis of housing

In the municipality of Ravenna, there were 78,745 houses in 2001 (distributed into 28,960 buildings for housing), each of 94.02 m^2 average, and 44.14 m^2 per person (ISTAT [14]).

In Table 4, the number of houses and relative square meters and volume, in the municipality of Ravenna, are reported with the equivalent amount of emergy for building manufacturing, use and maintenance. Values refer to the years 2001 and 2003.

Table 3: Emergy analysis of an urban system: the municipality of Ravenna, 2003.

Input	Quantity	Unit/yr	Specific Emergy (sej/unit)	Emergy (sej/yr)
PHYSICAL AGENTS				**9.11E+19**
1 Solar irradiation (already accounted as rain)	2.33E+18	J	1.00E+00	2.33E+18
2 Rain	4.15E+14	g	1.45E+05	6.01E+19
3 Wind (already accounted as rain)	1.24E+15	J	2.45E+03	3.03E+18
4 Geothermal heat	6.86E+14	J	1.20E+04	8.24E+18
5 Soil erosion	1.83E+14	J	1.24E+05	2.28E+19
WATER AND ENERGY				**9.62E+20**
6 Water use	1.61E+13	g	1.95E+06	3.13E+19
7 Natural gas	1.15E+16	J	8.11E+04	9.30E+20
8 Electricity use (already accounted as natural gas)	1.88E+15	J	2.05E+05	3.85E+20
MATERIALS				**1.64E+21**
9 Extracted materials (sand and and gravel)	9.78E+11	g	1.68E+09	1.64E+21
FUELS				**5.41E+21**
10 Gasoline and diesel	7.25E+15	J	1.11E+05	8.05E+20
11 Fuel oil and GPL	5.05E+16	J	9.12E+04	4.60E+21
AGRICULTURE AND ANIMALS				**5.22E+21**
12 Cereals	1.05E+16	J	2.67E+05	2.81E+21
13 Legumes	5.04E+14	J	1.75E+05	8.82E+19
14 Fruit	1.53E+14	J	4.82E+05	7.37E+19
15 Vegetables	3.38E+12	J	7.38E+05	2.49E+18
16 Seeds	1.36E+15	J	1.33E+06	1.81E+21
17 Spices	6.00E+12	J	1.75E+05	1.05E+18
18 Flowers	4.26E+10	g	4.74E+09	2.02E+20
19 Cuttle breeding	1.13E+13	J	5.33E+06	6.03E+19
20 Forestry	5.70E+08	g	1.68E+08	9.58E+16
21 Fishing and hunting	7.31E+11	g	2.27E+08	1.66E+20
EXTRACTIVE INDUSTRY				**6.60E+20**
22 Extractive industry (metal minerals)	1.31E+11	g	1.68E+09	2.19E+20
23 Extractive industry (non metal minerals)	2.62E+11	g	1.68E+09	4.41E+20
INDUSTRY				**1.40E+22**
24 Food industry	6.05E+11	g	2.52E+09	1.52E+21
25 Tobacco industry	1.06E+07	J	1.75E+05	1.85E+12
26 Leather industry	3.87E+12	J	1.44E+07	5.57E+19
27 Textile industry	9.89E+12	J	6.38E+06	6.31E+19
28 Furniture and clothing industry	7.86E+12	J	6.38E+06	5.01E+19
29 Wood and cork industry	5.87E+10	g	6.79E+08	3.98E+19
30 Paper industry	5.57E+15	J	3.61E+05	2.01E+21
31 Graphic industry	1.17E+14	J	3.61E+05	4.23E+19
32 Metallurgic industry	1.05E+12	g	5.81E+09	6.10E+21
33 Mechanical industry	1.27E+10	g	1.13E+10	1.43E+20
34 Mineral industry	1.71E+11	g	1.68E+09	2.86E+20
35 Chemical industry	5.76E+12	g	6.38E+08	3.67E+21
36 Rubber industry	1.88E+09	g	6.42E+09	1.20E+19
37 Other manufacturing industries	5.04E+09	g	5.81E+09	2.93E+19
TOTAL USED EMERGY				**2.80E+22**

Table 4: Emergy storage of housing in the municipality of Ravenna.

Up to year	Total houses	m^2	m^3	emergy building manufact.	emergy building use	emergy building mainten.
2001	78,745	7,403,605	22,210,815	2.49E+22	1.49E+20	2.52E+20
2003	80,204	7,557,288	22,671,865	2.55E+22	1.52E+20	2.57E+20

Table 5: Emergy assessment of new manufactured buildings: time series.

Year	houses	mq	mc	emergy building manufact.	emergy building use	emergy building mainten.
1991	420	68,192	204,576	2.30E+20	1.37E+18	2.32E+18
1992	513	75,327	225,981	2.54E+20	1.52E+18	2.56E+18
1993	422	71,398	214,195	2.41E+20	1.44E+18	2.43E+18
1994	458	70,719	212,157	2.38E+20	1.42E+18	2.41E+18
1995	378	37,906	113,718	1.28E+20	7.63E+17	1.29E+18
1996	308	44,615	133,844	1.50E+20	8.98E+17	1.52E+18
1997	740	90,503	271,516	3.05E+20	1.82E+18	3.08E+18
1998	552	61,073	183,218	2.06E+20	1.23E+18	2.08E+18
1999	873	108,389	325,166	3.65E+20	2.18E+18	3.69E+18
2000	654	89,222	267,668	3.01E+20	1.80E+18	3.04E+18
2001	550	72,431	217,293	2.44E+20	1.46E+18	2.46E+18
2002	592	66,835	200,505	2.25E+20	1.35E+18	2.27E+18
2003	**867**	**86,848**	**260,545**	**2.93E+20**	**1.75E+18**	**2.95E+18**
average per year	**564**	**72,574**	**217,722**	**2.45E+20**	**1.46E+18**	**2.47E+18**

The emergy amount for building manufacturing is an evaluation of the investment of environmental resources used for construction and thus maintained in buildings as in an emergy storage. The built environment is thus a storage of emergy equivalent to $2.55x10^{22}$ sej, up to 2003, that has been previously spent to provide the urban architecture.

The emergy amount for building maintenance ($2.57x10^{20}$ sej) is an annual emergy flow as well as the emergy for building use ($1.52x10^{20}$ sej) given by assessing water, natural gas and electricity consumption. These values refer to the annual emergy flow needed to supply and use the whole of buildings in the municipality of Ravenna.

In Table 5, a time series of the building industry is shown since 1991 to 2003. An average value (13 years database) is given in the last raw. The built environment grew with a rate of 217,722 m^3/yr that corresponds to an increase of equivalent $2.45x10^{20}$ sej/yr for building manufacturing; an emergy flow of $1.46x10^{18}$ sej/yr and $2.47x10^{18}$ sej/yr due to the new buildings has to be added to the annual cost for building use and maintenance respectively.

6 Conclusion

This paper presents new outcomes from the emergy assessment of building manufacturing, use and maintenance extending results from the case study of a specific building (as presented in Pulselli *et al.* [12]), more in general, to a traditional typology, a building block with a reinforced concrete frame and brick walls. Thus, emergy values for building construction, use and maintenance were allocated to a built cube metre as a sort of specific emergy of buildings in order to provide an emergy assessment of housing of an entire neighbourhood or urban area. An emergy evaluation of housing (and the building industry) in the municipality of Ravenna was achieved by comparing results with the emergy analysis of the entire territorial system (referring to: Pulselli *et al.* [13]), with its

physical agents, population, consumption, agriculture, industry and other activities. In particular, emergy of building construction, that refers to the resource use of the building industry, was compared to other human activities, in the local area in order to measure the environmental concern of housing.

In 2003, building manufacturing (2.93×10^{20} sej) was about 1% of the total used emergy (2.80×10^{22} sej) of Ravenna. The building industry (5.50×10^{20} sej) that involves the construction of new buildings (2.93×10^{20} sej) and the maintenance (2.57×10^{20} sej) of the existing ones is about 4% of the total emergy of industry.

Building use (1.52×10^{20} sej) - water, electricity and natural gas - was about 16% of the emergy for water, electricity and natural gas assessed for the entire urban region.

The entire building cubature in an urban area always increases due to new buildings. New buildings need resources for manufacturing (especially non renewable) and an increase of resource use of the local community for building maintenance and use (the rate of increasing of the building volume is about 1.2%/yr). Since the growth of built areas is unsustainable and an environmental policy of housing is strongly required, restoring and converting existing buildings is a good practice that needs minor investment of resources. The emergy analysis of housing can be a powerful tool for the evaluation of urban planning practices and making choices. For instance, it suggests restoring existing urban structures and dismissed built areas instead of planning new buildings and neighbourhoods; it measures material and energy saving through the restoration of existing buildings instead of manufacturing new ones. Furthermore, emergy analysis of housing can evaluate the energetic performances of buildings by assessing emergy for building use and an energy saving due to practices of eco-architecture can be assessed and measured for an entire neighbourhood (think, for example to the Bed Zed neighbourhood in London). These and other practices for decreasing the environmental impact of cities can be measured through the emergy analysis of housing.

Results presented above were probably expected to be higher but the area of Ravenna presents intensive industrial activities that have a high impact in terms of emergy use with respect to buildings. In the future, the emergy assessment of housing will be applied to other areas in order to compare new outcomes with the case study presented here.

References

[1] Odum E., The strategy of ecosystem development, *Science*, 164, 262-270, 1969.

[2] Evans, F.C., Ecosystem as the basic unit in ecology. *Science,* 23, 1127-1128, 1956.

[3] Odum, E., *Ecology and our endangered life-support systems*. Sinauer, Sunderland, Massachusetts, USA, 1989.

[4] Pulselli R.M., Ratti C., Tiezzi E., City out of chaos: social patterns and organization in urban systems, *International Journal of Ecodynamics*, 1(2), 125-134, 2006.

[5] Prigogine I. & Stengers I., *La Nouvelle Alliance*, Gallimard: Paris, 1979.
[6] Tiezzi E., *The Essence of Time*, WIT press: Southampton, 2003.
[7] Tiezzi E*., Steps Towards an Evolutionary Physics*, WIT press: Southampton, 2006.
[8] Odum H. T. & Odum E.C., *Energy basis for man and nature*. McGraw Hill: London, 1981
[9] Odum H. T., *Environmental accounting: emergy and environmental decision making*. Chichester Wiley: New York, 1996
[10] Odum H.T., *Environment, power and society*. Wiley, New York, US, 1971.
[11] Odum H.T., *Systems ecology*. Wiley, New York, US, 1983.
[12] Pulselli F.M., Pulselli R.M., Simoncini E., Environmental accounting of buildings: outcomes from the emergy analysis, (Eds.) Mander U., Brebbia C.A., Tiezzi E., *The Sustainable City IV, Urban Regeneration and Sustainability*. WIT Press, 489-498, 2006.
[13] Pulselli R.M., Magnoli G.C., Tiezzi E.B.P., Emergy flows and sustainable indicators: the Strategic Environmental Assessment for a Master Plan, (Eds.) Marchettini N., Brebbia C.A., Tiezzi E., Wadhwa L.C., *The Sustainable City III, Urban Regeneration and Sustainability*. WIT Press, 3-10, 2004.
[14] ISTAT 2001. Available on: www.istat.it

Energy and the environment

O. T. Inal, P. F. Gerity & D. D. H. López
New Mexico Institute of Mining and Technology, USA

Abstract

Energy is one of the major ingredients required for human sustenance and comfort. Yet, energy has an environmental impact at every stage of its acquisition as well as its use. We produce significant environmental waste and impact extracting energy, transporting the extracted material(s), and in processing it to render it useable. The global warming gases generated and the associated effects represent only a small sample of the serious by-products of our energy use, and only a fraction of the devastation that we presently bestow on our planet.

Presently, we are all aware that the price of one barrel of crude oil has exceeded the $75.00 level for the first time in history, and the negative impact on our individual economics as well as that affecting the industrialized world is such that every segment of our quality of life and living standard is threatened. These increased costs do not produce any direct benefit to the planet, but only contribute to the total global decline of living conditions and standards. The actual monetary benefit is only enjoyed by a very few as well.

Currently, the only fuel source alternative which seems to be considered politically acceptable involves the conversion of biomass (such as maize, sugar cane, other vegetable sources or use[s]) into bio-fuels to partially replace the shrinking supply of traditional sources of hydrocarbons. However, there is a finite and limited amount of arable land that can efficiently produce and successfully harvest biomass. All human needs for food and fiber must compete for access to this land area with any proposed use for energy production. The competition among food, fiber and fuel is already very stressed, and the use of bio-mass is not a viable, long-term solution to the problem.

This paper identifies fuel sources, the quantities currently available, and projections for the future. The areas of solar, wind, hydrogen, and renewed interest in nuclear power are all considered. It is our effort to put these alternative options into a more realistic light.

Keywords: energy, biomass, biodiversity, pollution, fusion, fission.

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060741

1 Introduction

For many, the state of the world is getting worse based on four major fears which they harbor. These four fears include:

1.1 Natural resources are running out

Their quantities are not of such an extent that their production could cope with the needs of a more industrialized and expanding population. Since the next million or so years will only net an accumulation of about one six hundredths of the amount built up thus far, it cannot be relied upon to replenish itself.

Although the amounts of these resources are limited, the limits are far greater than many project. The "main three" sources and the most recent assumed quantities are [1]:

1. Coal Reserves: Enormous; greater than 2000 years' supply

2. Natural Gas: (in trillion cubic feet)

Total	Middle East	Russia	North America	US Reserves (1999)	US Reserves (2000)
6000	2500	1600	269	167	189

US Annual Usage: 18-20, i.e. less than ten year domestic year supply, but reserve growth is outpacing depletion at the moment.

3. Producible petroleum reserves (Billions of barrels)

Conventional:		
	World	1,300
	Middle East	700
	North America	40
Tar Sands:		
	Canada alone	175
Oil Shale:		
	Recoverable today	200
Annual Usage:		
	USA	7
	World	30

There is greater than a 50 year supply at current usage rate. (It also puts into question as to why we have a sudden increase in per barrel crude oil prices).

The projected recoverable reserve today seems to be:

Estimate	Quads	Years Supply
Fossil Fuels	200,000	600
Nuclear	430,000	1400
Geothermal Fluids	160,000	500

i.e. no urgent cessation of energy delivery from any of these supplies at the present time. The environmental cost though is large.

1.2 Population is ever growing, leaving less and less to eat

Agricultural production in the developing world has increased by 50% per person since 1961 (as of 2001). The food intake in poor countries has shown the following increases: 1,932 (1961), 2,650 (1998) and expected to be 3,620 (2030) Calories; i.e. not a lessening but a steady increase in nourishment.

In contrast, starvation in developing countries was 45% in 1949 has reduced to 18% in 2001 and is expected to be 12% in 2010 [2]. Also, since year 1800 food prices have decreased by more than 90% and in year 2000 prices were lower than ever before [3]. It turns out, as people grow richer and healthier, they tend to have smaller families. While the population growth rate mentioned was greater than 2% per year in early 1960's, it became 1.26% in year 2001 and the expected the rise will be 0.46% in year 2050. UN estimates show that population growth will be over in year 2600 stabilizing at just 11 billion. It turns out that human ingenuity has boosted food production not merely in line with, but ahead of, population growth with a concomitant reduced pressure on land acquisition and thus biodiversity.

1.3 Biodiversity loss

With forced increase in arable land, and thus decrease in forested area, biodiversity is expected to suffer the greatest loss. This doesn't seem to be the case when one considers that the Eastern United States, that had its forested areas reduced to 1-2% of their original size, in the last two centuries, has resulted in the extinction of one forest bird. All but 12% of forested area being cleared in Brazil has lead to no verifiable extinction of any species. Further, UN reports that tropical forest losses in the last century is less than 0.5% rather than the 2-4% as specified by environmentalists.

1.4 The Planet's air and water are becoming ever more polluted

This is another area that is exaggerated. Although there is pollution related to every stage of energy use, this is far below what some project. Human endeavor is actually very small in comparison to natural activity on the planet and could only be the "last straw that broke the camel's back", as the saying goes. A volcano produces more SO_x and CO_2 than hundreds of industrial facilities. Air pollution is increasing in developing countries; they are merely duplicating the development route of the industrial ones. When they get to be economically stable they too will address industrial pollution.

1.5 Why a disjunction between perception and reality?

It is clear that a disjunction exits between what some are claiming and reality is. This possibly comes about, due to the following:

1.5.1 Research in any given area initiates with the definition of the existence of a problem in that area

This unfortunately leads to the impression that many more problems exist than is the case.

1.5.2 Activists need to be noticed

To achieve this interest they tend to exaggerate considerably (one extremist environmental scientist wanted to destroy, with a nuclear bomb, humanity (and bring it back to the stone age) so that he, without polluting instruments of modern life, could live without sin [4]. A good example in this regard is the estimation of the worldwide fund (1977) that two thirds of world's forests were extinct; the truth is this number is nearer 20%. Also there is the inherent suspicion of people; a trade organization is either altruistic – when it goes for stronger pollution controls – and is instantly seen as self-interested if it argues for weaker controls.

1.5.3 People are clearly more curious about bad news than good

Although 1997-1998 El Niño phenomena was blamed for damages amounting to $4 billion, the $19 billion benefits were generally not as widely reported.

1.5.4 There is also the wrong perception about the environment

After the infamous rubbish loaded barge made its travel through the Caribbean bringing refuse to unload at every port, unsuccessfully returning to its original port, New York, people began talking about the amount of waste created in the US. The actual case is that America would produce, through the entire 21st century, enough waste that will take up an area 1/12th of the country.

In this regard, ignorance matters only when it leads to faulty judgments; for example, cost of Kyoto, for US alone, will be higher ($1 trillion) than five times the money needed to cure a major problem for the planet: universal access to clean drinking water and sanitation. In this regard, it may be costly to be overly optimistic but more costly still to be too pessimistic.

2 Energy options for the future

As Sheik Yamani, Saudi Arabia's former oil minister and a founding member of OPEC, is said to have pointed out: "The stone age came to an end not for a lack of stones, and the oil age will end, but not for a lack of oil" [5]. Humanity stopped using stones because we found superior materials (iron and bronze) and likewise we probably will stop using oil when other energy technologies provide superior benefits. What we have said to this point is that it is not that we are going to run out of energy but that we are going to have it at a higher cost. The rate of our consumption of energy related materials has depended on human ingenuity in that when cereal consumption increased twofold since 1960, meat consumption increased by threefold since 1960, fish catch raised its take to six times between 1950 to 1997 this was balanced by the fact that half of the commercial fertilizer ever produced was applied to farming since 1984 [6]. Life on earth exists only because of the natural greenhouse effect, the ability of the atmosphere to retain enough heat for species to thrive (and no more).There is a finite amount of carbon on the planet: stored in fossil fuels, the sea, living matter and the atmosphere. Without human influence, transfers between these storage areas roughly balance each other; for example: plants take up carbon as they

grow but release it as they decay. Humans cut down trees and burn fossil fuels and thus release extra carbon into the atmosphere increasing the greenhouse effect. The greenhouse gases include carbon dioxide, emitted by fossil fuel burning and deforestation, and methane, released from rice paddies, landfill sites and animal waste. Warming that comes about, from the greenhouse effect, will trigger some processes which speed further warming and other effects which mitigate it (decreased ice cover produce exposed land which leads to more absorption of heat and speeds warming further; in contrast, plants' CO_2 intake is likely to increase as higher temperatures increase growth rates). A danger of increasing warming is to either alter or stop natural occurrences on the planet that equilibrate climate changes; for example, the "Great Ocean Conveyor", surface and deep sea movements, in the form of Gulf stream uses the wind driven surface currents and carry the milder tropic weather to Northern Europe and thus make life quite bearable [7].

The, so-called, "infinite" sources of power – sources that are continuously replenished – fall into two groups:

2.1 Group A

2.1.1 Hydrogen

Abundant and clean but must be processed for use as a fuel. The product of its use is water vapour. There is currently no economic way to dissociate water and thus we still have to develop an efficient means of doing this and to store and distribute it. Its use in transportation vehicles will have to be preceded by careful engineering of its containment due to its highly explosive characteristic.

