Studies in Computational Intelligence 348

Beniamino Murgante
Giuseppe Borruso
Alessandra Lapucci (Eds.)

Geocomputation, Sustainability and Environmental Planning

Springer

Studies in Computational Intelligence, Volume 348

Further volumes of this series can be found on our homepage: springer.com

Vol. 326. Sheryl Brahnam and Lakhmi C. Jain (Eds.)
Advanced Computational Intelligence Paradigms in Healthcare 5, 2010
ISBN 978-3-642-16094-3

Vol. 327. Slawomir Wiak and Ewa Napieralska-Juszczak (Eds.)
Computational Methods for the Innovative Design of Electrical Devices, 2010
ISBN 978-3-642-16224-4

Vol. 328. Raoul Huys and Viktor K. Jirsa (Eds.)
Nonlinear Dynamics in Human Behavior, 2010
ISBN 978-3-642-16261-9

Vol. 329. Santi Caballé, Fatos Xhafa, and Ajith Abraham (Eds.)
Intelligent Networking, Collaborative Systems and Applications, 2010
ISBN 978-3-642-16792-8

Vol. 330. Steffen Rendle
Context-Aware Ranking with Factorization Models, 2010
ISBN 978-3-642-16897-0

Vol. 331. Athena Vakali and Lakhmi C. Jain (Eds.)
New Directions in Web Data Management 1, 2011
ISBN 978-3-642-17550-3

Vol. 332. Jianguo Zhang, Ling Shao, Lei Zhang, and Graeme A. Jones (Eds.)
Intelligent Video Event Analysis and Understanding, 2011
ISBN 978-3-642-17553-4

Vol. 333. Fedja Hadzic, Henry Tan, and Tharam S. Dillon
Mining of Data with Complex Structures, 2011
ISBN 978-3-642-17556-5

Vol. 334. Álvaro Herrero and Emilio Corchado (Eds.)
Mobile Hybrid Intrusion Detection, 2011
ISBN 978-3-642-18298-3

Vol. 335. Radomir S. Stankovic and Radomir S. Stankovic
From Boolean Logic to Switching Circuits and Automata, 2011
ISBN 978-3-642-11681-0

Vol. 336. Paolo Remagnino, Dorothy N. Monekosso, and Lakhmi C. Jain (Eds.)
Innovations in Defence Support Systems – 3, 2011
ISBN 978-3-642-18277-8

Vol. 337. Sheryl Brahnam and Lakhmi C. Jain (Eds.)
Advanced Computational Intelligence Paradigms in Healthcare 6, 2011
ISBN 978-3-642-17823-8

Vol. 338. Lakhmi C. Jain, Eugene V. Aidman, and Canicious Abeynayake (Eds.)
Innovations in Defence Support Systems – 2, 2011
ISBN 978-3-642-17763-7

Vol. 339. Halina Kwasnicka, Lakhmi C. Jain (Eds.)
Innovations in Intelligent Image Analysis, 2010
ISBN 978-3-642-17933-4

Vol. 340. Heinrich Hussmann, Gerrit Meixner, and Detlef Zuehlke (Eds.)
Model-Driven Development of Advanced User Interfaces, 2011
ISBN 978-3-642-14561-2

Vol. 341. Stéphane Doncieux, Nicolas Bredeche, and Jean-Baptiste Mouret(Eds.)
New Horizons in Evolutionary Robotics, 2011
ISBN 978-3-642-18271-6

Vol. 342. Federico Montesino Pouzols, Diego R. Lopez, and Angel Barriga Barros
Mining and Control of Network Traffic by Computational Intelligence, 2011
ISBN 978-3-642-18083-5

Vol. 343. Kurosh Madani, António Dourado Correia, Agostinho Rosa, and Joaquim Filipe (Eds.)
Computational Intelligence, 2011
ISBN 978-3-642-20205-6

Vol. 344. Atilla Elçi, Mamadou Tadiou Koné, and Mehmet A. Orgun (Eds.)
Semantic Agent Systems, 2011
ISBN 978-3-642-18307-2

Vol. 345. Shi Yu, Léon-Charles Tranchevent, Bart De Moor, and Yves Moreau
Kernel-based Data Fusion for Machine Learning, 2011
ISBN 978-3-642-19405-4

Vol. 346. Weisi Lin, Dacheng Tao, Janusz Kacprzyk, Zhu Li, Ebroul Izquierdo, and Haohong Wang (Eds.)
Multimedia Analysis, Processing and Communications, 2011
ISBN 978-3-642-19550-1

Vol. 347. Sven Helmer, Alexandra Poulovassilis, and Fatos Xhafa
Reasoning in Event-Based Distributed Systems, 2011
ISBN 978-3-642-19723-9

Vol. 348. Beniamino Murgante, Giuseppe Borruso and Alessandra Lapucci (Eds.)
Geocomputation, Sustainability and Environmental Planning, 2011
ISBN 978-3-642-19732-1

Beniamino Murgante, Giuseppe Borruso and
Alessandra Lapucci (Eds.)

Geocomputation, Sustainability and Environmental Planning

Springer

Beniamino Murgante
Department of Architecture
Urban and Regional Planning
Transport Planning and Infrastructure
Laboratory of Urban and Territorial Systems
University of Basilicata
Viale dell 'Ateneo Lucano, 10
85100 - Potenza
Italy
E-mail: beniamino.murgante@unibas.it

Giuseppe Borruso
Department of Education Science and
Cultural Processes, University of Trieste
P.le Europa, 1
34127 Trieste
Italy
E-mail: giuseppe.borruso@econ.units.it

Alessandra Lapucci
Department of Civil Engineering
Laboratory of Territorial and Environmental
Systems Engineering
University of Pisa
Via Diotisalvi, 2
56126 Pisa
Italy
E-mail: alessandra.lapucci@ing.unipi.it

ISBN 978-3-642-19732-1 e-ISBN 978-3-642-19733-8

DOI 10.1007/978-3-642-19733-8

Studies in Computational Intelligence ISSN 1860-949X

Library of Congress Control Number: 2011923798

Typeset & *Cover Design:* Scientific Publishing Services Pvt. Ltd., Chennai, India.

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

Contents

Geocomputation, Sustainability and Environmental Planning

Sustainable Development: Concepts and Methods for Its Application in Urban and Environmental Planning

Beniamino Murgante[1], Giuseppe Borruso[2], and Alessandra Lapucci[3]

[1] University of Basilicata, Viale dell'Ateneo Lucano, 10 - 85100 Potenza, Italy
e-mail: beniamino.murgante@unibas.it
[2] University of Trieste, P.le Europa, 1 - 34127 Trieste, Italy
e-mail: giuseppe.borruso@econ.units.it
[3] University of Pisa, Via Diotisalvi, 2 -56126 Pisa, Italy
e-mail: alessandra.lapucci@ing.unipi.it

1 Sustainability: From Principles to Evaluation Methods

The idea of *sustainable development* may appear quite vague, fuzzy and evasive (Pearce et al. 1989). In fact, whereas sustainability is related to a status of maintenance and conservation of the existing conditions, both in space and time and is referred to the capacity to guarantee a support without causing decay, the concept of development implies, instead, an alteration and a transformation of actual status, then a condition of instability.

This semantic conflict induces to an idea of both improvement and preservation: in substance, the effective aim of a sustainable development is the possibility to guarantee a better life quality for an enduring period of time.

The Bruntland report (1987) systematized the definition of environmental sustainability even on a political level:

> "Sustainable development is development that meets the needs of the present without compromising the ability of future generations to meet their own needs. It contains within it two key concepts: the concept of needs, in particular the essential needs of the world's poor, to which overriding priority should be given; and the idea of limitations imposed by the state of technology and social organization on the environment's ability to meet present and future needs".

The definition above reported implies a series of concepts connected to that of Sustainable development. Two key elements in particular are highlighted, these being intergenerational and intragenerational equity. The first one refers to the need to manage present resources in order to allow future generations to meet their needs, while the second one aims at reducing differences in resources allocation between people in a same timeframe, thus recalling the need to tackle issues

B. Murgante et al. (Eds.): Geocomputation, Sustainability & Environ. Planning, SCI 348, pp. 1–15.
springerlink.com

referred to differences in development between industrialized and developing countries. A third element is time, as sustainability involves a care of the future and therefore to plan an evolutionary path of development, that therefore should be inserted in models.

Without recalling the overall evolution that brought to implement the concept of sustainable development, it is worth recalling the work carried on by Meadows et al (1972), reporting to The Club of Rome to investigate major trends of global concern at the beginning of '70s of the past century. The matter is not trivial, as we can notice a series of key points still important in to-date research. One key point is referred to modeling, as the work carried out was one of the first to model different variable connected to natural and human resources, although in a simplified manner. A second point referred to evolution over time, therefore considering the time dimension as important in designing scenarios. A third element is related to the critiques that such model brought over itself, that in a sense recalled the general criticism in the following years on quantitative measures as elements to provide discrete solutions to real problems.

'The limits to growth' was focused on a set of trends as accelerating industrialization, rapid population growth, widespread malnutrition, depletion of nonrenewable resources and a deteriorating environment. Authors used World3 model to simulate interactions between Earth humans as systems and tried to focus on possibilities to implement 'sustainable' – although this term was not used - actions that would alter growth trends among the variables considered. Although the major aim of the work carried out was to analyze the interactions of exponential growth with a finite set of resources and not to predict actual evolutions, criticism arose since the beginning from different domains.

Criticism focused on different aspects of the work, namely on the base of data used, considered weak, as well as a not clear procedure, as the details about the model World3 and assumption were made clear in 1974. Also, critiques were put on the lack of consideration of technological changes in the evolutionary model, although the same authors stated that the aim of the simplified model was to study possible interactions using a limited set of variables and that technology was not considered, and on the mainly Malthusian assumption of the different paces in growth of population and resources.

The work was re-elaborated in 1992 and 2004 – respectively 20 and 30 years after the original research – in order to consider in a more refined way the growing importance of concepts related to sustainable development, and particularity income distribution and intergenerational exchange, offering also alternatives, originally not considered in the model, as human creativity for improving quality of life and elements as energetic efficiency, recycling and rise in average endurance of human life.

In any case the importance of the research carried out lies on the attention it brought over the finiteness of resources and on environmental issues. The debate started in the years '70s of Twentieth century leads to the definition of Sustainable Development as well as to events as the Rio Conference and can be also rooted in such research on global and models of interactions, than at least in this sense played an important role. Also, 'Limits to growth' can be read to-date as a call

both for a use of quantitative modeling to elaborate complex and interactive amount of data and for a wider debate on one side between different disciplines and on the other side between quantitative and qualitative research and decision making processes. The Bruntland report, above all, then the Rio Conference and the Instambul Habitat II Conference definitely underlined the importance of social and economical dimension of environmental sustainability (O.N.U. 1993 1996). Since then the concept of sustainable development goes transversely across all scientific, social and economic disciplines and becomes the main objective of modern market economies. The necessity to include, in decision-making process, all the concerns connected with *environmental resources exploitation* emerges in all different sectors of public policies.

With respect to *ecological protection* demands, the conditions of sustainability are at first connected with the requirement that natural resources stock does not decrease in time (Odum 1989, Pearce 1990). Subsequently, a new concept of sustainability comes forward which underlines the importance of inserting environmental criteria in territorial development choices, rather than supporting specific protection policies, indicating boundary thresholds to resources use and consumption. In accordance with this idea, all decisions about new government interventions are taken with the respect of environment "carrying capacity" (Costanza 1991, Daly and Cobb 1990, Nijkamp 1990 1994, Nijkamp and Archibugi 1989, Pearce 1991).

The *economic efficiency* criterion in resources exploitation, as well as the *social dimension* of sustainability, become relevant issues at the same time and a special emphasis is given to social justness, even extended to an intergenerational perspective; low-income classes' protection and care represent the core of sustainable development subjects. (Sen 1992, Serageldin et al. 1995, Zamagni 1994).

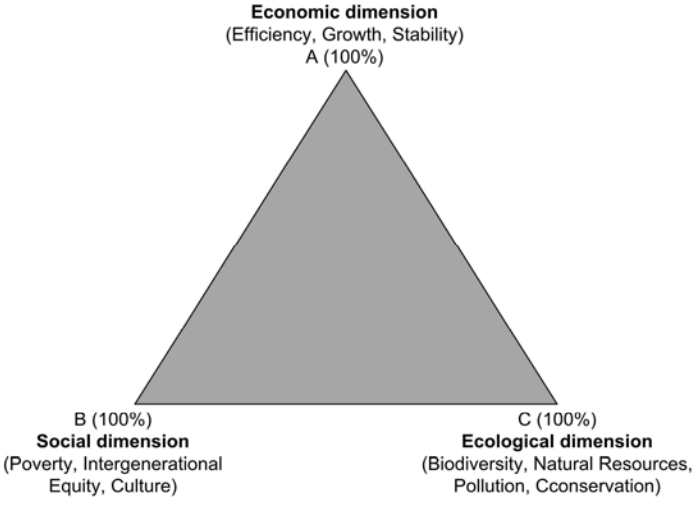


Fig. 1. Symbolic representation of environmental sustainibility (Giaoutzi and Nijkamp 1993)

Therefore a systemic approach stands out with clarity, reinforcing von Bertalanffy's theories (1967) focused on systems as realities more complex than the simple collection of their parts and characterized by openness and interactions: connections among economic, social and environmental systems have always to be considered; a sustainable development is in fact capable to integrate, settle and balance these three factors: social justice, economic utility and environmental integrity (Figure1) (Giaoutzi and Nijkamp 1993).

An effective sustainable development is then impossible without finding out integration among the above-mentioned dimensions; promoting sustainability means to pursue a balance among these different components (Munasinghe and Mc Neely 1995).

The combination of environmental sustainability fundamental values as well as their complete application in territorial context implies the identification of "what" a certain city and its territory aim to become, i.e. the *strategic vision of sustainable development* (Fusco Girard and Nijkamp 2000).

Such strategic vision has to be shared by public and private subjects and becomes the expression of an intergenerational social contract, which requires a debate among all stakeholders in order to find out a set of objective and shareable *values*. The value-oriented approach to sustainability (Keeney 1992) enables to evaluate the desirability of a development perspective and to focus on those strategic objectives that do not change day by day, but remain somehow stable in time.

In this context, a sustainable development becomes also a participative development: *sustainability and participation* result as two much correlated activities which allow identifying a long term development vision, the only one able to pursue and achieve the entire community general interest.

Long-period logic, proper to sustainability, imposes decision makers to *plan* temporal evolution of urban and territorial systems that appear too complex to be interpreted by articulated deterministic models. As a consequence, urban and territorial *evolution scenarios* have to be implemented: they represent "possible futures" that enable to focus on specific topics selected to better characterize a certain development hypothesis (Nijkamp 1994). Building a scenario means to single out a strategic vision of what a city and its territory might become, then to plan their effective sustainable development.

Since the Rio United Nation Conference in 1992, *urban and territorial planning* is considered as the effective instrument necessary to promote and put into practice the general principles of environmental sustainability, combining and weighing conservation with development in the most efficient way.

The possibility to assure a good quality of life to all citizens and the capability to guarantee social equity, biodiversity protection and environmental resources carrying capacity respect are considered urban planning main tasks.

The objective to carry out the so called "*strong*" sustainability is typical of strategic planning process, which aims to analyze potential territorial transformations and to identify those resources to be protected and exploited compatibly with cultural, historical and environmental values of a certain territory.

The idea of this type of sustainibility requires the identification of limit thresholds (Fusco Girard 1987, Nijkamp 1995) by means of appropriate *indicators* able to measure quantitative and qualitative characteristics of territorial resources.

Such analyses are almost accepted in planning practice, but the systemic feature of environmental phenomena calls for the construction of complex indices, which makes the specification of territorial transformation limits more difficult (Fusco Girard and Nijkamp 2000).

Sustainability defined as "weak" presents instead a different approach: not all forms of capital (natural, human, social, economical) are replaceable each other, therefore it is necessary to introduce some limits to their replaceability (Pearce and Turner 1990).

The application of weak sustainability principles within territorial planning belongs to the operative dimension of the plan that defines effective ambits of territorial interventions, previously established at a strategic level, and specifies functional contents of those transformations (Stanghellini 1999).

Urban planning is required to meet, above all, strong sustainibility principles in order to guarantee the conservation of unique and irreplaceable environmental resources, and, in second place, to observe weak sustainibility criteria according to which, in every urban and territorial transformation, the consumption of some environmental resources can be compensated by the increase of others (Lombardi and Micelli 1999).

Within urban and territorial planning process, *evaluation* is regarded as a core activity: it enables to face environmental sustainability complexity and permits to find out, for each specific territory, those solutions capable to best integrate environmental, social, cultural and economical objectives.

A *strategic and integrated assessment* is, in fact, the instrument by which it is possible to pass from theoretical *sustainable development* principles to their effective *application* in territorial space.

The complexity of a decision making process, involving territorial trasformations, reveals the necessity to carry out evaluations not only related to the feasibility (technical, economical, financial, etc.) of the proposed interventions, but also able to analyze their impacts on environmental, social, economical, cultural and political systems: this requires the devising and the construction of *multidimensional evaluation methods* (Las Casas 1992, Lombardo 1995).

Urban planning process can adopt a first type of evaluation which employs specific indicators, built "ad hoc" for the intervention under study, and aims to analyze the effects on the territory directly affected by the *project/plan impacts*: this type of evaluation can be called "*local*". On the other side, a model of *integrated assessment* focuses its attention on the study of systemic effects that an intervention may have on a vaste area.

In this context, the European debate relative to the definition of *territorial impact assessment* procedures, connected with national and EU policies, becomes very significant (CEC 2002, 2004, 2005). This discussion highlights the necessity to adopt integrated evaluation approaches able to overcome the sectoral valence of assessment techniques adopted since the second half of the nineties (cost benefit analysis, cost effectiveness methods, environmental auditing, etc.) in order to

improve the completeness and applicability level of the assessment process, as well as to indicate government action sustainability (Pearce and Secconbe-Hett 2000).

The evaluation process cannot be reduced to the function of audit, but it should aim to develop territorial sustainable development scenarios both adopting scientific methods and stimulating stakeholders' participation actions.

In this sense, territory has not to be strictly considered in physical terms, but as a place where economical, social and environmental changes converge and occur; such transformation characterizes development dynamics (Camagni 2006).

The analysis of European regulations, related to sustainable environmental assessment, highlights that, notwithstanding the EU efforts to widen objectives and significance of evaluation instruments, in particular by Strategic Environmental Assessment (Therivel et al. 1992), results obtained are still quite modest in comparison with the whole objectives established by regulation themselves.

This can be due to two main reasons (Camagni and Schizzerrotto 2004):

- complexity of evaluation process, sometimes considered as a limit for its complete application and, consequently, a potential pretext to start up ultra-simplified procedures;
- multidimensional and integrated nature of decision making process, connected to territorial transformations, in comparison with the sectoral character of adopted evaluation techniques.

Actually, some of the most used evaluation methods present a lack of balance in the direction of environmental variables or alternatively they have simply added or juxtaposed social and economical indicators to traditional and consolidated ecological and natural indices, without finding out a link among all the variables connected with the three dimensions (fig. 1) of sustainable development.

This fact is possibly due to the strictly "environmental" origin of most evaluation procedures: overlay mapping (Mc Harg 1969, Krauskopf and Bunde 1972), matrices and networks (Leopold 1971, Bereano 1972), weighting-scaling checklist such as the EES (Environmental Evaluation System) system (Dee et al. 1973) and the WRAM (Water Resource Assessment Methodology) method (Solomon et al. 1977).

The necessity to dispose of innovative evaluation methods (Camagni and Musolino 2004) able to consistently consider the three sides of sustainable development emerges. Social justice, economic utility and environmental integrity can be summarized in the *territorial dimension* that merges geographical and physical elements with environmental and natural features, with cultural social and economical characters.

Modeling and quantitative analysis in general suffered criticism after the enthusiasm of the years '50s and '60s of the past century, particularly for providing quite deterministic conclusions following the elaboration of data. The same models implemented by Meadows et al (1972) suffered such criticism, partially tackled in the following studies (1992 and 2004), although their effort played an important role in feeding the debate over environmental issues. To date, scholars involved in quantitative research, and particularly those interested in space, are aware of the issues raised by their object of study and of the limitations of the models and tools available. Particularly, scholars and researchers to-date do not limit the analysis to

quantitative aspects but even more make efforts both to test their conclusion on the real space and also to take into consideration qualitative aspects in the modeling, allowing the latter approach a renewed 'dignity' once neglected in favor of exact sciences.

Still, issues remain at stake when dealing with integration into models of the different elements that build the concept of sustainability. In fact, urban and economical models on one side and environmental ones on the other one have generally evolved in separate domains of knowledge, although maintaining a strong link with the relations between Earth and its inhabitants (Guhathakurta 2003). Urban models generally rely on elements as economic and spatial interaction with little attention on ecological dynamics, while environmental models focused mainly on non-urban - and therefore non-artificial subsets of the environment.

One of the common elements between the two 'families of models' lays on the limited capacity to follow rules of testing and accuracy that on the contrary characterize physical models; therefore social scientists rely on models that produce first insights to the studied object and focusing on some of its attributes.

Recently, these two aims have become closer and the debate over sustainability has probably helped in forcing these two sides to meet. Also, this convergence has helped in understanding social and environmental relationships and spillover effects can be observed in research in the different domains. This is particularly evident in the more recent advances and publications in geocomputational issues that, involving scholars from different research areas, have often made unsearched similarities to emerge. In fact, scholars interested on spatial issues, and sustainability heavily rely on space, generally need advanced geographical information instruments and tools and spatial analytical techniques to carry on their research. This involves them to improve knowledge on different theories and techniques, often not originally present in their 'home' research domain and therefore a contamination between disciplines arises. As a result, interdisciplinary areas as workshops and books play an important role in helping theories, ideas, methods, models and techniques to be exchanged and to understand different aspects and points of view of spatial issues, often noticing both a use of similar instruments to model different human or ecological characteristics or, still, the rising of common conclusions from different starting points and research agendas.

2 Geo-computational Approaches to Sustainability

The experience developed by Ian McHarg (1969) represents the first attempt to base environmental planning with more objective methods. More particularly he supposed that the real world can be considered as a layer cake and each layer represents a sectoral analysis (e.g. hydrology, geology, vegetation, transports, etc.). This metaphor represents the fundamental of overlay mapping. At the beginning, these principle applied only by hand, just considering the degree of darkness, produced by layer transparency, as a negative impact. In the following years, this craftmade approach, has been adopted for data organization in Geographical Information Systems producing analyses with a high level of quality and rigour.

Nowadays, great part of studies in environmental planning field have been developed using GIS. Campagna (2005) realized a complete survey of research topics and applications of GIScience methods to sustainable spatial planning. The next step than the simple use of geographic information in supporting environmental planning is the adoption of spatial simulation models which can predict the evolution of phenomena. As the use of spatial information has definitely improved the quality of data set on which to base the decision-making process, the use of Geostatistics, spatial simulation and more generally geocomputation methods allows the possibility of basing the decision-making process on predicted future scenarios. The aim of this book is to provide an overview of the main methods and techniques adopted in the field of environmental geocomputation in order to produce a more sustainable development. The authors have deliberately avoided to have a complete dissertation on geocomputation topics, because a complete illustration of the main concepts and the different approaches to geocomputation have been presented in a previous book Geocomputation and Urban Planning (Murgante at al. 2009).

In recent times, there is a growing attention to climate change, particularly due to the increasing awareness of this phenomenon. It is also more evident that coastal zones could be the more damaged by these changes, Starting from the assumption that climate changes are taking place, it is clear that coastal areas, with huge urbanization pressures, are most vulnerable to these changes. Unfortunately, planning documents, at regional and local level, do not take into account this phenomenon.

In the same way it is very strange that a discipline such as planning which programs the territory for the future years in great part of cases is not based on simulation models. Sectoral analyses, often based on surveys, are not enough to highlight the dynamics of the area. Better knowing urban and environmental changes occurred in the past, it is possible to provide better simulations predicting future scenarios.

The contribution of *Hansen* tries to solve both this issues adopting the model LUCIA - Land Use Change Impact Analysis (Hansen, 2007) - coupled with climate change scenarios. The integration of LUCIA, based on cellular automata land use simulation model, with climate change scenarios, produced detailed model of spatial dynamics identifying possible vulnerable zones and improving the decision making process.

The paper by *Vanegas et al.* presents a comparison between a mathematical formulation and a heuristic solution method with the ultimate objective to develop an high performance heuristic solution able to locate near to optimal sites, composed by a given number of cells (raster structure): these sites must be compact and maximize cells' intrinsic multiple criteria suitability.

The problem of location of optimal sites for afforestation of agricultural land considers levels of importance (weights) for the criteria involved in the analysis, therefore a MADM (Multi Attribute Decision Making) approach is applied. Since the paper focuses its attention on general formulations for site location, it does not adopt alternative methods for the computation of criterion weights, as they are considered as mathematical parameters from the optimization point of view.

The Integer Programming (IP) method for Multiple-Criteria Site Location develops linear programming formulation for acquiring optimal sites composed by a set of pixels forming a compact and contiguous site: this result is obtained maximizing its performance according to one or more weighted Normalized Criteria (NC). Then, an alternative solution is devised in order to locate feasible (near to optimal) sites that fulfill multiple criteria requirements: this method is a Heuristic Multiple-Criteria solution for Site Location (HMSL), which is divided in 3 stages: seeds generation, region growing and region ranking. The final goal of the entire algorithm is to construct a compact site maximizing the intrinsic multiple criteria suitability of the cells.

To validate the heuristic approach, a comparison with a mathematical formulation is performed with afforestation data of reduced areas within The Netherlands, Denmark, and Flanders, but the authors believe that the proposed approaches for site location can be applied on grids of cells with attribute data generated by any other application field. This experimentation reveals that heuristic method is considerably faster than the mathematical one, that objective values obtained with the two approaches are substantially similar and that the heuristic solutions are also spatially nearer to the optimal patch location. Whereas the HMSL has a higher performance, each whole territory is processed with this method and a sensitivity analysis is carried out to determine its behavior for different parameter values. This analysis shows that regions' homogeneity plays an important role for identifying the most suitable sites, but the obtained results show that HMSL is entirely applicable for real sized problems, even considering a high number of candidate solutions.

The research carried out by *Lombardo and Petri* implements a methodology able to evaluate both the opportunity to use wind energy and, in the affirmative case, the producibility (productivity?) and location of a wind farm. An evaluation framework is developed, which can be used both for wind energy and for other renewable energy sources exploitation.

As systematic planning and programming for renewable energy plants siting are still at embryonic stage, the present role of local authorities is not to develop new strategies, but it is mainly limited to the approval or the denial of private proposals. However, in Italy, none of the current procedures, norms or laws attempts to put into relation the different aspects involved in the exploitation of renewable energy sources and their environmental and visual impact, so that a reliable assessment is not possible (even not required) and rational planning of installation locations is quite impossible.

The general objective of the proposed research is to develop methods and techniques aimed to produce systematic and integrated knowledge on the possibility of production and sustainable use of renewable energy sources. The specific objective is to build a planning support system where sectoral modelling and evaluations can be integrated with advanced spatial analysis techniques in order to assess the possibilities of production and sustainable use of renewable energy on the base of natural vocations/potentiality of territorial systems. The conceived methodology is composed of two parts: the first concerns the analysis at the wide scale of the area probably affected by the wind blade impacts and the second part implements local impact assessment of each possible wind farm locations.

The entire methodology and its application on an Italian case study (S. Luce, Livorno) is an attempt to elaborate a procedural framework to evaluate the landscape impact of every new type of territorial element (wind farms, dumps, incinerators, etc.). So, landscape quality measure (an element clearly difficult to schematize) and its assessment are included into a process that starts from the impacted area characterization and ends with local impacts mitigation.

The paper by *Danese et al.* is also based on the analysis for wind farms location. More particularly, visual impact assessment has been enriched with a new type of viewshed analysis. Single, multiple and cumulative viewshed do not collect information concerning visibility of single objects distinguished in an evident way from the others. The Identifying Viewshed (Danese et al 2009) implemented in GRASS environment overcomes this limits allowing more accurate analyses and producing more detailed assessments.

The paper by *Mauro* tackles an issue related to sustainable development, particularly from the point of view of the safeguard and recovery of terraced landscape. In light of the pursuit for a sustainable development, the main aim of this study is to identify the recent evolution of terraced areas, representing the first step to plan their recovery. Actions taken to identify areas and to monitor changes are carried on using GIS analytical functions and an application to the province of Trieste (North-eastern Italy) is presented.

The research brought to the realization of different maps related to different timescales. In particular, four rural land-cover classes were realized, these being derived from MOLAND geographical database, realized for urban areas at European level and extended to the overall regional territory in Friuli Venezia Giulia Region. Maps of land cover likely to host terraced formation for the four considered periods (1955, 1975, 1985 and 2000) were crossed with a map of morphology derived from a DEM, selecting 5-45% slope areas. The combination of such intermediate maps, as rural maps with slopes for each considered year – were used to estimate past or current presence of terraced landscapes.

The contribution of *Novack et al*. is to construct and compare potential models for the estimation of population density for 212 census sectors of brasialian São Paulo city (10.886.513 inhabitants, living in an area of 1.523 km^2) using, as independent variables, landscape metrics calculated on a classified image from the QuickBird II sensor. As the traditional approaches to estimate population density are mainly based on field surveys, they become, specially for big cities, labor-intensive, time-consuming, costly and difficult to update: in this context the combination of high resolution remote sensing data and spatial regression techniques represents a very powerful tool. To this end the authors devised and implemented a new methodology structured through several consequential steps.

First of all, the QuickBird image is classified by the Maximum-Likelihood method and all independent variables are generated by FragStats software over this satellite image. Independent variables considered on each census sector are only four: number of polygons classified as ceramics roofs, percentage of class with dark roof, aggregation index of the streets class and the patch density of the vegetation class. As some of the selected variables could well explain population density, two ordinary linear regression models (called M1 and M2) are selected

and formal statistical tests are applied. For the analysis of the spatial dependency of residuals, a distance-based neighborhood matrix is created: for spatial dependency detection Global Moran index is used, nevertheless the strongest evidence on the existence of spatial dependency of residuals for both models is the visual inspection of LISA maps. The next step is to test which spatial regression model would be the most suitable for each model and to this aim Lagrange Multiplier test statistics are used. Finally, the running of spatial regression models is conducted as well as the inclusion of dummy variables in order to compare them in terms of spatial dependency elimination. With the purpose to create a simple but accurate enough model, which could be applied onto other areas and validated in other conditions, the most appropriate model for the test area's population density estimation is model M1, with the introduction of Spatial Lag parameter. Results of this experimentation prove that population density can be relatively well estimated by the use of spatial metrics calculated over a classified high resolution image when using the population density itself as independent variable in spatial regression models.

The concept of spatial autocorrelation has been applied using different techniques in the paper by *Lucia et al.*. The integration of chemical and physical parameters and geostatistical techniques has found many applications in recent years that have allowed the estimation of background values for contaminants in the investigated matrices. More particularly this paper investigates the distribution of some heavy metals on the geochemical characterization of the polluted site of national interest Tito (PZ) through the application of geostatistical techniques. All geostatistical analysis were carried out using the "geoR" spatial extension of the statistical software "R".

The paper by *Vivanco et al.* "Evaluating the impact of resolution on the predictions of an air quality model over Madrid area (Spain)" compares different model results and observations obtained using traditional standard evaluation statistics with different scales of domains, noticing an improvement when observing a coarsest domain, but noticing a much lower or inexistent one for finer domains. A better behavior of pollutant predictions should be related to an improvement of meteorological predictions in terms of parameterizations involved in the meteorological model and input data, such as land use information. In particular, the research carried on is aimed at comparing a same set of analytical tools at different scales, observing a better simulation envelope when working at a more refined and detailed local scale, rather than on a coarser one, therefore observing that resolution plays a significant role in modeling, especially when trying to simulate local effects.

Pollution issues have been faced from a different point of view also in the contribution by *Di Martino et al.*. More particularly this topic has been analyzed with a spatial Data Warehouse approach. Spatial OLAP (SOLAP) systems allow multidimensional analysis of geographical data. On the other hand, they lack of geovisualization techniques to take advantage of great analysis capabilities of georeferenced data such as 3D visualization, cartographic displays personalization, etc. Then, in this paper, authors present a new spatial decision support system that integrates well-established OLAP and geovisualization tools: Mondrian/JPivot and

Google Earth. Some experiments on a simulated multidimensional dataset concerning pollution of Italian regions show the functionalities of the proposed system.

Cellura et al. deal with "Nonlinear black-box models for short-term forecasting of air temperature in the town of Palermo", where their attention is focused on the implementation of a system for obtaining short-term forecasts of air temperature and on mapping them over the monitored area of the city of Palermo (Italy), elaborating data collected hourly by a network of weather stations. The interest is, among other issues, on the Urban Heat Island (UHI) affecting metropolitan areas. The system clearly highlights temperature differences between urban and extra-urban area and average intensity of UHI of Palermo. The research is the starting point for a forecasting model with a wider time horizon in order to obtain the future evolution of the temperature with a relevant advance and to use this information to study the evolution of urban comfort conditions.

The contribution of *Kanevski et al.* is about "Automatic Mapping and Classification of Spatial Environmental Data". Authors focus on the development and consequent application of a methodology for automatic processing of environmental data for mapping and classification purposes. The attention is drawn towards spatial regression and a method called General Regression Neural Network (GRNN) is considered in detail, as well as a Probabilistic Neural Network (PNN) as an automatic tool for spatial classifications.

The idea underneath the work presented here resides on the importance of the environmental automatic decision-oriented treatment of data, from exploratory data analysis to mapping and classification, for decision-oriented purposes. In general, the problem is to develop nonlinear, robust, adaptive and data driven methods which can detect spatial patterns in noisy data and which are able to produce prediction and probabilistic/risk maps which can be used in a real decision making process. Such methods are also important, as they should be able both to model multi-scale variability in data and to detect hot-spots.

More generally some data are said to be ambiguous if they can have at least two particular interpretation. Ambiguity leads to a discordance in data classification due to a different perception of the phenomenon. Inaccuracy produces uncertainty in the case of low quality of data, due to a certain degree of error (Murgante et al 2008, Murgante and Las Casas 2004). Vagueness (Erwig and Schneider, 1997) takes into account multi-valued logic and it is based on the concept of "boundary region" which includes all elements that cannot be classified as belonging to a set or its complement (Pawlak, 1998). Three theoretical approaches to vagueness exist: the first one is based on fuzzy set theory (Zadeh, 1965) which accounts for partial membership of elements to a set; the second is Egg-Yolk Theory (Cohn and Gotts 1996, Hazarika and Cohn 2001) based on the concepts of "egg", i.e. the maximum extension of a region, and "yolk", i.e. the inner region boundary; the third approach is rough set theory. These issues have been analyzed in the paper by *Taramelli*, which analyze uncertainty in landform boundary definition. More particularly semantic and geometric approaches to landform definition have been analyzed adopting Fuzzy Logic, which allows each pixel to have a degree of membership.

The theme of uncertainty is also treated by *Hagen-Zanker and Martens*. In particular, they analyze how geosimulation models are close to reality, testing quality and accuracy of results. Also, they highlight that in several cases constrained cellular automata model can lead to false conclusions.

References

Archibugi, F., Nijkamp, P. (eds.): Economy and Ecology: towards Sustainable Development. Kluwer, Dordrecht (1989)

Bereano, A.: A Proposed Methodology for Assessing Alternative Technologies. Cornell University, New York (1972)

Bruntland Report, Our Common Future. Oxford University Press, Oxford (1987)

Camagni, R.: Territorial Impact Assessment – Tia: a Methodological Proposal. Ita. J. Reg. Sci. 5/2, 135–146 (2006)

Camagni, R., Musolino, D.: Verso una valutazione d'impatto territoriale di politiche, piani e programmi. Edizioni 31, Trento (2003)

Camagni, R., Schizzerotto, A.: Metodi e applicazioni di ricerca valutativa per la pubblica amministrazione. Edizioni 31, Trento (2004)

Campagna, M.: GIS for Sustainable Development. Taylor & Francis Group – CRC, Boca Raton (2005) (United States)

CEC - Commission of the European Community, Towards a Global Partnership for Sustainable Development. Commission Staff Working Paper (Sec 82), Brussels (2002)

CEC - Commission of the European Community, Impact Assessment: Next Steps. Commission Staff Working Paper (Sec 1377), Brussels (2004)

CEC - Commission of the European Community, Impact Assessment Guidelines. Commission Staff Working Paper (Sec 791), Brussels (2005)

Cohn, A.G., Gotts, N.M.: The 'Egg-Yolk' representation of regions with indeterminate boundaries. In: Burrough, P.A., Frank, A.U. (eds.) Geographic objects with indeterminate boundaries. Taylor & Francis, London (1996)

Costanza, R. (ed.): Ecological Economics: the Science and the Management of Sustainability. Columbia University Press, New York (1991)

Daly, H.E., Cobb, J.J.: For the Common Good. Beacon Press, Boston (1990)

Danese, M., Nolè, G., Murgante, B.: Visual Impact Assessment in urban planning. In: Murgante, B., Borruso, G., Lapucci, A. (eds.) Geocomputation and Urban Planning. SCI, vol. 176, pp. 133–144. Springer, Berlin (2009)

Dee, et al.: Planning Methodology for Water Quality Management: Environmental Evaluation System. Battelle-Columbus Laboratories, Columbus, Ohio (1973)

Erwig, M., Schneider, M.: Vague Regions. In: Scholl, M.O., Voisard, A. (eds.) SSD 1997. LNCS, vol. 1262, pp. 298–320. Springer, Heidelberg (1997)

Fusco Girard, L.: Risorse architettoniche e culturali: valutazioni e strategie di conservazione. Franco Angeli, Milano (1987)

Fusco Girard, L., Nijkamp, P. (eds.): Le valutazioni per lo sviluppo sostenibile della città e del territorio, Franco Angeli, Milano (2000)

Giaoutzi, M., Nijkamp, P.: Decision Support Models for Sustainable Development, Aldershot, Avebury (1993)

Guhathakurta, S. (ed.): Integrated Land Use and Environmental Models – A Survey of Current Applications and Research. Springer, Berlin (2003)

Haggett, P.: Geography. A Global Synthesis. Pearson Education Limited, Harlow (2001)
Hansen, H.S.: An Adaptive Land-use Simulation Model for Integrated Coastal Zone Planning. In: The European Information Society. Lecture Notes in Geoinformation and Cartography (2007)
Hazarika, S.M., Cohn, A.G.: A Taxonomy for Spatial Vagueness, an alternative egg-yolk. In: Montello, D.R. (ed.) COSIT 2001. LNCS, vol. 2205, pp. 92–107. Springer, Heidelberg (2001)
Keeney, R.L.: Value focused thinking: a path to creative thinking decision making. Wiley, New York (1992)
Krauskopf, T.M., Bunde, D.C.: Evaluation of Impact through a Computer Modelling Process. In: Ditton, R., Goodale, T. (eds.) Environmental Impact Analysis: Philosophy and Methods. University of Wisconsin Press, Wisconsin (1972)
Las Casas, G.B.: Una ricerca di razionalità a-priori nella valutazione dei piani. Territorio 12, 7–21 (1992)
Leopold, L.B.: A procedure for Evaluating Environmental Impact. US Geological Survey Circular (45) (1971)
Lombardi, P., Micelli, E. (eds.): Le misure del piano: temi e strumenti della valutazione dei nuovi piani. Franco Angeli, Milano
Lombardo, S. (ed.): La valutazione nel processo di piano. Franco Angeli, Milano (1995)
McHarg, I.: Design with Nature. Natural History Press, New York (1969)
Meadows, et al.: The limits to growth. Potomac Associates Universe Books, New York (1972)
Meadows, et al.: Beyond the limits: confronting global collapse, envisioning a sustainable future. Chelsea Green Publishing Company, Post Mills (1992)
Meadows, et al.: Limits to Growth-The 30 year Update. Chelsea Green, Vermont (2004)
Munasinghe, M., Mc Neely, J.: Key Concepts and Terminology of Sustainable Development. In: Munasinghe, M., Shearer, W. (eds.) Defining and Mesuring Sustainability. The World Bank, Whashington (1995)
Murgante, B., Borruso, G., Lapucci, A.: Geocomputation and Urban Planning. In: Murgante, B., Borruso, G., Lapucci, A. (eds.) Geocomputation and Urban Planning. SCI, vol. 176, pp. 1–18. Springer, Berlin (2009)
Murgante, B., Las Casas, G., Sansone, A.: A spatial rough set for extracting periurban fringe. In: Revue des Nouvelles Technologies de l'Information, Cépaduès, Toulouse, France, vol. 857, pp. 101–125 (2008), ISSN: 1764-1667
Murgante, B., Las Casas, G.: G.I.S. and fuzzy sets for the land suitability analysis. In: Laganá, A., Gavrilova, M.L., Kumar, V., Mun, Y., Tan, C.J.K., Gervasi, O. (eds.) ICCSA 2004. LNCS, vol. 3044, pp. 1036–1045. Springer, Heidelberg (2004)
Nijkamp, P.: Sustainable Cities in Europe. Earthscan, London (1994)
Nijkamp, P.: Quantity and Quality: Evaluation Indicators for Our Cultural Heritage. In: Coccossis, H., Nijkamp, P. (eds.) Planning for Our Cultural Heritage. Aldershot, Avebury (1995)
Nijkamp, P. (ed.): Sustainability of Urban Systems. Aldershot, Avebury (1990)
O.N.U.: The Global Partnership. A Guide to Agenda 21, New York (1993)
O.N.U.: Habitat Agenda. Conference on Human Settlements (Habitat II), Istambul (1996)
Odum, E.P.: Ecology and our Endangered Life-Support Systems. Sinuaer Associates, Sunderland (1989)
Pawlak, Z.: Rough Sets theory and its applications to data analysis. Cybernetics and Systems An International Journal 29, 661–688 (1998)

Pearce, D.W.: Blueprint 2: Sustaining the World Economy. Earthscan, London (1991)
Pearce, D.W., Markandya, A., Barbier, E.: Blueprint for a Green Economy. Earthscan, London (1989)
Pearce, D.W., Secconbe-Hett, T.: Economic Valuation and Environmental Decision-Making in Europe. Environ. Sci. Technol. 34(8), 1419–1425 (2000)
Pearce, D.W., Turner, R.K.: Economics of Natural Resources and the Environment. Harvester Weathsheaf, New York (1990)
Sen, A.: Risorse, valori e sviluppo. Bollati Boringhieri, Torino (1992)
Serageldin, I., Barret, R., Brown, J.: The Business of Sustainable Cities. The World Bank, Whashington (1995)
Solomon, R., et al.: Water Resolurces Assessment Methodology (WRAM). Impact Assessment and Alternative Evaluation. Contract Report Y-77-1, Vickburb, Mississipi (1977)
Stanghellini, S.: Riforma urbanistica e domanda di valutazione. In: Lombardi, P., Micelli, E. (eds.) Le misure del piano:temi e strumenti della valutazione dei nuovi piani, Franco Angeli, Milano (1999)
Therivel, R., Wilson, E., Thompson, S., Heaney, D., Pritchard, D.: Strategic Environmental Assessment. Earthscan, London (1992)
Tinacci Mossello, M.: Politica dell'ambiente – Analisi, azioni, progetti. Il Mulino, Bologna (2008)
Vallega, A.: Geopolitica e sviluppo sostenibile – Il sistema mondo del secolo XXI. Mursia, Milano (1994)
von Bertalanffy, L.: General System Theory. Foundation, Development. George Braziller Inc., New York (1968)
Zadeh, L.: Fuzzy sets. Information Control 8, 338–353 (1965)
Zamagni, S.: Global Environmental Change, Rationality and Ethics. In: Campidoglio, L., et al. (eds.) The Environment after Rio. Graham and Trotnam, London (1994)

Urban Land-Use Projections Supporting Adaptation Strategies to Climate Changes in the Coastal Zone

Henning Sten Hansen

Department of Development and Planning, Aalborg University, Fibigerstræde 11, DK-9200 Aalborg East, Denmark
e-mail: hsh@land.aau.dk

Abstract. Climate change has received much attention during the last decennium and especially various mitigation and adaptation strategies. Particularly the coastal zone will feel the consequences of climate change and the associated effects like sea level rise, increased storminess and flooding. Thus there is an urgent need for local and regional spatial planners to include climate change in their planning efforts. Using modelling and simulation, we can increase our understanding of the future land-use system under influence of a changing climate and accordingly reduce uncertainty concerning decisions. The current paper describes how land-use simulations combined with climate change scenarios represented by the SRES narratives can facilitate the definition of adaptation strategies to counteract the consequences of potential climate changes.

Keywords: Land-use modelling, climate change, SRES scenarios, sea level rise.

1 Introduction

There is growing evidence that our climate is changing. These climatic changes are manifesting themselves in different ways. Global average temperature has already increased by 0.7 °C and in Europe by 0.95°C above pre-industrial levels, and we can expect Europe's climate to continue to change over the coming years. In its Third Assessment Report, the Intergovernmental Panel on Climate Change - IPCC (2001) drew on a series of modelling approaches to estimate the global parameters of future change. These suggest that over the next 100 years, a rise of between 1.4 ºC and 5.8 ºC might be expected in average near-surface temperatures, bringing a range of associated changes including increase in sea level and alterations in patterns of precipitation and storm events. Particularly in the coastal zones the potential consequences of climate change are a cause of mounting concern. Coastal areas are perceived as particularly vulnerable to the impacts of climate change because they are subject to changes both in the marine environment and in the terrestrial environment. They would be affected by sea level rise, and any changes in storm surges and wave heights, and they would also be affected by

B. Murgante et al. (Eds.): Geocomputation, Sustainability & Environ. Planning, SCI 348, pp. 17–34.
springerlink.com

changes inland, including alterations in river flow regimes. It is estimated that 9 % of all European coastal zones (12 % for EU Member States), which can be defined as a 10 km strip, lie below a 5 m elevation and are potentially vulnerable to sea level rise and related inundations and for Denmark this figure is 22 % (European Environment Agency, 2005). Adapting to climate change is therefore an essential part of ensuring our communities to remain desirable places to live and work.

Measures to reduce emissions are only part of the climate change challenge. Even if we make significant reductions in emissions tomorrow, the lag in the climate system means that emissions we have already put into the atmosphere will continue to affect the climate for several decades to come. The impact on towns and cities and their inhabitants will be significant, and consequently adaptation strategies are needed.

Climate-change vulnerability assessment has become frequently employed with the purpose of informing policy-makers attempting to adapt to global change conditions. The European nature and agricultural areas - particularly in the coastal zone - are under increasing pressure from urbanisation, and the global warming will further enhance this pressure. Accordingly spatial and environmental planners have urgent needs for scenario tools analysing the impact of possible land use changes. Several bodies have called for a more integrated management of the coastal zone as a fundamental prerequisite for sustainable development, and one of the most recent efforts is the EU Recommendation for a European Strategy for Integrated Coastal Zone Management (European Commission, 2000). Following the European strategy several projects financed by INTERREG were launched. The conclusions from the NORCOAST (2000) project can serve as common recommendations for most of the projects. The general recommendations focus on improvement of the planning and decision-making processes to create better-integrated and more sustainable solutions for the development of coastal areas. This should be accomplished through for example involving all stakeholders and politicians from an early stage in the planning process, and by developing tools for assessing the consequences for various planning initiatives – here under setting up scenarios for finding the best balance between different interests (Nakicenovic et al., 2004). The INTERREG IIIB project Forum Skagerrak aimed at developing policies and measures supporting a sustainable development around the Skagerrak Sea. As an important part in this project we developed a modelling framework for Land Use Change Impact Analysis – LUCIA (Hansen, 2007).

Using modelling and simulation, we can reduce uncertainty and increase our understanding of the land-use system. Spatial planning is a future-oriented activity, strongly conditioned by the past and present, and the planners need to enhance their analytical problem solving and decision making capabilities. The help of land-use models can facilitate scenario building and provide an important aid in the decision making process.

The aim of the current project has been to develop urban land-use scenarios based on the narratives for the future societal development as described in SRES - Special Report on Emissions Scenarios (Nakicenovic, N. et al., 2004), and describe the potential use of the scenarios in an integrated climate change and coastal zone management context. The case area is the Region of Northern Jutland

where particularly the low land around Limfjorden is most sensitive to the effects of climate change. Built-up areas represent huge societal investments and accordingly a rather long life expectancy - for example 100 years. Therefore, new built-up areas must be established considering what we call a 100 years flood risk zone.

The paper is divided into 5 sections. After the introduction a brief description of climate change scenarios and their effect on sea level rise and flooding follows. In part 3 we describe the land-use modelling framework and data applied. Part 4 presents the scenario approach applied and discusses the results of the simulations. The paper ends with some conclusions and an outline for subsequent work.

2 Climate Change Scenarios

Emission scenarios are a crucial part of climate change research. Emission scenarios are derived from population, economic and technology scenarios, which also shape vulnerability and therefore impacts of climate change. Emission scenarios are an important component of IPCC assessments. SRES (Special Report on Emissions Scenarios) is an international framework, which proposes a range of possible futures that integrate socio-economic and climate change driving forces.

2.1 Emission Scenarios

The SRES were developed as quantitative interpretations of four alternative storylines which represent possible futures with different combinations of driving forces. All four storylines and scenario families describe future worlds that are generally more affluent compared to the current situation. They range from very rapid economic growth and technological change to high levels of environmental protection, from low to high global populations, and from high to low GHG emissions. What is perhaps even more important is that all the storylines describe dynamic changes and transitions in generally different directions. The storylines are summarized as follows (Nakicenovic, N. et al. 2004):

- The A1 storyline and scenario family describe a future world of very rapid economic growth, low population growth, and the rapid introduction of new and more efficient technologies. Major underlying themes are convergence among regions, capacity building and increased cultural and social interactions, with a substantial reduction in regional differences in per capita income.
- The A2 storyline and scenario family: a very heterogeneous world. Fertility patterns across regions converge very slowly, which results in high population growth. Economic development is primarily regionally oriented and per capita economic growth and technological change are more fragmented and slower than in other storylines.
- The B1 storyline and scenario family describe a convergent world with the same global population as in the A1 storyline but with rapid changes in

economic structures toward a service and information economy, with reductions in material intensity, and the introduction of clean and resource-efficient technologies. The emphasis is on global solutions to economic, social, and environmental sustainability, including improved equity, but without additional climate initiatives.

– The B2 storyline and scenario family describe a world in which the emphasis is on local solutions to economic, social, and environmental sustainability, with continuously increasing population (lower than A2) and intermediate economic development and less rapid and more diverse technological change than in the B1 and A1 storylines. While the scenario is also oriented towards environmental protection and social equity, it focuses on local and regional levels.

Based on the SRES scenarios and spatio-temporal global circulation models the international climate research community has developed several global climate change scenarios. When you will estimate the effect of climate change at the local or regional level it is necessary to downscale the results of the global simulations by introducing various local circumstances.

2.2 Expected Sea Level Rise

Sea level rise is one of the most apparent and widespread consequences of climate change. Predictions of global mean sea level rise can be made with more confidence than many other aspects of climate change science. The thermal expansion of the world's water bodies and the melting of glaciers is slow, but nevertheless a process with far-reaching consequences. Thus the sea level rise will increase the exposure of coastal populations to storm surges and storm waves. The destructive power of such events will increase as a consequence of higher mean sea-level, and higher waves will be capable of reaching the original shoreline (defined as the shoreline prior to the rise in sea level) and areas further inland will become exposed to wave action.

Local changes in sea level depart from the global mean trend due to regional variations in oceanic level change and geological uplift/subsidence; it is the relative sea-level change that drives the impacts and is of concern to coastal planners. Rapid population growth, urban sprawl, growing demand for waterfront properties, and coastal resort development have additional harmful effects on protected coastal ecosystems. The rate of coastal erosion is primarily a function of mean water level (driven by relative sea-level rise), storm and wave forcing, sediment supply, and the form and response of the shore zone.

Simulations based on the A2 and B2 emission scenarios are explored in the current research project. The A2 scenario represents the more severe case corresponding to a dramatic increase in equivalent CO_2 content from 353 ppm in 1990 to 1143 ppm for the future climate in 2100. The less severe B2 scenario represents an increase of the equivalent CO_2 content from 353 ppm to 822 ppm. Based on SRES the global average sea level is projected to rise from 1990 to 2100 between 0.09 m

and 0.88 m (Meier et al., 2004). According to the Danish Coastal Authority (2007) the Danish coasts will partly be affected by the rising sea level and partly by increased storminess, because of changed wind forces and wind directions. Generally this will result in increased erosion of the coasts and reduced safety against flooding of low-lying areas, of which many today are protected by sea walls.

In order to estimate the local sea level rise, you must take the post-glacial uplift into consideration. For the northern part of Jutland the global sea level rise will accordingly be reduced of 9 cm in year 2100 (Danish Coastal Authority, 2007). Finally, an expected sea level rise due to storm surges must be added, and in the Limfjorden area this corresponds to up to 200 cm in the current situation. Based on these figures and on a local digital elevation model we created impact maps to estimate the spatial extent of inundation and flooding in year 2100 for the most severe case (see fig.1).

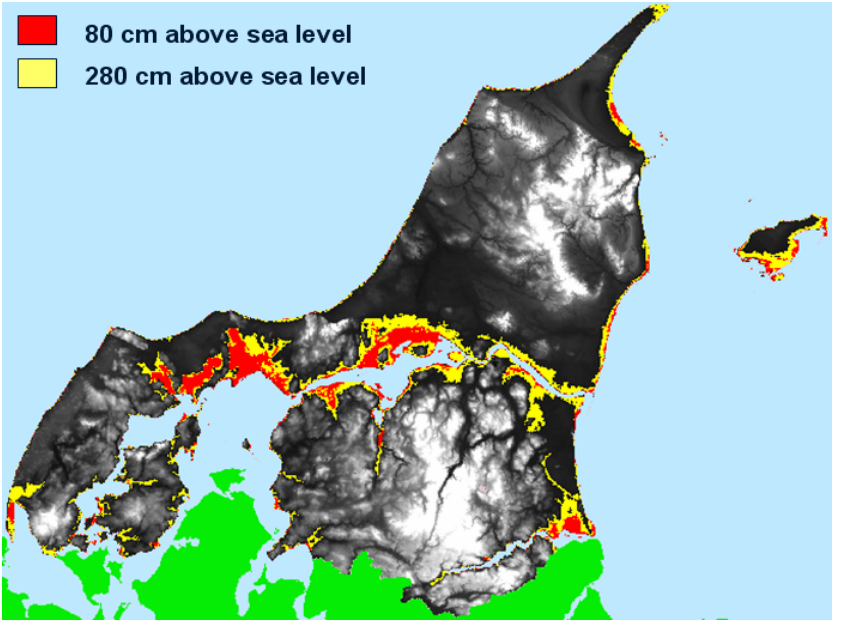


Fig. 1. Elevation model with potential flooding in year 2100

2.3 Adaptation Strategies through Spatial Planning

Given the long lifetime – 50 – 100 years - and high cost of the built environment it makes climate change a current, rather than a future topic Spatial planning and land management provide various tools to prevent natural hazards. The prevention of catastrophes in general is a consideration of spatial planning and land management on the regional and local level. Therefore a more active role of planning and land management is necessary in order to adapt to future climate changes.

However, spatial planning can only react to gradual changes like sea level rise with time horizons of decades, whereas spatial planning has rather limited effects on sudden events like extreme precipitation events or storm surges.

Generally there has always been a trend to locate new housing and associated services near seashores and riverbanks, and the attraction of the coast is even more pronounced for summer cottages and other buildings for leisure. Industries are similarly often located near harbours, with easy access to ship transport. This will inevitably enlarge the risk of flooding due to future sea level rise and storm surges and enhance the necessity and challenges for integrated coastal zone planning and management. Accordingly it is imperative that we plan for and create communities that are robust in the face of climate change.

Several EU Directives and other legislation consider the consequences of climate change. Under the EU SEA Directive (European Commission, 2001) planners are legally obliged to consider climate change when developing spatial plans. There are two climate change issues to address – the impact/ constraints set by climate change on the plan, and the plan's effects on future GHG emissions. Besides, the Floods Directive - the European Directive on the Assessment and Management of Flood Risks (European Commission, 2007) - is designed to support the Member States preventing and limiting floods and their damaging effects on human health, environment, infrastructure and property. The Floods Directive came into force on 26 November, 2007 and Member States have 2 years during which to transpose the Directive into domestic law. Finally, on 29 June 2007, the European Commission has adopted its first policy document - a Green Paper - on adapting to the impacts of climate change. In addition to the EU regulation most countries have set up their own adaptation strategy against climate change.

They have to support a sustainable settlement development and a sustainable land use in consideration of the different public and private interests because of their important influences on environmental disasters. There is a need for comprehensive vulnerability analysis to be undertaken for risk-prone areas.

According to the Spinning-Top Decision-Evaluation framework the decision-making process can be divided into 5 steps: Decision preparation, Decision, Implementation, Evaluation, and Feedback (Vedung, 2000). Particularly, the decision preparation phase can be assisted by land-use modelling. Simulations of future land-use provide planners and policy makers with descriptions and knowledge about possible spatial developments paths in the decision preparation process. Particularly scenario studies are appropriate for assessing the consequences of alternative policies. Typically a baseline scenario with business as usual settings and several policy scenarios are produced to support decision preparation.

The various SRES are constructed using very different postulated future world economic and social conditions without assigning any probabilities to any of them. For planners, this causes a difficult situation, as the projections of future climate made by GCMs using the SRES differ considerably between storylines by the end of the current century. What amount of sea level rise should therefore be assumed for planning purposes?

3 The Land-Use Modelling Framework

Within the current research we used the LUCIA modelling framework (Hansen, 2007). LUCIA is a traditional decision support system with factors and constraints, where the spatial dynamics are modelled through constrained cellular automata (CA).

Cellular automata is an obvious way to take spatial interaction into account and CA based models have been a very popular way of implementing dynamic land-use models. Basically, cellular automata models determine the number of cells to be changed in the next time step endogenously based on the defined transition rules. However the pure CA approach is not appropriate for land-use simulation, and like other recent CA models (Engelen et al., 2002; Barredo et al., 2003) LUCIA is based on constraint cellular automata being driven by external forces.

3.1 Land-Use Simulations

LUCIA has a multi-level structure, where the upper regional level represents drivers, whereas the detailed lower level represents land-use (figure 2). The driving forces for the amount of rural-urban change are basically population and economic growth. These drivers represent what we call macro-level drivers, and they are modelled externally to our model in various sector models, and basically define land demand from each active land-use type. Statistics Denmark makes every year national level projections for population, and these national figures are afterwards distributed to the local level (municipalities).

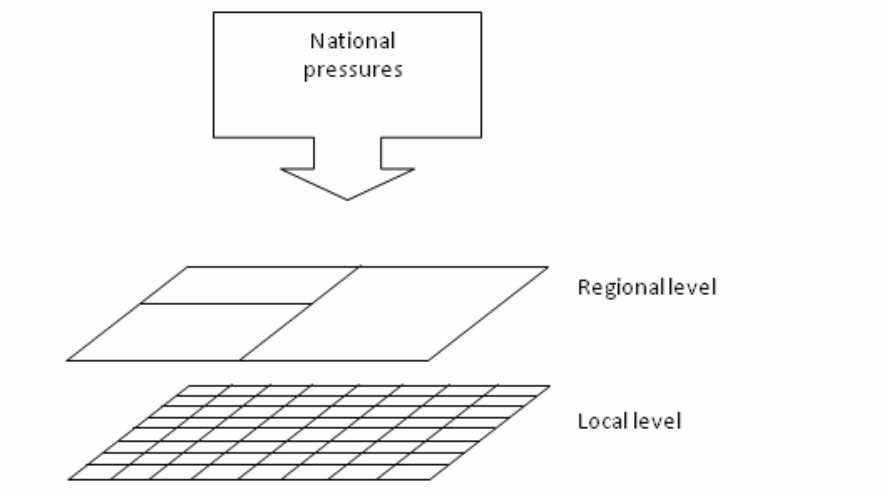


Fig. 2. The LUCIA multi-level structure

At the micro level, we deal with constraints and factors often used in various land-use modelling efforts. Policy making at national and local level has a strong influence on land-use – particularly policies having a spatial manifestation like creation of conservation areas or designation of areas for subsidised development (Verburg et al., 2004 a). However even more general legislation like the EU Common Agricultural Policy has a strong indirect influence on the spatial development in rural areas. However, the current version of the model does only involve policies and legislation with an explicit spatial aim under the headline *Zoning*.

LUCIA can incorporate 5 factors. The first factor involved in the model is the neighbouring effect, which represents the attractive or repulsive effects of various land-uses within the neighbourhood. It is generally well known that some land-use types, for example private service (shopping), tend to cluster, whereas others – e.g. recreation and industry - tend to repel each other. This is often referred to as the First Law of Geography (Tobler, 1970). However cells, which are more remote, will have a smaller effect. Traditional CA models most often adopted this effect using either the 3×3 Moore neighbourhood or an extended 5×5 neighbourhood. However, in the case of land-use modelling, the idea of neighbourhood may be much larger, since people and organisations are aware of their surroundings in a wider space. Therefore, it is desirable to define a neighbourhood large enough to capture the operational range of local processes. Similarly to Engelen et al. (2002) we decided to use an extended ring-shaped neighbourhood structure – currently with ten rings. Within the model we refer to this effect by the term *proximity*. Often the proximity effect has been estimated through the calibration phase, but recently we have developed a new method to quantify and analyse the neighbourhood effect in land-use modelling (Hansen, 2008). Through access to urban land-use maps for several consecutive years, we can easily identify neighbouring land-uses for new urban cells. The empirically derived neighbourhood functions confirm the expected positive attraction between existing residential areas and new cells with residential land-use. Similarly, our expectations regarding the repellent effect between existing industrial land-use and new residential cells are confirmed. The proximity effect Q for land-use L at cell (i,j) is calculated according to the following equation:

$$Q_{L,i,j} = \sum_{r=1}^{10} \sum_{K} w_{L,K,r} \delta_{r,K}$$

where r refers to ring number and $w_{L,K,r}$ is a weight representing the degree of interaction between land-use L and land-use K at a certain distance (ring). $\delta_{r,K} = 1$ if land-use of the cells in the ring r is equal to K. Otherwise $\delta_{r,K} = 0$.

Besides proximity, LUCIA can handle up to 4 additional factors. The second factor is the *suitability* of each grid cell – i.e. how the specific characteristics of each cell can support a given land-use. The third factor is *accessibility* – i.e. access to the transportation network. Some activities, like shopping, require better accessibility than recreational activities, for example. Often the latter activity even feels attracted to areas with low accessibility for example due to lower noise levels in

such areas. These three headline factors – proximity, suitability and accessibility - define the basic preconditions for cell ability to support a given land-use, and are, in some degree, fixed, although the accessibility can be changed by improving the infrastructure for example. Currently, we have not used the possibility of using the last two optional factors. The overall modelling concept is illustrated in figure 3.

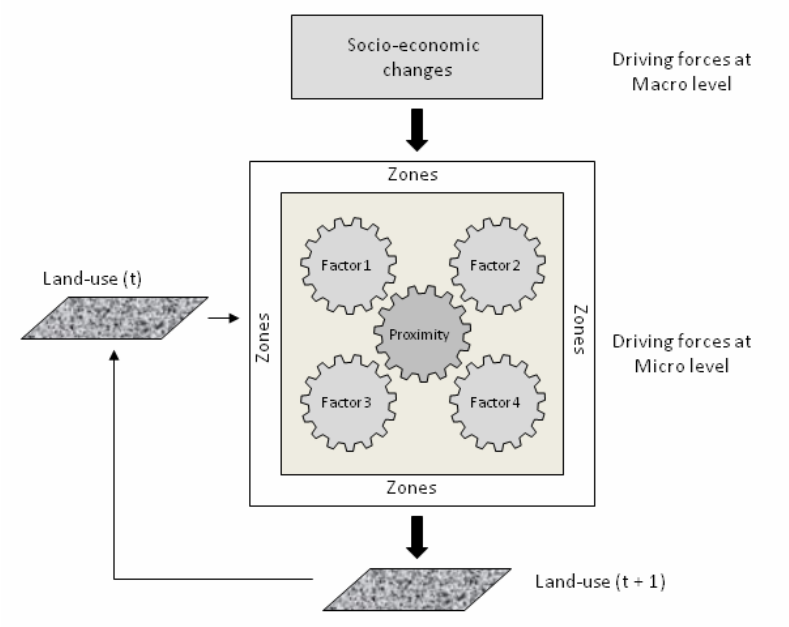


Fig. 3. Principles of the LUCIA land-use simulation model

The factors (Proximity, Suitability, and Accessibility) have dimensionless values between 0.0 and 1.0, whereas the constraints (Zoning) most often have binary values – 0 *or* 1. But if needed, continuous constraints are supported. By combining factors and constraints for each *active* (se table 1) land-use type L^*, we can estimate the transition potential P for each cell to change land-use from one type to another. Additionally, we need to incorporate the spatial distribution of the socio-economic drivers.

The possibility for each cell to change land-use type at the next time step is given by the function below:

$$P^{L^*}(t+1) = C_1^{L^*}(t) \times C_2^{L^*}(t) \times .... C_n^{L^*}(t) \times \sum_{i=1}^{5} (W_i^{L^*} \times F_i^{L^*})$$

Where

P = Transition potential
C = Constraints (values between 0.0 and 1.0 – but often 0 or 1)
F = Factors (values between 0.0 and 1.0)
W = individual weight factor between 0 and 1
L^* = land-use type.

Initially w is set to 1.0 for all factors, but during calibration the w value can be lowered to obtain a better agreement between the simulated and the real land-use for historical years.

The number of cell values to be changed during the iterations is determined by the external drivers. Once the transition potential has been calculated for all active land-uses the cell transformation process can start. Cell changes start with the cell having the highest transition potential for a certain land-use and this process proceeds downwards until the predetermined cell number changes for each active land-use category has been reached. This way of doing is in accordance with von Thünen's assumption, that (in equilibrium) land is devoted to the use that generates the highest potential profitability (Verburg et al., 2004 b). A deeper discussion of the factors and constraints involved in the conceptual model follows below.

3.2 Land-Use Simulations

Land-use simulation involves a wide range of data, and providing the data needed as well as the pre-processing it is a rather time consuming effort. The data set used in the current project is land-use data, soil type data, road network, prices of land, spatial planning regulations, population development, and a regional economic growth index.

Land-use data

The basic source for land-use information in the model is CORINE land-cover for the years 1990 and 2000. Unfortunately, the level of thematic detail in CORINE land-cover does not satisfy our requirements for the built-up areas and protected nature. Therefore we introduced two auxiliary data sets. First – and most important – we used the Danish Building and Housing register, which contains detailed information about each building in Denmark, and this register has been in operation for about 30 years. The register is geo-referenced by using the Danish Address database, and the register allocates all buildings to one of 25 use categories (Daugbjerg & Hansen). Currently we aggregate the 25 categories into five – residential, industry, service, summer cottages, and (other) recreation. Using the Danish national 100-meter square grid we summarised the built-up area for each use category within each grid cell and assigned the use having the biggest area to the cell. A further criterion is that at the number of urban cells within a 3 x 3 Moore neighbourhood must be greater than 1 - otherwise it is not considered built-up. Second, we used detailed nature type registrations to improve the spatial resolution of these sensitive areas. Several nature types were aggregated into three categories: semi-nature, wetlands and lakes. The next step was to merge the new data sets with CORINE, in order to produce the final and improved land-use layers. Thus eleven new land-use grids for each of the years 1990 until 2005 were produced. The land-use types are divided into three categories, similarly to the MOLAND model (Engelen et al., 2002; Barredo et al., 2003) for example.

Table 1. Land-use divided into categories

Active land-use classes	Passive land-use classes	Static land-use classes
Residential areas	Grass and arable land	Harbour
Industry areas	Forest	Airport
Private & public, service areas	Semi-nature	Waste and extraction sites
Summer cottage areas	Wetlands	Lakes
	Recreational areas	Sea

The most important category is active land-use types, which are forced by the demands generated externally. Another category is passive land-use types, which are not driven by an external demand, but on the other hand enter into the calculations, because cells can disappear by being transformed into one of the active land-uses. The final category is the static land-uses, which cannot be transformed into one of the active land-uses, but will nevertheless affect the land-use simulation by attracting or repelling land-use transformation within their vicinity.