2.1.2 Wind

Wind power can be obtained by careful design and with knowledge of the stream flow pattern(s) of air. Large scale use is very impractical. It is a good source for exposed areas, is environmentally friendly and is a relatively endless source. Noise of its operation and the obtrusiveness of its appearance are its major distractions. Few seem to understand the metal fatigue and mechanical failure which occurs. The likelihood of a windmill having extended life, and thus be cost effective, is, therefore, rather small.

2.1.3 Water

Safe and pollution free source of energy that is limited by location and by the upheaval it can cause since it requires big sites for collection. Water power is assumed to have a capacity of three trillion watts, which is equivalent to present total energy use in industry. Only 8.5% of this power though is utilized at present because of the lack of industrial development in regions such as South America, Africa, and Southeast Asia which are of highest potential for hydroelectric power.

The efficiency of use of these "infinite" sources is low and they have restrictions of locations, seasons and large capital investment is required [8].

2.2 Group B alternatives

2.2.1 Solar energy

Solar energy is derived from the fusion reactions that take place in the core of the sun. The main obstacle to date in efficiently harnessing this power has been finding ways or collectors to economically concentrate the available low energy density of the solar radiation. Also, the intermittent nature of solar energy requires an efficient storage capability in utilizing this source.

Collector designs are based on a thermal process in which the temperature of a working fluid, typically water or alcohol, is increased by absorbed solar radiation. Techniques utilized in collection of this energy include:

2.2.3 Concentrating solar collectors

A parabolic mirror system concentrates solar irradiation onto a pipe with a selective absorber coating placed at the focal point containing the heat transfer medium. Temperatures in excess of 500ºC have been achieved with this design but high installation costs have limited their use at present. To attain highest efficiencies, a tracking system that follows the sun is needed. The collector tube assembly includes a one and a quarter inch stainless steel tubing, 12 feet long, that has a coat of conductor sulfamate nickel – and a double coat of absorber metal – black chrome or black nickel – to gain absorbtivity efficiencies above 95%. In the actual design of this tubing the arrangement of the above layers are: stainless tube, nickel coating, black chrome or black nickel layer followed by a vacuum jacket and a glass cover to reflect back all reflections from the absorber surface. As can be imagined, this is extremely costly together with being an extremely bulky assembly.

2.2.4 Flat plate collectors

A cheaper version consisting of a black coating over a flat surface that is attached to copper tubing carrying the exchange fluid – usually water. The absorbed thermal energy by the black surface heats the exchange medium that flows in copper pipes to the storage facility. The idea in this process is to supplement the regular heating source with solar energy. This design is cheaper and has been incorporated into the construction of many houses already. According to a study done at Los Alamos National Laboratories, this design could produce up to 75% of the total energy needs of a household in a geographic region like New Mexico. Therefore, the main problem with this approach is the fact that, to achieve the power consumption of U.S. in 1970, the area of collection for solar irradiation would be approximately 24,500 square kilometers [9].

It is evident at this point that except for solar energy, all the other forms mentioned are not plentiful enough to supply the future demand of energy but can and should all be utilized as limited contributions to the total picture. More efficient collection and storage of solar energy can lead to an unlimited source of clean energy for future.

3 Nuclear energy: fusion and fission

To understand the basic principles involved in Fission and Fusion, nuclei and the arrangement of neutrons and protons within these have to be studied. In a nuclear plant the energy is released from the nuclei of the atoms while energy derived from fossil fuel burning (coal, oil, or gas) comes from the atoms or molecules composed of atoms.

The binding energy for nucleons is much greater than the binding energy of atoms; for example helium atom with two protons has a binding energy per atom of 19 eV (electron volts) while helium nucleus with two protons and two neutrons (four nucleons) has a binding energy of 7.0 MeV (million electron volts) per nucleon.

Energy per nucleon goes through a minimum for elements with 50 to 130 nucleons rising sharply for lighter nuclei and gradually toward the heavy nucleus end. This indicates that it is possible to decrease internal energy and thereby release the excess energy by fusing lighter nuclei together and move towards the middle of this curve or splitting an unstable heavy nuclei into two intermediate nuclei and again release large quantities of energy. The former is the principle behind fusion and the latter is utilized in fission reactions. Although it might seem that more energy release occurs in fusion, this is energy per nucleon and the number of nucleons participating in fission is much more. Actual energy available from fusion is only about ten percent of the energy derivable from fission [10].

Nuclear breeding is achieved with the neutrons released by nuclear fission. The fissioning of each atom of a nuclear fuel, such as ^{235}U, liberates an average of more than two fast (high energy) neutrons. One of the neutrons must trigger another fission to maintain the nuclear chain reaction and there is at least one neutron available to breed new fissionable atoms, that is, to transform "fertile" isotopes of heavy elements into fissionable isotopes. This is illustrated as follows [11-13].

Uranium Cycle:

$$2.\ {}^{238}_{92}U_{146} + \mathrm{n} \rightarrow {}^{239}_{92}U_{147} \rightarrow {}^{239}_{93}Np_{146} \rightarrow {}^{239}_{94}Pu_{145}$$

and, Thorium cycle:

$$3.\ {}^{232}_{290}Th + \mathrm{n} \rightarrow {}^{233}_{290}Th \rightarrow {}^{232}_{291}Pa \rightarrow {}^{233}_{92}U$$

The two man-made isotopes ^{239}Pu and ^{233}Pu are fissionable like ^{235}U and this way the energy source material is increased tremendously.

The reactor core assembly, consists of fuel elements (clad in thin high melting point, high thermal conductivity metal containers), control rods and the cooling medium surrounding the assembly.

When cooling is done with water, dual purpose is served. Water removes the heat of fission from the containers and also acts as a moderator. To slow down the neutron of 1 MeV energy to less than the 1 eV is accomplished by some 20

or so collisions before it is absorbed by a ^{235}U and induce fission. Control rods used in the core area regulate the neutron flow. The heat carried away by the cooling media is converted to electricity through a generator.

4 Environmental considerations

To summarize, with a power system based on neutrons the following environmental advantages are attainable:

- With the fusion and/or fission reactors no burning of the world's oxygen or hydrocarbon resources and hence no release of carbon dioxide or other combustion products are brought about for power generation.

- Fusion creates no wastes and the high energy neutrons can be utilized to burn radioactive wastes of the fission reactors.

- Fusion reactors are also inherently incapable of a "run away" accident, there is never enough fuel present at any one time to support a nuclear excursion.

These concepts are only concepts and years of research based on the most advanced technologies are required to make these reactors operate with efficiency and be failsafe. The increasing demand for energy and decreasing amount of fossil fuels have today made the topic of nuclear reactors very prominent, since this is the only energy form in magnitude and long range usability that is at hand.

5 Fuelish?

A logical conclusion of this study would be to question the merits of a spiraling increase of our energy needs. Since there are no "pure" forms of energy with no dangers associated or no pollution resulting there from, shouldn't the goal be to curb this energy hunger and go for a zero-based population growth rather than finding new ways of satisfying this increasing need? The prospects we are faced with if the present trend continues are as follows.

5.1 If the source is coal

- It would require 33 tons of coal/person/year; which we might be able to pay for but the environmental cost might be too large.

- Would result in 20 trillion tons of CO_2 emitted into the earth's atmosphere which could involve enough atmospheric changes that melting of the polar caps would become reality. The excavations for coal would leave the countryside an environmental mess and thousands of vehicles required to transport the coal could congest traffic in every major city.

5.2 If the source is nuclear

- Fissionable U-235 is of very limited quantity and would only last about fifty years.
- Breeder reactor that would have to take the slack-breeding ^{238}U and ^{232}Th into ^{234}Pu and ^{233}U – are the worst offenders in thermal pollution. The product elements are also the major components of nuclear weapons.
- 11,000 megacuries of long-lived radioactive isotopes that would be produced each year would have to be transported between the reactor site and the processing plants – where they are made into bricks – and ultimately stored at a disposal site.

This leads to only one conclusion and that is to control energy consumption and population explosion together since there are no "absolutely clean" forms of energy production.

References

[1] Oil and Gas Journal, Jan 1, 2005.
[2] B. Lomborg, The Economist, Aug 4, (2001), 63-65.
[3] World Bank, 2001.
[4] Dixy Lee Ray with L. Guzzo, "Trashing the Planet", Regnery Gateway, Washington D.C. 1987.
[5] Greider, William 2000 "Oil on Political Waters". Nation, 10/23/2000 271 (12):5-6.
[6] Alex Kirby, http://newsuote.bc.co.uk/mpapps/pagetools/print/news.bbc/co.us/z/.
[7] "Guide to Climate Change", http://news.bbc.co.uk/z/shared/spl/hi/sci-nat/04/climate_change/html/climate.stm
[8] M. King Hubert. Scientific American, 224, #3, Sept(1971) 60.
[9] M. King Hubert. Scientific American, 224, #3, Sept(1971) 62.
[10] D.R. Inglis, Nuclear Energy, Its Physics and its Social Challenge, Addison Wesley, Pub. Co. Reading, Mass (1973).
[11] G.I. Scaborg and J.L. Bloom, Scientific American, Vol. 223, No.5, No. (1970) 13.
[12] The Complete Ecology Handbook; Philip Nobile and John Dealy, Eds., Doubleday and Co. Inc, NY (1972).
[13] W.H. Jordan, Physics Today, 23, #5, May (1970) 32.

This page intentionally left blank

Renewable resources of energy in northern Baja California, Mexico

M. Quintero-Núñez[1], A. Sweedler[2] & S. Tanaka[2]
[1]*Instituto de Ingeniería, Universidad Autónoma de Baja California, Baja California, México*
[2]*Center for Energy Studies, San Diego State University, San Diego, CA, USA*

Abstract

In Mexico, the Comisión Federal de Electricidad (CFE) is a government entity created to generate and distribute energy in Mexico. CFE is operating a 720 MW capacity geothermal power plant at Cerro Prieto, located in the valley of Mexicali, Baja California, Mexico. Wind energy in Baja California is not now being exploited. It is potentially productive in some areas of the Rumorosa mountain range and at the Cañon de San Martin in the Valle de la Trinidad. Wind pattern measurements have been carried out and investors' interest is growing for wind farms of different sizes (1500 MW, 250 MW, 50 MW). Solar radiation in Baja California is quite significant. Solar photovoltaic (PV) electricity is widely used for lighting, communications, and appliances throughout the state of Baja California in rural areas and small towns not connected to an electric grid. In order to exploit the full potential of renewable energy resources in this border region with California, EUA, a series of barriers and drivers are addressed. While significant technical potential for renewable energy exists in Baja California, additional study is required to fully characterize the resources and refine estimates.
Keywords: renewable energy, Baja California-California border region, Mexico, geothermic, wind energy.

1 Introduction

The Secretariat of Energy [1] encouraged the National Energy Savings Commission (CONAE) to promote the development of renewable energies in

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060751

Mexico to reduce its dependence on hydrocarbon fuels. In 1966 CONAE, along with the National Association for Solar Energy (ANES), organized a forum to discuss the potential steps to promote renewable energy sources. The result was the creation of the Advisory Council for the Promotion of Renewable Energies (COFER). This group is made up of representatives from industry, commerce, academia, government, and development banks. The aim of COFER is to promote the use of renewable sources of energy in Mexico within a market framework. It also serves as an advisory group to identify projects and for the design and development of programs and policy related to renewable energy, including small hydro, solar, biomass, and geothermal. SENER (1998) estimates that by 2008 close to 559 MW will be installed in such systems, producing approximately 1,836 GWh/y.

2 Characteristics of Baja California

The state of Baja California (Fig. 1) is located in the northwest part of the peninsula of the same name with approximately 500 km of length from the international line with California, USA (33° latitude North) down to parallel 28° in the south. It has a surface of 70, 113 km^2 which represent 3.75% of the total surface of the country. It is formed by mountains, hillsides, central valleys and coasts, besides a very discontinuous litoral on the Pacific Ocean.

The state of Baja California is characterised by the existence of two main meteorological regimes that divide the climate: a warm and dry during the spring, summer and the beginning of fall, and the other in winter with relatively humid and slightly cold [2].

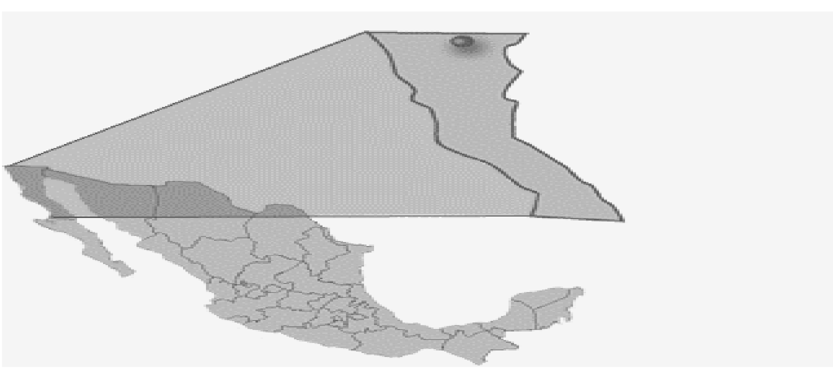


Figure 1: The state of Baja California is located in the northwest of Mexico.

3 Legal framework

Each renewable energy source requires an adequate legal and regulatory environment that favors its development. Mexico's existing legal framework allows power generation projects using renewable energy under self-supply,

small production, independent construction, and export schemes. The Public Service for Electric Energy law does not constrain power generation to a specific technology. Even though environmental costs are not expressly considered when pricing the Mexican electric market, there are a few legal provisions that promote the use of renewable energy.

4 Geothermal technical potential

The Federal Electricity Commission (CFE) is a federal government entity created to generate and distribute electric energy in Mexico (Constitucion Politica de los Estados Unidos Mexicanos, 1988) [3]. In 1982 CFE created the Geothermal Electrical Projects division whose specific function is to coordinate all geothermal-related activities in Mexico. The following is a brief summary of geothermal activity in the valley of Mexicali, currently under development by the Mexican government for the production of electricity**.**

Cerro Prieto (Mexicali, Baja California) is the most important site being developed by the Mexican government. It is located in the Mexicali Valley (115.16 longitude west, 32.25 latitude north) between the Pacific and American tectonic plates and near the San Andreas Fault. The plain is a delta, and the geological area is composed of unconsolidated clays, sand, and gravel that rest on sedimentary rocks of sandstone, lutites, and limonites.

Currently the site has a capacity of 720 MW distributed over four plants, Cerro Prieto I-IV (180, 220, 220, and 100 MW respectively). According to Alonso 1988 [4], the Cerro Prieto geothermal reservoir has an estimated capacity of 1,200 MW and a proven capacity of 840 MW. Since 720 MW of the 840 MW of proven capacity have been developed, the technical potential for new capacity is 120 MW.

5 Energy sources in Mexico

To date, geothermal energy, along with other alternative sources such as solar, wind, marine, and biomass, have contributed marginally to the energy balance in Mexico. In the National Energy Balance of Mexico for 2004 [5], geothermal energy represents only 0.6% and wind energy is not significant.

5.1 Wind energy

The Center for Higher Education and Research of Ensenada (CICESE) carried out a project titled, "Wind Energy in Potentially Productive Areas in Baja California." This study was based on available data for the Rumorosa mountain range in relation to meteorology, climatology, geography, and topography.

The objective of the study was to determine wind patterns for a one year period, possible locations for the establishment of wind farms, recommendations for the construction of sustainable and profitable wind farms.

The evaluation criteria to determine the potential production areas included the following:

* Average annual wind magnitude (minimum of 5.5 to 6 m/s)
* Proximity of energy transmission lines
* Proximity of centers of energy consumption

Daily wind registries (velocity and direction) were obtained from seven meteorological stations within the study area. Monthly and annual averages were projected after processing the data. The result was that high intensity winds are produced during the day, coinciding with the time of greater demand of electricity in urban areas. Table 2 shows a list of the potential areas where this resource can be developed. The study area is shown if Figure 2.

Table 1: Wind energy locations in Baja California, Mexico.

Station	Ave. Wind Speed (knots)	Ave. Wind Speed (M/s)	Standard Deviation	Energy (watts/m^2)
Pino Suárez	20.2	10.9	4.5	1299.6
Jacumé	15.5	8.3	3.7	581.6
La Rumorosa	14.9	8.0	4.0	516.4
El Hongo	12.0	6.5	2.6	274.0
El Pinal	11.7	6.3	2.9	254.7
La Puerta	11.5	6.2	2.5	238.2
El Centinela	17.2	9.3	4.9	793.7

Figure 2: The wind study area in Baja California, Mexico [6].

To determine the location of a wind farm one must measure the daily and seasonal wind velocity at various locations. Wind velocity is also affected by the terrain and elevation. For this reason it is important to install one or several automatic meteorological stations that collect data such as humidity, solar radiation, and other climatic elements that could be relevant to each site.

Based on those studies Mexican specialists have concluded:

- Baja California has the potential to produce electric energy using wind power.
- Jacume and Pino Suarez are the areas with the highest potential for energy production.
- Specific studies for these areas must be carried out.

Several wind projects are being proposed in Baja California:

A) Short-term: A pilot project in the area described, with major potential that possesses the following characteristics:

- Four 1,500 kW wind turbines
- Annual production of 15,768 MWh/y
- Plant factor of 0.3
- Two hectares of land
- An approximate investment of U.S. $6 million

B) Medium-term: Development of a wind farm of 50 MW capacity to supply several state government agencies:

- 341,500 kW generators
- Annual production of 134,000,000 kWh/y
- Plant factor of 0.3
- 200 hectares of land
- An approximate investment of U.S. $ 50 million

The Baja California 2000 project was proposed by Fuerza Eolica S.A. de C.V. in partnership with Enron Wind Corporation (U.S.), with an estimated cost of U.S. $170 million [7]. The project aimed at producing and supplying non-polluting, reliable electric energy with a peak production capacity of 120 MW in the town known as La Rumorosa. This project would allow the five state municipalities of Baja California (Tijuana, Mexicali, Ensenada, Tecate, and Rosarito) a savings equivalent to 15 percent of the electricity costs for public lighting in its first twenty years of operation. The project planned to have an initial capacity of 60 MW, made up of two 30 MW capacity modules and two future modules of the same capacity for a combined capacity of 120 MW. The wind generators would produce in excess of 300 million kWh per year.

In 2003 Spanish investors visited Baja California to analyze the feasibility of constructing a 250 MW wind farm at the Cañon de San Martin in the Valle de la Trinidad [8]. This project could potentially generate 4,000 MW with an investment of U.S. $250 million.

The rugged topography of the La Rumorosa area, with several canyons and many ravines, dictates extensive and highly site specific wind surveying in order to assess the overall wind power potential. Given the natural secrecy and reluctance of any commercial developer active in the area to share data, further resource assessment of the wind resource will require public funding and coordination with Mexican government agencies.

5.2 Solar energy

Solar photovoltaic electricity is widely used for lighting, communications, and appliances such as refrigerators throughout the peninsula in rural areas and small towns not connected to an electric grid. Some fishing cooperatives have also installed solar-based and hybrid solar-wind systems in isolated fishing camps. CFE, in collaboration with the Instituto de Investigaciones Electricas (IIE), has collected information on the maintenance requirements and long-term availability of PV systems. In the long term, the CFE-IIE collaboration may be expanded to develop several hundred MW of solar electricity nationwide within the context of a distributed generation project. A shorter term project will include the development of a grid-connected 1-MW photovoltaic array at a Mexicali substation. There is also a great potential for using solar photovoltaics for water pumping, which is carried out in a very small scale in the valley of Mexicali.

To take advantage of the excellent insulation in the state of Baja California, CFE studied the technical and economic feasibility of integrating a solar steam system to a conventional gas-fired combined cycle generating plant. A field of parabolic trough solar thermal collectors would be used to produce the steam. The concept, known as the Integrated Solar Combined Cycle System (ISCCS), was incorporated into the tender requirements issued by CFE with a 25 solar component at the Rosarito III generation plant, at Rosarito, Baja California. It is scheduled to enter in service by April of 2011.