The macro level driver data

The socio-economic drivers at the macro level (here municipal level) comprise factors such as: population change, industrial structure, economic development, technological change, policies and legislation. However, these conceptual drivers must be converted into demand for land for all the active land-use types, and this process is not straightforward at all. Generally we shall expect that a growing population will increase the demand for residential purposes, and this is usually correct. But what about static and even declining populations – will this situation free cells from residential to other purposes? Not necessarily. Within the study area the number of people rose marginally from 571956 to 580190 – corresponding to 1,4% - during the calibration period from 1990 to 2000, but some sub-areas have experienced a decline in population – and at the same time an enlarged residential area! This reflects the so-called thinning out effect, where each dwelling unit houses fewer and fewer people.

Thus the demand for more space for residential purposes should not only consider population growth but also this thinning out effect. Similarly, the relationships between economic growth and the demand for land for industry and service facilities are not easy to resolve. The economic growth will normally require bigger factories, but often production processes become more effective, or the factory moves from a central location (often near the harbour) to a new location at the urban fringe and near motorway junctions. This requires new space for industrial purposes, but at the same time frees their original central locations for other purposes – often residential. A parallel process can be observed for many service facilities. The economic growth in the current study is based on the regional economic growth index from Denmark Statistics. It is available from 1993 onward. The index figures for 1990 – 1992 are estimated by linear extrapolation.

4 Land-Use Scenarios

The world is full of uncertainty, and dynamic processes in the coastal zone – both man-made and natural - are interdependent and complex. Furthermore, uncertainty regarding climate change, its impacts and adaptive processes is so impressive that very little can be said yet with confidence about adaptive capacity and therefore also about vulnerability to long-term climate change. It is possible, however, to make statements about the expected outcomes with a reasonable level of certainty. Scenario testing can bring the complexity of coastal interactions into focus and provide a better knowledge base for decisions. Scenarios can also help to incorporate a long-term view and to illustrate and explain issues to stakeholders and the general public during the planning process. We carried out three different scenarios for land-use development in Northern Jutland, and analysed the impact of future climate change and sea level rise on the expected future urban land.

4.1 Scenario Descriptions

The first (baseline) scenario is just a projection of current trends in land-use development, whereas scenarios 2 and 3 simulate an urban development as presumed in story lines A2 and B2, which – regarding climate change - represent a pessimistic and an optimistic SRES scenario, respectively. Although it is never a trivial task to make detailed projections on future land-use, we can at least simulate the future urban development by a simple linear projection based on past trends. It is much more complicated to transform the SRES storylines for the future societal development into expectations about future urban development. Solecki and Oliveri (2004) have analysed the SRES scenarios A2 and B2 to identify 'spatial' elements, which can be used in urban land-use modelling. Regarding A2 they conclude that there will be a growth in per capita land-use conversion (Development along road corridors and growth associated with new sub-urban, peri-urban employment centres as important drivers). For B2 they expect a decrease in per capita land-use conversion and minimal minimum spontaneous growth as well as fewer new spreading centres. Compact growth with infilling and edge growth will take place adjacent to existing urban spaces.

Baseline scenario: This scenario is a 'simple' projection based on model parameters, which are calibrated using the period 1990 – 2000. The future demands for land needed for housing, industry and service use are based on existing official population projections at municipality level until 2030. The expected demand for new summer cottages is basically dependent on the economic growth, and – of-course - no official projections are available. Therefore we assume the future yearly growth to be equal to observed average demand for the period 1990 – 2000. The constraints represented by the Danish spatial planning regulations are crisp (values 0 or 1).

Policy scenario A: This scenario corresponds to the A2 emission scenario. Generally we have very strong spatial planning regulations, but it is expected that spatial

planning is weak under the A2 narrative. Therefore, we have changed the spatial planning constraints, so it is partly legal to create new settlements in areas, which are prohibited according to the base line scenario. The 0 and 1 constraint values in the baseline scenario are replaced by 0.4 and 0.6 respectively in policy scenario A. Furthermore, the weight associated with the proximity factor is reduced of 50% compared to the baseline scenario in order to impede edge growth. Finally, we have added a random raster (values between 0 and 0.3) to the suitability map to support spontaneous growth. The drivers are similar to the baseline scenario, but the estimated demands for land are increased by 10% to reflect the expected higher per-capita land consumption as expected in A2.

Policy scenario B: This scenario is rather similar to the baseline scenario, but it represents a more sustainable societal development as outlined by SRES scenario B2. As mentioned earlier we have already very strict spatial planning rules, which are in line with the B2 narrative. Thus, compared with the baseline scenario the only difference is an expected decrease of 10% in per-capita land consumption.

4.2 Land-Use Simulations

The simulation period is 25 years from 2005 to 2030, and the cell size applied is 100 x 100 meter. Calibration of factor weights was carried out using a ten-year period from 1990 to 2000. The model contains 1905 x 1370 = 2,609,850 grid cells, and using a powerful PC it takes about 15 minutes to carry out a 25 years land-use simulation. The LUCIA modelling framework is an integrated stand-alone application supporting several aspects of land-use modelling and developed in Delphi. Below (figure 4 and 5) you can see the differences between the two policy scenarios for Aalborg city, which is the regional capital of Northern Jutland.

Policy scenario A (figure 4) shows a much more scattered urban development compared to scenario B. We can observe as well a more branched urban structure as real new settlements in the countryside. This is in accordance with what we can expect from the A2 narrative. Some obvious examples are enclosed in ellipses. On the contrary, figure 5 shows the future land-use according to policy scenario B, and we can observe a more nucleated city produced by filling in gabs and holes. Referring to narrative B2 this is clearly what we can expect.

Thus we are able by using LUCIA to simulate the future urban development in ways, which have obvious similarities to our expectations in the two SRES emission scenarios. Nevertheless, it is only in scenarios for possible future development, and with changed parameters and input that the simulations will be different. However, we can conclude that the simulations are in accordance with our expectations qualitatively, and in these way useful inputs in the decision-making concerning future land-use and spatial planning. Below, we will demonstrate how the developed scenarios can be used in a coastal planning context.

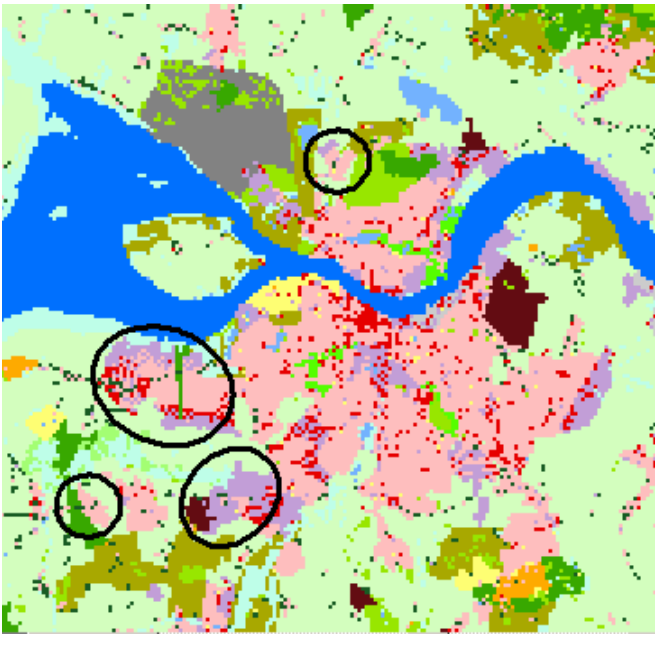


Fig. 4. Land-use in 2030 according to Policy scenario A

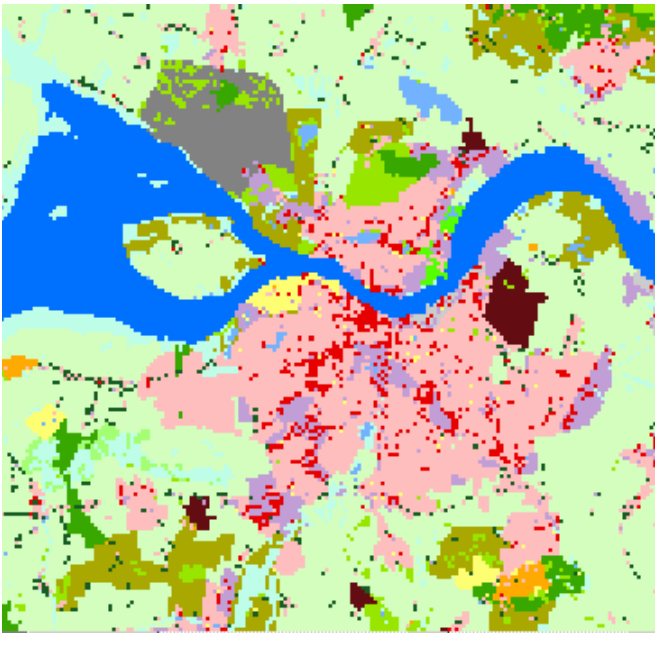


Fig. 5. Land-use in 2030 according to Policy scenario B

4.3 Impact Analysis

Aiming at analysing the possible consequences of climate change, the policy scenario A were overlaid by the two flooding zone maps representing sea level rise (80 cm) and sea level rise plus the effect of storm surges (80 + 200 cm), respectively. On figure 6 the 80 cm sea level rise is shown as cross-hatched polygons, whereas the combined effect of sea level rise and storm surge is illustrated by hatched polygons. Remark, that the hatched and cross-hatched areas represent areas that might be flooded in year 2100! Overall we can conclude that only minor patches of land along the fjord can expect permanent flooding due just to sea level rise, but with addition of storm surges much larger areas are in danger. From the map on figure 1 you can see that the impact has its maximum in some low-lying wetlands areas, which can be severely damaged.

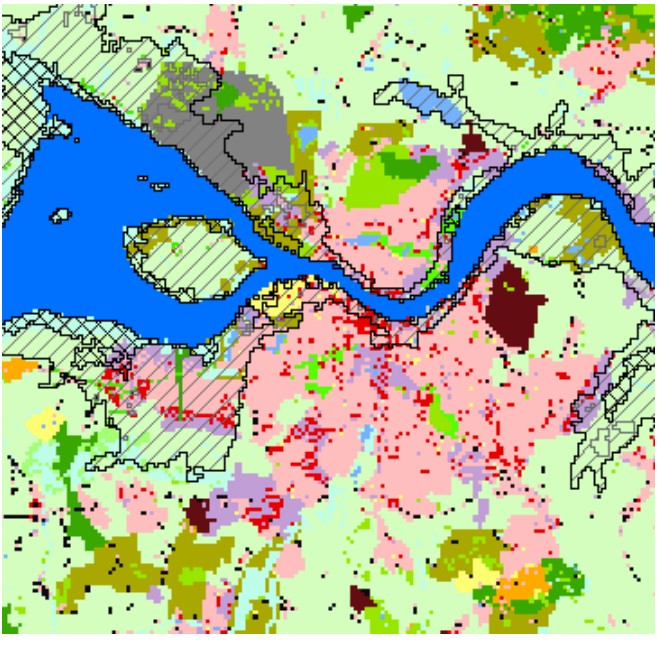


Fig. 6. The effect of sea level rise and storm surge around Aalborg city. The cross-hatched polygons represent sea level rise, and the hatched polygons represent the combined effect of sea level rise and storm surge.

But if you zoom in on the major city of Aalborg, you can clearly see that particularly the central part of the city is located in a flood-prone zone (fig. 6). The storm surge impact on the central administrative and shopping district will increase the vulnerability and threaten its function as capital for the Northern Jutland Region. Clearly, this area must be protected for example by dikes, when it is needed. However another potential threat can be seen just west of the city. Here

we can expect extensive urban development during the next 25 years- and this development has already started! Nevertheless, it is an area vulnerable to frequent flooding at the end of this century. Therefore, spatial planners need to take action to prevent the potential risks, and in this adaptation process the described land-use modelling effort can be a very useful tool.

Table 2. Percentage of vulnerable new cells for the two policy scenarios

Land-use Scenario	Land-use	Sea Level Rise	SLR + Storm surge
Policy	Housing	0 %	17 %
Scenario A	Industry	2 %	19 %
	Service	2 %	27 %
	Summer house	9 %	24 %
Policy	Housing	0 %	29 %
Scenario B	Industry	0 %	2 %
	Service	0 %	0 %
	Summer house	9 %	24 %

The number of vulnerable buildings within each land-use type was summarised in table 2. It is obvious that generally rather few new built-up areas in year 2030 will be threatened by permanent flooding through mere sea level rise. This is the case for both policy scenarios. Only summer cottages which are often located in low-lying areas near the coast are more vulnerable to flooding. However, taken storm surges into consideration will enlarge the vulnerability substantially for all urban land-uses when we consider policy scenario A. Thus about one fifth of the new built-up areas are located in flood prone areas. Regarding policy scenario B we can see that especially residential buildings and summer cottages will be located within the risk zone. Results like this are useful for spatial planners and coastal zone managers in the preparation phase for new land-use plans.

5 Conclusions

Climate change has received much attention during the last decades and not at least various mitigation and adaptation strategies. The coastal zone will clearly feel the consequences of climate change and the associated effects like sea level rise, increased storminess and flooding. In order to mitigate the negative consequences of this development the European Union has defined a set of recommendations for integrated coastal zone management, and several pilot studies have emphasised the use of spatial models and scenarios to support the decision-making concerning a sustainable coastal zone. Besides this, the European Commission has adopted a so-called Green Paper on adapting to the impacts of climate change.

Land-use models are useful for unravel the complex collection of socio-economic and biophysical forces which determine the rate and spatial pattern of land-use change. Using modelling and simulation, we can increase our understanding of the

future land-use system under influence of a changing climate and accordingly reduce uncertainty concerning decisions. Minimising the impacts of sea-level rise and flooding from storm-surge events can be achieved through implementing adaptation strategies - including a wide spectrum of approaches, from policy and law to engineering and technology.

The current paper has demonstrated how to integrate a cellular automata based land-use simulation model with climate change scenarios, and thus facilitate the identification of future vulnerable areas. Currently, we have transformed the A2 and B2 SRES narratives to quantifiable elements to be handled in a land-use simulation model. We have found that the LUCIA modelling approach is well suited for this kind of modelling.

This is our first step to integrate the climate change scenarios with land-use modelling within an operational tool. However, we are aware that the extent of future climate change is still uncertain, and not at least our translation of the SRES scenarios to changed conditions for urban development. Nevertheless, governments at various levels must be aware of the potential effects of climate change, and consider the consequences in their planning efforts. The current research has been primarily theoretical and lacking any practical planning connection. Therefore, the next step will be to use the developed methods and tools in a practical coastal zone planning context in Northern Jutland. This will hopefully give us feedback with ideas for further developing our modelling framework. Not at least it is our hope that we can use the developed methods in a public participation context.

References

Barredo, J.I., Kasanko, M., McCormick, N., Lavalle, C.: Modelling dynamic spatial processes: Simulation of urban future scenarios through cellular automata. Landscape and Urban Planning 64, 145–160 (2003)

Danish Coastal Authority, Consequences of Climate Change for the Danish Coasts (2007) (in Danish)

Daugbjerg, P., Hansen, K.V.: Property Data. In: The Danish National Survey and Cadastre. Copenhagen (2000) (in Danish)

Engelen, G., White, R., Uljee, I.: The MURBANDY and MOLAND models for Dublin. Final report, RIKS (2002)

European Environment Agency, Vulnerability and adaptation to climate change in Europe. EEA Technical report no. 7 (2005)

European Commission, Communication from the Commission to the Council and the European Parliament on integrated Coastal Zone Management: A Strategy for Europe (2000)

European Commission Directive, 2001/42/EC of the European Parliament and of the Council of June 27 2001 on the assessment of the effects of certain plans and programmes on the environment (2001)

European Commission Directive, 2007/60/EC of the European Parliament and of the Council of October 23 2007 on the assessment and management of flood risks (2007)

Hansen, H.S.: A quasi-four dimensional database for the built environment. In: Westort, C.Y. (ed.) DEM 2001. LNCS, vol. 2181, pp. 48–59. Springer, Heidelberg (2001)

Hansen, H.S.: An Adaptive Land-use Simulation Model for Integrated Coastal Zone Planning. In: The European Information Society. Lecture Notes in Geoinformation and Cartography, pp. 35–53 (2007)

Hansen, H.S.: Quantifying and Analysing Neighbourhood Characteristics Supporting Urban Land-use Modelling. In: The European Information Society. Lecture Notes in Geoinformation and Cartography, pp. 283–299 (2008)

Intergovernmental Panel on Climate Change. Climate Change 2001. The scientific basis. Cambridge University Press, Cambridge (2001)

Meier, H.E.M., Broman, B., Kjellström, E.: Simulated sea level in past and future climates of the Baltic Sea. Climate Research 27, 59–75 (2004)

NORCOAST, Recommendation on improved Integrated Coastal Zone Management in the North Sea Region. County of Northern Jutland (2000)

Nakicenovic, N., et al.: Special Report on Emissions Scenarios: A Special Report of Working Group III of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge (2004)

Selecki, W.D., Oliveri, C.: Downscaling climate change scenarios in an urban land use change model. Journal of Environmental Management 72, 105–111 (2004)

Tobler, W.R.: A computer movie simulating urban growth in the Detroit region. Economic Geography 46, 234–240 (1979)

Vedung, E.: Public Policy and Program Evaluation. Transaction Books (2000)

Verburg, P.H., Schot, P., Dijst, M.J., Veldkamp, A.: Land use changes modelling: current practice and research priorities. Geo. Journal 61, 309–324 (2004a)

Verburg, P.H., van Eck, R., Nijs, T., Dijst, M., Schot, P.: Determinants of land-use change patterns in the Netherlands. Environment and Planning B 31, 125–150 (2004)

A Multiple Criteria Heuristic Solution Method for Locating Near to Optimal Contiguous and Compact Sites in Raster Maps

Pablo Vanegas[1,2], Dirk Cattrysse[1], and Jos Van Orshoven[3]

[1] Centre for Industrial Management, Katholieke Universiteit Leuven,Celestijnenlaan 300A, 3001 Heverlee – Leuven, Leuven - Belgium
e-mail: {Pablo.Vanegas,Dirk.Cattrysse}@cib.kuleuven.be
[2] Facultad de Ingeniería, Universidad de Cuenca, Cdla. Universitaria s/n, Cuenca - Ecuador
e-mail: Pablo.Vanegas@ucuenca.edu.ec
[3] Department of Earth and Environmental Science, Katholieke Universiteit Leuven, Celestijnenlaan 200E, 3001 Heverlee – Leuven, Leuven - Belgium
e-mail: Jos.VanOrshoven@biw.kuleuven.be

Abstract. A high performance heuristic solution method is proposed able to locate near to optimal sites composed by a given number of cells (raster structure). These sites must be compact and maximize levels of the sites intrinsic multiple criteria suitability. To validate the heuristic approach, a comparison with a mathematical formulation is performed with afforestation data of regions within the Netherlands, Denmark, and Flanders. This reveals that the heuristic is considerably faster than the mathematical method and the objective values obtained with the two approaches are substantially similar. A sensitivity analysis shows that the region's homogeneity plays an important role in the performance of the process identifying most favourable sites. Moreover, computation time follows a power model in the number of cells forming the site.

Keywords: Site Location, Heuristic, Exact Methods.

1 Introduction

In the practice of environmental conservation and land use planning sites are searched for fulfilling shape (e.g. contiguity, compactness) and size requirements, satisfying also particular attribute criteria. This paper presents a mathematical formulation and a heuristic solution method for identifying compact and contiguous sites formed by a set of cells maximizing the intrinsic multiple criteria suitability of each site. The notion of compactness is associated with firmly packed sites. In this paper we adhere to *the earliest attempts to develop a compactness index based on perimeter to area ratios* (Maceachren, 1985). On the other hand, a

B. Murgante et al. (Eds.): Geocomputation, Sustainability & Environ. Planning, SCI 348, pp. 35–56.
springerlink.com

site is contiguous if one can walk from an identified parcel to another without leaving the site (Xiao, 2006). Therefore, compactness implies contiguity, but not the opposite.

Contiguity and compactness requirements and intrinsic multiple criteria suitability of the cells are all involved in finding optimal sites. Thus, Multiple Criteria Decision Analysis (MCDA) is included as part of the proposed mathematical and heuristic approaches for site location.

Although *mathematical optimization methods have been used for about 30 years in areas like forest planning* (Williams and ReVelle, 1997), the size of the problems that can be handled in a practical way remains limited. Nevertheless, these methods can be very useful as part of an adaptative, learning process of problems at hand (Hof and Bevers, 2000). Therefore, the proposed heuristic for site location is evaluated taking the mathematical formulation results as a reference. Even though the proposed mathematical formulation by itself does not ensure contiguity, it makes use of the benefit gained through the boundaries of contiguous pixels, tending to minimize the perimeter while the area is constant. A weight configuration giving the same importance to each criterion fulfills the contiguity and compactness requirements.

This paper is organized as follows: Section 2 reviews techniques applied in site location problems; Section 3 describes the data and approaches used in this study,; Section 4 analyzes the results and Section 5 summarizes the conclusions.

2 Literature Review

2.1 Exact Methods for Site Location Problems

Exact methods include enumeration and mathematical programming, as well as many specialized algorithms that have been developed for particular optimization problems (Williams and ReVelle, 1997).

2.1.1 Mathematical Programming Methods

Several approaches for site location problems make use of Linear Programming (LP) models where the variables are integers (Integer Programming - IP). LP/IP attempts to maximize (or minimize) a linear function, constraining the values of the decision variables. Hof and Bevers (2000) formulated four linear programming examples including constraints to avoid adjacency in order to account for biological dispersal. The constraints relate population in a habitat area in a time period (*t*) to the populations in other areas in a previous time period (*t-1*), while taking into account the population growth and the immigration dispersion. Church and ReVelle (1974) introduce the Maximal Covering Location Model (MCLM), which minimizes the number of facilities to cover each and every demand point on a network. MCLM is modified (Dimopoulou and Giannoikos, 2001) to determine the optimal deployment of available fire-fighting vehicles. In order to reduce the vulnerability of elements like species, communities, or endemic plants,

Church et al, 1996 develop a mathematical model for selecting sites for diversity conservation (Biodiversity Management Areas - BMAs). Since the solutions are composed of isolated planning units, to avoid fragmentation, Fischer and Church, 2003 formulate a mathematical model including the objective of minimizing the outside perimeter of the selected areas. The idea to represent a mosaic of *n* cells as a planar graph with vertices and edges is presented by Williams (2001; 2002). Each cell is equated with a vertex, and each adjacency relation between a pair of cells is equated with an edge. Shirabe, 2004 also applies this idea to formulate the necessary and sufficient conditions for assembling a connected region with a desired degree of perforation, from no hole to a largest possible number of holes.

2.1.2 Enumeration Methods

Enumeration methods evaluate all candidate solutions (explicit enumeration - brute force), or identify a set of efficient solutions (implicit enumeration), and select the one that optimizes specific criteria. Since the computational cost of this sort of search is proportional to the number of candidate solutions, it is typically used in problems of limited size. Gilbert et al (1985) develop an interactive multiobjective algorithm for allocating an area of land (set of cells). This approach partially generates a set of efficient solutions that achieve four objectives: minimize development cost, minimize the distance to desirable cells (amenity distance), maximize distance to undesired cells (detractor distance) and minimize a shape objective expressing contiguity. Similarly, Diamond and Wright (1991), apply an implicit enumeration method, which is based on irregular grids to generate multiobjective contiguous sites. The tracts of sites maximize the level of suitability regarding cost, area, and shape.

2.2 Approximate Methods for Site Location Problems

Heuristic programming techniques have been developed to solve problems of which the solutions each constitute just one point or a small set of points in a very large, and possibly infinite space, the search space (Siklossy and Marinov, 1971). A heuristic is a problem-specific way of directing problem solving. *It seeks to obtain good, that is, near-optimal solutions at relatively low computational cost without being able to guarantee the optimality of solutions* (Dorigo, 2004).

2.2.1 Heuristic Approaches

To deal with the problem of generating contiguous and compact districts while providing population equality and retaining jurisdictional boundaries, Mehrotra et al (1998) develop an optimization-based heuristic capable of considering many potential districts. The problem is represented with a graph, where each node is associated with the population of a county (unit), and an edge exists when two geographical units are neighbours. A penalty cost is assigned to every potential district, measuring its *deviation* from an ideally compact district. To explicitly

manage the shape in site location problems, Brookes (2001b) proposes the Parameterized Region-Growing (PRG) process, which starts from a seed cell. This algorithm is a fusion of two ideas: Simple Region-Growing (SRG) and Parameterized Shape-Growing (PSG). The SRG algorithm iteratively adds the most suitable neighboring cell. If two or more cells have equal suitability then the one closest to the seed is chosen. The PSG algorithm uses the same incremental process as the SRG but with a shape-suitability score determined by the distance and direction of the cell to the seed. PRG combines PSG and SRG through a weighted average of the two scores. The suitability of a cell S with shape score *Ss* and underlying cell suitability *Sc*, when the trade-off is T, is given by *S=[T*Ss]+[(1-T)*Sc]* (Brookes, 1997). PRG generates promising regions with a specific shape when an operator chooses the approximate location, shape and orientation of the regions. Nevertheless, an appropriate parameter setting is required. Church et al (2003) develop a Patch Growing Process (PGP) to generate feasible patches with reference to a seed cell. Once the seed patch is defined, the neighbors to the patch are placed on a list in a random order. Each cell in the list is analyzed in terms of the number of edges (*e*) that it shares with the current patch (from 1 to 4). The composite suitability of the i_{th} cell (CS_i) is defined by $CS_i = Suit_i + N.e_i$, where $Suit_i$ is the suitability value of the cell itself, *N* the weight attached to the number of edges shared with the existing patch, and e_i the number of edges that the i_{th} cell shares with the current growing patch. Then the list of neighboring cells is ordered according to the composite suitability, and the top *X* percent of the cells on this list are added to the patch.

2.2.2 Metaheuristic Approaches

When the heuristics are general-purpose methods that can assist in finding high quality solutions, those are called metaheuristics. Genetic algorithms and simulated annealing are frequently applied metaheuristics in site location problems.

Stewart et al (2005) propose the integration of Genetic Algorithms (GA) with a formulation for Reference Point Goal Programming (RPGP). The general formulation of the objective function for RPGP considers in the first part simple additive attributes which associate costs or benefits with the allocation of any particular land use to a specific cell. In the second part it considers spatial attributes indicating the extent to which the different land uses are connected or fragmented across the region. The PRG approach proposed by Brookes (2001b), described earlier in this paper, is combined with a Genetic Algorithm for Patch Design (GAPD) (Brookes, 2001a) in order to explicitly handle dynamic and static criteria. Additional approaches applying Genetic Algorithms for the site location problem are proposed by Li and Yeh (2004), to deal with multiple criteria evaluation; and by Xiao (2006) to search locations represented under vector structures.

Other metaheuristics have also been applied in site location, Aerts and Heuvelink (2002) apply simulated annealing to allocate *N x M* pixels with *K* different types of land use. McDonnell et al (2002) compare greedy search and simulated annealing methods to construct spatially cohesive reserve areas. While the simulated annealing approach is similar to the one by Aerts and Heuvelink, 2002, the greedy search algorithm updates an existing solution by adding one unreserved site.

2.3 Multiple Criteria Decision Analysis

The term 'Multiple Criteria Decision Analysis (MCDA)' is used by Belton and Stewart (2002) as an umbrella term to describe a collection of formal approaches that seek to take explicit account of multiple criteria in helping individuals or groups exploring decisions that matter. Many site location problems involve also criteria coming from different actors, and selected sites must comply with more than one objective. To deal with these issues, several approaches for site search include Multiple Criteria Decision Analysis (MCDA); *where a criterion is a generic term including the concepts of attribute and objective* (Malczewski, 1999). While attributes are the properties of elements of a real world, an objective is a statement about the desired state of a system under consideration (Malczewski, 1999). Xiao (2006) states that when objectives conflict, it is often impossible to find a single optimum that dominates all other solutions. One of the dichotomies present in MCDA is Multiple Attribute Decision Making (MADM) versus Multiple Objective Decision Making (MODM). While MADM obtains levels of attributes through preferences in the form of objective functions and attribute weights, in MODM these levels are derived from the preferences among objectives and from the functions relating attributes to objectives (Malczewski, 1999). Several approaches for site location make use of MCDA. Li and Yeh (2004) apply MADM while Gilbert et al (1985), Stewart et al (2005) and Xiao (2006) apply MODM.

3 Materials and Methods

First part describes the process to prepare the data in order to make it usable by the Integer Programming (IP) as well as for the Heuristic formulation. Next, both approaches are described. Since the problem at hand considers levels of importance (weights) for the criteria involved in the location of optimal sites for afforestation of agricultural land, a MADM approach is applied in the current work. However, since the paper is dealing with general formulations for site location, it does not evaluate alternatives for weights computation. The weights are in fact considered as mathematical parameters from the optimization point of view. This condition does not restrict the applicability and generality of the proposed approach.

3.1 Data Preparation

Both, the IP formulation and the HMSL are tested with data (criteria) generated and validated in the AFFOREST project (EU 5th Framework Programme for Research and Technological Development), pertaining to the Netherlands, Denmark, and the region of Flanders in the North of Belgium. In this sense, a validation of the input data is out of the scope of the present contribution. Moreover, it is believed that the proposed approaches for site location can be applied on grids of cells with attribute data generated by any other application field.

In AFFOREST, the Environmental Performance (EP) is defined (Gilliams et al, 2005) as the combination of three Environmental Impact Categories (EIC): *total carbon sequestration* (composed by carbon sequestration in biomass and in soil), *nitrate leaching*, and *ground water recharge*. The EP for each land unit, represented by a class of cells *or pixels*, as a function of *time* after afforestation, is computed by means of a metamodel (details of the Afforest project can be found in Heil et al, 2007). While a *pixel class* represents all the cells that have the same characteristics (initial agricultural land use, soil type, annual average of precipitation amount, annual average of nitrogen deposition), *time lag* is a period (year) after afforestation. In this sense, *afforestation* is defined as the transformation of agricultural land into forest.

Environmental Performance (EP) criteria are applied to test the mathematical and heuristic approaches for site location. Each criterion is represented through a raster map upholding information of one EIC (carbon sequestration, nitrate leaching or ground water recharge), at a specific year after afforestation, and for a specific afforestation strategy (e.g. total carbon sequestration after 10 years of afforestation with oak under medium stand preparation and medium stand tending levels).

For the purpose of this study, each criterion (raster map) is first *normalized* (NC) in order to avoid the influence of the differences in value ranges; each pixel upholds a value between 0 and 1 (1 = the best). Since the Environmental Performance (EP) can be defined by one or more EICs, a multi-criteria objective function is required in order to find a set of pixels maximizing the involved EICs. It is essential to consider that *maximization of EP* implies *maximization* of carbon sequestration (CS) and ground water recharge (GWR), but *minimization* of nitrate leaching (NL) levels. Accordingly, two normalization functions are defined: equation 1 is applied when the highest EP is obtained as a maximization, and equation 2 when the highest EP is obtained as a minimization. In these equations, u refers to a specific EIC layer. While $f(u)$ is the pixel value in the u layer, γ_u and σ_u are respectively the maximum and minimum value within the u layer (i.e. within each criteria).

$$M_{\max} = \frac{f(u) - \sigma_u}{\gamma_u - \sigma_u} \tag{1}$$

$$M_{\min} = 1 - \frac{f(u) - \sigma_u}{\gamma_u - \sigma_u} \tag{2}$$

3.2 Integer Programming (IP) Formulation for Multiple-Criteria Site Location

In a first stage, a 0-1 linear programming formulation is developed for acquiring optimal sites composed by a set of pixels forming a compact and contiguous site with maximal performance according to one or more weighted *Normalized*

Criteria (NC). Compactness is achieved by increasing the number of boundaries that a selected cell shares with other cells that are also selected as part of the solution. This formulation implies the reduction of the patch perimeter, while the area is constant.

In the first part of the objective function (equation 3), $s_{i,j}$ is a binary variable upholding a value 1 if the cell in row *i* and column *j* is part of the target site. While *P* is the number of criteria considered in the decision problem, w_k is the importance weight for the k_{th} criterion, and $c_{k,i,j}$ upholds the value for the k_{th} NEIC layer in the *i,j* cell.

The second part of the objective function deals with the compactness criterion, where w_N is the importance weight assigned to this requirement. To construct a compact site, the model considers the Von Neumann neighborhood of a cell (upper, down, left and right cells). The binary variable $u_{i,j}$, will uphold a value 1 if the *i,j* cell is selected as part of the target site ($s_{i,j}=1$), and at the same time the upper neighbor cell is *also* selected. The same judgment is applied for assigning values to the binary variables $d_{i,j}$, $l_{i,j}$, $r_{i,j}$, which are associated to the *d*own, *l*eft, and *r*ight neighbor cell respectively. Since the objective function is dealing with normalized values, and applies the Von Neumann neighborhood, the sum of $u_{i,j}$, $d_{i,j}$, $l_{i,j}$, and $r_{i,j}$ is multiplied by a 0.25 factor. In this manner, while the maximum level regarding the compactness contribution of the *i,j* cell is 1 (four neighbors of $s_{i,j}$ are *also* selected), the minimum level is 0 (no neighbors of $s_{i,j}$ are *also* selected as part of the solution).

maximize:

$$\sum_{k}^{P}(w_k * \sum_{i}^{m}\sum_{j}^{n} s_{i,j} * c_{k,i,j}) * w_N * \sum_{i}^{m}\sum_{j}^{n} 0.25 * (u_{i,j} + d_{i,j} + l_{i,j} + r_{i,j}) \quad (3)$$

Subject to:

$$s_{i,j} - x_{i,j} \geq 0 \quad \forall_{i,j}; \quad \forall x_{i,j} \text{ where } x_{i,j} \in \{u_{i,j}, d_{i,j}, l_{i,j}, r_{i,j}\} \quad (4)$$

$$s_{i-1,j} - u_{i,j} \geq 0 \quad \forall_{i,j} \quad (5)$$

$$s_{i+1,j} - d_{i,j} \geq 0 \quad \forall_{i,j} \quad (6)$$

$$s_{i,j-1} - l_{i,j} \geq 0 \quad \forall_{i,j} \quad (7)$$

$$s_{i,j+1} - r_{i,j} \geq 0 \quad \forall_{i,j} \quad (8)$$

$$m_{i,j} - x_{i,j} \geq 0 \quad \forall_{i,j}; \quad \forall x_{i,j} \text{ where } x_{i,j} \in \{u_{i,j}, d_{i,j}, l_{i,j}, r_{i,j}\} \quad (9)$$

$$\sum_{i}^{m}\sum_{j}^{n} s_{i,j} = M \quad (10)$$

$$s_{i,j} \in \{0,1\} \tag{11}$$

$$u_{i,j} \in \{0,1\} \tag{12}$$

$$d_{i,j} \in \{0,1\} \tag{13}$$

$$l_{i,j} \in \{0,1\} \tag{14}$$

$$r_{i,j} \in \{0,1\} \tag{15}$$

The first set of constraints (equation 4) guarantee that $u_{i,j}$, $d_{i,j}$, $l_{i,j}$, or $r_{i,j}$ can have a value 1 if and only if the *i,j* cell is *also* selected ($s_{i,j}=1$). Constraint 5 assures that if $u_{i,j}$ is equal to 1, the matching cell in the set of binary variables *s*, will be necessarily equal to 1 ($u_{i,j} = s_{i-1,j}$). Constraints 6 to 8 articulate the previous rule for variables $d_{i,j}$, $l_{i,j}$, and $r_{i,j}$. The integrality constraints in equations 5 to 8 tend to effectively reduce the perimeter in order to form compact and contiguous areas. These constraints evade fragmentation. Since some cells are not eligible (no agricultural land, so no EP-data available) as part of the target site, a mask indicating the availability of the cells is expressed with the set of variables *m*; therefore if the *i,j* cell is available, the variable $m_{i,j}$ will uphold a value 1, and 0 otherwise. The set of constraints in equation 9 avoid selecting unavailable cells as part of the solution, and equation 10 restricts the number of selected cells to be equal to the predefined number *M*.

3.3 Heuristic for Multiple-Criteria Site Location (HMSL)

An alternative solution method is developed to locate feasible (near to optimal) sites that fulfill multiple criteria requirements. This solution method is based on the heuristic approach developed by Church et al (2003), introducing three main differences with respect to the original process: (*1)* In order to remove the influence of value ranges, all criteria involved in the current problem, including the suitability rewarding compactness, are normalized. After this process, all raster maps representing criteria use the same measuring scale, each cell upholding values between 0 and 1. Normalization allows all criteria compete among them under the same conditions; (*2)* The proposed heuristic for site location (HMSL) makes use of objective functions considering weighted multiple criteria (MADM). In equations 16 and 17, the denominator act as a weight normalization factor. They control that the increase in importance (weight) of one criterion implies also the decrease of importance of the others; (*3)* A number of seed patches are automatically generated, and the region growing process is repeated to produce several candidate solutions.

The method developed here is a Multiple-Criteria Heuristic solution method for Site Location (HMSL), which is divided in 3 stages: *1)* seeds generation, *2)* region growing, and *3)* region ranking. The final goal of the entire algorithm (figure 1) is

to construct a compact site maximizing its intrinsic multiple criteria suitability as derived from the suitability of the member cells. Environmental Performance (EP) criteria are applied to test this method as well as the mathematical formulation.

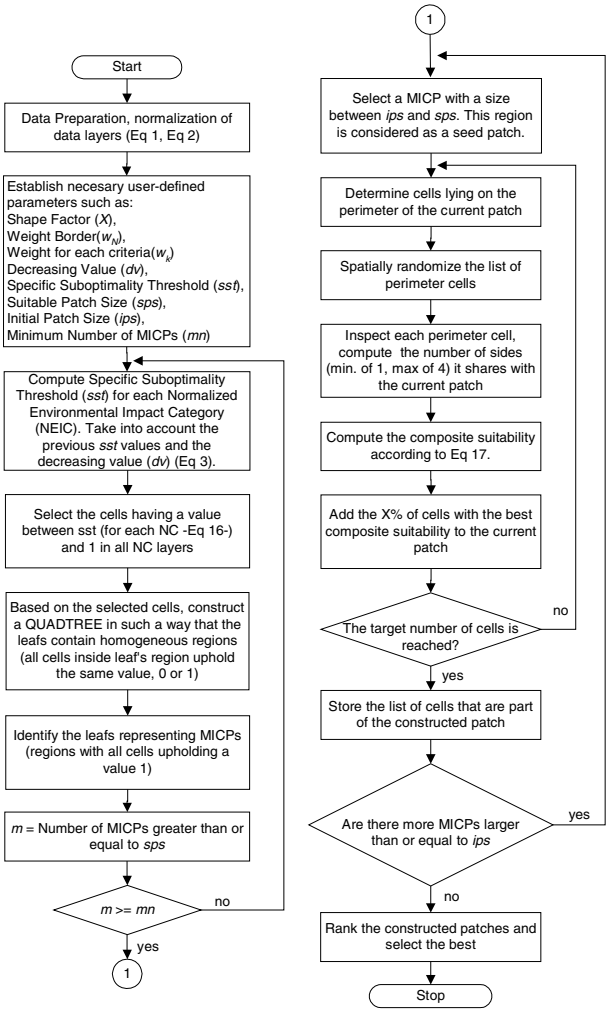


Fig. 1. General Flow Diagram of the Heuristic for Multi-Criteria Site Location (HMSL)

3.3.1 Generation of Seed Patches

The foundation of seed patch generation is with interval goal programming and on quadtree-based search. While the first deals with multiple criteria analysis, the second technique is applied for searching Maximum Initial Compact Patches (MICPs). An individual MICP is defined as a square area composed by a set of *n x n* cells.

Multiple Criteria Analysis

The first step to generate seed patches (MICPs) is selecting a set of individual cells with maximal intrinsic multiple criteria suitability. To achieve this objective, goal programming with intervals (Charnes and Collomb, 1972; Ignizio, 1974) is applied. A pixel is considered optimal when it upholds the highest levels for each Normalized Criteria (NC) considered as part of a specific decision problem. But when the maximum values do not coincide in the same pixel, suboptimal pixels need to be selected through an iterative process.

Those suboptimals are identified by a gradual increase of a suboptimality tolerance with respect to 1 (1 = maximum NC value). Therefore, a *specific suboptimality threshold (sst)* for each NC is computed, starting at an initial value 1, and iteratively decreasing in a delta value *dv*. The *sst* is computed according to equation 16, where $psst_{nc}$ refers to the previous *sst* value (started at 1), w_{nc} is the weight of a specific NC, *dv* the decreasing value, and $\sum_k w_K$ is a normalization factor summing all weights. Equation 16 shows that the weight of the NC determines how fast sst_{nc} decreases. This process allows selecting all the cells having each NC level between the corresponding *sst* value and 1 (suboptimal cells). Consequently, the increase in tolerance implies the decrease of the *specific suboptimality threshold* (*sst*). In every iteration, the multi-criteria analysis produces a map with selected cells upholding a value 1, and with unselected cells upholding a value 0. This procedure is iterated until a set of suboptimal cells (*S*) is selected, in such a way that those cells form a *minimum number* of MICPs, each one having a size greater than or equal to a predefined *suitable patch size (sps)*. Thus, if *s* is a suboptimal cell from *S*, a set $O(s) \subset S$ is a set of cells that form a MICP sizing at least *sps*. The *minimum number* of MICPs is referred as *mn* in the algorithm (figure 1), and the number of cells pertaining to each MICP measures its size. A quadtree-based search is applied in order to find the MICPs within the current selected cells.

$$sst_{nc} = psst_{nc} - \left(\frac{dv * w_{nc}}{\sum_k w_k} \right) \tag{16}$$

Quadtree Based Search of Maximum Initial Compact Patches (MICP)

In a quadtree, the nodes branch off to four children. Trees have been widely applied in search algorithms, and particularly in the present approach the quadtree allows finding homogeneous MICPs composed by cells upholding a value 1 in a binary map. While the root of the quadtree corresponds to the entire binary map, each one of the four children corresponds to one of the four quadrants within this map. Each quadrant is recursively divided in four new quadrants until all cells within a quadrant are homogeneous (all cells upholding a value either 1 or 0). Figure 2 shows an example of the process applied to build the quadtree in figure 3. The process starts using the lines marked with 1 to divide the entire region in *four* quadrants (north west -a-, north east -b-, south west -c-, south east -d-). These

lines correspond to the root node in figure 3. Since 3 (out of 4) quadrants are not homogeneous, they are divided again with lines marked as 2. The homogeneous region that is not divided corresponds to the white leaf node located left most in the second level of the quadtree in figure 3. This process is iterated until quadrants are homogeneous, and do not require to be divided again. Once this recursive process ends, homogeneous quadrants must correspond to the quadtree's leaves. Each leaf linked with regions upholding a value 1 (nodes in gray in figure 3) is considered as a potential MICP. Since the entire map can be a rectangular area instead of a squared area, each recursive child can be rectangular as well. If it is the case, a MICP corresponds to the biggest square inside the leaf's rectangular area; otherwise the complete leaf's region is a MICP. Thus, the multi-criteria analysis, together with the quadtree-based search are iterated until at least *mn* MICPs are generated, with a size not smaller than the *suitable patch size (sps)* parameter.

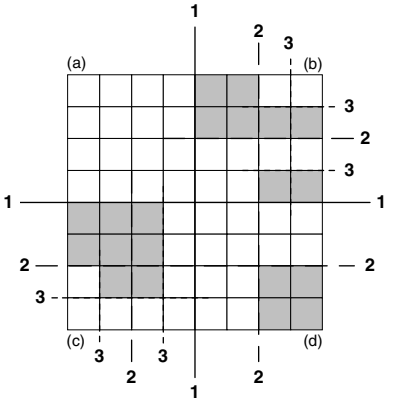


Fig. 2. Binary Image. Lines 1,2 and 3 show how the image is iteratively divided in quadrants

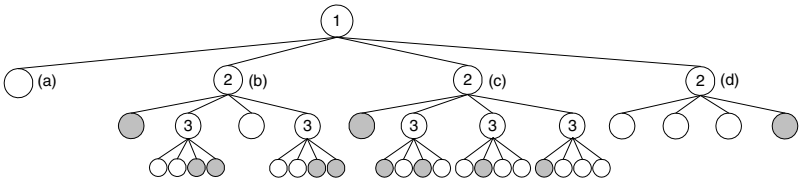


Fig. 3. Quadtree structure result of the iterative process illustrated in figure 2

3.3.2 Region Growing

Although the size of every MICP must be larger than or equal to a predefined suitable patch size (*sps*), other homogeneous regions smaller than *sps* could also be present in the quadtree's leaves. To increase the size of the search space, the region

growing process considers as seeds all regions (MICPs) greater than or equal to an *initial patch size (ips)*. The parameters *ips* and *sps* control the number of candidate solutions to be created.

The algorithm proposed by Church et al (2003) is applied to each generated seed patch. As in the mathematical formulation, the Von Neumann neighborhood allows constructing a list with the cells neighboring the seed patch, and the order within this list is randomized so that the order does not play a role in the final selection of the cells to be added to the patch. Next, a composite suitability is computed for each neighboring cell, and the list is ordered according to this suitability. Although the cells with the same composite value appear as a group, they are in random order inside the group. Finally, the best *X* % of neighboring cells are added to the seed patch to form a new patch which is used in a next iteration to add more cells until a predefined number (*M*) is achieved as part of the final site. The composite suitability equation proposed by Church et al (2003) is substituted by equation 17; where: *i* refers to each i_{th} neighboring cell; *P* is the number of Normalized Criteria (NCs) considered as part of the problem; $f(k_i)$ is the k_{th} NC value for the i_{th} neighboring cell; w_k the weight associated with this NC; and e_i is the normalized value characterizing the number of edges that the i_{th} neighbouring cell shares with the current growing patch. Since the maximum number of shared edges (*se*) is 4, the value e_i is equal to *se* multiplied by a constant value of 0.25 ($e_i = se * 0.25$); the value e_i is weighted according to w_N. As in Church et al (2003), the parameters w_N and *X* control the growing process; *X* is the percentage of cells to be added to the current patch in each iteration, and w_N weights the number of edges shared with the current patch. For convenience, a slightly different notation is used here: (1) w_N stands for weight border, and is equivalent to the parameter *N* in Church et al (2003) and (2) *X* is referred as a shape factor and has the same meaning as in Church et al (2003).

$$suit_i = \frac{\sum_{k=1}^{P} w_k * f(k_i) + w_n * e_i}{\sum_{k=1}^{P} w_k + w_N} \tag{17}$$

3.3.3 Region Ranking

Once several patches (candidate solutions) have been generated, the best solution is selected through a multi-criteria ranking procedure. Each candidate solution is represented as a binary map, where cells upholding a value 1 are part of the final site, and cells with a value 0 are not. The accumulated composite suitability for every site *R* is computed with equation 18; *M* is the total number of cells that are part of the final site, *P* is the total number of layers (NCs) taken into account, g_i the number of edges that the i_{th} cell shares with adjacent cells upholding a value 1, and $f(k_i)$ is the normalized value for the impact category *k* in the cell *i*. As before, a constant value ($c=0.25$) is applied to normalize g_i. The accumulated composite suitability is calculated for each patch, and the one corresponding with the highest value is chosen as the best.

$$suit_R = w_N * \sum_{i=1}^{M} g_i * c + \sum_{i=1}^{M} \sum_{k=1}^{P} w_k * f(k_i) \quad (18)$$

4 Results and Discussion

In a first stage the Heuristic for Multiple-Criteria Site Location (HMSL) and the Integer Programming (IP) formulation are compared using reduced areas (*70 x 70 cells*) within the Netherlands, Denmark and Flanders. Afterwards, every entire territory is processed with the HMSL, and a sensitivity analysis is carried out to determine its behavior for different parameter values. The compactness criterion and two normalized intrinsic criteria are considered in the tests: carbon sequestration (*CS*) and nitrate leaching (*NL*) after 10 years of afforestation with oak under intensive stand preparation and intensive stand tending levels. The impact categories are normalized according to equation 1 in the case of *CS*, and according to equation 2 in the case of *NL*. The tests are carried out in a computer with a Pentium-3GHz processor and 1Gb in RAM.

4.1 Comparison of the Mathematical and Heuristic Approaches

During the comparison, the tests are configured in order to select patches of 30 cells. In both approaches, the weight border (w_N), and the weights for CS and NL are constant, the value 0.33 is assigned to each criterion. Whereas the mathematical formulation is solved in LPSolve-IDE (Gourvest et al, 2006), the heuristic solution method is implemented in Java.

The Shannon entropy (equation 19), based on the number of pixel classes (table 2) is applied to measure the amount of information in each reduced region (raster map). In equation 19, n is the number of pixel classes, and $p(x_l)$ is the probability of occurrence of pixel class x_l. The Shannon entropy is used here as a measure for homogeneity, or level of uniformity in the composition of the maps. The lowest amount of information corresponds to the highest homogeneity, and the highest amount of information with the opposite. Consequently, Flanders is the least homogeneous areas and Denmark the most homogeneous.

$$H(X) = -\sum_{l=1}^{n} p(x_l) * Ln(p(x_l)) \quad (19)$$

To evaluate the HMSL, 144 tests are carried out in each reduced area using different values (table 1) for shape factor (*X*), suitable patch size (*sps*), initial patch size (*ips*), and decreasing value (*dv*) (section 3.3. explains the meaning of these parameters). The weights for CS and NL, and the weight border (w_N) stay constant in 0.33, and *mn* is kept in 1. This configuration gives *6 x 3 x 8 = 144* tests for each reduced region, the results of which are summarized in table 3.

In tables 2 and 3 i the IP results are compared with the summary of the 144 tests performed with the HMSL. Regarding computation time required for reaching the solutions, the heuristic (table 3) is considerably faster than the IP model

(table 2). While the mathematical approach requires time in the order of minutes and even hours as in the cases of the Netherlands and Denmark, time average of the 144 HMSL tests is around one second.

In Flanders, 29.17% of HMSL tests give solutions deviating less than 0.5% from the objective value obtained with the mathematical approach, and 99.3% deviate less than 2% (table 3). Concerning spatial location, 55% of the solutions have a centroid less than 4 cell apart from the centroid of the site generated by the mathematical approach; this percentage represents the number of patches that are very near to the mathematical objective value, but that at the same time are located in the same region where the mathematically generated site is located. Since *30* cells are selected as part of the final patch, a *4* cells distance is suitable to know whether the solutions are spatially near.

Table 1. Values for *sps-ips*, *dv* and *X* applied in the reduced tests

sps-ips	*dv*	*X*
6 - 2	0.002	0.1
8 - 2	0.01	0.2
15 - 2	0.02	0.3
		0.4
6 - 3		0.5
8 - 3		0.6
15 - 3		0.7
		0.8
		0.9

In The Netherlands, 4.17% of the tests give solutions deviating less than 0.5% from the mathematical objective value, 81.94% deviate less than 1%, and all solutions deviate less than 2%. On the other hand, 63% of HMSL tests produce a site with a centroid less than 4 cells apart from the mathematical centroid.

25.69% of HMSL solutions in Denmark deviate less than 0.5% from the objective value obtained with the mathematical approach, and all solutions deviate less than 1%. Homogeneity plays an important role locating the final patch with the heuristic approach, only 2% of the tests produce solutions having a centroid less than 4 cells apart from the optimal centroid (x=64,y=31.5; table 2). This low percentage is obtained because of the high homogeneity present in Denmark, where a huge number of near to optimal solutions likely exist.

Table 3 shows that while in Denmark, the most homogeneous region, it is easier to obtain solutions near to the optimality reference given by the IP results, the distribution of the solutions in Flanders shows that it is harder to obtain spatially near to optimal solutions. These results are in the agreement with the average of cells overlapping the mathematical solution (table 3).

A more detailed analysis of the overlapping cells can be performed with the results shown in figures 4a and 4b. Although the location and shape of the final site vary in every HMSL test case because of the random ordering of equal valued

candidate cells in the *region growing* step, figure 4 gives some thoughts about the influence of *sps-isp* configurations over the results. In Flanders (figure 4a), small *sps-ips* intervals can find sites more closely located to the one from the mathematical solution when the delta interval (*dv*) is large (*0.01, 0.02*). Since Flanders is the least homogeneous region, small values of *sps* and *ips* guarantee good quality seed patches, which in turn generate the best results not only from the intrinsic suitability but also from the location point of view. Taking into account that large *sps* implies the generation of larger seed patches, not high quality cells in this heterogeneous region could be included as part of these seeds. Accordingly, larger *sps-ips* ranges do not guarantee finding best seeds and nor best solutions. Regarding the delta value (*dv*) for small *sps-ips* ranges, it is feasible to conclude that the smallest *dv* is generating less seed patches although they could be also good quality ones. Therefore the final solutions are of lower quality than the ones obtained for larger *dvs*. Nevertheless, location quality for small *dv* values are increased for large *sps-ips* intervals, where more seed patches are.

Table 2. Results obtained by the mathematical approach

Region				***Mathematical Model***			
	Size (Cells)	**Pixel Classes**	**Shannon Entropy**	**Objective Value**	**Centroid**		**Time (sec)**
					X	**Y**	
Flanders	70 x 70	43	2.51	25.02	64.5	12	548
The Netherlands	70 x 70	117	2.41	24.77	17.5	27	28450
Denmark	70 x 70	26	2.3	26.74	64	31.5	13204

Table 3. Results obtained by the heuristic approach and comparison with the mathematical solution

Summary of the 144 Tests performed with HMSL in each Reduced Region	**Deviation from the Objective Value Obtained with the Mathematical Model (Less than)**				**HMSL Solutions less than 4 cells apart from the centroid of the mathematical solution**	**Average of cells overlapping the mathematical solution**	**Time avg. (sec)**	
	0.5%	***1.0%***	***1.5%***	***2.0%***			***Min***	***Max***
Flanders	29.17%	60.42%	95.14%	99.3%	55%	12	0.02	0.6
The Netherlands	4.17%	81.94%	95.14%	100 %	63%	15	0.3	1.5
Denmark	25.69%	100 %	100 %	100 %	2%	0	0.8	0.8

In line with the Shannon Entropy (table 2) according to which the Netherlands is more homogeneous than Flanders, the results show a more uniform behavior for larger *sps-ips* intervals (figure 4b). This is explained by the fact that although in more homogeneous areas good quality seed patches can be generated with small *sps-ips* intervals, the homogeneity condition implies generating more seed patches in order to generate higher quality solutions from the location standpoint. This argumentation is also valid for explaining the behavior of *dv* and *sps-ips* ranges in Denmark, the most homogeneous region.

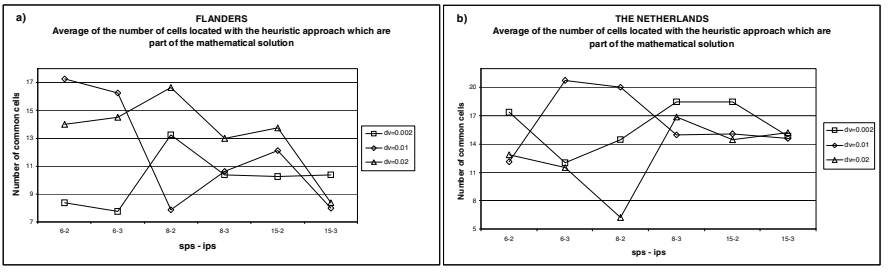


Fig. 4. Average number of common cells obtained with the heuristic and mathematical approaches

4.2 Heuristic Approach

The Heuristic method for Multiple-Criteria Site Location (HMSL) is applied to locate 3000 cells forming compact patches within Denmark, Flanders, and The Netherlands. The criteria are the same to the ones applied for the reduced areas (section 4.1): maximize Carbon Sequestration (CS) and minimize Nitrate Leaching (NL).

A sensitivity analysis is carried out to determine how the *weight border* (w_N) and the *shape factor* (X) control the heuristic site search process. Twenty-four tests are performed varying w_N with values in the set {0.2, 0.5, 0.9}, and X with values in the set {0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9}. The difference between 1 and w_N is equitably distributed for weighting CS and NL. The *Shape Index* (*SI*) in equation 20 (Maceachren, 1985) measures the compactness of each constructed site, taking values between 0 and 1, with the highest values corresponding to the highest compactness. Perimeter and area are computed taking into account the cells part of the final patch. On the other hand, the *Intrinsic* or *Core Suitability* (*CoS*) of the patches is measured according to equation 21, where $f(k_i)$ is the k_{th} NC value in the i_{th} cell part of the final site. The heuristic sensitivity analysis is shown in figures 4, 5, and 6. The sub-figures (a) show the *Core Suitability* (*CoS*) and the sub-figures (b) show the *Shape Index* (*SI*) of the sites obtained in the Netherlands, Flanders, and Denmark.

$$SI = \frac{\sqrt{area}}{0.282 * perimeter} \tag{20}$$

$$CoS = \sum_{i=1}^{M} \sum_{k=1}^{P} f(k_i) \tag{21}$$

As in the reduced tests, the Shannon Entropy measures the amount of information in each entire region. According to the measures summarized in table 4, Flanders has the highest amount of information (highest heterogeneity), and Denmark the lowest (highest homogeneity).

Table 4. Shannon Entroy in The Netherlands, Flanders and Denmark

Region	Original Resolution	Size (Cells)	Pixel Classes	Shannon Entropy
Flanders	100m x 100m	1000 x 2500	1072	2.36
The Netherlands	250m x 250m	1268 x 1076	1742	2.23
Denmark	100m x 100m	3590 x 4740	1413	1.30

Regarding the influence of the Shape Factor parameter (X), figures 5 to 7 show a common behavior, sites with the highest Shape Index (SI) and the highest accumulated Core Suitability (CoS) are obtained with the smallest X values. Since a low X guarantees that the best pixels are added to the current patch, large X values will result in the addition of sub-optimal pixels. This characteristic of the HMSL explains the lack of clarity in the influence of w_N for X values larger than 0.6.

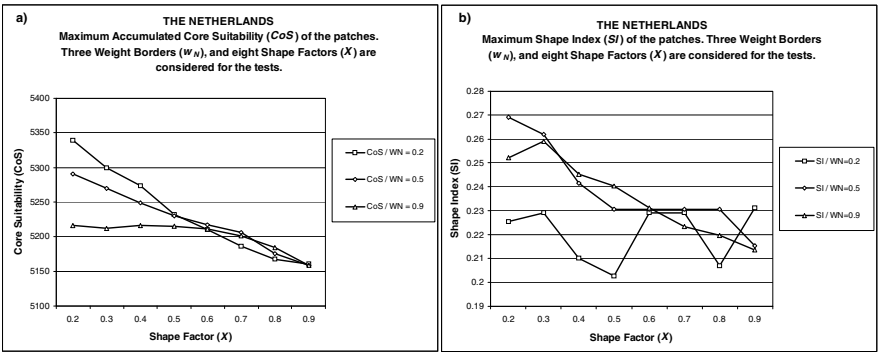


Fig. 5. The Netherlands: Test results for three Weights for Border (W_N) and eight Shape Factors (X). a) Core Suitability; b) Shape Index.

Figures 5 to 7 also show that while SI is favored by higher values of w_N, CoS is favored by lower values of w_N. This behavior is in agreement with the logic of the region-growing algorithm, where w_N is in competition with the weights applied to the criteria involved in the CoS computation (Carbon Sequestration and Nitrate Leaching in the tests). There is also evidence that w_N has a substantial impact for the CoS in Denmark and in The Netherlands. In the tests for these regions (figures 6a and 6b), the differences in CoS introduced by w_N are clearly identified. Indeed the differences are most consistent in Denmark, the most homogeneous area (lowest Shannon Entropy -table 4-).

On the other hand, the influence of w_N is less apparent in Flanders (figures 6a, 6b) which is the region with the lowest homogeneity. Therefore, it is feasible to assert that w_N has a higher significance for CoS in more homogeneous regions. This can be explained by the fact that in homogeneous regions the candidate pixels to be included as part of the patch are more or less similar from the intrinsic

suitability point of view (*CoS*), which implies that, the best pixels are strongly supported by w_N. The opposite behavior is found in less homogeneous areas (e.g. Flanders).

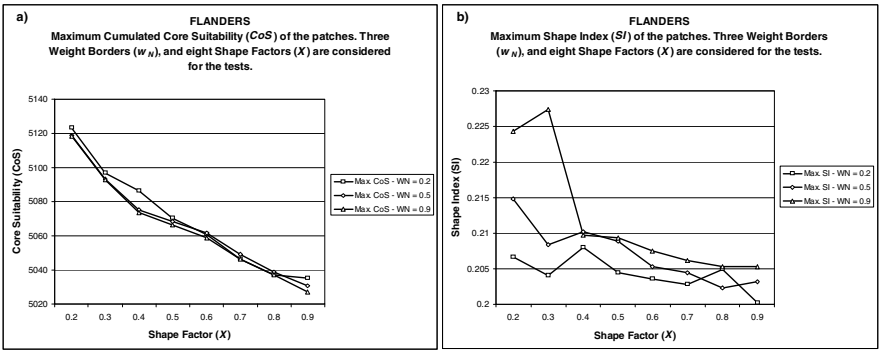


Fig. 6. Flanders: Test results for three Weights for Border (W_N) and eight Shape Factors (X). a) Core Suitability; b) Shape Index.

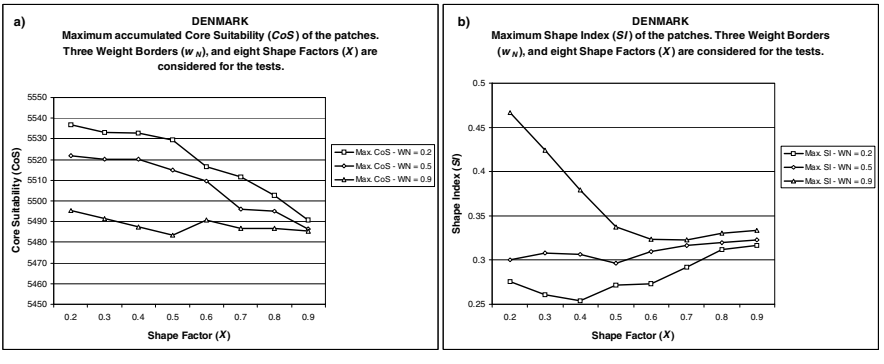


Fig. 7. Denmark: tests results for three Weights for Border (W_N) and eight Shape Factors (X). a) Core Suitability; b) Shape Index.

Figure 8 shows the average of computation time required for the proposed heuristic in order to complete the region-growing process for an individual seed patch. Time required for building a complete site depends on the number of cells required to form it, as well as on the X parameter. Figure 8a shows that the data follow a power model, in which computing time decreases exponentially as X increases.

Data in figure 8b correspond to the computation time for growing different sized sites (from 200 to 3000 cells) in the three testing regions, but keeping $X=0.2$. These results also follow a power model, where computation time increases with the number of cells. Figure 8b shows that computation time for growing patches in Denmark is slightly higher than times recorded for the Netherlands and Flanders.

Since these small differences are produced when the final site is composed by a high number of cells, this behavior can have an explanation in the performance of the sorting algorithm included as part of the region-growing process. The results show that the sorting algorithm likely has inferior performance for the regions with the lowest Shannon entropy. The differences are not substantial though.

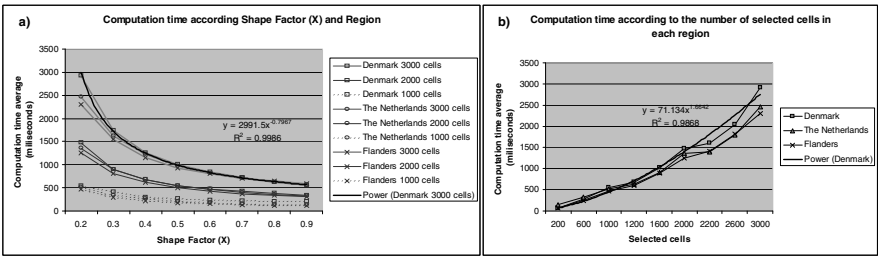


Fig. 8. Average of computing time for the region-growing process included as part of the proposed heuristic. a) Time according to Shape Index (*X*), number of cells, and region. b) Time according to region, and number of selected cells, keeping *X=0.2*.

Figure 9 shows the most suitable site of 3000 cells generated with HMSL assigning the same weight to compactness as well as to two intrinsic criteria: Carbon Sequestration (*CS*) and Nitrate Leaching (*NL*). As shown in figure 9, the site is composed by pixels maximizing *CS* (pixels in dark gray) and minimizing *NL* (pixels in bright gray).

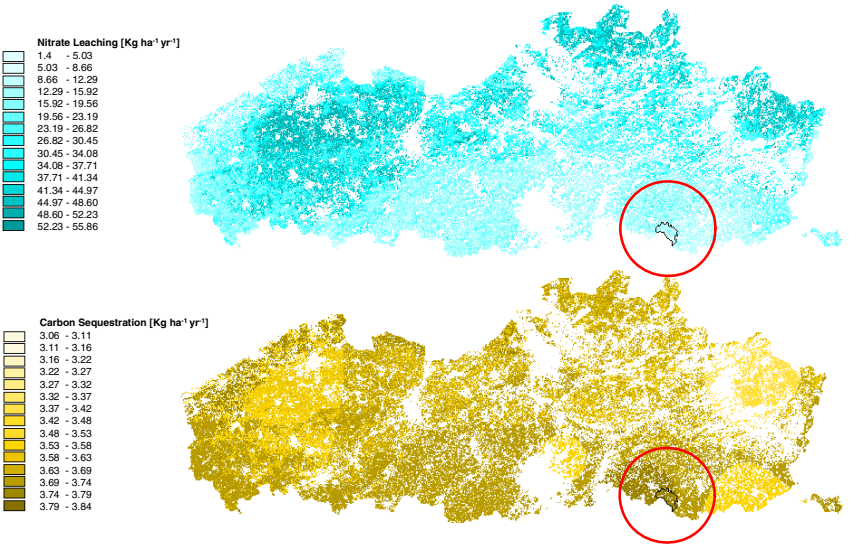


Fig. 9. Most suitable site of 3000 cells for afforestation with oak in Flanders

5 Conclusions and Further Work

Two methods were described to generate compact sites maximizing the intrinsic or core multiple criteria suitability of the set of cells forming these sites: a reference mathemical method and a heuristic approach (HMSL). Both approaches were compared using datasets of limited size. It was found that the HMSL heuristic finds solutions which are very much in line with the mathematical solutions. According to the tests, many solutions near to the optimal objective value can exist, and these solutions can even be spatially distributed. Nevertheless, most of the heuristic tests generate solutions that are also spatially near to the optimal patch location. The comparison also revealed that although the suitable patch size (*sps*), and the initial patch size (*ips*) parameters control the size of the search space, large search spaces (large *sps*, small *ips*) do not necessarily improve the quality of the solutions. Selection of *sps*-*ips* ranges should consider the level of homogeneity of the information at hand knowing that search space size has a direct impact on computation time.

HMSL tests in real sized data for growing one site of 3000 cells and keeping $X=0.2$, required a computation time around 3 seconds. These results show that HMSL is entirely applicable for real sized problems, even considering a high number of candidate solutions. The similarity of the results with respect to the mathematical approach, and the low computation time allow concluding that HMSL has a high performance.

The sensitivity analysis shows that the parameter X (percentage of cells added to the growing patch in each iteration) has a predominant influence on the computation time and on the quality of the results. Low values for X guarantee high quality solutions. Values lower than 0.6 are seem to be adequate for this parameter since this value range does not only guarantee near to optimal solutions, but also allows keeping a clear behavior of the heuristic when w_N (weight given to the compactness criterion) is changed. On the other hand, high values in w_N generate more compact areas, while low values increase core suitability.

Taking the Shannon entropy as a homogeneity approximation, it is feasible to assert that the heuristic is more sensitive to changes of criteria weights in areas with less information (more homogeneous), and for X values lower than 0.6.

HMSL can be improved by including goal programming with intervals, as part of the region growing stage. This improvement will guarantee that each pixel added to the current patch is effectively the best for one or more criteria. The computation time will rise though.

References

Aerts, J., Heuvelink, G.: Using simulated annealing for resource allocation. Geogr. Inf. Sci. 16, 571–587 (2002)

Belton, S., Stewart, T.: Multiple Criteria Decision Analysis. An Integrated Approach. Kluwer Academic Publishers, Dordrecht (2002)

Brookes, C.: A genetic algorithm for locating optimal sites on raster suitability maps. Trans. in GIS 2, 201–212 (1997)

Brookes, C.: A genetic algorithm for designing optimal patch configurations in gis. Geogr. Inf. Sci. 15, 539–559 (2001a)
Brookes, C.: A parameterized region growing program for site allocation on raster suitability maps. Int. J. of Geogr. Inf. Sci. 11, 375–396 (2001b)
Charnes, A., Collomb, B.: Optimal economic stabilization policy: Linear goal-programming models. Socioecon. Plan. Sci. 6, 431–435 (1972)
Church, R., Gerrard, R., Gilpin, M., Stine, P.: Constructing cell-based habitat patches useful in conservation planning. Ann. of the Assoc. of American Geogr. 93, 814–827 (2003)
Church, R., ReVelle, C.: The maximal covering location model. Reg. Science Assoc. 32, 101–118 (1974)
Church, R., Stoms, D., Davis, F., Okin, B.: Planning management activities to protect biodiversity with a gis and an integrated optimization model. In: Proceedings of the Third Int. Conf. on Integrating GIS and Environ. Model (1996)
Diamond, J., Wright, J.: An implicit enumeration technique for the land acquisition problem. Civ. Eng. Syst. 8, 101–114 (1991)
Dimopoulou, M., Giannoikos, I.: Spatial optimization of resources deployment for forest-fire management. Int. Trans. Oper. Res. 8, 523–534 (2001)
Dorigo, M., Stutzle, T.: Ant Colony Optimization. MIT Press, Cambridge (2004)
Fischer, D., Church, R.: Clustering and compactness in reserve site selection: An extension of the biodiversity management area selection model. For. Sci. 49, 555–565 (2003)
Gilbert, K., Holmes, D., Rosenthal, R.: A multiobjective discrete optimization model for land allocation. Manag. Sci. 31, 1509–1522 (1985)
Gilliams, S., Raymaekers, D., Muys, B., Orshoven, J.V.: Comparing multiple criteria decision methods to extend a geographical information system on afforestation. Comput. and Electron. in Agric. 49, 142–158 (2005)
Gourvest, H., Eikland, K., Notebaert, P.: LPSolve IDE v5.5 (2006)
Heil, G.W., Muys, B., Hansen, K. (eds.): Environmental Effects of Afforestation in North-Wester Europe: From Field Observations to Decision Support. Springer, Heidelberg (2007)
Hof, J., Bevers, M.: Direct spatial optimization in natural resource management: Four linear programming examples. Ann. of Oper. Res. 95, 67–91 (2000)
Ignizio, J.: Interval goal programming and applications. Pennsylvania State University Working paper (1974)
Li, X., Yeh, A.G.: Integration of genetic algorithms and GIS for optimal location search. Int. J. of Geogr. Inf. Sci. 19, 581–601 (2004)
Maceachren, A.: Compactness of geographic shape: Comparison and evaluation of measures. Geogr. Ann. 67, 53–67 (1985)
Malczewski, J.: GIS and Multicriteria Decision Analysis. John Wiley & Sons, Chichester (1999)
McDonnell, M., Possingham, H., Ball, I., Cousins, E.: Mathematical methods for spatially cohesive reserve desing. Environ. Model. and Assess. 7, 107–114 (2002)
Mehrotra, A., Johnson, E.: An optimization based heuristic for political districting. Manag. Sci. 44, 1100–1114 (1998)
Shirabe, T.: Modeling topological properties of a raster region for spatial optimization. In: Proceedings of the 11th Int. Symp. on Spat. Data Handl (2004)
Siklossy, L., Marinov, V.: Heuristic search vs. exhaustive search. In: Proc. Second Int. Joint Con. on AI (1971)

Stewart, T., Janssen, R., Herwijnen, M.V.: A genetic algorithm approach to multiobjective land use planning. Comput. & Operat. Res. 31, 2293–2313 (2005)
Williams, J.: A linear-size zero-one programming model for the minimum spanning tree problem in planar graphs. Netw. 39, 53–60 (2001)
Williams, J.: A zero-one programming model for contiguous land acquisition. Geogr. Anal. 34, 330–349 (2002)
Williams, J., ReVelle, C.: Applying mathematical programming to reserve site selection. Environ. and Model. Assess. 2, 167–175 (1997)
Xiao, N.: An evolutionary algorithm for site search problems. Geogr. Anal. 38, 227–247 (2006)

Renewable Energy Sources: The Case of Wind Farms Analysis

Silvana Lombardo and Massimiliano Petri

University of Pisa, Department of Civil Engineering, Via Diotisalvi 2, 56126 Pisa, Italy
e-mail: {s.lombardo,m.petri}@ing.unipi.it

Abstract. The state of art about the assessment of renewable energy sources productivity and sustainable use shows the lack of robust methods and techniques and the inadequacy of applied instruments. The present research implements a methodology that allows to evaluate both the opportunity to use wind energy and, in the affirmative, the producibility and location of a wind farm, introducing elements and quantitative techniques for a correct support decision aiding. An evaluation framework is developed, usable not only for wind but also for other renewable energy sources.

Keywords: Renewable energy integrated assessment, visibility analysis, landscape.

1 Introduction

Following the UE guidelines as to energy, one of the objectives to be reached by 2020 is the 20% of consumed energy coming from renewable sources. Up to now, there are neither sufficient estimations nor adequate methods to assess the feasibility of such a quantified objective, let alone the consideration of the technological, land use, social, economic evolution which is going to take place during the next 15 years.

In Europe there are good examples of intense use of renewable energy and there are examples of national planning and regulations about renewable energy sources (such as solar energy in Spain or France).

In Italy, the lack of national guidelines led the different Italian Regions to emanate local laws and norms which may vary, in some cases even significantly, from a Region to another and do not allow for a wide, sustainable and differentiated diffusion of the various plant typologies, nor take into account the energetic potentials of different territories. These norms are rapidly evolving, and, at present, few Regions have adopted an Energy Plan while several of them are developing it, pressed by the requirement of reduction of greenhouse gases emissions imposed by the Kyoto protocol (Donatini 2007).

In other terms, systematic planning and programming for renewable energy plants siting are still at embryonic stage, and currently the initiative is still left to

B. Murgante et al. (Eds.): Geocomputation, Sustainability & Environ. Planning, SCI 348, pp. 57–72.
springerlink.com

proposals of private companies and subjects. The present role of local authorities is not in developing strategies and is mainly limited to the approval or denial of private proposals. However, none of the current procedures, norms or laws attempts to put in relation the different aspects involved in the exploitation of renewable energy sources, that are the expected energy production and the environmental and visual impact, so that a reliable assessment is not possible (nor required) at present and a systematic and rational planning of the location of the installations over the regional (and, of course, national) territory is impossible.

In this research we analyze the case of wind power energy, but the methodological framework can be extended to other renewable energy sources.

2 Objectives

The *general objective* of this research project is to develop methods and techniques aimed to produce systematic and integrated knowledge on the possibility of production and sustainable use of renewable energy sources. The *specific objective* is to build a planning support system where sectorial modelling and evaluations can be integrated with advanced spatial analysis techniques in order to assess the possibilities of production and sustainable use of renewable energy on the base of natural vocations/potentiality of territorial systems without exceeding its carrying capacity.

Such system can effectively support planners to build production and use scenarios, not only by identifying the more suitable areas for the new installations location on the basis of technical and economic criteria, but also by assessing their effects in socio-economic and environmental terms.

These scenarios must be built and evaluated in an environmental, economic and social sustainability system context and they derive from the application and integration of advanced techniques of territorial data analysis, of simulation and evaluation of the energy-territory interactions. In such direction, it is possible to estimate the "second order" interactions connected with production and use of renewable energy by integrating and evaluating, also from the spatial point of view, information of different nature, otherwise confined in sectorial assessments.

In this context, we implemented a detailed methodology for supporting the wind farm location choices.

3 Methodology

3.1 Preliminary Methodological Remarks

The evaluation associated with wind farms location is not simple, because it involves a great number of environmental factor evolving during the time and interacting one another. From the point of view of energy productivity, the available wind power is influenced not only by meteorological factors, but also by land morphology, terrain roughness and so on. All these elements make not appropriate the classical GIS interpolation techniques for punctual anemometric data.

On the other hand, also the assessment of wind turbine generators impact on the landscape is not simple, taking into account the meaning of landscape as described at the first article of the European Landscape Convention:

> *"Landscape" means an area, as perceived by people, whose character is the result of the action and interaction of natural and/or human factors.*

Existing legislation analyzes the problem from every side but does not draws a clear and shareable methodology.

For instance, the Tuscany Regional Guide (Tuscany Region 2004) for the evaluation of environmental impact of wind farm proposes to evaluate the historical landscape quality using indexes derived from the field of landscape ecology, but the link between all the indexes/criteria to be considered for the final landscape impact assessment results too simplified: only a weighted overlay technique is proposed to take into account the criteria relative importance.

This solution anyway is not satisfactory considering the first article of the European Landscape Convention; in fact, the Regional Guide suggests a two-dimensional and aerial territorial view (figure 1a), while the real landscape perception is three-dimensional and is linked to the vertical individual view. The 3D view superimposes all the objects belonging to the human visual radius and located at different depth/distances from the observer.

Then, landscape quality maps are not enough (can be also misleading, as demonstrated below) and reconstructions of landscape perception from points/observers located on the territory are necessary (McBurney and Collings 1984).

The landscape quality will be different when the observer position changes and depends on all the territorial elements that are interposed between the observer and the wind farm (Danese et al. 2009); these elements are those which build the "*perceived landscape*" (figure 1b).

The wind farm location based on the two-dimensional view does not then allows an appropriate landscape impact assessment (Bishop et al. 2000): the location of the three wind turbine generators in the above example could be considered of high impact because of the presence of the wood and the lake (fig. 1.a), but the

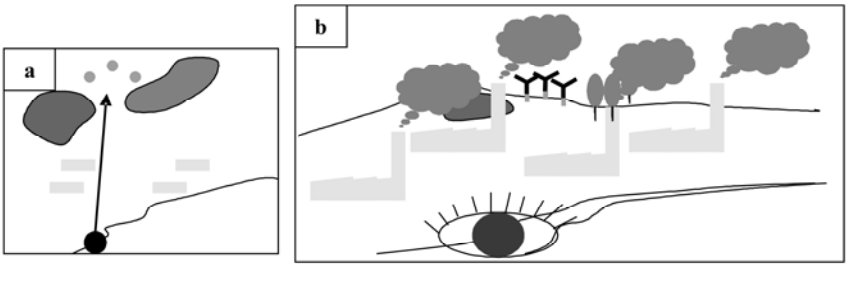


Fig. 1. Aerial two-dimensional view (a) and perceived view (b)

perceived landscape from possible observers hasn't a big quality for the presence of an industrial area in the vertical individual view (fig. 1.b). Of course, the opposite case can occur.

Finally, other two documents that analyze the wind farm landscape impact are the "Linee Guida per l'inserimento paesaggistico degli interventi di trasformazione territoriale" edited by the Ministry for Cultural Heritage and Activities, Department for cultural heritage and landscape (Di Bene and Scazzosi 2006) and the "Linee Guida regionali per la realizzazione di impianti eolici e l'inserimento nel paesaggio" edited by the A.P.E.R. - Producers of Renewable Energy Sources Association (Folgheraiter 2007).

In both cases it is not proposed an assessment method but a group of suggestions to take into account in the evaluation process.

In the present work we build a methodology [Manwell et al. 2002, Maller 2006, Rodman and Meentemeyer 2006) that allows to assess the possible location of a wind farm based on the above-mentioned remarks and introducing elements and techniques for a correct evaluation and decision support.

In this work the methodology is applied for the environmental impact assessment of possible wind farm location in the municipality of S.Luce, in Pisa Province (Tuscany).

3.2 Methodology Description

The methodology is composed of two parts: the first concerns the analysis at the wide scale and the second part the local impact assessment of each possible locations. The methodological phases are illustrated in figure 2. It is important to underline the introduction of the participatory process step at the end of the first part (phase AB); in this step all the environmental and socio-economic data of the study area are known and stakeholders contribute their preferences on the possible locations for each wind turbine generator (evaluated in the second part).

3.2.1 Phase A: Wide Area Analysis

Phase A1: Visual impact area identification (AVI)

In this phase the maximum visibility radius of a wind turbine generator (Bishop 2002) is analyzed through different methods:

1. Visibility analysis based on existing wind farm (observers at three distances: 23, 20 and 17.5 km);
2. Human eye real perceptive capacity analysis;
3. Earth's curvature influence analysis;
4. Visibility analysis of data measured at the Pisa Airport (correlation of visibility with relative humidity).

In the first case it is analyzed the visibility of an existing wind farm located on Monte Vitalba in the municipalities of Chianni (at a distance of 25/30 km from S.Luce) (Brusa and Lanfranconi 2004). In this wind farm there is the same generators type as those expected in S.Luce.

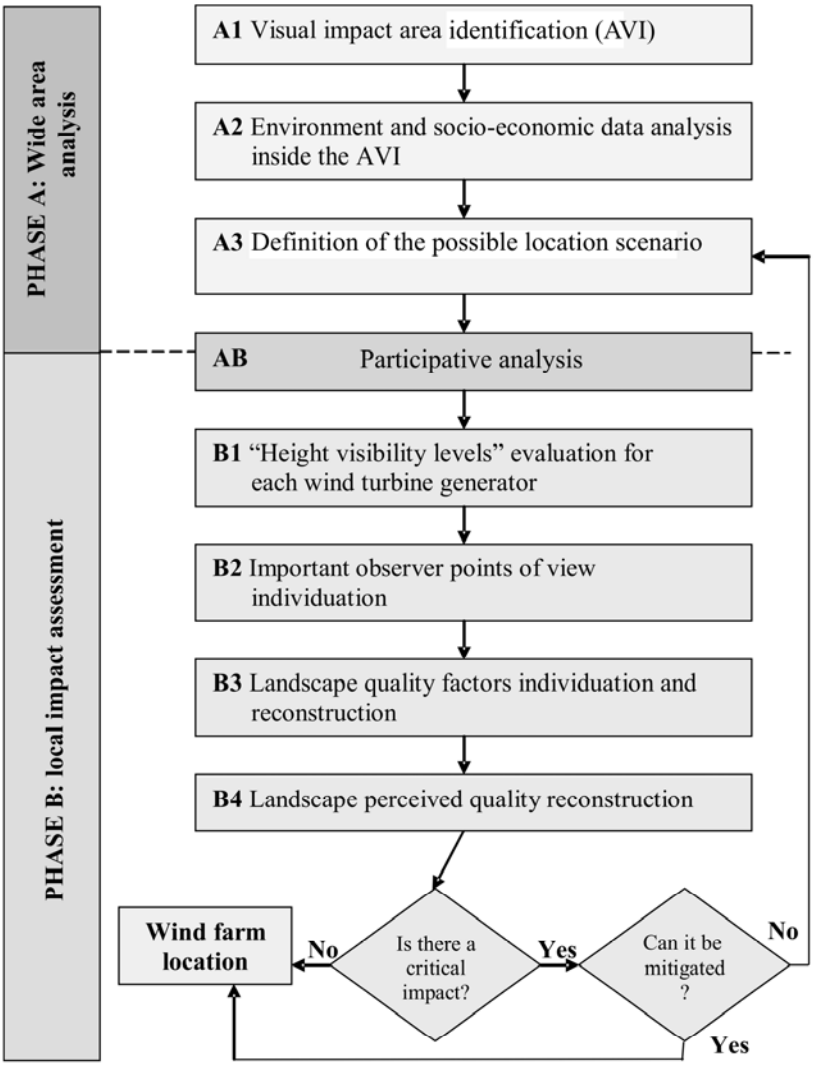


Fig. 2. Methodology general framework

Figure 3 shows the three observer points and the relative wind farm visibility.

The photographic analysis results indicate that the wind farm is hard to see at the distance of 23 km and only the pole supports are visible, from a distance of 20 km it is possible to see also the blades but it is difficult to discern their rotational movement while at a distance of 17,5 km also the blade rotation is visible.

In the third way the method used by Italian Navy Hydrographic Institute for taking into account the influence of Earth's curvature was applied.

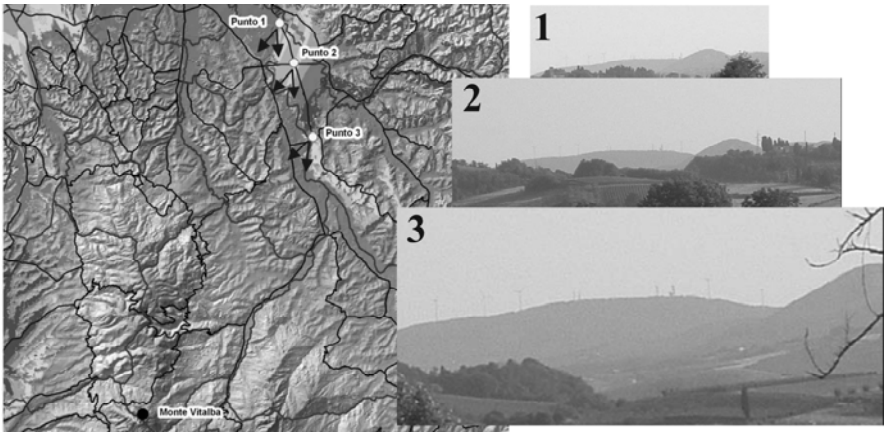


Fig. 3. Visibility from the wind farm located in Monte Vitalba

In the Nautical Tables the method is used to calculate the maximum distance value between a boat and a lighthouse, but in the present case study this analysis does not provide useful information.

The second and fourth analysis derive from the guide edited by the Ministry for Cultural Heritage and Activities.

For all these reasons it was chosen a radius of 25 km around S.Luce hills and, in synthesis, the area shows in figure 4 was adopted as Visual Impact Area (AVI).

Phase A2: Environmental and socio-economic data analysis inside the AVI

The environmental and socio-economic data analysis in the study area was carried considering the following elements:

1. Wind maps:
2. National, regional and provincial prescriptive constraints;
3. Environmental sensitivity relating to:

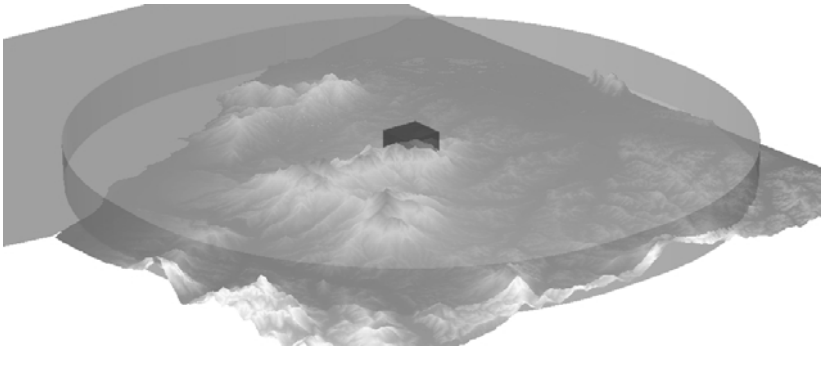


Fig. 4. The Visual Impact Area around S.Luce hills

I. Fauna
- Presence of migratory course;
- Presence of important wintering species (relating to valor and/or vulnerability);

II. Flora
- Presence of important species (relating to valor and/or vulnerability);

III. Air (sound pollution, presence of buildings, etc..);
IV. Soil and subsoil (geomorphologic and hydraulic dangerousness)
V. Landscape and historic/architectonic public property
Presence of landscape attractive/detractive elements;
- Historical Index;
- Shannon's Index (landscape mosaic fragmentation level);
- Hill's Index (predominance of particular land use);
- Sharpe's Change Index (land use dynamics);
- **Visibility analysis** ("visual frequency" and visual field analysis)

VI. Infrastructure (distance from roads and from the power station)

The wind maps were elaborated by the Department of Energetic of the University of Pisa (figure 5).

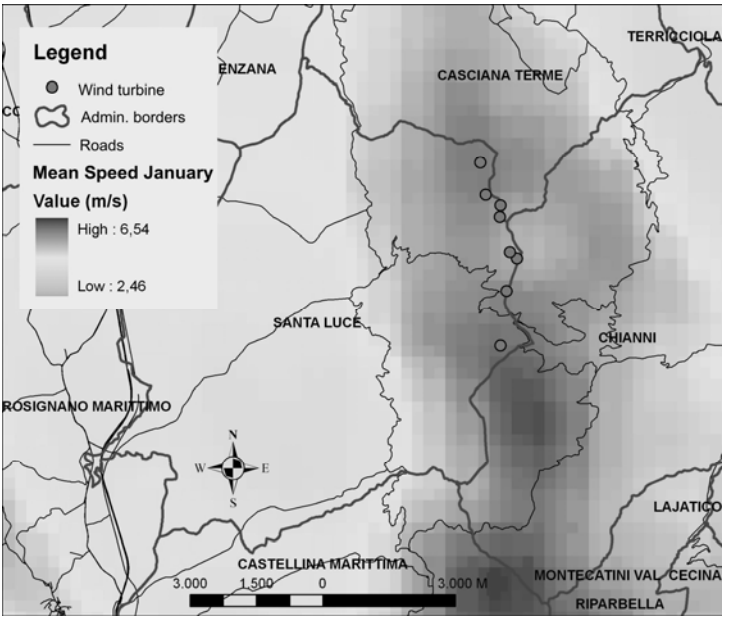


Fig. 5. Wind map relative to the mean speed in January at the height of 60 meters above the ground level

Among all the performend environmental sensitivity analyses, we describe here only that concerning the visibility, because it introduces complex GIS elaborations and the most relevant innovative elements.

Visibility sensitivity analysis

Firstly, the "raised" Digital Terrain Model (DTM) was elaborated, taking into account the height of obstruction elements can influence the visibility (figure 6).

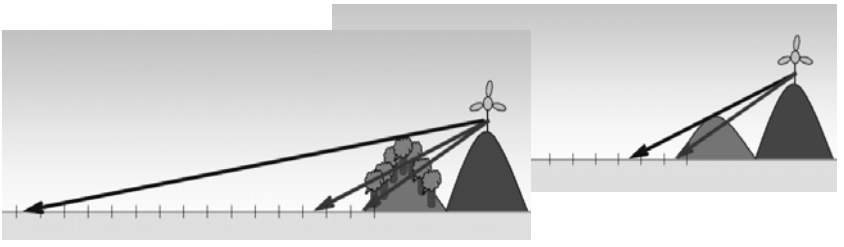


Fig. 6. Effects on visibility of possible obstructions

From the analysis of the study area, its natural and anthropic presences, the main element to add to the original DTM is the wood; all the vegetation types were analyzed (Pertugi et al. 1995) and a height range related to the different vegetation age was elaborated for each one. Then, the original DTM quote was added to the smaller vegetation height value, assuming to have everywhere young plants (Fisher 1993, 1994).

A visual sensitivity index was calculated for the whole study area and on the base both of this index and of the wind maps it will be possible afterwards to identify the preferable locations for wind turbine generator; each of these locations will be evaluate in the phase B. In other words, the visual "capacity" of each territory cell (squared cells of 20x20 meters) was measured; in particular the number of observers that can see it (figure 7) for each cell was calculated (Kidner et al. 1999). These observation points (Franklin 2002, Franklin and Ray 1994) are the centroids of the cells where there is an high probability of visual perception (town residents, train passengers, drivers, etc..).

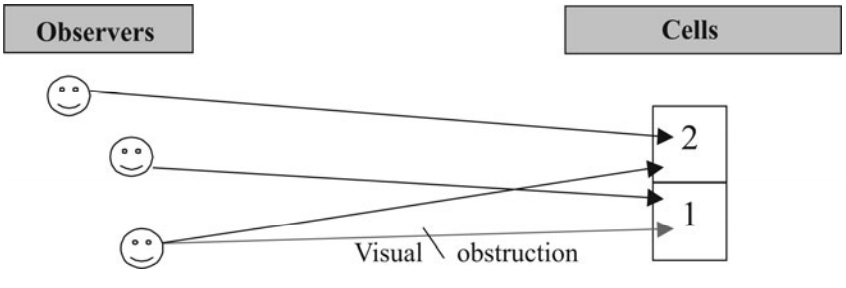


Fig. 7. Evaluation of the number of observers for each cell of the territory

The observers are divided in three main groups:

1. Observers located in the main streets;
2. Observers located in the touristic and natural streets (pathways, streets for trekking, muletracks, etc..);
3. Observers located in urban centers

Observer points are located every squared cell of 100 meters which contains one of the three above categories (figure 8).

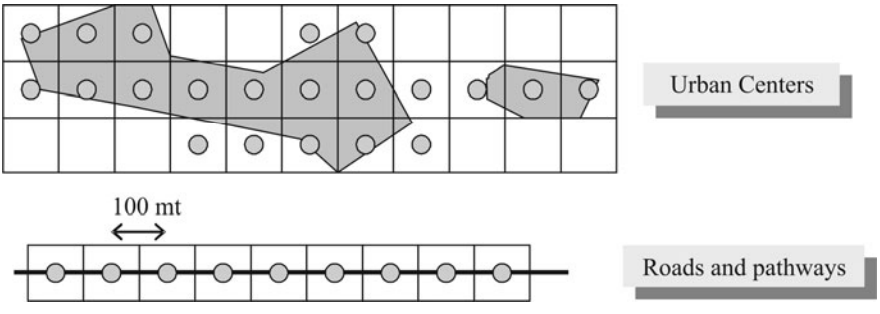


Fig. 8. Observer points representation

Three visibility maps, relative to each observer type, were elaborated; in each of them the number of observers that can view each cell (size 20x20 meters) is measured. The visibility of the single generator can be also partial so the visibility was calculated for different height levels (every 20 meters), as illustrated in figure 9; for each height level the previous three visibility maps were elaborated (figure 10).

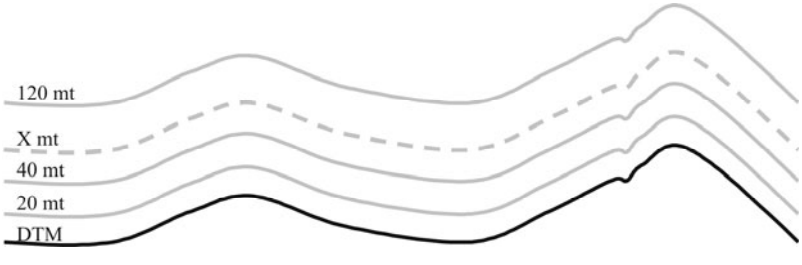


Fig. 9. Height visibility levels construction

Now the aggregation phase is started; at first it is elaborated an "horizontal" aggregation by merging, for each height level, the three visibility maps. The linear aggregation was weighted on the base of the potential observer visibility attention assuming a greater landscape attention for the observers located on pathways or tourist places, then for urban inhabitants and a lower weight has been assigned to drivers.

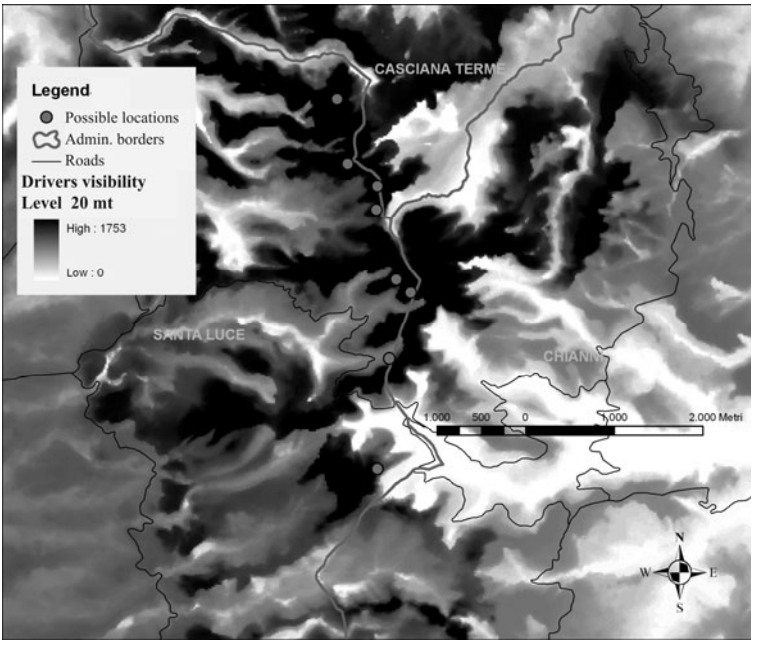


Fig. 10. Height visibility levels construction

This horizontal aggregation is characteristic of multicriteria assessment methods. In the following table 1 the different weight values are showed.

Table 1. Weights assigned to the different observers type

	Visibility from pathways	Visibility from urban centers	Drivers visibility
Weights	0.5	0.35	0.15

Finally it was elaborated the "vertical" aggregation, merging the visibility maps related to each height level. For this final operation the utility function showed in figure 11 was used.

The most part of the studies concerning the visual impact of wind turbine generator assumes that the element of greater impact is the blade rotation because it is in high contrast with the landscape immobility. Then, it was assigned a greater weight to the visibility of the whole blade.

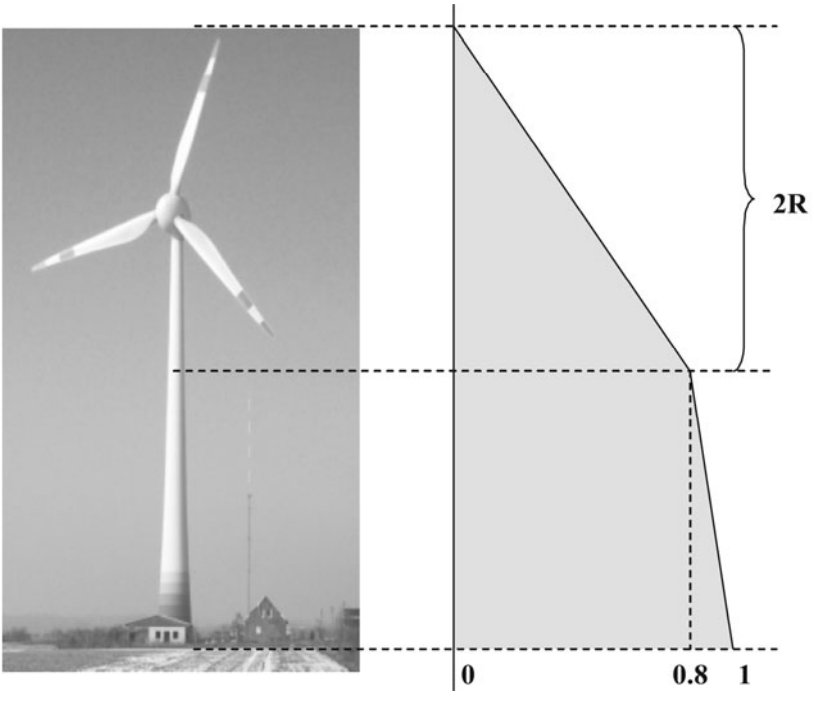


Fig. 11. Utility function for the "vertical aggregation"

The Visibility Index and the wind maps were the main criteria for the choice of the wind farm location areas (see the points in figure 10), even if all the other territorial elements were taken into account (environmental criteria, infrastructure, etc..). For example, some considerations about local road network brought the exclusion of some possible location areas.

3.2.2 Phase B: Local Impact Assessment

Phase B1: "Height visibility levels" evaluation

The wind farm visual effect depends on "how much" of the single generator is visible, therefore it was calculated the generator's visible height from each cell.

A prototypical generator 100 meters high was simulated by means of the GIS software in each location area resulting from Phase A and every ten meters along each generator a target point has been located (figure 12).

For each cell inside the AIP the number of visible target points was calculated so, for example, a number of three means that from this cell it is possible to see the generator only for 30 meters from the top of the turbine while with a value of 11 all the generator is visible.

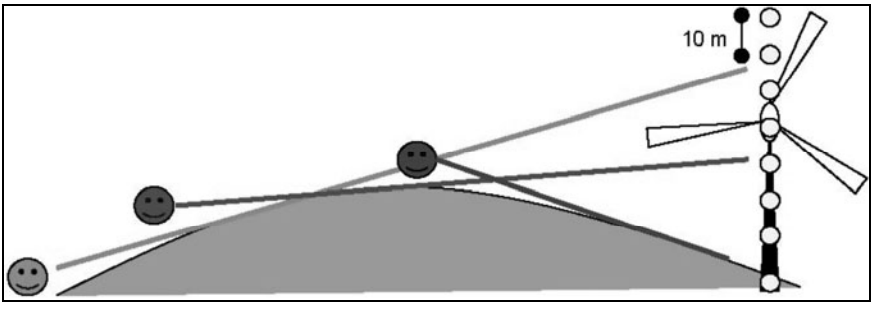


Fig. 12. Different generator visibility degrees

Phase B2: Important observer points of view individuation

From the overlay of the visibility maps and the spatial distribution of observers inside the AIP it was possible to identify the areas having a greater visual impact.

Phase B3: Landscape quality factors individuation and reconstruction

It is then necessary to analyze the whole landscape impact, besides the visual one (Hurtado et al. 2004). To this end, some indexes about landscape quality was evaluated:

- Indices deriving from the field of landscape ecology like Shannon Index, Hill Index, Sharpe's Change Index; these values are capable to represent quantitatively different landscape features;
- Historical Index able to estimate quantitatively the value of the landscape historical elements.

A problem that arises is related to the input data necessary to evaluate all the previous indexes; these input data are mainly multi-temporal land use maps and these informations are often not available and very expensive to elaborate.

Moreover a right landscape quality analysis should include a participatory process where each citizen and each "land user" specifies the local landscape characteristic elements; in this analysis it hasn't been possible to start a participatory process mainly for the short available time.

So the landscape quality has been evaluated with the individuation of detractive and attractive landscape elements (indicated in the Regional Plan and in the local legislation) (table 2).

After the individuation of the relevant visual points it is possible to analyse in detail the observer landscape perception in order to reconstruct the individual visual scene.

Table 2. Example of some extracted elements

Env. Factors		Criteria	Data
Landscape quality elements	Attractive elements	**Anthropic elements**	Urban area built before 1954
			Urban area built before 1878
			Crumbling ruins
			Civil architecture of great value
			Religious architecture
			Rural architecture
			Monuments
			Others
		Protected areas	Parks
			Nature reserves
		Agricultural elements	Olive grove (and rows)
			Vineyard (and rows)
			Orchard (and rows)
			Chestnut grove
			Tree Rows
		Natural elements	Geologic high value elements
			Ridge grasslands
			Rocky outcrops
			Riparian associations
			Wood boundaries
	Detractive elements	**Anthropic elements**	Shacks
			Hothouses
			Power stations
			Electrical boxes
			Filling station
			Abandoned quarries
			Working quarriess
			Special isolated buildings
			Cableways
			Purification plants
			Chimneys
			Others
		Transport elements	Main streets
			Railways

Phase B4: Landscape perceived quality reconstruction

The first step is the calculation of the visual fields of each "relevant observer" by means of geoprocessing elaborations. After the superimposition of the visible area with the detractive and attractive landscape elements, the value of each landscape view and then the impact value of the wind farm can be evaluated. An example is represented in figure 13 with the observer placed in Vada town.

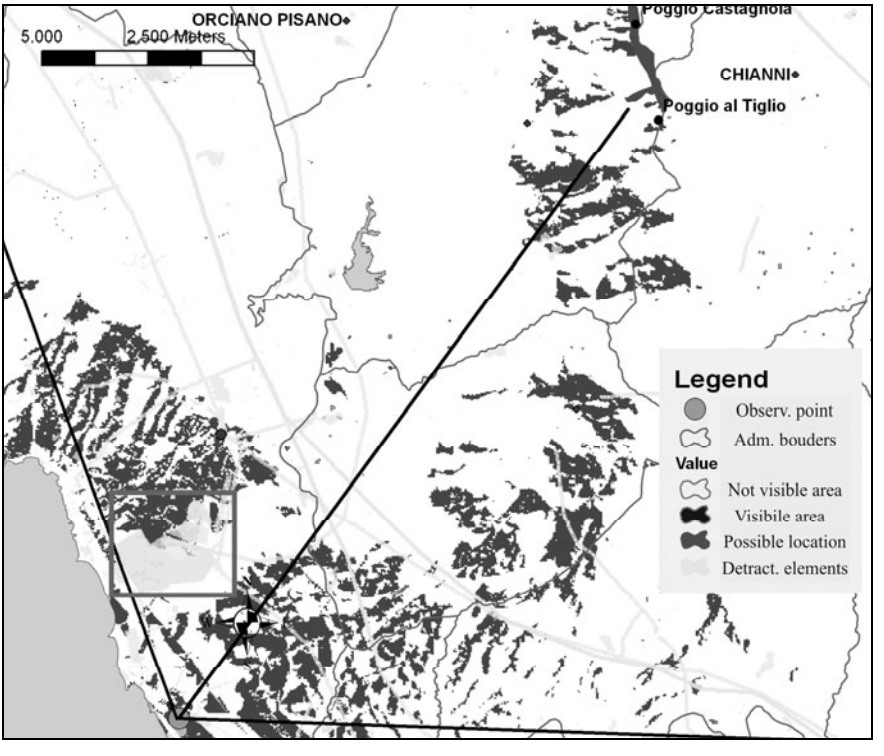


Fig. 13. Visual scene reconstruction for an observer placed in Vada

In this example the presence of a visible area where it is located the big industrial area called Solvay (dark grey rectangle on the bottom left of figure 14) shows that the impact of the wind farm is low for the low value of the landscape perceived quality. There is also the possibility to have a critical impact; in this case it is necessary to find some solutions to mitigate the impact. For example it is possible to locate a vegetation zone (grey circle in figure 14) to hide the wind farm. Also this analysis is elaborated by means of GIS geoprocessing tools able to construct the altimetrical profile graph (figure 14).

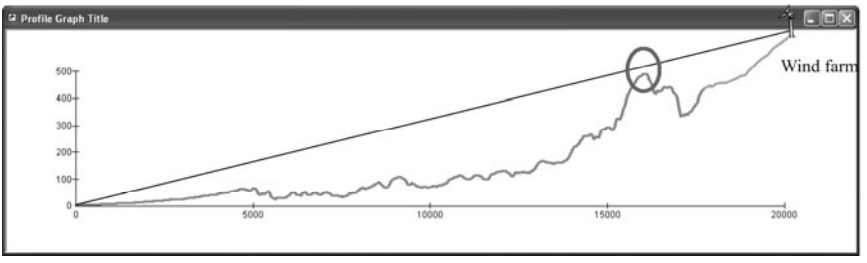


Fig. 14. Altimetrical profile graph and an example of impact mitigation

3.3 Conclusions and Future Developments

The proposed methodology is an attempt to elaborate a procedural framework to evaluate the landscape impact of every new type of territorial elements (wind farms, dumps, incinerators, etc..). So the landscape quality measure (elements clearly difficult to schematize) and its assessment is included inside a complete process that starts from the wide area and ends with the local impacts mitigation. The methodology starts from the construction and analysis of the wide area probably interested from impacts and, after the evaluation of possible wind farm locations, goes to analyse in details each possible location.

Then this framework can be useful to local authorities in developing strategies and when it needs to express an opinion inside the Environmental Impact Assessment (EIA) procedure (approval or denial of private proposals) and can be used also by private firms in their environmental impact study to identify the location of wind farms.

The most difficult step of this methodology is the evaluation of the perceived landscape quality. This is due to subjective elements that need of a proper participative phase to be individuate.

So the future developments of the research regard a better evaluation of landscape quality, not only influenced by the presence of detractive or attractive elements but also from cultural and historical features (Agostini 2007) and by their interactions in the individual subjective view. For this last point the present research activity is based on the use of Multi Agent Participatory Systems (Guyot et al. 2007) to be inserted in the phase AB relative to the Participatory analysis.

References

Agostini, R.: Il paesaggio nel regolamento urbanistico di Fauglia. Energy Landsc, Locus 2, 70–80 (2007)

Bishop, I.D.: Determination of thresholds of visual impact: the case of wind turbines. Environ. Plan B 29, 707–718 (2002)

Bishop, I.D., Wherrett, J.R., Miller, D.: Using depth variables as predictors of visual quality. Environ. Plan B 27, 865–875 (2000)

Brusa, A., Lanfranconi, C.: Linee guida regionali per la realizzazione di impianti eolici e l'inserimento nel paesaggio. A.P.E.R. – Associazione Produttori Energia da fonti Rinnovabili web page (2004)

Danese, M., Nolè, G., Murgante, B.: Visual Impact Assessment in urban planning. In: Murgante, B., Borruso, G., Lapucci, A. (eds.) Geocomputation and Urban Planning. SCI, vol. 176, pp. 133–144. Springer, Berlin (2009)

Donatini, F.: Fonti rinnovabili. Il quadro generale in Italia. Energy. Energy Landsc, Locus 2, 17–25 (2007)

Fisher, P.: Algorithm and implementation uncertainty in viewshed analysis. Int. J. GIS 7, 331–347 (1993)

Fisher, P.: Stretching the viewshed. In: Proc. Symposium on Spatial Data Handling, pp. 725–738 (1994)

Folgheraiter, W.: Parco Eolico Monte Vitalba (Chianni). Galileo, 2 year II pagine (2007)

Franklin, W.R.: Siting observers on terrain. In: Symposium on Spatial Data Handling pagine (2002)

Franklin, W.R., Ray, C.: Higher isn't necessarily better: Visibility algorithms and experiments. In: Symposium on Spatial Data Handling, pp. 751–763 (1994)

Guyot, P., Murakami, Y., Platon, E., Quenum, J.: Partecipatory Design of Multi-Agent System. In: AAMAS 2007 Tutorial (2007),
http://cormas.cirad.fr/forum/cormas/archives/att-0181/01-CallForParticipation.pdf

Hurtado, J.P., Fernandez, J., Parrodo, J.L., Blanco, E.: Spanish method of visual-impact evaluation in wind farms. Renew Sustain Energy Rev. 8, 483–491 (2004)

Kidner, D., Sparkes, A., Dorey, M.: GIS and Wind Farm Planning. In: Stillwell, J., Geertman, S., Openshaw, S. (eds.) Geographical Information and Planning. Springer, Heidelberg (1999)

Manwell, J.F., McGowan, J.G., Rogers, A.L.: Wind-energy systems: environmental aspects and impacts. Wind-energy explained. Wiley & Sons, New York (2002)

McBurney, D.H., Collings, V.D.: Introduction to Sensation/Perception, 2nd edn. Prentice-Hall, Englewood Cliffs (1984)

Maller, B.: Changing wind-power landscapes: regional assessment of visual impact on land use and population in Northern Jutland, Denmark. Appl. Energy 83, 477–494 (2006)

Pertugi, A., Gabellini, A., Acciai, A.: Bosco di Santa Luce Carta naturalistica ed escursionistica. Santa Luce Comune web page (1995)

Rodman, L.C., Meentemeyer, R.K.: A geographic analysis of wind turbine placement in Northern California. Energy Policy 34, 2137–2149 (2006)

Tuscany Region, Linee guida per la valutazione dell'impatto ambientale degli impianti eolici. Tuscany Region Editor, Florence (2004)

Di Bene, A., Scazzosi, L.: Gli impianti eolici: suggerimenti per la progettazione e la valutazione paesaggistica. Linee Guida per l'inserimento paesaggistico degli interventi di trasformazione territoriale. Gangemi editor, Rome (2006)

Identifying Viewshed: New Approaches to Visual Impact Assessment

Maria Danese[1], Gabriele Nolè[2], and Beniamino Murgante[2]

[1] Institute for Archaeological and Monumental Heritage (CNR-IBAM)
e-mail: m.danese@ibam.cnr.it
[2] University of Basilicata, Viale dell'Ateneo Lucano, 10 - 85100 - Potenza - Italy
e-mail: beniamino.murgante@unibas.it, gab.n@libero.it

Abstract. Location of a wind farm needs a visual impact assessment procedure considering different points of view, such as natural-environmental sides and landscape aspects. About the first point, used evaluation approaches are nearly consolidated in literature. Instead, debate is still open on what concerns the second point, especially on establishing an as objective and effective as possible method. Some instruments already exist, such as viewshed analysis (single, multiple, cumulative, identifying), allowing to follow step by step visual evaluation by supporting wind farm location process from site choice to single turbine location. Consequently, this paper aims to implement an open source model to identify a wind farm in Laurenzana municipality (Basilicata, southern Italy).

Keywords: Wind farm, Visual impact assessment, Viewshed, Cumulative Viewshed, Identifying Viewshed.

1 Introduction

Decrease in fossil fuel availability, rising oil prices and increase of air pollution emissions have led to the growth of renewable energy sources, allowing natural and quick replenishment and reduced environmental pollution. Particularly, in last years, green economy success coupled with the need to rapidly increase the amount of energy derived from traditional sources has led to the establishment of a huge number of wind farms with weak environmental impact assessments. Landscape protection has not been sufficiently taken into account for such evaluations. Wind farms produce soil consumption, noise due to electromagnetic interference, impacts on birdlife, vegetation and visual aspects.

In typical approaches to wind farms environmental impact assessment a lot of features are computed, such as effects on flora, fauna, protected areas, soil, geology, but only in recent times visual aspects have become essential requirements in order to pursue landscape preservation. This feature is the most evident one for population, generating a public resistance in the case of huge impacts. It is not easy to objectively assess visual impacts of wind turbines, because of various

B. Murgante et al. (Eds.): Geocomputation, Sustainability & Environ. Planning, SCI 348, pp. 73–89.
springerlink.com

factors to consider, such as distance from turbines, size, number, colour, weather conditions, amount of time during which the population is exposed at the sight of turbines and how it can really observe their presence (Möller 2006).

Traditional approaches based on cartographic maps in a first period have been extended with a lot of multimedia tools, photomontages, quick time VR, video sound enriching the potentiality of geographical data visualization (Dransch 2000). Despite such tools may offer a support in assessing impacts, they have limits in accounting for a certain degree of subjectivity. A more objective way to achieve more objective evaluations is to adopt spatial information technologies in visual impact assessment. Compared to typical multimedia tools, such as renderings, photomontages or videos, results are less attractive and suggestive, but represent the only way to assess, from each part of the study area, if an object is visible or not.

This kind of analysis is based on viewshed computation, which identifies every cell visible from one or more observation points (Burrough and McDonnell 1998). Viewshed analysis has been improved realizing multiple viewshed (Ruggles et al 1993; Kim et al 2004) and cumulative viewshed (Wheatley 1995). Multiple viewshed combines all layers achieved for each target point with the union operator. Cumulative viewshed allows to account for many points of view, summing all grids by means of map algebra (Tomlin 1990, DeMers 2000), providing further information on how many objects can be seen in a certain zone. In this paper a new viewshed operator, Identifying Viewshed, has been developed in GRASS (Geographic Resources Analysis Support System) GIS software. This extension allows to understand how many and which objects are visible in several areas.

2 Viewshed Analysis

Viewshed analysis is a common method adopted in spatial analysis. In recent times, following the great diffusion of GIS, viewshed analysis has been used in a lot of fields: in several archaeological applications (Fisher et al 1997; Lake et al 1998; O'Sullivan and Turner 2001; Ogburn 2006), in the study of protection of extinguishing species (Camp et al 1997), in visual impact assessment of wind turbines (Kidner et al 1999), in marble quarry expansion (Mouflis et al 2007), etc. Planners adopted viewshed analysis predominantly in landscape planning (Aguilo´ and Iglesias 1995; Hanna 1999; Hernández et al 2005), but also in supporting decisions in urban planning (Lee and Stucky 1998; Danese et al 2009) and in urban design applications (Yang et al 2007).

Viewshed analysis allows to understand which pixel is visible from an observation point, located in a certain area with known morphology. Therefore, in this analysis the input datum is a point pattern, while the output is a grid containing numerical information about target visibility.

Nevertheless, in most cases, there is the need to evaluate visibility from more than one point, at the same time. For this reason several authors (Ruggles et al 1993; Kim et al 2004), in order to increase viewshed functionality, developed multiple viewshed which produced a binary grid, where 0 means target not visible, while 1 means target visible, from the union of single viewshed rasters (figure 1).

Another approach considers cumulative viewshed analysis (Wheatley 1995). This is calculated using the union operator and inserting a counter. Consequently, the final result is a no more binary grid, but it indicates, for each cell, the number of observation points or (inverting the problem) the number of observed objects (figure 1). The same result can be obtained summing single viewshed binary grids coming from each observation point.

Despite the fact cumulative viewshed seems to be a fairly complete analysis, this approach lacks in a fundamental aspect. Cumulative viewshed does not show which target is visible from a certain cell. Evaluating the impacts of a new neighbourhood, it is possible that only one building generates a huge visual impact. It is fundamental to identify the object which generates great part of the obstruction. On this purpose, a new viewshed analysis has been developed, the Identifying Viewshed, which shows which target is visible for each cell (figure 1).

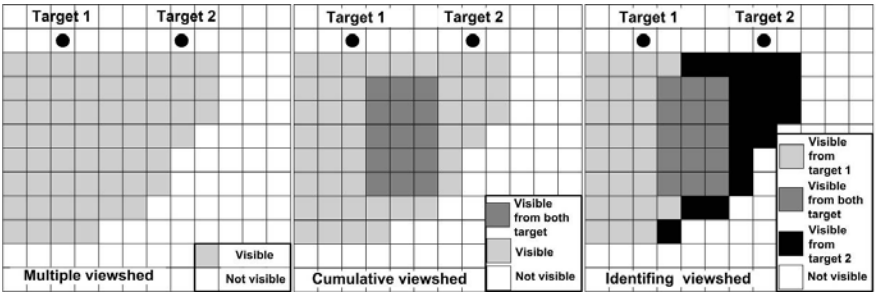


Fig. 1. Main differences among Multiple, Cumulative and Identifying Viewshed in the case of two targets

Clearly, many factors influence results of visual impact. These can be summarized in three main points (Ogburn 2006):

- observer properties;
- observed object properties;
- environment properties.

For what concerns observer properties, two main aspects should be considered:

1. Limits of human sight. Consequently, there is the need to consider a limited observation radius around the observer to take into account these factors.
2. Altitude of the observer above the terrain. As the common sense suggests, visibility can change completely, sometimes even only if we consider two differently tall people. When they observe an object, this can be seen only from the taller and not from the shorter one, because of obstacles to sight which can exist between shorter observer and object.

Observed object properties are strongly related to geometric properties. Increasing object dimensions, the possibility that the object is visible will grow. Also, object colour could influence its visibility.

About environment properties, it is important to consider surrounding atmospheric factors, like the presence of humidity and mist, because they can alter the visibility level. But most of all, it is important to implement a detailed digital elevation model, closer to reality. On this purpose, it is important to account for the third dimension of land cover and other artefacts as bridges and dams, which could substantially obstruct the visibility.

Two other important aspects in viewshed interpretation could be considered: edge effects and cell size. Wheatley and Gillings (2000) analyzed edge effects in archaeological visibility, showing how results could be affected by a certain degree of inaccuracy in areas close to the edge of the study zone. Particularly, in the case of cumulative and multiple viewsheds, errors could be combined with other errors producing meaningless results. This problem can be solved analyzing a wider zone than the study area. Another problem could be generated by cell size. If Digital Elevation Model can be generated from a detailed map, it is important to produce a very fine grid in order to increase accuracy level.

2.1 Identifying Viewshed

The developed viewshed has been called "Identifying viewshed" because it is possible to understand, in an univocal synthetic map, how many and which landmarks are seen from each point of the territory, instead to have different maps, difficult to consult and to analyse at the same time.

In order to obtain it, considering n binary viewshed maps V_i, respectively calculated for each observation point, the Identifying map **I** can be obtained applying the following expression:

$$I = \sum_{i=1}^{n} V_i \cdot 10^{i-1} \tag{1}$$

In simpler words, if we have three matrixes A, B, C (figure 2), each corresponding three viewshed maps, the Identifying map **I** can be obtained combining the simple binary viewshed maps according to equation 1. Consequently, in each pixel of the map **I** there will be a binary number, which is formed by 3 digits, each representing the capability to see one observation point; so it allows to understand which point is visible (digit equal to 1) and which is not (digit equal to 0). As an example, considering the pixel highlighted in the picture: the first digit represents the ability that point A has to see that pixel (which is 0), so point A cannot see it; the second digit represents point B (in this case it is 1), so the pixel is visible from B; finally the third digit represents point C (in this case 0), so the pixel considered is not visible from C.

Obviously, the problem can be inverted, since the identifying map does not only show which points are visible from each pixel, but also which points can be seen from each pixel. This means that, considering again the highlighted pixel, point B not only can see the pixel, but it can also be seen from it; in the same way points A and C cannot see and cannot be seen from the pixel.

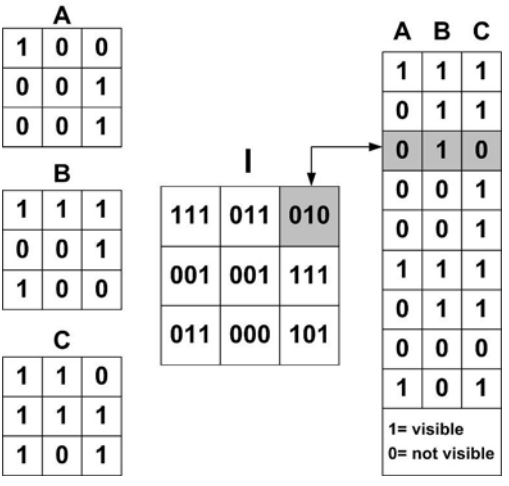


Fig. 2. Scheme of identifying viewshed map implementation

This procedure has been implemented in Grass by creating a main module, called *r.visibilità.sh*, useful to calculate Cumulative and Identifying Viewsheds. This module is supported by two other scripts, called *v.to.dem.sh* and *v.impact.sh.* Such scripts will be described in the following paragraphs.

3 The Open Source Implemented Model

In order to improve a new quantitative approach to visual impact assessment, a model has been implemented in GIS open source software GRASS. In this way it is possible to reduce processing time, increasing the level of objectivity and transparency of analysis.

This model allows to understand:

- Current visibility of landscape and cultural heritage from critical viewpoints. More particularly, it is possible to understand how much and from which observation point each monument and natural resource can be visible.
- Visibility after wind farm location. It is possible to simulate if, how much and which are the wind turbines in huge contrast with landscape property. The great advantage is the possibility to assess if just one wind turbine represents an environmental threat. Consequently, moving just this wind turbine, it is possible to greatly reduce visual impact.

Intervisibility analysis developed in this application is mainly based on a new viewshed analysis, called “Identifying Viewshed”, because it provides information about which cell is visible from each target.

On this purpose, three modules have been realized; their main properties can be synthesized as follows:

- *v.to.dem.sh* script allows to build a DEM from vector data accounting at the same time for contour lines, spots height, buildings and woods, and other elements, which that can be added to 3D model and that could obstruct sight of an object. When the user chooses these elements, he/she can select or not a flag, depending on if he/she wants to consider the height of the element above the sea level or not.
- GRASS software adopts *r.los* (line of sight) module to create single viewshed maps. Each pixel of these maps contains the vertical angle (measured in degrees) from which the target is visible. An extension of *r.los* module, *r.visibilità.sh,* has been developed in order to produce: as a first step, the basic binary viewshed, where the pixel is 0 if the target is not visible and 1 if the target is visible; as a second step, Cumulative and Identifying Viewshed.

4 The Case Study

The case study has been applied in a region located in southern Italian Apennine, with elevated wind power and at the same time excellent landscape from a naturalistic point of view. Sited in the central part of Basilicata region (southern Italy) Laurenzana municipality is located on the top of the ridge at 850 m a.s.l.. This feature coupled with low inhabitants density (22 in/km²) is the main reason why this area is considered particularly suitable for wind energy production.

4.1 Construction of the DEM

The first important step for site selection of wind factory and for visual impact assessment procedure, was the construction of a DEM. To this aim, two kinds of parameters were considered.

Cell size was the first one. It was chosen equal to 20m to find a compromise between the great extent of the studied area and a good accuracy level of the DEM.

The presence of vegetation was the second one. It can obstruct the visibility of a wind farm, so, depending on the land uses, different height values were chosen. Even though vegetation has a variable height, depending on the age of the topsoil and other environmental factors like altitude, exposure, etc., constant mean values have been adopted, according to Bernetti (1995):

- Conifer woods: 20-30m;
- Broad-leaf woods: 15-25m;
- Bushes: 3-5m;
- Orchards: 5-8m;
- Olive groves: 5-8m;
- Fences: 15-20m;
- Riparian vegetation: 20m.

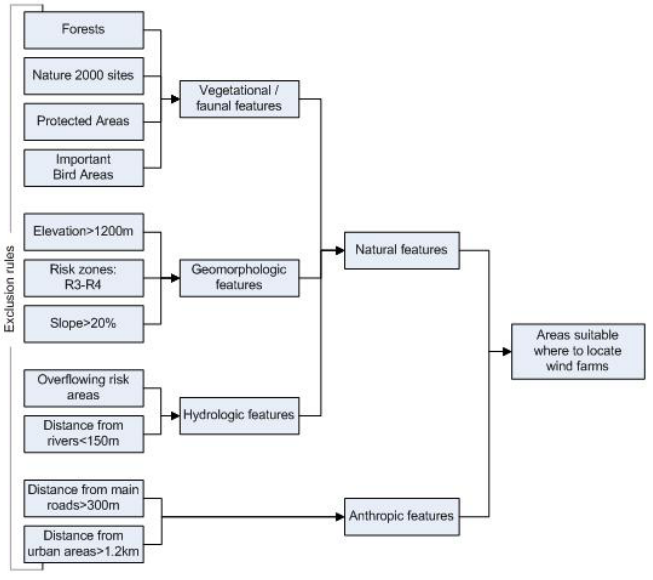


Fig. 3. Land Suitability Analysis procedure scheme

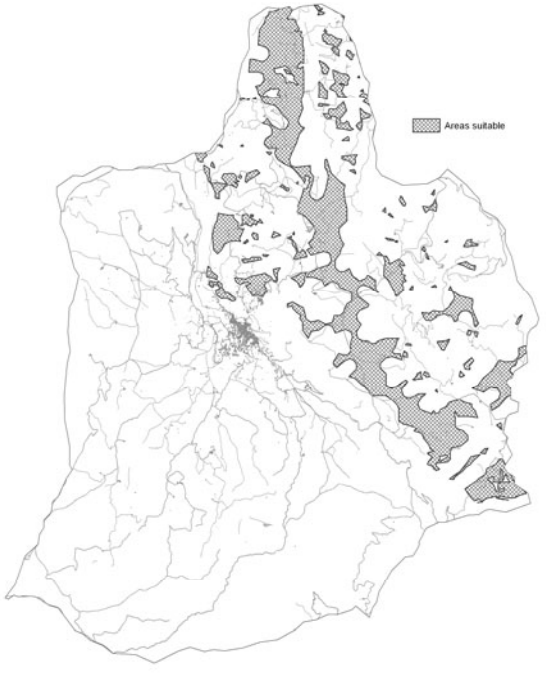


Fig. 4. Areas compatible to wind turbines

4.2 Land Suitability Analysis for the Best Site Choice

In the second step, according to McHarg's (1969) Land Suitability Analysis, some excluding rules have been chosen in order to understand which are areas inside Laurenzana Municipality that are suitable for wind farm location. From these areas we excluded areas characterized by the presence of natural or particular anthropic features. Within the first one, there are areas included within a distance of 150 m from rivers, streams and springs, slopes higher than 20%, Nature 2000 sites, hydro-geological risk zones, areas higher than 1200 m a.s.l., landslides; within the second one there are areas close to railways and road networks.

From this procedure (figure 3) we derived a first selection of sites which can be considered compatible with the location of wind farms (figure 4).

4.3 Simple Viewshed Analysis and Geometric Rules for Wind Turbines Location

At this point, wind turbines were located considering a combination of geometric rules and of results of the simple viewshed analysis.

Firstly, a simplification was done in already selected sites, excluding from the following analysis areas smaller than $650m^2$, because this is the minimum extension necessary to realize additional infrastructures in a farm, like accessibility roads, as well as tubular base and foundations.

Secondly, another geometric rule was considered. The aim was to locate turbines and to avoid both the spread effect of wind farm, caused by elements too far from each other, and the forest effect, caused, on the contrary, by elements too close to each other. Consequently, because the suggested distance dividing two turbines is generally about 3÷5 times the diameter of wind blades, suitable areas were divided by a squared grid with a side of 270 m (each turbine has 90m diameter wind blades) and a wind turbine was placed in the centroid of each square, for a total of 125 turbines.

Thirdly, actual area of visual impact was determined considering limits of human sight. In fact, since object perception is a function of the distance between observer and observed object, a wind farm is perceptible only within a certain distance, and this is further stressed because the more is the distance, the more the object mixes with the background.

Considering classical descriptive geometry (figures 5 and 6) it is possible to observe the linearity of the relationship between the dimension of perceived object and the distance between object and observation point.

More particularly, perceived height of an object is inversely proportional to that distance (Benson et al., 1998; Bishop et al., 2004). Consequently, perceptible height of the object is the result of the distance between the observation point and the tangent of the perception angle.

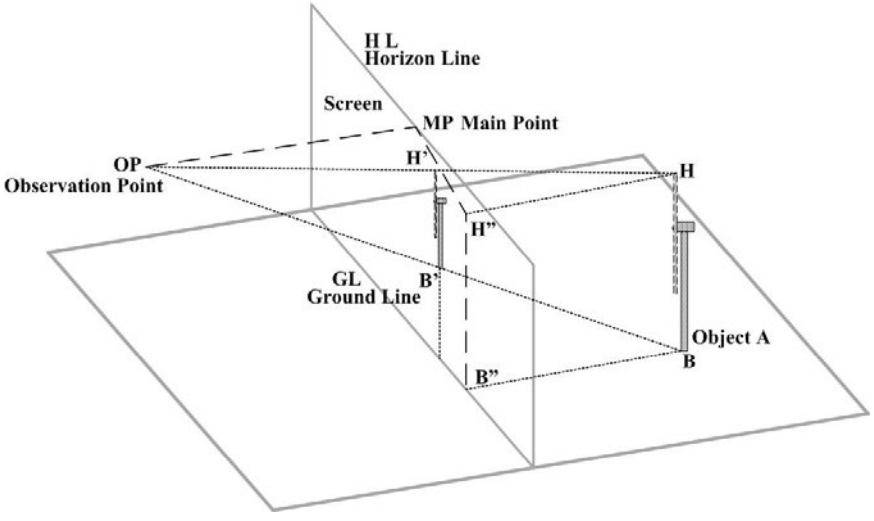


Fig. 5. Perspective of a wind turbine

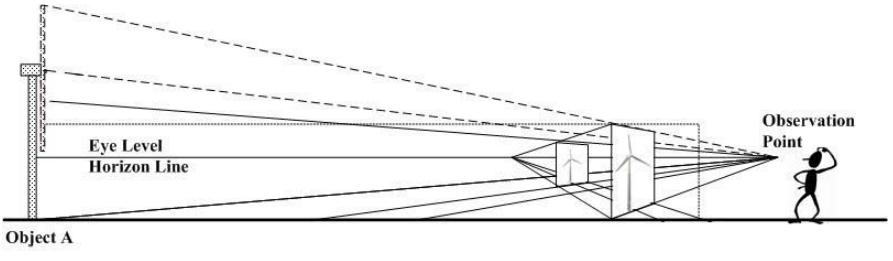


Fig. 6. Relationship between the dimension of perceived object and the distance between the object and the observation point

A lot of studies on landscape planning are focused on attempting to reduce impacts, also trying to hide objects (Gimblett et al 1987). Such experiences are based on the definition of a threshold (Shang and Bishop 2000; Cau and Cocco 2004) beyond which visual impacts could be considered. In this case study, the height of the wind turbine is 80m, so, considering a perception angle of 0.6 degree, actual area of visual impact has a radius of 6.4km. This radius was introduced in simple viewshed analysis calculated for each of 125 turbines.

In visibility percentages, minimum and maximum values where calculated with the aim to fix a threshold of visibility tolerance and to decide which turbines can be accepted and which should be instead rejected. Such threshold was fixed equal to the average value between minimum and maximum percentage.

Lastly, a final consideration was done. Turbines that do not create a group with at least four elements were excluded from the analysis, because it is not economic and neither sustainable to locate a wind farm in a small region.

Table 1. Visibility percentages for each wind turbine

wind turbine	visibility (%)	wind turbine	visibility (%)	wind turbine	visibility (%)	wind turbine	visibility (%)
1	5,9	33	5,0	65	2,7	97	5,1
2	6,0	34	4,7	66	9,3	98	5,6
3	7,9	35	6,8	67	4,5	99	3,8
4	6,3	36	11,5	68	5,2	100	2,8
5	8,7	37	4,2	69	5,7	101	3,6
6	7,0	38	2,3	70	7,4	102	5,0
7	4,9	39	10,9	71	5,6	103	4,0
8	3,9	40	7,6	72	9,0	104	3,9
9	8,3	41	12,2	73	7,6	105	4,2
10	5,3	42	3,6	74	5,1	106	3,2
11	3,0	43	3,3	75	2,4	107	3,4
12	5,1	44	4,1	76	5,4	108	3,7
13	6,8	45	3,8	77	5,2	109	3,7
14	6,9	46	3,3	78	6,7	110	4,1
15	4,4	47	3,4	79	5,3	111	3,8
16	6,2	48	3,7	80	2,3	112	4,6
17	5,3	49	2,7	81	4,7	113	4,3
18	4,9	50	1,9	82	4,0	114	5,4
19	5,3	51	6,7	83	6,8	115	5,1
20	4,2	52	3,1	84	4,4	116	4,9
21	5,6	53	3,1	85	4,2	117	4,4
22	4,0	54	3,2	86	3,4	118	3,8
23	5,2	55	5,0	87	3,3	119	5,1
24	11,0	56	5,9	88	2,9	120	3,8
25	9,1	57	2,9	89	2,5	121	6,1
26	3,1	58	2,8	90	4,4	122	5,1
27	9,1	59	6,4	91	3,5	123	4,5
28	4,2	60	7,5	92	8,0	124	3,1
29	5,2	61	2,6	93	4,0	125	3,2
30	6,7	62	1,6	94	5,9		
31	6,6	63	5,8	95	3,7		
32	4,4	64	8,3	96	2,6		

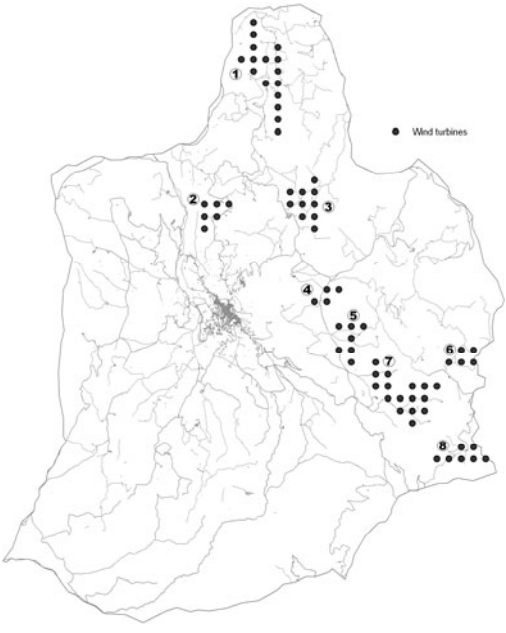


Fig. 7. Result of the combination of geometric rules and simple viewshed analysis

Finally, eight groups of wind turbines were selected as the best choice (figure 7).

5 Visual Impact Assessment

At this point, for each group of wind turbines cumulative and identifying viewshed were calculated in order to quantify visual impact assessment.

Two kinds of visual impacts were evaluated:

- visual impact from some more sensitive observation points, such as monuments and cultural heritage areas;
- global visual impact of each wind farm and how much each turbine contributes to it.

5.1 Visual Impact Assessment from Sensitive Observation Points

About monuments and cultural heritage areas, with the aid of identifying viewshed, it was possible to understand which turbines produce the greatest visual impact. More in particular, groups 3-6-7-8 do not produce visual impact on monuments, while the following results were obtained for the other groups:

Wind farm number 2 has the worst impact because the ratio between maximum number of turbines of the group seen from observation areas and total wind turbines number of the group is 100%, so it is entirely visible. Because of its great visual impact it was excluded from the analysis.