The National Meteorological Service of Mexico (SMN) has a net of monitoring stations measuring and recording solar irradiance. Data is averaged over 10 minute intervals. Six stations in Baja California have been operational since 2000:

- Emilio Lopez Zamora dam (Ensenada)
- Abelardo L. Rodríguez dam (Tijuana)
- Mexicali
- San Quintín
- Bahia de los Angeles
- Gustavo Diaz Ordaz

The data in Figure 3 was collected from all stations except the Gustavo Diaz Ordaz station for 2003 [9]. The Bahia de Los Angeles site shows the highest solar radiation in the region, followed by Mexicali.

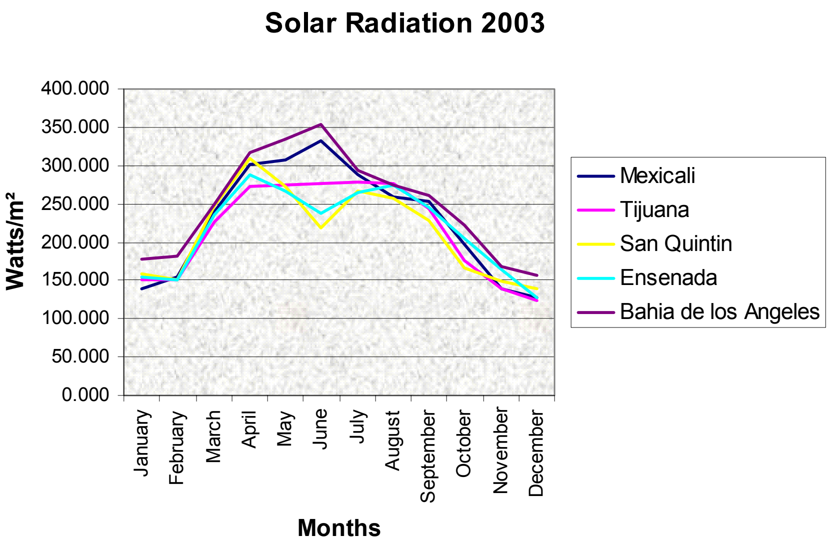


Figure 3: Solar radiation from several monitoring stations located in Baja California for 2003.

5.3 Micro-hydroelectrics in the Valley of Mexicali

Twenty years ago national studies were carried out [10] to explore the possibility of installing micro-hydroelectric plants in Mexico. In this analysis, the valley of Mexicali was assigned 7.5 MW Technical Potential utilizing the irrigation canal system in Mexicali.

A later study [11] was conducted to evaluate the supply of electricity to the Autonomous University of Baja California's (UABC) academic and administrative buildings located in the city of Mexicali using micro-hydroelectric units installed in selected locations on the valley's irrigation canals. The study showed that once established the system could provide the 5 MW demand required by UABC.

The two projects have a Technical Potential of 12.5 MW.

6 Barriers and drivers

In order to exploit the full potential of renewable energy resources in Baja California, a series of barriers and drivers will need to be addressed

6.1 Cross border transmission capacity

At the current time, only about 800 MW of transmission capacity exists between Baja California, Mexico and California, USA. Most of this capacity is already utilized. In order to take advantage of new renewable energy generation in Mexico, increased cross border transmission capacity will be needed.

6.2 Renewable Energy Credits

California utilities, including San Diego Gas and Electric (SDG&E), are required to meet a portion of their demand from approved renewable energy sources. They receive a Renewable Energy Credit (REC) that counts towards meeting this requirement. At the present time, renewable energy sources in Mexico do not count towards the renewable energy credit. This is a significant barrier for companies to develop renewable projects in Mexico and a major barrier for SDG&E to buy energy from renewable sources originating in Baja California.

6.2.1 Production tax credit

The production tax credit (PTC) is very important for wind projects to be cost effective. Currently, the PTC can only be applied to US-based projects. This is clearly a major barrier to the development of wind power in Baja California, even though the resultant energy would be sold to California. This requires a change in federal legislation.

6.3 Binational energy planning

Although energy facilities located on one side of the border could have environmental impacts on the other side, there is no mechanism to incorporate these cross border effects. Some mechanism is needed to take into account cross border impacts from energy production and distribution, even if the energy resource is a renewable one. Some headway has been made in this area via the San Diego Association of Governments' (SANDAG), the Border Energy Issues Group and the Border Governor's Energy Working Group, but more needs to be done.

Despite these barriers, there are definite advantages to develop enhanced cross border energy capacity to better utilize the renewable potential in Baja California. Chief among these is the large renewable potential that exists in northern Baja California in the form of solar, wind and geothermal. In addition to the renewable resources themselves, land may be less expensive than in California and there could be certain tax advantages to locating a project in Mexico.

The above discussion focuses primarily on cross border transfer of renewably generated electricity from Baja California to California. From an energy, environment and economic perspective, northern Baja California is a vital part of the greater San Diego Region. As indicated earlier in this chapter, the need for electricity in Baja California is growing rapidly and renewable technologies have a role to play to meet this demand. However, before this potential can be developed, certain barriers need to be addressed.

6.4 CFE and renewables

CFE, as the only entity allowed to generate electricity in Mexico for general distribution, should consider increasing the portion of renewable energy in its mix of resources for Baja California. The main issue for CFE – as it is on the US side of the border – is the relatively high capital cost of installation for solar and wind technologies. However, in Baja California, where the primary fuel for power generation is natural gas that must be imported, renewable technologies may be more cost competitive than in other regions of Mexico. It may also be possible for CFE to develop a renewable portfolio standard, similar to what exists in California.

7 Full renewable potential for Baja California

This paper has not addressed the full potential for renewable energy development in Baja California. For example, we have not analyzed the potential contribution of solar hot water to replace electricity, or the possible contribution of rooftop PV in commercial and residential buildings. Nor have we done an analysis of the regions appropriate for central station solar generation. There are potential wind sites that need further analysis. These issues we hope to address in a future study, but clearly from our present study, there is considerable potential for significant renewable energy development in Baja California.

The main drivers for renewable energy development in Baja California are:

- The high cost on imported natural gas.
- The abundant solar and wind potential of the region.
- The resulting reduction of air pollution by increasing the amount of renewable energy utilized.
- The potential creation of new industries.

References

[1] Advisory Council for the Promotion of Renewable Energies (COFER). Sener. www.sener.gob.mx.
[2] Alvarez, M. “Climatología de la Sierra de San Pedro Martir”. Desert Fishes Council 15º Annual Simposio, November. 1983
[3] Constitución Politica de los Estados Unidos Mexicanos, ed. Porrua, México, D.F. 2005
[4] Alonso, E. H., Cerro Prieto :Una alternativa en el desarrollo energético, Memoria de la Reunión Nacional sobre la Energia y el Confort, eds. Garcia F.C., Instituto de Ingenieria, UABC, Mexicali, B.C., pp. 314-319. 1988.
[5] Scott, A., Bialek, T., Geier, D., Houston, J.D., Quintero N.M., Resley, R., Rohy D.A., Swedler, A., Tanaka, S., Gin C. & Zeng, K., Potential for renewable energy in the San Diego Region, ed. San Diego Regional Renewable Energy Group, August 2005, Sand Diego, CA, USA.

[6] Energy Information Systems (SIE). Sener, www.sener.gob.mx.
[7] Gottfried Carlos. "Baja California 2000, Energía eólica para el Estado de Baja California", Fuerza Eólica S.A de C.V.7 Noviembre. Mexicali, Baja California. 1999
[8] Rivero, M. "*Promueven Planta eolica*". La Crónica, Mexicali Baja California, 16 de junio, pp 2/F. 2002
[9] SMN. Base de datos sobre radiación solar del 2003. Sistema Meteorológico Nacional. México. 2003.
[10] Ferran, R. F. "Estimación del potencial hidroenergético de los distritos de riego del país", Instituto de Investigaciones Eléctricas (IIE). Informe IIE/42/3988/101/F. Marzo. Palmira, Morelos, México. 1986
[11] Quintero, N.M López, R.M., Microhydroelectric plants in the valley of Mexicali (Chapter 15) *Energy and Environment in the California Baja-California Border Region*, ed. A. Sweedler, Paul Ganster, and P. Bennett, IRSC, SDSU, pp. 129-132, 1995.

Waste to energy as a contribution to ravage elimination

P. Stehlík & M. Pavlas
Institute of Process and Environmental Engineering,
Brno University of Technology, UPEI – VUT, Brno, Czech Republic

Abstract

The fact that mankind has been facing serious environmental and global problems during recent years is generally known and it is not necessary to emphasize it. However, every concrete way providing us with a method leading to elimination of the ravage of the planet represents a valuable contribution. This paper is aimed at renewable energy in terms of utilizing waste as alternative fuel. The benefit is quite obvious – waste is disposed and its energy content utilized. This approach is called waste to energy

Currently, an increasing amount of waste produced (both municipal solid waste - MSW and industrial and/or hazardous waste - IHW) can be considered as a driving force for finding efficient ways for the thermal processing when waste is removed and at the same time used as a valuable fuel. It represents a challenge for potential investors and operators as well as researchers.

It is necessary to take into account the following needs:

- waste disposal as a necessity
- waste-to-energy as the main aim
- limitations given by environmental legislation
- locality – big cities or villages

An idea about building new regional Waste-to-Energy Centers (WTEC) is outlined. It is based on a convenient combination of experience, know-how and sophisticated approach. WTEC is equipped with:

- a unit for the thermal treatment of hazardous industrial wastes
- a unit for the thermal treatment of common MSW
- a special incinerator for the thermal treatment of sewage sludge if more waste water treatment plants are situated in the region
- a system for sorting and separation of wastes connected with a production of alternative fuels
- a system for utilizing ash and flying ash for secondary building materials

WTEC has to be based on up to date technology which enables a further extension and innovation initiated usually by increasing population and standard of living as well as more and more sweeping environmental limits.

Keywords: waste, energy, sustainable development, emissions reduction.

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060761

1 Introduction

Serious problems are currently solved in the field of environmental protection. The challenge facing concerned citizens and decision-makers is a formidable one: To identify and implement long-term solutions that are safe, socially acceptable, and cost-effective. Such amounts of waste, which are produced either by inhabitants (municipal solid waste – MSW, see Figure 1, [1]) or by industrial companies (industrial and hazardous waste - IHW) require one to use efficient ways of waste disposal. The recent focus on incineration has been on environmental consequences, not on performance.

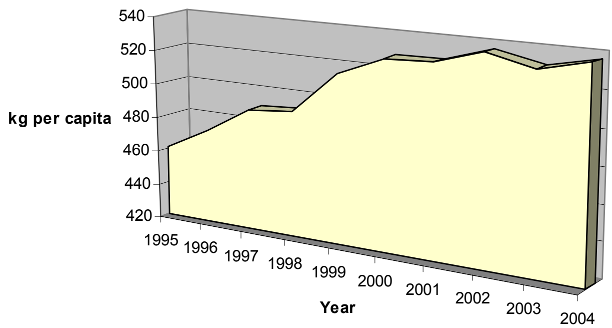


Figure 1: Municipal waste collected per capita (kg per capita) in the European union [1].

One of other global problems of the human race is the greenhouse effect. Principles and methods of reaching the given targets in this field are well known and do not need further attention:

- reducing process energy consumption
- increasing the effectiveness of energy utilization stored in fossil fuels
- strong use of renewable *energy sources*.

The preferred sources of renewable energy are, apart from biomass, also thermal treatment of waste with heat recovery. It can be said, that waste stops to be a problem and becomes an available fuel, which brings two advantages – waste is treated and at the same time energy is produced. Then we speak about waste-to-energy technology (*WTE*). WTE is also referred to as thermal processing of wastes including energy utilization [2–4].

2 Primary Energy Saving by WTE

Due to the more and more sweeping environmental limits on the quality of side products of waste combustion (flue gas, solid residues and waste water) WTE systems can provide us with clean and reliable energy in the form of heat as well as power (Q_{exp}, Figure 2). In a number of countries energy from WTE systems is considered to be renewable energy.

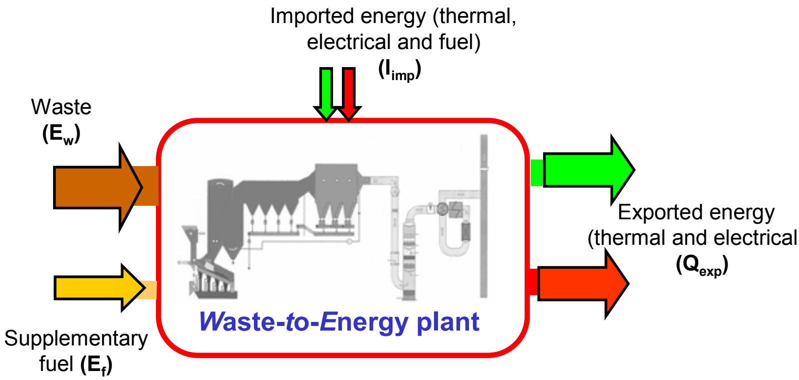


Figure 2: Thermal treatment of waste as a process.

The controlled combustion process is not energy self-sufficient, i.e. it cannot be processed without the consumption of supplementary energy (E_f, I_{imp}). The consumption (E_f, I_{imp}) depends on the type of waste incinerated, its calorific content and the thermal regime used. For example for incineration plants it is typical that the lower heating value of waste is sufficient to meet operational requirements without the need of supply of auxiliary energy to the combustion chamber (e.g. combustion of natural gas). On the other hand for IHW incinerators and units, which are not optimized or are not operated efficiently, the specific consumption of energy can be much higher, thus bringing negative economic and environmental effects.

2.1 Criteria for evaluation

To compare the effectiveness of energy production in different incineration plants several criteria have been developed by several organizations [5,6]. As an example the criterion called *Plant Efficiency* defined in the BREF [6] can be used. It defines the ratio between energy exported by incinerating the waste and energy consumed by the process itself.

2.2 WTE and cogeneration

Thermal treatment of waste is accompanied by the release of a considerable amount of heat, which has to be, if we talk about modern WTE systems, efficiently utilized. The most frequent method of utilizing heat content of flue gas is production of steam in waste heat boiler (heat recovery steam generator, HRSG). The produced steam can be directly consumed on-site but a much more efficient way of utilizing it is production of power and subsequent utilization of the remaining energy for heat supply. In that case a WTE system is also a cogeneration plant (see Figure 2).

3 Factors influencing technology selection and plant performance

For a selection of waste processing technology various criteria have to be respected. First, in most cases it is necessary to "hand-tailor" the technology from case to case. For a customer there is usually a choice of best available technologies (BAT, [6]). However, only few of them (sometimes none of them) are applicable for the given purpose. Then a creativity and all-round potential of the technology supplier (to be selected) plays a key role. There is not a general prescript/instruction how to proceed, however, some general and specific criteria have to be taken into account. It is necessary to always bear in mind the main aim of the incinerator, which is a safe and environmentally friendly disposal of a given amount of waste. At the same time suppliers of 21st century's incinerators have to respect maximum utilization of energy from the waste incineration

An optimum choice of a waste processing technology is subjected not only economic requirements but especially it is limited by environmental constraints. Then it is necessary to combine all the following factors and make an acceptable trade-off:

- Waste disposal as a necessity
- Waste-to-energy as the main aim
- Limitations given by environmental legislation
- Locality – big cities or villages.

Before looking at the specific factors in more detail it is important to mention a typical unit for municipal solid waste (MSW) incineration (Figure 3, Figure 4). Waste incineration is performed in a combustion chamber equipped by a moving grate under temperatures ranging between 850 and 1000°C. The heat released in this process is utilised in a heat recovery steam generator (HRSG), where it is cooled down to approximately 250 to 280°C. Mechanical cleaning of the flue gas is performed in electrostatic precipitator (ESP). Then it enters the block of off-gas cleaning, comprising a dioxin filter (DEDIOX) and a wet scrubber.

3.1 Waste disposal as a necessity

To successfully solve a waste treatment process we have to be familiar with the waste character, composition, state (whether solid, liquid, gas and/or a mixture) etc. It is necessary to emphasize important influence of the thermal regime in the combustion chamber on the auxiliary energy consumption [7]. It is based mainly on the consumption of supplementary fuel used to achieve and maintain the conditions for complete oxidation of all combustibles (with an emphasis on the so-called "3T" Time, Temperature, Turbulence). It is generally known that the minimal temperature of gaseous products after the last air input is given by legislation and it differs according to the type of processed waste [8].

A typical example of waste with a very low heating value, the combustion of which is linked with a higher consumption of supplementary fuel, is sludge from

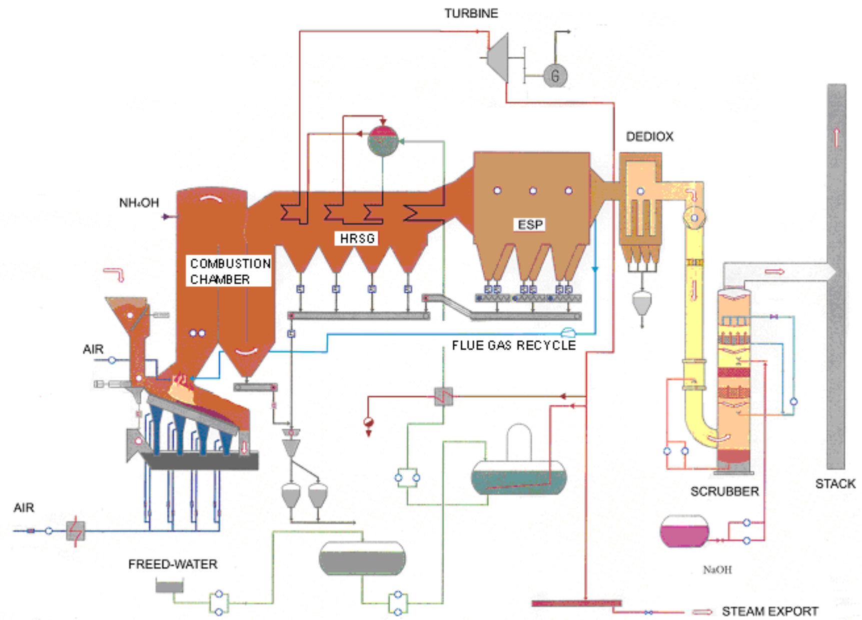


Figure 3: Typical configuration of up-to-date municipal solid waste incinerator (courtesy of TERMIZO Liberec Ltd.).

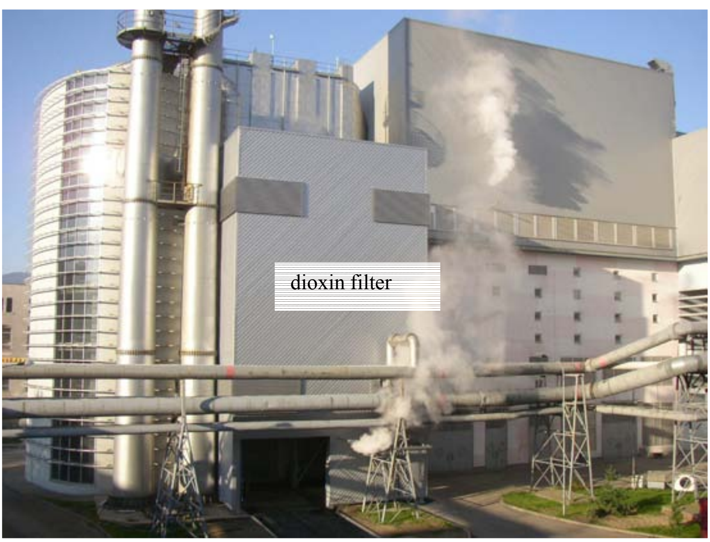


Figure 4: Up-to-date MSW incinerator with capacity of 300 t/day and new dioxin filter (courtesy of EVECO Brno Ltd.).

industrial processes and wastewater treatment plants. Processing of this type of waste requires specific technologies that are “hand-tailored” from case to case [9,10]. The following example shows thermal treatment of industrial sludge in unit with a multiple hearth combustion chamber with fluidized bed (see Figure 5:). An interesting feature of this system consists in substituting natural gas by mining gas (from closed coal mines), which means using alternative fuel. It required development of a specific original burner [11].

Figure 5: Unit for thermal treatment of sludge from pulp production (courtesy of EVECO Brno Ltd).