Table 2. Visual impact assessment of each wind farm on monuments and cultural heritage areas

Group number	Monuments and cultural heritage areas affected by visual impact	Wind turbines number of the group	Maximum number of turbines of the group seen from the observation areas	Maximum percentage of the wind farm that has an impact on monuments areas	Turbines seen
1	4	16	5	31.25%	1, 2, 14, 15, 16
2	12	6	6	100%	1, 2, 3, 4, 5, 6
4	2	4	3	75%	1, 3, 4
5	1	7	3	42.86%	5, 6, 7

For what concerns wind farms number 1-4-5, identifying viewshed was used in order to understand which turbines produce impacts, with the aim to remove them from the group.

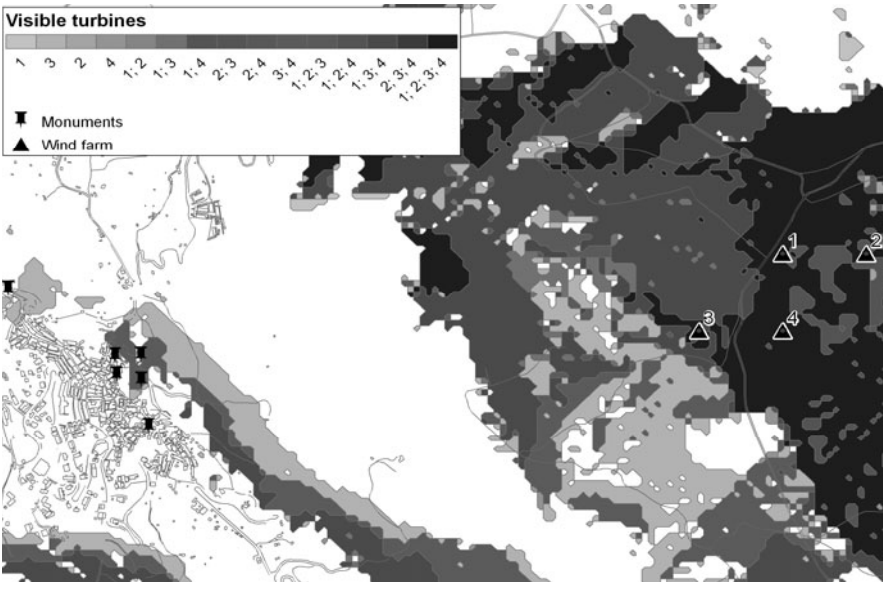


Fig. 8. Identifying viewshed for wind farm option number 4 to evaluate monuments visibility

In this way, even wind farm number 4 resulted excluded from the analysis, because three of its four elements should have been removed and we just decided that the minimum number of turbines in a wind farm had to be 4.

In this way, after the evaluation of visual impact on monuments and cultural heritage areas, only six groups remained; among these, two groups had less elements than at the start of the analysis (n. 1-5) as it was possible to see comparing figures 8 and 9.

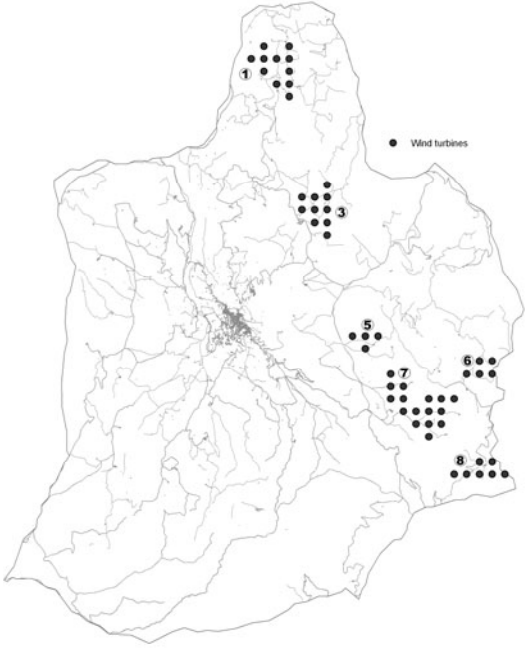


Fig. 9. Remaining wind farms after visual impact assessment on monuments and cultural heritage areas

5.2 Global Visual Impact Assessment

Concerning global visual impact of each wind farm, since each farm has a different number of turbines, it could happen that a little wind farm, does not have an "absolute" less visual impact on landscape, but it is caused by the lower number of elements. Consequently, we looked for a method that allowed us to compare, in an absolute way, not dependent only on turbines number, their impact, using quantitative viewshed.

Quantitative viewshed contains two information: number of turbines visible from each pixel $\mathbf{n_t}$, total area from where it is possible to see $\mathbf{n_t}$, obtained multiplying the total number of pixels $\mathbf{n_t}$ by pixel size.

This total area was then multiplied by a correction factor f, obtained according to the following formula:

$$f = \frac{n_t}{\max(n_t)} \tag{2}$$

where max($\mathbf{n_t}$) is maximum number of turbines visible in all wind farms (in this case 16). A corrected visible area for each wind farm quantitative viewshed was found. Results are reported in a table, where each column represents the corrected visible area calculated for each class of $\mathbf{n_t}$ found in each wind farm. Comparisons were done by rows, so for equal $\mathbf{n_t}$ values, best wind farm resulted always group 11.

Table 3. Corrected visible areas obtained for each quantitative viewshed raster (one for wind farm) and best group choice

n_t	Corrected visible area (mq)						Best group
	Group 3	Group 5	Group 9	Group 10	Group 11	Group 12	
2	290550	394350	272800	100600	80975	171725	11
3	360800	747800	494150	356400	161100	253150	11
4	357975	785700	495600	299925	201150	219750	11
5	357400	865800	664100	781400	262100	504900	11
6	380500	1828375		3937750	266875	530875	11
7	674550	949500			296250	739050	11
8	966000	651000			353850	6288100	11
9	1045000	579000			541200		11
10	1275975	957150			426375		11
11	1769250	5588000			775500		11
12	1542750				811525		11
13					753600		11
14					746200		11
15					1152900		11
16					1791750		11

6 Conclusions

In order to reach the European Union target of 30% reduction of CO_2 emissions within 2020, according to Kyoto Protocol aims, the only way is to reduce the waste of energy and to balance the production of expanding renewable energies. These kind of energies, if on one side should be the cornerstone of emission reductions, on the other side could produce soil consumption and visual impacts. The real problem is that, generally, windiest areas are located on mountain ridges, often with high environmental value. While the assessments of environmental impacts have reached a good level of reliability, visual aspects not always are properly taken into account. Too often wind turbines cover entire ridges, creating a barrier with a visual impact comparable to 50 floors buildings. Despite visual aspects are the most evident ones, evaluating visual impact assessment is not simple,

because it is affected by subjectivity. Analyzing significant and not significant visual impacts, Stamps (1997) considers three main problems: subjectivity, magnitude of impacts and vagueness of language in design. Since many years, every kind of technology has been adopted in order to improve decision making quality. Visual aspects take into account three main factors (Bishop and Karadaglis 1996):

- observer: any person who can be affected by object perception;
- object: human artefacts to be included in landscape;
- environment: every natural or anthropic element between object and observer and behind the object.

These three aspects are strongly related and can never be considered separately (Laurie 1975). Also, Kant (1791) in his "Critique of judgment", considers physical relation between people and environment. Visual impact assessment accounts for relationships between human viewer and landscape properties. Observer location relative to objects defines if the object is visible and how big it appears (Stamps 1997).

Nowadays, computer graphics and GIS can be considered a cornerstone of this analysis in order to increase level of objectivity. The use of multimedia tools in visual impact evaluation shows the following limits:

- such techniques are strongly related to the aesthetic sense of the designer, affecting choices by a certain degree of subjectivity;
- they do not give any information about from where the new object is visible and, consequently, if it obstructs view of town symbols;
- carrying the analysis is onerous, so it is possible to understand if the landmark is visible or obstructed only for a limited number of observation points and for a limited number of design alternatives;
- it is not possible to quantify and compare in a more objective way different alternatives;
- it is not possible to compare visual aspects to other important factors, in new residential zone locations, such as geology, hydrology, accessibility, etc..

The adoption of viewshed analysis represents an important step in improving quality of visual impact analyses, but the most used viewshed models do not collect information concerning single object visibility.

In this paper, starting from the consolidated literature concerning viewshed analysis, a more detailed kind of viewshed has been developed. Identifying Viewshed implemented in GRASS environment overcomes limits and weaknesses of single, multiple and cumulative viewshed allowing more detailed analyses.

The application illustrated in this paper highlights how just one wind turbine is the responsible of the great part of the impacts of a wind farm and its elimination can be enough to remove all visual intrusions producing more detailed assessments.

References

Aguiló, M., Iglesias, E.: Landscape inventory. In: Martínez-Falero, E., González-Alonso, S. (eds.) Quant. Tech. in Landsc. Plan, pp. 47–85. Lewis Publishers Boca Raton, FL (1995)

Benson, E.D., Hansen, J.L., Schwartz, A.L., Smersh, G.T.: Pricing residential amenities: The value of a view. Journal of Real Estate Finance and Economics 16, 55–73 (1998)

Bernetti, G.: Selvicoltura speciale. UTET (1995)

Bishop, I.D., Karadaglis, C.: Combining GIS based environmental modeling and visualization: another window on the modeling process. In: Proceedings of Third Int Conf./Workshop on Integrating GIS and Environ Modeling, Santa Fe, NM, USA, January 21–25 (1996)

Bishop, I.D., Lange, E., Mahbubul, A.M.: Estimation of the influence of view components on high-rise apartment pricing using a public survey and GIS modeling. Environment and Planning B: Planning and Design 31, 439–452 (2004)

Burrough, P., McDonnell, R.: Principles of geographical information systems. Oxford University Press, Oxford (1998)

Camp, R.J., Sinton, D.T., Knight, R.L.: Viewsheds: a complementary management approach to buffer zones. Wildl Soc. Bul. 25, 612–615 (1997)

Cau, G., Cocco, D.: L'impatto ambientale dei sistemi energetici. SGE editoriale, Padova (2004)

Danese, M., Nolè, G., Murgante, B.: Visual Impact Assessment in urban planning. In: Murgante, B., Borruso, G., Lapucci, A. (eds.) Geocomputation and Urban Planning. SCI, vol. 176, pp. 133–144. Springer, Berlin (2009)

DeMers, M.N.: GIS modeling in Raster. John Wiley & Sons, Chichester (2000)

Dransch, D.: The use of different media in visualizing spatial data. Computers & Geosciences 26, 5–9 (2000)

Fisher, P., Farrelly, C., Maddocks, A., Ruggles, C.: Spatial Analysis of Visible Areas from the Bronze Age Cairns of Mull. J. of Archaeol. Sci. 24(7), 581–592 (1997)

Gimblett, H.R., Fitzbibbon, J.E., Bechard, K.P., Wightman, J.A., Itami, R.M.: Procedure for assessing visual quality for landscape planning and management. Environ. Manag. 11(3), 359–367 (1987)

Hanna, K.C.: GIS for land architects. ESRI Press, Redlands (1999)

Hernandez, J., Garcia, L., Ayuga, F.: Assessment of the visual impact made on the landscape by new buildings: a methodology for site selection. Land and Urban Plan 68(1), 15–28 (2004)

Kant, I.: Critique of judgment. Edizione Italiana a cura di Leonardo Amoroso. Fabbri Editore, Milano (1996)

Kidner, D., Sparkes, A., Dorey, M.: GIS and Wind Farm Planning. In: Stillwell, J., Geertman, S., Openshaw, S. (eds.) Geogr. Inf. and Plan., pp. 203–223. Springer, London (1999)

Kim, Y.H., Rana, S., Wise, S.: Exploring multiple viewshed analysis using terrain features and optimisation techniques. Comput. and Geoscis. 30(9-10), 1019–1032 (2004)

Lake, M.W., Woodman, P.E., Mithen, S.J.: Tailoring GIS software for Archaeol Applications: An Example Concerning Viewshed Analysis. J. of Archaeol. Sci. 25, 27–38 (1998)

Laurie, I.C.: Aesthetic factors in visual evaluation. In: Zube, E.H., Brush, R.O., Fabos, J.G. (eds.) Landscape Assessment: Values, Perceptions and Resources, Dowden, Hutchinson and Ross Incorporated, Stroudsberg (1975)

Lee, J., Stucky, D.: On Applying Viewshed Analysis for Determining Least-cost Paths on Digital Elevation Models. International Journal of Geographical Information Science 12(8), 891–905 (1998)

McHarg, I.L.: Design with Nature. John Wiley & Sons Incorporated, Chichester (1969)

Möller, B.: Changing wind-power landscapes: regional assessment of visual impact on land use and population in Northern Jutland, Denmark. Applied Energy 83, 477–494 (2006)

O'Sullivan, D., Turner, A.: Visibility graphs and land visibility analysis. Int. J. of Geogr. Inf. Sci. 15(3), 221–237 (2001)

Ogburn, D.E.: Assessing the level of visibility of cultural objects in past landscapes. J. of Archaeol. Sci. 33, 405–413 (2006)

Ruggles, C.L.N., Medyckyj-Scott, D.J., Gruffydd, A.: Multiple viewshed analysis using GIS and its archaeological application: a case study in northern Mull. In: Andresen, J., Madsen, T., Scollar, I. (eds.) Computing the past: CAA92. Aarhus University Press, Aarhus (1993)

Shang, H.D., Bishop, I.D.: Visual thresholds for detection, recognition and visual impact in landscape settings. J. of Environ. Psychol. 20(2), 125–140 (2000)

Tomlin, C.D.: Geographic information systems and cartographic modeling. Prentice-Hall, Englewood Cliffs (1990)

Wheatley, D.: Cumulative viewshed analysis: a GIS-based method for investigating intervisibility, and its archaeological application. In: Archaeology and GIS: A European Perspective. Routlege, London (1995)

Wheatley, D.W., Gillings, M.: Visual perception and GIS: developing enriched approaches to the study of archaeological visibility. In: Lock, G. (ed.) Beyond the map: Archaeology and Spatial Technologies. IOS Press, Amsterdam (2000)

Yang, P.P., Putra, S.Y., Li, W.: Viewsphere: a GIS-based 3D visibility analysis for urban-design evaluation. Environ. and Plan B: Plan and Design 34, 971–992 (2007)

Agricultural Terraced Landscapes in the Province of Trieste (Northeastern Italy)

Giovanni Mauro

Department of Geographical and Historical Sciences, University of Trieste,
Via Tigor, 22, 34124 Trieste, Italy
email: gmauro@units.it

Abstract. Terraced agricultural areas represent a beautiful but problematic landscape, because their current maintenance involves great diseconomies. However the recent agricultural European policy, more careful to the environmental problems, could mean several new opportunities for this kind of rural settlement. This paper analyzes the evolution of the terraced landscapes in the Province of Trieste (North-eastern Italy) in the 1955-2000 periods. So, we studied the recent development of agriculture for this territory and the local traditional agricultural terraced settlement, here generally called *pastini*. In order to locate and estimate the size of terraced areas, we applied a GIS overlaying analysis, using an official European cartography, MOLAND, and a digital elevation model (DEM). MOLAND (*Monitoring Land Use / Cover Dynamics*) is a research project supported by the European Commission, giving as result several land-cover maps at medium spatial resolution (scale 1:25,000). Considering slope parameters, we estimated the recent development of terraced areas, identifying wider settled hill zones in Trieste Province. Finally, we briefly described three rural areas well-characterized by a stepped appearance landscape.

Keywords: MOLAND, DEM, GIS analysis, Vineyards, Olive grove.

1 Terraced Agricultural Areas: Problems and Opportunities

In Italy and in all the Mediterranean basin terraced landscapes are very spread, so they are estimated as being among the most important and characteristic anthropological settlements on relief (Trischitta 2005). Terraced landscape areas symbolize an important European cultural heritage, a sort of "middle landscape" as result of an environmental, social, economic and aesthetic mediation (Varotto 2008). They represent a traditional land use, sometimes highly adapted to the natural landscape structure: in fact, they are the only way to prevent the erosion and to enable the cultivation of the hill or mountain land. However, in the last century the maximization of productivity changed our agriculture system, so the maintenance of these particular agricultural areas has been very problematic. The hard mechanization of the primary sector and hence the reduction of people in agriculture produced great

B. Murgante et al. (Eds.): Geocomputation, Sustainability & Environ. Planning, SCI 348, pp. 91–109.
springerlink.com

diseconomies in the maintenance of terraced areas. Landslides, erosion and natural reforestation are the common problems causing the decline of agricultural terraces. Another problem is the urban sprawl: actually, terraced areas are frequently located in beautiful places near to the urban areas, so often they are transformed into residential neighborhoods (Cori et al. 1993).

Nevertheless, in the last decades the European policy gave an important relevance to sustainability development principles also in the farming, as an ecological answer to the great environmental impacts produced by the primary sector. So, the relevance admitted to landscape for the tourism in a sustainable development perspective could represent an important role in order to recover terraced landscapes. Maintenance, care and development of the cultivated zones could mean promotion of nature and landscape, or rather promotion of typical local and regional products. Then, terraced areas could become an identity landscape for the coastal area of Trieste Province too: "a whole of characteristics of the landscape structure and its symbolic meanings, by which each individual landscape area can be recognized and identified" (Hudoklin 2008).

To plan these regions in a sustainable way it is first at all very important to define their location, but the quantitative estimation of terraced areas represents a key problem: difficulties in realising maps are put in evidence by different authors. Brancucci and Paliaga (2008) underline issues to identify terraced areas, also when update cartography or orthophotos are available in a GIS system. The reasons could be various: working at a large scale (1:2,000), vegetated plants or infrastructures could cover terraced areas or, sometimes, the analyser could muddy these settlements with geological structures. Also mistakes in the official mapping process could mean difficulties in map visual rendering. Direct census could not solve the problem too, because it could be difficult to approach with private owner or to reach the terrace owing to the luxuriant vegetation, etc. Analytical studies are carried out by Terranova (1989) about linear lengths of walls for hectare, but this approach is very expensive, mainly for the survey activity.

So many authors tried to define a common methodology to quantify these geographical elements. Varotto (2006), for example, proposed using different indices - like extension index or intensity index[1] - in a GIS environment in order to compare different terraced landscapes in Veneto region using the same methodology. Also Mautone and Ronza (2006) suggest a GIS application (overlaying orthophotos, digital terrain models, land uses, official cartography of Campania Region) to define terraced areas in the Amalfi Coast. Scaramellini (2005, 2006) proposed a common research protocol[2] for geographers studying this kind of landscape. It is an empirical methodology that classifies terraced landscape in areas as:

[1] Extension index: terraced surface on total surface ratio. Intensity index: wall linear length on terraced surface ratio.

[2] Briefly, this protocol research prefigures five steps: 1) a careful analysis of the general building factors (phisical, morphological, agronomic, etc.); 2) informations about materials used in the terracing processes; 3) sizing of the terraced areas; 4) analysis of terraced areas typology (informations about building typology, land use, etc.); 5) preservation and utilization state.

micro-terraced areas, when only a little share of total, no more then one third of total, is set in this way; meso-terraced areas when a consistent, less than two thirds, settlement involved hill side; macro-terraced areas when almost the total study area is terraced.

In this study we applied a GIS overlaying analysis, in order to estimate the size of terraced areas in the Trieste Province (North-eastern Italy) in the 1955-2000 periods, mainly using an official European cartography, as the MOLAND (see paragraph 5).

In the rural landscape of Karst there are different cultural settlements of hill side: considering the overall low slope conditions of this territory, generally there are terrain embankments following the contour line, but terraced areas or "step" rural areas (shorter) are also well spread. On the coastal zone, between Duino and Contovello (near Trieste), these kind of agricultural settlements (with vineyards and olive grove) are widespread, but they are present in the Karst plateau too[3] (Panjek, 2003). In the past, farmers created terraced cultivation areas on the steep fall, building supporting walls. Women brought worldly using a panier, named *nečka*. On the terraces, local known as *pastini* or *paštni*[4], rural people mainly grow vineyards and olive tree.

2 The Study Area: The Province of Trieste

The Province of Trieste (figure 1) is the smallest province in Italy (about 212 Km^2): it is a long (about 30km) and narrow (about 5-6 km) territory, placed in a NW-SE direction, characterized by contemporaneous presence of two main geographic elements: the calcareous plateau, the Karst (about 300 metres above sea level) and the close coastal area (locally called *Costiera Triestina*), enlarging in south direction and having a promontory in the municipality of Muggia (Valussi 2000).

Karst, in Italian "Carso", is the well know Cretaceous limestone plateau region between Northeastern Italy and Southwestern Slovenia and it is famous for its caves[5]. Karstic landforms and processes characterize the plateau from Duino municipality at North to Basovizza - a little village near Trieste - at South. The Karst of Trieste area is composed of a limestone series with a 1,800m depth in the Doberdò area (Martinis, 1971). However, in the neighborhoods of Trieste the flysch appears, the sedimentary rock composed from marl and sandstone alternate layers: this geological formation is extended on the entire narrow coastal area, from Sistiana to Muggia (figure 2).

---

[3] In particular, sometimes in the past sinkholes also are terraced, because in these natural depressions soil and climate conditions are favourable to cultivation and this was very relevant in a land, like the Karst, generally not favourable to agriculture (Panjek, 2003).

[4] This word derived from Latin "*pastinum*", "ground worked with the marra (a sort of hoe)". In Latin the verb "*pastinare*" means also "work the ground to plant vineyards" (AAVV 1990).

[5] In the Province of Trieste there is the largest cave in the world, the "Grotta Gigante" (the name means "Giant Cave").

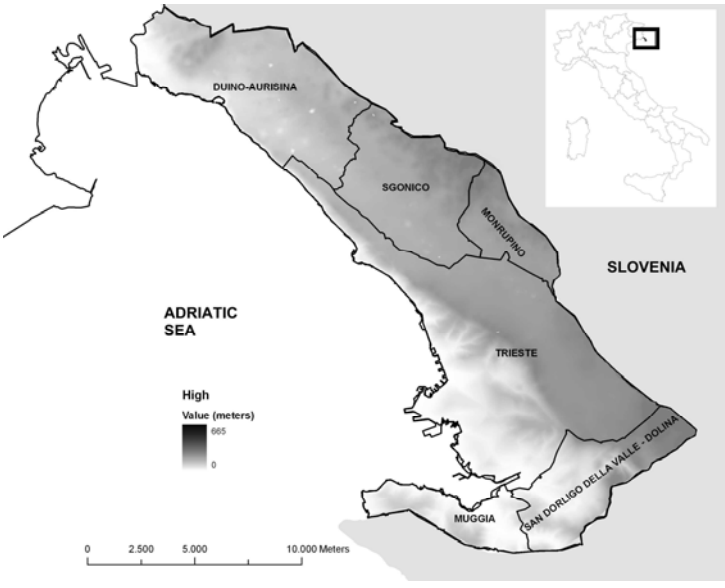


Fig. 1. Location, municipalities and morphology of the Trieste Province

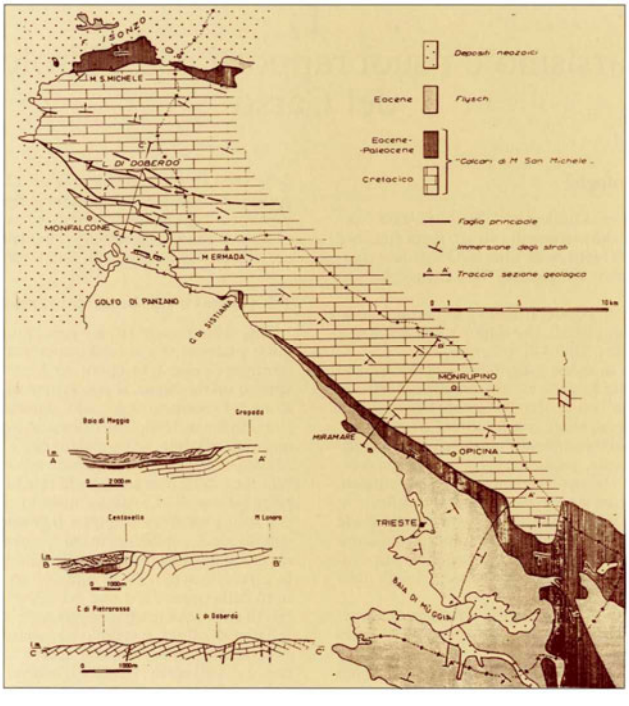


Fig. 2. Geological map of Trieste Province (source: Martinis, 1971)

With regards to the climate, the Karst is a plateau geographically located at the end of a warm sea like the Adriatic and surrounded by high mountains to the East and North directions. For these reasons the sea influence is very accentuated in the narrow coastal area (average temperature 13.8 °C), but also in the plateau - although it is less exposed to the beneficial climatological effects of the sea[6] (Fakin et al. 2004). A climatologically characteristic element is a very intensive freezing and dry Northeast wind, blowing in the winter season: the "Bora", a wind draining the karstic soil, wich is just a few centimeters deep and barren.

These soil and climate differences set the conditions for a hornbeam (*Ostrya carpinifolia*) and pubescent oak (*Quercus pubescens*) forest on the Karst and a Mediterranean forest (mainly holm oak, *Quercus ilex*, and hornbeam, *Ostrya carpinifolia*) on the coastal zone (Del Favero 1998).

Several different ecological conditions characterize this narrow territory, so in the Trieste Province there is a high density of protected areas[7]. However, in a "delicate balance", there is also a high density of road network (like highway, railway, primary and secondary roads), connecting urban areas and local villages to the Friuli Venezia Giulia Region or the neighbouring Slovenia.

Concerning to the settlement, while in the Karst plateau the small size settlements with up to 200 inhabitants prevails, Trieste[8] and his 205,000 inhabitants (ISTAT website, http://www.demoistat.it) is the major urban area of Friuli Venezia Giulia Region. Besides, in the Trieste Province lives the most important Slovenian community of the Region: it is mainly located in the karstic municipalities (Slovenians people are the majority in Sgonico, Monrupino and Dolina-San Dorligo della Valle), but also in the peripheral districts in the karstic areas of Trieste - as the village of Opicina (Battisti 1979). So, the Karst and the urban area of Trieste represent two different districts for many reasons: terrain morphology, geology, climate, forestry, human settlement, etc. Morris (2003) suggests the idea of the Karst like a "wild front", a sort of natural presence surrounding the Trieste city district.

2.1 The Agriculture in Trieste's Province

With reference to agriculture, it does not represent an utmost important economic activity in the six municipalities of the Province, although some quality produc-

[6] The position close to the sea influences the precipitations amount, in particular in the coastal area (1,075 mm/year), while in the internal zone they increase remarkably (1,400-1,500 mm/year) (Fakin et al. 2004).

[7] In Trieste Province there are the following protected areas: "Falesie di Duino", "Monte Lanaro", "Monte Orsario", "Val Rosandra" (Azienda Parchi e Foreste Regionale 1999).

[8] Regarding to Trieste urban plan, it follows the hill morphology of the surrounding landscape. Starting from the second half of the eighteenth century to 1955 - a year characterized by maximum number of inhabitants ever reached, about 280,000 - the urban growth developed in the Southeast direction, actually on the neighbouring rural landscape. In the last decades, Trieste's demographic trend has been decreasing.

tions have been set in the last few years: the karstic phenomena and the consequent water lack represent all serious problems in farming and forestry activities. Besides there was a strong decrease in the second Post-war, as already Battisti (1979; 2000) put in evidence (Table 1). The reasons can be referred to the morphology of this land, as supporting economically a not-mechanized agriculture represent a difficult task. Besides, the urban sprawl of Trieste and the agricultural activity defection of local rural population are other relevant processes setting a great acceleration to the changes in the cultural landscape (Panjek 2003)[9].

However, the Province's morphology warrants slope conditions very suitable to the wine and olive tree farming. So, also following the "new" agricultural European policy, in the last decades vineyards and olive grove are ever-growing, supporting ancient local agriculture's habits.

In fact, the relevant role of viticulture in this land was put in evidence also from Plinius "The Old", in his famous quotation in the *Historia naturalis* about *vinum nigerrimum* ("the very black wine"). Someone relates this wine to the Pucino, a table wine that was very famed for its peculiar properties (Plinius "The Old" suggests that this wine is a sort of "source of longevity" for the wife of emperor Octavianus Augustus, Livia Drusilla) (Covaz and Turel 2006).

Table 1. Land use in Trieste Province (source: census of ISTAT,1951-2001)

Land Use	1951	1961	1971	1981	1991	2001
	Ha	Ha	Ha	Ha	Ha	Ha
Arable land	2,526	1,658	701	328	306	155
Vineyards, Orchards	1,517	1,280	560	396	343	259
Pastures	7,738	6,313	4,918	2,981	2,259	1,785
Forest	4,927	6,086	3,061	2,826	3,033	2,359
Uncultivated areas	3,722	2,744	876	900	714	112
Total	20,430	18,081	10,116	7,431	6,655	4,670

With reference to the olive grove, until the cold winter of 1929[10] farmers also grew olive tree on the coastal zone (as suggested, for example, from the name of a little suburb near Aurisina, Oljščica). The mass extinction and, maybe, political condition changes[11] caused the defection of this cultivation. Since 1970, the Regional Government of Friuli Venezia Giulia promoted new olive tree plants

[9] For example, the end of animal husbandry is the main reason of the recent reforestation process in the Karst.

[10] The winter of 1929 was very cold, causing the death of olive trees in all the high Adriatic area.The very cold winter of 1956 represented the deathblow for this kind of cultivation.

[11] In Italy Fascist Government introduced a new onerous tax to the farmers replanting the olive grove (sometimes linked to Slovenian agricultural world traditions) (Covaz and Turel 2006).

Fig. 3. Olive tree cultivation on the terraced areas near Crociata di Prebenico (Dolina-S. Dorligo della Valle municipality) (source: Mauro G.)

projects, in Muggia and Dolina – San Dorligo della Valle[12], above all (RAFVG, 2007) (figure 3). Today olive grove farming is growing and it represents one of the most important reality in the local rural context.

The relevance of olive cultivation has been recently officially recognized: in fact, since 2001 the olive oil produced in this area is labelled as a Protected Designation of Origin (PDO), named "Tergeste DOP", a label set to protect the name of this regional food. Actually, this trend follows the ever-growing relevant role of the olive tree cultivation in Istria peninsula, today representing an important phenomenon[13].

3 Sources of Data

As suggested before, in order to learn where there are wider terraced areas in Trieste Province, we compared some landcover maps of different years, named MOLAND, and the digital elevation model (DEM) in a GIS application.

The recent map product named MOLAND (*Monitoring Land Use / Cover Dynamics*) consists in a research project supported by the European Commission and

[12] In the period 1982-2004 olive areas are increased from 18 hectares till 100 hectares (the growing involved especially the Dolina – S. Dorligo della Valle municipality).

[13] In fact in Istria there are more than 31,000 hectares (30,000 Ha in Croatia and 1,600 Ha in Slovenia) of olive tree cultivated areas.

coordinated by the Institute for Environment and Sustainability of the European Commission's Joint Research Centre (Barredo et al. 2003). As suggested from the name, the main aim of MOLAND is "to provide a spatial planning tool that can be used for assessing, monitoring and modeling the development of urban and regional environments". The Friuli Venezia Giulia Region promoted this study, concerning land-cover changes and land-use evolution over a period of forty-five years, at four dates: 1955, 1975, 1985, 2000. The resulting maps (scale 1:25,000, coordinate reference system: Gauss Boaga, Rome 40, East) are available on the official Website of Friuli Venezia Giulia Region (http://www.regione.fvg.it).

The regional digital elevation model (DEM) was created from the quoted point of official digital cartography of Friuli Venezia Giulia Region. Its spatial resolution is 25 meters and the coordinate reference system is the same as MOLAND. As these previous maps, it is available on the official site of the administration.

Besides, in order to verify our results we also analysed old maps (IGM, Italian Military Geographical Institute 1962), some old ortophotos (1974), the official digital cartography of Friuli Venezia Giulia Region (1988-1989) and the spatial high resolution satellite images of GoogleEarth[14] (2003, 2004, 2005).

4 Methodology

The MOLAND land use legend is very exhaustive, it is similar to CORINE Land Cover classification, but more detailed. It includes about a hundred land use classes, classifiable in five macro-classes concerning human settlement, rural areas, forests, marshes and hydrosphere (rivers, lakes and sea). In this case, we considered only the macro-classes of rural areas, in detail ("name"): "not irrigated arable land" (code 2.1.1); "vineyards" (code 2.2.1); "pastures" (code 2.3.1); "complex cultivation patterns without settlement" (code 2.4.2.1); "complex cultivation patterns with scattered settlement" (code 2.4.2.2); "land principally occupied by agriculture, with significant areas of natural vegetation" (code 2.4.3).

In light of the pursuit for a sustainable development, the main aim of this study is to identify the recent evolution of terraced areas, representing the first step to plain their recover. So, we evaluated slope edge values useful for this agricultural settlement: technical references put in evidence that the terraced areas are developed from a minimum of 5% slope (under the land is plan) to a maximum of 45% (in higher slope conditions the width of terraced areas is very reduced) (Bonarrigo 2006).

We analyzed the hill side aspect, also: to warrant better sunlight conditions, the terraced areas are generally oriented southwest. However the morphology and the displacement of Trieste Province (a long and narrow territory in a NW-SE direction) warrant good sunlight condition almost always. So, we often verified

[14] We desired to have satellite images of the 2000 period, but the high resolution images available on GoogleEarth for the Trieste Province are Quickbird images (spatial resolution: 0,7m), dated 28 october 2003, 2th march 2004 and 10 february 2005.

the presence of terraced areas exposed north-east, therefore we did not consider this variable.

For these reasons, we created the following maps: four rural land-cover class (extracted from MOLAND) maps for the four considered periods (1955, 1975, 1985 and 2000) and a map of morphology derived from DEM, selecting the 5-45% slope areas. Then we crossed our resulting maps (a rural map with slope map for each considered year), to estimate the past or the current presence of terraced landscapes.

Fig. 4. Overlay of terraced area maps of 1975 on the ortophoto (1974) near Prosecco (Trieste municipality). It is easy to note that our MOLAND derived maps not cover all the terraced landscape (for example, on the left between the sea and the main road)

In order to verify the quality of the estimates deriving from our GIS application, we overlaid our results on several not homogeneous maps. Estimated terraced areas maps of 1955 have been compared with old IGM maps (1962, 1:25,000 scale). Old ortophotos (1974) are used to verify the 1975 terraced areas map. Official digital cartography of Friuli Venezia Giulia (1989-1990) helps us to analyse 1985 map results. Finally, to understand 2000 terraced area map quality, we overlaid our results on the GoogleEarth raster maps[15]: for the Trieste Province, in the latest version of this software (GoogleEarth 5.0) high resolution images acquired in 2003-2004 are available.

We compared our resulting maps on the reference maps in different ways. For the IGM maps or the official digital cartography, for example, we considered land use symbology and the distance between contour line (to study the slope of hills). For the images (orthophotos and GoogleEarth satellite images) we studied their texture, colours (where it is possible), the size, shape and the vegetation coverage, too. In addition, we applied the 3D effects of the GoogleEarth virtual globe to better understand the location of terraced landscapes (figure 5).

Overall, we must underline that these are empirical methodologies, based on visual interpretation: we could not made a "classic"[16] accuracy assessment, not being able to define a common set of points to compare all our resulting maps. In fact several changes, like reforestation and urban growth, involved this region in the 1955-2000 period, so it is problematic to identify many points where terraced areas existed in all along this period. Besides, a symbology for terraced areas unfortunately does not exist in any of considered maps, so we had any reliable maps as reference. However, MOLAND already is an official map, created following the European's quality standards. So the following analysis on these derived maps must be considered only a further authentication, in order to understand drawbacks and attributes of our study.

Following this approach we realized that our methodology provides good quality results, also if terraced areas generally could be undersized for the old maps (1955 and 1975), whereas they seem overvalued for the most recent, in particular for the 2000. In fact, if we overlay our resulting maps on the ortophotos (figure 4), we simply notice that our MOLAND derived products do not cover all the terraced landscape. On the contrary, when we overlay our maps on the high resolution satellite images of GoogleEarth (figure 5), sometimes we identify reforested areas as terraced landscapes.

The reasons behind this problem could be found in the land-cover class of MOLAND code 2.4.3 ("Land principally occupied by agriculture, with significant

[15] We reprojected our files in the WGS84 datum, Plate Carée projection then we transformed our shapefiles into .kml files.

[16] This procedure mainly consists of the preparation of a classification error matrix, in order to compare, on a category-by-category basis, the relationship between know reference data (ground truth) and the corresponding results of a generated thematic map (Lillesand and Kiefer 1999).

areas of natural vegetation") does not allow a sharp separation of cultivated fields from naturally vegetated areas. This means an addictional analysis to exactly quantify the reforestation processes in the past - maybe overestimated - and at present – possibly underestimated.

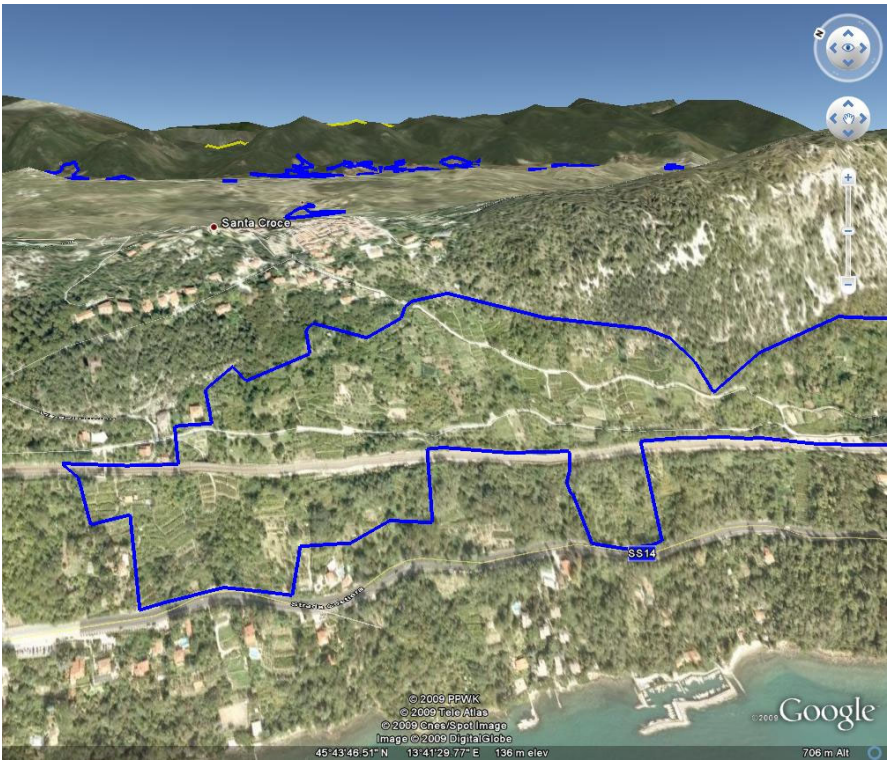


Fig. 5. Overlay of the terraced areas map in 2000 on the satellite image (2003) of GoogleEarth (improved with 3D effect). Some errors occourred, as we identified a reforested areas (visible on the satellite image) as terraced landscapes (on the MOLAND derived map).

5 Results

First of all, we discovered an overall decrease of terraced landscape in the 1955-2000 period (Table 2), following the agriculture negative trend of the study area (see paragraph 4). However, while areas of the primary sector fell from 12,000 hectares in 1955 to 2,200 hectares in 2000 (see table 1), the reduction of terraced landscape is smaller and higly centered in two decades, between the years 1955 and 1975.

Moreover, the changes did not involve all the municipalities in the same way (figure 6). The urban growth of Trieste[17], for example, is the main reason of terraced landscape decrease in this municipality: the agricultural settlements neighbouring the town were gradually reduced by new peripheral working-class neighbourhoods, mainly in the south-east (figure 7), near to the industrial zone. In opposite, Dolina - San Dorligo della Valle municipality experienced a terraced areas growth between 1975 and 1985, probably linked to the recovery of the olive cultivation.

Moreover, from 1955 to 2000 we also noticed an overall, slow and continuous partitioning of this kind of landscape in all the study area, Karst and coastal zone.

Table 2. Terraced areas (estimated surfaces in hectares) for the municipalities of Trieste Province in the 1955 – 2000 period

Year	Dolina	Duino	Monrupino	Muggia	Sgonico	Treieste	Total
1955	321	104	32	430	47	582	1,516
1975	251	102	25	294	41	354	1,067
1985	259	97	22	291	39	317	1,052
2000	245	96	22	293	39	314	1,009

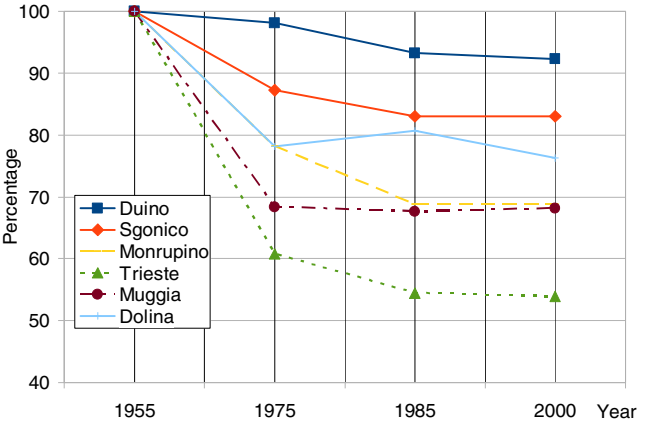


Fig. 6. Percentage decrease of estimated terraced areas in the municipalities of Trieste Province in the 1955-2000 period

[17] Already Battisti (1979) put in evidence the twofold increase of the urban area surface in the 1947-1977 period.

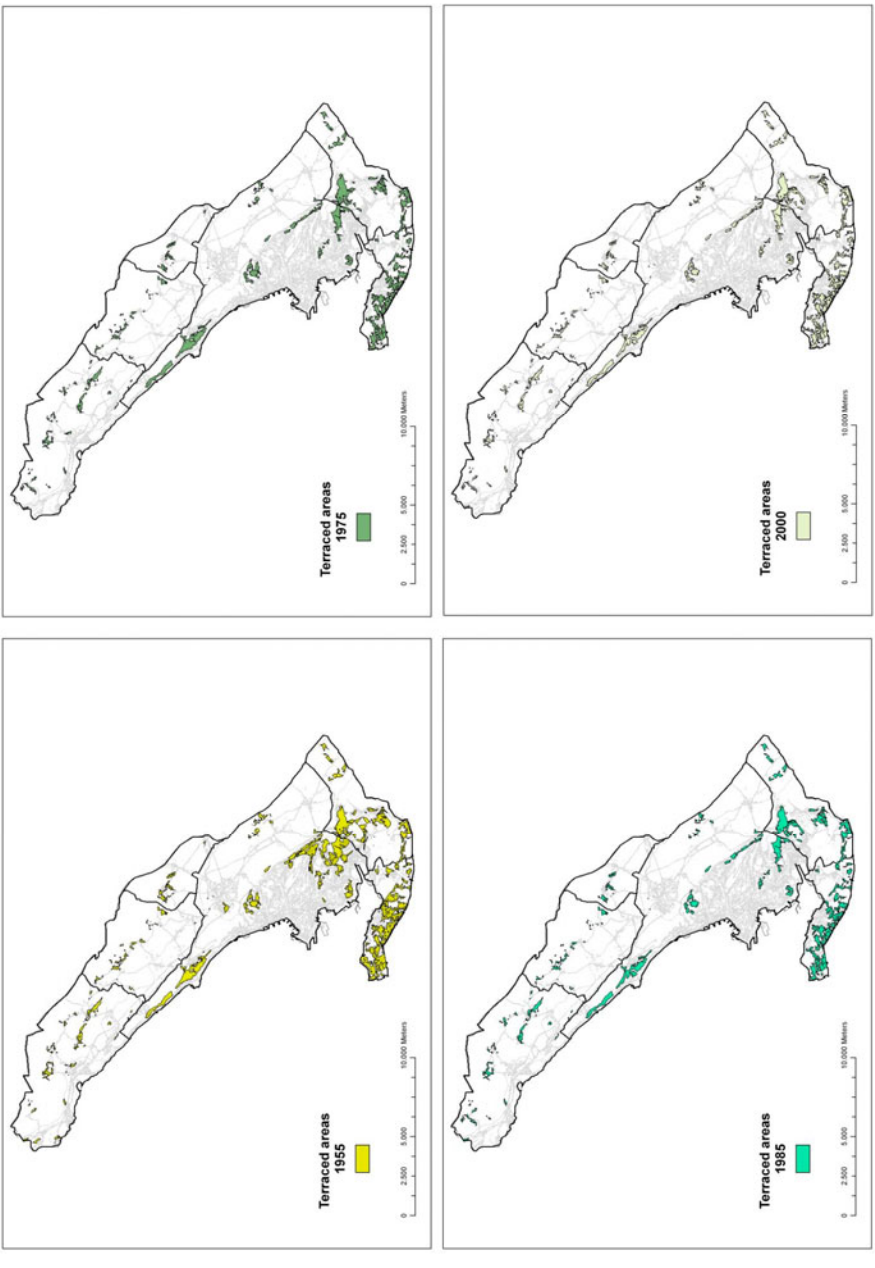


Fig. 7. Estimated location and development of terraced landscapes in Trieste Province (period 1955-2000)

Nonetheless, some wide hill settlements remain quite constant in terms of size and presence: we analyzed their density[18], so we identified three relevant high density terraced areas (figure 8). They are located on the promontory of Muggia (1), in an area between Trieste and Dolina – San Dorligo della Valle (near Bagnoli della Rosandra) (2) and on the costal zone (a long and narrow area between Prosecco and S. Croce villages, Trieste municipality) (3).

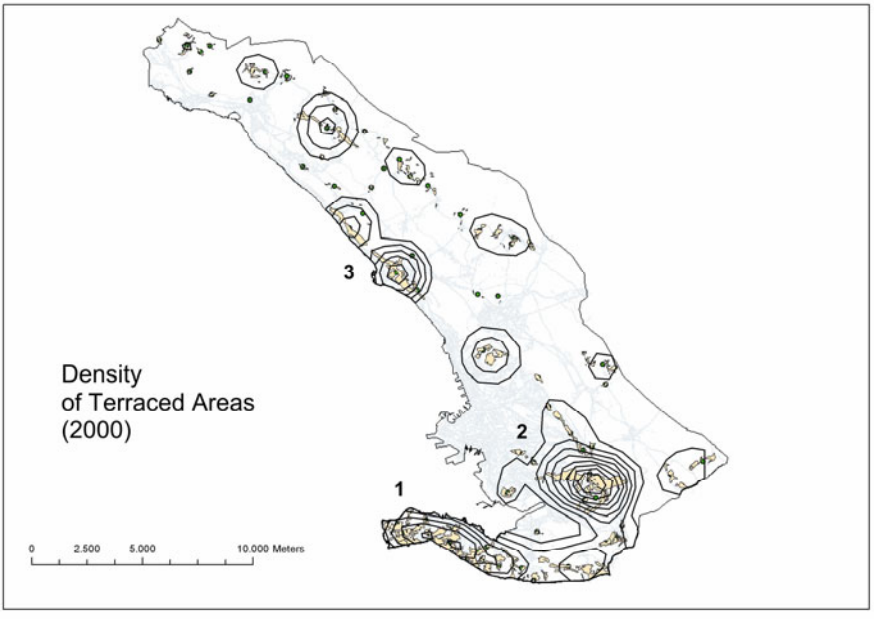


Fig. 8. Estimated terraced areas in the Province of Trieste (2000)

Placed at south of Trieste Province and near to the Italy-Slovenia State border, the first settlement consists in a meso-terraced area (Scaramellini 2006), very vegetated. Actually, it is a landscape characterized by a low slope, so the agricultural hill settlements generally use embankments, disposed in parallel along the same high contour line. Sometimes, sandstone walls split house's gardens in this high residential density area (so different from the bordering Slovenian landscape, characterized by open vineyards; figure 9). On average the field surface is very little, therefore the length of the embankment (or wall) is a few tens of meters long. Olive grove is the main farming, in particular the local variety (the *bianchera-belica*), characterized by high productivity and climate resistance. However, the simultaneous growing of other fruit trees is very common.

[18] We applied a Kernel density calculation for points features (each representing a terraced area) and we considered also the surface of each terraced area.

Fig. 9. On the background the meso-terraced area on the promontory of Muggia (Italy), in opposite with the Slovenian extended vineyards (source: Mauro G.)

The area near Bagnoli della Rosandra is wider and rural, where terraced a landscape put in evidence both traditional settlements (figure 3, Dolina - S.Dorligo della Valle municipality) and also new "standard" settlements – close to an industrial zone - realized using a mechanized system. This is a meso-terraced vegetated area and, as the previous case, the low slope warrants more simple morphology's conditions, so usually hill settlements are embankments. Rare walls bordering terraced areas are built up using sandstone, but more frequent calcareous rocks. The residential density is lower than on the Muggia promontory, so this landscape surrounded by the Karst plateau appears very natural. The gradual slope warrants a field's surface (on the terraced areas) wider than in the previous case. Olive grove is the most spread cultivation, although it is not infrequent to find some short vineyard lines.

A different context is represented by terraced areas on the coastal zone. Worrying building speculation and urban sprawl regards this area of Trieste Province, sometimes compromising the delicate ecological balance of the beautiful landscape. In a territory geologically characterised by flysch, frequent micro-landslides run-into new constructions, at times quitted for this reason. Reforestation is in great progression, erasing all the recent past human evidences, mainly for the rural world. In this bad context, the rural landscape surrounding the karstic

small villages of Santa Croce and Prosecco represents outstanding conditions. Connected to Trieste[19] by a scenic and little used road, named *Pucino* street (recalling the ancient vine celebrated by Plinius "The Old", see paragraph 2.1), the first village is located just on the Karst edge, nearly located at 250 meters above the sea. A similar situation regards also Prosecco, a rural village located in the plateau, but closer to Trieste urban area.

Till the end of the Second World War, their residents were fishermen and/or farmers mainly. They took out calcareous rocks from the ground, then they harvested little crops on the Karst plateau and the *Costiera Triestina*. Farmlands along the hill-side are preferable, because the soil and the climate warrant better conditions to vineyards growing (Semerani et al. 1970). Recent social and economic changes involved this community also, inducing a marked decline of agricultural landscape, above all in the Karst, and its rural society. It also means a give up of the terraced areas, especially for the most faraway crops from rural villages. As suggested from our visual analysis (figure 4 and figure 5), the road network presence - like a high-used road or the railway - could represent another reason of the cropping release: at present, the relevant terraced zones located between the road named *costiera* and the sea at the 1950 has quite disappeared.

However, as suggested before, a renewed interest, connected to the sustainable agriculture concept and a new tourism model, is involving this traditional rural settlement. So, in the last years, different recovering processes are carried out also in these micro-terraced vegetated areas near Santa Croce and Prosecco, and therefore it is not infrequent to spot a terrace rebuilding or a deforestation of a field recently forested. Spontaneous enterprises are also supported from the current local planning regulations[20], although they are ruled by a rigorous bureaucratic procedure: for example, to rebuild a terrace a detailed (scale 1:100) recovering project is requested[21].

High walls bordering very little terraced areas and surrounding local network roads are very spread, but they are short; calcareous rocks are the material commonly used (figure 10), but, when the terrace is faraway from the village - therefore from the Karst plateau - calcareous rocks are replaced by sandstone. Vineyards are the main farming of these areas: *terrano* is the local well-know wine, sold in the peculiar and spread local taverns called *osmizza*[22] .

---

[19] The Pucino road connects S. Croce to Miramare Castle, a well-know monument located at Trieste urban areas beginning.

[20] In the last town-plan, under a general recover policy of vernacular architecture, there are foreseen incentives for the restoration of the karstic walls and terraced cultivation areas. However the municipal Council did not implement the "Semerani" plan (2000), setting serious restrictions to the building speculations in coastal area, because it was "in conflict" with the General Plan of Trieste (Covaz and Turel, 2006).

[21] This project can be carried out only after a favourable judgement of a special municipality commission and of the local forestry authority, too.

[22] In the "osmizze" the wine is directly sold by the farmer, often in his wine cellar. The name derived from "osmica", pointing out ancient short opening period, only few days (height or "osem" days in Slovene language) allowed by local authority.

Fig. 10. Rebuilt wall bordering a terraced area near S. Croce, Trieste municipality (source: Antonini M.)

6 Conclusions

The terraced landscape put in evidence problems and opportunities also in the Trieste Province: property speculation, micro landslides and, mainly, reforestation represent worrying and diffused events. However new investments, connected to the olive oil production - also by a certificate of origin, the "Tergeste DOP" - and to the wine production - the *terrano* - highlight the new potentialities.

To plan the recovering of terraced landscape in a sustainable development key, this study proposed an overlay GIS methodology to identify and to quantify this particular land use. So, we pointed out some drawbacks of the methodology, mainly connected to the land-cover map (MOLAND), but very useful to put in evidence recent development of terraced landscapes and to locate wider settled areas. In a region where the economic role of agriculture fell in the last decades, we pointed out an overall decrease of terraced landscape, caused mainly by the urban growth. However, we recognized three relevant agricultural terraced areas very vegetated, but fairly constant in the 1955-2000 period, two located at South and near to the Trieste urban area and another placed on the coastal zone. Several general building factors (low slope and vegetated area in a high residential density area), technical characteristics (mainly hill settlements realized by embankments) and cultivations (olive tree) are the same in the two areas at south. The terraced landscape on the coastal zone is quite different, mainly for terrain morphology (slope very high), building characteristics (mainly short walls) and cultivation (vineyards). An update interest regards this traditional rural settlement also in the Trieste Province, so different recovering processes of embankments or rebuilding activities regarding the walls are frequently carried out.

Acknowledgements. We would like to thank Antonini Mauro for precious information about the rebuilding walls activities on the terraced landscapes of coastal area.

References

AAVV.: Dizionario di toponomastica: storia e significato dei nomi geografici italiani, UTET Torino (1990)

Azienda dei Parchi e delle Foreste Regionali: Aree naturali protette – Parchi, Riserve e Biotopi nel Friuli Venezia Giulia, Regione Autonoma Friuli Venezia Giulia, Udine (1999)

Barredo, J.I., Lavalle, C., Demicheli, L., Kasanko, M., McCormick, N.: Sustainable urban and regional planning: the MOLAND activities on urban scenario modelling and forecast, Institute for Environment and Sustainability, Join Research Center. European Commission, EUR 20673 EN (2003), http://moland.jrc.ec.europa.eu/index.htm

Battisti, G.: Una regione per Trieste, Studio di Geografia Politica ed Economica, Del Bianco Udine (1979)

Battisti, G.: La città, il porto, la regione. Gli elementi fondativi di un'economia urbana. In: Dipartimento di Scienze Geografiche e Storiche, Trieste, vol. 2 (2000)

Bonarrigo, C.: Fattori agronomici dei terrazzamenti e in generale della sistemazione a pendio, Geotema, 29, Anno X, Patron Editore Bologna (2006)

Brancucci, G., Paliaga, G.: The problems with mapping: the case of Liguria, Terraced landscapes of the Alps. In: Scaramelli, G., Varotto, M. (eds.) Atlas. ALPTER Project. Marsilio, Venezia (2008), http://www.alpter.net

Brščić, K., Poljuha, D., Krapac, M.: Olive residues – Renewable source of energy. Management of Technology – Step to Sustainable Production, Sibenik, Croatia (2009)

Cori, B., Corna Pellegrini, G., Dematteis, G., Pienotti, P.: Geografia urbana, UTET Torino (1993)

Covaz. R., Turel, A.: La Costiera triestina, Storia e misteri di una strada. MGS Press Trieste (2006)

Del Favero, R. (ed.): La vegetazione forestale e la selvicoltura nella regione Friuli-Venezia Giulia, Regione Autonoma Friuli-Venezia Giulia, Direzione Regionale delle Foreste e della Caccia, Servizio della Selvicoltura, Udine (1998)

Fakin, J., Jazbec, T., Legiša, Z., Pertot, M., Tuta Ban, V.: Il Carso da Štanjel a Duino/Devin. Comune, Komen (2004)

Hudoklin, J.: Terraced landscapes: areas of landscape identity in changing processes. In: International Conference Living Terraced Landscapes, Project ALPTER, Ljubljana (2008), http://www.alpter.net

ISTAT: Censimenti dell'Agricoltura, 1951-2001, Analisi Provinciali, Roma (1951-2001)

Lillesand, T.M., Kiefer, R.W.: Remote sensing and Image Interpretation. Wiley & Sons, NYC (1999)

Martinis, B.: Geologia generale e geomorfologia, in Enciclopedia monografica del Friuli-Venezia Giulia. In: Istituto per l'Enciclopedia del Friuli-Venezia Giulia, Udine, vol. I (1971)

Mautone, M., Ronza, M.: Versanti terrazzati tra tufi e calcari: valenze ambientali, destinazioni agronomiche, scale di osservazione. Casi di studio in Campania. Geotema, 29, Anno X, Patron Editore, Bologna (2006)

Mrak, I., Repe, B.: Vine and vine growing in the area of Kras (Slovenia). Geoadria 9(2) Zadar (2004)

Morris, J.: Trieste o del nessun luogo, Il Saggiatore Milano (2003)

Panjek, A.: Il paesaggio agrario del Carso, Caratteri, elementi ed evoluzione in età moderna e contemporanea. Quaderni Ezio Vanoni, 1-2 Trieste (2003)

Poldini, L.: La vegetazione del Carso Isontino e Triestino. LINT, Trieste (1989)

Regione Autonoma Friuli Venezia Giulia (RAFVG): L'olivicoltura nelle Province di Trieste e Gorizia. Direzione centrale delle risorse agricole, naturali, forestali e montagna, Trieste (2007), http://www.regione.fvg.it

Scaramellini, G.: Il paesaggio agrario e il paesaggio culturale dei terrazzamenti artificiali nelle Alpi. In: Il paesaggio terrazzato, Un patrimonio geografico, antropologico, architettonico, agrario, ambientale (2005)

Scaramellini, G.: Paesaggi terrazzati e ricerca geografica - Un progetto di indagine sistematica. Geotema, 29, Anno X, Patron Editore Bologna (2006)

Semerani, L., Celli, L., De Rosa, D.: Il Carso triestino: Santa Croce. Libreria Italo Svevo Trieste (1970)

Terranova, R.: Il paesaggio costiero terrazzato delle Cinque Terre in Liguria. Studi e Ricerche di Geografia, XII (1) (1989)

Trischitta, D.: Il paesaggio dei terrazzamenti: tra cultura e natura. In: Il paesaggio terrazzato, Un patrimonio geografico, antropologico, architettonico, agrario, ambientale. Atti del seminario di Studi Taormina (2005)

Valussi, G.: Il confine nordorientale d'Italia, Nuova edizione. In: Nodari, P. (ed.) ISIG Gorizia (2000)

Varotto, M.: Le "masiere" del Canale di Brenta: origini, crisi e rilancio di un paesaggio culturale. Geotema, 29, Anno X, Patron Editore Bologna (2006)

Varotto, M.: Towards the rediscovery of the "middle landscapes". Terraced landscapes of the Alps. Atlas. In: Scaramellini, G., Varotto, M. (eds.) ALPTER Project, Marsilio, Venezia (2008), http://www.alpter.net

Estimation of Population Density of Census Sectors Using Remote Sensing Data and Spatial Regression

Tessio Novack, Hermann Kux, and Corina Freitas

National Institute for Space Research (INPE) – Brazil, Department of Remote Sensing
e-mail: {tessio,Hermann}@dsr.inpe.br, corina@dpi.inpe.br

Abstract. Assuming that urban planning aims the optimization of urban functioning and the well-being of citizens, questions like "how many people are living in the city?" and "where do they live?" become key issues. In this work we utilized landscape metrics generated by the FragStats software for the estimation of population density out of census sectors in the mega city of São Paulo, Brazil. The metrics were calculated over an image from the *QuickBird* II sensor classified by the Maximum Likelihood algorithm. The accuracy of the classified image was analyzed qualitatively. Ordinary linear regression models were generated and formal statistical tests applied. The residuals from each model had its spatial dependency analyzed by visualizing its LISA Maps and by the Global Moran index. Afterwards, spatial regression models were tried and a significant improvement was obtained in terms of spatial dependency reduction and increase of the prediction power of the models. For the sake of comparison, the use of dummy variables was also tried and it became a suitable option for eliminating spatial dependency of the residuals as well. The results proved that some landscape metrics obtained over high resolution images, classified by simple supervised methods, can predict well the population density at the area under study when using it as independent variable in spatial regression models.

Keywords: Population density, Spatial regression, High resolution remote sensing.

1 Introduction

The 21st century up to now can be described from a historical perspective as the time of the establishment from the urban civilization. Presently half of the world population lives in urban areas. The new global economics has attracted man people from rural to urban areas, causing a strong migration flux. The urban expansion is a process that started decades ago but increased at this new phase of global capitalism. The United Nations estimated that in 2030, the percentage of persons living and working in cities can reach up to 60 to 70% (UN 2006). A considerable

B. Murgante et al. (Eds.): Geocomp., Sustainability & Environ. Planning, SCI 348, pp. 111–122.
springerlink.com

part of that proportion is and probably will be concentrated in big cities in developing countries where urban violence, air pollution, traffic problems and a lack of planning are today strong obstacles to the improvement of life quality.

Assuming that urban planning aims the optimization of urban functioning and the well-being of citizens, questions like "how many people are living in the city?" and "where do they live?" become key issues. Population projections for inner-city domains are more and more required to allow the evaluation and monitoring of social programs, constituting the denominator of several social indicators established periodically (Jannuzzi 2004). Nevertheless the population census in Brazil is decennial and presently there is no flexible and low cost methodology available to estimate the number and the density of inhabitants during an inter-census time frame. This problem is worsened by the fact that large cities in developing countries present an explosive growth dynamic, with constant changes such as urban sprawl, increase on the number of residential and commercial vertical dwellings throughout the city, increase of population living in slums, etc. The traditional approaches to estimate total population or population density are mainly based on field surveys. These methods are labor-intensive, time-consuming, costly, and its database cannot be easily updated. In this context high resolution remote sensing data (pixel size < 2.0m) and spatial regression techniques are powerful tools, indicating simple solutions relatively easy to implement and to overcome the problems mentioned. The main advantages of using orbital remote sensing data for population density estimation are the free availability of some of the robust software for processing these data, systematic generation of data and the increasing availability of high resolution imagery from several commercial satellites which will reduce its still high price in the near future.

The objective of this paper is to describe and comment the methodological steps for constructing a spatial regression model for census sectors of population density estimation, using as independent variables landscape metrics obtained from a *QuickBird* II classified image. The document is organized as follows: in session 1.1 we discuss briefly the relevant recent studies using remote sensing data for population density estimation and emphasize the importance of considering the spatial dependency at the estimations. In session 1.2 we formally present the objectives of our contribution; session 2 describes the test area and its context; session 3 informs on the data and software used; session 4, the heart of this article, details the methodology and gives some theoretical insights. Finally at session 6 conclusions are made and suggestions given for future works.

1.1 The Use of Remote Sensing Data for Total Population and Population Density Estimation

Many methods for population estimation have been reported in the GIS and remote sensing literature. Sutton et al. (2003) and Balk et al. (2005) used nighttime images as one of the inputs to regional and global population density grids. Satellite images with medium spatial resolution (15 – 30 m) and linear regression techniques were already used for the construction of population estimation models

using as explanatory variables (e.g. independent variables) the average spectral radiance (or reflectance) values associated to image pixels at several bands of a sensor (Harvey 2002; Reis 2005).

Lo et al. (2006) used an impervious surface fraction image obtained by spectral un-mixing techniques, GIS data and regression techniques to estimate total population with ETM+/Landsat-7 data. Li and Weng (2007) integrated remote sensing and census data for assessing urban quality of life. Presently the increasing availability of high spatial resolution imagery represents a new paradigm for the total population or population density estimation of an area, especially in urban areas. Liu and Clarke (2002) explored the relation between texture of an IKONOS image measured by spatial metrics and semi-variograms of population density in census sectors using linear regression methods. Liu and Herold (2007) used high resolution imagery, linear regression techniques and geographically weighted regression to estimate the population density of Santa Barbara city (USA) using as independent variables spatial metrics calculated by the FragStats software package. In spite of the remarkable contribution of these works, there have been few discussions on the spatial dependency of linear regression residuals. This study is innovative in the sense that it intends to consider the spatial dependency structure of the regression model residuals as a way to increase the efficiency of the coefficients confidence interval. This means that the Mean Square Error (MSE) of the ordinary linear regression as well as its coefficients variance is under-estimated when there is a dependence of the residuals. Furthermore, confidence intervals and tests using the t and F distributions are no longer strictly applicable. Hence, considering the dependence of the residual terms through spatial regression, it will make the coefficients of the regression and its confidence intervals less optimistic but more realistic (Netter and Wasserman 1974).

1.2 Objectives

The objective of this paper is to construct and compare potential models for the estimation of population density at 212 census sectors from São Paulo city (Brazil) using as independent (explanatory) variables landscape metrics obtained by the FragStats 3.3 software and calculated on a classified image from the QuickBird II sensor. The prediction power of the proposed models was compared on an index basis before and after considering the spatial dependency of the residuals. After all considerations a final model was chosen as the most applicable.

2 Test Area

The estimated total population for the municipality of São Paulo (Brazil) in 2007 was 10.886.513 inhabitants, living in an area of 1.523 km^2 (IBGE 2000). Considering that the majority of these people live in the urban area, which corresponds to only a fraction of the municipality, we get an idea on the size of the social and economic phenomenon of São Paulo city. The district of the city that corresponds to our test area has a size of 49 km^2 and is located at the SW of it (Figure 1). The

site's central geographical coordinates are: W 46° 43' 30''and S 23° 36' 30''. This specific region was chosen as a test area because it is very heterogeneous in terms of land use and land cover classes. Different types and sizes of roofs and vegetation covered areas as well as parking lots, swimming pools, streets and avenues and bare soil areas occur at this site. These urban entities configure industrial and residential occupation of different structures, densities and social-economic levels. This area also contains the second largest slum of São Paulo (known as Paraisópolis) as well as vertical condominiums of residential use and entire blocks occupied by houses of very high standard.

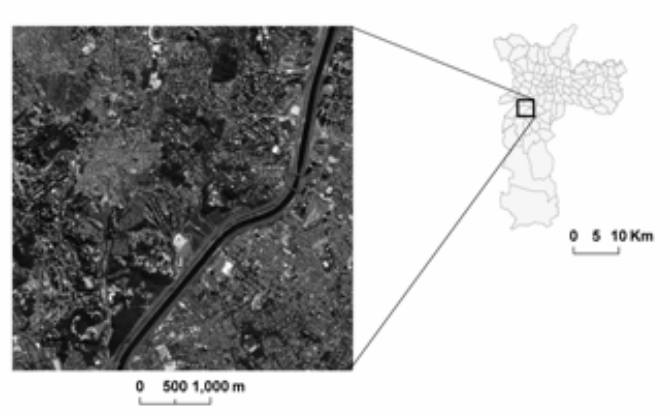


Fig. 1. *QuickBird* II image of the test area and the municipality of São Paulo city on the right

3 Materials

The following materials were used: 212 census sectors in vector format with socio-economic and demographic data from the Demographic Census of IBGE for the year 2000; a *QuickBird* II image from 2002 (Table 1); and the following software packages: (1) SPRING 4.3.3 for the automatic classification of the *QuickBird* II image; (2) FragStats 3.3 for the calculation of metrics of both the census sector as a whole and all the land cover classes in it; (3) Statistica 6.0 for the ordinary least square regression model construction and formal diagnostics and fitness tests; (4) GeoDA 0.9.5.1 (Beta version) for the spatial regression model construction and diagnostics tests.

4 Methodology

The methodological steps followed to construct spatial regression models for the estimation of population densities in São Paulo were: (1) land cover classification of the QuickBird II image; (2) generation of independent variables using FragStats 3.3 software package; (3) logarithm transformation of the dependent variable

population density; (4) selection of independent variables; (5) selection of potential ordinary least square models; (6) tests on the quality and fitness of these models; (7) analysis of the spatial dependency of the residuals for each model; (8) spatial regression model decision; (9) running the spatial regression models and (10) considerations on the model constructed and its appropriateness. In the following sub-session each step will be explained and commented.

Table 1. *QuickBird* II spectral and spatial resolutions

Mode	Spectral Resolution (nm)	Spatial Resolution (cm)	Radiometric Resolution
Panchromatic	450 – 900	62 to 71	11 bits
Multi-spectral	450 – 520 520 – 600 630 – 690 760 – 900	244 to 288	

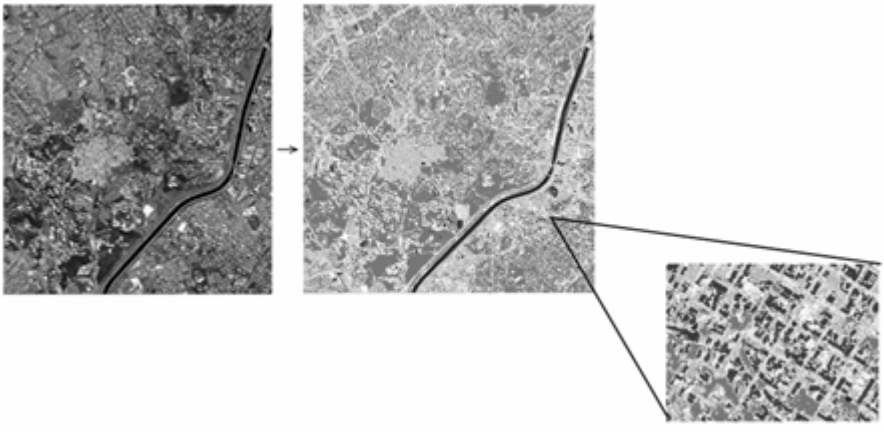


Fig. 2. *QuickBird* II image classified by the Maximum-Likelihood method

4.1 Image Classification

Image classification is a process where image pixels are grouped and associated with concepts based on their digital numbers (and/or object shape and topological relations) for the n bands of a sensor (Mather 2004). The QuickBird image was classified by the Maximum-Likelihood method (Ribeiro and Centeno 2001) considering the following classes: ceramics roofs, dark roofs, vegetation, streets, shadow and non-classified (Fig. 2). This is a simple but efficient pixel-by-pixel supervised classifier. The problem of single and isolated pixels (or a very small

group of pixels) becoming an object from a class different from its surrounding classes was eliminated, using a 3x3 median filter which applies to every pixel the median value of the nine pixel window. This was done on the SPRING 4.3.3 image processing free package developed by the National Institute for Space Research (Brazil). We took care that the classes had low confusion among each other when collecting samples for the supervised classification. The classification accuracy was evaluated qualitatively by visual inspection.

4.2 Generation and Selection of the Independent Variables

All independent variables were generated by the software FragStats 3.3 over the QuickBird II image. FragStats generates spatial metrics for the landscape as a whole (i.e. in this case every census sector of the test area), for each class belonging to each landscape and for every patch in every landscape. We did not use patch metrics because it could not be associated directly to population density of the sectors. At the end, sixty-four independent variables were obtained for both the landscapes (census sectors) and the land cover classes in it. A logarithm transformation on the dependent variable population density was applied to make its distribution closer to the normal one (the purpose of that was to enable eventual confidence interval calculations for the predictions of the model). This operation was successful and it also increased the correlation between the independent variables and the dependent one. All independent variables with a correlation below 0.4 with the dependent variable (i.e. population density) were discarded. Following, we took care that, among others, no pair of independent variables had a correlation above 0.7. This was done to avoid multiple co-linearity problems (Netter and Wasserman 1974). It eliminated most of the sixty-four variables available and made the selection of the model much simpler. Due to the fact that the metrics generated by FragStats were not made specifically for urban applications, all four of the selected variables were quite simple. The variables considered on the census sector were: *number of polygons classified as ceramics roofs*, *percentage of class with dark roof*, *aggregation index of the class streets* (i.e. aggregation index (Eq. 1.) equals the number of similar adjacencies involving the corresponding class, divided by the maximum possible number of similar adjacencies involving the corresponding class, which is obtained when the class is maximally clumped into a single, compact patch) and the *patch density of the class vegetation* (i.e. the number of patches of a certain class divided by the total sector area).

$$\text{Aggregation Index} = \left[\frac{g_{ii}}{\max. \rightarrow g_{ii}} \right] \quad (1)$$

where,

g_{ii} = number of similar adjacencies (joins) between pixels of patch type (class) *i* based on the single-count method.

$\max. \rightarrow g_{ii}$ = maximum number of similar adjacencies (joins) between pixels of patch type (class) *i* based on the single-count method.

4.3 Selection of the Classic Regression Model

We evaluated that some of the variables could explain well the population density, so to the quality tests step we took only the models that contained those variables where all variable coefficients were statistically significant (i.e. none of the coefficients equals zero).

Two models were finally selected. We called them M1 and M2. Table 2 shows the variables at both models with its respective R^2. The formal quality tests for the models showed that the normality of the residuals was acceptable for both models (e.g. p-value below 0.05 for the Kolmogorov test) but non-constancy of variance was a problem for model M1 and M2 as shown by the Breusch-Pagan test (e.g. p-value below significance level of 0.05) on tables 3 and 4.

The area under study is very heterogeneous in terms of size of census sector and land cover classes. Despite of that we did not carry out outlier analysis which would certainly increase the R^2, but eliminating outlying cases in the other hand would not consider completely the complexity of the reality in the test area.

Table 2. Models chosen for the classic linear regression

Model	Variables	R^2
M1	Number of polygons from class ceramics roofs; Percentage of class dark roof; Aggregation Index of class streets; Patch Density of class vegetation.	0.75
M2	Percentage of class dark roof; Number of polygons from class ceramic roofs.	0.65

4.4 Spatial Analysis of the Residuals

For the analysis of the spatial dependency of residuals, a distance-based neighborhood matrix was created. The distance threshold was optimally defined when generating and fitting the semi-variogram. The range detected in this operation was of 1.700 m. For spatial dependency detection the Global Moran index, calculated by the GeoDA software, was used. A z-value and an associated p-value is reported in the diagnostics output. The *z*-value is based on a normal approximation and takes into account the fact that these are residuals. See Anselin and Bera (1998) for details on this issue. The values were low but statistically significant (*p*-values close to zero) for both models (0,0427 for M1 and 0,0825 for M2). The statistical significance of the Global Moran index would already justify the use of a spatial regression model. Nevertheless, the strongest evidence on the existence of spatial dependency of the residuals for both models was the visual inspection of the LISA maps. LISA maps show which cases have their residual difference from the mean

residual of the valid neighbor statistically significant and it showed very clearly spatial groupings of the residuals at both models (Figure 3). The next step was then to test which spatial regression model would be the most suitable for each model.

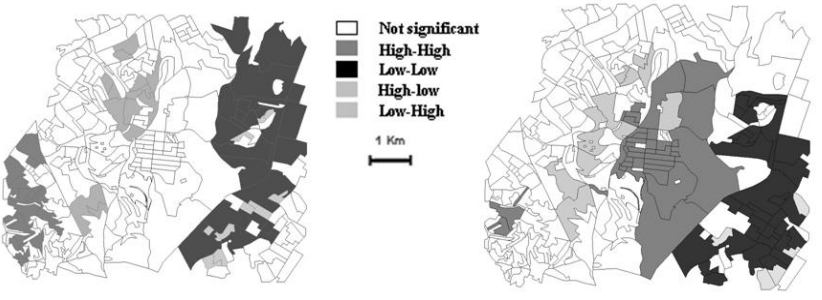


Fig. 3. LISA maps for models M1 (left) and M2 (right) for the ordinary least square regression

4.5 Spatial Regression Model Construction

While Moran's *I* statistic has a great power to detect other misspecifications in the model such as heterogeneity or non-normality (and not only spatial autocorrelation), it is less helpful in suggesting which type of spatial regression model is best suited. To this end, we used the Lagrange Multiplier test statistics. The Lagrange Multiplier (LM) and the Robust Lagrange Multiplier (RLM) tests indicated whether the Spatial Error or the Spatial Lag spatial regression model should be applied for M1 and M2. These statistical tests are distributed as X^2 distribution with one degree of freedom (Anselin 1996 2004; Anselin and Florax 1995). After the diagnostics we made the decision to apply the Spatial Lag model (2) on M1 and the Spatial Error model (3) on M2 (again, lower p-value).

The Spatial Lag model is given by:

$$y = \rho Wy + X\beta + \varepsilon \tag{2}$$

where y is a vector of observations on the dependent variable, Wy is a spatially lagged dependent variable for weights matrix W, X is a matrix of observations on the explanatory variables, ε is a vector of error terms, and β and ρ are parameters.

The Spatial Error model is given by:

$$y = X\beta + \lambda W\varepsilon + \mu \tag{3}$$

where y is a vector of observations on the dependent variable, W the spatial weight matrix, X a matrix of observations on the explanatory variables, ε a vector of spatially auto-correlated error terms, μ a vector of errors, and λ and β are parameters.

The next step was to run the spatial regression models and to evaluate its results.

5 Results

The diagnostics for these spatial regression models indicated that the non-constancy of variance was solved for model M1 (Breusch-Pagan test) and the reduction of the Akaike index attested the improvement of the models, which justifies the inclusion of the new variable: the spatial auto-correlation parameter in both cases (Table 3).

The Global Moran index, although still statistically significant, reduced at both models (-0,017 for M1 and 0,0075 for M2) and the LISA maps showed that the spatial dependency at model M1 was almost eliminated, while it was clearly reduced at M2 (figure 4).

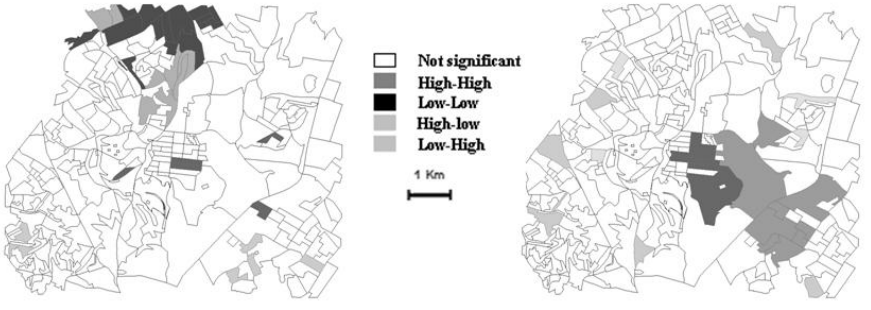


Fig. 4. LISA maps for the two spatial regression models – M1 on the left and M2 on the right

Fig. 5. Collection of dummy variables over the simple regression LISA Map of M2

As the model M2 still presented some spatial dependency and non-constancy of variance problems, we decided to try the inclusion of dummy variables and compare them with the spatial regression model in terms of spatial dependency elimination. Three dummy variables were generated. This operation gives the model three different offset coefficients depending on the spatial domain the cases are located (defined by the dummy variables). The cases associated with values 1 and 0

at the dummy variables were selected over the classical regression LISA map (Fig. 5). The ordinary linear regression model was again applied but this time with local regimes.

We could as well have multiplied the Beta coefficients on all possible combinations by two for the improvement of the prediction power of the model with the dummy variables. This would imply in not only different offsets but also on different gains for each spatial domain (Netter and Wassermann 1974). But since the objective was to come up with a model as simple as possible (for operational purposes and to export the model to other areas) this was not done. The analysis of the prediction power of the ordinary least square regression model with dummy variables had the purpose of evaluating the need for automatic selection of dummy variables (yet to be implemented) for the case study in São Paulo in areas with complex spatial dependency.

An increase of the prediction power on model M1 was achieved by the insertion of the Spatial Lag parameter. The non-constancy of variance which was a problem for M1 was solved when using the Spatial Lag regression as it is shown on Table 3 with the calculated Breusch-Pagan values (e. g. p-value higher than significance level 0.05). At model M2 the Spatial Error Model has slightly improved the R^2 and decreased the Akaike index. On the other hand, the calculated Breusch-Pagan values show that non-constancy of variance is still a problem even on the model with dummy variables. The insertion of these three new variables has enhanced the prediction power as it is shown by the R^2 and the Akaike index on Table 4.

Table 3. Results obtained by the classical regression and by the Spatial Lag regression for model M1

Regression Type (M1)	R^2	Akaike	Global Moran	Breusch-Pagan (p-value)
Ordinary Least Square	0.74	366	0.042	0.0192
Spatial Lag Model	0.75	359	-0.017	0.0781

Table 4. Results obtained for classical regression, Spatial Error regression and for classical regression with local regimes (dummy variables) for model M2

Regression Type (M1)	R^2	Akaike	Global Moran	Breusch-Pagan (p-value)
Ordinary Least Square	0.65	423	0.08	0.00085
Spatial Lag Model	0.68	411	0.007	0.0033
Ordinary Least Square – Local Regimes	0.77	339	0.01	0.0035

Taking into account all these considerations, our choice as the most appropriate model for the test area's population density estimation was model M1 with the Spatial Lag parameter. Not only has the inclusion of the autocorrelation parameters solved non-constancy of variance problems but also discarded implications of residual dependence existence cited in session 1.1. M1 with Spatial Lag parameters has also satisfactory R^2 and Akaike index. The final model has the following expression:

Ln (Population Density) = 0,304 *W* Ln (Population Density) + 0,0857 + 0,0016 Number of polygons from class ceramics roofs + 0,000057 Patch Density of class vegetation + 0,0201 Percentage of class dark roof + 0,038 Aggregation Index of class streets

where *W* is the standardized neighborhood weight matrix.

6 Conclusions

The analysis of the ordinary least square model residuals has shown that although numerically subtle, the spatial dependency of the residuals is clearly observed at the LISA maps on the two proposed models. The spatial regression approach was numerically justified and compensated as the R^2 and Akaike index values show. The use of dummy variables has also proved to have a potential for population density estimation at the test area. When choosing a regression model to be used periodically for population density estimation by the urban planning organizations, one must take into account limitations and advantages of each of the potential models. For this reason, the aim here was to create a model simple but accurate enough which could be applied on other areas and validated in other conditions in the future.

It was proven that population density can be relatively well estimated by the use of spatial metrics calculated over a classified high resolution image. More accurate results obtained in the classification procedure could increase considerably the prediction power of the models proposed. An object-oriented classification approach will be tried in future works and information on block land use can also be considered as spatial domain or independent variables in future occasions. We suggest also that the variables used in our model are applied on a larger area which will test the predictive capacity of these variables are and how solid the spatial regression model is to reduce the dependency of residuals and resulting in more efficient and realistic coefficients.

References

Anselin, L.: The Moran scatterplot as an ESDA tool to assess local instability in spatial association. In: Fischer, M., Scholten, H., Unwin, D. (eds.) Spatial analytical perspectives on GIS in environmental and socio-economic sciences. Taylor and Francis, London (1996)

Anselin, L.: GeoDa 0.95i Release Notes. Spatial Analysis Laboratory (SAL). Department of Agricultural and Consumer Economics, University of Illinois, Urbana-Champaign (2004)

Anselin, L., Bera, A.: Spatial dependence in linear regression models with an introduction to spatial econometrics. In: Ullah, A., Giles, D.E. (eds.) Handbook of applied economic statistics. Marcel Dekker, New York (1998)

Anselin, L., Florax, R.J.: Small sample properties of tests for spatial dependence in regression models: some further results. In: Anselin, L., Florax, R.J. (eds.) New directions in spatial econometrics. Springer, Berlin (1995)

Bailey, T.C., Gatrell, A.C.: Interactive Spatial Data Analysis. John Wiley and Sons, New York (1995)

Balk, D., Pozzi, F., et al.: The distribution of people and the dimension of place: methodologies to improve the global estimation of urban extent. In: Proceedings of the Urban Remote Sensing Conference, Tempe, AZ (2005)

Harvey, J.T.: Estimating census district populations from satellite imagery: some approaches and limitations. Int. J. Remote Sens. 23(10), 2071–2095 (2002)

IBGE – Fundação Instituto Brasileiro de Geografia e Estatística. In: Censo demográfico de 2000 (2002), http://www.ibge.br (access June 23, 2008)

Jannuzzi, P. M.: Projeções populacionais para pequenas áreas: métodos e aplicações (2004), http://www.abep.nepo.unicamp.br/docs/anais/pdf/2000/Todos/prot20_1.pdf (access August 18, 2008)

Liu, X.H., Clarke, K.: Estimation of residential population using high resolution satellite imagery. In: Proceedings of International Symposium on Remote Sensing of Urban Areas, Istanbul, Turkey (2002)

Li, G., Weng, Q.: Integration of remote sensing and census data for assessing urban quality of life: model development and validation. In: Weng, Q., Quattrochi, D. (eds.) Urban remote sensing. Taylor and Francis, Boca Raton (2007)

Liu, X., Herold, M.: Population estimation and interpolation using remote sensing. In: Weng, Q., Quattrochi, D. (eds.) Urban remote sensing, Taylor and Francis, Boca Raton (2007)

Lo, D., Weng, Q., et al.: Residential population estimation using a remote sensing derived impervious surface approach. Int. J. Remote Sens. 27(16), 3553–3570 (2006)

Mather, P.M.: Computer processing of remotely-sensed images: an introduction, 3rd edn. John Wiley & Sons, Chichester (2004)

Netter, J., Wasserman, W.: Applied linear statistical models. Irwin, Homewood (1974)

Reis, I. A.: Estimação da população dos setores censitários de Belo Horizonte usando imagens de satélite (2005), http://marte.dpi.inpe.br/col/ltid.inpe.br/sbsr/2004/11.18.18.39/doc/2741.pdf (access December 18, 2007)

Ribeiro, S. R. A,. Centeno, J .S.: Classificação do uso do solo utilizando redes neurais e o algoritmo MAXVER (2001), http://urlib.net/dpi.inpe.br/lise/2001/09.20.17.56 (access December 18, 2007)

Sutton, P.C., Elvidge, C.D., et al.: Building and evaluating models to estimate ambient population density. Photogramm Eng. Remote Sens. 69(5), 545–553 (2003)

UN-HABITAT (2006), http://www.unchs.org/programmes/guo/statistics.asp (access February 28, 2008)

Using Environmental Geostatistics for the Geochemical Characterization of Soils from the Polluted Site of National Interest of Tito (PZ – Italy)

Pietro Lucia[1], Achille Palma[1], Beniamino Murgante[2], Carmela Miriam D'Alessandro[3], Adriano Sofo[3], and Antonio Scopa[3]

[1] Metapontum Agrobios s.r.l., S.S. Jonica 106, Km. 448.2 - 75010 Metaponto di Bernalda, Matera, Italy
e-mail: {plucia,apalma}@agrobios.it

[2] Department of Architecture, Urban and Regional Planning, Transport Planning and Infrastructure, University of Basilicata, Via dell'Ateneo Lucano, 10 - 85100 Potenza, Italy
e-mail: beniamino.murgante@unibas.it

[3] Department of Sciences Crop system, Forestry and Environmental Sciences, University of Basilicata, Via dell'Ateneo Lucano, 10 - 85100 Potenza, Italy
e-mail: {carmela.dalessandro,adriano.sofo, antonio.scopa}@unibas.it

Abstract. The aim of this work was to provide information concerning the distribution of heavy metals in soils of a polluted ecosystem, in order to predict potential environmental risks and to provide a tool for the decision maker. This paper focuses on characterizing the industrial area of Tito (PZ, southern Italy), a site included among national interest sites to decontaminate, according to D.M. 8/7/2002. Soil contamination was monitored by means of a chemical-physical evaluation, coupled with a modelling approach using geostatistic techniques. A multistep sequential acid extraction technique was used to determine partitioning and levels of heavy metals in soil samples. Results showed that concentrations of analyzed elements are high in the whole area and above legislative admissible limits. A high spatial variation of heavy metals was observed in the studied area, with higher levels of heavy metals beside active and abandoned industrial areas. The adopted approach highlighted that anthropogenic industrial pressure may have detrimental repercussions on the surrounding environment and that recovering contaminated areas by implementing decontamination or permanent making safe interventions becomes necessary.

Keywords: Heavy metals, soil contamination, industrial area, geostatistics, kriging.

1 Introduction

The increase in environmental metal concentrations is primarily due to erosion phenomena and anthropogenic activities; since metals are very persistent

B. Murgante et al. (Eds.): Geocomputation, Sustainability & Environ. Planning, SCI 348, pp. 123–144.
springerlink.com

pollutants they accumulate in the soil. One of the most closely monitored areas is the soil ecosystem, where great attention is paid to hazardous elements. Knowledge of the processes driving pollutants migration and availability in the soil is important to evaluate their effect on environment and to forecast their environmental impact. In particular, soil represents one of the matrices receiving many polluting inputs, such as heavy metals. Moreover, soil is a distribution centre towards other environmental compartments due to pollutant flows in different matrices, which may compromise the equilibrium of many ecosystems.

Soil pollution by heavy metals arises strong concerns among developed countries, since it is mainly related to anthropogenic activities performed in industrial zones and in areas with high population density, representing the most important sources of heavy metals.

Decontamination of polluted sites represents one of the most relevant issues in debating environmental themes which particularly interest communities and decision makers. In fact, art. 17 of D. Lgs. 22/97 and D. Lgs. 152/06 regulate decontamination of polluted sites, introducing decisional participation of local agencies, to which most important tasks of human health and environmental safety are demanded. Moreover, decision makers must take into account EU guidelines concerning the analysis of effective or potential risk or different expectations for reusing and controlling abandoned and/or still used areas.

The need of providing a more and more detailed territorial knowledge has led to the development of techniques, which are reaching a broad diffusion thanks to their integration in GIS software (Bailey 1994; Burrough 2001), able to achieve geographic visualization of experimental data.

Spatialization of geochemical data has a great practical relevance in territorial management and may be absolutely necessary to evaluate background levels of polluting species in different investigated matrices. A correct determination of background values is of particular importance in environmental surveys, since law in force fixes intervention limits concerning metal elements noxious for human health (Cicchella *et al.* 2004). If these limits are exceeded, recovering of contaminated areas by means of decontamination or permanent making safe interventions becomes absolutely necessary. Starting from punctually surveyed experimental data, the elaboration of geochemical maps may provide opportune information about numerical models of transport, diffusion and pollutant levels as well as their temporal evolution in analyzed matrices. Defining if a determinate element is present at an anomalous concentration, and therefore it has to be considered a pollutant of that definite matrix, means first knowing its natural background grade, which may significantly vary from zone to zone. The main aim of geochemical survey is the characterization of anomalous areas (*e.g.* due to biotic or abiotic transformations or to anthropogenic pollution). In order to reach this aim, it is necessary to evaluate the values of natural background grades in the study area (Lima *et al*. 2004).

Integration of chemical-physical parameters and geostatistical techniques has been largely applied during the last years, allowing to estimate background values of contaminants in investigated matrices (Liang *et al*. 2003; Aboal *et al*. 2005; Liu *et al*. 2006; El Sebai *et al*. 2006; Rodriguez *et al*. 2008; Wu *et al*. 2008).

In environmental modelling studies, especially for the distribution of xenobiotic molecules in soils, the reconstruction of concentration fields is carried out using simple interpolation procedures. In this way, it is difficult to estimate how measure uncertainty and eventual stochastic characteristics in concentration field due to the small scale complexity of contaminant mobility in soils, affect the reconstructed datum in a point where measures are not available. In order to obtain this kind of results, various geostatistic techniques have been developed and widely applied in the field of mining survey and underground fluxes modeling. These techniques mainly aimed to spatially reproduce a phenomenon above the whole study area, on the basis of a limited number of punctual observations. Particularly, kriging techniques are based on a statistic model of the phenomenon and not on the model of the interpolation function (Chiles *et al*. 1999).

The use of such interpolators, included in the so-called "regionalized variables theory" (Matheron 1965), allows to get an estimate of the characteristics of a stochastic field in the points where it has not been measured (Bocchi *et al*. 2000; Castrignanò *et al*. 2002; Castrignanò *et al*. 2003; Haining 2003; Lloyd *et al.* 2006). In order to perform the geostatistic interpolation, it is necessary to preliminarily carry out the estimate of the so-called variogram, which allows to identify a covariance structure of the stochastic field consistent with the variability of studied data. Once defined a covariance model consistent with the data, it is possible to continue with a reconstruction of the field on a uniform grid (Madeo *et al.* 2005).

The aim of the present study was to investigate the distribution of selected heavy metals for the geochemical characterization of the national interest polluted site of Tito (PZ, southern Italy) by applying geostatistical techniques to heavy metal concentration data in different fractions of sampled soils. The area is of considerable ecological relevance, and a detailed knowledge of heavy metals distribution in soils sampled in this ecosystem it is necessary to predict the potential environmental risks and to provide a tool for the decision maker. The evaluation of heavy metal distribution in different soil fractions has been performed both by *multistep* sequential extraction, in order to evaluate leaching processes, and by total mineralization with acid attack for the geochemical characterization. Starting from the initial dataset, we identified the presence of outliers, in order to separate data representative of the natural background by those representative of populations whose values have been determined, and successively we interpolated them.

2 Materials and Methods

2.1 Study Area

The experiment was carried out during summer 2005, and soils were collected in 70 sampling stations at 0-40 cm depth. UTM (Universal Transverse Mercator) coordinates were determined for each position of sampling stations (their locations are shown in Figure 1) by using GPS map (Garmin, Taiwan). The sampling was carried out in stations located on a regular grid with a 200 m side, within the perimeter of the industrial area pertaining to ASI Consortium in Tito municipality

(Potenza, Italy, 40°36'11" N, 15°42'54" E), located at 763 m a.s.l. and 4.5 km distant from the built-up area. Soils were classified as clay loam (sand 432.2 ± 22.2, silt 245.0 ± 25.7, clay 322.8 ± 27.1, g kg^{-1} ± SD). The study area lays along slopes with precarious stability due to the presence of mainly clay soils, strongly eroded by exogenetic agents. Outcropping formations are clay marl, sandstone, siliceous schist and limestone. Drillings highlighted the presence of slimy soils tending to clay, with very low permeability up to 26 m depth. The water table has an average level of about 8 m, which reaches maximum levels of 2-3 m below the pattering plane in winter.

Fig. 1. Map of the Industrial Area of Tito (PZ, Italy). Locations of sampling stations are shown.

Studies carried in this area are among those provided by the national decontamination and environmental restoration program, drawn up according to art. 1, comma 3 of 426/98 law for industrial sites of national interest. According to such rules, Basilicata Region proposed to the Ministry of Environment to delimitate the national interest site. Partly demolished or evidently abandoned buildings and plants, silos and crumbling tanks, vats containing refluing waters and depuration mud, scattered piles of waste products are present in the area.

2.2 Chemical Analysis

All reagents used were of analytical grade (Carlo Erba, Italy), while extra pure grade acids (Merck, Darmstadt, Germany) were used to dissolve solid samples. Chemicals used for preparation of calibrating standard solutions were at heavy metals grade (Merck, Darmstadt, Germany). All solutions were prepared with ultrapure water at 18MΩ supplied by a Milli-Q RG unit (Millipore, Bedford, MA, USA). All plastic bottles, Teflon vessels and glassware materials were cleaned thoroughly with detergent solution, soaked in HNO_3 20% (v/v), rinsed with deionised water and dried.

The selective chemical *multistep* extraction was in accord with the sequential extraction procedure proposed by Tessier *et al.* (1979). The microwave-digestible residual concentration of heavy metals was determined utilizing a Teflon digestion vessel (1 g) and the residual was digested in 3 ml of HNO_3 70%, 4 ml of HF 40% and 1 ml of $HClO_4$ 70%. Samples were first heated from room temperature to 200 °C for 10 min and kept at this temperature for 15 min in a closed high-pressure microwave system Ethos SEL (Milestone, Italy). The instrument was equipped with a 10-vessel position carousel and a temperature controller. After digestion, vessels were allowed to cool to ambient temperature. Another microwave digestion system ETHOS TC (Milestone, Italy) equipped with VAC-4000 was utilized, and samples were boiled to near dryness in order to evaporate the exceeding acids of HF and $HClO_4$. Operating parameters and power of microwave digestion systems were set according to the procedures reported in the instrumentation manual. The residual solution was subsequently transferred into a 25 ml volumetric flask, and 5 ml of 50% HNO_3 (trace metal grade) were added to preserve the sample for trace metal analyses. Samples were kept at –18 ºC until analyses were carried out.

Determination of Al, As, Be, Cd, Co, Cr, Cu, Hg, Mn, Ni, Pb, Sb, Se, Sn, Tl, V and Zn was achieved with inductively coupled plasma mass spectrometer ICP-MS ELAN 6000 DRCe with a cross flow nebulizer (Perkin Elmer, USA). ^{103}Rh and ^{187}Re at the concentration 10 $\mu g\ l^{-1}$ were used as internal standards. Blank solutions were prepared to correct for contaminants contained in reagents used during sample dissolution. Calibrating standard solutions were used for the calibration of the ICP/MS instrument at different concentrations of heavy metals. The calibration curves were linear in the whole calibrating range ($r \geq 0.9996$). Measures have been replicated thrice and recoveries have been determined using certified soils.

2.3 Statistical and Geostatistical Analyses

Both spatial and numerical statistical analyses have been applied in order to identify and to handle possible outliers and data distribution, and to obtain indications about population behavior. Identification of outliers has been performed according to EPA (2006), while background values of single metals have been calculated according to APAT-ISS (2006). Graphical representation, used to search for outliers, has been performed using box-plot and normal QQ plots. The presence of anomalous data has been verified by means of Grubbs test (maximum normalized residual test). Probability distribution better approximating the whole of available data has been checked for applying Lilliefors (Kolmogorov-Smirnov) test. All statistical analyses have been carried out using opensource "R" software (Ihaka *et al.* 1996; Ribeiro *et al.* 2001a; Iacus *et al.* 2003; R Development Core Team 2006).

Spatial variability, which interpolates semivariance of observed values in pairs of points at definite distances (lag) along a certain direction through the semivariogram function, described the structure of a regional variable of known value in a discrete number on points (Journel *et al.* 1978). We used a statistical interpolation method of kriging, which allowed to measure the average error and the standardized error variance through cross-validation. Compatibility between the set of

experimental data and their structural model has been evaluated by the vicinity of previous statistics, at values of 0 and 1, respectively. All geostatistical analyses have been performed using "geoR", a spatial extension of "R" statistical software (Ribeiro *et al.* 2001b; Reimann *et al.* 2008; Bivand *et al.* 2008).

3 Results and Discussion

Characteristics and utilization of contaminated sites, character of environmental impact and target-oriented protection objectives determine different criteria and methods for site investigation, evaluation and remedial concepts. Knowledge regarding physical, chemical, and biological characteristics of the soil is important for assessing its capacity to act as sink and source of heavy metals. In particular, remobilization processes may reintroduce heavy metals in a natural or terrestrial system in different chemical forms or can be strongly retained in function of soil parameters, such as pH, redox potential, and ionic strength. In the present study, organic matter content has not been examined, according to D.M. 471/99 and D.Lgs. 152/2006. Nevertheless, it is well known that availability of heavy metals in soils is strongly affected by organic matter content and quality. Many studies have highlighted the importance of microbial biomass, an important fraction of soil organic matter, in metal mobilization equilibriums. The determination of heavy metals total concentrations is not sufficient for predicting the potential toxicity or their behaviour in the soil.

Soil contamination with heavy metals is often neglected or underestimated, both for the slowness of progression of the phenomenon and for the difficulty of measuring (Capri *et al.* 2002). Evaluating the level of heavy metals and metalloids in soils is extremely important since they are potentially toxic for life forms, as assessed by organizations such as U.S. EPA (*Environment Protection Agency* of USA) and World Health Organization.

The most used method for determining soil metal concentration is the mineralization in turpentine oil, which is not able to solubilize the silicate component (Gupta *et al.* 1996). Many authors studied fractioning of soil constituents with different reagents and operative conditions. Currently, different chemical extracting methods are used to study the distribution of metals in soils, based on partition techniques using selective chemical solutions able to remove heavy metals, according to their bond strength (Leita *et al.* 2000). Selective sequential extraction techniques provide indications on mobility and availability of different chemical forms of the metals in the soil. The basic principle is that various soil constituents are attacked by different solutions in a selective way, so that the different phases are separately solubilized and it is possible to determine the amount of heavy metal adsorbed on them. The experimental work allowed to consolidate and define operative modalities and the best experimental conditions (Hlavay *et al.* 2004). Geochemical characterization has been performed on fine ground, in order to ensure an effective contact of extracting solutions. Furthermore, granulometry has been determined and cluster analysis performed (data not shown).

ICP-MS allowed to simultaneously determine elements with few spectral interferences. Disadvantages due to overlapping of isotopes or polyatomic species have been eliminated by the use of DRC (*Dynamic Cell Reaction*), located between the ion source and the spectrometer. This consists of a quadrupole enclosed in a reaction chamber.

Determined elements were present in soils with amounts in a wide order of magnitude. Normally, Al and Mn are held in soils in the order of some g kg^{-1}, whilst concentration of most of the other elements is in the order of mg kg^{-1}. Hg, Cd, Tl, Sb and others generally reach concentrations of μg kg^{-1}. Therefore, it has not been possible to determine all elements by operating with a single dilution. Elements in great amounts tend to saturate the spectrometric detector and to compromise its regular functioning; therefore, it has been necessary to operate opportune dilutions. As for trace elements, dilution has been reduced to the minimum in order to maintain a good signal/noise ratio. Isotopes of more abundant elements, or alternatively the less interfered ones (^{27}Al, ^{9}Be, ^{114}Cd, ^{202}Hg, ^{205}Tl, ^{75}As, ^{52}Cr ^{59}Co, ^{63}Cu, ^{60}Ni, ^{51}V, ^{208}Pb, ^{64}Zn, ^{121}Sb, ^{120}Sn, ^{55}Mn, ^{80}Se) have been used. Detectability limits have been calculated on the basis of blank standard deviation, of instrumental variability and of volumes of reagents used to treat samples. Distribution of individual metals is shown in Fig. 5 and 6.

The study of background values and distribution has been carried out on elements with legally fixed limits (except for Al and Mn). Mercury has been analyzed as well, since it was detected in two samples only. Nevertheless, in our samples this element showed concentration values strongly below law limits.

In order to identify anomalous points corresponding to much higher or lower values relative to the others, box plots (Figure 2) and normal Q-Q plots (Figure 3) have been used as data graphical representations. Box plots support table presentation, aiming to simplify analysis and reasoning. They summarize the main aspects of a value distribution: lower and upper bases of the rectangle represent 25^{th} and 75^{th} percentile, respectively; the internal line reproduces the median. Nearby these parameters, the box plot identifies two limits (*upper* e *lower fence*), corresponding to 75^{th} percentile + 1.5 times interquartile difference and to 25^{th} percentile – 1.5 interquartile difference, respectively. Values beyond these limits might be outliers. Normal Q-Q plots can discriminate the presence of gaps or "jumps", as well as slope variations in the obtained curve of a dataset, which may be considered as threshold values for identifying two or more populations. They compare quantiles of a (standardized) dataset to quantiles of a standard normal distribution (average 0 and variance 1). If the distribution is normal, all points would fall along the straight line, while points deviating from it are possible outliers.

Nevertheless, the presence of anomalous data has been confirmed by the Grubbs test (Table 1). Outliers have been excluded from the geostatistical analysis since they were probably due to punctual anthropogenic inputs (analyzed data refer to surface sampling) deriving either from contaminant substances pouring or from landfill unrelated to the study area (e.g. ST72 site proved to be an outlier for Cd, Co, Cu, Se and Zn). Outlier number kept around a value ≤ 3 of 70 analyzed samples.

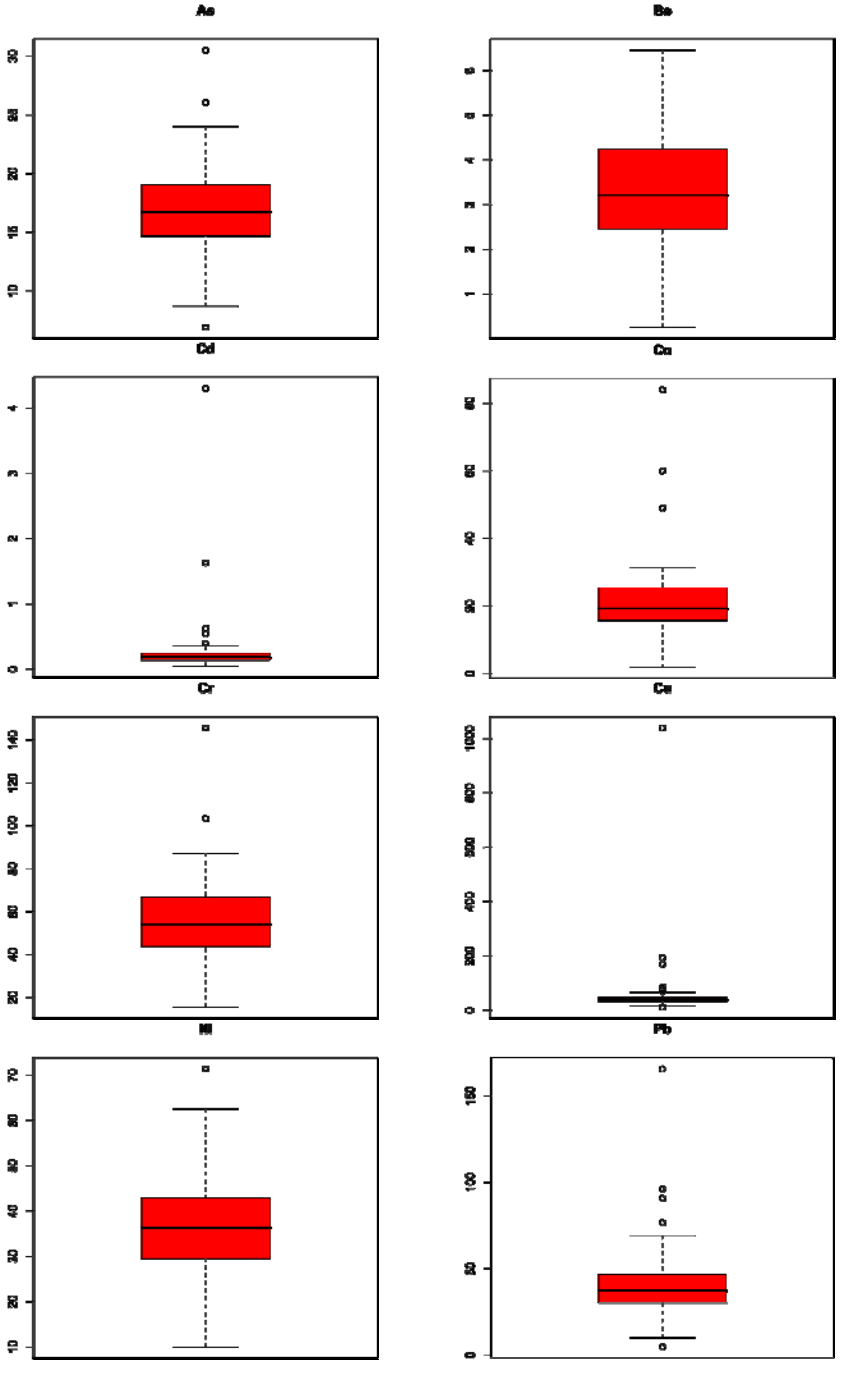


Fig. 2. Box plots

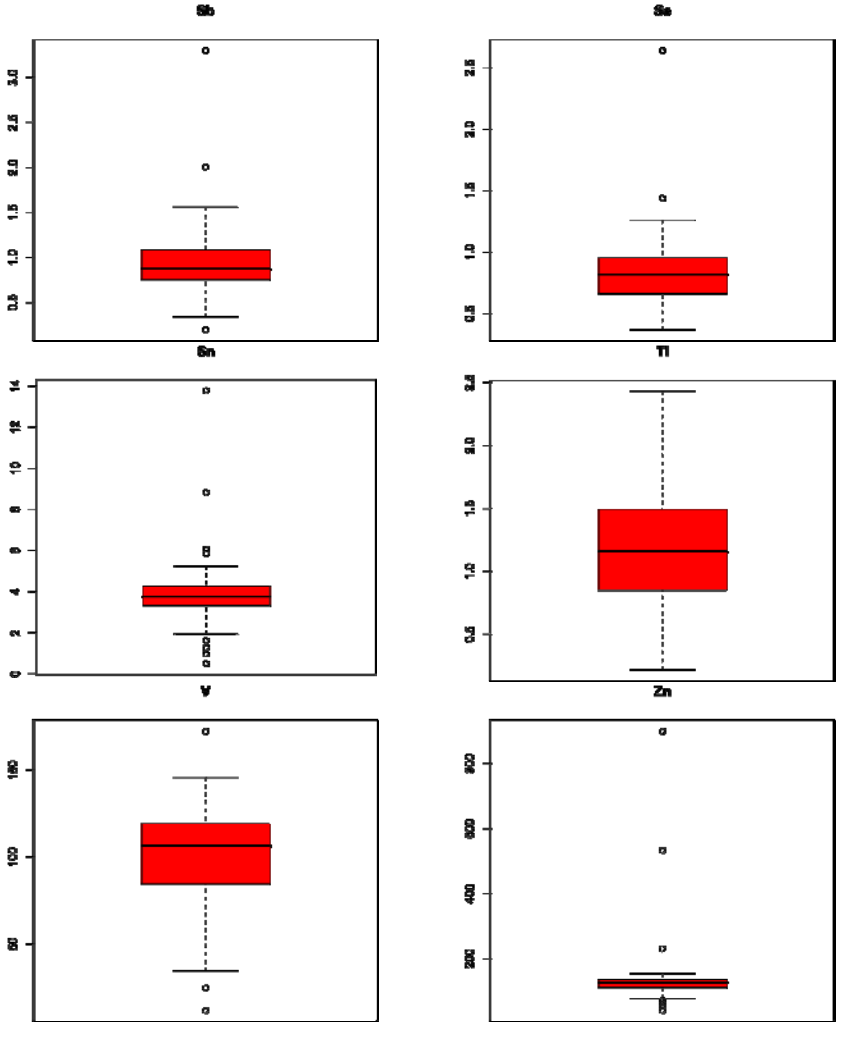


Fig. 2. (*continued*)

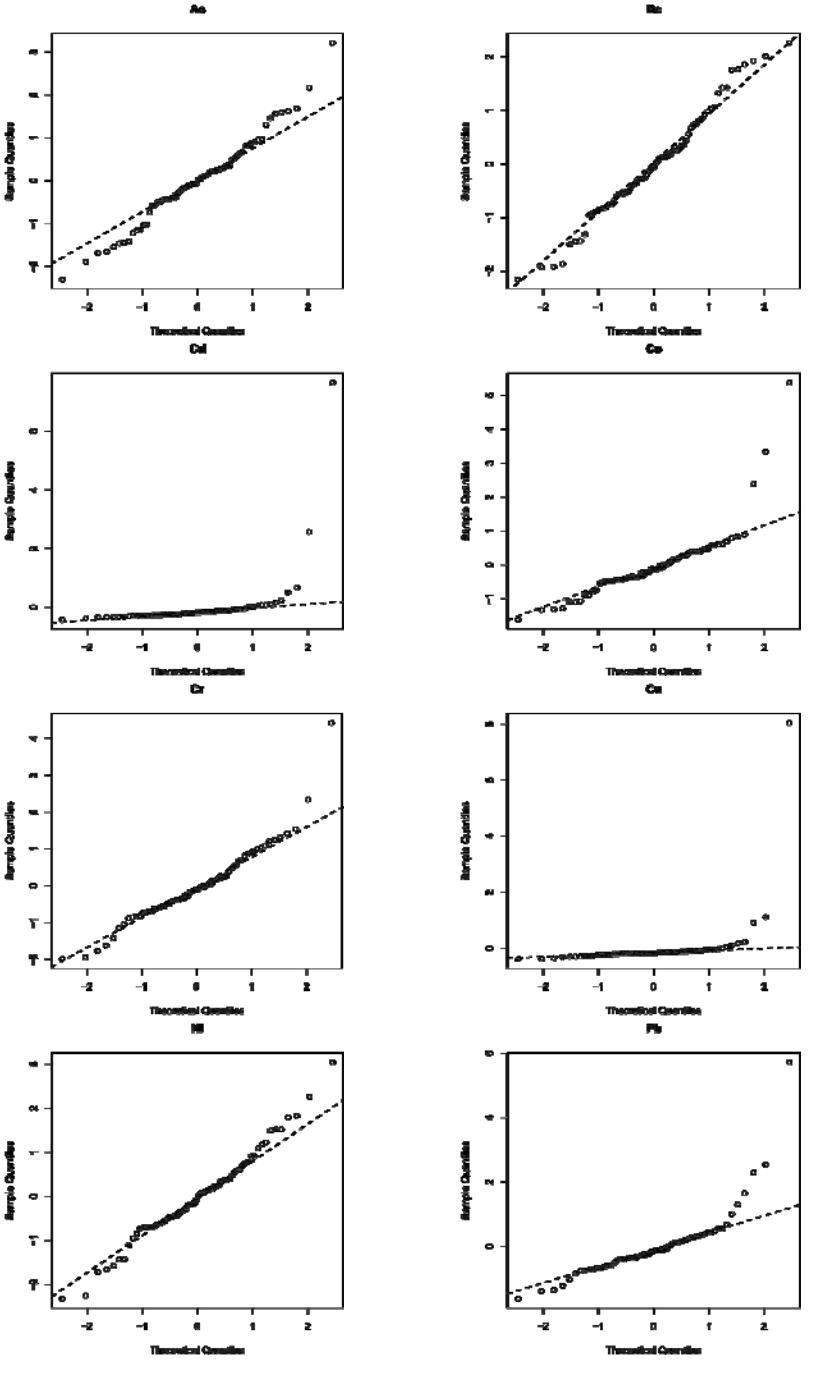


Fig. 3. Normal Q-Q plots

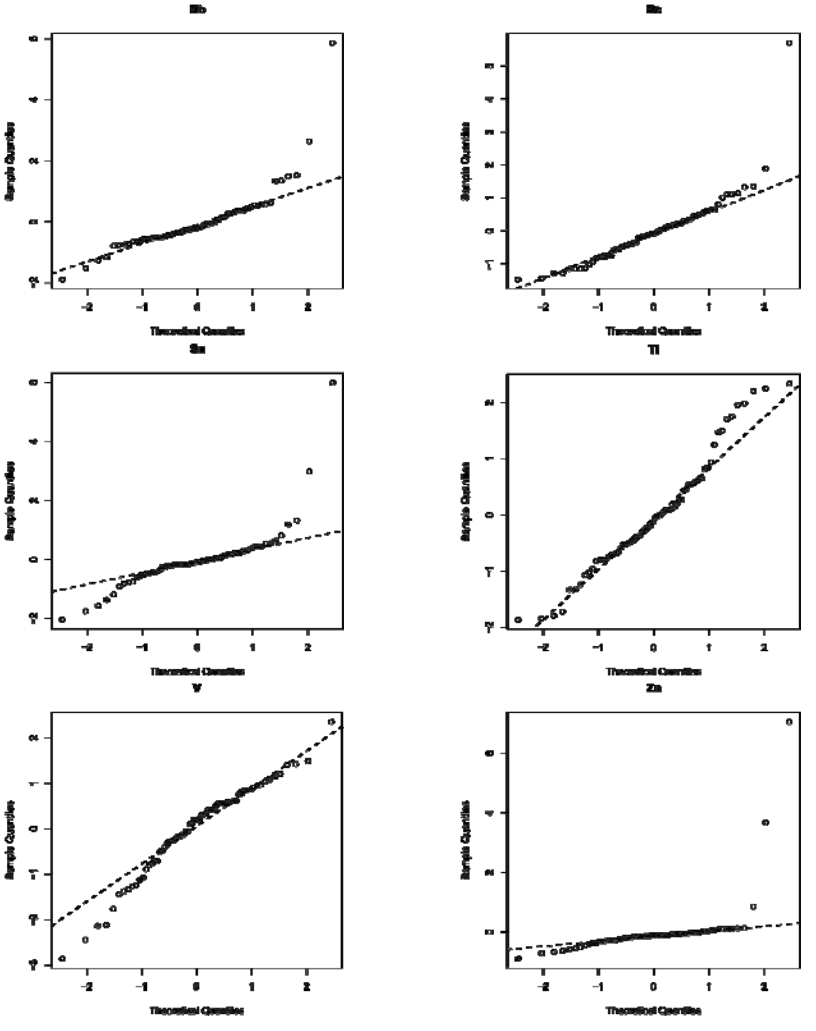


Fig. 3. (*continued*)

Data population without outliers has been considered as a background representation and the 95th percentile has been calculated on the basis of this value as an indication of the background value. As indicated by APAT-ISS document, we suggest that concentrations of samples collected within the enclosed area might be used, once provided that they are external to areas occupied by industrial plants, dumps or, in general, by zones where actual or past contamination sources are located. Nevertheless, our data indicate that, for many metals, values exceeded threshold levels listed in Table A of D.Lgs. 152/2006 and, in some cases, those reported in Table B (Table 2) as well.

Table 1. Grubbs test, for each element G statistic and *p*-value for the most extreme element are shown, the test has been repeated several times till the identification of outlier number

Heavy metals	As	Be	Cd	Co	Cr	Cu	Ni
G	3.2127	2.2555	7.6516	5.3649	4.4217	8.0374	3.0380
p-value	0.03004	0.769	2.2e-16	3.821e-08	6.026e-05	2.2e-16	0.05961
n. outliers	0	0	2	3	2	3	0
Heavy metals	Pb	Sb	Se	Sn	Tl	V	Zn
G	5.7168	5.8591	5.6967	5.9886	2.3337	2.8590	7.0601
p-value	1.024e-09	1.974e-10	1.281e-09	3.96e-11	0.6165	0.1149	2.2e-16
n. outliers	2	3	2	1	0	0	3

Table 2. Provided limits by D.Lgs 152/2006 for public, private and residential green zones (Tab A), for commercial and industrial zones (Tab B) and calculated background value

D.lgs. 152/2006	As	Be	Cd	Co	Cr	Cu	Ni
Tab A	20	2	2	20	150	120	120
Tab B	50	10	15	250	800	600	500
Background	20.52	4.60	0.35	28.69	81.07	54.04	55.72
D.lgs. 152/2006	Pb	Sb	Se	Sn	Tl	V	Zn
Tab A	100	10	3	1	1	90	150
Tab B	1000	30	15	350	10	250	1500
Background	49.93	1.17	1.19	4.61	1.55	140.69	151.39

Therefore, we need to discriminate if calculated values should be considered mostly as a “current background tenor” (baseline), so also including the presence of anthropogenic elements (contamination) in soils, than as a “natural background tenor” (background), which is a natural background value the presence of mineralization (Lima *et al.* 2004).

The population representing the background shows with good approximation a normal distribution for all elements, as showed by the istograms (Figure 4), where the class number has been calculated with Sturges method and confirmed with Lilliefors (Kolmogorov-Smirnov) test (Table 3). Descriptive statistics are reported in Table 4.

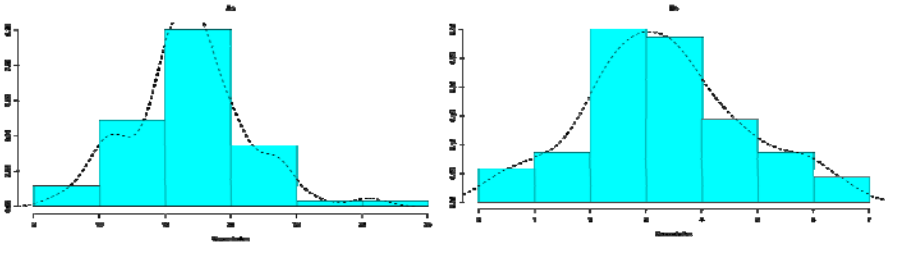


Fig. 4. Istograms

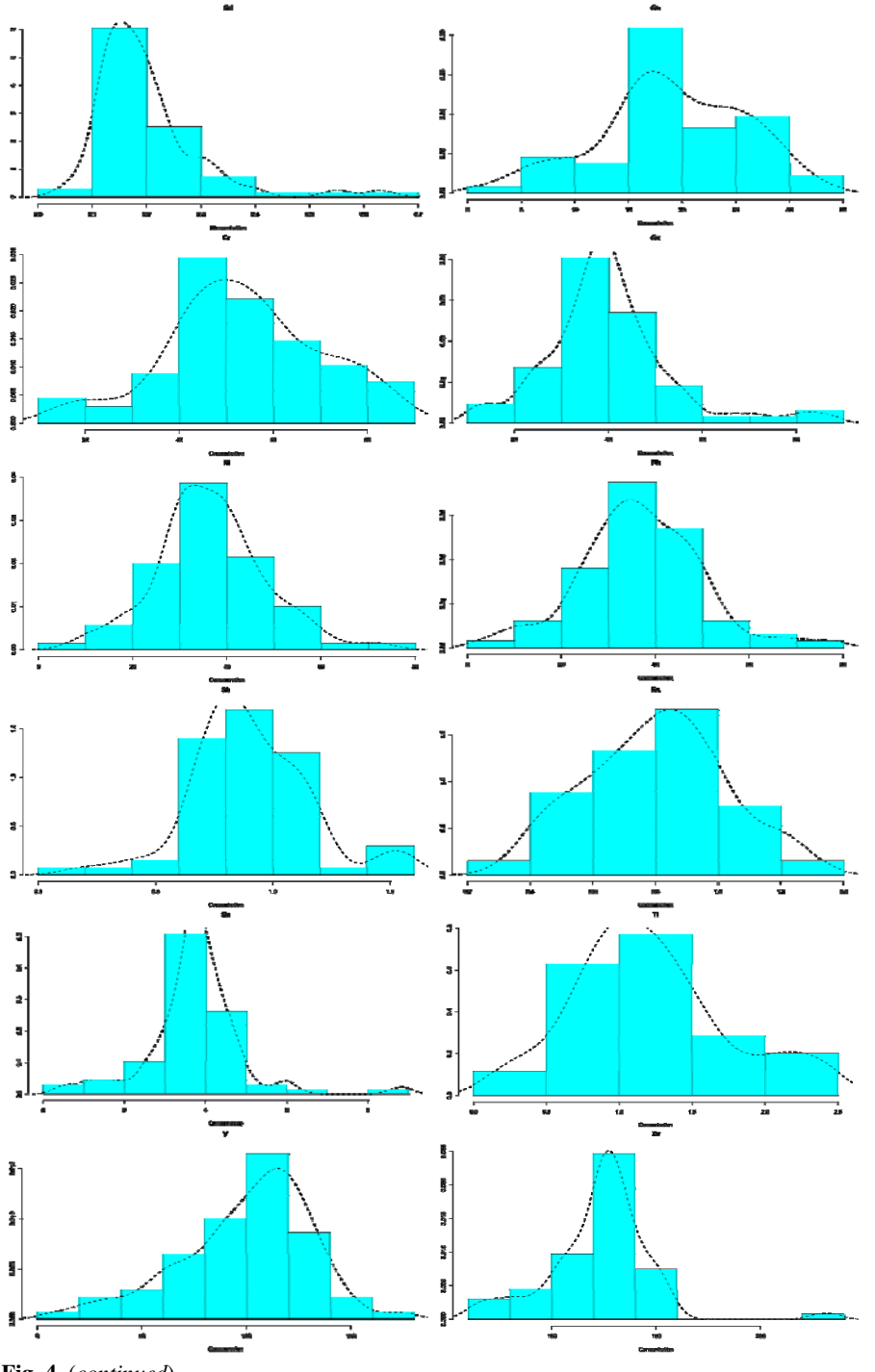


Fig. 4. (*continued*)

Table 3. Lilliefors (Kolmogorov-Smirnov) test's results

Heavy metals	Sample number	D	p-value
As	70	0.09	0.18
Be	70	0.07	0.48
Cd	68	0.08	0.28
Co	67	0.10	0.07
Cr	68	0.06	0.76
Cu	67	0.12	0.02
Ni	70	0.09	0.16
Pb	67	0.07	0.60
Sb	68	0.09	0.18
Se	68	0.05	0.93
Sn	69	0.13	0.01
Tl	70	0.09	0.13
V	70	0.11	0.05
Zn	67	0.13	0.01

Table 4. Statistical descriptors of data population (mg kg^{-1})

	As	Be	Cd	Co	Cr	Cu	Ni
Count	70	70	68	67	68	67	70
Min	6.95	0.26	0.05	1.86	15.47	12.42	9.98
Max	30.50	6.46	0.63	31.34	87.22	85.77	71.42
Mean	16.78	3.29	0.20	18.93	53.82	39.87	36.57
Std.Dev.	4.27	1.41	0.10	6.73	16.34	13.26	11.47
Skewness	0.33	0.16	2.21	-0.34	-0.09	0.98	0.31
Kurtosis	0.84	-0.19	6.89	-0.25	-0.11	2.59	0.76
1-st quartile	14.73	2.46	0.14	15.51	43.57	33.59	29.54
Median	16.69	3.21	0.18	18.87	53.47	39.12	36.25
3-rd Quartile	18.98	4.20	0.22	24.07	64.18	44.71	42.60
	Pb	Sb	Se	Sn	Tl	V	Zn
Count	67	68	68	69	70	70	67
Min	4.74	0.20	0.37	0.50	0.22	11.87	41.70
Max	76.82	1.56	1.26	8.83	2.43	172.44	153.85
Mean	36.82	0.90	0.80	3.73	1.20	100.05	120.21
Std.Dev.	12.92	0.25	0.22	1.15	0.53	30.84	22.85
Skewness	0.27	0.30	0.04	0.73	0.52	-0.67	-1.20
Kurtosis	1.17	1.26	-0.49	5.98	0.04	0.48	1.71
1-st quartile	29.47	0.75	0.66	3.31	0.85	84.61	109.69
Median	36.87	0.87	0.81	3.73	1.16	106.30	125.72
3-rd Quartile	44.59	1.06	0.94	4.19	1.49	118.88	133.41

Spatial variability for all the analyzed elements has been evaluated by interpolating the spherical model with experimental semivariograms (Table 5). Parameters deduced from the semivariogram (nugget, partial sill, range and Q index), concerning the distribution of the analyzed metals, showed a spatial dependence up to a distance between 600 m and 900 m. The Q index for almost all the metals gave evidence for an average spatial structure, so allowing to explain great part of the variance with the adopted model. The Q values were definitely low for Cu, Pb and Sb, indicating the permanence of a significant residual variance at distance zero. This suggests to thicken sampling points around some sites in order to deepen the study of the small scale variability.

Figg. 5 and 6 show the experimental semivariograms with their interpolates model and the maps obtained by kriging. Results of cross validation (Table 6) indicate the compatibility between the experimental dataset and the adopted structural model. In fact, the average of errors and standardized errors was close to zero for all elements, whilst standard deviation of standardized errors was close to 1.

Table 5. Parameters of adepte semivariograms models

	Model	Nugget	Partial Sill	Range	Q
As	spherical	8.33	11.40	890.90	0.58
Be	spherical	0.96	1.11	583.04	0.54
Cd	spherical	0.08	0.17	643.10	0.68
Co	spherical	25.87	21.76	724.54	0.46
Cr	spherical	131.64	156.72	737.35	0.54
Cu	spherical	132.73	19.43	837.61	0.13
Ni	spherical	79.29	53.32	993.09	0.40
Pb	spherical	102.66	56.05	856.65	0.35
Sb	spherical	0.05	0.01	900.00	0.19
Se	spherical	0.02	0.04	900.00	0.66
Sn	spherical	0.75	0.51	800.00	0.40
Tl	spherical	0.13	0.15	583.20	0.54
V	spherical	338.65	653.49	626.34	0.66
Zn	spherical	210.32	279.41	942.70	0.57

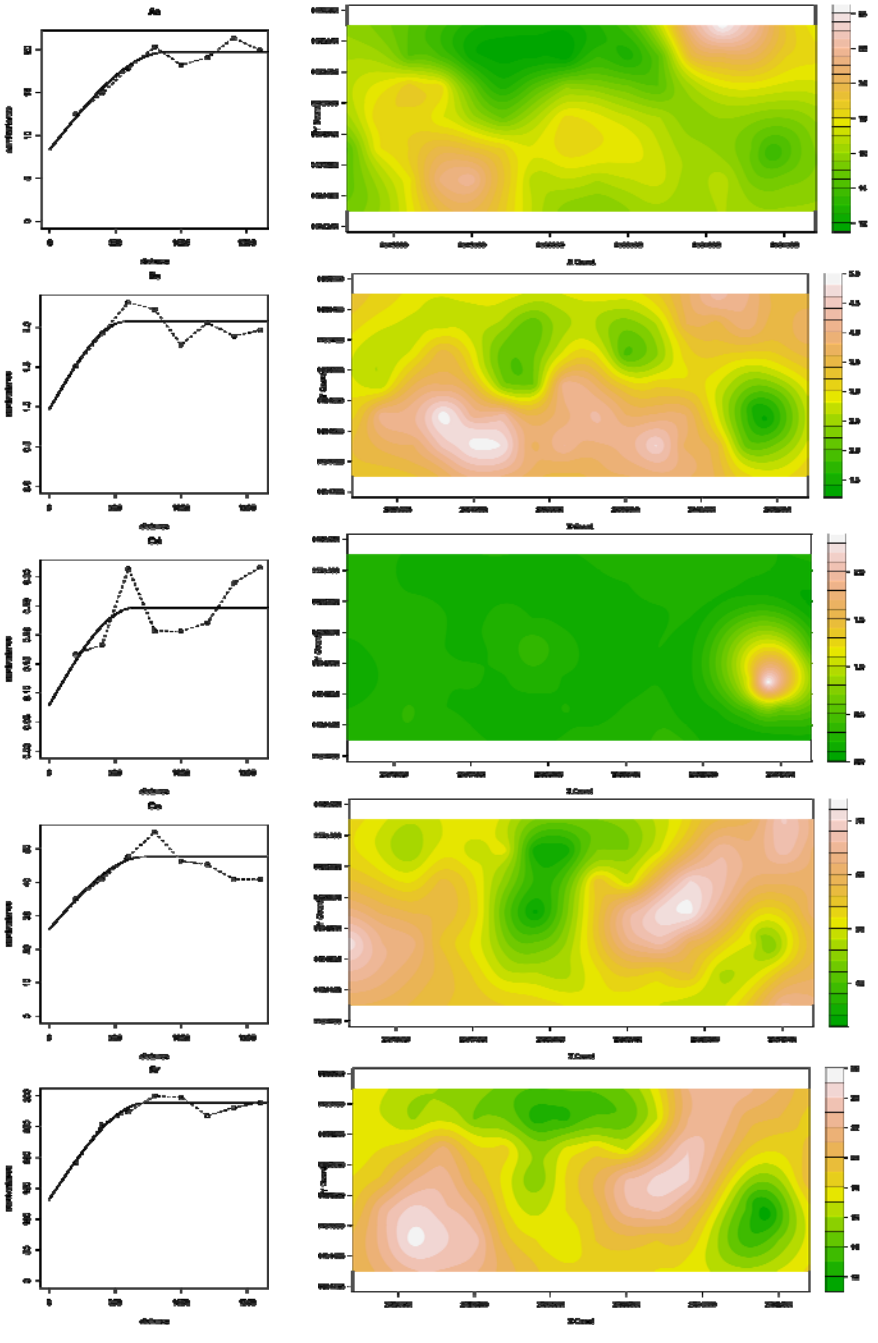


Fig. 5. Experimental semivariogram,background value model and map of elements As, Be, Cd, Co, Cr, Cu and Ni

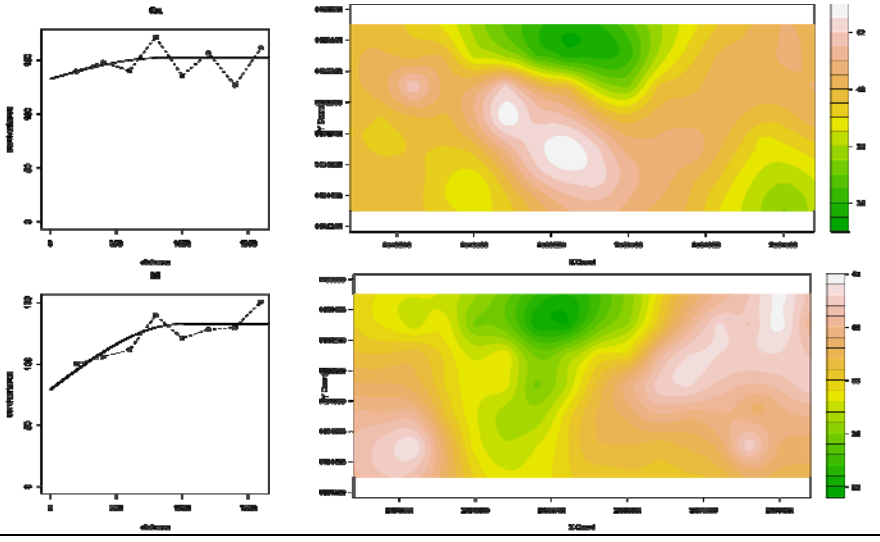


Fig. 5. (*continued*)

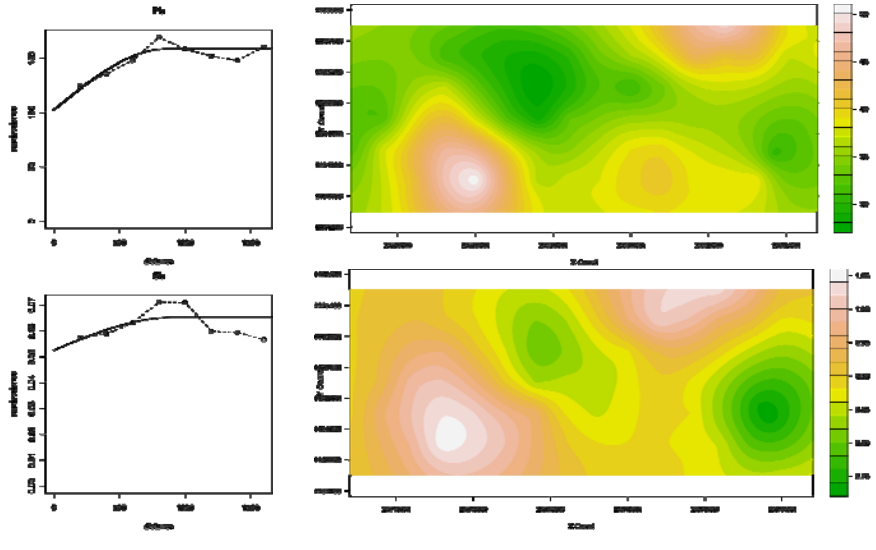


Fig. 6. Experimental semivariogram, background value model and map of elements Pb, Sb, Se, Sn, Tl, V, Zn

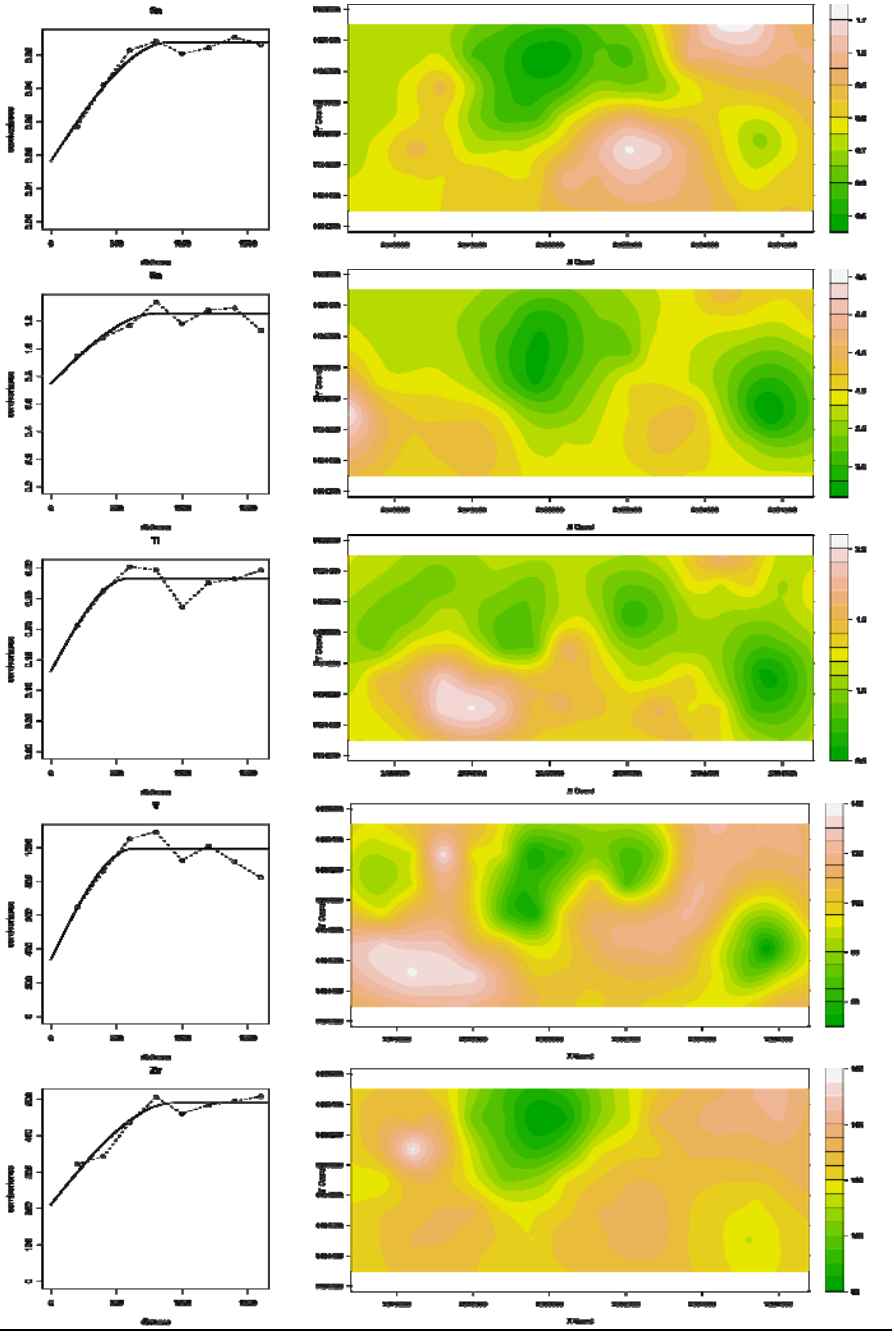


Fig. 6. (*continued*)

Table 6. Cross validation results

		Min	1st Qu.	Median	Mean	3rd Qu.	Max	sd
As	errors	-7.87	-2.54	0.44	-0.02	1.96	1067	3.77
	std.error	-2.03	-0.70	0.12	0.00	0.53	2.89	1.02
Be	errors	-2.67	-0.88	-0.06	0.00	0.68	3.03	1.28
	std.error	-2.11	-0.68	-0.04	0.00	0.52	2.36	0.99
Cd	errors	-1.10	-0.08	-0.03	-0.01	0.03	3.60	0.50
	std.error	-2.64	-0.19	-0.06	-0.01	0.08	9.31	1.28
Co	errors	-13.47	-4.42	0.74	0.01	3.75	12.84	6.18
	std.error	-2.22	-0.72	0.12	0.00	0.61	2.11	1.00
Cr	errors	-39.54	-7.35	-1.70	0.13	7.97	39.87	15.29
	std.error	-2.76	-0.50	-0.12	0.00	0.54	2.76	1.04
Cu	errors	-27.02	-6.72	-1.42	-0.02	5.85	49.16	13.77
	std.error	-2.21	-0.55	-0.12	0.00	0.48	4.03	1.13
Ni	errors	-23.37	-6.70	-1.01	-0.02	5.45	37.09	11.46
	std.error	-2.31	-0.65	-0.10	0.00	0.53	3.64	1.11
Pb	errors	-33.94	-4.71	0.66	-0.01	6.67	34.75	12.48
	std.error	-2.92	-0.40	0.06	0.00	0.58	2.98	1.07
Sb	errors	-0.62	-0.15	-0.01	0.00	0.11	0.69	0.24
	std.error	-2.50	-0.59	-0.06	0.00	0.44	2.80	0.98
Se	errors	-0.34	-0.13	0.01	0.00	0.11	0.39	0.17
	std.error	-1.91	-0.71	0.08	0.00	0.62	2.20	0.93
Sn	Errors	-2.82	-0.44	-0.05	0.01	0.45	5.24	1.13
	std.error	-2.82	-0.44	-0.04	0.00	0.44	4.85	1.09
Tl	Errors	-1.01	-0.35	-0.03	0.00	0.29	1.33	0.47
	std.error	-2.15	-0.72	-0.05	0.00	0.62	2.79	1.00
V	Errors	-70.12	-13.33	5.60	0.22	17.23	81.35	26.92
	std.error	-2.81	-0.50	0.21	0.00	0.62	3.04	1.02
Zn	Errors	-61.25	-12.07	1.58	-0.05	10.09	116.30	24.19
	std.error	-3.29	-0.65	0.09	0.00	0.56	6.41	1.32

4 Conclusions

The role of geostatistics in soil science and its future perspectives have been widely studied. The use of geostatistical techniques in soil geochemical characterization has proved to be very useful for at least two main reasons. The former is due to the possibility of distinguishing pollution double nature: “natural”, related to intrinsic qualitative characteristics of rocks (weathering), and “anthropogenic”, deriving from industrial activities. The latter, since these techniques allow to produce distribution maps of background values, which may represent an efficient

decisional tool for natural resources managers and planners, since they allow to identify risk areas and to plan the most effective recovery actions.