3.2 Waste-to-energy as the main aim

Heat recovery in units for the thermal processing of various types of wastes can be without any doubts considered as one of the most important parts of these processes. Design of equipment for utilization of energy contained in the flue gas and their placement in the process is one of key factors in these technologies. At the same time it is necessary to tackle the question of integration of a WTE system, in the sense of finding a suitable consumer of heat released by the combustion process. The main result of an analysis of the heat recovery system lay-out and influence of the main parameters were presented in [12].

The highest efficiency expressed in saving of primary energy sources is reached when backpressure turbine is applied. In this case all the steam going through the turbine and used for power generation is then utilised for heating purposes, which results in higher efficiency. However the disadvantage of this

system is the direct dependence of the incinerator on the grid conditions, to which steam is distributed. Suitable consumers of heat for this application are for example industrial processes with steady heat consumption. On the other hand these processes usually require higher parameters of steam to supply heat to the cold process streams. Of course this has a negative effect on plant performance, since less power is produced, which is considered to be a more valuable type of energy than heat, as a result of reduction of the heat drop over the turbine.

On the other hand steam with relatively low parameters can be effectively utilized for hot water production in systems of district heating in the field of municipal energy supply. In that case it is necessary ensure that the incinerator heat output can be utilized all year round. This could be quit serious problem, if WTE system is integrated into a relatively small district heating system in an area with a mild climate, where the heat consumption in summer months is very low. A very flexible solution for such unsteady heat consumption is application o bleeding condensing turbine, where excess steam is utilized for power production. However the consequences of so-called summer operation mode with a reduced heat output and increased power output on plant performance is evident - the efficiency of the cycle as well as primary energy saving are considerably lower.

3.3 Limitations given by environmental legislation

The system of flue gas cleaning is another important part of units for thermal processing of waste. Off gases polluted by NO_x, SO_2, HCl, HF, fly ash and dioxins and furans (PCDD/F) flow through a number of equipment with the aim of reducing the concentration of harmful compounds to the level below emission limits. In connection with more and more sweeping emission limits, the requirements on each equipment unit, which the off-gas cleaning system consist of, are increasing. At the same time, its integration into the technology is just as important.

As an example of possible technology extension in the future it can be shown that a baghouse filter with a special material for removal not only solid particles but also PCDD/F (dioxins) will be necessary in the nearest future and MSW incinerators have to be adapted accordingly to enable an easy way of potential retrofit. Dioxins (collective name for polychlorinated - dibenzo-p-dioxins and dibenzofurans) belong among the most dangerous pollutants, elimination of which from the products of incineration (solid residues, flue gas) is a necessity strictly required by legislation. For example in the EU, the present limit on dioxin emissions is 0.1 TEQng/m_N^3 (whereas e.g. in China the current limit is 1 TEQng/m_N^3). It is therefore very attractive to employ technologies that actually destroy dioxins instead of only collecting them. The latest achievement in this respect is the catalytic filtration technology REMEDIA™ of the company W.L. Gore & Associates Inc., which combines fabric filtration (collection of particulate matter) with catalytic destruction of dioxins. A recent information on the performance of this technology applied in a municipal waste incinerator may be found e.g. in [13].

3.4 Locality

Last but not least factor, which influence technology selection is the locality. In agglomerations and large cities with high density of population and/or industrial areas there is possible to solve the problem by a standard way, in small towns and villages it is necessary to find a specific way.

4 Waste to Energy Centre

Thermal processing of various types of wastes is not only waste disposal and/or treatment at present. It has to be regarded as waste to energy process wherever and whenever possible and feasible. Therefore an idea about founding and building new regional Waste-to-Energy Centers (WTEC) is outlined. It is based on a convenient combination of experience, know-how and sophisticated approach.

The main fact influencing the concept of WTEC is the necessity to treat different types of wastes. Typical types of wastes as well as main products of the processes which proceed in the WTEC are shown in Figure 6.

In case of establishing such a WTEC the following principle advantages are of primary importance:

- substantial contribution to environmental protection
- saving natural raw material and energy resources
- increasing employment
- source of regions' income.

Of course general common principles have to be respected like e.g. the following ones:

- sorting of waste
- waste recycling and utilizing all known processes for the treatment of biodegradable wastes
- permanent developing the environmental legislative framework
- necessity of people's education (in this field) starting from early youth, and under support of media and public administration.

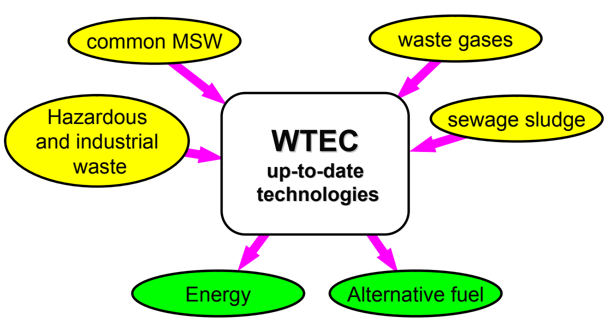


Figure 6: Types of waste processed in WTEC.

Then it is necessary to evaluate the situation in terms “what is available”. WTEC should be equipped by basic items such as:

- unit for the thermal treatment of hazardous industrial wastes
- unit for the thermal treatment of common MSW
- special incinerator for the thermal treatment of sewage sludge if more waste water treatment plants are situated in the region
- system for sorting and separation of wastes connected with a production of alternative fuels
- system for utilizing ash and flying ash for secondary building materials
- cogeneration (Combined Heat and Power System – CHP)

To put it into practice a view from both user and supplier has to be taken into account, i.e. “what do we need” and “what we are able to offer”, respectively. Last but not least it is necessary to adapt the WTEC system according to local conditions, region’s character, agricultural or industrial area etc.

5 Conclusions

Up-to-date incinerator plants cannot be considered only as units for waste disposal, but they have to be looked at as modern energy sources producing renewable energy, The potential for primary energy saving can be seen as the difference in generated and consumed energy. An optimum choice of a waste processing technology is influenced by many factors. Some of them are mentioned in more details in this paper. Last but not least ideas about potential organizing WTEC are outlined provided latest achievements would be utilized.

Acknowledgements

We gratefully acknowledge financial support of the Ministry of Education, Youth and Sports of the Czech Republic within the framework of research plan No. MSM 0021630502 "Waste and Biomass Utilization focused on Environment Protection and Energy Generation" as well as support from the Czech Science Foundation within project No. 101/05/2469.

References

[1] Eurostat, *Generation and collection of municipal waste*, available on http://epp.eurostat.ec.europa.eu, 2006.

[2] Bebar L., Martinak P., Hajek J., Stehlik P., Hajny Z. & Oral J., Waste to Energy in the Field of Thermal Processing of Waste, *Applied Thermal Engineering,* **22**, pp. 897-906, 2002.

[3] Stehlik P., Progress and Challenges in Thermal Processing of Various Types of Waste, *Proc. of the 1st International Conference & Exhibition on Thermal Treatment and Resource Utilization of Wastes IWTR 2005*, Proceedings on CD ROM, Beijing, 2005. (Key-note Lecture.)

[4] Stasta P., Boran J., Bebar L., Stehlik P. & Oral J., Thermal Processing of Sewage Sludge, *Applied Thermal Engineering,* **26**, pp. 1420 – 1426, 2006

[5] CEWEP, *Discussion paper on criteria for energy recovery in waste incineration plants*, Brussels, available on http://www.cewep.com/energy/energy_recovery, 2005.

[6] European IPPC Bureau: *Reference Document on the Best Available Techniques for Waste Incineration*, Brussels, available on http://eippcb.jrc.es, 2005.

[7] Pavlas M., Bébar L. & Stehlík P. Waste-to-energy systems and their benefit. *Proc. of the 1st International Conference & Exhibition on Thermal Treatment and Resource Utilization of Wastes IWTR 2005*, Proceedings on CD ROM, Beijing, 2005.

[8] Council Directive 2000/76/EC of the European parliament and of the council of 4 December 2000 on the incineration of waste, *Official Journal of the European Communities*, 28. December 2000.

[9] Oral J., Stehlik P., Sikula J, Puchyr R., Hajny Z. & Martinak P., Energy Utilization from Industrial Sludge Processing, *Energy*, **30**, pp. 1343 – 1352, 2005.

[10] Boráň J., Houdková L., Ucekaj V., Šťasta P. & Stehlík P., The Analysis of Energy Utilization in Processes for Sewage Sludge Treatment, *Applied Thermal Engineering*, 2005, (accepted for publication).

[11] Boran J., Houdkova L. & Stehlik P.. Waste as alternative fuel. *Proc. of the 1st International Conference & Exhibition on Thermal Treatment and Resource Utilization of Wastes IWTR 2005*, Proceedings on CD ROM, Beijing, 2005.

[12] Pavlas M., Bébar L., Urban L. & Stehlík P. Analysis of utilizing energy from thermal processing of waste, *Proc. of the 9th Conference on Proces Integration, Modelling and Optimisation for Energy Saving and Pollution Reduction PRES 2006,* Proceedings on CD ROM, Prague, 2006.

[13] Bebar L., Puchyr R., Pranghofer G., Parizek T., Bernat P. & Stehlik P., Optimum involvement of dioxin catalytic filter in a MSW incinerator. *Proc. of the Filtech 2005*, Wiesbaden, 2005.

Expansion of the Brazilian refining industry and its local requirements: critical factors for siting a new refinery in Rio de Janeiro State, Brazil

A. Magrini, A. Szklo, G. Machado & R. Schaeffer
Energy Planning Program, COPPE,
Federal University of Rio de Janeiro, Brazil

Abstract

The trade-off between welfare and environmental quality is clear in the refining industry. Economic development and well-being are still related to oil product demands in contemporary society. However, the refining industry is usually aggressive to the environment, bringing externalities out to localities where facilities are built. Therefore, there is a dilemma between meeting national oil product demands growth and preventing environmental damages in the locations eligible for a green field project. This paper shows that Brazil should expand its refining industry over the next twenty years by two to four new refineries. One of these refineries, focused on petrochemicals, might be located close to major consumption centers. In this case, Rio de Janeiro State is an interesting option. The analysis indicates the trade-offs between the positive and negative external effects of a new refinery in the Itaguai Municipality. The advantages of good logistics and infrastructure might be offset by the cumulative environmental problems related to air, water and solid wastes management in the region.
Keywords: refinery expansion, location, environmental pressure, Brazil.

1 Introduction

The trade-off between welfare and environmental quality is particularly clear in the refining industry. Economic development and well-being are still closely related to oil product demands in contemporary society, despite relevant efforts to substitute and/or to save oil products after the Oil Shocks in the 70's.

www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060771

However, the refining industry is usually very aggressive to the environment, bringing externalities out to localities where facilities are built. In this sense, there is a dilemma between meeting national oil product demands growth and preventing, or at least controlling, environmental damages in the locations eligible for a green field project (local communities might say “gray” field project). As a matter of fact, environmental requirements have become one of the major issues in refinery green field projects. For instance, new refinery capacity will probably not be installed in the USA in the next 30 years, mainly due to environmental requirements for green field projects (oil products supply increase will come mostly from imports and conversion capacity) – [1]. In the case of Brazil, evidences brought out by [2–7] suggest that Brazil might reach a bottleneck in its refining capacity for supplying the oil products market in the coming years, particularly for key fuels such as diesel, which has presented high levels of dependence during the past two decades.

Brazil's refining industry includes thirteen refineries with the operational processing capacity of 1.96 million barrels/calendar day [8]. The main refineries are located alongside or close to the consumer centers at the southeast region of the country, except for the RLAM refinery, which is located in Bahia State where crude oil production outputs paraffinic crude oil (KUOP of 12.6), allowing the production of lube oils. There are seven refineries in the Southeast, whose output meets some 50% of the demand for oil products in Brazil [8]. The total capacity of these refineries was equivalent to 62% of the operating capacity of Brazil's refining segment in 2003.

Brazil is a net importer of diesel. In 2003, 30% of total oil product imports consisted of diesel oil [8]. Although its thirteen refineries have been revamped in order to handle the average lower quality processed crude, while the oil products market has shifted towards medium cuts, they are nearing production capacity. Short and medium-term investments planned for conversion and treatment units at existing refineries might not be enough to meet the diesel market growth. In this case, total refinery capacity expansion might be required. However, while meeting oil products demand is a national need, identifying the location where the facility will be built requires the definition of who will hold the local environmental burdens. Refineries result in economic benefits for the regions where they are installed, but put environmental pressures, particularly due to intensive water use and atmospheric emissions. Developing countries such as Brazil must also deal with complex logistics and infrastructure difficulties when taking decisions on the best location for a new refinery.

2 Brazil’s refinery expansion through 2025

Oil products demand and supply forecasts are based on two studies: one performed to the Ministry of Mines and Energy [5]; and another jointly concluded with the International Atomic Energy Agency [9]. These studies used the following tools:

- The Model for Analyses of Energy Demands (MAED) developed by the International Atomic Energy Agency (IAEA) - for more details see [9].

- For the transportation sector, the forecast of the light vehicle fleet by type followed a scrapping curve. For the industrial sector, the Integrated Energy Planning Model (IEPM) was used – for more details see [10] and [11]. For the residential sector, logistics curves were sized to forecast the penetration and use of stoves in households, by income and location (rural and urban areas).

In addition, two technical and economic scenarios were drawn up – see Table 1:

1. Business as Usual Market Scenario. This scenario does not include qualitative changes in the development of Brazil's production segment other than those already delineated over the past few years. Shifts in the Brazilian economy – such as the incorporation of technical progress and alterations to the production structure – take place at a pace compatible with that of the past decade.
2. Alternative Scenario. This scenario includes qualitative changes to Brazil's production development, through upgrading exports and “dematerializing” the economy from 2010 onwards.

Table 1: Real GDP growth rate and sector added value assumptions for the Brazilian economy (% p.a.).

	2005/10	2010/15	2015/20	2020/23
GDP	4.26	4.11	4.05	4.05
Business as Usual Market Scenario				
Agriculture	3.58	3.55	3.38	3.26
Industry	5.42	5.16	5.03	4.98
Services	3.60	3.45	3.42	3.46
Alternative Scenario				
Agriculture	4.19	4.13	4.88	4.93
Industry	4.00	3.00	1.86	1.80
Services	6.00	6.00	3.00	3.00

In terms of energy prices, the basic assumption was the convergence of oil and oil products prices in the Brazilian and the international markets. The scenarios took the oil, oil products and natural gas price forecasts drawn up by the US Department of Energy as the basis for the international prices [1]. Technical assumptions were made for potential inter-energy substitutions and crude oil production curves [12]. In terms of oil output quality, increasingly heavy acid oil production is expected. This affects the technological options for the development of Brazil's refining industry. Three options were assessed:

- Expansion of Brazil’s current refineries in terms of processing capacity and conversion and treatment capacities. Modifications currently under way in Brazil's existing refineries are designed to boost the processing of domestic heavy crudes while reducing the supply of heavy oil products (fuel oil) and fostering higher and better supplies of medium and light oil products (diesel, jet fuel, and LPG). Investments are channeled to bottom-of-barrel units, such

as delayed coking, vacuum residue catalytic cracking, and hydrocracking, and to hydrotreatment. These projects alter the oil products output profile of the existing refineries (see Table 2), but do not represent a significant expansion in the atmospheric distillation capacity as a whole.

- A new refinery optimized for diesel (see Table 3), processing mainly the Brazilian Marlim crude oil (API 19, sulfur content of 0.7% and TAN of 1.1), which is the main stream produced in Brazil [13]. The key bottom-of-barrel units of this refinery are delayed coking and hydrocracking. Due to its versatility, this refinery can also output higher yields of light cuts, and is flexible to step up its production of heavy oil products as well, if necessary.
- A new refinery focused on propylene (see Table 3), which is produced in a DCC unit, LPG and distillates (mainly diesel). This petrochemical integration is in line with technological innovation of the oil refining industry worldwide [14] and [15]. This venture would solve the diesel dependence issue, also guaranteeing petrochemical feedstock supplies.

Table 2: Estimated production profiles for Brazil's current refining segment (%).

Oil Product	2001	2010	2015-2020
LPG	8.1	8.9	9.1
Gasoline	19.5	21.9	23.8
Naphtha	10.5	11.3	9.6
Kerosene	2.6	5.4	5.9
Diesel	36.4	40.7	41.6
Residuals	22.9	11.9	10.1

Table 3: Production profiles for new refining schemes.

Production Focus	Diesel	Fuel-Propene
Capacity (10^3 barrel per day)	250	200
Yield (%)		
LPG	9.5	2.4
Gasoline	34.1	22.0
Naphtha	5.3	3.4
Propene	-	16.4
Kerosene	0.0	-
Diesel	44.6	21.0
Heavy oil products	11.6	14.0

Finally, the expansion refining criteria was based on reducing the vulnerability of Brazil's oil chain. In this case, a ceiling of 10% dependency was stressed for the key oil product on Brazil's fuel market: diesel. This figure corresponds to the average ceilings adopted in previous studies – [3, 5, 7].

Therefore, among the refinery options described above the modeling strives to select those whose versatility or industrial integration potential would justify investments in Brazil. The findings of the simulations indicated that the expansion of Brazil's refining sector between 2002 and 2025 should take place through four refineries in the Business as Usual Scenario, and two refineries in the Alternative Scenario – see Table 4.

Table 4: Forecast of entry schedules by refinery profile in Brazil.

	Business As Usual Scenario		Alternative Scenario	
Refinery type	Diesel	Petro-chemicals	Diesel	Petro-chemicals
2010	0	1	0	1
2015	1	0	1	0
2020	1	0	0	0
2025	1	0	0	0

3 Critical factors for sitting a new refinery in Brazil: Rio de Janeiro State Case Study

Both scenarios indicate that Brazil should expand its refining segment by two to four refineries over the next twenty years. Implying an additional 500 to 1,000 thousand barrels per day of refining capacity, this is fully compatible with our forecast increase in Brazil's crude oil production output (equal to 1,730 kbpd between 2002 and 2025) – see [12].

However, the need for expanding refining industry in the country level raises the question of finding a proper location for the new facilities. At local-level, fuel and petrochemicals market and environmental requirements become major issues. As pointed out in the first section, the oil products market and the refinery capacity are concentrated in the Southeast region of Brazil. This region accounts for around 60% of Brazil's GDP and Manufacturing Value Added [16]. It represents around 50% of the country's oil products market [8], and presents fuel surpluses that are channeled to other Brazilian regions (Mid-West and Northeast, basically) and to other countries (Brazil's fuel surpluses include, especially, gasoline and high-grade petcoke). Therefore, the analysis of the regional location of new refineries in Brazil usually suggests the Northeast region for a fuel-focused refinery and the Southeast region for a petrochemical-focused refinery. Actually, one of the new refineries simulated in both scenarios is integrated with petrochemicals (coming on-stream in 2010). In this particular case, the Southeast region is the most suitable location, concentrating the country's petrochemicals market and an increasing demand for propene [17]. It is also heavily dependent on naphtha from domestic refinery production and from imports.

At the Southeast region, Rio de Janeiro State holds 92.0% of Brazil's proven offshore reserves [8], is located close to major oil product consumption centers, particularly for petrochemicals, and has an advantageous location for both oil

products imports and exports. In addition, the Campos Basin at Rio de Janeiro holds around one-third of Brazil's total probable reserves (probability of 50%), and Rio de Janeiro accounts for 82.6% of Brazil's total production.

Therefore, Rio de Janeiro State is a good macro-location candidate for a new refinery focused on petrochemicals. However, this industrial complex imposes pressures on the environment and the infrastructure in the site where it is located. This refinery is complex (see Table 5). It requires 600 MJ per barrel of feed, and 5,000-13,000 m^3 per hour of water in open cycle. The project implementation schedule time is 4 years.

Table 5: Main refinery processing units.

Atmospheric Distillation Unit – ADU (1,000 bpd)	200.0
Vacuum Distillation Unit (%)	65.2
Delayed coking (%)	31.6
DCC unit (%)	65.6
Hydrocracking (%)	10.0

Note: downstream units' capacity is related to the ADU capacity.