If the geostatistical approach fits to reproduce a spatial structure (correlation) of pollutant data, the interest of such approach remains unaltered even when this correlation is low (e.g. Cu and Sb where the Q index is low).

Nevertheless, the detection of no spatial continuity in concentrations is an important result and suggests that the whole predictive model will be uncertain concerning pollutant species positioning. The knowledge of a heavy metal chemical form is important for future recovery actions of contaminated soils or for future interventions of contamination attenuation. Evaluation of such parameters, essential for planning a recovery project, represents the necessary preamble to deepen both investigations concerning most salient aspects of this work and studies concerning biological aspects.

Vastness of the study area, danger related to pollutant typology and to presence of waste products disposal, together with side instability, have led to consider as extremely interesting the study of the chemical-physical parameters of the site, which showed high environmental risk characteristics..

Due to the extreme anthropogenic impact, levels of the analyzed elements were high in the whole area, and generally higher than admissible limits and values characteristic of polluted soils. Our study showed a high spatial variation of heavy metals in soils from the industrial area analyzed. Nevertheless, the study of spatial distribution highlighted the highest metal concentrations nearby industrial plants. In fact, these plants are the main pollution sources due to their present productive activity and to the long-period contamination deriving from abandoned enterprises. Heavy metal mobility is a problem to be considered for its noxious environmental repercussion, since migration of heavy metals in the soil compartment may lead to dangerous situations for the surrounding environment.

Acknowledgments. This study was supported by Regione Basilicata, Project "Caratterizzazione Geochimica dei Siti Inquinati di Interesse Nazionale di Tito e della Val Basento finalizzata agli interventi di messa in Sicurezza e di Bonifica".

References

Aboal, J.R., Real, C., Fernàndez, J.A., Carballeira, A.: Mapping the results of extensive surveys: The case of atmospheric biomonitoring and terrestrial mosses. Sci. Total Environ. 356, 256–274 (2006)

APAT–ISS Protocollo operativo per la determinazione dei valori di fondo di metalli/metalloidi nei suoli dei siti d'interesse nazionale, Roma, Italy (2006), http://www.apat.gov.it/site/_files/Suolo_Territorio/TEC_valori_di_fondo.pdf

Bailey, T.C.: A review of statistical spatial analysis in geographical information systems. In: Fotheringham, S., Rogerson, P. (eds.) Spatial analysis and GIS. Taylor & Francis, London (1994)

Bivand, R.S., Pebesma, E.J., Gómez-Rubio, V.: Applied spatial data analysis with R. Springer, Heidelberg (2008)

Bocchi, S., Castrignanò, A., Fornaro, F., Maggiore, T.: Application of factorial kriging for mapping soil varation at field scale. European Journal of Agronomy 13(4), 295–308 (2000)

Burrough, P.A.: GIS and Geostatistics: Essential partners for spatial analysis. Environmental and Ecological Statistics 8(4), 361–377 (2001)

Capri, E., Trevisan, M.: I metalli pesanti di origine agricola nei suoli e nelle acque sotterranee. Collana Quaderni di tecniche di protezione ambientale. sezione Protezione delle acque sotterranee. Pitagora Editrice, Bologna (2002)

Castrignanò, A., Maiorana, M., Fornaro, F., Lopez, N.: 3D spatial variability of soil strength and its change over time in a durum wheat field in Southern Italy. Soil & Tillage Research 65(1), 95–108 (2002)

Castrignanò, A., Stelluti, M.: Analisi delle caratteristiche fisico-chimiche dei suoli. In: Flagella, Z., Tarantino, E. (eds.) Caratterizzazione agroecologica del territorio garganico, Claudio Grenzi Editore, Foggia, Italy (2003)

Chiles, J.P., Delfiner, P.: Geostatistics modeling spatial uncertainty. John Wiley & Sons, Chichester (1999)

Cicchella, D., Albanese, S.: Cartografia geochimica con l'uso di sistemi informativi geografici (GIS). In: Lima, A., De Vivo, B., Siegel, F.R. (eds.) Geochimica ambientale, Liguori Editore, Napoli, Italy (2004)

Lgs, D.: 152/2006 (Decreto Legislativo 3/4/2006, n° 152) Norme in materia ambientale. Gazzetta Ufficiale Repubblica Italiana n. 88 del 14/04/2006 (S.O. n. 96)

D.M. 471/99 (Decreto Ministeriale 25/10/1999, n° 471) Regolamento recante criteri, procedure e modalita' per la messa in sicurezza, la bonifica e il ripristino ambientale dei siti inquinati, ai sensi dell'articolo 17 del decreto legislativo 5/02/1997, n. 22, e successive modificazioni e integrazioni. Gazzetta Ufficiale Repubblica Italiana n. 293 del 15/12/1999 (S.O. n. 218)

El Sebai, T., Lagacherie, B., Soulas, G., Martin-Laurent, F.: Spatial variability of isoproturon mineralizing activity within an agricultural field: Geostatistical analysis of a simple physicochemical and microbiological soil parameters. Environ. Pollut. 145, 680–690 (2007)

EPA (Environmental Protection Agency) QA/G-9S, Data quality assessment: Statistical methods for practitioners. Office of Environmental Information, Washington, DC 20460, USA (2006)

Gupta, S.K., Vollmer, M.K., Krebs, R.: The importance of mobile, mobilisable and pseudo total heavy metal fractions in soil for three-level risk assessment and risk management. Sci. Total Environ. 178, 11–20 (1996)

Haining, R.: Spatial Data Analysis: Theory and Practice. Cambridge University Press, Cambridge (2003)

Hlavay, J., Prohaska, T., Weisz, M., Wenzeland, W.W., Stingeder, G.J.: Determination of Trace Elements Bound to Soils and Sediment Fractions. Pure Appl. Chem. 76, 415–422 (2004)

Iacus, S., Masarotto, G.: Laboratorio di statistica con R. McGraw-Hill, New York (2003)

Ihaka, R., Gentleman, R.: R: A language for data analysis and graphics. Journal of computational and Graphical Statistics 5(3), 299–314 (1996)

Journel, A.G., Huijbregts, C.J.: Mining geostatistics. Academic Press, London (1978)

Leita, L., Petruzzelli, G.: Metalli pesanti. In: Violante, P. (ed.) Metodi di Analisi Chimica del Suolo. Cap XI. Collana di metodi analitici per l'agricoltura. Franco Angeli, Milano (2000)

Liang, Y., Wong, M.H.: Spatial and temporal organic and heavy metal pollution at Mai Po Marshes Nature Reserve, Hong Kong. Chemosphere 52, 1647–1658 (2003)

Lima, A., De Vivo, B., Siegel, F.R.: Geochimica ambientale. Liguori Editore, Napoli, Italy (2004)

Link, D.L., Kingstron, H.M.: Use of microwave assisted evaporation for the complete recovery of volatile species of inorganic trace analytes. Anal. Chem. 72(13), 2908–2913 (2000)

Liu, X., Wu, J., Xu, J.: Characterizing the risk assessment of heavy metals and sampling uncertainty analysis in paddy field by geostatistics and GIS. Environ. Pollut. 141, 257–264 (2006)

Lloyd, C.D., Lloyd, L.D.: Local Models for Spatial Analysis. CRC Press, Boca Raton (2006)

Madeo S., Penati, M.: Ricostruzione di campi di concentrazione di inquinanti sulla regione Lombardia mediante interpolatori geostatistici. Tesi di laurea, Politecnico di Milano (2005), http://www.mate.polimi.it/biblioteca/tesiview.php?id=11&L=i

Matheron, G.: Les variables regionalisees et leur estimation: une application de la théorie des fonctions aléatoires aux sciences de la nature. Masson, Paris (1965)

R Development Core Team: R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria (2006), ISBN 3-900051-07-0, http://www.R-project.org

Reimann, C., Filzmoser, P., Garrett, R., Dutter, R.: Statistical Data Analysis Explained: Applied Environmental Statistics with R. John Wiley & Sons, Chichester (2008)

Ribeiro Jr., P.J., Brown, P.E.: Some words on the R project. The ISBA Bulletin 8(1), 12–16 (2001a)

Ribeiro Jr., P.J., Diggle, P.J.: geoR: A package for geostatistical analysis. R-NEWS 1(2), 15–18 (2001b), ISSN 1609-3631

Rodríguez, J.A., Nanos, N., Grau, J.M., Gil, L., López-Arias, M.: Multiscale analysis of heavy metal contents in Spanish agricultural topsoils. Chemosphere 70, 1085–1096 (2008)

Tessier, A., Campbell, P.G.C., Misson, M.: Sequential extraction procedure for the speciation of particulate trace metals. Anal. Chem. 51(7), 844–850 (1979)

Wu, C., Wu, J., Luo, Y., Zhang, H., Teng, Y.: Statistical and geostatistical characterization of heavy metal concentrations in a contaminated area taking account soil map units. Geoderma 144, 171–179 (2008)

Evaluating the Impact of Resolution on the Predictions of an Air Quality Model over Madrid Area (Spain)

Marta G. Vivanco[1], Oier Azula[2], Inmaculada Palomino[1], and Fernando Martín[1]

[1] Atmospheric Pollution Unit, Environmental Department, CIEMAT, Avda. Complutense 22, 28040 Madrid, Spain
[2] Departamento de Química Fisica, UPV, Paseo de la Universidad, 7, 01006 Vitoria-Gasteiz, Spain.
e-mail: m.garcia@ciemat.es

Abstract. Modelling has become a very useful tool in air quality management. The accurate prediction of air quality using numerical models involves correctly simulating both the meteorology and chemical processes. Resolution plays an essential role in the quality of model predictions. Choosing an appropriate resolution is important to obtain good air quality forecasts. In this paper the influence of four different model resolutions on model predictions has been analyzed over the Madrid area for a summer period in 2004. The comparison between model results and observations using traditional evaluation statistics for some pollutants indicate that a significant improvement is found when comparing the coarsest domains, but it is much lower or inexistent for the finest domains. A spectral analysis of a topographic profile crossing the area and the evaluation of predicted meteorological parameters, such as wind speed and temperature, suggests that transitioning from 36-to 19-km grid spacing allows the definition of the major mesoscale topographic features and therefore, the atmospheric circulations. A better behaviour of pollutant predictions should be related to an improvement of meteorological predictions in terms of the parameterizations involved in the meteorological model and the input data, such as land use information. Also the use of more accurate emissions is needed to improve the traditional objective verification scores.

Keywords: ozone, photochemical modeling, air quality, model evaluation, grid resolution.

1 Introduction

A great deal of research is presently focused on air quality modeling. The European Legislation has incorporated several directives regarding air pollutant levels for the last decades. In 2008 the Directive on Ambient Air Quality and Cleaner Air for Europe (Directive 2008/50/EC) has been approved. This directive mentions modeling techniques as a supplement of mandatory measurements in areas where the long-term objectives for ozone or the assessment thresholds for

B. Murgante et al. (Eds.): Geocomputation, Sustainability & Environ. Planning, SCI 348, pp. 145–162.
springerlink.com

other pollutants are exceeded. Model results provide a geographical distribution of concentrations that can support interpretation from observed point data.

In Spain, ground-level ozone and particulate matter concentrations are still exceeding thresholds established in EU legislation to protect human health and prevent damage to ecosystems (Baldasano et al., 2003; EEA, 2005). Monitoring data indicate that ozone concentrations are above European standards at many locations, representing a potential damage on human health. Several studies based on air quality models have been recently carried out over Spain (Jimenez et al., 2006; Pérez et al., 2006; Palacios et al., 2005).

Resolution plays a significant role in modeling, especially when trying to simulate local effects. Higher resolution allows a more accurate definition of pollutant distribution in the atmosphere, although a higher computing time and more computational resources are required. Besides the more realistic picture obtained in a high-resolution domain, a better model performance is also expected. Some studies have examined the effects of horizontal resolution on forecast accuracy. Gego et al. (2005) evaluated temperature and ozone predicted values against observations for two different resolutions. Their study could not confirm the expected improvement of model predictions at higher resolution. Mass et al. (2002) suggested that a decrease of grid spacing in mesoscale models to less than 10-15 Km generally improves the realism of the results but does not necessarily improve the accuracy of the forecasts.

In this paper, an evaluation of the CHIMERE model performance for ozone considering four different resolutions was carried out for a summer period in 2004 over Madrid area. The CHIMERE is a chemistry-transport model that has been extensively used over Europe (Schmidt et al. 2001, Bessagnet et al. 2004, Hodzic et al., 2005) and Spain (Vivanco et al. 2008). Simulations over the Iberian Peninsula at a 0.2º horizontal resolution covering the Iberian Peninsula indicated a significant underestimation of ozone levels in Madrid area. An analysis of the impact of model resolution for this area presents a great interest as an attempt of improving ozone predictions over this area.

Madrid region is one of the most populated zones of Spain with more than 6 millions of inhabitants over an area of more than 8000 Km^2. Emissions are mainly affected by road transport. The industrial sector is not so significant over this region, although several power plants are located in the neighbor provinces. High ozone level episodes are quite frequent. Some of them were reported in 2004 at some air quality stations, especially during June and July.

A summer period in 2004 starting on June 1 and ending on July 27 was evaluated.

2 Model Simulations

Simulations of photochemical compounds were carried out using the regional V200603par-rc1 version of the CHIMERE model. This version calculates the concentration of 44 gaseous species and both inorganic and organic aerosols of primary and secondary origin, including primary particulate matter, mineral dust, sulfate, nitrate, ammonium, secondary organic species and water.

To investigate the effect of the horizontal grid spacing on model predictions, four model configurations were chosen. The largest domain (EUR in Figure 1), at European scale, covers an area ranging from 10.5W to 22.5E and from 35N to 57.5 N with a 0.5 degree horizontal resolution and 14 vertical sigma-pressure levels extending up to 500 hPa. A second domain (SP) is focused over the Iberian Peninsula, with a 0.2 degree resolution. The other two domains were located around Madrid and present a horizontal resolution of 0.07º and 0.04º. All these domains are presented in Figure 1. A nesting procedure was used; coarse-grid simulations forced the fine-grid ones at the boundaries without feedback.

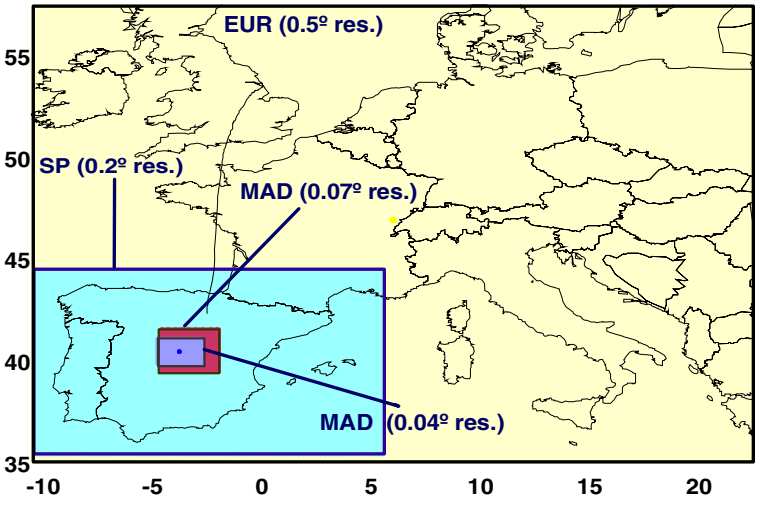


Fig. 1. Simulation domains for CHIMERE model

The emissions for all the simulations were derived from the annual totals of the EMEP database for 2004 (Vestreng et al., 2005). Original EMEP emissions were disaggregated taking into account the land use information, in order to get higher resolution emission data. The spatial emission distribution from the EMEP grid to the CHIMERE grid was performed using an intermediate fine grid at 1km resolution. This high-resolution land use inventory comes from the Global Land Cover Facility (GLCF) data set (http://change.gsfc.nasa.gov/create.html). For each SNAP activity sector, the total NMVOC emission is split into emissions of 227 real individual NMVOC according to the AEAT speciation (Passant, 2002), and real species emissions are aggregated into model species emissions. Biogenic emissions are computed according to the methodology described in Simpson et al. (1995), for alpha-pinene, NO and isoprene. Boundary conditions for the coarse domains were provided from monthly 2003 climatology from LMDz-INCA model (Hauglustaine et al. 2004) for gases concentrations and from monthly 2004 climatology from GOCART model (Chin et al., 2002) for particulate species.

The MM5 model (Grell et al., 1995) was used to obtain meteorological input fields. The simulations were carried out also for four domains. Horizontal grids of

36 km, 19 km, 7 km and 3 km were chosen. All the modeling options such as the number of vertical levels and the parameterization for the treatment of different physical processes were common between all the simulations. The two coarsest MM5 simulations were forced by the National Centres for Environmental Prediction model (GFS) analyses. The third domain was nested to a 21 km resolution MM5 simulation and the fourth was obtained from a 3-nest setup, with horizontal grids of 27 km, 9 km and 3 km. Figure 2 illustrates the setups used for the four simulations.

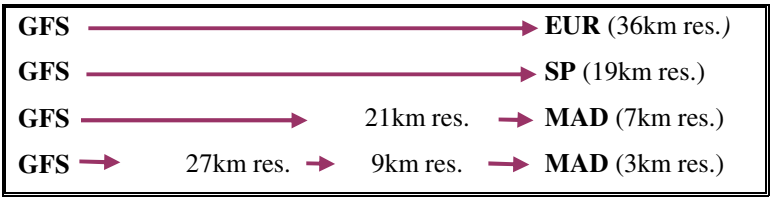


Fig. 2. MM5 simulation schemes for the four domains. Res: resolution

The CHIMERE model was run for a summer period in 2004, from June 1 to July 27. Although the four domains cover different areas, just the geographical region included in MAD (0.04º res.) domain (Figure 1) was analyzed.

3 Spectral Analysis of a Topography Profile

In order to find a way of numerically predict the impact of model resolution a spectral analysis of a topographic profile crossing the area was done. The objective of this analysis is based on the assumption that atmospheric circulations are primary forced by topography and that a frequency spectrum can provide the spatial dimension of the most frequent topographic features. Consequently, the frequency spectrum of a topographic profile can provide some information about the minimum resolution necessary to capture the most significant topographic features and therefore the minimum resolution required to reproduce atmospheric circulations. The topographic profile used for this spectral analysis is presented in Figure 3. The normalized spectral density obtained is shown in Figure 4. On the assumption that at least two points are needed to represent a certain feature, we first calculated the smallest scale that can be resolved at the four different resolutions: topographic accidents affecting distances of 100 km can be resolved at a 0.5º horizontal resolution.; topographic features of 40 km can be captured at a 0.2º resolution; features of 18 km can be represented at a 0.07º resolution and structures of 8 Km can be resolved at a 0.04º horizontal resolution. Figure 4 indicates that features affecting distances of 40 km contribute significantly to the spectral density. Below 40 km, there is a decrease in the spectral density, indicating that the most relevant improvement of predicted atmospheric processes should be found when comparing 0.5º to 0.2º resolution. Nevertheless, this improvement should not be so large when comparing 0.2º to 0.07º horizontal

resolution, as features of 18 km contribute almost ten times less than those of 40 km in the spectrum. Topographical accidents covering distances of 9 km contribute almost ten times less than those covering distances of 18 km. For this reason, a very small improvement should be also expected at 0.04º resolution.

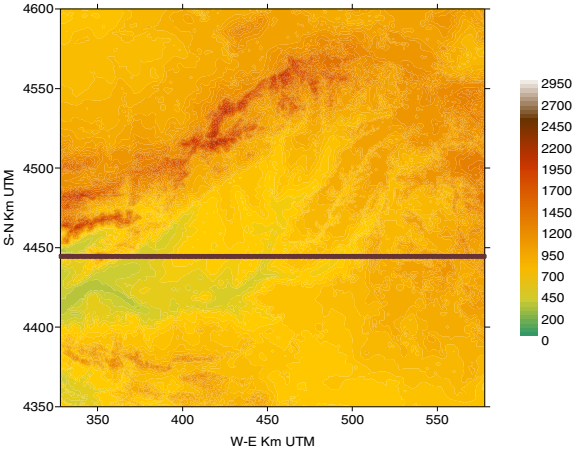


Fig. 3. Location of the profile used to calculate the spectral analysis

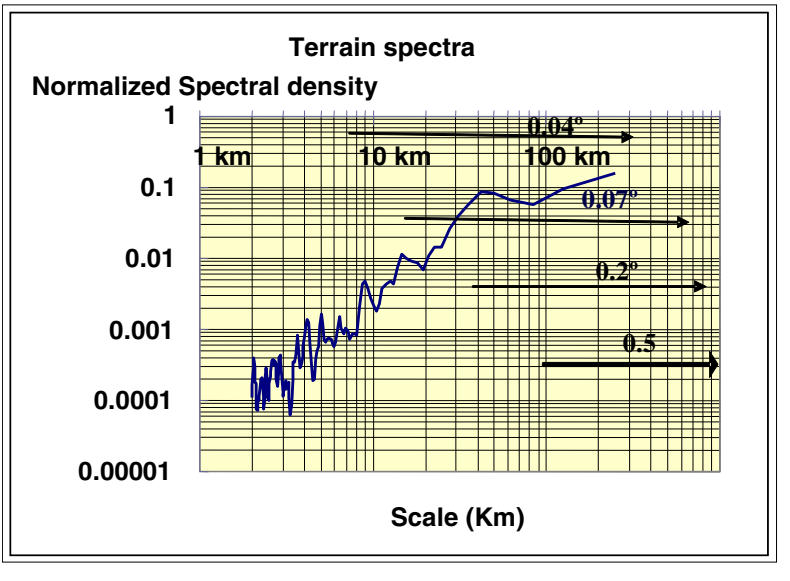


Fig. 4. Spectrum of a topographic profile in Madrid area

4 Model Results

In order to illustrate the effect of model resolution on the spatial distribution of predicted concentration, O_3 and NO_2 contour maps for June 29 2004, 5 p.m. have been included in figures 5 and 6. Higher resolution allows a better description of

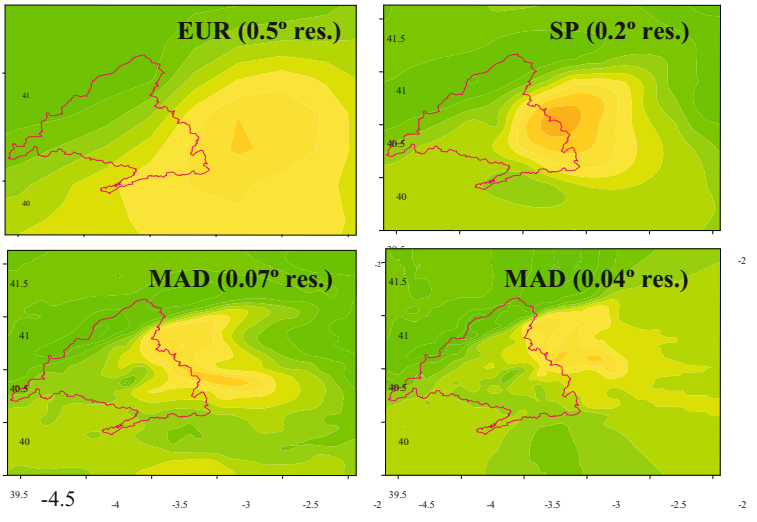


Fig. 5. Ozone contour maps for June 29, 17:00. An increase of resolution drives to a higher definition of ozone transport.

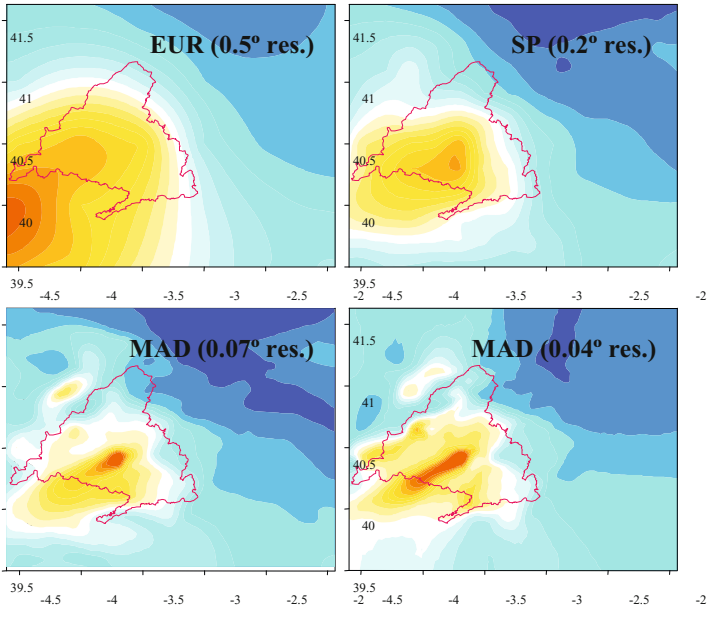


Fig. 6. NO2 contour maps for June 29,17:00

pollutant transport and a finer texture of model predictions, a fact that is very important for air quality management purposes. The increased grid resolution drives to a more realistic appearance of pollutant concentrations in the atmosphere. This fact is specially observed for NO_2 as it can be inferred from the highest-resolution contour map. The area with maximum values is concentrated in a smaller area, presenting higher values.

5 Evaluation of Model Predictions

First of all, an evaluation of MM5 predicted meteorological variables was done. Temperature and wind speed predicted values were compared to observations at nine meteorological sites. To evaluate air quality CHIMERE predictions, some statistics were calculated. Table 1 presents the metrics used and their definition. Parameters such as mean bias (B_{MB}), mean normalized bias (B_{MNB}), mean normalized absolute error (E_{MNAE}), root mean square error (E_{RMSE}) and root mean normalized square error (E_{RMNSE}) were estimated for O_3 and NO_2. Regarding ozone, only statistics for moderate-to-high ozone concentration cases (more important for human health protection) were considered by selecting predicted-observed value pairs when hourly observations were equal to or greater than the cutoff of 80 μgm^{-3}. For NO_2 and NO_x a cutoff value of 5 μgm^{-3} was used. 45 air quality sites were taken into account to estimate ozone statistics. For NO_2 information from 41 stations was considered and 38 for SO_2.

Table 1. Definition of the metrics used in the evaluation of the CHIMERE model performance

Mean bias	$B_{MB} = \frac{1}{N}\sum(M_i - O_i) = \overline{M} - \overline{O}$
Mean normalized bias	$B_{MNB} = \frac{1}{N}\sum\left(\frac{M_i - O_i}{O_i}\right) = \left(\frac{1}{N}\sum\frac{M_i}{O_i} - 1\right)$
Mean normalized absolute error	$E_{MNAE} = \frac{1}{N}\sum\left(\frac{\lvert M_i - O_i\rvert}{O_i}\right)$
Root mean square error	$E_{RMSE} = \left[\frac{1}{N}\sum(M_i - O_i)^2\right]^{\frac{1}{2}}$
Root mean normalized square error	$E_{RMNSE} = \left[\frac{1}{N}\sum\left(\frac{M_i - O_i}{O_i}\right)^2\right]^{\frac{1}{2}}$

N: pairs of modeled and observed concentrations M_i and O_i. The index i is over time series and over all the locations in the domain.

5.1 Meteorological Model Evaluation

An evaluation of temperature and wind speed for the four domains was carried out in order to determine if a higher resolution provides a better simulation of atmospheric circulation. MM5 predictions were compared to information recorded at 9 meteorological sites. The location of these stations is presented in Figure 7. Some statistical parameters were evaluated.

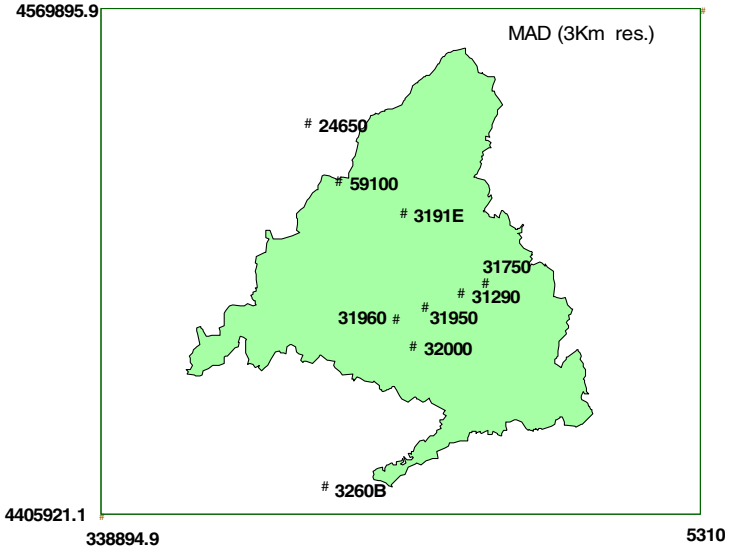


Fig. 7. Location of the meteorological sites inside the 3-km resolution domain

Figure 8 illustrates, the mean normalized absolute error obtained for temperature and wind speed at all the individual meteorological sites. Regarding wind speed there is a general improvement of model accuracy at higher resolution, with the exception of 59100 station (Navacerrada), located on a mountain saddle. In some of the stations, especially those located in the middle of the domain, there is a significant improvement of wind speed when transitioning from 36 to 19-km grid spacing. It seems that the major mesoscale topographical features of the region and their corresponding atmospheric circulations are better defined at the 19-km resolution, producing a positive effect on the verifications. In some of these stations, such as 31290 and 32000 the improvement at higher resolution is also observed for 9 and 3 km resolution. In others (31750, 31950), there is not an improvement in verification statistics as grid spacing decreased from 9 to 3 km.

The highest error for temperature is found for 59100 stations, presenting the largest error for the highest-resolution domain.

Time series of wind speed and temperature are included in figures 9 and 10. These figures contain hourly observed and predicted meteorological parameters. The behavior of modeled temperature has a similar pattern for all the domains.

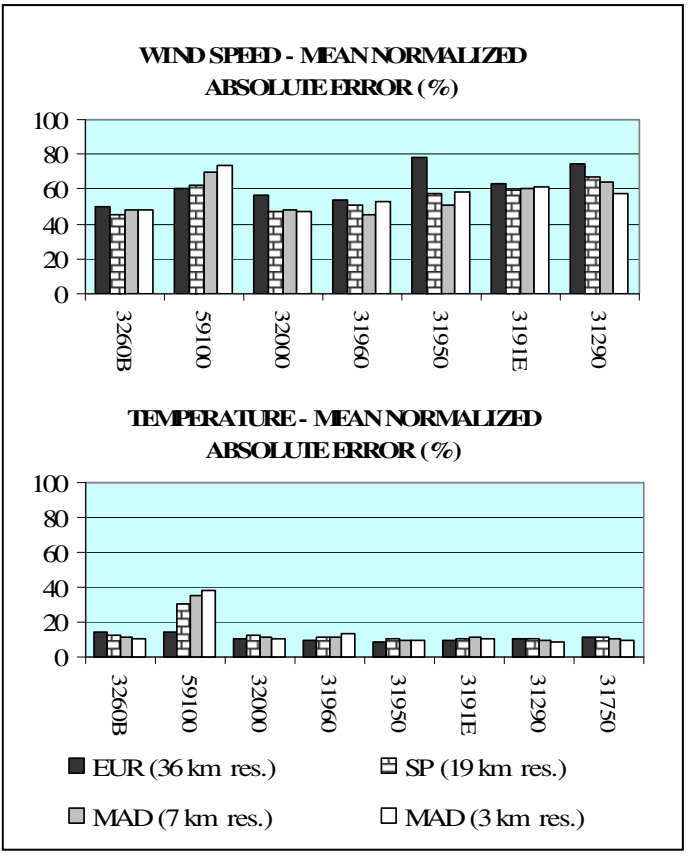


Fig. 8. Mean normalized absolute error for wind speed and temperature at all the meteorological stations in the domain

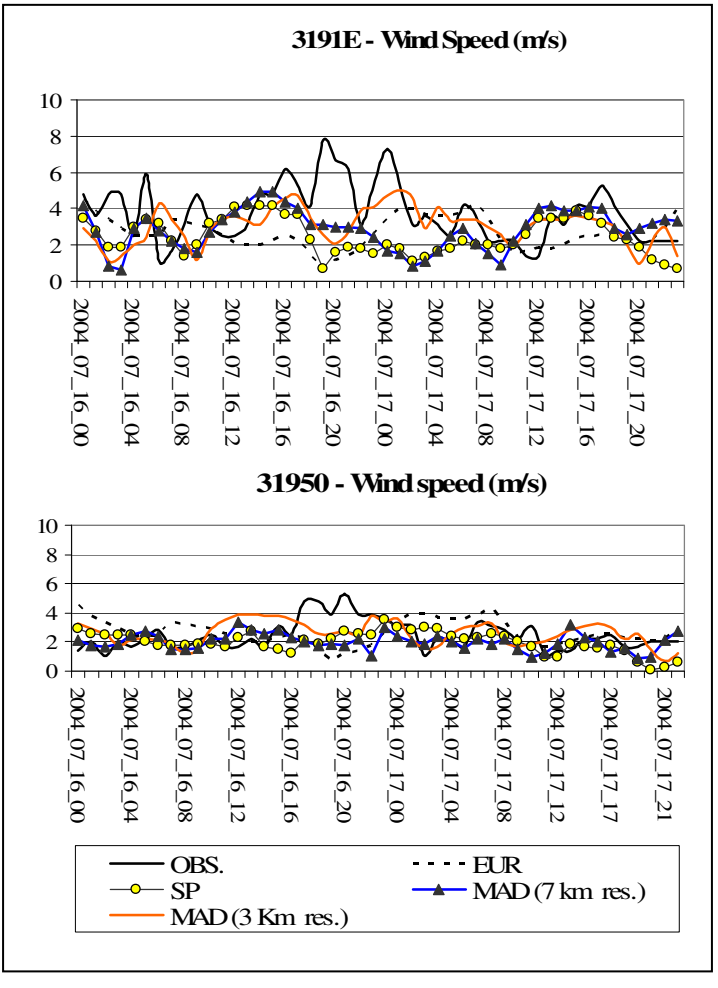


Fig. 9. Time series of predicted and observed hourly wind speed at two meteorological sites

5.2 Evaluation of Air Quality Predictions

In order to compare the model predictions for the four resolutions, hourly ozone concentrations were compared to observations for all the available air quality stations located in the area. Taking into account the conclusions in Palacios et al. (2005) about the ozone transport from the Greater Madrid to regions located 100 km away from Madrid , a large set of stations covering a wide area was used, (45, 41and 38 stations for ozone, NO_2 and SO_2 respectively). A statistical evaluation of the model performance was done, taking into account the metrics described in the previous section.

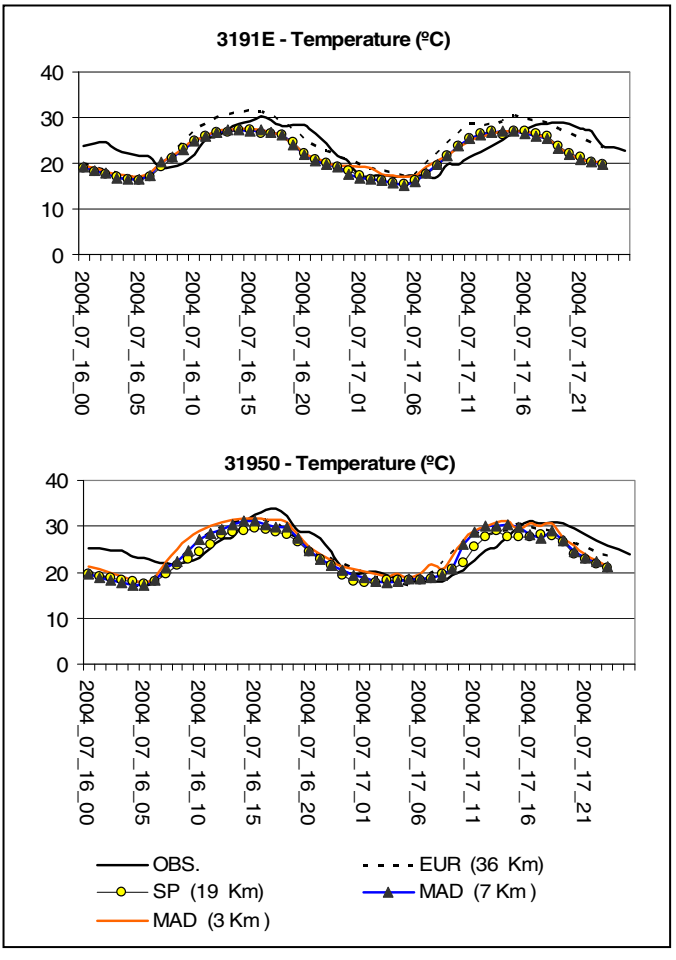


Fig. 10. Time series of predicted and observed hourly temperature at two meteorological sites

Table 2 reports the statistical results for ozone, using hourly ozone values for the entire period. For all the domains, mean normalized bias and mean normalized absolute error present values inside the range proposed in Tesche et al (1990) to consider an acceptable model performance (<15% and <30% respectively). An improvement of statistic scores is found when downscaling to 0.07º res. Nevertheless, comparing 0.07º versus 0.04º horizontal resolution the scores are very similar, indicating that a better agreement between observations and model predictions is not obtained.

Table 2. Analysis of model performance for ozone. Statistics for ozone evaluation for 2004. Based on hourly values higher than 80 μgm^{-3}.

Statistical measure	EUR (0.5º res.)	SP (0.2º res.)	MAD (0.07º res.)	MAD (0.04º res.)
Mean bias ($\mu g\ m^{-3}$)	5.9	0.9	-1.1	0.1
Mean normalized bias (%)	7.9	3	1.1	2.4
Mean normalized absolute error (%)	20.9	17.7	15.6	15.2
Root mean square error ($\mu g\ m^{-3}$)	24.1	21.7	19.8	19.3
Root mean normalized square error (%)	24.9	21.3	18.9	18.5

A plot showing the mean normalized absolute error for the individual stations and for all the domains is presented in Figure 11. The improvement observed for the two finest domains is significant at some monitoring sites, especially at traffic stations. However, for both 0.07º and 0.04º resolutions the MNAE presents very similar values.

Statistics for NO_2 calculated when the observed value was higher than 5 μgm^{-3} are presented in Table 3. Chemiluminescence $NO/NO_x/NO_2$ analyzers can respond to other nitrogen containing compounds, such as peroxyacetyl nitrate (PAN), which

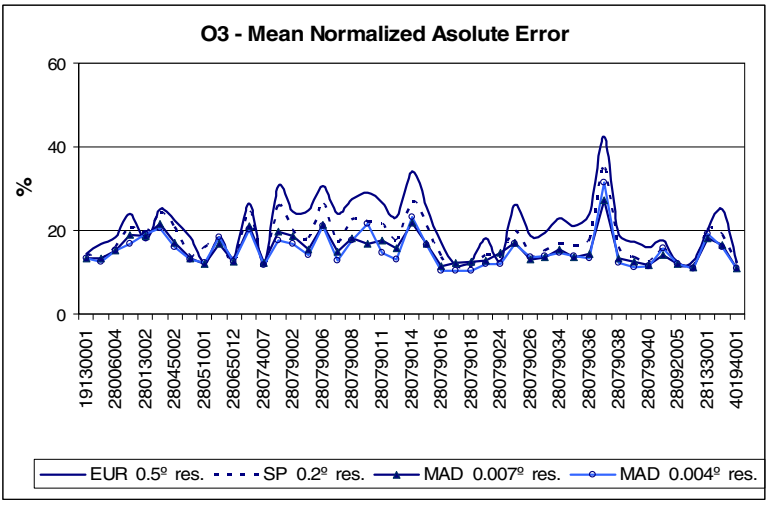


Fig. 11. Mean normalized absolute error for ozone at all the air quality sites in the Madrid area

Table 3. Analysis of model performance for NO2

Statistical measure	EUR (0.5º res.)	SP (0.2º res.)	MAD (0.07º res.)	MAD (0.04º res.)
Mean bias ($\mu g\ m^{-3}$)	-33.0	-32.7	-30.6	-31.7
Mean normalized bias (%)	-45.4	-49.5	-46.1	-50
Mean normalized absolute error (%)	66.7	61.8	58.7	58.7
Root mean square error ($\mu g\ m^{-3}$)	43.6	41.4	39.8	40.4
Root mean normalized square error (%)	76.3	68.3	65.8	64.2

might be reduced to NO in the thermal converter (Winer, 1974). Although atmospheric concentrations of these potential interferences are generally low relative to NO_2 (especially in urban areas) we have added modeled PAN and HNO_3 compounds to modeled NO_2, in order to compare model results to NO_2 and NO_x observations.

The improvement with resolution is no so clearly observed. The highest improvement is found for the SP domain when comparing with EUR domain. In general an improvement of scores is observed when increasing resolution from 0.5º to 0.07º. Again, scores for 0.07º and 0.04º are very similar, indicating that model predictions at both scales are very similar. A considerably high underestimation is found for all the domains.

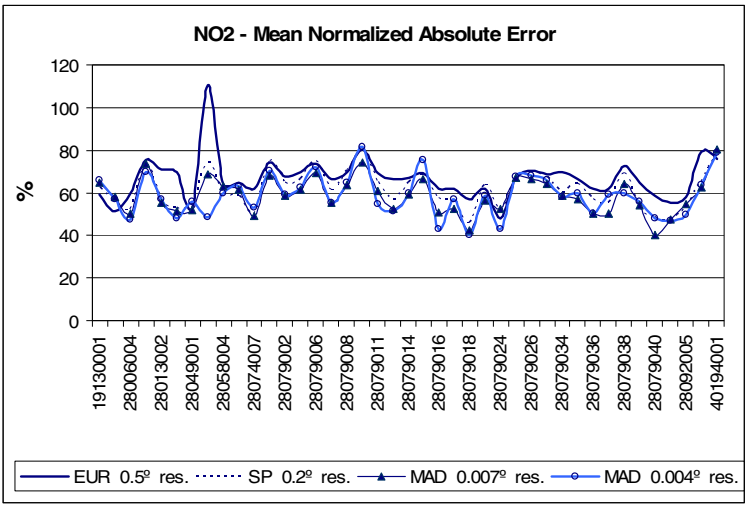


Fig. 12. Mean Normalized Absolute Error for NO2 the air quality sites in Madrid area

Figure 12 shows the mean normalized absolute error for NO_2 at all the individual sites. Some stations, especially those with high errors, present a decrease in normalized absolute error at higher resolutions.

For SO_2, again predictions for the two finest domains present the best scores, although there is no noticeable difference between them. Mean statistical results are presented in Table 4; Figure 13 shows the behavior at the individual monitoring sites.

Table 4. Analysis of model performance for SO_2

Statistical measure	EUR (0.5º res.)	SP (0.2º res.)	MAD (0.07º res.)	MAD (0.04º res.)
Mean bias ($\mu g\ m^{-3}$)	-5.5	-5.1	-4.3	-4.3
Mean normalized bias (%)	-66.8	-61.5	-51.9	-52
Mean normalized absolute error (%)	68.1	62.6	57.0	57.8
Root mean square error ($\mu g\ m^{-3}$)	5.8	6.3	5.8	5.5
Root mean normalized square error (%)	70.7	65.2	61.3	62.4

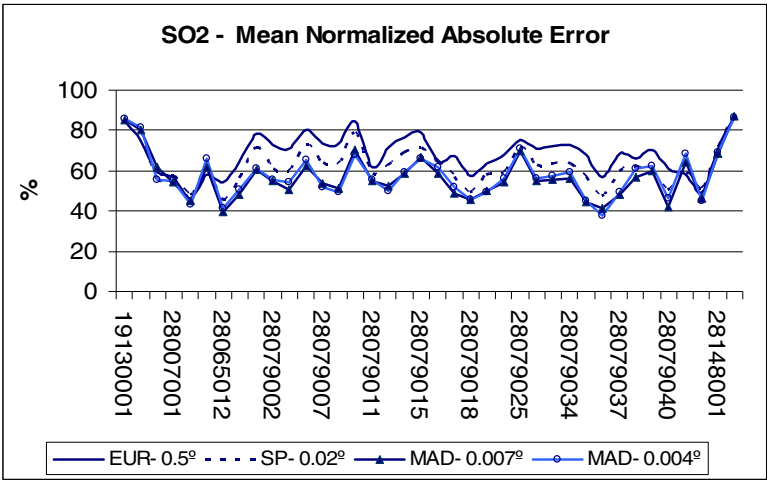


Fig. 13. Mean Normalized Absolute Error for SO2 the air quality sites in Madrid area

Time series showing model and observed hourly values (O_3, NO_2) at some of the stations are also included in order to illustrate the temporal behavior of the model. Figure 14 illustrates the time series at two rural background stations for a two-days episode in 2004 (June 16-17). Figure 15 shows time series at two urban background stations. Some underprediction is observed for the maximum values. A very similar behavior is found for all the simulations. Finally, time series for an urban traffic stations are presented in Figure 16. Some differences can be distinguished between the 0.07º and 0.04º NO2 predictions. In this particular monitoring site and for the two days considered, it seems that the better agreement between observations and predictions is obtained at a 0.04º resolution.

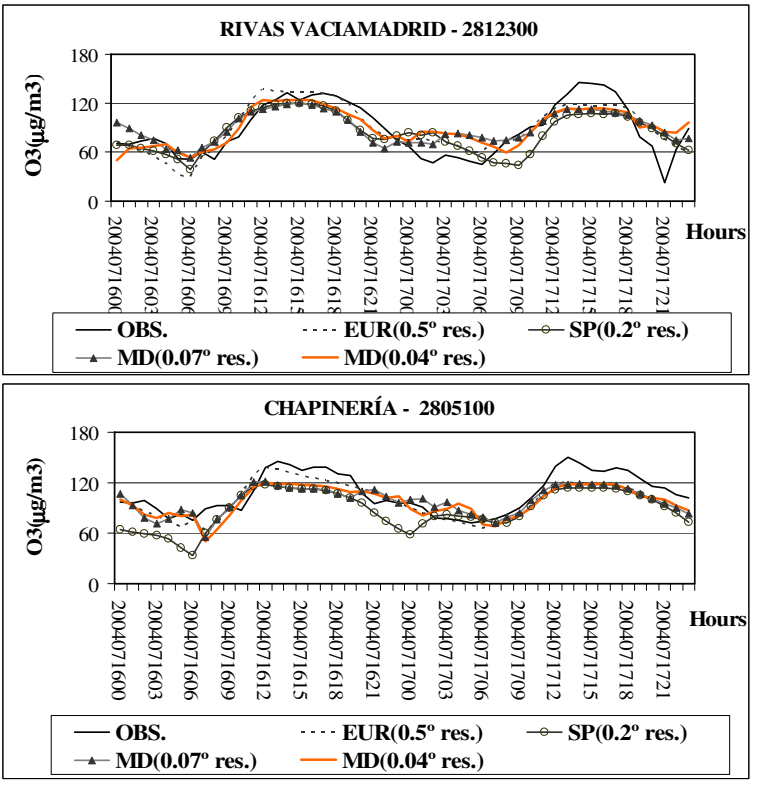


Fig. 14. Time series of observed and predicted ozone at two rural background sites for 2004 June 16-17

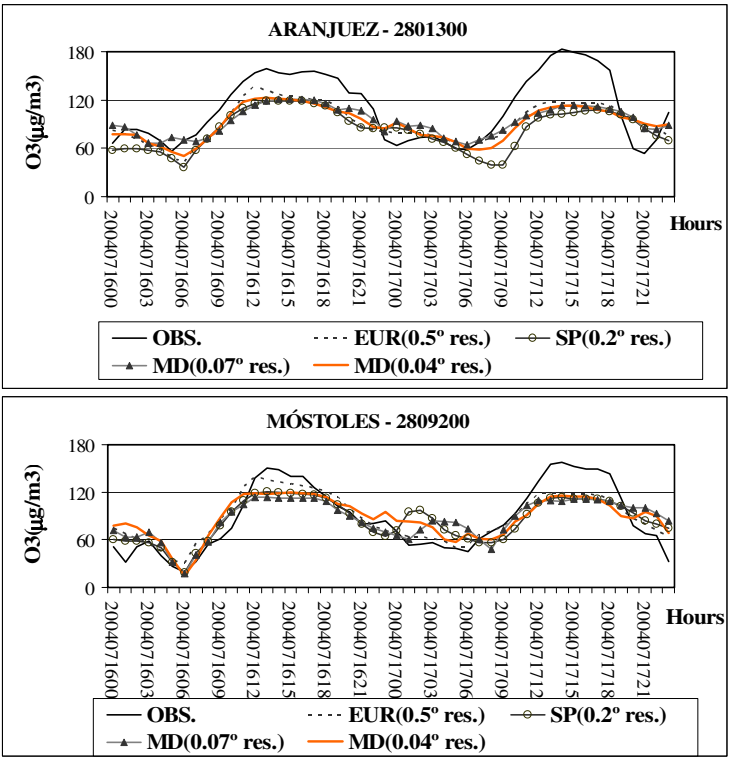


Fig. 15. Time series of observed and predicted ozone at two urban background sites for 2004 June 16-17

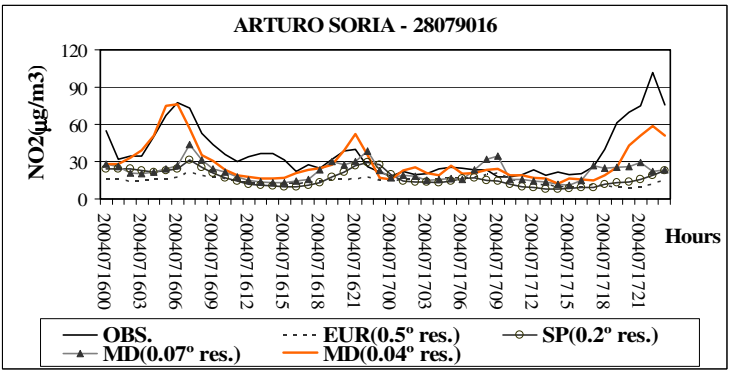


Fig. 16. Time series of observed and predicted NO2 at an urban traffic station for 2004 June 16-17

6 Conclusions

The impact of the spatial resolution on the simulations of meteorological conditions and air quality has been analyzed for the MM5-CHIMERE models. Four domains were used: 1) EUR, at a 0.5º horizontal resolution, 2) SP, at a 0.2º resolution, 3) MAD (0.07º res.) and 4) MAD (0.04º res.) The analysis was carried out for a summer period between June 1 and July 27. The analyzed pollutants were O_3, NO_2 and SO_2.

Traditional evaluation statistics such as mean bias, mean normalized bias, mean normalized absolute error, root mean square error and root mean normalized square error were used to compare observations and predictions. Results show that there is an improvement of the statistical scores at higher resolution. However, the superiority of the finest grid could not be confirmed. The 0.04º ozone predictions appear to be the same accurate as 0.07º predictions. It must be pointed out that the most important improvements are observed between the 0.5º and 0.2º resolution simulations. Regarding the meteorological variables, the major differences in wind speed are found between the 36-km and 19-km resolution simulations. It seems that the differences observed in the CHIMERE results could be very related to the performance of the MM5 model.

Acknowledgments. This work is funded by the Caliope project of the Spanish Ministry of the Environment (441/2006/3-12.1 and A357/200/2-12.1).

References

Baldasano, J.M., Valera, E., Jiménez, P.: Air quality data from large cities. The Science of the Total Environment 307, 141–165 (2003)

Bessagnet, B., Hodzic, A., Vautard, R., Beekmann, M., Rouil, L., Rosset, R.: Aerosol modeling with CHIMERE –first evaluation at continental scale. Atmospheric Environment 38, 2803–2817 (2004)

Chin, M., Ginoux, P., Kinne, S., Holben, B.N., Duncan, B.N., Martin, R.V., Logan, J.A., Higurashi, A., Nakajima, T.: Tropospheric aerosol optical thickness from the GOCART model and comparisons with satellite and sunphotometer measurements. J. Atmos. Sci. 59, 461–483 (2002)

EEA. Air Pollution by Ozone in Europe in summer 2005 Overview of exceedances of EC ozone threshold values for April-September 2005. EEA Technical Report No 3/2006 (2006)

Gego, E., Hogrefe, C., Kallos, G., Voudouri, A., Irwin, J.S., Rao, S.T.: Examination of model predictions at different horizontal grid resolutions. Environmental Fluid Mechanics 5(1-2), 63–85 (2005)

Grell, G.A., Dudhia, J., Stauffer, D.R.: A Description of the Fifth-Generation Penn State/NCAR Mesoscale Model (MM5). NCAR/TN-398 + STR. NCAR TECHNICAL NOTE (1995)

Hauglustaine, D.A., Hourdin, F., Jourdain, L., Filiberti, M.A., Walters, S., Lamarque, J.F., Holland, E.A.: Interactive chemistry in the Laboratoire de Météorologie Dynamique general circulation model: Description and background tropospheric chemistry evaluation. J. Geophys. Res., 109 (2004), doi:10.1029/2003JD003957

Hodzic, A., Vautard, R., Bessagnet, B., Lattuati, M., Moreto, F.: On the quality of long-term urban particulate matter simulation with the CHIMERE model. Atmospheric Environment 39, 5851–5864 (2005)

Jiménez, P., Lelieveld, J., Baldasano, J.M.: Multiscale modeling of air pollutants dynamics in the northwestern Mediterranean basin during a typical summertime episode. Journal of Geophysical Research 111, D18306 (2006), doi:10.1029/2005JD006516

Mass, C., Ovens, D., Albright, M., Westrick, K.: Does Increasing Horizontal Resolution Produce Better Forecasts?: The Results of Two Years of Real-Time Numerical Weather Prediction in the Pacific Northwest. Bull. Amer. Meteor. Soc. 83, 407–430 (2002)

Palacios, M., Martín, F., Aceña, B.: Estimate of potentially high ozone concentration areas in the centre of the Iberian Peninsula. Int. J. Environment and Pollution 24(1/2/3/4), 260–271 (2005)

Passant, N.R.: Speciation of UK emissions of non-methane volatile organic compounds. AEAT/ENV/R/0545 Issue 1 (2002)

Pérez, C., Nickovic, S., Baldasano, J.M., Sicard, M., Rocadenbosh, F., Cachorr, V.E.: A long Saharan dust event over the western Mediterranean: Lidar, sun photometer observations and regional dust modeling. Journal of Geophysical Research 111, 15214 (2006), doi:10.1029/2005JD006579

Schmidt, H., Derognat, C., Vautard, R., Beekmann, M.: A comparison of simulated and observed ozone mixing ratios for the summer of 1998 in western Europe. Atmospheric Environment 35, 6277–6297 (2001)

Simpson, D., Guenther, A., Hewitt, C.N., Steinbrecher, R.: Biogenic emissions in Europe 1. Estimates and uncertainties. Journal of Geophysical Research 100(D11), 22875–22890 (1995)

Tesche, T.W., Georgopoulos, P., Seinfeld, J.H., Cass, G., Lurmann, F.W., Roth, P.M.: Improvement of procedures for evaluating photochemical models. Draft final report prepared for California Air Resources Board, Sacramento, CA by Radian (1990)

Vestreng, V., Breivik, K., Adams, M., Wagener, A., Goodwin, J., Rozovskkaya, O., Pacyna, J.M.: Inventory Review 2005, Emission Data reported to LRTAP Convention and NEC Directive, Initial review of HMs and POPs, Technical report MSC-W 1/2005 (2005), ISSN 0804-2446

Vivanco, M.G., Palomino, I., Vautard, R., Bessagnet, B., Martín, F., Menut, L., Jiménez, S.: Multi-year assessment of photochemical air quality simulation over SPAIN. Environmental Modeling & Software (2008)

Winer, A.M., Peters, J.W., Smith, J.P., Pitts, J.N.: Response of commercial chemiluminiscent NO-NO2 analyzers to other nitrogen-containing compounds. Environmental Science & Technology 8, 1118–1121 (1974)

Spatial OnLine Analytical Processing of Geographic Data through the Google Earth Interface

Sergio Di Martino[1], Sandro Bimonte[2], Michela Bertolotto[3], Filomena Ferrucci[4], and Vincenza Leano[1]

[1] University of Naples "Federico II", Napoli, Italy
e-mail: dimartino@na.infn.it
[2] Cemagref, UR TSCF, 24 Avenue des Landais, 63172 Clermont-Ferrand, France
e-mail: sandro.bimonte@cemagref.fr
[3] University College Dublin, Belfield, Dublin 4, Ireland
e-mail: michela.bertolotto@ucd.ie
[4] University of Salerno, Fisciano (SA), Italy
e-mail: fferrucci@ unisa.it

Abstract. OnLine Analytical Processing (OLAP) tools act as support systems for Decision Makers to discover new knowledge hidden within data warehouses. In the spatial domain this capability is crucial. However, notwithstanding the pressing need for Spatial OLAP (SOLAP) tools, only very few are currently available. Such tools present several limitations in terms of their flexibility in the functionality and the analytical properties they provide. To overcome these limitations, we have developed a web-based SOLAP tool, which relies on the integration of a standard Geobrowser (Google Earth) with a freely available OLAP engine, namely Mondrian. Our system allows a Decision Maker to perform exploration and analysis of spatial data both through the Geobrowser and a Pivot Table in a seamlessly fashion. In this paper, we illustrate the main features of the system we have developed, together with the underlying architecture, using a simulated case study.

Keywords: Spatial OLAP, Spatial Data Warehouses, Geovisualizaiton, Geobrowser.

1 Introduction

Thanks to current technologies for data integration, enterprises are enabled to collect even more heterogeneous information in data warehouses. From a business point of view, these repositories can contain very precious, but often hidden, information for improving the competitiveness of the enterprise.

To take advantage of this information, OLAP (*OnLine Analytical Processing*) solutions allow Decision Makers to gain better insight on phenomena via an

B. Murgante et al. (Eds.): Geocomputation, Sustainability & Environ. Planning, SCI 348, pp. 163 –182.
springerlink.com

interactive exploration of data. Indeed, decisional applications provides different analysis axes and granularity, exploiting the multidimensional model of data (Inmon, 1996), enabling Decision Makers to discover unknown phenomena, patterns and data relationships without mastering either the underlying complex structure of the database, or multidimensional query languages (Stolte et al. 2002). As a consequence, a crucial role for the success of OLAP solutions is played by the adopted interaction/visualization techniques, which should effectively support the mental model of the Decision Maker, in order to take advantage of the unbeatable human abilities to perceive visual patterns and to interpret them.

On the other hand, even if it is recognized that most of these data have a spatial component (Franklin, 1992), these solutions do not take into account "*what is special about spatial' data spatial*" (Anselin 1989), i.e. the heterogeneity of physical spaces and the strong spatial correlation of thematic data. As stated by the first Tobler law ("everything is related to everything else, but near things are more related than distant things") (Tobler 1970), thematic data are correlated if they are spatially near. This means that to really understand spatial phenomenon, the spatial analyst need the visual representation of thematic data on the space using for example a map or an orthophoto, as it allows to "see" and understand spatial relations, trends and knowledge. As a consequence, it is recognized in the literature that OLAP systems lacking of cartographic visualization and spatial analysis functionalities, neglect the special nature of geographic data (Caron 1998).

In order to introduce spatial aspects of the geographic information into the decisional process, new kinds of decision-support systems (Goodchild and Denshman, 1990) has been developed: *Spatial OLAP* (SOLAP).

SOLAP has been defined by Bedard as a "as a visual platform built especially to support rapid and easy spatiotemporal analysis and exploration of data following a multidimensional approach comprised of aggregation levels available in cartographic displays as well as in tabular and diagram displays" (Bédard 1997). The idea underlying the SOLAP approach is to explicitly take into account the spatial nature of data in the decision-making process thanks to advanced *Geovisualization* techniques, coherently merged with traditional OLAP interaction techniques.

Indeed, Geovisualization integrates scientific visualization, cartographic, image analysis, data exploration, data analysis to provide theory, methods and tools to identify relevant geospatial information, data, and knowledge (MacEachren et Kraak 2001). Geovisualization methods and tools improve humans mental analysis capabilities through adequate visualizations and interactive techniques allowing users to "see" knowledge inside data which usually is complex (i.e. 3d data, multimedia data) and huge.

In spite of the importance of this field, to date very few tools have been developed that integrate OLAP and geovisualization techniques (see Section 2). In any case, they still miss some key points for a more flexible and effective spatial analysis process, and they could be enriched by the following aspects:

- 3D visualizations of geographic data, in some contexts (i.e. environmental data, etc.) are mandatory for detecting and understanding geo-spatial phenomena (Bleisch and Nebiker 2008) as 3D displays help user in orientation and provide a more natural description of landforms and spatial aspects than traditional 2D displays. Then, SOLAP tools integrating 3D displays will stimulate spatio-multidimensional analysis, and facilitate SOLAP results queries interpretation.
- SOLAP tools usually rely on proprietary Geographic Information Systems (GIS) technologies (Bimonte 2008). This implies that it is difficult to integrate external geographic data sources into the spatio-multidimensional analysis process (Bédard et al. 2006). By this way, contextualization (Pérez et al. 2007) of spatial data warehouse information is not possible. This represents a main drawback because the spatial analysis process is interactive and iterative (Mitchel 2005), in the sense that spatial analyst cannot define useful geographic information a priori, but she/he should need adding and changing surrounding geographic and thematic elements of interest on the land (e.g. roads, industries, photos, wikipedia description, etc…) to validate hypotheses, understand causes and effects. SOLAP tools will benefit of abilities of geobrowsers to import geographic information from other (potentially remote) data sources is fundamental for this kind of tools.
- Another drawback of SOLAP tools is the personalization of map visual encodings. Indeed, contrary to GIS cartographic visualization techniques of SOLAP tools, geobrowser allow defining personalized advanced visual encodings adapted to analyzed geographic phenomena providing alternative insights on the data, and so revealing additional knowledge.
- Moving from desktop to web-based technologies SOLAP tools could highly improve the spreading and the flexibility of these kinds of solutions.

Since to the best of our knowledge, currently no SOLAP tools are able to provide the above defined features, in this paper, we propose a Web-based system, called Goolap, trying to overcome these limitations. Indeed, it integrates the functionalities provided by a commonly used geobrower and a traditional OLAP system. Goolap enriches geographic exploration and visualization functionalities of a freely available OLAP server, *Mondrian* and an OLAP client, *JPivot,* with the geobrowser *Google Earth*. The main advantage of this solution is to provide a web-based OLAP framework for geographic and environmental analysis. Indeed, Goolap is able to support OLAP analysis, and to render in 3D spatial data stored in different data repositories, with a high degree of personalization of the visual encodings of the information. The main features of our system are validated by an environmental case study concerning the air pollution of Italian regions.

The paper is structured as follows. Section 2 describes the state of art of SOLAP and Geovisualization tools and introduces our environmental case study. In Section 3 we describe the main features of the proposed system, introducing the

user interface of the tool and an example of multidimensional analysis using a simulated case study. In Section 4 we describe the architecture and the technological solutions we adopted for the system we propose to support SOLAP tasks. Some final remarks and future work conclude the paper.

2 Related Work

In this section, we present main SOLAP concepts and systems, and we introduce Geovisualization systems based on Geobrowser tools.

2.1 Geovisualization Systems

Geovisualization systems lie on the integration of GIS, which allows storing, managing, querying, analyzing and visualizing geographic information (Longley et al., 2001), and Visualization systems. The latters are aimed to support analysis and exploration of thematic data using advanced visualization and interaction techniques (Keim et al. 2006).

To allow discovering, analysing and exploring spatial information, Geovisualization tools should provide different spatial analysis techniques into a unique interoperable environment (MacEachren and Kraak 2001; Andrienko et al. 2007). Indeed, these tools should provide visual interfaces that guide user in the analysis process showing relevant data through several techniques for the visualization of complex spatio-temporal constructs. User can explore data, validate and/or reformulate his/her hypothesis using spatial analysis data methods. These data are often heterogeneous and very large. Then, scalability and the support of heterogeneous information are fundamental for Geovisualization systems. Furthermore, Geovisualization tools should provide an important degree of interactivity, which is fundamental for the analysis and exploration of data (MacEachren et al. 2004), supporting significant computational costs in a transparent way for users.

A geobrowser is a software that provides access to rich spatial data sets and sophisticated and intuitive interfaces through which they may be explored. Recently, these geovisualization facilities are being used for exploration and analysis of geographic data. (Zhe-Ping 2007) uses Google Earth as a spatial platform to display the findings about disaster potential of hillside areas around Taipei. In the same way, (Girardin, et al. 2008) present a system for the analysis of tourists movements. (Compieta et al. 2007) exploit Google Earth to display the mining results of 3D spatial data and combine them with other geographic layers, and (Andrienko et al. 2007b) use the 3D capabilities of Google Earth to visualize trajectories data according the *Space Time Cube* metaphor (Hägerstrand 1970). (Bleisch and Nebiker 2008) show the synchronization of 2D graphic displays with spatial data represented in the 3D virtual environment of Google Earth, for the interactive exploration of spatial data. (Slingsby et al. 2007) add different

exploration and analysis functionalities to Google Earth in order to provide an effective Geovisualization system. Thematic mapping is integrated with hypertext descriptions, and ancillary data, such as photos and videos, have been used to better understand data. The tools allow also analyzing data at different scales. Data at lower scales are aggregated using external tools. (Wood et al. 2007) present a mush-up architecture on the base of Google Earth for exploring 3D geographic data, collecting geographic data from the web, supporting multimedia data on top of geographic layers, using advanced visual encoding for thematic data, and representing temporal data. Moreover, using Google Earth, it provides "query and reasoning" spatial analysis methods (Longley et al. 2001) such as filtering by time, space and attributes. Spatial analyst can select a particular dataset using spatial and temporal predicates in order to refine its analysis. (Wood et al. 2007) also introduce new visual encodings, and it permits to aggregate data by using server-side scripts and geographic layers containing pre-aggregated data.

However, none of the above-mentioned projects integrate such a geobrowser with a data warehouse and/or OLAP solutions, thus heavily limiting the multidimensional analysis.

2.2 Data Warehouse and OLAP Systems

Data warehouse and OLAP systems are intended as business intelligence tools supporting multidimensional analysis of huge amount of data. These are organized according to the *multidimensional model.* This one organizes data along different axes (dimensions) and facts. Facts represent the subjects of analysis. They are described by numerical measures, which are analyzed at different granularities associated to the levels of hierarchies composing the dimensions. The analyst explores the data warehouse through OLAP operators obtaining detailed and aggregated measures according to several dimensions. Classical OLAP operators (Rafanelli, 2003) are: *Slice* which selects of a part of the data warehouse, *Dice* which projects a dimension, *RollUp* which aggregates measures climbing on a dimension hierarchy and *DrillDown* which is the reverse of RollUp.

Spatial OLAP redefines multidimensional model's concepts (Rivest et al. 2001; Malinowski and Zimányi 2007; Bimonte 2008b). It defines spatial dimensions as dimensions where the spatial data is explicitly stored and spatial measures as geometries and/or numerical data directly derived from them.

Figure 1 presents an example of a model of spatio-multidimensional application for the environmental analysis of air pollution of Italian regions. This model, that will be used in Section 4 to describe a case study, presents a spatial dimension "Location" which groups cities into departments and regions. The other dimensions are the temporal dimension ("Time") and the pollutants ("Pollutants"). The measure is the pollution value which is aggregated using the average, the minimum and the maximum. The model allows answering queries like: "What are the pollution values per year, region, and pollutant?" (Bimonte et al. 2007).

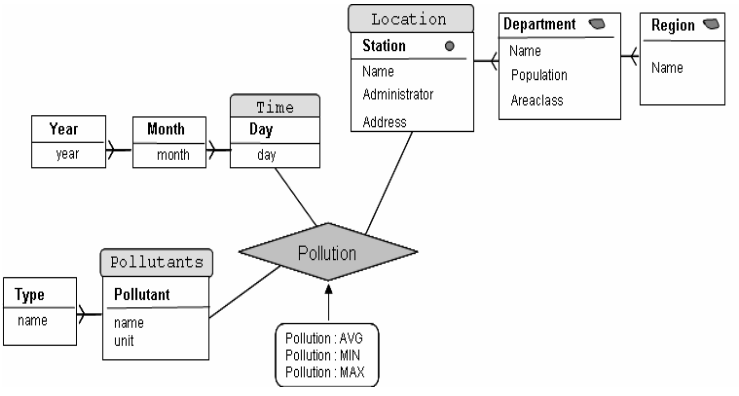
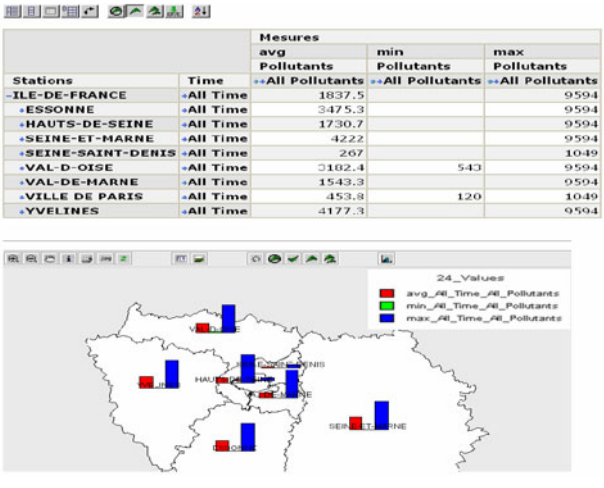


Fig. 1. A SOLAP application (by Bimonte et al. 2007)

SOLAP systems should integrate advanced OLAP and GIS functionalities (Rivest et al. 2005), in a visual interface that fuses pivot table, graphic displays and maps. Maps objects represent spatial dimension members and measures are visualized on maps using classical visual encodings (bars, pie charts, etc.) (Bertin and Bonin 1992).



		Mesures		
		avg	min	max
		Pollutants	Pollutants	Pollutants
Stations	Time	All Pollutants	All Pollutants	All Pollutants
ILE-DE-FRANCE	All Time	1837.5		9594
ESSONNE	All Time	3475.3		9594
HAUTS-DE-SEINE	All Time	1730.7		9594
SEINE-ET-MARNE	All Time	4222		9594
SEINE-SAINT-DENIS	All Time	267		1049
VAL-D-OISE	All Time	3102.4	543	9594
VAL-DE-MARNE	All Time	1543.3		9594
VILLE DE PARIS	All Time	453.8	120	1049
YVELINES	All Time	4177.3		9594

Fig. 2. User interface of a SOLAP tool integrating OLAP and GIS functionalities (Bimonte et al. 2007)

This allows to "see" measures at different spatial granularities, and to visualize spatial relations between spatial dimension members in the studied phenomenon (Rivest et al. 2005). SOLAP tools allow triggering multidimensional operators through the simple interaction the maps. Here, the different components of the SOLAP client are synchronized (Rivest et al. 2005): an action on one component is applied to all other component, for example clicking on the map to perform a drill-down operation corresponds to a drill-down operation in the pivot table. Using the spatio-multidimensional model presented above, we show in figure 2 the visualization of a SOLAP query in the SOLAP tool GeWOlap (Bimonte et al. 2007). Some SOLAP tools have been developed (Rivest et al. 2005; Scotch & Parmanto, 2006; Escribano et al. 2007). These tools provide 2D synchronized interactive maps using classical GIS visual encodings (Bertin and Bonin 2002) with classical OLAP displays. They are based on a server OLAP, which provides all advanced multidimensional analysis functionalities. Moreover, spatial analysis and geographic data management techniques are also provided, such as control management layers, cartographic visualization, geostatistic methods, etc.

Few efforts have been made to enrich SOLAP facilities with basic geovisualization functionalities.

(Bédard et al. 2005) present a SOLAP tool based on the Proclarity OLAP system and the ERSI ArcGIS GIS. The latter allows visualizing multidimensional data with non-interactive 3D maps. (Bedard et al. 2006) add multimedia elements such as photos, videos, etc. to spatial data warehouse data. This approach allows contextualizing SOLAP analysis (Pérez et al. 2007), in the sense that additional and complementary data can be used to better understand and explain SOLAP query results. On the other hand, multimedia elements are statically associated to dimensions and measures during the construction of the spatial data warehouse. Then, existing SOLAP tools do not permit to search and add to the spatial data warehouse geographic and multimedia data from the Web.

Finally, (Hernandez et al. 2005) present a system which adapts a geovisualization system, CommonGis, for the exploration of multidimensional databases.

CommonGis provides multi-criteria and spatial analysis functionalities, and visualization techniques for the analysis of spatio-temporal data using temporal series. The user interface of CommonGis is flexible and interactive, and it provides several advanced cartographic visualization tools such as Space-Time Cube, multi-maps, etc. CommonGis does not use an OLAP server. It is simulated using a relational database modeled using the star or the snow-flake schema. This approach lacks in terms of advanced multidimensional functionalities (Bédard et al. 2005b) such as derived measures, complex hierarchies, etc. Consequently, it limits spatio-multidimensional analysis capabilities.

Finally to the best of our knowledge, no SOLAP tool that provides advanced OLAP functionalities offers geovisualization methods. In particular, no work tries to integrate Geovisualization functionalities of geobrowsers to OLAP systems into a unique framework.

3 The *GooLAP* System

In this section we describe the main features offered by the GooLAP system.

3.1 Main Features

The key idea underlying the design of the user interface of our system was to provide two different visual representations of information, by complementing a widely-adopted tabular rendering of the data with a flexible, web-based Geobrowser, providing a 3D visualization of spatial dimensions data, in order to exploit the advantages of both the textual and visual solutions.

Thanks to this approach, a Decision Maker can perform a multidimensional exploration and analysis of information in an integrated environment, interacting seamlessly on textual and spatial representation of the data. Moreover, he/she can import different layers of spatial data from many freely available repositories, to complement the information coming from the data warehouse.

In the current implementation of this system, we have adopted Google Earth as Geobrowser while, as pivot table, we have developed our own solution, based on JPivot. Therefore, our system can be considered as an OLAP dominant solution, which integrates some geovisualization functionalities.

In the following we first describe these two technologies and then present how they have been arranged together to form a coherent user interface. As for SOLAP operations, our system provides drill-down, roll-up, slice and dice operators using a pivot table and a cube navigator. Moreover it allows triggering drill operators through the simple interaction with the cartographic component of the Geobrowser.

3.2 Mondrian-JPivot

For the OLAP features, we employed two widely-adopted, freely available tools, namely Mondrian (Mondrian 2008) and *JPivot* (JPivot 2008).The former is a software package designed to provide OLAP functionality in an open and extensible framework, on top of a relational database. This is achieved by means of a set of JAVA APIs, which can be used for writing applications, such as a graphical interface, for browsing the multidimensional database. These APIs can also be invoked by JSP/Servlets, within a web environment. Mondrian includes a Calculation layer, that validates and executes MDX (Multidimensional Expressions) queries, and an Aggregation layer that controls data in memory and request data that is not cached. MDX is a standard language to query multidimensional databases, just like SQL for the relational ones. In order to guarantee the greatest

flexibility, to interface the relational data, an XML description of the multidimensional application has to be written.

JPivot is a software package designed for providing a web-based, graphical presentation layer on top of Mondrian. It provides specific JSP tags for easily building powerful graphical interfaces, suited to explore the data warehouse. JPivot provides functionality to modify the visualization of the pivot table and triggers the desired OLAP operators: *drill-down replace, drill-down position, expand-all,* and *drill-through*. The drill-down replace operator enables drilling from one pointed member to its child members in the dimension hierarchy, hiding the parents, whereas the drill-down position shows the parents. The expand-all operator enables drilling from all visible members in the table to child members.

3.3 Google Earth

Google Earth (shortly GE) is a virtual globe, currently freely available for personal use. It is provided in two versions: as a stand-alone application, for PC running on Windows, Mac OS, Linux and FreeBSD, and as a browser plug-in (released on June 2008) for Firefox, Internet Explorer 6 and 7. GE combines satellite raster imagery, with vector maps and layers, in a single and integrated tool, which allows users to interactively fly in 3D from outer space to street level views. It currently incorporates data about almost every place in the world, with a typical resolution of 15 meters per pixel (although most datasets from USA and Europe are available at 1 meter resolution). A very wide set of geographical features (streets, borders, rivers airports, etc.), as well as commercial points of interest (restaurants, bars, lodging, shopping malls, fuel stations, etc…), can be overlaid onto the map. A key characteristic of this tool is the fact that the spatial datasets are not stored on client computers, but they are streamed, upon request, from Google's huge server infrastructure, ensuring fast connections and almost 100% up time. This guarantees that data are always up-to-date. Another remarkable feature implemented by GE is the ability to render a Digital Elevation Model (DEM) of the terrain, mainly using data collected by NASA's Shuttle Radar Topography Mission (SRTM). The internal coordinate system of Google Earth is geographic coordinates (latitude/longitude) on the World Geodetic System of 1984 (WGS84) datum. From an application development point of view, GE offers two complementary ways of interaction: an ad-hoc file format, to present spatial features, and a set of APIs. More specifically, GE supports an XML grammar and file format, named Keyhole Markup Language (or KML) (KML 2008), suited to model one or more spatial features to be displayed. Through KML files, a developer can assign icons and labels to a location on the planet surface, specify camera positions to define views, add basic geometrical shapes, and so on. At the same

time, the GE browser Plug-in offers a set of JavaScript APIs, allowing developers to place and control GE into web pages. In our proposal, we exploited the 3D capabilities provided by the GE browser Plug-in to combine information from the data warehouse with real world infrastructures and geographic features. Moreover, currently many repositories and communities, containing a very broad amount of informative layers are available on the Web and can be freely and effortlessly integrated in any application exploiting Google Earth.

3.4 The User Interface

The user interface we have developed is composed of two main panels, coherently integrated into a dynamic web page: a *Geobrowser* on the left and a *Pivot Table* on the right. The pivot table, representing the OLAP hypercube, is responsible for showing the textual data and providing the hierarchical navigation across the dimensions of the data warehouse. On the other hand, the Geobrowser is responsible for rendering in 3D the spatial information over a geo-referenced satellite image, potentially enhanced by additional informative layers. The Geobrowser is enriched by histograms or pie charts to show data values. A mouse click onto these visual encodings opens a balloon, showing detailed information on the phenomenon. Moreover, by clicking on a geographical area, the system performs a drill-down on that dimension. Below these two panels, there is an area suited to customize the way the information is rendered. In particular, the Decision Maker can choose the visual encoding of the information (in this version he/she can switch between bars and pie charts, as well as their colours), and the set of standard informative layers provided by Google Earth to show onto the map, such as roads, 3D terrain, 3D buildings, and so on. Moreover, there is a form allowing the user to load any KML file available over the web. In the following we present an example of multidimensional exploration and analysis of a spatial data warehouse using the GooLAP interface. This description is supported by a preliminary case study that uses a simulated dataset on pollution values in various Italian regions.

3.4.1 Multidimensional Navigation

In order to perform a multidimensional navigation, a Decision Maker would analyze the average value of the pollution in Italy by year and by pollutants. Of course, the pivot table and the Geobrowser show the same information, in two different fashions. In particular, the latter displays the boundaries of Italy and the pollution value using a bar. Thus, all the information is aggregated (Fig. 3).

Let us suppose that subsequently the Decision Maker is interested in knowing the pollution values for all the Italian regions. Then, by simply clicking on the nation inside the Geobrowser, he/she performs a drill-down operation of the spatial dimension “Location” moving to the “Region” level.

Fig. 3. Pollution value of Italy

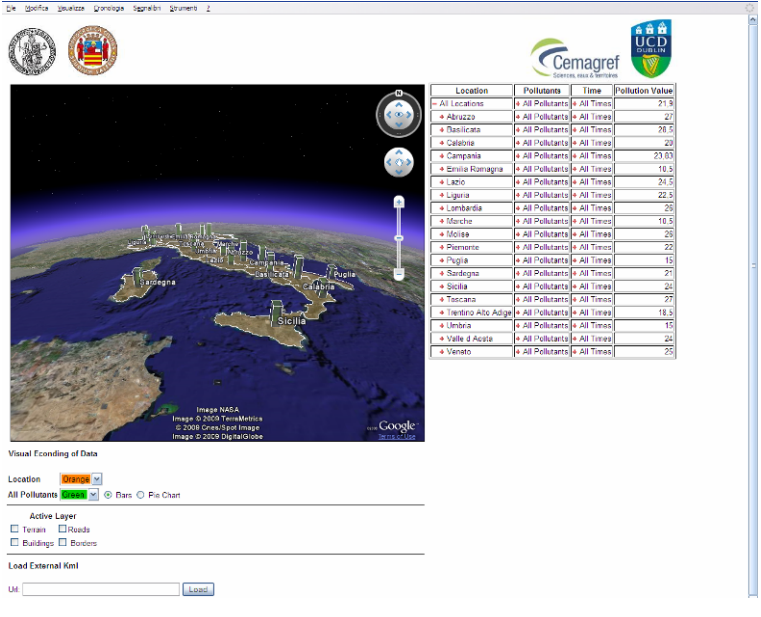


Fig. 4. Drill-Down on Italy: Pollution value per region

As a consequence of this action, GooLAP updates the pivot table and the map at the same time, accordingly (Fig. 4): the pivot table shows the average pollution value for Italy and for all its regions. In the same way, the Geobrowser displays a map highlighting the boundaries of all Italian regions, and one bar is placed onto each region to visually represent the pollution value. As further analysis of the data warehouse, let us suppose that the Decision Maker is then interested in exploring the pollution values per type of pollutant. To this aim, he/she can perform the drill-down operator on the "Pollutants" dimension by simply clicking on the member "All Pollutants" in the Pivot table. GooLAP triggers this OLAP operator and thus both the Pivot table and the Geobrowser are updated accordingly. Now two bars, corresponding to organic and inorganic pollutants, are displayed for each Italian region (**Fig. 5**).

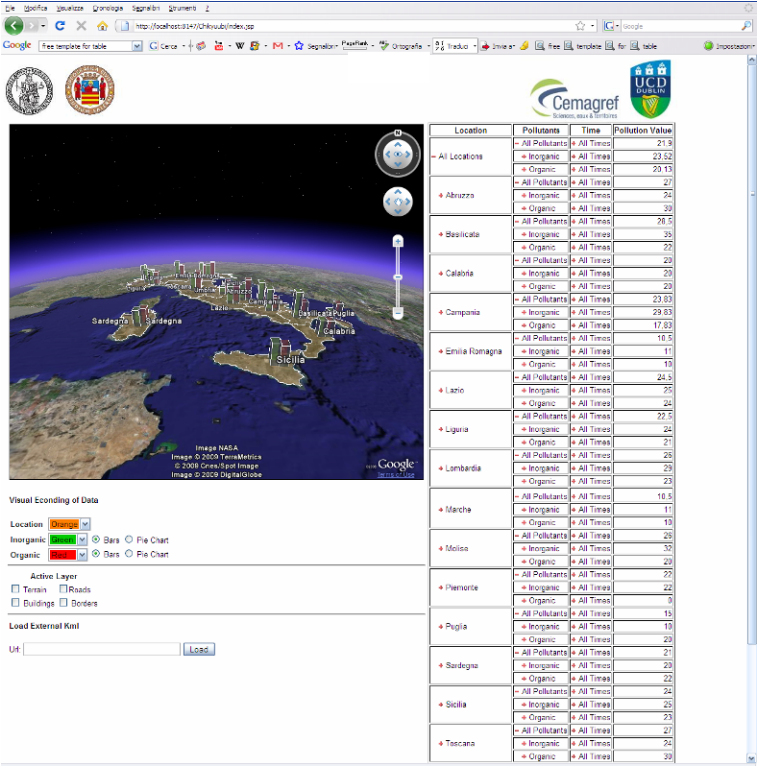


Fig. 5. Drill-down on Pollutants dimension: Pollution value per region and types of pollutants

The two previously examples show navigation in spatial ("Location") and thematic ("Pollutants") dimensions. Pivot table and the Google Earth are synchronized, in the sense that SOLAP operators can be trigged by the simple interaction with them (Rivest et al. 2005). Then, GooLAP links 3D and textual information.

This is fundamental for the exploration and analysis of spatial data (Bleish & Neibeker 2007) as 3D displays could provide new deep insights into geographic data, complementing the textual description that offers a general and abstract visualization of data. Furthermore, let us note that, in our approach, the geobrowser Google Earth is enriched with interaction capabilities for the exploration of the spatial data warehouse. This allows performing analysis using the paradigm "*zoom and filter and the details on-demand*" (Shneiderman 1996). Indeed, Decision Maker can zoom and filter on spatial and alphanumeric data using drill-down and slice operators. Then, details about dimensions members and (aggregated) measures are provided. As further analysis action, the let us suppose that the Decision Maker is interested in two particular districts, for instance Napoli and Avellino, within the Italian *Campania* region. Using the *Cube Navigator* feature provided by the pivot table, she/he can select these two districts and the visualization tool will show only the pollution values for organic and inorganic pollutants associated with those two districts (**Fig. 6**). Thus, in this way, he/she performs a slice operation on the spatial dimension restricting its analysis domain at two geographic districts.

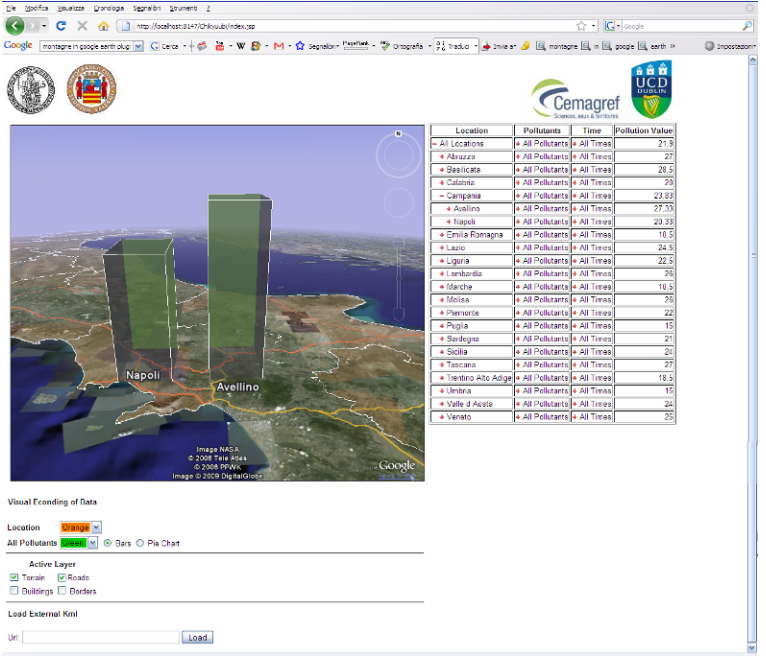


Fig. 6. Slice: pollution value for a specific region

Please note that in **Fig. 6** the road layer has been activated, highlighting the main Italian roads. This could be useful to understand possible correlation of phenomena. For instance, in order to understand pollution environmental phenomenon of the two districts, she/he contextualizes spatial information with external geographic data representing highways road, on the fly. Decision Maker is looking

for possible correlations between pollution and traffic. By this way, user can validate his/her hypothesis, by analyzing multidimensional data according with additional geographic data which improve the spatial analysis context of his/her SOLAP application. Finally, GooLAP integrates zooming capabilities of SOLAP drill operators (Roll-Up and Drill-Down) and of Google Earth. This allows performing multi-scale spatial analysis (Andrienko et al. 2007). Indeed, measures values as well as data of geographic external layers are provided at different granularities: measures are aggregated, and layers information are added or removed in function of the zoom factor used.

3.4.2 Customizing the Visualization

The developed system is able to support also different ways to visually represent data.

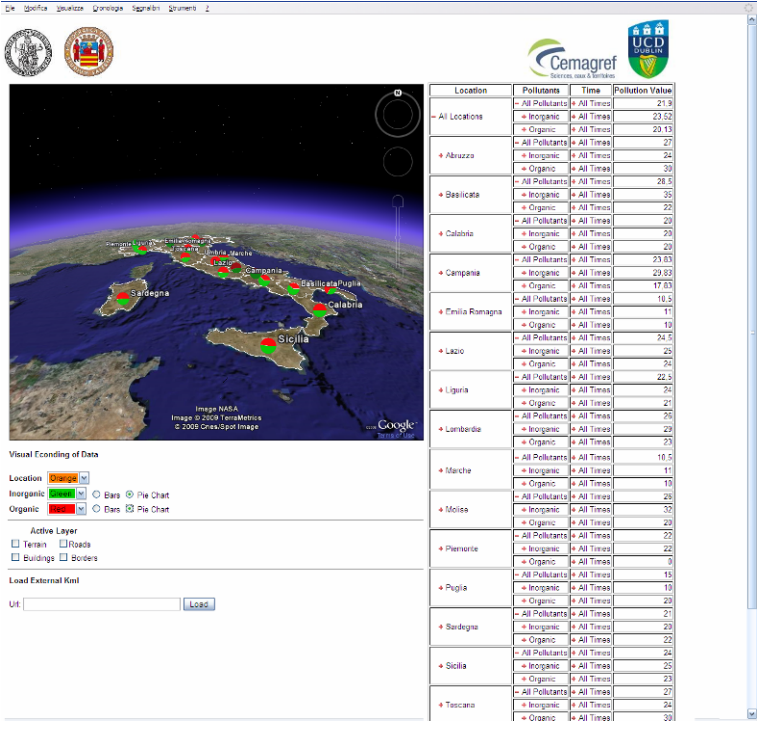


Fig. 7. Thematic map with pie charts

To this aim, we exploited Google Chart, which is a set of APIs to easily create different chart from some data and embed it in a webpage. Indeed, it simply requires sending the data to render and some parameters into an HTTP request to a specific Google Server, and it will return a PNG image of the chart.

By using this technology, the Decision Maker can customize the way information is show. For instance, Fig. 7 represents the same phenomenon of **Fig. 5**, but here the pollution values are rendered using pie charts. This metaphor conveys information in a different fashion, highlighting different kinds of relationships among data. As further example of flexibility of the visual encodings provided by our solution, we choose the *balloons,* i.e. pop-up contextual panels containing charts to visualize values directly on the map. This representation turns out to be useful to surmount map readability problems when too much visual variables are used (Bertin and Bonin 1992). Each balloon can contain a HTML fragment (in **Fig. 8** it is a simple PNG image generated by Google Chart), but of course it could be straightforward to associate multimedia elements, such as photos or videos, to members of spatial dimensions, improving the information contextualization.

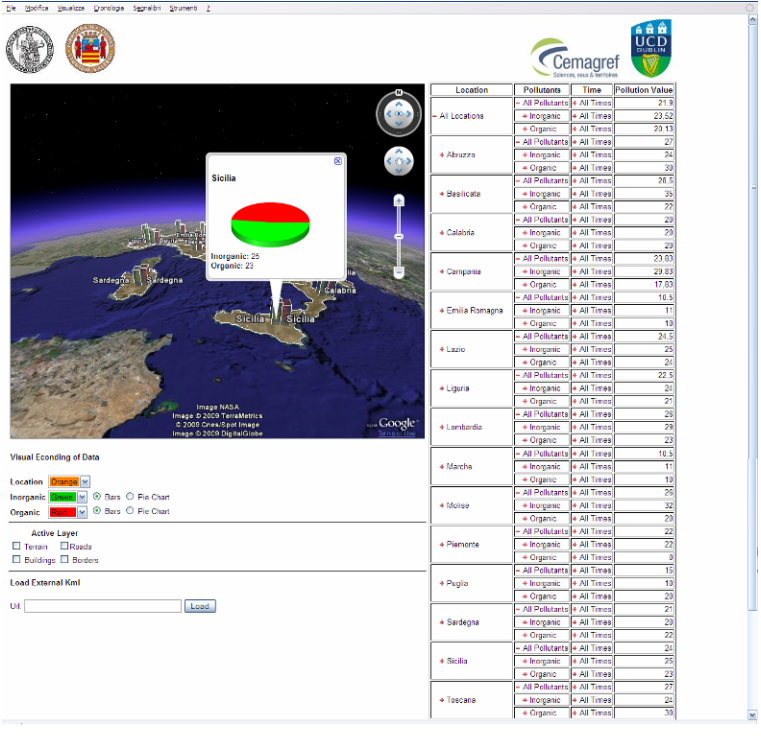


Fig. 8. Balloon representation of a measure

4 System Architecture

In this section we describe the architecture of the web application we have developed, which is responsible to merge together all the different technological solutions we have adopted. The main rationale was to obtain a flexible architecture, where each module is characterized by a loose coupling, in order to permit to replace the adopted technologies with potential novel solutions. For instance, the

system could be easily modified to adopt different Geobrowsers, such as Microsoft Virtual Earth. In particular, the *GooLAP* proposal relies on a three-tier architecture, composed by a data storage at the back end, a business logic layer, and two visualization components at the front end, arranged as shown in **Fig. 9**. To achieve modularization, all the communication among or intra modules are carried out through standard protocols and file formats, such as XML, HTML and KML.

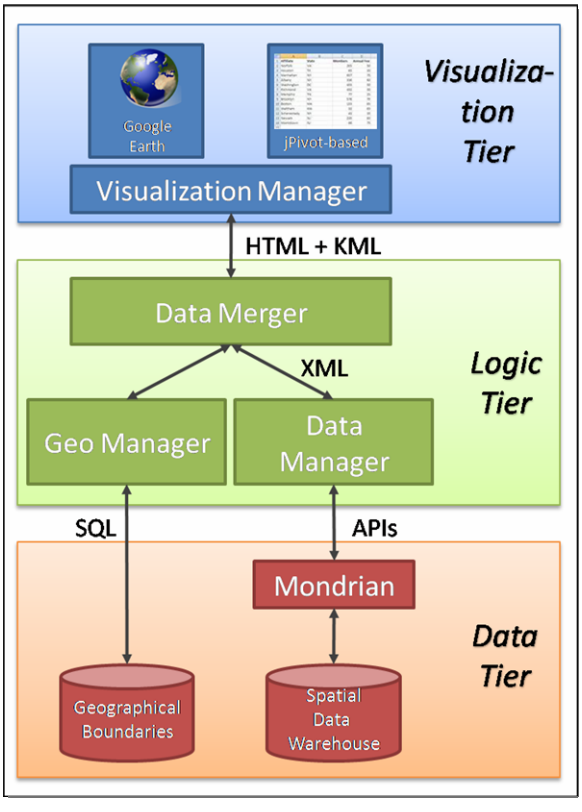


Fig. 9. GooLAP Architecture

The GooLAP architecture is structured as follows:

1. ***Visualization Tier.*** The main goal of this layer is to present both the spatial and textual information to the Decision Maker, and to notify the underlying layers with the OLAP operations he/she performed. This tier encompasses three main components: the Geobrowser, able to render user's selected information onto a map, the Pivot table, to show textual information and the Visual Encoding Bar which allows the user to customize the way data is visualized. They are suitably integrated through some JavaScript code and a Java Servlet, which logically forms the *Visualization Manager* module. In particular the JavaScript module handles events

on the user interface (clicking on the globe or on the pivot table, or changing viewing preferences) and communicates them to the Java Servlet via AJAX. Using AJAX to handle the communication between the user interface and the *Logic Tier* allows us to not reload the page at every request, which could be very time- and resource-consuming, especially for reloading the Google Earth Plug-in. The Java Servlet notifies the *Logic Tier* about the operation the user wants to do (e.g.: roll-up), the data element on which perform the operation and the Visual Encoding of chosen data. As a response the *Visualization Manager* receives data to refresh the Pivot table, and a KML file showing the spatial information according to the user actions.

2. ***Logic Tier*:** This layer is responsible for performing the (S)OLAP operations required by the user, and integrating data coming from the data warehouse with spatial information coming from external data stores in a single data structure, to be fed to the Visualization Layer. This tier is composed of three main modules, the *Data Manager*, the Geo *Manager* and the *Data Merger*.

   a. **Data Manager.** This module interacts with Mondrian (the OLAP server) through the JPivot APIs. These APIs offer an easy way to translate user action (roll-up, drill-down, etc) into MDX query for Mondrian in order to retrieve information in Data Warehouse. The resulting information is formatted into a XML file, suitable for further processing.
   b. **Geo Manager.** This module queries a database containing the geographical boundaries of the regions and states involved in the data analysis.
   c. **Data Merger.** This component is responsible for both handling the communication from the *Visualization Tier*, and merging the information coming from the two other managers to feed the top layer. *Data Merger* receives the XML file representing OLAP data, the spatial information about involved regions and the user visual encoding of data preferences, and then integrates them for the *Visualization Tier*. In particular, in this module we designed and implemented some routines for an on-the-fly generation of KML files, containing the spatial information required by the user, which are shown in the Geobrowser. Spatial information is shown according to user visual encoding of data preferences (e.g.: type of chart: bar or pie and colors). This component also generates HTML code to update the Pivot Table and the Visual Encoding Bar.

   Data Manager and *Data Merger* interact through XML files. In this way they are totally decoupled. Moreover, since they are intended as interfaces, they could be implemented in various ways for different technological solutions (e.g. a different *Data Merger* to work in conjunction with *Microsoft Virtual Earth* instead of *Google Earth*). This allows us to achieve a less tight coupling between the data warehouses and the visualization tools, making them independent and replaceable.

3. ***Data Tier*:** The third layer is responsible to store and retrieve information from both the data warehouse and the database of geographical boundaries. To this aim, an OLAP server is included, *Mondrian*, interacting with a relational DBMS (in our example we adopted *Postgres*). At the same time, also the data store containing data on the geographical boundaries of regions and nations is stored in a relational database. . In order to make the application less time and resource consuming, we also stored information about involved regions geographic coordinates instead of making continuous on-line request to Google Earth geocoder.

5 Conclusions and Future Work

To date there is a strong need for augmenting OLAP tools with geovisualization features, in order to provide Decision Makers with powerful instruments to get insight on hidden patterns, relations, and knowledge in the data stored in spatial data warehouses. In this work, we presented a Web-based SOLAP tool, GooLAP, which integrates the OLAP system Mondrian-JPivot with Google Earth, in order to enhance OLAP analysis with the geovisualization techniques provided by the geobrowser. Therefore, our proposal presents a new SOLAP architecture which integrates OLAP and geovisualization systems.

This proposal enhances existing SOLAP tools in several aspects. GooLAP offers a 3D visualization of spatial dimensions data. It allows integrating information in the data warehouse with geographic information on the Web, in order to improve the geo-spatial context, which is mandatory for Spatial Decision Support Systems. Moreover, different visual encodings of the information can be used for the multidimensional analysis, to better fit the Decision Maker mental model of information. Currently the main technical open issue we are dealing with is the handling of temporal aspects in the proposed framework. At the same time, we are starting to experiment the system on a real data warehouse. Indeed, even if it is recognized that geovisualization tools should be provided to perform SOLAP operations, only an empirical usability study, with real data and a sample of Decision Makers could provide real feedbacks on the effectiveness of the solution.

Acknowledgement. Research presented in this paper was funded by a Strategic Research Cluster grant (07/SRC/I1168) by Science Foundation Ireland under the National Development Plan. The authors gratefully acknowledge this support.

References

Anselin, L.: What Is Special about Spatial Data? Alternative Perspectives on Spatial. Data Analysis, Technical Report 89-94, National Center for Geographic Information and Analysis, Santa Barbara, CA (1989)

Andrienko, N., Andrienko, G., Jankowski, P., et al.: Geovisual analytics for spatial decision support: Setting the research agenda. Int. J. Geographical Inf. Sci. 21(8), 839–857 (2007)

Andrienko, N., Andrienko, G., Wrobel, S.: Visual analytics tools for analysis of movement data. SIGKDD Explor. 9(2), 38–46 (2007)

Bédard, Y.: Spatial OLAP. In: Proc. of 2nd Forum annuel sur la R-D, Géomatique VI: Un monde accessible, Montréal, Canada, November 13-14 (1997)

Bédard, Y., Merrett, T., Han, J.: Fundaments of Spatial Data Warehousing for Geographic Knowledge Discovery. In: Geographic Data Mining and Knowledge Discovery. Taylor & Francis, London (2001)

Bédard, Y., Proulx, M., Rivest, S.: Enrichissement du OLAP pour l'analyse géographique: exemples de réalisation et différentes possibilités technologiques. R. Nouvelles Tech. de l'Inf., Entrepôts de données et l'Analyse enligne, 1–20 (2005)

Bédard, Y., Proulx, M., Rivest, S., et al.: Merging Hypermedia GIS with Spatial On-Line Analytical Processing: Towards Hypermedia SOLAP. In: Geographic Hypermedia: Concepts and Systems. Springer, Berlin (2006)

Bertin, J., Bonin, S.: La graphique et le traitement graphique de l'information. Flammarion, Paris (1992)

Bleisch, S., Nebiker, S.: Connected 2D and 3D visualizations for the interactive exploration of spatial information. In: Proc. of 21th ISPRS Congress, Beijing, China (2008)

Bimonte, S.: On Modelling and Analysis of Geographic Multidimensional Databases. In: Data Warehousing Design and Advanced Engineering Applications: Methods for Complex Construction. Idea Group Publishing, Hershey (2008)

Bimonte, S.: Des entrepôts de données, l'analyse en ligne et l'information géographique. J. of Decis. Syst. 17(4), 463–486

Bimonte, S., Tchounikine, A., Miquel, M.: Spatial OLAP: Open Issues and a Web Based Prototype. In: Proc. of 10th AGILE International Conference on Geographic Information Science, Aalborg, Denmark, May 8-11 (2007)

Compieta, P., Di Martino, S., Bertolotto, M., et al.: Exploratory spatio-temporal data mining and visualization. J. Vis. Lang. and Comp. 18(3), 255–272 (2007)

Caron, P.: Étude du potentiel de OLAP pour supporter l'analyse spatio-temporelle. Université Laval, Laval (1998)

Escribano, A., Gomez, L., Kuijpers, B., et al.: Piet: a GIS-OLAP implementation. In: Proc. of the ACM 10th International Workshop on Data Warehousing and OLAP, pp. 73–80. ACM Press, New York (2007)

Franklin, C.: An Introduction to Geographic Information Systems: Linking Maps to databases. In: Database (1992)

Girardin, F., Calabrese, F., Dal Fiore, F., et al.: Digital footprinting: Uncovering tourists with user-generated content. IEEE Pervasive Comp. 7(4), 36–43 (2008)

Goodchild, M., Densham, P.: Spatial Decision Support Systems. Scientific Report for the Specialist Meeting, TR-90-5. National Center for Geographic Information and Analysis, Santa Barbara, California, USA (1990)

Hägerstrand, T.: What About People in Regional Science? Regional Science Association 24, 7–21 (1970)

Hernandez, V., Voss, A., Göhring, W., et al.: Sustainable decision support by the use of multi-level and multi-criteria spatial analysis on the Nicaragua Development Gateway. In: Proc. of From pharaohs to geoinformatics Proceedings of FIG Working Week 2005 and 8th International Conference on the Global Spatial Data Infrastructure, Le Caire, Egypte (2005)

Inmon, W.: Building the Data Warehouse. Wiley, New York (1996)

JPivot, The JPivot Project web site (2008), `http://jpivot.sourceforge.net/` (last visited on December 07, 2008)

KML, The KML file format specifications (2008), http://code.google.com/intl/en/apis/kml/documentation/ (last visited on December 07, 2008)

Keim, D., Mansmann, F., Schneidewind, J., et al.: Challenges in Visual Data Analysis. In: Proc. of Information Visualization Symposium, pp. 9–16. IEEE, New York (2006)

Longley, P., Goodchild, M., Maguire, D., Rhind, D.: Geographic Information Systems and Science. John Wiley & Sons, New York (2001)

MacEachren, A., Kraak, M.: Research challenges in geovisualization, Cartography and Geographic. Inf. Syst. 6(1), 3–12 (2001)

MacEachren, A., Gahegan, M., Pike, W., et al.: Geovisualization for Knowledge Construction and Decision Support. IEEE Comput. graphics and appl. 24(1), 13–17 (2004)

Malinowski, E., Zimányi, E.: Logical Representation of a Conceptual Model for Spatial Data Warehouses. In: Geoinformatica. Springer, Berlin (2007)

Mitchell, A.: The ESRI Guide to GIS Analysis. Spatial Measurements and Statistics, vol. 2. ESRI Press, Redlands (2005)

Mondrian, The Mondrian Project web site (2008), http://mondrian.pentaho.org/ (last visited on December 07, 2008)

Pérez, J., Lavori, R., Cabo, M., et al.: R-Cubes: OLAP Cubes Contextualized with Documents. In: Proc. of International Conference on Data Engineering, pp. 1477–1478. IEEE, New York (2007)

Rafanelli, M.: Operators for Multidimensional Aggregate Data. In: Multidimensional databases: problems and solutions. IGI Publishing, Hershey (2003)

Rivest, S., Bédard, Y., Marchand, P.: Towards better support for spatial decision-making: defining the characteristics of Spatial On-Line Analytical Processing. J. of the Canadian Inst. of Geomatics 55(4), 539–555 (2001)

Rivest, S., Bédard, Y., Proulx, M., et al.: SOLAP: Merging Business Intelligence with Technology for Interactive Spatio-Temporal Exploration and Analysis of Data. J. Int. Soc. for Photogrammetry and Remote Sensing 60(1), 17–33 (2005)

Scotch, M., Parmanto, B.: Development of SOVAT: a numerical-spatial decision support system for community health assessment research. Int. J. Med. Infor. 34(10), 771–784 (2006)

Shneiderman, B.: The eyes have it: A task by data type taxonomy for information visualizations. In: Proc. of IEEE Symposium on Visual Languages. IEEE, New York (1996)

Slingsby, A., Dykes, J., Wood, J., et al.: The Visual Exploration of Insurance Data in Google Earth. In: Proc. of Geographical Information Systems Research, UK, Manchester, pp. 24–32 (2007)

Stolte, C., Tang, D., Hanrahan, P.: Polaris: A System for Query, Analysis and Visualization of Multi-dimensional Relational Databases. IEEE Trans. on Vis. and Comp. Graphics 8(1), 52–65 (2002)

Tobler, W.: A computer movie simulating urban growth in the Detroit region. Econ. Geogr. 46, 234–240 (1970)

Zhe-Ping, S.: Geovisualization of Disaster Potential Evaluation on Hillside Areas Around Taipei. In: Proc. of Asian Association of Remote Sensing (2007)

Wood, J., Dykes, J., Slingsby, A., et al.: Interactive visual exploration of a large spatio-temporal data set: reflections on a geovisualization mashup. IEEE Trans. on Vis. and Comp. Graphics 13(6), 1176–1183 (2007)

Nonlinear Black-Box Models for Short-Term Forecasting of Air Temperature in the Town of Palermo

Maurizio Cellura, Simona Culotta, Valerio Lo Brano, and Antonino Marvuglia

Dipartimento di Ricerche Energetiche ed Ambientali (DREAM),
Università degli Studi di Palermo, viale delle scienze, edificio 9 -90128
Palermo, Italy

Abstract. Weather data are crucial to correctly design buildings and their heating and cooling systems and to assess their energy performances. In the intensely urbanized towns the effect of climatic parameters is further emphasized by the Urban Heat Island (UHI) phenomenon, known as the increase in the air temperature of urban areas, compared to the one measured in the extra-urban areas. The analysis of the heat island needs detailed local climate data which can be collected only by a dedicated weather monitoring system. The Department of Energy and Environmental Researches of the University of Palermo (Italy) has built up a weather monitoring system that works 24 hours per day and makes data available in real-time at the web site: http://www.dream.unipa.it/meteo. The data collected by the system have been used to implement a set of nonlinear black-box models aiming to obtain short-term forecasts of the air temperature and map them over the monitored area. By using the data recorded during the 2008 summer, the daily profiles of the hourly average temperature have been plotted for each weather station of the monitoring system, thus clearly highlighting the temperature differences between the urban and extra-urban area and the average *intensity* of the UHI of Palermo.

Keywords: artificial neural networks, nonlinear black-box models, MLP, temperature short-term forecasting; Urban Heat Island.

1 Introduction

The knowledge of the climatic parameters like hourly mean values of relative humidity, air temperature and wind velocity and direction are useful in the thermal building simulation, heating and cooling load calculations to decide the correct sizing of an air conditioning system and to optimize the design of many solar energy systems (Beccali et al. 2007; Ardente et al. 2005).

The UNI standards (UNI 10349 1994) provide Heating, Ventilation and Air Conditioning (HVAC) plants designers with average climate data, but their temporal resolution is inadequate to appreciate the real dynamics of weather variations.

B. Murgante et al. (Eds.): Geocomputation, Sustainability & Environ. Planning, SCI 348, pp. 183–204.
springerlink.com

In fact they provide the average climate year of a geographic location and so they are not suitable for time series studies and for the investigation of short-term weather sensitivity of buildings energy consumption.

Nowadays, the influence of microclimatic conditions on the energy behavior of buildings draws the attention of many researchers (Lopes et al. 2001). Concerning the civil sector, in addition to the multiplicity of human factors affecting the energy demand, several studies have explained the influence of weather variables on energy consumption and especially on electricity demand (Beccali et al. 2007; Beccali et al. 2004).

The enhanced urbanization, occurring both in developed and developing countries, led to the appearance of the Urban Heat Island (UHI) phenomenon, notably known for the air temperature increase in the wide urban areas, compared to the conditions measured at the meteorological stations located in the extra-urban areas (Wong and Yu 2005).

The main distinctive features of the heat islands are (Gartland 2008):

1. they are often warmest, compared to rural surroundings, after the sunset, and coolest after the sunrise. Urban air in the "canopy layer", below the top of the trees and buildings, can be as much as 10°C warmer than the air in the rural areas (Santamouris et al. 2001);
2. air temperatures are driven by the heating of urban surfaces, since many man-made surfaces absorb more of the sun's heat with respect to the natural vegetation;
3. these differences in air and surface temperatures are enhanced when the wind is low and the sky is clear;
4. areas with the least vegetation and greatest development tend to be hottest, and heat islands tend to become more intense as cities grow larger;
5. heat islands also display warmer air in the "boundary layer", a layer of air up to 2000 meters high.

The presence of UHI has the effect of enhancing the cooling load in commercial and residential buildings (Oke et al. 1991; Papadopoulos 2001). In fact, especially in the Mediterranean countries, peaks in the electricity demand occur nowadays more frequently during the summer period than during the winter.

An accurate analysis of the spatial and temporal evolution of the UHI needs a detailed collection of local climate data which can be accomplished only through an efficient weather monitoring system. In this way, the comparison between the weather data collected within the urban area and those related to the rural suburbs nearby the city allows the estimation of the "heat island intensity" ΔT_{u-r}, i.e. the temperature difference between a central urban site and a rural site.

The Department of Energy and Environmental Researches (DREAM – Dipartimento di Ricerche Energetiche ed Ambientali) of the University of Palermo has built up a network of weather stations, whose positions are showed in Fig. 1.

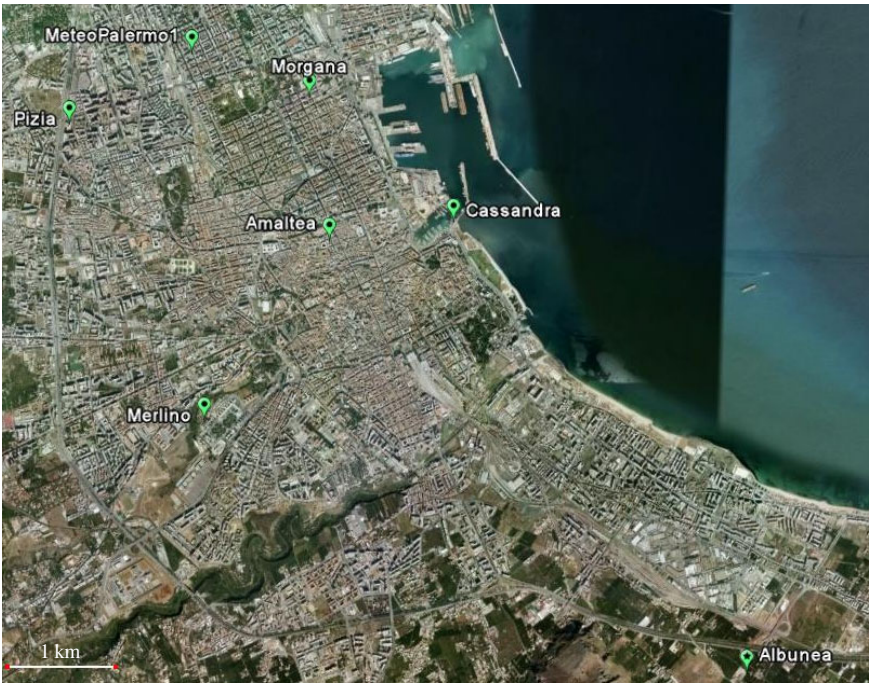


Fig. 1. Locations of the weather stations constituting the weather monitoring network created and operated by DREAM. The map was created by using Google Maps API.

In particular, the network includes a weather station located in the proximity of the sea (named Cassandra), one located within the core part of the town (Amaltea), two stations closer to the outskirts (Merlino and Pizia), two in an intermediate position (MeteoPalermo1 and Morgana) and a weather station located in a rural suburb of the town (Albunea). Concerning the station named MeteoPalermo1 it has to be said that it has been integrated in the DREAM network, but it is operated and owned by a private association called MeteoPalermo (http://www.meteopalermo.it). The weather stations are equipped with:

- Air Temperature sensor;
- Barometer;
- Hygrometer (to measure humidity and dew point);
- Anemometer and weather vane (to measure wind speed and direction);
- Rain gauge;
- Radiometer[1] (to measure solar global radiation and ultraviolet component).

The monitoring system works 24 hours per day and data are immediately available at the web site: http://www.dream.unipa.it/meteo. Data plots are automatically generated every thirty minutes. In a special section of the web site, the actual

[1] "MeteoPalermo1" weather station is not equipped with this sensor.

weather situation and the fluctuations (daily, monthly and yearly) of the main weather variables are showed, along with several statistical elaborations. This system was designed to provide an efficient way to tackle significant amounts of data and information, subject to the following specifications: 1. Data acquisition occurs autonomously and automatically, 2. Data management, processing and presentation are easy and automatic, 3. Users are able to access all data and information remotely through a web browser.

The monitoring system has currently been in operation for over two years. By using the data collected in this period the authors implemented a set of nonlinear black-box models for a short-term temperature forecasting. The temperatures related to a generic summer week were used to validate the model and assess its forecasting performances. Some considerations have also been made about the UHI in the town of Palermo.

2 Data Retrieval, Pre-processing, Storage and Analysis

Every 30 minutes each weather station automatically generates an ASCII file containing the last 336 collected data and immediately transfers it (via GSM) to a MS Windows PC located at the DREAM building, in which proprietary software is installed. Afterwards, the Linux server of DREAM connects to the shared folder of this PC where the file is stored and copies it into a local folder. The procedure is automated by a bash script and is repeated for each weather station. Once the ASCII file has been acquired by the server, it has to be modified and re-formatted in order to allow the updating of the MySQL server where all data are recorded. This procedure is realized through a PERL script.

The text file thus obtained is ready to be read by the database server. The chosen Data Base Management System (DBMS) is MySQL, the most popular among the open source DBMSs. This database server allows an excellent integration among Linux operation system installed in the server, Apache web server used to publish web pages and PHP scripting language used for the creation of dynamic web pages. The complete set of the above mentioned software for management and online publication of large amounts of data is also known as LAMP package (Linux-Apache-MySQL-PHP). The database is updated by using an SQL command. Every command is executed by CRON.

In this way the MySQL database is automatically updated every thirty minutes and contains tables of data related to the whole set of weather stations. The database is thus readable and interactively available over TCP/IP network.

Two kinds of queries can be run by the web visitors: temporal queries and spatial queries. The time query allows, once chosen the station(s) of interest, to extract data related to a specific time interval that could even span over the whole period in which the station has been operating. The spatial query allows the investigation about the occurrence of a particular meteorological event in the sites of one or more weather stations.

By exploiting the flexibility of the LAMP platform it was possible to create a set of dynamic html pages, which allow the visualization of the current weather

situation and of the statistical analysis accomplished with data, both in graphical and table format. All data are automatically published on the web page http://www.dream.unipa.it/meteo, which displays the locations where the weather stations are installed by exploiting Google Maps API utilities.

In the section related to the current weather situation, two different data representations are available:

- a numerical representation that automatically shows the current values of the weather parameters and displays the maximum and minimum value reached by them in the last 24 hours, in the last month and since the beginning of the current year;
- a graphical representation that shows the evolution of the main weather parameters in the last 24 hours.

These pieces of information are continuously updated by the system in an automatic way. In the specific section of the website devoted to the statistical elaborations[2] these are reported in a graphical form as well. The statistical elaborations concern temperature, rainfall, solar radiation, wind speed and direction. A more detailed description of the content of the statistical section of the DREAM website can be found in (Beccali et al. 2008).

3 Neural Networks and Temperature Forecasting

Accurate forecasting of hourly air temperatures can be useful in many practical situations. For example, the knowledge of variation in the air temperature has a considerable importance for the hourly energy consumption and cooling load estimation in buildings (Mihalakakou et al. 2002; Ben-Nakhi and Mahmoud 2004), room air temperature prediction (Yang and Kim 2004), urban outdoor comfort estimation (ASCE 2004).

In all the above mentioned applications one of the approaches very often used by researchers is represented by the Artificial Neural Networks (ANNs). ANN based techniques can be used as an alternative method in the analysis of complex and/or ill-defined engineering problems. ANN does not require the formulation of a physical relation describing the concerned problem. Moreover, one of the important properties of ANN-based models is their capability in detecting the underlying structure existing in a complex natural system and adapting themselves to recurrent changes.

The ability of the ANN to approximate large classes of non-linear functions with sufficient accuracy makes them very appropriate for the representation of dynamic non-linear systems. The fact that static and dynamic back propagation can be used for the adjustment of their parameters also makes them suitable for systems identification (Chen and Billings 1992).

González and Zamarreño (2002) used short-term hourly temperature forecaster based on a state space neural network (ssNN) in building electric load forecasting. The authors used two different real databases and compared the forecasting

[2] The access to this section is protected by password.

performances of their model with those of a linear ARMAX predictor and a radial basis function neural network (RBFNN) trained with the same data. In the best case, the mean absolute error (MAE) obtained with the ARMAX model, the RBFNN and the ssNN were 0.52 °C, 0.45 °C and 0.42 °C, respectively. An interesting application is also described by Abdel-Aal (2004), who developed 24 models to forecast the temperature at the next hour (*h*) during day (*d*) using the full hourly temperature data on day (*d-1*) together with all available hourly temperatures on day (*d*) up to the hour (*h-1*). The MAE and mean absolute percentage error (MAPE) obtained over the evaluation year were respectively 0.58 °C and 2.14% if the minimum and maximum value of the temperature measured during the forecasting day (up to the hour (*h-1*)) were also used as model inputs; 0.59 °C and 2.18% if extreme temperatures of day (*d*) were not used. The paper also shows the results (MAE≈1.73 °C and MAPE=6.42%) obtained with a next-day, day (*d*), forecasting obtained by sequential use of the next-hour model. As it was normal to expect, in this case the authors obtained higher error values, because the next-hour models are simpler and easier to apply than the corresponding next-day hourly models, which are obviously affected by the accumulation of forecasting errors. Another next-hour RBFNN forecaster is presented in Lanza and Cosme (2001), where a MAE≈0.44 °C is obtained without using extreme temperature forecasts. Finally, in Hippert et al. (2000) a hybrid model based on a linear ARMA predictor and an ANN was applied to temperature forecasting, obtaining a MAPE=2.66% on a five days test period. The minimum MAPE obtained with a hybrid AR+ANN was 2.82%.

In this paper, the authors treated the short-term temperature forecasting problem as the identification of a linear time-varying system and tackled it by the utilization of nonlinear black-box models[3].

The underlying hypothesis is that the system is signified by transfer function characterizations. As the behaviour of the system changes, the ANN model developed keeps track of the changes in the characteristics and parameters of the system. Thus, at any instant of time, it correctly simulates the given time-varying system, despite any significant change in its properties.

The approximation capability of the neural network is used to identify the relationship between system variables and parameters.

4 Nonlinearity Detection in Time Series

The application of a nonlinear approach to model a time series should be motivated by an accurate analysis of the data structure in order to efficiently detect an underlying nonlinear component that would not be accounted for with a linear model. In fact, the presence of nonlinear components in a system does not prove

[3] A short description of the models applied will be provided further in the paper. An interesting overview of the nonlinear models used for system identification is contained in Sjöberg et al. (1995).

that this nonlinearity is also reflected in a specific signal we measure from that system (Schreiber and Schmitz 2000). Several methods for detecting nonlinearity and determinism in experimental time series have been proposed in the scientific literature. One of the most popular methods to address such a topic is the method of surrogate data (Schreiber and Schmitz 2000; Theiler et al. 1992).

Surrogate data are artificially generated data which mimic some of the properties of the data under study, but not the property which we are testing for. In the case of testing for nonlinearity, the surrogate data should have the same spectrum and autocorrelation function ("linear properties") as the original data under study, however, they are generated as realizations of a *linear stochastic process*.

The basic idea in the surrogate-data based nonlinearity test is to compute a nonlinear statistic from the studied data and from a set of realizations of the above mentioned linear stochastic process. If the computed statistics for the original data is significantly different from the values obtained for the surrogate set, it is possible to infer that the data at hand are not realizations of a linear process; otherwise the null hypothesis (that a linear model fully explains the data) is accepted and the data can be further analyzed and characterized by using linear models.

In the case study dealt with in this paper, the novel approach proposed in (Gautama et al., 2004), called Delay Vector Variance (DVV) method, was applied to the time series recorded in each weather station. It examines the predictability of a time series by virtue of the observation of the variability of the targets. In practice a time series is characterized on the basis of its predictability and comparing the results to those obtained for the linear surrogates. Due to the standardisation within the algorithm, the method is robust to the presence of noise. Given a time series $\boldsymbol{y}$, its *time-delay embedding representation*

$$\boldsymbol{y} = \left\{ \boldsymbol{y}\left(k\right) \middle| k = 1, \ldots, N \right\} \tag{1}$$

is based upon a set of *N Delay Vectors* (DVs) defined as:

$$\boldsymbol{y}\left(k\right) = \left[y_{k-m}, \ldots, y_{k-1} \right] \tag{2}$$

The generic vector $\boldsymbol{y}\left(k\right)$ is therefore a vector containing *m* consecutive time samples, where the scalar *m* is called *embedding dimension*. Every DV $\boldsymbol{y}\left(k\right)$ has a corresponding *target*, namely the following sample, y_k.

For a given embedding dimension *m*, the mean target variance, σ^{*^2} is computed over all sets Ω_k generated by grouping those DVs that lay within a certain distance from $\boldsymbol{y}\left(k\right)$, which is varied in a standardized manner with respect to the distribution of pairwise distances between DVs. The DVV method can be summarised as follows for a given embedding dimension *m*:

1. The mean, μ_d, and standard deviation, σ_d, are computed over all pairwise distances between DVs, $\left\| \boldsymbol{y}(i) - \boldsymbol{y}(j) \right\| \ (i \neq j)$
2. The sets Ω_k are generated such that $\Omega_k = \left\{ \boldsymbol{y}(i) \ \middle| \ \left\| \boldsymbol{y}(k) - \boldsymbol{y}(i) \right\| \leq \tau_d \right\}$, i.e., sets which consist of all DVs that lye closer to $\boldsymbol{y}(k)$ than a certain distance τ_d, taken from the interval $\left[\min\{0, \mu_d - n_d \sigma_d\}; \mu_d + n_d \sigma_d \right]$, e.g., uniformly spaced, where n_d is a parameter controlling the span over which to perform the DVV analysis.
3. For every set Ω_k, the variance of the corresponding targets, σ_k^2, is computed. The average over all sets Ω_k, normalised by the variance of the time series, σ_x^2, yields the measure of unpredictability, σ^{*^2}:

$$\sigma^{*^2} = \frac{(1/N)\sum_{k=1}^{N} \sigma_k^2}{\sigma_x^2} \tag{3}$$

By using the DVV method, the linear or nonlinear nature of the time series is examined by performing DVV analyses on both the original and a number of surrogate time series, using the optimal embedding dimension of the original time series. A variance measurement is considered *valid*, only if the set Ω_k contains at least 30 DVs. As a result of the standardisation of the distance axis, the resulting "DVVplots" are easy-to-interpret. The presence of a strong deterministic component will lead to small target variances for small spans. At the extreme right, the DVV plots smoothly converge to unity, since for maximum spans, all DVs belong to the same set, and the variance of the targets is equal to the variance of the time series. If this is not the case, the span parameter n_d should be increased.

Due to the standardisation of the distance axis, the DVV plots can be conveniently combined in a *scatter diagram*, where the horizontal axis corresponds to the DVV plot of the original time series, and the vertical to that of the surrogate time series. If the surrogate time series yield DVV plots similar to that of the original time series, the "DVV scatter diagram" coincides with the bisector line, and the original time series is likely to be linear. The deviation from the bisector line is, consequently, an indication of nonlinearity.

As an example, in Fig. 2 it is showed the result of the DVV scatter diagram for the multivariate time series (humidity, dew point, wind speed, atmospheric pressure, solar radiation, temperature) recorded at Amaltea station.

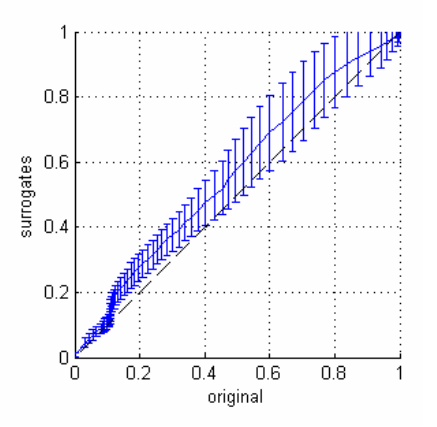


Fig. 2. DVV *scatter diagram* for the multivariate time series recorded at Amaltea station. The error bars showing the confidence level of the data are also plotted.

The iterative Amplitude Adjusted Fourier Transform (iAAFT) method (Theiler et al., 1992) was applied to generate the surrogate. This method yields time series with amplitude spectra identical to that of the original time series, and approximately identical signal distributions (retaining approximately the same autocorrelation structure as the original signal). From the figure it is possible to see that the DVV scatter diagram does not coincide with the bisector line, thus revealing a nonlinear underlying structure of the time series.

5 Selection of the Embedding Parameters

In order to understand and predict future values of nonlinear time series, it is important to analyze them to extract knowledge of the underlying dynamical system. Time series are generally sequences of measurements of one or more observable variables of an underlying dynamical system, whose state changes with time as a function of its current *state vector*. The notion of *state* plays a very important role in the mathematical formulation of a dynamical system. The state of a dynamical system is formally defined as a set of quantities that summarizes all the information about the past behaviour of the system that is needed to uniquely describe its future behaviour, except for the purely external effect arising from the applied input (excitation) (Haykin, 1999).

The *embedding theorem* developed by Takens (Takens, 1981) guarantees that, with a sufficiently long time series $\mathbf{y}(t)$, we can obtain the following f which has one-to-one correspondence to the original system and recreates its underlying dynamics:

$$\mathbf{v}(t+1) = f\left[\mathbf{v}(t)\right] \tag{4}$$

$$\mathbf{v}(t) = \left(y(t), y(t+\tau), \ldots, y\left(t+(m-1)\tau\right)\right) \tag{5}$$

where *f* denotes reconstructed dynamical system, **v** is the *time delay coordinate vector*, *m* is the *embedding dimension* and τ is called *delay time*.

Unfortunately embedding theorem does not provide any values for the *embedding parameters*, i.e. delay time τ and embedding dimension *m*.

Often, more important than the selection of the network architecture is the optimal selection of the embedding parameters. A wrong choice of the *lag space*, i.e. the number of delayed signals used as regressors of the nonlinear black-box model, can influence the model performances in a very negative way. For instance, if the temporal span of $(m \cdot \tau)$ is too small, the signal variation within the delay vector is mostly governed by noise and either *m* or τ should be increased. Notice that if τ were too small to cover the minimal time span needed to capture the dynamics of the signal, the number of regressors would become rather large, resulting in an increased complexity of training. In turn, if τ is greater than optimal, the nature of the resulting model becomes too discrete, resulting in a failure of the model to capture the underlying dynamics of the signal.

For this reason, the estimation of the optimal embedding parameters for reconstruction of the nonlinear dynamics has been studied as an important problem in the scientific literature. There are several heuristic techniques for estimating the embedding parameters (Abarbanel 1996; Alligood 1997; Kantz and Schreiber 1997).

The selection of the optimal embedding parameters to be used in the case study described in this paper was accomplished by applying the procedure described in (Gautama et al., 2003) and using the related Matlab toolbox (Gautama 2007). The approach described in Gautama et al. (2003) is based on the minimization of a "cost" function depending on the two variables *m* and τ. The coordinates of the minimum of this cost function yield the optimal embedding parameters, which best reflect the dynamics of the underlying signal production system.

In order to make the method robust with respect to dimensionality of the time series under study, the cost function is standardised with respect to an ensemble of surrogates of signal *y*.

The surrogate data are generated using the iAAFT method. This way, the serial correlations are present in both the original and the surrogate time series.

In the case study tackled in this paper, the above described method was applied to the temperature time series recorded in every weather station and the corresponding optimal embedding parameters were determined.

6 Black-Box Structures for Time Series Prediction

According to Ljung (1999) a single-output system with input *u* and output *y* is called linear if it is possible to describe it by a model that takes the form:

$$y(t) = G\left(q^{-1}\right)u(t) + H\left(q^{-1}\right)e(t) \tag{6}$$

where G and H are transfer functions in the time delay operator, q^{-1}; $e(t)$ is a white noise signal that is independent of past inputs; $u(t)$ is an exogenous signal.

The delay operator works on a signal in the following way:

$$q^{-d}x(t) = x(t-d) \tag{7}$$

where d is a multiple of the sampling period.

All the common models used can be summarized by the general equation:

$$A\left(q^{-1}\right)y(t) = \frac{B\left(q^{-1}\right)}{F\left(q^{-1}\right)}u(t) + \frac{C\left(q^{-1}\right)}{D\left(q^{-1}\right)}e(t) \tag{8}$$

where

$$A\left(q^{-1}\right) = 1 + a_1 q^{-1} + \ldots a_n q^{-n}$$
$$B\left(q^{-1}\right) = b_0 + b_1 q^{-1} + \ldots b_m q^{-m}$$
$$C\left(q^{-1}\right) = 1 + c_1 q^{-1} + \ldots c_k q^{-k}$$
$$D\left(q^{-1}\right) = 1 + d_1 q^{-1} + \ldots d_l q^{-l}$$
$$F\left(q^{-1}\right) = 1 + f_1 q^{-1} + \ldots f_r q^{-r}$$

According to the different values of the polynomials in Eq. (8), the following models are defined: Finite Impulse Response (FIR) model (A=F=C=D=1); *Box-Jenkis* (BJ) model (A=1); AutoRegressive Moving Average with eXogenous inputs (ARMAX) model (F=D=1); *Output-Error* (OE) model (A=C=D=1) and AutoRegressive with eXogenous inputs (ARX) model (F=C=D=1).

The predictor associated with Eq. (8) can be expressed in "pseudo-linear" regression form as (Ljung and Söderström 1983):

$$\hat{y}\left(t|\theta\right) = \theta^T \varphi\left(t,\theta\right) \tag{9}$$

where θ is the *parameter vector* and φ is the *regression vector*.

For example, in the ARX model the *regression* and *parameter vectors* are defined respectively by:

$$\varphi(t) = \left[y(t-1) \ldots y(t-n_a), u(t-n_k) \ldots u(t-n_k-n_b) \right]^T \quad (10)$$

$$\theta = \left[-a_1, \ldots -a_{n_a}, b_0, \ldots b_{n_b} \right]^T \quad (11)$$

For the ARMAX model thy take the form:

$$\varphi(t|\theta) = \begin{bmatrix} y(t-1) \ldots y(t-n_a), u(t-n_k) \ldots u(t-n_k-n_b), \\ \varepsilon(t-1,\theta), \ldots \varepsilon(t-n_c,\theta) \end{bmatrix}^T \quad (12)$$

$$\theta = \left[-a_1, \ldots -a_{n_a}, b_0, \ldots b_{n_b}, c_1, \ldots c_{n_c} \right]^T \quad (13)$$

where $\varepsilon(t,\theta) = y(t) - \hat{y}(t|\theta)$ represents the *prediction error* or *residual*.