Due to its available basic infrastructure and proximity to industrial market, Itaguai Municipality is frequently eligible for new industrial facilities in Rio de Janeiro State. This region comprehends utilities, roads and ports, and it is up to 500 km close to Sao Paulo, Minas Gerais and Rio de Janeiro industrial markets (see Figure 1). Still, it already has an area originally earmarked for the Rio de Janeiro Petrochemical Complex owned by PetroRio. This is the area that was assessed in the present study. In addition, Itaguai offers fiscal exemption from land duties, work levies and occupancy permits for ten years, to companies employing more than thirty people. Finally, the assessed site presents low-density urban areas (less than 0.1 inhabitant per km^2), despites the vicinity to the town of Itaguai.

Figure 1: Sepetiba Bay location in Rio de Janeiro and Brazil.

The following site factors were assessed in Table 6: energy, water, road, rail, port and pipeline infrastructure, labor-force, land, conservation units, air quality, water quality, solid wastes.

Table 6: Local Site Factors for Itaguai.

Site Factor	Facility Requirements (demand)	Site Availability (supply)	Comments
Energy	The refinery is self-sufficient. The CHP plant of 200 MW generates surplus power.	The site is also served by natural gas-fired thermo-power plants.(1)	The refinery does not impose a local pressure.
Water	Water requirement hovers between 5,000 and 13,000 m^3/h (1.17 – 3.60 m^3/s).	Gross water availability is 78.68 m^3/s. A saline wedge penetrates 5 km up the Guandu River that serves Itaguai and reduces water availability to 28.68 m^3/s.	There is water availability. However, there is also a problem of a saline wedge penetration in Guandu River.
Transportation Infrastructure	The location of the refinery should be as close as possible to ports, roads, railroads.	There are two ports: the Port of Sepetiba (83 millions of tones) and the TEBIG Terminal owned by Petrobras in Angra dos Reis. The assessed area is 2.5 km away from highways; the Mangaratiba rail spur runs through this area.	The region is well provided by road and rail infrastructure. The Port of Sepetiba has a large spare capacity, but a bulk liquids terminal needs to be built.
Labor-Force	The new Refinery will demand roughly 30,000 workers under temporary contracts (refinery construction), while some 1,200 workers under permanent contracts (refinery operation)	The labor force in the Rio de Janeiro Metropolitan Region in December 2003 was around 5 million people aged 18 to 49 years old	There might be migration from other Municipalities, with large number of workers brought in for the implementation phase. However, Itaguai has the advantage of lying close to the Metropolitan Area of Rio de Janeiro.

Table 6: Continued.

Land	The refinery requires 7-10 km^2 of area	The area under analysis covers 10.76 km^2. However, 2.79 km^2 can be flooded. The available area options are: 5.95 km^2 (without drainage project) and 8.74 km^2 (with drainage project)	Itaguai has already a site that could be devoted to a new refinery. However, a drainage system should still be implemented.
Conservation Units	The refinery cannot be installed in the conservation units.	The area has a large space of grazing land and meadows, in addition to rainforest and secondary vegetation. To the Northeast, it has a tract of floodlands. The refinery should respect the mangrove swamp, forest and secondary vegetation lands.	The site does not lie in conservation units. However, the mangrove swamp, forest and secondary vegetation should be preserved
Air Quality	The main pollutants from the refinery operation are: Sox, Nox, Volatile Organic Compounds, CO and Particulate Matters (PM).	The atmospheric pollutants concentration monitored in the region are below the ceiling amounts stipulated by Brazilian law (2); however, close attention should be paid to the sharp increase over the past few years, due to industrial expansion in this region	The carrying capacity of Itaguai will be affected by a new refinery, mostly due to region cumulative effects (impacts of other industrial facilities located or planned in the area).

Table 6: Continued.

Water Quality	The new refinery generates 1-2 million m^3 per month of wastewater with high BOD, COD and contaminants concentration (phenol, sulfides, ammonia, toxic metals).	Two rivers are important. The Guandu River complies with the specification stipulated by Brazilian law. The Guarda River does not meet these specifications, with strong indications of degradation.	Due to heavier industrial concentration, Itaguai is few resilient to water pollution.
Industrial Wastes	The new refinery generates 500-750 kg of dangerous residues (spent catalysts, moods, waxes, heat exchangers fouling etc.)	The area lies alongside the Santa Cruz Industrial District, consisting of fourteen widely diversified industrial firms.	The refinery might interact with companies in the Santa Cruz Industrial District for treating and/or recycling their industrial liquid wastes, and for treating and disposing their solid wastes

(1) Santa Cruz Power plant (766 MW) and Eletrobolt (379 MW). (2) PM_{10} 44 $\mu g/m^3$, SO_2 2 $\mu g/m^3$, NO_2 12 $\mu g/m^3$ and O_3 25 $\mu g/m^3$.

In sum, from the logistical and infrastructure standpoints, Itaguai offers several advantages, including transportation infrastructure, skilled labor force and logistical synergies with neighboring industrial districts. However, from the environmental and urban planning point of view, Itaguai faces environmental issues such as polluted water-bodies, slums (*favelas*) and improper land use and occupancy. Itaguai presents:

- High vulnerability to contamination by heavy metals
- High water vulnerability (pollution of the Guarda River and the saline wedge in the Sao Francisco Channel)
- High possibility of saturation of the air basin due to the emission of atmospheric pollutants from the operation of the refinery;
- High level of urbanization and nearby townships with dense populations.

Interestingly enough, the case presented here shows a quite common situation of urban areas in developing countries, where sites with good infrastructure (especially transportation infrastructure) also face environmental vulnerabilities. It is worth noting that industrial projects seek for areas with developed public infrastructure, but those projects create additional local

environmental pressures. In this sense, there is always an environmental opportunity cost to take in consideration: a project “X” might excludes a project “Y” due to their relevant cumulative effects over the local carrying capacity. Furthermore, such environmental opportunity cost also implies in socio-economic opportunity costs: the environmental pressures might bring more (less) socio-economic benefits for the region (income, jobs, and local taxies).

This might be the case for Itaguai Municipality, since its regional neighborhood places other large industrial facilities (such as the Santa Cruz thermo-power plant, and the Industrial District of Santa Cruz) and is eligible for other large industrial projects besides the petrochemical refinery.

4 Conclusions

Brazil might be called to expand its refining industry over the next twenty years by two to four new refineries, in addition to modifying its current facilities. One of these new refineries is focused on petrochemical products and might be located close to major consumption centers. In this case, Rio de Janeiro State is a good option. However, the decision on where to locate a refinery within Rio de Janeiro State is no trivial. It opposes the country’s requirement for a green field project to the local perception of a “gray field” project. The analysis of a possible site in Rio de Janeiro State shows the trade-offs between positive and negative external effects of a new refinery. In the Itaguai Municipality, the advantage of good logistic and infrastructure are offset by the cumulative environmental problems related to air, water and solid wastes management. Moreover, there are other large industrial projects for the region. The refinery will have to compete with other projects for the “right” to pollute, such as new steel mills under consideration for the area analyzed here.

Finally, this paper shows the need for improving the site-location study. This will request detailed information about regional economies and environments, which is a great challenge for developing countries. Building a detailed regional database has to become a top priority in countries such as Brazil.

Acknowledgments

To IBP, CAPES, CNPQ, ANP, and to Amaro Olímpio, Amanda Aragão, André Simões, Carla Achão, Francisco Mendes, Janaina Sala, Jacqueline Mariano, Jeferson Soares, Lilian Veiga, Lucy Guimarães, Marina Tavares and Milena Scheffer, for their collaboration in the early stages of this study.

References

[1] EIA/DOE. Annual Energy Outlook. Washington, DC; February 2005. www.eia.doe.gov/oiaf/aeo/

[2] ANP. Perspectivas para o Desenvolvimento do Refino no Brasil. Séries ANP **3**, pp. 1-45, 2002.

[3] MME (Ministry of Mines and Energy): Plano de Longo Prazo: Projeção da Matriz 2022 – Sumário Executivo. Brasília/Brazil: MME Press; 2002.
[4] Schaeffer R, Szklo A, Machado G, Sala J, Mariano J, Tavares M, Magrini A. Evolução do mercado brasileiro de derivados de petróleo e perspectivas de expansão do parque de refino. Technical report presented in IBP Seminar on The Future of Oil Refining in Brazil. Rio de Janeiro, November 23; 2004.
[5] Schaeffer R, Szklo A, Machado, G. Brazil's Long Term Energy Plan Revision – 2002-2023. Technical report for the Ministry of Mines and Energy. Brasília/Brazil: MME Press; 2005.
[6] Luaces, A Brazil Refined Products Market Outlook. Technical Paper (Purvin & Gertz Inc) presented in IBP Seminar on The Future of Oil Refining in Brazil. Rio de Janeiro November 23; 2004.
[7] Tavares M, Szklo A, Schaeffer R, Machado G. Oil Refining Expansion Criteria for Brazil. Energy Policy **In press** (accepted), 2005. Available on line (website Science Direct, since May 2005).
[8] ANP. Brazilian oil yearbook database. 2005. Available at www.anp.gov.br.
[9] International Atomic Energy Agency [IAEA]. Brazil: A Country Profile on Sustainable Energy Development. International Atomic Energy Agency/ United Nations (IAEA/UN): Vienna (in press); 2006.
[10] Tolmasquim M, Szklo A. Energia na Virada do Milênio: A Matriz Energética do Brasil - 1998-2010. Rio de Janeiro/Brazil: Editora da COPPE/UFRJ; 2000.
[11] Tolmasquim M, Cohen C, Szklo A. CO_2 Emissions in the Brazilian Industrial Sector According to the Integrated Energy Planning Model (IEPM). Energy Policy **29,** 641-651, 2001.
[12] Szklo A. Machado G. Schaeffer R. 2005. Perspectiva da produção de petróleo no Brasil. Modelagem a partir de Hubbert. In: 3° CONGRESSO BRASILEIRO DE P&D EMPETRÓLEO E GÁS, 2005, Salvador. 3° Congresso Brasileiro de P&D em Petróleo e Gás. 2005.
[13] Szklo A, Machado G, Schaeffer R. Simões A, Mariano J. Placing Brazil's heavy acid oils on international markets. Energy Policy, **34**, pp. 692-705, 2006.
[14] Worrell E, Galitsky C. Profile of the Petroleum Refining Industry in California. California Industries of the Future Program. Energy Analysis Department. Environmental Energy Technologies Department. Berkeley; 2004.
[15] Zai-Ting L, Chao-Gang X, Zhi-Gang Z, Jiu-Shun Z. Olefin production technology with adjustable propylene/ethylene ratio by catalytic cracking route. 17th World Petroleum Congress. Rio de Janeiro. Brazil; 2002.
[16] IBGE. Brazil's socio-economic database; 2005. Available at www.ibge.gov.br.
[17] Gomes G, Dvorsak P, Heil T. Indústria Petroquímica – Situação Atual e Perspectivas. Technical Paper – February 2005. Rio de Janeiro: BNDES press.

This page intentionally left blank

Utilization of coal and waste for ecological fuel production

P. Sedlacek, J. Vales & M. Safarova
Department of Research and Development,
Brown Coal Research Institute, j.s.c., Czech Republic

Abstract

When processing brown coal mined in the North Bohemian Basin, dusty types of low heating coal containing lots of ash are produced. This coal is difficult to sell. On the other hand when bituminous coal is produced, the waste is very fine, with a high calorific value, which is impossible to sell. World wide continuously increasing amounts of waste and their further processing is a big problem. For processing waste combined with coal the most suitable types of waste are those individual types of coal which use favourable qualities of reprocessed high-energy waste.
Keywords: brown coal, waste, desulphurization, sulphur emissions.

1 Introduction

Reliable and economically available energy sources are a key issue for the good economic function of any state. Therefore the basic aim of energy policies of European states is to ensure energy self-sufficiency and at the same time reduce the energy consumption of consumers.

Globally, coal is the second most-used raw energy material in the world (after crude oil). In 2002 26% of world energy was produced by burning coal (more than 4.74 milliard tons of coal). According to an estimation by the IEA, world demand for coal will rise by approximately 1.4% per year until. It is expected that the amount of electricity produced in coal power plants will rise by 4% per year over the same period. Coal as a fuel has a number of economic advantages. It is situated in rich deposits that are more evenly located when compared to crude oil, and even less-developed countries are able to handle its, which is not too economically demanding.

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060781

A significant disadvantage of producing electricity from coal is the operation of ecologically-damaging combustion equipment, which generates a huge amount of pollution, including so-called greenhouse gases which cause global warming and climate change.

One possible way of using remaining relatively large coal deposits and lowering levels of air pollution are so-called Clean Coal Technologies, e.g. separation of carbon dioxide from the combustion products and depositing into underground reservoirs (e.g. closed mines) [2].

The main energy source in the Czech Republic is coal. The main areas of coal mining in the Czech Republic are in the Ostrava region and in the Ore Mountains region. In the Ostrava region there is mined coking bituminous coal in deep mines and in the opencast mines of the Ore Mountains region there is brown coal.

Burning coal, particularly brown coal, is accompanied by the production of carbon dioxide, sulphur oxides and other substances, which are emitted into the air. The issue of pollution produced by burning dusty coal, such as those used in heating and energy production industries, is remedied by the subsequent desulphurisation of combustion products or construction or reconstruction of existing combustion equipment to cope with more modern fluid technologies that meet the valid legislation regarding pollution.

Sorted coal has a higher sulphur content and when burned, pollution limits laid out in legislation are exceeded. The additive procedure, which is used mainly for dusty types of coal, was in the past verified for sorted types as well. However, efficiency was very low when burning. Suitable fuel for coal boilers with a low heat output, which burn the sorted types of coal, and which are often the biggest local producers of pollution, could become multi-component fuel made from brown coal and biomass waste as well as some combustible waste materials modified by admixtures so that burning would produce significantly less pollution, especially SO_2 (compared to burning just brown coal.) The aim of this project was to verify the production possibilities of such fuel [1].

Table 1: Total annual production of brown coal in the Czech Republic in the years 2002 – 2004.

Mining company	**2002**		**2003**		**2004**	
	Thousands of tons	**%**	**Thousands of tons**	**%**	**Thousands of tons**	**%**
MU, j.c.s.	16,6	33,1	16,9	32,6	16,6	33,3
SD, j.s.c.	21,8	43,8	23,4	45,2	22,0	44,9
SU, j.s.c.	11,3	23,1	11,4	22,2	10,3	21,8
Total	**49,7**	**100**	**51,7**	**100**	**48,9**	**100**

2 Mining of brown coal in the Czech Republic

In the Czech Republic there are three companies that mine coal in two areas; the North Bohemian Coal Basin and the Sokolov Coal Basin. In the Table 1 the

overall production of brown coal in the Czech Republic and market share of the companies at the market are summarised.

3 Components of ecological fuel

The verification of ecological fuel production was based on the use of three components; brown coal, waste and biomass.

3.1 Brown coal

From the very beginning of the project, coal had been considered a major constituent of the new fuel type. The fuel should always consist of at least 30% coal. Brown coal with a high sulphur content from MU, j.s.c. was used for this purpose. The problem, which is being solved by the treatment and processing of brown coal, is using low heating dusty types of coal. The types of coal that are created during the treatment of the sorted types are mostly used for burning in large power and heating plants. Their calorific value is about 11 $MJ.kg^{-1}$. Another problem is the higher sulphur content, which is one of the reasons for the limited use of these types for direct burning in boilers of small thermal performance regarding the $SO_{2\ pollution}$ [4].

3.2 Biomass

Industrial crops grown for energy are called energy crops. The final products of energy crops processing are bio-fuels (phyto-fuels), which are categorised according to their states as solids (briquettes, pellets), liquids (vegetable oils, bio-diesel oil) or gases (biogas). Currently, in CR energy, biomass utilization ranges from 1,5% to 2% of the total primary resources. However, the target for the year of 2010 is to ensure that 8% of power and 6% of heat energy come from renewable sources. Table 2 shows an overview of biomass to ensure such targets [3].

Table 2: Energy production from biomass in 2010.

Biomass type	Energy (%)	Total (PJ)	Out of this heat (PJ)	Power (GWh)
Wood and wood waste	24	33,1	25,2	427
Grain and oil plant straw	11,7	15,7	11,9	224
Energy crops	47,1	63	47,7	945
Biogas	16,3	21,8	15,6	535
Total	**100**	**133,6**	**100,4**	**2231**

3.3 Combustible waste for energy purposes

Recently there have been increasing efforts to increase energy production from renewable resources. Approved state power industry conception assumes an increase of a portion of renewable resources for the production of electrical energy by the year 2010, up to a total of 8%. In the Czech Republic the energy

from renewable resources accounts for approximately 2.3% of the present consumption of primary energy resources. The largest contribution to power from renewable resources of energy (RRE) has solid biomass (74%) followed by water power (12%) and biologically decomposable portion of incinerated waste (5%). Liquid bio-fuels reach over 6% of total power production from RRE [5]. Usage of biogas for power production is about 3% and wind energy usage is so far totally negligible. From this proportion, it is clear that the biomass will be the principal article used to achieve the binding goals in 2010. It would be wrong not to take into account another class of potential power resources, i.e. waste. These are the substances that are not recyclable under the present available technical possibilities, while keeping balance of environmental, economical and social aspects. Because this waste is useless even as a secondary raw material, it usually ends up on tips, with a negative impact on living environments, even though it has considerable power potential. Such waste substances are however, pre-treated for fuel production, so harmful substances are eliminated and its energy potential may be fully used.

4 Experimental works

In terms of the first phase of experimental work a few samples of multi-component fuel were prepared from the following components; brown coal, waste biomass, waste paper. For the preparation of ecological fuels, the input coal was high-sulphur coal. The coal was sampled in the Komořany preparation plant prior to a production of graded ranks (i.e. nut coal, cobbles) [2]. The first sample was prepared from coal and a high proportion of biomass. The second sample was prepared from coal and scrap paper. The final sample was prepared from coal and a mixture of energy-utilisable wastes, mainly crushed plastic bottles and paper. Table 3 displays a recipe according to which briquette samples have been prepared.

Table 3: Fuel recipes.

	Content of fuel constituents-press mixture (in wt per cent)		
Fuel constituent	**Sample 1**	**Sample 2**	**Sample 3**
Brown coal	45,05	42,37	42,38
Energy sorrel	45,05	0	16,95
Dry hydrate	2,70	2,54	2,54
Ground scrap paper	0	42,37	16,95
Crushed plastics	0	0	8,47

The fuel was prepared in the form of briquettes on laboratory press KAHL-14-175. The pressing procedure consists of rolling the input material on the matrix by crushing wheels, which causes precompacting, and subsequently the material is pushed into the matrix pressing channels. These channels are shaped so that the pressed material is compacted again and a cylindrical cord of pressed material is pushed out of the matrix [6]. This cord is cut into lengths by means of a cutting machine.

In terms of the next phase of experimental work, some samples of ecological fuel were prepared from the following components: dusty brown coal, solid product from the pyrolysis of waste rubber; waste fraction from the production of charcoal; products of a wastewater treatment plant. In the accredited test laboratory of VÚHU a.s., a simplified laboratory testing pyrolysis unit was developed with retort made from special material for the thermal processing of brown coal with a weight of approximately 1000g. The testing unit is equipped with a cooling and separating circuit and a burner for the gas being originated (with properties similar to coal gas).

The temperature program of the laboratory pyrolysis unit (retort) was set to a final temperature of pyrolysis of 750°C with a temperature rise and time delay as required, and set temperatures. Figure 1 shows the product of the wastewater treatment plant. Figure 2 shows the ecological fuel produced from this product.

Figure 1: Product of wastewater treatment plant.

4.1 Combustion tests

The combustion tests with prepared briquettes were carried out at the Research Power Engineering Centre of the Mining College at the Technical University of Ostrava. Based on the previous experience during combustion tests, a Viadrus Ling 25 boiler was used. During the course of the burning tests, no substantial problem with fuel combustion occurred. In the case of sample 1 combustion, SO_2

emissions reached approximately 1,328 mg.m^{-3}, which corresponds to a desulphurization of 62,4%. With this sample, a high concentration of emitted CO was measured, which however, complies with the knowledge of biomass combustion, and of sorrel in particular. In the case of Samples 2 and 3, the levels of desulphurization were 63.5% and 63.3% respectively. Figure 3 shows combustion of sample 1. In Table 4 the results of combustion tests are shown.

Figure 2: Ecological fuel prepared from the product of wastewater treatment plant from a combination of brown coal and biomass.

5 Conclusion

The aim of this project was to verify experimentally and in semi-operation the possibility of using reprocessed waste as components of a new type of fuel for small and middle-size sources including combustion tests with pollution measuring. We can consider the following points as the main benefits of the project:

- Utilisation of fine fractions, impossible to sell, produced by treatment of quality types of brown coal
- Utilisation of brown coal with high sulphur content for production of ecological fuel using biomass and additives

- Utilisation of selected types of burnable waste as another component of ecological fuel based on brown coal with a high sulphur content
- Utilisation of renewable energy sources based on biomass
- Burning of new types of products in existing modern types of combustion boilers without having to undertake any constructional modifications.

Based on the results, it is possible to say that such ecological fuel can be used as a substitute for burning pure brown coal with a higher sulphur content.

Figure 3: Combustion of sample 1.

Table 4: Results of combustion tests.

Fuel	CO $[mg.m^{-3}{}_{n}]$	CO_2 $[g.m^{-3}{}_{n}]$	SO_2 $[mg.m^{-3}{}_{n}]$	NO_x $[mg.m^{-3}{}_{n}]$	Efficiency of desulphurization [%]
Brown coal	**1 186**	**251**	**3 531**	**432**	**-**
Sample 1	5 318	263	1 328	313	62,4
Sample 2	1 356	270	1 290	351	63,5
Sample 3	1 885	269	1 294	386	63,3

Acknowledgements

This work has been created in terms of the research intention MSM 4456918101 "Research of physico-chemical properties of materials involved in mining and utilisation of coal and their influence on the environment in the region of north-west Bohemia" with the assistance and contribution of the Ministry of Education, Youth and Physical Training of the Czech Republic.

References

[1] Sedláček, P., Maček, S., Šafářová, M., Kusý, J.: Využití pyrolýzy odpadů pro zvýšení užitných vlastností paliv ČR, Sborník Mezinárodní konference Recyklace odpadů VIII, s. 345-350, ISBN 80-248-0560-X (Using waste pyrolysis for increasing fuel qualities in the Czech Rep., Collection of International Waste Recycling Conference VIII)

[2] Sedláček, P., Valeš, J., Fečko, P., Čablík, V.: Výroba směsných pelet z hnědého uhlí a biomasy, Publikace: Zpravodaj Hnědé uhlí 4/2003, s.5-10, ISSN 1213-1660 (Production of mixed briquettes from brown coal and biomass)

[3] Obnovitelné zdroje energie v roce 2003, Stručný přehled zpracování výkazu ENG (MPO) 4-01), MPO 2004 (Renewable Energy Sources- Brief overview)

[4] Roubíček, V., Buchtele, J.: Uhlí - zdroje, procesy, užití, Ostrava 2002, Sborník. 6th Conference on Environment and Mineral Processing, part II,s.797-802, ISBN 80-248-0072-1 (Coal-sources, Processes, Utilisation, Ostrava 2002, Collection)

[5] Nováček, J.: Technologie úpravy uhlí II, VŠB TU Ostrava, (Coal Processing Technology)

[6] Švehla, E, Bernáth, V.: Úpravnické stroje a zařízení, SNTL 1964 (Treatment Machines and Equipment)

Injection technology for sustainable environmental protection in the petroleum industry

B. Muvrin, Z. Kristafor, K. Simon, L. Maurovic
& D. Karasalihovic
University of Zagreb,
Faculty of Mining, Geology and Petroleum Engineering,
Department of Petroleum Engineering, Croatia

Abstract

Since the petroleum industry is considered to be one of the main environmental polluters, increased attention is put on environmental protection during petroleum activities. During exploration, well stimulation and oil and gas production a certain amount of waste is produced. By using injection technology petroleum industry waste is injected into suitable underground geological formations, from where it originated and it is safely removed from the human biosphere. With the injection technology the liquid, solid and gaseous wastes are deposited simultaneously and safely into the selected formation. Compared to the injection technology, surface disposal of petroleum industry waste is much more expensive and more environmentally intensive. For the last two years the same technology has also been used in Croatia. At this moment, petroleum industry waste in Croatia is injected at just one oil field, but there are several projects that consider wells and geological structures suitable for disposal of this kind of waste. With the negative publicity regarding environmental impact in the past, and considering the new set of environmental laws and standards, the waste injection technology could be the solution for the petroleum industry on how to achieve sustainable waste management and diminish its impact on environmental balance that is predicted to be even greater. Furthermore, implementation of injection technology could also be the solution for the greenhouse gas problem.
Keywords: petroleum industry waste, injection technology, geological formation, CO_2 sequestration.

WIT Transactions on Ecology and the Environment, Vol 99, © 2006 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060791

1 Introduction

In last few decades more strict and resolute regulations regarding environmental protection have been posed to all energy related industries among which the petroleum industry is considered to be one of the main polluters. All this led to significantly increased attention surrounding environmental aspects in petroleum companies and to the fact that environmental considerations have attained a central position in the formulation of petroleum and energy companies's policy. Principles of environmental protection have successfully integrated into all segments of upstream operational procedures and activities, from exploration work, through drilling operations, field development, production of oil or natural gas and ending by field abandonment.

One of the emerging problems of petroleum companies, due to ever more strict environmental protection standards and laws, in last few decades is the problem of management of waste generated during exploration, drilling, production and refining activities. Most of all petroleum companies have developed waste management systems, which are more or less based on re-use, recycling, reduction and disposal of waste throughout every stage of operations. Many aspects of waste management are covered by regulations at national, regional and international levels. Regulatory requirements are usually embodied in licenses granted for specific operations. Due to mentioned environmental standards, safe disposal of waste, along with the limitation of world oil and gas resources, its disproportional global distribution and costs of production, has become one of the limiting factors and its impact on the future development of petroleum industry is predicted to be even greater.

2 Petroleum industry waste

Drilling during oil and natural gas exploration and production can produce considerable quantities of rock cuttings coated with residual amounts of drilling mud, which is lubricating fluid that, among other functions, prevents well bore collapse while drilling and transports drilled cuttings to the surface. Drilling mud is a complex mixture containing a number of additives which can be either water-based or oil-based. The drilling mud is considered to be primary technological waste from petroleum industry and, as opposed to secondary fluids, does not represent a big environmental problem. The bigger problem lies in the secondary fluids which include inorganic and organic acids, cellulose and oil gels, cement slurry, technological wastewater and salt solutions that cannot be used again, but should be neutralized with a proper neutralizer like lime, limestone or dolomite, and deposited in the environmentally friendly manner. Also, waste like brine, sand, CO_2, reservoir and process facilities residua, oily water, paraffin etc., is generated during oil and natural gas production.

According to international standards technological waste from petroleum industry generated during exploration, drilling, production, refining and distribution activities can be, considering its generation site, divided into several groups as shown in table 1.

Table 1: Types of petroleum industry technological waste.

Waste	Content and physical and chemical characteristics
Brine produced during hydrocarbon production	High content of chlorides, oil in traces, small size particles of clay, other dissolved salts
Sand produced during hydrocarbon production	Small to middle size granulated sand produced along with heavy oil, > 85% SiO_2, possible high content of chlorides, 4–5% wg. oil residua on sand particles
Waste mud	Rock cuttings, chlorides, sand, clay, water, chemicals, 8% of oil in emulsion, high pH,
Slime from storage tank bottom, stable emulsions	Slime form separator tanks, chlorides, clay, brine, asphaltens, variable oil content from 30–50% stabile emulsion
Slime	Waste water, oil with particles of sand and clay, soil polluted with oil (surface oil spills), chlorides, cycle hydrocarbons

The amount of the waste produced during exploration and production activities depends on range of exploration and reservoir engineering activities, workover operations, production conditions and formation characteristics.

As hydrocarbon reservoirs are developed, water, which co-exists with oil and natural gas in the reservoirs, is brought to the surface. In oil fields in particular, quantities of produced water increase as reservoirs are depleted. In some cases, eventually more water is produced than oil. Since water production directly affected the profit of oil companies it was the first waste problem that was addressed and solved. The solution for water production was the development of injection technology which was introduced by petroleum industry only several decades after the beginning of commercial petroleum production. By the injection of water, today commonly known as the secondary recovery method of hydrocarbon production, into the hydrocarbon reservoir the production of hydrocarbons is enhanced and at the same time the problem of water as technological waste is solved.

With time, pressurized with waste problems, petroleum industry has improved the injection technology to the point where it can be used, not only for the water injection, but also for the injection of other types of waste. The petroleum industry waste injection technology is nowadays commonly implemented by most of the large world oil companies with approval and under the supervision of authorized state institutions.

The US Environmental Protection Agency (EPA) was one of the first regulatory organizations that set standards and rules for underground injection under Underground Injection Control Program. According to EPA injection activities are grouped into five classes. Class I includes injection of hazardous and non-hazardous fluids (industrial and municipal wastes) into geological formations beneath the lowest underground source of drinking water. Because Class I includes injection of hazardous waste, Class I wells are the most strictly regulated. Class II includes injection of fluids produced during oil and gas production. Class III encompasses injection of fluids associated with mineral

mining activities. Injection of hazardous or radioactive waste into or above underground sources of drinking water is included in Class IV. Class V includes all underground injection not included in Classes I–IV [4–6].

3 Petroleum waste injection technology

The injection of different kind of petroleum industry waste is done using specially selected and registered wells. Wells, commonly used for waste injection, are negative wildcat wells, former production wells or injection wells, which are equipped especially for the injection purposes. Wells are located in areas where the waste is originally generated due to costs reduction of temporary deposition and transport. By injection technology, liquid, solid and gaseous waste can be simultaneously and safely deposited into the selected geological formation. Waste injection involves processing solid materials, produced during exploration, production and workover activities, to particles of suitable size and blending them with a fluid to make the slurry that can be injected into selected geological formations. Injection of slurry can be continuous or as a series of smaller volume cycles. Injection of waste can be conducted continuously for a few weeks, even months, after which follows non-injection period with pressure measurements. This period should be as long as to allow total dispersion of pressure and repeated estimation of reservoir parameters. During the period of injection, the formation characteristics are affected and changed by the injected waste material. Due to that fact, injection of waste into underground formations requires, along with standard requisite and obligatory monitoring, also constant monitoring and analysis of the formation characteristics and injection process. Injection of petroleum waste should be conducted according to recommended physical and rheological characteristics of waste due to feasibility specifications of injection facilities. The characteristics of waste to be injected are set by injection facility performances.

As regards technical aspects injection could be done through casing string or the injection tubing inside the long string casing either through perforations in the long string or in the open hole below the bottom of the long string. The annulus between injection tubing and casing string is filled with an inert, pressurized fluid and is sealed with packer preventing injected wastewater from backing up into the annulus. The completion of injection wells is shown in figure 1. The injection of the slurry is performed by injection facility that consists of a facility for preparation of the slurry, injection pumps and injection well (figure 2).

There are several advantages of injection technology over commonly used surface disposal of petroleum industry waste, including following:

- Less area intensive – the area that is occupied during injection activities is much smaller than the area occupied in case of surface disposal.
- Minimal chemical treatment of waste
- Negligible impact on the environment during normal work – the deep well injection is, at the moment, considered to be only permanent on-

site disposal method available that can fully comply with zero discharge to the surface environment. By the injection cuttings are disposed in their native environment.

- Risk of environmental impact is at lowest rate compared to all know surface disposal methods (table 2).

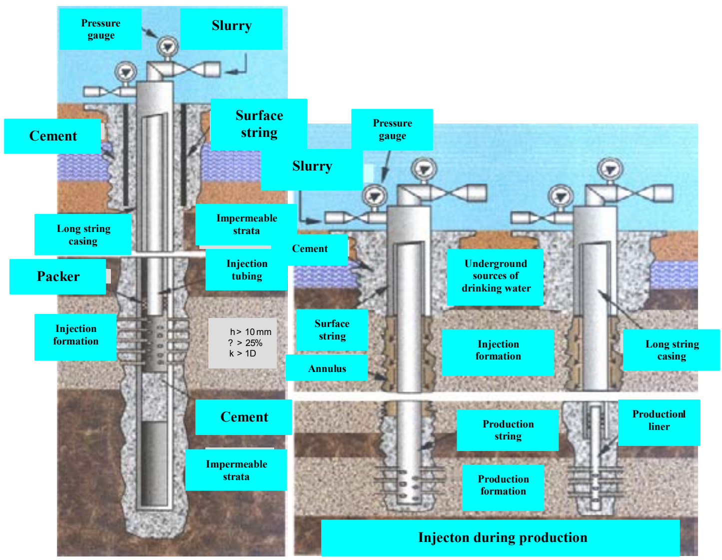


Figure 1: Completion of injection wells.

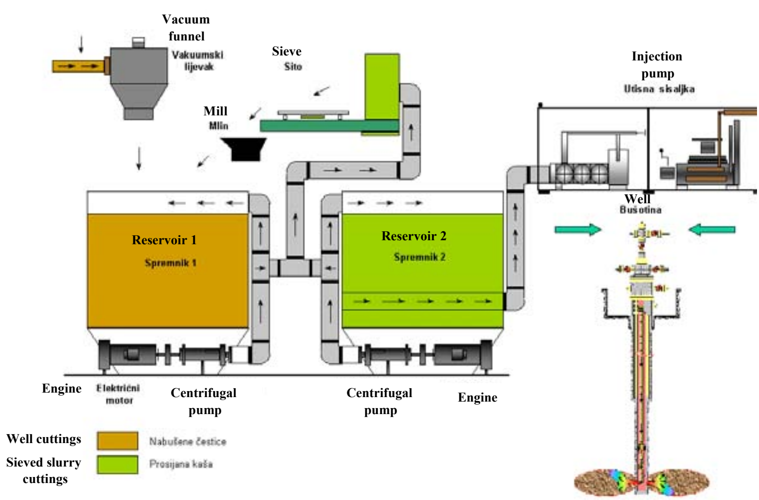


Figure 2: Mobile facility for petroleum industry waste injection.

Table 2: Environmental risk of waste disposal methods.

Methods of waste disposal	Environmental risk
Surface landfill	**high**
Shallow burring on the spot	**high**
Disposal into salt mines	**low**
Land Spreading	**high**
Deep well injection	**low**
Disposal into barrels on the spot	**Very high**

In addition to lower capital investments, reduced operating costs and minimal risk of environmental impact, the advantage of the injection technology, in comparison with the surface waste disposal, is a permanent disposal of waste into underground geological formation, from where the most part of the waste (water, CO_2, cuttings, send, etc.) originated, by which the waste is permanently, safely and irretrievably removed from the immediate environment of the people (biosphere).

On the other hand, there are some factors that limit the applicability and effectiveness of petroleum waste injection like potential seismic activity, compatibleness of injected slurry with the existing fluid in the injection pore space and content and characteristics of the injected waste. In general, injection wells have to be located in seismically stable areas that are fee of transmissive fractures or faults through which injected fluids could vertically migrate.

The geological formation into which the waste is to be injected has to be sufficiently porous and permeable so that the waste can enter the rock formation without an excessive build up of pressure. The selection of injection wells and geological formations suitable for waste disposal is done based on tectonic and stratigraphic underground relations, physical characteristics and hydrogeological content of formation, physical and chemical characteristics and content of the waste, completion of the well and working regime during the field production. The formations suitable for disposal (injection) of waste are those isolated from their environment by impermeable cap and bed-rock, preventing uncontrolled dispersion and migration of waste out of boundaries of selected formation and pollution of underground water in shallower underground layers. Also, important factor of injection well selection represents economical warranty of injection activities. There are three critical factors that affect cost effectiveness of petroleum industry waste injection- volume of waste to be disposed, the availability of disposal infrastructure and the state regulatory system.

4 Environmental impact of injection activities

Since the injection activities are conducted in closed system, during normal working conditions, the impact on the environment is not expected. For ensuring a safe system, degree of safety during the injection activities is put on the highest

level with the scope of additional environmental protection and safety measures. Prior to waste injection the environmental impact analysis of the injection activities should be done. Furthermore, it is necessary to analyze all other relevant characteristics of selected injection well, selected permeable formation, wellhead, content and characteristics of waste, well stability for injection purposes. Environmental protection measures and monitoring of injection activities should be conducted during the injection of waste into geological formation with high importance put on analysis of possible impacts of the injection on underground water.

For avoiding possible accidental situations the rules for construction, operating, monitoring, record keeping and reporting have been set.

Approved injection well, due to its construction and completion does not represent environmental threat during the non-injection period. However, during injection activities the environmental protection measures are conducted according to standard working procedures and set measures implemented in technical documentation of injection facility. After injection period the injection well should be stabilized.

Environmental protection measures, which are compulsory in Croatia, are operational safety measures, soil and underground water protection measures, fire and explosion protection measures and protection measures of the working area in accidental situations [1].

Along with environment protection measures taken, on injection well there should be constant monitoring and measurement of [1]:

- Static temperatures in specific time periods
- Spreading of waste injected into the formation (radioactive markers, oxygen activation method)
- Pressure measurement at wellhead and in annular space of the injection well
- Analysis and monitoring of pressure on nearby wells
- Analysis of integrity of the injection well
- Well logging
- Monitoring of chemical, physical and rheological characteristics of the injected waste
- Induced seismic and seismic monitoring

5 Sequestration of CO_2 by using injection technology

Every year, human activities emit around 25 billion tons of CO_2 into the atmosphere. About 2 to 3 billion tons of this annual output is absorbed by forests and another 7 billion tons is absorbed by the ocean [3]. Petroleum industry is one of GHG emission intensive industries. Due to increased global consciousness regarding climate change problems and increased emissions of greenhouse gases, carbon dioxide, as one of the wastes generated during hydrocarbon production, has gained a special place in petroleum industry's hierarchy of priorities. Beside direct CO_2 and CH_4 emissions with oil and gas production, petroleum industry is also characterized with high-energy intensity and final products with high

emission potentials. The petroleum industry has been injecting carbon dioxide into geological formations for about 30 years. This injection has not been done in order to sequestrate CO_2 but to displace or dissolve oil for increased/enhanced oil recovery (IOR or EOR). The pressure created by injection of carbon dioxide into the reservoir rock drives crude oil towards the production well. The injected CO_2 emulsifies and partially dissolves the crude oil, which becomes more fluid and runs more easily to the production well.

Based on the Kyoto Protocol committed parties are obliged to decrease CO_2 emissions for certain amount. Therefore, it is necessary to decrease industrial CO_2 production and to manage produced CO_2 in a safe manner. One of possible solutions for managing CO_2 problem, not only in petroleum industry business, is its sequestration into underground formations by using injection technology. The first geological CO_2 sequestration project was conducted by Statoil in the mid 1990s at the Sleipner Vest field. From the beginning of the project nearly 1Mt CO_2/year is removed from natural gas and injected in to a water saturated sandstone formation deep under the North Sea [2]. Since over the past decade underground CO_2 sequestration has emerged as one of the most promising methods for CO_2 emission reduction, today there are several more commercial projects underway: the In Salah project in Algeria, the Gorgon Project in Australia, and the Snohvit Project in the Norwegian continental shelf. In addition to these, more are under development.

Key factor in process of disposal of CO_2 represents adequate defining and choice of suitable geological formations. Today, four principle types of geological formations are considered to have significant potential for storing large amounts of CO_2 [2]: active and depleted oil reservoirs, active and depleted gas reservoirs, saline formations and deep coal seams and coalbed methane formations.

Summary of worldwide storage capacity estimates is shown in table 2.

Table 3: Summary of worldwide storage capacity estimates.

Formation type	Capacity estimate (Gt CO_2)
Depleted oil and gas reservoirs	~450
Coalbed methane reservoirs	60–150
Salt – water saturated formations	300–10,000

Carbon dioxide can be stored in these geological formations by four principle processes [2]:

- It can be trapped as gas or supercritical fluid in the pore zone of the injection formation
- It can dissolve into the fluid phase-solubility trapping
- It can become trapped a residual, non-wetting phase in the pore spaces of the formation
- It can react with the minerals and organic matter and become part of the solid mineral matrix.

Injection of CO_2 into deep underground geological formations, as a method to eliminate emissions of CO_2, will not solve the problem of global warming itself, but it could dramatically reduce the amount of heat-trapping greenhouse gases released into the atmosphere.

6 Implementation of injection technology in Croatia

Like the others, so the Croatian national petroleum company, in last decade was also developing techniques and technologies to manage waste produced during oil and natural gas exploration and production by injection of this kind of waste into isolated, integral hydrodynamic underground geological formations by using deep well injection technology. During the intensive exploration and production of hydrocarbons in Croatia, nearly 5000 wells were drilled. Since, some of the drilled wells were wildcat and a great majority of rest was cut out of production during the time, there are a large number of unused wells. This fact enables Croatian petroleum industry to select wells suitable for safe and efficient petroleum industry waste disposal. Currently, petroleum industry waste is injected in only one well (Kal-6) on Kalinovac gas-condensate field located in the northern part of the country, in Koprivnica-Krizevac County. The selection of Kal-6 well and its transformation from production into injection well has been done based on hidrogeological characteristics of reservoir, well completion and field production regime.

Pilot EOR project in Croatia started on mature oilfield Ivanic in 2001. The project was conducted in 5 phases that include successively injection of brine and CO_2. The expected result of CO_2 injection into mature oil field Ivanic in Croatia is 7–8% increase in oil production and prolonged production for at least 15 more years. Beside the potential opportunity of enhanced oil recovery from Ivanic and Zutica fields, important for domestic energy supply, the EOR project in Croatia also offers sustainable solution for the climate change problems and the challenges posed by the Kyoto Protocol commitments. By this project it is possible to sequestrate the 0,7 million tones per year of CO_2 produced on gas field Molve, the biggest CO_2 producer in Croatian petroleum industry and one of the biggest stationary emitters in the country.

7 Conclusion

By using the injection technology for disposal of waste generated during petroleum industry exploration and production activities into selected geological formations through technically suitable wells, the process of petroleum industry waste disposal will be, not only accelerated, but also conducted in a safe and efficient manner with total injection costs reduction. One of the main advantages of injection technology is permanent disposal of waste deep beneath the surface with “zero” pollution of the environment. The formations suitable for disposal (injection) of waste are those isolated from their environment by impermeable

cap and bed-rock, preventing uncontrolled dispersion and migration of waste out of boundaries of selected formation and pollution of underground water in shallower underground layers.

Over the past decade underground CO_2 sequestration has emerged as one of the most promising methods for CO_2 emission reduction. Injection of CO_2 into deep underground geological formations, will not solve the problem of global warming itself, but it could dramatically reduce the amount of heat-trapping greenhouse gases released into the atmosphere.

References

[1] Group of authors, Environmental Impact Study for the Petroleum Waste Injection in Osjecko- baranjska County in the Republic of Croatia, Faculty of Mining, Geology & Petroleum Engineering, Zagreb, 2003

[2] Benson, S. M., Overview of Geologic Storage of CO_2, Carbon Dioxide Capture for Storage in Deep Geologic Formation, Vol. 2, pp. 665-670, 2005

[3] Lawrence National Laboratory, Science and Technology Review, pp. 20-22, 2005

[4] Environmental Protection Agency, Statement of Basis and Purpose of Underground Injection Control Regulations, 1980

[5] Environmental Protection Agency, Class I Underground Injection Control Program: Study of the Risks Associated with Class I Underground Injection Wells, 2001

[6] Environmental Protection Agency, Code of Federal Regulations Protection of Environment 40, parts 139-149, 1998

[7] Apps, J. A., Tsang, C., Deep Injection Disposal of Hazardous and Industrial Waste, 1996

[8] www.ina.hr

Desulphurization of flue gases using waste of a water treatment plant

J. Vales, P. Sedlacek, S. Macek, L. Chytka & J. Durik
Department of Research and Development, Brown Coal Research, Institute, j.s.c., Czech Republic

Abstract

All activities in a technically developed society are accompanied by waste generation. The question of its removal or consequent environment friendly and economic use has been dealt with by mankind since the early 1990s. The target of a common international project worked on by the Investex Group, .a.s. companies and the Research Institute for Brown Coal, the joint-stock company, is to look for ways to use the sludge from waste water treatment plants as an alternative resource for additive agent fabrication. Suitable waste treatment sludge for the additive agent preparation applied in the combustion of lower quality fossil fuels with a higher content of sulphur is generated as waste in ammonia fabrication from the sludge produced by municipal waste water treatment plants. The procedure in the agent fabrication from waste treatment sludge, the course of the combustion tests and the results of a theoretical level of desulphurization form the content of this paper.
Keywords: water treatment plant, lignite, desulphurization limestone.

1 Introduction

One of the important sources of energy in the Czech Republic is still the burning of coal. In order to reduce pollution, it is necessary to reduce the amount of harmful substances from burning coal, in different ways. Large energy sources, such as power plants, use combustion gas cleaning. One of the main harmful substances that is being removed is sulphur dioxide. Its concentration in combustion gas is directly proportional to the sulphur content in the coal being burned. During the burning of coal in fluid boilers, a desulphurization additive is added to the fluid bed – mostly limestone or dry hydrate which causes part of the

doi:10.2495/RAV060801

sulphur dioxide to react with limestone ions to produce calcium sulphate. This calcium sulphate becomes a part of the ash. When using sufficient amounts of a desulphurization agent, the concentration of sulphur dioxide in the combustion gas will be reduced accordingly. Using the already-mentioned desulphurization additives, this represents a method for reducing the sulphur dioxide load on the environment, but also increases costs for energy production. In terms of cost reduction it is necessary to find new types of additives, which have a similar efficiency, but are not so financially demanding. One such additive could be the modified sludge of wastewater treatment plants which is stabilized by means of limy ions. Such sludge contains a certain amount of useful energy and can contain a significant proportion of limy ions, depending on the modification method used. The aim of this study was to verify in the lab the desulphurization activity of three types of stabilized sludge (prepared using different stabilization methods) and to evaluate their suitability as a desulphurization additive for fluid burning of the Most region brown coal.

2 Properties of the modified sludge of a waste water treatment plant

Sludge from a waste water treatment plant has limited potential for further use so far. To be able to use it further, the sludge stabilization method is used, with the help of limy ions. Due to the increasing alkalinity of sludge as a result of adding lime (pH values mostly over 12) pathogenic organisms are liquidated during sludge processing.

Appropriate sludge for the preparation of additives applicable by combustion processes of poorer quality fossil fuels with increased levels of sulphur is produced as waste of ammonia production of sludge which is produced in communal wastewater treatment plants. Ammonia is produced as a result of the chemical reaction when ammonia salt contained in produced sludge and hydrated lime (which is oversupplied) are mixed. After this reaction, the waste sludge containing useful energy substances and a surplus of lime ions is used as material for the preparation and production of additives. During the combustion process, the additives are necessary for binding the rising gaseous sulphur dioxide and converting it to solid calcium sulphate, which stays in the ash.

For the lab tests three wastewater treatment plant sludge samples were prepared in cooperation with a Slovak company INVESTEX-GROUP s.r.o (marked as ADK1 to ADK3), stabilized by means of limy ions. The prepared samples differed in their production method, amount and type of limy agents used. The recipe of sample preparation is obvious in Table 1. For preparation of samples, common types of agents were used for addition of Czech brown coal.

As the suggested new additive contains a certain amount of high-energy particles lab analyses were made of the prepared samples, whereby the basic parameters of these samples were determined; calorific value, ash content, water content and sulphur content. The results of these analyses are shown in the Table 2.

After burning the contents of effective components in the prepared samples were determined. It was the content of free CaO and total content of carbonates. These values of individual samples were compared with values of normally used additives, mostly then of limestone from Čížkovice, which is added at the coal treatment plant Komořany. Due to different recipes, the modified samples differ in the content of limy ions. The highest concentration of limy ions was found in sample ADK 1 and was comparable with the concentration of these ions in a commonly-used desulphurization additive used in fluid combustion which is the limestone from Čížkovice with a of CaO of approximately 40%. In the second sample of modified sludge ADK2, the concentration of limy ions was approximately a third lower than in the first sample. In the third sample ADK3 the concentration of limy ions was approximately three times lower than in the sample ADK1. The free CaO and total carbonate content analysis results are stated in the Table 3.

Based on this information, recipes for mixtures for lab burning tests were suggested. An overview of suggested recipes is stated in Table 5. In Figure 1 prepared sludge marked as ADK2 is shown.

Table 1: Receipt of sludge modified for addition.

Sample	Sludge [% weight]	Limestone [% weight]	Dry Hydrate [% weight]
ADK1	62	-	38
ADK2	53	40	7
ADK3	50	50	-

Table 2: Basic technological analysis of sludge modified for addition.

Sample descr.	W^a	W_t^r	A^r	A^d	Q_s^d	Q_s^{daf}	Q_i^r	Q_i^d	Q_i^{daf}	S^d
	%	%	%	%	MJ/kg	MJ/kg	MJ/kg	MJ/kg	MJ/kg	%
ADK 1	5,09	19,65	50,68	62,84	4,62	12,42	2,89	4,23	11,39	0,46
ADK 2	3,19	26,47	39,46	53,67	3,89	8,38	2,02	3,64	7,86	0,36
ADK 3	3,26	28,30	38,12	53,16	3,36	7,17	1,56	3,16	6,76	0,44

Table 3: Properties of modified sludge for addition.

Sample	C^{daf}	H^{daf}	N^{daf}	O^{daf}	S^{daf}	CO_2	CaO free in solid	Carbonates in solid
	%	%	%	%	%	%	%	%
ADK 1	31,81	4,74	2,74	59,47	1,24	<0,01	36,88	1,62
ADK 2	36,00	2,38	2,22	58,62	0,78	<0,01	6,08	22,30
ADK 3	34,48	1,90	2,16	60,52	0,94	3,19	4,01	10,00

Figure 1: Prepared sample ADK2.

3 Laboratory verification of effectiveness of new additives

It is possible to verify the effectiveness of desulphurization additives by practical tests of coal burning in an industrial facility for which the additive is specified. Such tests are quite expensive and often also require careful management. In order to compare the effectiveness of different additives and to check the optimum amount of the additive to achieve the necessary grade of desulphurization, it is possible to use a model burning test using a smaller sample amount. The efficiency of the additive at a given concentration can easily be determined by finding out the sulphur content in the ash produced by burning the sample under certain conditions. In this way the additives can be compared from an efficiency point of view, but their optimum consumption can also be guessed (i.e. optimum ratio of the additive and coal of certain quality). If the laboratory conditions are similar to real conditions, especially the temperature at which the processes are taking place, the results of desulphurization can be used for the prediction of additive behavior under real conditions. A similar principle of desulphurization condition modeling was successfully applied to compare the desulphurization of various types of additives in the past.

3.1 Sample preparation for verification

First of all homogenization and adjustment of all input materials was made for additive adding to a granularity of less than 0.2mm. For modified samples,

parameters were set which are needed to determine the efficiency according to the methodology of accredited labs of the Brown Coal Research Institute, j.s.c. The results of analyses made are shown in the Table 4.

Table 4: Analysis results of samples for lab burning tests.

Sample	W^a	W_t^r	A^r	A^d	Q_s^d	Q_s^{daf}	Q_i^r	Q_i^d	Q_i^{daf}	S^d
	%	%	%	%	MJ/kg	MJ/kg	MJ/kg	MJ/kg	MJ/kg	%
M-1	6,16	23,82	18,67	25,38	23,05	30,89	15,59	22,09	29,60	1,43
M-2	6,13	26,98	17,03	23,32	23,66	30,86	15,81	22,57	29,44	1,29
M-3	6,02	27,38	16,91	23,29	23,80	31,03	15,82	22,71	29,61	1,39
M-4	4,69	25,56	19,99	26,86	21,61	29,54	14,77	20,70	28,30	1,41
M-5	5,84	26,17	18,20	24,65	22,86	30,34	15,52	21,90	29,07	1,45
M-6	4,79	26,87	16,75	22,91	23,74	30,79	16,62	22,72	29,47	1,51
M-7	4,34	23,23	21,78	28,37	19,81	27,66	13,98	18,96	26,47	1,11
M-8	4,27	23,98	20,98	27,60	20,46	28,25	14,30	19,60	27,07	1,25
M-9	4,30	24,62	19,97	26,49	21,31	28,98	14,75	20,38	27,73	1,25
M-10	4,35	26,38	18,05	24,52	22,69	30,06	15,99	21,72	28,77	1,29
M-11	4,58	27,16	16,92	23,23	23,59	30,73	15,85	22,70	29,56	1,28
M-12	4,62	27,43	16,37	22,56	23,91	30,88	15,90	22,86	29,52	1,35
M-13	4,63	29,28	11,63	16,44	26,32	31,50	17,11	25,22	30,19	1,44

Based on the comparison of information about additive used so far and previously tried additive which is limestone from Čížkovice and information about the modified sludge from wastewater treatment plant, recipes of samples for lab burning tests were suggested. With every newly-checked additive (ADK1 to ADK3), three samples were always prepared, each with a different ratio of additive to coal. For comparison three samples were prepared with previously-tried additive, i.e. Čížkovice limestone. For comparison, another coal sample without additive was prepared. For every recipe, ash fusibility was determined for control in order to ascertain the influence of the additive amount to possible sintering of additive samples. Recipes of the samples are again shown in Table 5.

3.2 Lab burning tests

The prepared samples were analyzed and model lab burning of analytic sample with known input chemical composition was done at a temperature of approximately 850°C. This temperature is achieved when burning coal in a fluid boiler.

The samples with additives which were prepared according to recipes were burnt in a muffle furnace under controlled conditions at a temperature of 850°C until it attained a constant weight. The partial samples with additive were at first dried in an oven at 105 – 110°C until they attained a constant weight. The dried samples with additive were annealed in a muffle furnace at 850°C with gradient rise of temperature and time pause (30 min.) at 700°C. Total annealing time of samples with additive was 10 hours. After annealing and ash cooling-down, a visual check of unburned particles was made. At the same time, the loss caused

by annealing of every sample with additive was determined. In the ash, the total content of sulphur S_t^d by method Eschka acc. to ČSN 44 1379 was determined. These are the results of theoretical desulphurization level.

4 Desulphurization efficiency calculation procedure

If there is a reaction as a result of additive being checked with sulphur dioxide being created, calcium sulphate is produced, which is bound in the ash. From sulphur concentration in ash produced by burning samples and other information, it is possible to determine how many per cent of sulphur were released into the air and how much was bound into the ash. This shows desulphurization efficiency of additive under set conditions. If no sulphur was bound into the ash, the efficiency is zero and all sulphur released by burning comes into the combustion gas. If, on the contrary, all the sulphur stays in the ash, no sulphur dioxide would be released into the air and desulphurization efficiency would be 100%.

The calculated values of desulphurization efficiency are shown in Table 5. From the desulphurization efficiency values, it is obvious that all used additives desulphurize effectively in the suitably-selected dosing ratio under set lab conditions. In spite of the fact that certain efficiency variances were observed, it can be said that the desulphurization efficiency achieved at maximum used additive was about 64% in one case and in other cases equal or higher than 68% which means that about two-thirds of the created amount of sulphur dioxide can be bound by the added additive into the ash. In case of burning the sample of pure coal a part of sulphur (approximately 4%) combined with alkalic components (CaO, MgO) contained in the ash (so called self- desulphurization).

Table 5: Desulphurization efficiency found out at lab tests.

Sample	**Weight part [Coal Hp2] [%]**	**Type of additive**	**Weight part of additive [%]**	**Desulphurization efficiency [%]**
M-1	100	ADK1	10	68,0
M-2	100	ADK1	8	85,0
M-3	100	ADK1	6	69,0
M-4	100	ADK2	15	63,9
M-5	100	ADK2	12,5	69,7
M-6	100	ADK2	10	59,9
M-7	100	ADK3	30	71,4
M-8	100	ADK3	25	55,4
M-9	100	ADK3	20	65,9
M-10	100	ČÍŽK. VÁP	10	73,2
M-11	100	ČÍŽK. VÁP	8	70,6
M-12	100	ČÍŽK. VÁP	6	61,1
M-13	100	ŽÁDNÉ	0	4,1

5 Conclusion

In terms of this study there was laboratory checked activity of new desulphurization additives. The new additives were prepared based on the sludge from wastewater treatment plants stabilized by adding components containing limy ions (dry hydrate or limestone). After stabilizing the sludge shows heating power in original state according to a recipe approximately 1.5 to 2.9 MJ/kg. In terms of lab burning tests, the desulphurization activity of the additives is comparable today with Čížkovice limestone. It can be assumed that to substitute the amount of this limestone for desulphurization completely at fluid burning, it would be possible to use the same amount of additive ADK1 or approximately for 50% more additive ADK2 or three times the amount of ADK3. Sludges from wastewater treatment plants are often placed in depots. The results of this study show that after stabilizing sludge by adding limy ions, there is the possibility that such modified sludge can also be used for industrial use as substitute for limestone that is currently used at fluid burning of brown coal. The implementation of the results of this study makes it possible in reality to use potential waste, which would otherwise be thrown to a depot. At the same time, the energy contained in the sludge of a wastewater treatment plant would be used as well. This fulfills the requirement of these days – that is to say, using waste for making energy rather than throwing it in a depot.

Acknowledgements

This study was made in terms of research work of Research Center 1M06007 “Research center of integrated system of using auxiliary products of mining, modification and processing energy base materials” in support of and contribution from the Ministry of Education, Youth and Physical Education of the Czech Republic.

References

[1] Sedláček, P., Fečko, P., Valeš, J., Čablík, V., Martinek, T.: Verification of theoretical calculation for reduction of SO_2 emissions upon production of brown-coal pellets, Sborník: 7th Conference of Environment and Mineral Processing, VŠB-TU Ostrava 26.6.-28.6.2003, s.355-358, ISBN 80-248-0248-1

[2] Valeš, J., Sedláček, P.: Zajištění, provedení a vyhodnocení spalovacích zkoušek aditivovaného paliva (hp, tříděné druhy) na vybraných průmyslových spotřebičích (Děčín, Bezdružice) – TR8 050, Technická zpráva VÚHU a.s., 1999

[3] Sedláček, P., Valeš, J., Šafářová, M., Riedlová, S.: Potential production of ecological pellets with biomass or wastes, 5. European Conference Coal Research and its Applications, 6th - 8th September 2004, Edinburgh, Scotland

[4] Roubíček, V., Buchtele, J.: Uhlí - zdroje, procesy, užití, Ostrava 2002, Sborník: 6th Conference on Environment and Mineral Processing, part II, s.797-802, ISBN 80-248-0072-1 (Coal-sources, Processes, Utilisation, Ostrava 2002, Collection)
[5] Vidlář J., Hodek, O.: Základy úpravy užitkových surovin, I. A II. Díl, VŠB Ostrava, Most 1984

The use of the Mock-up-Cz physical model in the design of engineered barriers

J. Pacovský
Centre of Experimental Geotechnics,
Czech Technical University in Prague, Czech Republic

Abstract

A fundamental challenge concerning research into high-level radioactive waste disposal is the construction design of natural and engineered barriers capable of preventing the leakage of hazardous radionuclides over a period of roughly 100 thousand years. In order to obtain the information enabling the design of such a construction, it is necessary to employ all the experimental tools and procedures available including physical modelling. The most relevant model types for this task have been found to be those made at a scale of 1 : 1, referred to as mock-up models. Information on the construction, operation and dismantling of a vertical model of a bench-scale buffer mass test of Czech smectitic clay employing the KBS-3V system modification forms the main subject matter of this paper.
Keywords: *radioactive waste, engineered barrier, bentonite block, Mock-up-Cz experiment, dismantling procedures.*

1 Introduction

Solving the problem of the safe isolation of radioactive waste requires a multidisciplinary approach. It is generally accepted that underground repository waste isolation must be assured by a multi-barrier system consisting of the container with the high level radioactive waste itself, an engineered barrier - currently intended to be based on bentonites - and a natural barrier formed by a suitable rock environment.

As part of the research into engineered barriers based on highly compacted bentonite blocks, a physical model called Mock-up-Cz, was built at the Centre of Experimental Geotechnics, CTU in Prague from 2000 to 2002, and operated from May 2002 to January 2006. In February 2006, dismantling commenced

www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RAV060811

which lasted a total of 47 days. The multidisciplinary research of all the samples taken from the bentonite barrier has been underway since March 2006. A full evaluation of the experiment will be available by the end of 2007.

2 Design and construction of Mock-up-Cz

Mock-up-Cz took the form of a vertical model of a bench-scale buffer mass test of Czech smectitic clay. It was a long-term experiment which aimed to model the behaviour of a sealing barrier composed of highly compacted blocks.

2.1 Basic description

The model was placed in an underground test silo with dimensions of 3 x 3 x 3m. The Mock-up-Cz model itself consisted of a steel tank or tube with an inside diameter (d) of 800mm, a height of 2240mm and a wall thickness of 8mm. The system was designed to withstand an internal pressure of up to 5MPa, fig. 1.

Figure 1: Mock-up-Cz experiment.

The vessel containing high-level radioactive waste was simulated by using a heater placed in an experimental bin. The heating mode was fully automated, and the temperature range could be set between 20°C and 300°C.

2.2 Buffer material

Based on an evaluation of research into Czech bentonite, the most appropriate buffer material was found to be a mixture of treated bentonite (from the Rokle deposit), silica sand and graphite [1]. Sand was added in order to facilitate the control of the swelling abilities and plasticity of the mixture. Graphite was added to the mixture to increase its thermal conductivity so that the heat produced by the container was more easily redistributed towards the natural barrier. The basic mixture contained 85% bentonite, 10% silica sand and 5% graphite.

2.3 Hydration system

The potential inflow of granitic water from the natural barrier was simulated by using a flooding system installed on the inner side of the bin featuring two concentric filters, the one closer to the bin being coarse-meshed and the other having voids of less than 100μm. Water was let in through four vertical perforated tubes. A hydration system which allowed an increase in hydration pressure of up to 1MPa was placed outside the experimental silo. Synthetic granitic water was used for saturation; the amount of water and the saturation rate were continuously recorded.

2.4 Model instrumentation

The Mock-up physical model was equipped with 52 thermometers, 50 hydraulic pressure cells, 37 humidity sensors of varying construction and 20 resistive tensometers. Temperature, swelling pressure and hydration measurements were taken continuously at six measurement profiles, fig. 2.

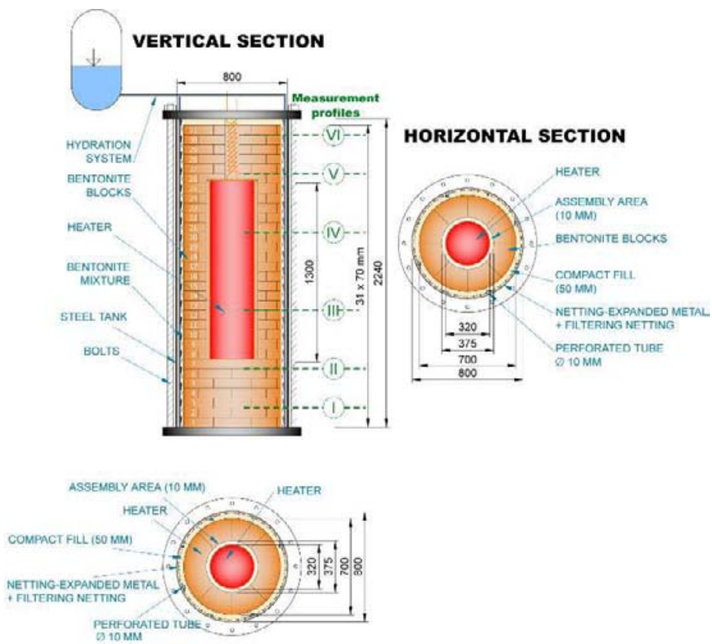


Figure 2: Vertical and horizontal scheme of Mock-up-Cz.

The whole system of registration, evaluation and transfer of data was based on the use of a small portable AD-SYS data logger connected to a CEG server.

The measurement of changes in moisture content using sensors was the most difficult and financially the most demanding of all the basic measurements of the various geotechnical parameters. Due to the notoriously inconsistent nature

of the results of moisture content measurement provided by sensors, an alternative method for monitoring on-going saturation was sought from the very launch of the experiment. This resulted in a decision to combine automated measurement using sensors with regular core sampling (drilling) used for the gravimetric evaluation of the moisture content.

3 Experimental procedures

The Mock-up-Cz was designed as a long-term experiment. Following consultation with foreign experts, the experiment was divided into 3 operational phases followed by the dismantling phase and final post mortem analysis [2].

3.1 Phase 1

The filter was kept dry and the power switched on to reach a maximum temperature within the bentonite of 90±5°C. Temperature, swelling pressure and hydration recordings were taken continuously for 6 months. Buffer evolution in Phase 1 consisted of a redistribution of initial pore water as a result of the thermal gradient across the buffer and the homogenisation and subsequent consolidation of the filling mixture due to the swelling pressure exerted by the hydrating and expanding blocks.

3.2 Phase 2

Power to the heater was maintained at a constant level and the filter filled with synthetic granitic water by connecting two of the perforated pipes to a water reservoir and using the other two for air release. Temperature, swelling pressure and hydration recordings were taken continuously as the buffer evolved towards a state of full water saturation.

Core samples were taken at pre-set intervals throughout this phase in order to physically determine saturation progress and the effect of the various processes at work on the permeability and swelling capacity of the bentonite mixture.

3.3 Phase 3

Saturation and temperature loading of the buffer by the heater were stopped during this phase. When the heater was switched off, the cooling phase commenced which lasted approximately 2 weeks.

3.4 Dismantling and post-mortem analysis

The dismantling process was conducted according to a detailed timetable and included the following:

1 scientific programme (tests and analysis on samples)
2 sampling plan (location and specific requirements of each sample)
3 scenario (guidelines and procedures for dismantling and sampling operations, work organisation as well as records and documents)

Detailed tests focused on the filling material, corrosion and instrumentation.

4 Selected results from the experiment

4.1 Phase 1

Loading the barrier with temperature (before the introduction of external water) caused a redistribution of the original moisture content. In the vicinity of the heater, the moisture content dropped from the original 8-10% to 2%, while, conversely, in those areas more remote from the heater, it rose to a maximum of 22%. Exact moisture content distribution on completion of the redistribution process was determined by analysing the moisture content of core samples.
Together with a redistribution in moisture content, a new equilibrium temperature gradient was created.

4.2 Phase 2

After the commencement of the saturation process, swelling pressure inside the experimental bin gradually increased. The swelling bentonite blocks gradually filled in all the empty spaces within the bentonite barrier (joints and technological gaps). This led to a reduction in the original volume density of the highly compacted blocks thus accounting for the unexpectedly low swelling pressure results at this stage in the experiment. Maximum measured swelling pressure reached a value of approximately 2MPa as compared to the input swelling ability of the blocks of 4MPa, fig. 3.

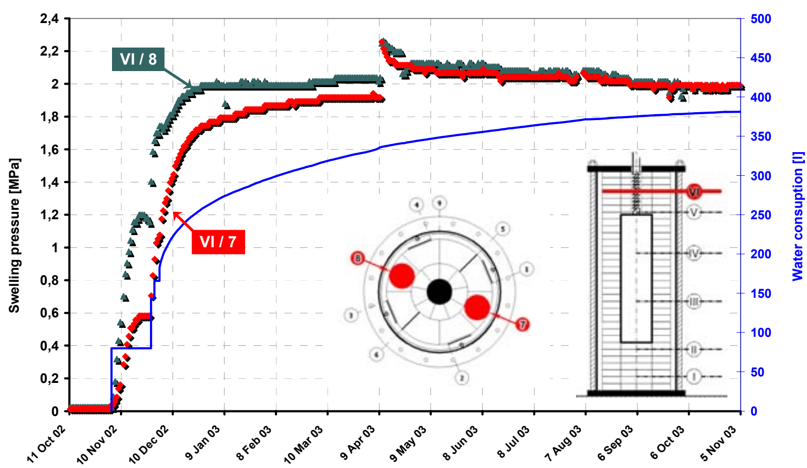


Figure 3: Maximum measured swelling pressure.

On-going saturation increased the barrier's thermal conductivity so that the maximum temperature within the barrier fell (from 95°C to 85°C) even though heater output remained constant.

Barrier sampling, which took place at regular intervals (by means of core drilling), allowed the precise determination of the saturation process. An analysis

of samples subsequently revealed that practically complete saturation of the barrier (S_r = 0.9) was reached 780 days after the initiation of the saturation process, fig. 4 [3]. Phase 2 continued until the end of 2005.

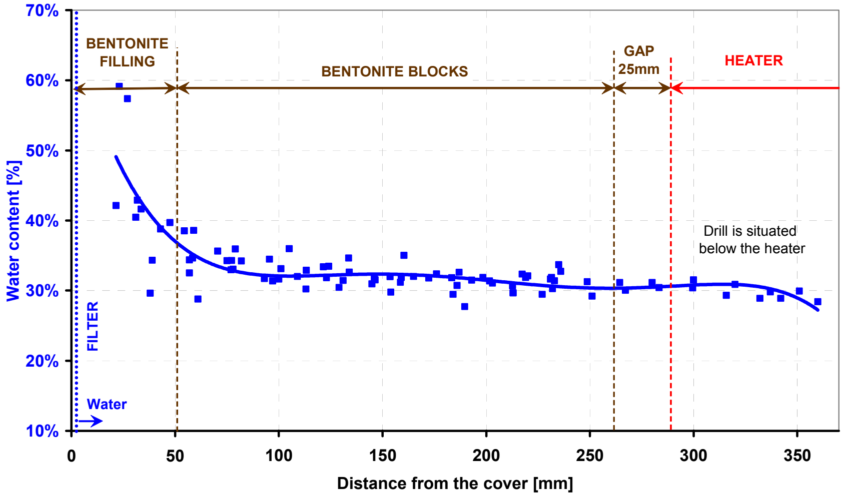


Figure 4: Water content redistribution after 24 months saturation - Oct 2004.

4.3 Phase 3

The experimental heater was switched off on 3.1.2006, and the saturation medium inlet was left open for the entire cooling period. Measurements of temperature decrease and changes in swelling pressure were taken continually during Phase 3.

5 The Dismantling project

The originally mono-disciplinary (geotechnically oriented) Mock-up-Cz experiment was more successful than any of those involved dared hope and was set to provide a wealth of important data and experience. Consequently, at the beginning of 2005, it was agreed with the principal sponsor of the experiment (RAWRA - The Czech Radioactive Waste Repository Authority) that the final part of the experiment, i.e. dismantling and post-mortem analysis, would be carried out on a multidisciplinary basis. It was agreed that in addition to the CEG CTU in Prague, the final phases would be performed with the participation of the Institute of Chemical Technology in Prague, the Faculty of Science - Charles University in Prague and Masaryk University in Brno. Geotechnical research was extended to include geochemical, corrosion, bacteriological and microscopic research.

In conjunction with these partners, CEG designed a detailed dismantling project which included a sampling plan which took into account the various individual research areas and a detailed time schedule.

The dismantling process itself commenced on 30th January 2006 with the opening of the experimental bin. The full dismantling of the bentonite barrier, the taking of samples and the extraction of the heater, measuring gauges and cabling was completed on 17th March 2006.

6 Selected results from post-mortem analysis

Extensive multidisciplinary research commenced during the dismantling stage the results of which will provide information on the rheological stability of an engineered barrier consisting of highly compacted blocks moulded from a mixture based on Czech bentonite (Rokle deposit).

The results of liquid limit tests on the barrier material and a full evaluation of changes in volume densities are just two of the interesting sets of results obtained to date.

6.1 Changes in liquid limit w_L of barrier material

One of the most significant characteristics of the bentonite-based barrier material is its high plasticity. In soil mechanics, material plasticity (plastic material behaviour) is defined as the range of moisture content values between the liquid limit w_L and the plasticity limit w_P (Atterberg limits). The distance between these limits specifies by how many percent the moisture content of the material at the plasticity limit may be increased so that it will still behave as a plastic material. It generally holds true that the higher the plasticity index of a material, the more suitable it is for use in an engineered barrier. This material remains plastic even after a considerable increase in moisture content, and is easily able to fill all the discontinuity faults (gaps, joint, cracks).

Research of the rheological stability (Atterberg limits) of the Mock-up-Cz experiment was aimed at verifying whether the long-term thermal stress from the heater acting on the barrier, accompanied by barrier saturation, causes material degradation and thus a decrease in plasticity signalled primarily by a decrease in the liquid limit w_L.

Fig. 5 shows the decrease in the liquid limit in relation to long-term temperature effects. Individual sections of the bentonite barrier were loaded with various temperatures depending on the distance from the heater for a period of over 3 years. The effect of a temperature of 77°C caused a decrease in the liquid limit from the original 160% to 110%. As the plasticity limit tends to show only minimum variations, the decrease in plasticity (plasticity index) can be stated as being approximately 50%. It will require the results of additional tests (permeability, swelling ability, mineralogy etc.) to show whether extensive material degradation has occured in this case.

6.2 Changes in the input dry density ρ_d of the barrier material

A further important characteristic of the barrier material is its swelling ability. In contact with water, bentonite swells, filling up any gaps and sealing discontinuity faults. The commonly accepted swelling pressure value required of barrier

material ranges from 5 to 10MPa. In order to obtain such high swelling ability, natural bentonite must be compacted into so-called prefabricates. The dry density of compacted prefabricates ranges from 1700 - 1800kg/m^3 (the dry density of the natural material is around 1200kg/m^3). The swelling ability of bentonites, therefore, largely depends on their dry density, fig. 6.

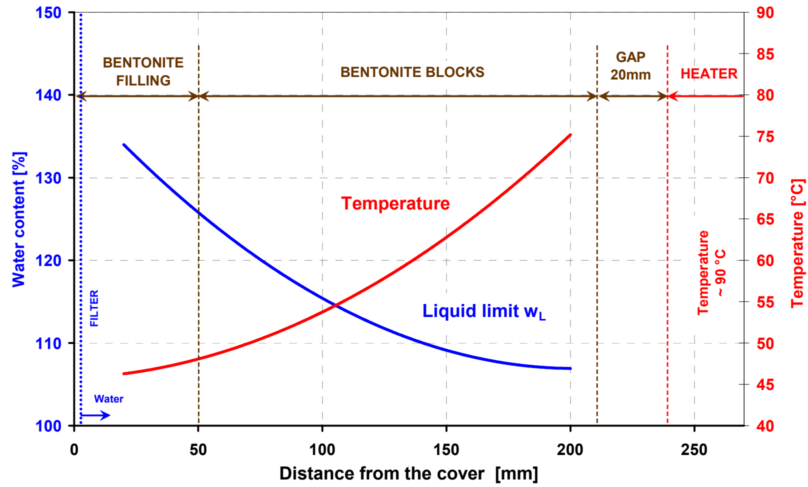


Figure 5: Decrease in liquid limit due to long-term heating.

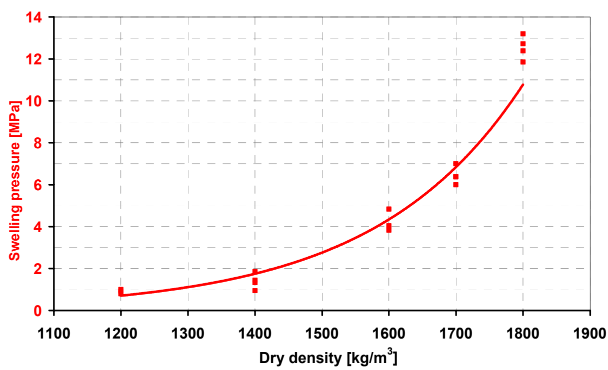


Figure 6: Swelling ability of natural Rokle bentonite in relation to dry density.

Unless compacted bentonite (highly compacted block) is allowed to expand into empty spaces, it exerts a maximum swelling pressure of 5 to 10MPa. All the experimental barriers to date, however, have, to a greater or lesser extent, contained construction joints between the prefabricates, technological gaps for operating the heater etc. The smaller the number of discontinuity faults, the higher the swelling pressure exerted by the prefabricate and the higher the pressure by which existing joints, cracks etc. are sealed.

Prefabricates for the Mock-up-Cz experiment were designed so that their swelling pressure would not exceed 5MPa (for reasons of the safety of the experimental bin). The maximum swelling pressure measured, however, only slightly exceeded 2MPa. Furthermore, the swelling pressure in the experimental bin of the Mock-up-Cz experiment displayed a highly non-uniform distribution pattern, even after reaching maximum saturation level.

The very first dry density test results showed considerable decreases in the original dry density; this corresponded to the lower swelling pressure values measured, fig. 7.

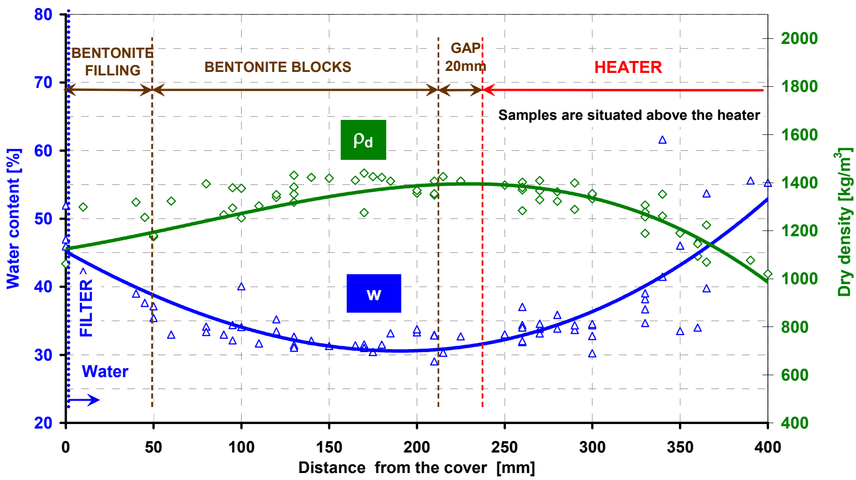


Figure 7: Changes in dry density along the barrier profile after its saturation.

7 Conclusion

The Czech Republic joined international research on methods of isolating radioactive waste only after 1990. Thanks to assistance particularly from the International Atomic Energy Agency in Vienna and the European Community, Czech institutions have been asked to participate in numerous international projects through which they have been able to familiarise themselves with the research results from experiments carried out at various foreign institutions. The Mock-up-Cz experiment benefited greatly from the assistance of foreign, particularly Swedish and Belgian, experts (SKB Sweden, SCK.CEN Belgium).

Mock-up-Cz is the first large-scale Czech experiment aimed at studying the behaviour of bentonite-based engineered barriers. The original intention was to obtain new experience in this highly specialised research field and create a team of young Czech experts. Results from both the construction and performance of the experiment, however, have exceeded all expectations. Consequently, a decision was made in 2005 to include other research areas in this initially mono-disciplinary geotechnical experiment. Further, the experiment, which has been

highly praised by foreign institutions, should result in a confirmation of the suitability of Czech bentonites for the isolation of radioactive waste. A full evaluation of all the stages of the Mock-up-Cz experiment will be completed by the end of 2007.

Acknowledgement

The experiment could not have taken place without the support of the CTU - Faculty of Civil Engineering, Prague - CIDEAS and RAWRA.

References

[1] Pacovský, J., Selected Results from Geotechnical Research on Bentonite. *Proc. of the 8th Int. Conference on Radioactive Waste Management and Environmental Remediation,* eds. R. Trávníček & R. Vašíček, The American Society of Mechanical Engineers ASME, Bruge, session 44, pp.1 - 5, 2001.

[2] Pacovský, J., Experimental Research on Engineered Barriers. *Proc. of 4th Int. Symposium on Environmental Aspects in Civil Engineering Education,* Formacao Contínua e Publicacoes do Departamento de Engenharia Civil da FEUP, Porto, pp.14-20, 2002.

[3] Pacovský, J., Saturation Development in the Bentonite Barrier of the Mock-up-Cz Geotechnical Experiment. *2nd Int. Meeting Clays in Natural & Engineered Barriers for Radioactive Waste Confinement,* eds. J. Svoboda & L. Zapletal, Andra Agence nationale pour la gestion des dechets radioactifs, Tours, pp.69 - 70, 2005.

Author Index