When widening the focus to also include identification of *nonlinear* dynamic systems, the problem of selecting model structures becomes more difficult. By exploiting the typical capability of MultiLayer Perceptrons (MLP) to learn *nonlinear* relationships from a set of data, nonlinear extensions of the most common linear structures for time series prediction have been created and successfully used in literature (Sjöberg et al. 1995).

Nonlinear counterparts to the linear time series forecasting model structures are thus obtained by:

$$y(t) = g\left[\varphi(t,\theta), \theta \right] + e(t) \quad (14)$$

or on predictor form:

$$\hat{y}(t|\theta) = g\left[\varphi(t,\theta), \theta \right] \quad (15)$$

where $\varphi(t,\theta)$ is again the regression vector, while θ is the vector containing the adjustable parameters in the neural network known as *weights*. g is the function realized by the neural network and it is assumed to have a feed-forward structure. Depending on the choice of the regression vector, different nonlinear model structures can be obtained. If the regression vector is selected as for ARX models, the model structure is called NNARX (the acronym of Neural Network ARX). Likewise, NNFIR, NNARMAX, NNOE structures there exist.

In the case study described in this paper, the models yielding the best results are the NNARX and NNARMAX models. The structures of NNARX and NNARMAX models are depicted in Fig. 3.

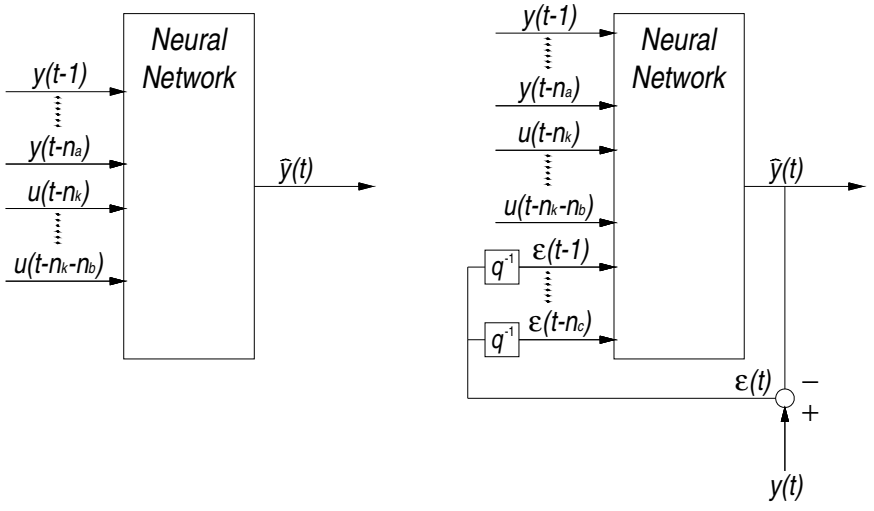


Fig. 3. The NNARX (left hand side) and NNARMAX (right hand side) model structures

7 Case Study: Temperature Short-Term Forecasting in Palermo

By using the Neural Network based System Identification Toolbox (NNSYSID) (Norgard, 2000) a nonlinear regression model was implemented for each weather station in order to accomplish a short-term temperature forecasting. The used model is, for each station, a two layer neural network with 10 neurons in the hidden layer, trained with the Levenberg-Marquardt algorithm (Hagan and Menhaj, 1994).

The activation functions employed are the hyperbolic tangent for the first layer and the linear function for the output layer. The embedding parameters of the predictors were determined by using the method described in section 5 (in every model it was set $n_a = n_b$).

The models were validated by observing the differences between the measured and the predicted temperatures for the week ranging from July 23rd to 29th 2008, which had not been used for the training phase.

The variables used for all the stations, except MeteoPalermo1, are: humidity, dewpoint, wind speed, atmospheric pressure and solar radiation. For MeteoPalermo1 only the variables humidity, dewpoint, wind speed and atmospheric pressure were used, because this weather station is not equipped with the solar radiation sensor.

The maximum values of the MAPE and MAE computed for the validation week are respectively 1.24% and 0.30 °C and they were obtained for Albunea

station. Table 1 shows the black-box model that yielded the best forecasting performances in each station, the corresponding embedding parameters and the values of MAPE, MAE and error variance obtained for the validation week.

Table 1. Mean absolute prediction error, percentage error and error variance obtained by each model for the week 23/07/08 – 29/07/08

Weather station	Model	Parameters	MAPE (%)	MAE (°C)	Error variance
Albunea	NNARMAX	n_a=2;n_b=2;n_c=2;n_k=1	1.24	0.30	0.16
Amaltea	NNARMAX	n_a=2;n_b=2;n_c=2;n_k=1	0.80	0.20	0.08
Cassandra	NNARX	n_a=2; n_b =2; n_c =1	0.86	0.22	0.11
Merlino	NNARMAX	n_a=3;n_b=3;n_c=1;n_k=1	0.70	0.18	0.05
MeteoPalermo1	NNARMAX	n_a=2;n_b=2;n_c=1;n_k=1	1.00	0.26	0.15
Morgana	NNARMAX	n_a=2;n_b=2;n_c=1;n_k=1	0.74	0.19	0.07
Pizia	NNARMAX	n_a=2;n_b=2;n_c=2;n_k=1	1.00	0.26	0.13

For the same week, the upper part of Fig. 4 shows the evolution of the output temperature (namely the temperature which represents the target of the system, i.e. the measured temperature) and of the predicted one at Morgana station. The lower part of Fig. 4 shows the corresponding evolution of the prediction error. Analogous graphs are showed in Figure 5 for Albunea station.

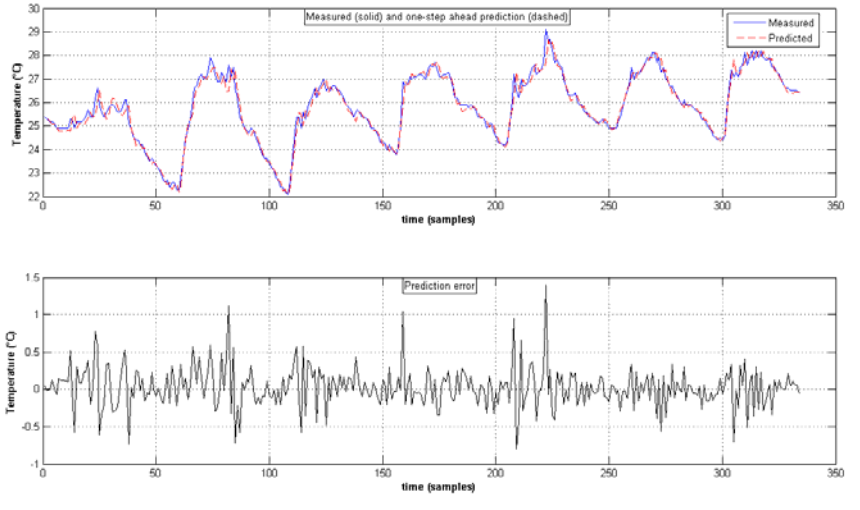


Fig. 4. Evolution of the measured and predicted temperatures for the period 23/07/2008 – 29/07/2008 (upper) and corresponding trend of the prediction error (lower) for Morgana station

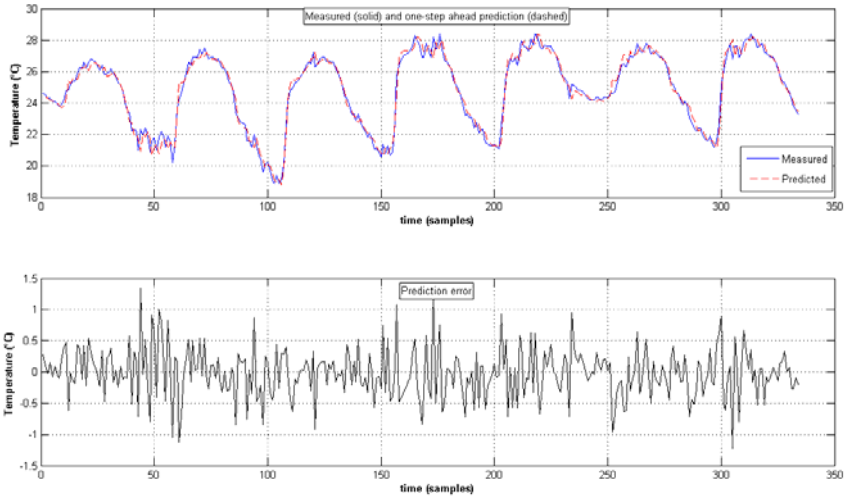


Fig. 5. Evolution of the measured and predicted temperatures for the period 23/07/2008 – 29/07/2008 (upper) and corresponding trend of the prediction error (lower) for Albunea station

These results can be considered good if compared to the error rates generally obtained in short-term temperature forecasting applications (see section 3).

7.1 Discussion of the Results

In order to study the UHI it can be very useful to create temperature contour maps to observe the spatial variation of the temperature within the urban texture. Especially if created starting from a significant number of monitoring locations, they can be enough accurate to easily point out the difference, in terms of temperature, between the urban area and the surroundings of the town.

In Figures 6 and 7 are respectively depicted the temperature contour map obtained by interpolating (using the inverse distance weighting method) the measured and the forecasted temperatures for 25th July 2008 at 4.00 a.m. (one of the coolest hours of the test week). As it is possible to observe, the two maps are very similar. In Figures 8 and 9 are showed the analogous maps referring to 26th July 2008 at 4.30 p.m. (one of the hottest hours of the test week).

It is interesting to note that during the diurnal hours the temperatures in the extra urban area are higher than within the city, whilst during the night the situation is opposite. This is easily understandable by bearing in mind that the thermal balance of urban areas is determined by the net balance between the solar gains and the heat loss caused by emitted long-wave (infrared) radiation. Because the radiant heat loss is slower in the urban areas than in the extra-urban neighborhood, the net balance determines the presence of higher temperatures in the former.

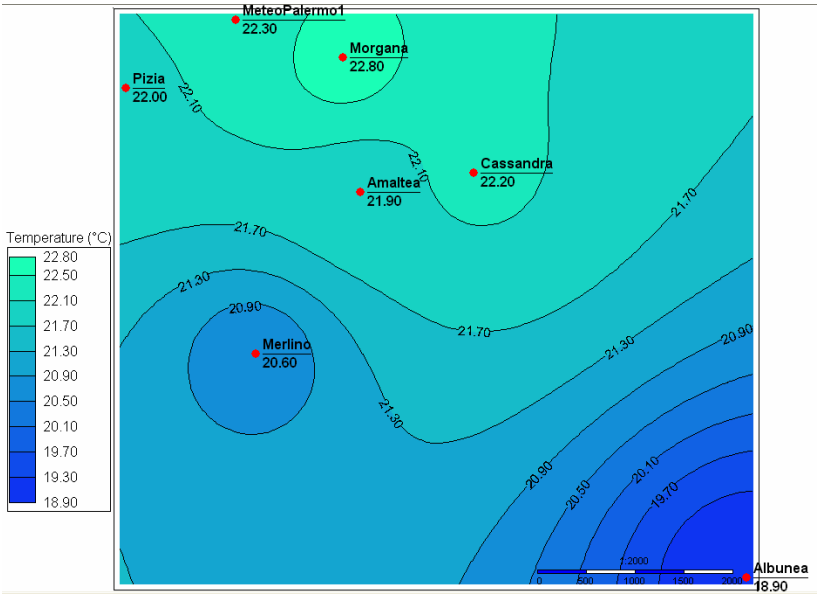


Fig. 6. Contour map of the measured temperature for 25th July 2008 at 4.00 a.m.

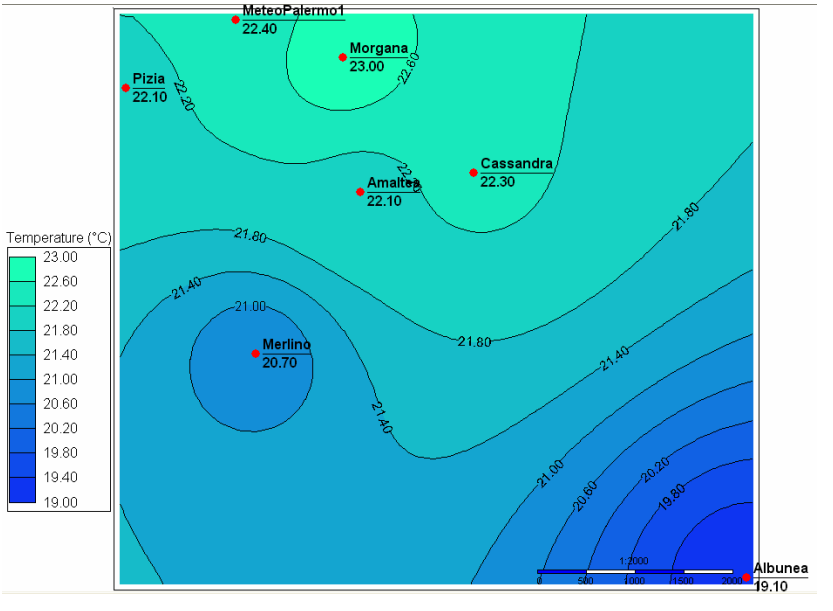


Fig. 7. Contour map of the forecasted temperature for 25th July 2008 at 4.00 a.m.

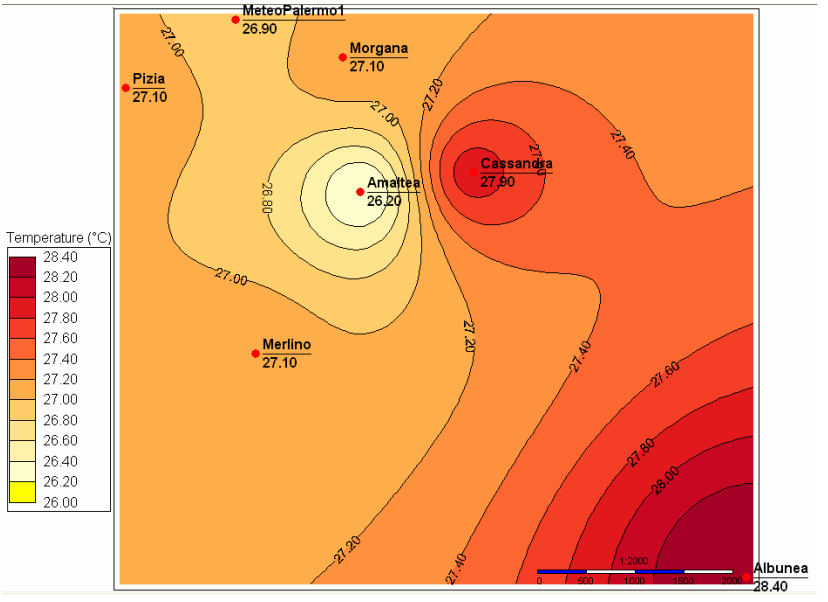


Fig. 8. Contour map of the measured temperature for 26th July 2008 at 4.30 p.m.

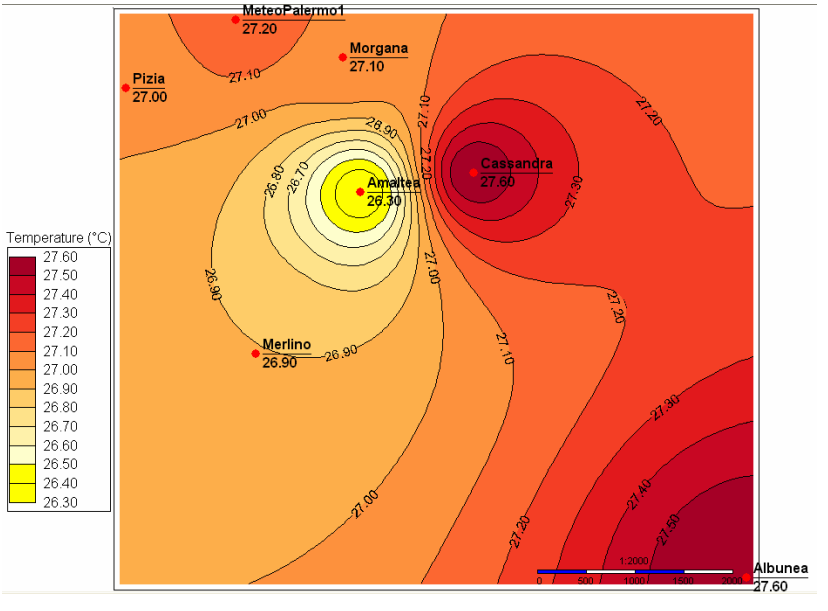


Fig. 9. Contour map of the forecasted temperature for 26th July 2008 at 4.30 p.m.

These considerations are strengthened by the observation of the summer daily average profile of the temperature difference between each of the urban stations and the extra-urban one (see Fig. 10).

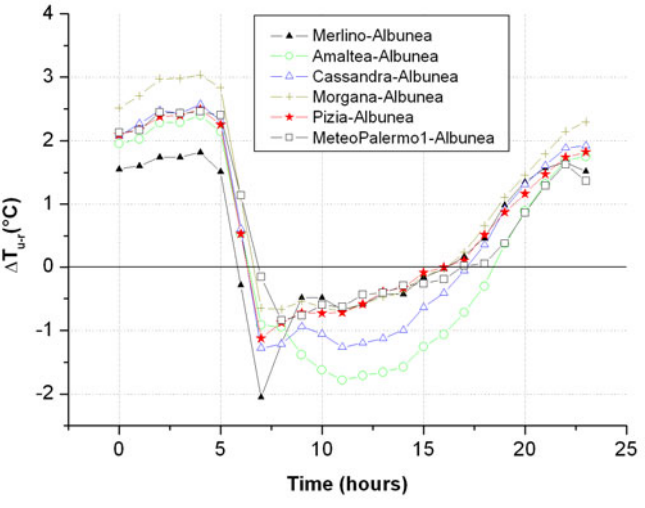


Fig. 10. Summer average temperature difference $(\Delta T_{u\text{-}r})$ between each of the urban weather station and the extra-urban weather station Albunea (summer 2008)

The diagrams showed in Fig. 10 were obtained by averaging the data recorded from June to August 2008. From the figure it is possible to infer that the extra-urban station is always hotter than every urban station from sunrise till about 6 p.m. (but in Merlino and Morgana stations the temperature matches the one in Albunea station about 2 hors earlier than in Amaltea station) and it is cooler from 6 p.m. till the sunrise of the day after. The maximum temperature difference ("heat island intensity") is about 3°C during the night and about 2°C during the day. However, it has to be remarked that the presence of lower dry bulb air temperature in the urban area during the hours ranging from the sunrise to the late afternoon does not imply that better comfort conditions are achieved here, because the comfort sensation felt by people depends upon the interaction of several factors, among which the combined effect of air temperature and humidity plays an important role. At this aim, different outdoor comfort indices have been conceived by researchers to assess human thermal stress and thermal comfort conditions of living environments (Epstein and Moran 2006).

Figure 11 shows the daily profiles of the hourly average summer temperatures in each of the weather stations. It is possible to observe that from midnight till the sunrise, there are four approximately parallel temperature profiles (Albunea;

Merlino; the group made up of Amaltea, Cassandra and Pizia; the couple Morgana and MeteoPalermo1) and, as previously noted, the lowest temperatures are attained at the extra-urban station. From sunrise till about 2 p.m. the lowest temperatures are attained at Amaltea station and the highest at Albunea (the extra-urban station). From about 2 p.m. to about 6 p.m. the temperature profiles are very close to each other, except for Amaltea station (which remains the coolest) and Cassandra station, which is slightly cooler than the remaining stations. From about 6 p.m. to midnight the temperature profiles of all the urban stations tend to get very close to each other, while the temperatures at the extra-urban station are again the lowest.

The presence of four approximately parallel temperature profiles from midnight to the sunrise can be better explained if one observes the frequency distribution diagram of the nocturnal wind directions related to the same period (June-August 2008). This diagram is showed in Fig. 12. Observing the figure, it is clear that during night the prevailing wind blows from the South-West direction. This is the direction of a typical "land breeze" in Palermo, due to differential heating of the earth's surface during the daily radiation cycle. During the night, sea is hotter than land and thus warm air rises from the surface of the sea drawing cooler air from the land to take its place, thus creating convection cells. The effect of this phenomenon is the so called "land breeze". It is generally very light, and its speed usually ranges from 2.5 to 4 m/s.

Along the path from land toward the sea, the air mass driven by the land breeze gradually tends to increase its temperature because it stores the heat removed from the land coverage. This could be one of the reasons for which the higher nocturnal temperatures are those recorded by the weather stations which are closer to the sea (along the direction of the land breeze).

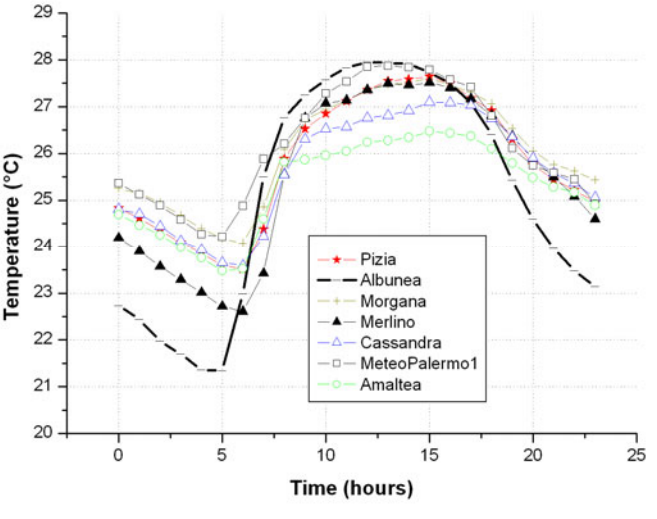


Fig. 11. Daily profiles of the hourly average summer temperatures (summer 2008)

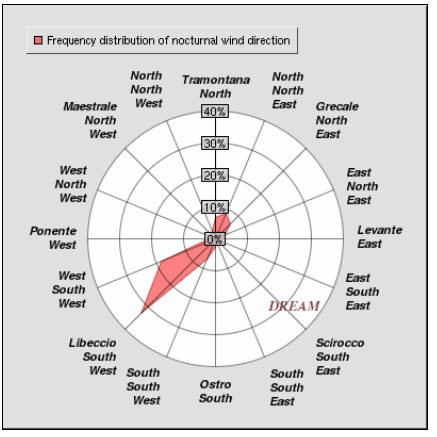


Fig. 12. Frequency distribution of nocturnal wind directions for the period ranging from June to August 2008

8 Conclusions

In the paper a weather monitoring system was described and set of nonlinear black-box models for short-term air temperature forecasting was applied on the data recorded trough the system. The system was designed and put into operation by the Department of Energy and Environmental Researches of the University of Palermo (Italy) and is currently measuring a large set of weather variables in the town of Palermo. In this moment it is made up of six weather stations dislocated in an area of about 13.5 km^2 and it was originally designed to study the spatial and temporal variation of the temperatures within the urban area and investigate the influence of various factors on the UHI.

The time span of the forecast is the same as the data sampling (30 minutes) and the results obtained are very good. The effects of the UHI were also detected and quantified, finding out an average intensity of the UHI of about 3°C during the night and about 2°C during the day. In a future research activity a forecasting model with a wider time horizon will be implemented, in order to obtain the future evolution of the temperature with a relevant advance and use this information to study the evolution of urban comfort conditions. By using some of the existing outdoor comfort indices it will be possible to create dynamic maps of the actual and forecasted thermal hygrometric comfort conditions at urban scale and make them available on the web site of the Department.

References

Abarbanel, H.D.I.: Analysis of Observed Chaotic Data. Springer, New York (1996)

Abdel-Aal, R.E.: Hourly temperature forecasting using abductive networks. Eng. Appl. of Artif. Intell. 17, 543–556 (2004)

Alligood, K., Sauer, T., Yorke, J.A.: Chaos: An Introduction to Dynamical Systems. Springer, New York (1997)

ASCE - American Society of Civil Engineers, Aerodynamics Committee, Outdoor human comfort and its assessment: State of the Art Report. Boston, VA, USA (2004)

Ardente, F., Beccali, G., Cellura, M., Lo Brano, V.: Life cycle assessment of a solar thermal collector: sensitivity analysis, energy and environmental balances. Renew. Energy 30(2), 109–130 (2005)

Beccali, M., Cellura, M., Lo Brano, V., Marvuglia, A.: Forecasting daily urban electric load profiles using artificial neural networks. Energy Convers. and Manag. 45(18/19), 2879–2900 (2004)

Beccali, M., Cellura, M., Lo Brano, V., Marvuglia, A.: Short-term prediction of household electricity consumption: assessing weather sensitivity in a Mediterranean area. Renew. & Sustain. Energy Rev. 12(8), 2040–2065 (2007)

Beccali, G., Cellura, M., Culotta, S., Lo Brano, V., Marvuglia, A.: A web-based autonomous weather monitoring system of the town of palermo and its utilization for temperature nowcasting. In: Gervasi, O., Murgante, B., Laganà, A., Taniar, D., Mun, Y., Gavrilova, M.L. (eds.) ICCSA 2008, Part I. LNCS, vol. 5072, pp. 65–80. Springer, Heidelberg (2008)

Ben-Nakhi, A.E., Mahmoud, M.A.: Cooling load prediction for buildings using general regression neural networks. Energy Convers. & Manag. 45, 2127–2141 (2004)

Chen, S., Billings, S.A.: Neural Networks for Nonlinear Dynamic System Modelling and Identification. Int. J. Control 56(2), 319–346 (1992)

Epstein, Y., Moran, D.S.: Thermal comfort and the heat stress indices. Ind. Health 44, 388–398 (2006)

Gartland, L.: Heat islands: Understanding and Mitigating Heat in Urban Areas. Earthscan Publications, London (2008)

Gautama, T., Mandic, D.P., Van Hulle, M.M.: A differential entropy based method for determining the optimal embedding parameters of a signal. In: Proceedings of ICASSP 2003, Hong Kong, vol. VI, pp. 29–32 (2003)

Gautama, T., Mandic, D.P., Van Hulle, M.M.: The delay vector variance method for detecting determinism and nonlinearity in time series. Physica D 190(3-4), 167–176 (2004)

Gautama, T.: Optimal Embedding Parameters - A differential entropy-based method for determining the optimal embedding parameters of a signal (2007), http://webscripts.softpedia.com/developer/Temu-Gautama-15893.html (accessed October 1, 2008)

González, P., Zamarreño, J.M.: A short-term temperature forecaster based on a state space neural network. Eng. Appl. of Artif. Intell. 15, 459–464 (2002)

Hagan, M.T., Menhaj, M.: Training feedforward networks with the Marquardt algorithm. IEEE Trans. on Neural Netw. 5(6), 989–993 (1994)

Haykin, S.: Neural Networks: A Comprehensive Foundation, 2nd edn. Prentice-Hall, Englewood Cliffs (1999)

Hippert, H.S., Pedreira, C.E., Souza, R.C.: Combining neural networks and ARIMA models for hourly temperature forecast. In: IEEE-INNS-ENNS International Joint Conference on Neural Networks, Como, Italy, July 24-27, vol. 4, pp. 414–419 (2000)

Kantz, H., Schreiber, T.: Nonlinear Time Series Analysis. Cambridge University Press, Cambridge (1997)

Lanza, P.N., Cosme, J.M.: A short-term temperature forecaster based on a novel radial basis functions neural network. Int. J. of Neural Netw. 11, 71–77 (2001)

Ljung, L.: System Identification – Theory for the User, 2nd edn. Prentice Hall, Upper Saddle River (1999)

Ljung, L., Söderström, T.: Theory and Practice of Recursive Identification. MIT Press, Cambridge (1983)

Lopes, C., Adnot, J., Santamouris, M., Klitsikas, N., Alvarez, S., Sanchez, F.: Managing the Growth of the Demand for Cooling in Urban Areas and Mitigating the Urban Heat Island Effect. In: European Council for an Energy Efficient Economy (ECEEE) Congress, Mandelieu, June 11-16, vol. II (2001)

Mihalakakou, G., Santamororis, M., Tsangrassoulis, A.: On the energy consumption in residential buildings. Energy and Build 34, 727–736 (2002)

Norgard, M.: Neural Network Based System Identification TOOLBOX, version 2 (2000), http://www.iau.dtu.dk/research/control/nnsysid.html

Oke, T.R., Johnson, G.T., Steyn, D.G., Watson, I.D.: Simulation of surface urban heat islands under "ideal"conditions at night: part 2. Diagnosis of causation. Bound. Layer Meteorol. 56, 339–358 (1991)

Papadopoulos, A.M.: The influence of street canyons on the cooling loads of buildings and the performance of air conditioning systems. Energy and Build 33, 601–607 (2001)

Santamouris, M., Papanikolaou, N., Livada, I., Koronakis, I., Georgakis, C., Argiriou, A., Assimakopolous, D.N.: On the Impact of Urban Climate on the Energy Consumption of Buildings. Sol. Energy 70(3), 201–216 (2001)

Schreiber, T., Schmitz, A.: Surrogate time series. Physica D 142, 346–382 (2000)

Sjöberg, J., Zhang, Q., Ljung, L., Benveniste, A., Delyon, B., Glorennec, P., Hjalmarsson, H., Juditsky, A.: Nonlinear black-box modeling in system identification: a unified overview. Autom. 31(12), 1691–1724 (1995)

Takens, F.: Detecting strange attractors in turbulence. In: Rand, D.A., Young, L.A. (eds.) Dynamical Systems and Turbulence, pp. 366–381. Springer, New York (1981)

Theiler, J., Eubank, S., Longtin, A., Galdrikian, B., Farmer, J.D.: Testing for nonlinearity in time series: the method of surrogate data. Physica D 58(1-4), 77–94 (1992)

UNI 10349 Heating and cooling of buildings. Climatic data (1994)

Wong, N.H., Yu, C.: Study of green areas and urban heat island in a tropical city. Habitat Int. 29(3), 547–558 (2005)

Yang, I.H., Kim, W.K.: Prediction of the time of room air temperature descending for heating systems in buildings. Build. and Environ. 39, 19–29 (2004)

Automatic Mapping and Classification of Spatial Environmental Data

Mikhail Kanevski, Vadim Timonin, and Alexei Pozdnoukhov

Institute of Geomatics and Analysis of Risk (IGAR),
University of Lausanne (UNIL),
CH-1015 Lausanne, Switzerland

Abstract. The paper deals with the development and application of the generic methodology for automatic processing (mapping and classification) of environmental data. General Regression Neural Network (GRNN) is considered in detail and is proposed as an efficient tool to solve the problem of spatial data mapping (regression). The Probabilistic Neural Network (PNN) is considered as an automatic tool for spatial classifications. The automatic tuning of isotropic and anisotropic GRNN/PNN models using cross-validation procedure is presented. Results are compared with the k-Nearest-Neighbours (k-NN) interpolation algorithm using independent validation data set. Real case studies are based on decision-oriented mapping and classification of radioactively contaminated territories.

Keywords: automatic cartography, General Regression Neural Networks, Probabilistic Neural Networks, decision-oriented mapping and classification.

1 Introduction

Analysis, modelling, and visualisation of environmental data are crucial tasks for the risk assessment and risk management. With the development of sensor network technologies, data on pollution, levels of radiation, meteorological parameters can be continuously captured by the automatic stations. These large volumes of information need an efficient automatic methods and algorithms in order to produce decision-oriented results in a real time.

Recently a Spatial Interpolation Comparison (SIC2004) exercises were organised by the Joint Research Centre (Dubois 2005) in order to evaluate different methods and their efficiency with an objective of "automating spatial interpolation algorithms for environmental monitoring systems designed for routine and emergency situations". Needs and basic problems concerning automatic environmental data analysis and mapping were explained and justified in the introductory chapter of Dubois (2005). In these exercises participants have produced their spatial predictions using different interpolation methods. The quality of different predictions was estimated using independent validation data sets. One of the best performances in these exercises, especially in the case of an emergency situation, was demonstrated by GRNN model. The description of the models used in SIC2004, including geostatistical ones, can be found in Dubois (2005). The details about GRNN

B. Murgante et al. (Eds.): Geocomputation, Sustainability & Environ. Planning, SCI 348, pp. 205–223.
springerlink.com

application in SIC2004 can be found in Timonin and Savelieva (2005). Until now there were no special competitions of methods dedicated to spatial classification.

The current research includes both topics: automatic mapping/regression and classification based on GRNN and PNN. GRNN and PNN models have the same origin – nonparametric modelling of joint distribution function using Parzen window (Aha 1997, Specht 1991). In the present study automatic means an application of data driven approach where parameters of the models can be tuned automatically using training data set.

In general, automatic decision-oriented mapping/classification deals with an integration and processing of different kinds of data, models and expert knowledge, including several GIS thematic layers, science-based models (e.g. meteorological models, migration of heavy metals in the environment), decision criteria etc.

In the present study decision-oriented data processing is concentrated mainly on some topics of automatic cartography, namely: automatic exploratory spatial data analysis, including hot-spot detection; description of spatial predictability (presence/absence of spatial structures); fast and efficient mapping of anisotropic patterns discovered; probabilistic/risk mapping/classification of decision levels; quantification of spatial prediction uncertainties; visualisation of the results and corresponding uncertainties.

The results obtained using GRNN and PNN models are compared with a k-Nearest Neighbours method which is considered as a "benchmark" model for automatic mapping and classification. Case study considered is based on real data on soil pollution by Chernobyl radionuclides. The present research is an extended study first presented in Kanevski et al. (2008) and completed by automatic spatial classification.

2 Automatic Mapping of Spatial Data: Tasks and Models

2.1 k-Nearest Neighbours Algorithm

k-Nearest Neighbours (k-NN) algorithm is widely used in data analysis including geospatial data. A particular case is a one-nearest neighbour which is known under different names: Voronoi polygons, Dirichlet cells, Thiessen tessellation, etc. Often this method is used for the quick visualisation (a preview) of data or as a benchmark tool for the comparison with other models.

k-NN is an example of the so-called "lazy learning" algorithms (Aha 1997). With such approach, the function is modelled locally and all computations are made directly during the prediction step. Therefore there is no actual training phase – all training examples are just stored in the memory for further predictions. To make a prediction at some point in a space one finds the first k nearest training points, according to some predefined distance measure. The prediction is a mean over the values of its k neighbours. To run the k nearest neighbours algorithm, one needs to define the distance measure and the number of neighbours to be used. In general, Minkowski p-norm distance can be used:

$$d_s\ (x,y) = \left(\sum_i |x_i - y_i|^s \right)^{1/s} \quad (1)$$

where s is a parameter equal or greater than one. So, with $s = 1$ it is a Manhattan (or city block) distance, with $s = 2$ – Euclidean distance, and $s = \infty$ corresponds to the infinity-norm distance, that is, $max\ |x_i - y_i|$. The selection of s parameter is task dependent. Euclidean distance is the most widely used because of its straightforward interpretability.

For different datasets the optimal number of neighbours (parameter k) is different. It is necessary to tune this parameter for every particular task, for instance, by means of the cross-validation procedure.

2.2 Tuning of Models Using Cross-Validation

Cross-validation is a common approach for tuning hyper-parameters of the data driven models. In n-fold cross-validation, the original training data set is partitioned into n subsets. One of the n subsets is used as a validation data set to calculate the error, and the remaining $n-1$ subsets are used as training data. The cross-validation process is repeated n times (the number of folds), with each of the n subsets used exactly once as the validation data. The n validation errors from all folds then can be averaged (or combined otherwise) to produce a single cross-validation error estimation for the specified set of parameters (the single parameter k in the case of k-NN model). This procedure is repeated for different values of the parameter k . The model with the lowest cross-validation error is chosen as an optimal one.

As a special case, the number of folds n can be set equal to the number of observations in the original training data set. This special case of the n-fold cross-validation is called leave-one-out cross-validation (or sometimes simply cross-validation). It uses a single observation from the original data as the validation point, and the remaining (n-1) observations are used as the training data. This is repeated n times such that each sample is used as the validation datum.

In a more general setting neighbours can be weighted according to the distance to the estimation point (for example, by using inverse distance weighting) instead of simple averaging. Therefore the simplest model based on averaging of neighbours can be considered as a “vanilla” version of k-NN.

2.3 General Regression Neural Networks for Regression

2.3.1 Theoretical Background

General Regression Neural Network is a well-known statistical nonparametric regression method - Nadaraya-Watson Kernel Regression. It was proposed independently in 1964 by Nadaraya and Watson . In 1991 it was interpreted by Specht

in terms of neural networks . This approach is based on kernel nonparametric density estimation proposed by Parzen (1962).

GRNN prediction *Z(x)* is given by the following analytical formula:

$$Z(\mathrm{x}) = \frac{\sum_{i=1}^{N} Z_i K\left(\frac{\mathrm{x}-\mathrm{x}_\mathrm{i}}{\sigma}\right)}{\sum_{i=1}^{N} K\left(\frac{\mathrm{x}-\mathrm{x}_\mathrm{i}}{\sigma}\right)} \qquad i = 1, 2\ldots, N \tag{2}$$

where N is a number of training points, and Z_i are the measurements (training data).

The core of this method is a kernel $K(\cdot)$ which in the simplest case depends only on two parameters: the distance to the predicted point and model dependent parameter σ - a positive number called kernel bandwidth or just width. Application of kernels with larger than optimal values of σ leads to over-smoothing of data; smaller than optimal produces over-fitting of data.

Note that x_i, in fact, is a centre of the i-th kernel. Different kinds of kernels can be selected from the kernels' library (Hardle 1989, Fan and Gijbels 1997).

In practice a Gaussian kernel is the most widely used:

$$K\left(\frac{\boldsymbol{x}-\boldsymbol{x}_i}{\sigma}\right) = \frac{1}{(2\pi\sigma^2)^{p/2}} \exp\left(-\frac{\left\|\boldsymbol{x}-\boldsymbol{x}_i\right\|^2}{2\sigma^2}\right) \qquad i = 1, 2\ldots, N \tag{3}$$

where p is a dimension of the input vector x.

Gaussian kernel is a usual choice of kernel for GRNN as well. Finally, the GRNN prediction using Gaussian-type of kernel and without the normalization term is given by

$$Z(\boldsymbol{x}) = \frac{\sum_{i=1}^{N} Z_i \exp\left(-\frac{\left\|\boldsymbol{x}-\boldsymbol{x}_i\right\|^2}{2\sigma^2}\right)}{\sum_{i=1}^{N} \exp\left(-\frac{\left\|\boldsymbol{x}-\boldsymbol{x}_i\right\|^2}{2\sigma^2}\right)} \tag{4}$$

Note that GRNN is a linear estimator (prediction depends on weights linearly), but weights are estimated non-linearly according to the non-linear kernel (3).

One of the useful improvements is to use multidimensional kernels instead of one-dimensional like in (3). When σ parameter is a scalar we deal with an isotropic model. In a more general case, parameter σ may be presented as a covariance matrix Σ. Covariance matrix is a squared symmetrical matrix with dimension

p by p and with the number of parameters equals to $p(p+1)/2$. So in a more general (anisotropic) setting the equation (3) can be rewritten in the following way:

$$K\left(\frac{x-x_i}{\sigma}\right)=\frac{1}{(2\pi)^{p/2}(\det\Sigma)^{1/2}}\exp\left(-\frac{1}{2}(x-x_i)\Sigma^{-1}(x-x_i)\right) \quad (5)$$

where *det* means determinant and Σ^{-1} – is an inverse of Σ matrix.

Model with a kernel (5) is an anisotropic one and is much more flexible to model real-world data. It is especially useful in case of complex multidimensional data. For example, for 2D spatial mapping we can use the following parameterisation $\sigma=(\sigma_x, \sigma_y, \sigma_{xy})$.

Simple Gaussian kernel depends only on one hyper-parameter σ - the width of the kernel which can be tuned in an automatic manner by using the same cross-validation techniques as described for k-NN model.

2.3.2 GRNN Training Using Cross-Validation

In order to estimate the hyper-parameters of the GRNN usually a cross-validation technique already used for k-NN tuning is applied. The difference lies only in the type of tuning parameter: discrete for k-NN (k –value) and continuous for GRNN (σ-values). The optimal σ value corresponds to the model with the smallest cross-validation error. In case of general anisotropic GRNN model an optimisation procedure is performed in a $p(p+1)/2$-dimensional cube in order to find corresponding optimal σ-values.

2.3.3 Estimating Uncertainty with GRNN

In addition to the prediction of conditional mean GNRR can provide an estimate of higher moments. Moreover, measurement errors in data also can be taken into account.

The process of the uncertainty estimation, which in most cases is based on the analysis of the training residuals, can be implemented as a part of the automatic process using GRNN itself. It is based on a secondary training procedure when the squared residuals at the training points are used as a target function. The optimal value of σ for this uncertainty measure of GRNN model can be found by the same cross-validation procedures as described above. Then, the trained model can be used to estimate (predict) the uncertainty at the unknown locations (Fan and Gijbels 1997, Kanevski and Maignan 2004, Kanevski 1999).

It is necessary to keep in mind that such estimations of the uncertainty depend on monitoring network (distribution of measurement points in space), and are biased (due to the biased estimation of the mean).

Using this uncertainty estimation and by applying the hypothesis of the local normality different decision-oriented maps, including probability/risk maps of being above of some decision levels, can be produced.

The concept of “thick” isoline (Kanevski and Maignan 2004) can be used to combine predictions and uncertainties. “Thick” isoline is an isoline for a given

decision level surrounded by the regions of the uncertainty (for example, plus/minus one or two standard deviations). Such isolines can be treated as an area where real isoline of a given level should be located with a given probability. As usually, all measures of uncertainties and corresponding confidence intervals depend on the modelling hypotheses, e.g. local Gaussian distribution. Nevertheless, in a more general sense they can be considered as indicators of mapping uncertainties.

2.4 Probabilistic Neural Networks for Classification

The Probabilistic Neural Networks are based on the same nonparametric foundations (Parzen 1962). The PNN models were introduced by Specht (1991) for the multi-class classification problems. The multi-class classification problem deals with the construction of a classifier capable to decide to which of K classes does the unknown sample $\boldsymbol{x}$ belongs. In general, the Bayesian optimal or maximum posteriori (MAP) decision rule is applied:

$$C(\boldsymbol{x}) = \{c_1, c_2, ..., c_K\} = \underset{c_i}{\operatorname{argmax}}\ P(c_i)\, p(\boldsymbol{x} \mid c_i) \qquad i = 1, 2, ..., K \tag{6}$$

To implement this approach, the density of each class $p(x|c_i)$ has to be estimated first. The kernel density estimator already used to construct GRNN model can be applied. The difference between the two approaches is in the final objective.

The density of each class can be estimated independently by the following formula:

$$p(\boldsymbol{x} \mid c_i) = \frac{1}{\left(2\pi\sigma^2\right)^{p/2} N_i} \sum_{n=1}^{N_i} \exp\left(\frac{-\left\| \boldsymbol{x} - \boldsymbol{x}_i^{(n)} \right\|^2}{2\sigma^2} \right) \tag{7}$$

where N_i is the size of the class, i.e. the number of samples belonging to class c_i, and $\boldsymbol{x}_i^{(n)}$ represents the n^{th} sample of class c_i.

The main difference between regression and classification is only in the data set used for the estimation. In the case of regression, all N samples are used, but here only the samples belonging to a specified class c_i .

Now the decision on class membership can be made by selecting a winner by comparing the densities of different classes taking into account the prior probability $P(c_i)$. Furthermore, due to assumption that any sample x belongs to one of K classes, the Bayesian confidence (posterior probability of belonging $\boldsymbol{x}$ to class c_i) can be estimated in the following way:

$$P(x \mid c_i) = \frac{P(c_i)\, p(x \mid c_i)}{\sum_{k=1}^{K} P(c_i)\, p(x \mid c_k)} \tag{8}$$

2.4.1 Advanced Mapping with PNN

How can the result of modelling in the case of classification tasks be presented? The obvious answer which one expects from the classifier is a class to which the unknown sample belongs. Some primitive models like k-NN produce that result and nothing more. With more sophisticated models like PNN, more useful information for decision-making can be obtained.

The output of the PNN model includes the following features (advanced ones are listed in italic).

For every data sampled at a point x:

- the probabilities of belonging to each of the K classes;
- the decision on the class membership is based on the latter probabilities.

For any region of the input space, the map(s) of:

- predicted class;
- *densities* for all classes – K maps;
- *maximum* probabilities (probabilities of the winner class), which provide confidence information.

Traditional error estimations include the classification rates computed as the number (or percent) of misclassified points. But it may be not enough, especially for the number of classes more than two. Let us remind that the most general description of the (multi)classification performance deals with the confusion matrix. It is a $K \times K$ square matrix whose elements are the counts of the number of actual samples versus those predicted to be in the class. For example, a non-zero value in row n and column m ($n \neq m$) means that a given number of points with the actual class n are predicted by the model as belonging to the class m. An ideal classification performance would provide a confusion matrix which is diagonal (the elements are zeros everywhere except the main diagonal). In this case the diagonal values correspond to the class sizes.

For a classification task the so-called reject category is very often used in statistics. When one decides what class the given sample belongs to, not only the winner class (max probability among all other) but also the value of this probability can be taken into account. A threshold corresponding to a confidence probability level has to be defined. If model correctly predicts some sample (predicted class is really actual class) but the winner class has a probability less than this threshold, we cannot accept this result with enough confidence – it is a rejected value for this class. One additional column is added to the confusion matrix and for each class a reject value if exists is added. Note that if threshold is less than $1/K$ the reject category will always be empty, since the winner class probability is always not less than $1/K$.

To provide some problem-specific information on the model performance, the confusion matrix can be modified to introduce problem-specific costs to the produced errors. To test the quality (or for the comparison with other models) of the classification model the so-called cost matrix can be used as well. This is a $[K \times K]$ square matrix with zeros on the diagonal. The other elements are the values of the penalty for misclassification (larger values corresponding to more dangerous mistakes). For example, consider row n and column m: the value in this cell is the

penalty given if the actual class for some point is *n* but the model predicted it to be class *m*. With such a matrix one can distinguish between different types of errors.

A confusion matrix (without a reject category – threshold is zero for this case) and a cost matrix can be combined to calculate the performance of the classification model taking into consideration the level of "danger" of different errors. This can be done by taking an overall sum of the values of the element-by-element product of the two matrices. Let us note that these error measures are quite general and are applicable not only to PNN but others classification algorithms based on estimations of probabilities.

3 Automatic Mapping of Spatial Data: Real Case Study

3.1 Data Description

The data on radioactively contaminated territories by ^{137}Cs radionuclide after the Chernobyl accident are considered as a real case study (Parkin et al. 2002). Data base contains 660 measurements in the region under study. The 500 points were randomly selected as a training subset used to tune the parameters of the model, and the remaining 160 points were only used for the validation purposes, i.e. only to estimate generalization abilities of models and to compare them.

Descriptive statistics of data are presented in Table 1. Voronoi polygons of the data and spatial distribution of the training and validation points are presented in Figure 1.

Table 1. Basic statistics of all, training and validation data

Data set (size)	Min	Median	Max	Mean	StdDev
Total (660)	1.0	6.6	193.0	19.8	26.2
Training (500)	1.0	6.2	193.0	19.5	26.9
Validation (160)	1.2	9.0	140.0	20.7	24.2

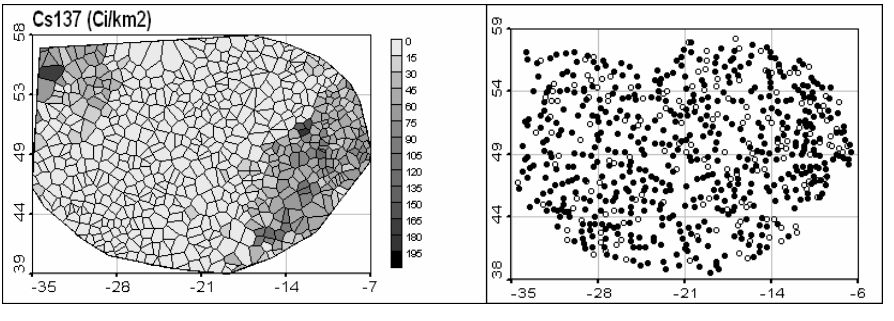


Fig. 1. Voronoi polygons of the raw data set (all data) - (*left*), spatial locations of training (filled circles) and validation (empty circles) points - (*right*)

In order to characterize spatial structures of raw data and the efficiency of modelling a well known geostatistical tool - variography was applied in the present study. Variograms are basic tools in geostatistical modelling and are widely used to characterise anisotropic spatial correlation patterns (Kanevski and Maignan 2004, Chiles and Delfiner 1999). It is an important exploratory spatial data analysis tool. In the present study the variography was not used to control an automatic mapping procedure.

The anisotropic variogram roses of training and validation data sets are presented in Figure 2. Variogram roses are composed of variogram calculation in several directions (directional variograms) and at different distances taking into account corresponding tolerances (Kanevski and Maignan 2004). Experimental variograms (Figure 2) demonstrate clear anisotropic structure – spatial correlations depend on the direction.

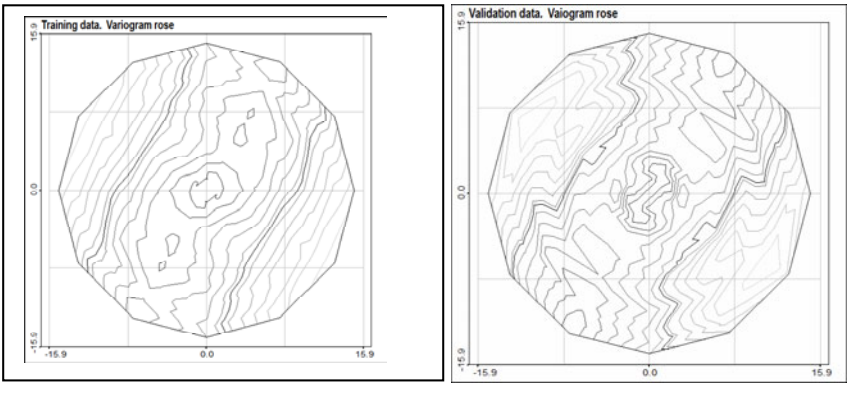


Fig. 2. Variograms roses for training (left) and validation data (right)

3.2 Regression Task

3.2.1 k-NN Modelling

First, let us apply a k-NN model for the prediction using the Euclidean metrics. The searching interval for *k* parameter was selected from 1 to 20 neighbours.

In Figure 3 the leave-one-out cross-validation error curve is presented. The optimal number of neighbours (*k* value corresponding to a minimum of cross-validation curve) is 5.

The result of mapping on a regular grid using 5-NN model is given in Figure 4. The root mean squared error (RMSE) and the coefficient of correlation (Ro) were used to measure the quality of the predictions. Error statistics (for all models applied in this case study) are shown in Table 2. Scatterplots of measured data

versus estimated values are presented in Figure 5. Corresponding omnidirectional variograms of the predictions on a regular grid (compared with original data and other models) are shown in Figure 11 (see more discussions below).

It is important to note, that k-NN cross-validation curve can be used to control the predictability of data, or in other words to detect the presence of spatial structures in original or training data. When there are no spatial structures variograms demonstrate pure nugget effect – variogram fluctuates around a priori variance for all distance lags considered. In such cases data are not correlated at all distances and the "best" prediction is a mean value of all data (we do not complicate explanations by considering clustering of measurement points and representativity of data).

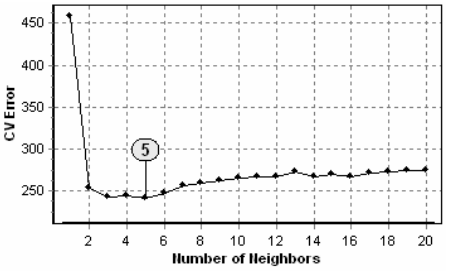


Fig. 3. Cross-validation curve for k-NN model. An optimal number of neighbours equals to 5

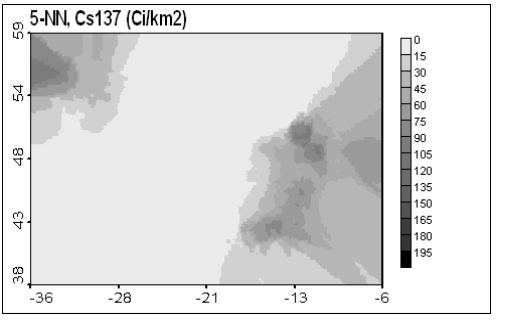


Fig. 4. Prediction mapping using 5-NN model

When there is no spatial correlation cross-validation error curve of k-NN model has no minimum. It means that there are no local patterns and the "optimal" map is a constant corresponding to the global mean value. Thus, the analysis of k-NN cross-validation curve can be used to check the predictability and locality of spatial patterns. More detailed discussions can be found in Kanevski et al. (1999).

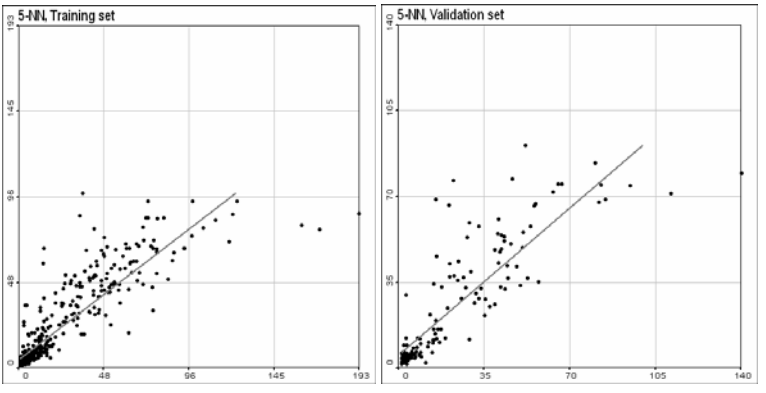


Fig. 5. Scatterplots of measured data versus estimations for 5-NN model: training data set (left), validation – (right). Solid line is a linear correlation line

3.2.2 GRNN Modelling

Now let us consider the application of GRNN model for the purposes of automatic mapping. In the previous studies the General Regression Neural Networks were efficiently applied for mapping of pollution (Timonin and Savelieva 2005, Kanevski and Maignan 2004, Kanevski 1999). In the present study the methodology is extended for automatic mapping by an application of isotropic and anisotropic GRNN models and characterization of the corresponding uncertainties. Training procedures are based on the cross-validation technique. During the study the variography is used as an important tool to control the quality of mapping produced by neural networks as it was proposed in Kanevski et al. (1996): the variograms of the residuals should demonstrate no spatial structures (pure nugget effect) with a variance fluctuating around nugget value of raw data.

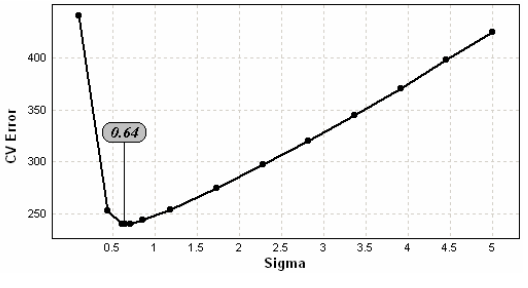


Fig. 6. Cross-validation curve for isotropic GRNN model. Optimal σ value is 0.64 km

The initial interval for the search of the optimal kernel bandwidth (σ value) was fixed in the interval [0.1 km; 5 km], and the number of searching steps equalled to 10. The result of kernel bandwidth search by using cross-validation procedure is given in Figure 6. The curve has a well-defined minima, with an optimal value of $\sigma = 0.64$ km.

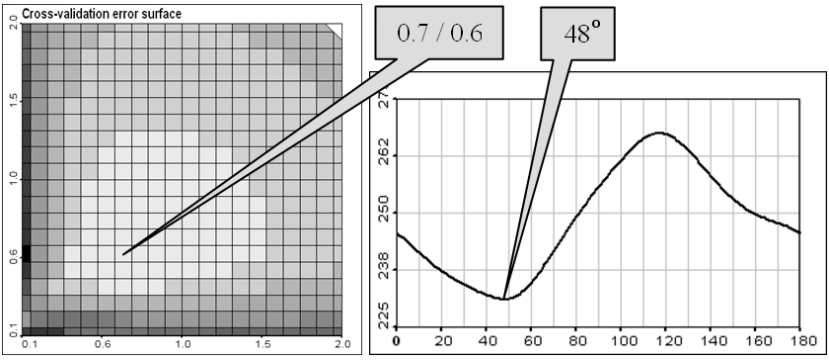


Fig. 7. Cross-validation surface for anisotropic GRNN model. Optimal σ value is 0.7 / 0.6 km – (left). Cross-validation curve for rotation angle – (right). Optimal value of anisotropy is 48 degrees.

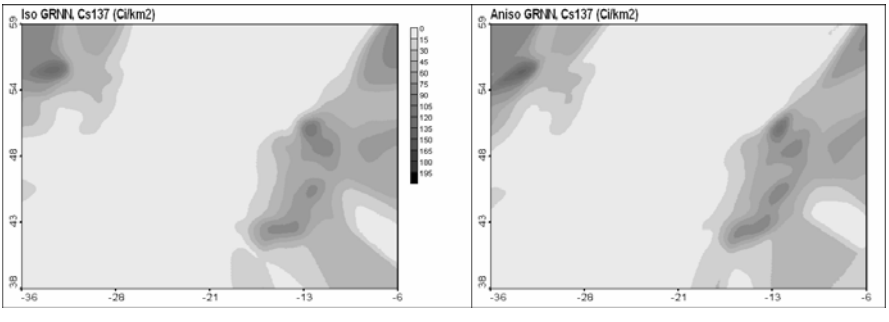


Fig. 8. Prediction mapping using GRNN models: isotropic – (left), and anisotropic – (right). Colour scale is the same for both maps.

Now, let us examine if a more complex anisotropic model can improve the results of the isotropic model. The grid search procedure to find best anisotropic kernel bandwidths was applied. The initial interval for the search of the optimal kernel bandwidths was fixed as [0.1 km; 2 km] for both directions, and the number of steps [10 x 10]. In Figure 7 (left) the result of the search is presented as a cross-validation error-surface. The optimal value (the minimum on the error surface corresponding to the x and y directions) is 0.7 km in x-direction and 0.6 km in y-direction. In order to detect the direction of the anisotropy an optimal data rotation angle (in x-y space) was performed by using the same cross-validation procedure: cross-validation was carried out for all [0; 180] degrees with step of 1 degree (Figure 7 (right)). The optimal angle of the anisotropy detected is 48 degrees from the WE direction. It should be noted that the anisotropy detected is rather weak (the difference between kernel bandwidths is rather small).

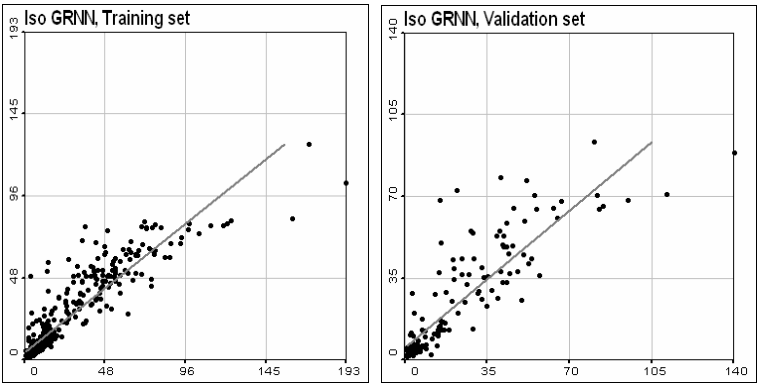


Fig. 9. Scatterplots of measured versus estimated values for isotropic GRNN model: training data set (left), validation – (right). Solid line is a linear correlation line.

Error statistics are presented in Table 2. Scatterplots of measured values versus estimates are presented in Figure 9 (isotropic model) and Figure 10 (anisotropic model). Omnidirectional variograms for GRNN predictions compared with the original data and 5-NN model are shown in Figure 11.

In case of the absence of spatial structures (pure nugget effect) the cross-validation GRNN curve has no minimum. The "optimal" solution in this case is reduced to the global mean value (see corresponding considerations for k-NN model). This fact can also be used to detect spatial patterns and to quantify the predictability of data. Thus, the analysis of cross-validation surfaces can significantly contribute to the exploratory analysis of data and to study structures of spatial patterns.

Root mean square errors (RMSE) observed are quite small and corresponding correlations between data and predictions are high enough both for training and validation data sets and for all models examined. As it was expected, scatterplots of measured versus estimated values (Figures 5,8,9) demonstrate some smoothing effects both for training and validation data subsets.

The quality of modelling (extraction of patterns from training data and prediction of validation data) can be studied by using variograms: models of good quality are able to extract spatially structured information by leaving only normally distributed and spatially uncorrelated residuals (Kanevski and Maignan 2004, Kanevski et al. 1996). Omnidirectional variograms presented in Figure 11 for the original data and developed models quantify these considerations. Models have filtered out nugget/"noise" (measurement noise and unexplained small scale variability) in the original data and correctly modelled spatial patterns. The range of the spatial correlation for all variograms is almost the same and approximately equals to 20 km.

GRNN model looks better from the point of view of error statistics and from the visual presentation of the predicted maps (see Figure 4 for k-NN model, Figure 8 for GRNN models).

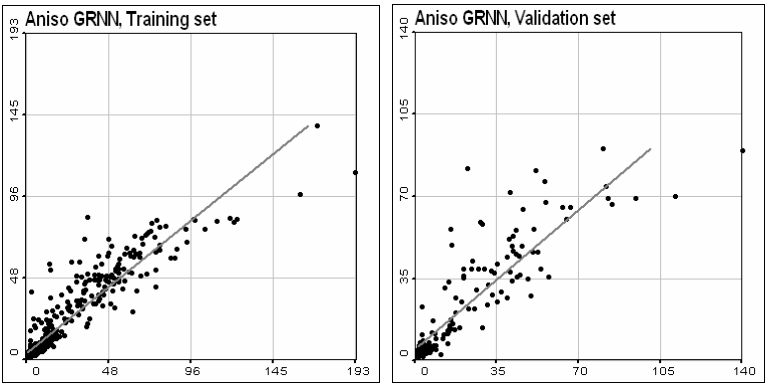


Fig. 10. Scatterplots of measured values versus estimated ones for anisotropic GRNN model: training data set (left), validation – (right). Solid line is a linear correlation line.

Anisotropic GRNN model produced lower RMSE and higher correlation than the isotropic one.

Taking into account the results presented above, GRNN can be considered as a promising modelling tool adequate for automatic nonlinear and adaptive spatial data analysis and mapping tool. A simpler k-NN method may be equally used, especially for a quick visualisation of data. In case of high global anisotropy an anisotropic type of GRNN should be preferable.

Table 2. Error statistics of the predictions for all models

Model		RMSE		Ro	
		Training	*Validation*	Training	*Validation*
5-NN		12.5	*22.1*	0.88	*0.86*
GRNN	isotropic	10.6	*12.4*	0.92	*0.87*
	anisotropic	9.8	*11.9*	0.93	*0.88*

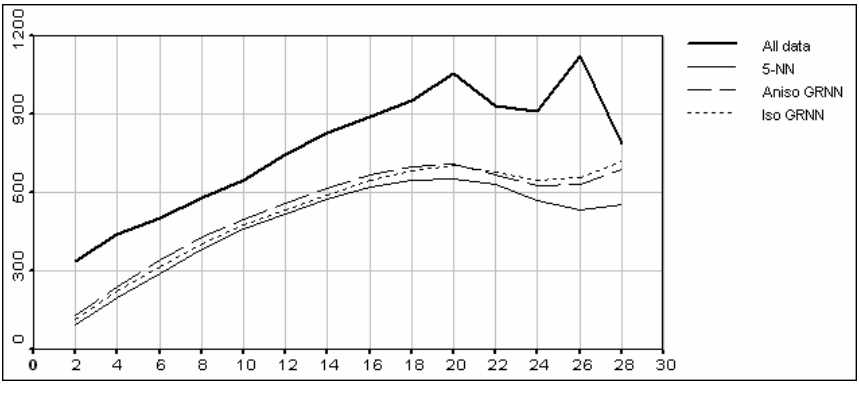


Fig. 11. Omnidirectional variograms for predictions of all studied models: original data (*thick solid*), 5-NN (*solid*), GRNN anisotropic (*dash*), and GRNN isotropic (*dot*)

3.2.3 Decision-Oriented Mapping

Using background and assumptions described in paragraph 2.3.3, in addition to the results obtained two more decision-oriented maps can be produced. First map (Figure 12 left) is a map of probability of exceeding a given contamination level - 70 Ci/km^2. This map was produced applying the hypotheses of local Gaussian distribution and by using conditional local mean and variance values [probability = 1 – (error function)]. The second map is a map of "thick" isoline produced for the same decision level and which represents an isoline surrounded by a gray zone. This zone characterises the uncertainty (plus/minus standard deviation) around the selected level. Confidence intervals can be derived from the available results by taking into account and accepting some hypotheses on models of local distributions, e.g. Gaussian distributions.

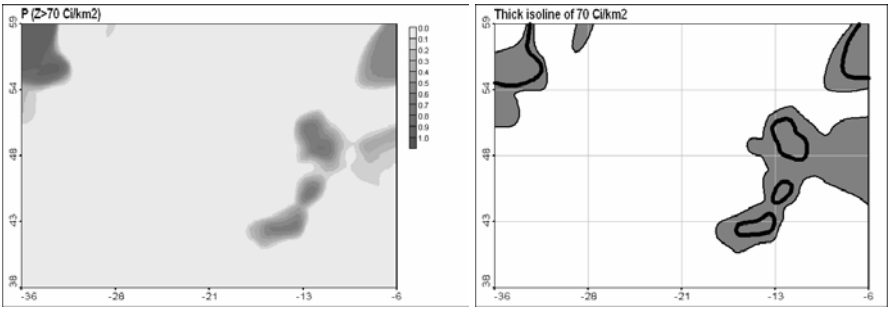


Fig. 12. Decision-oriented mapping: probability of exceeding level 70 Ci/km^2 - (*left*), "thick" isoline estimated for the same level - (*right*)

3.3 Classification Task

To demonstrate PNN model capabilities and to compare with GRNN modelling results the same data set is used. In order to compare with the result of advanced GRNN mapping (thick isoline and a map of probability of exceeding a given level of 70 Ci/km^2) an indicator transform to the continuous data was applied in order to pose a two class classification problem. One class corresponds to the values higher than 70 Ci/km^2, and second one to the values below this level. The same procedure was applied both to training and to validation data subsets.

3.3.1 k-NN Classification

First, like for GRNN model, a k-NN classification model was applied. The searching interval for *k* parameter was selected from 1 to 20 neighbours. In Figure 13 the leave-one-out cross-validation error curve is presented. The optimal number of neighbours (*k* value with a minimum on cross-validation curve) equals to 5.

The result of mapping on a regular grid using 5-NN model is given in Figure 14. The percent of misclassified points was used to measure the quality of the predictions (Table 3).

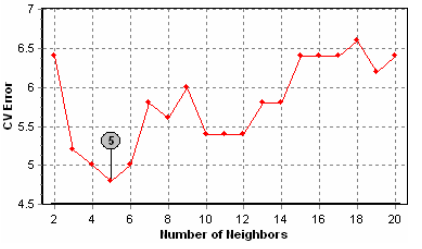


Fig. 13. Cross-validation curve for k-NN classification model ($k_{optimal} = 5$)

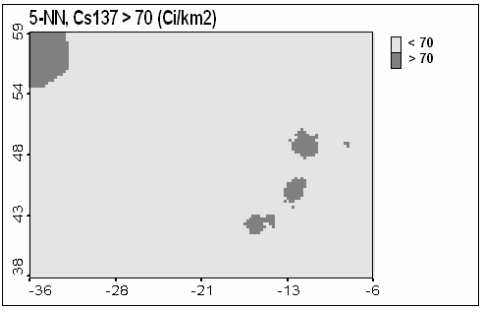


Fig. 14. Prediction mapping using 5-NN model

3.3.2 PNN Classification

Now let us use PNN model for automatic classification purposes. Training procedures are based on the cross-validation technique and are the same that were used for the GRNN model. Omitting the details, the optimal (anisotropic) value of kernel bandwidths were found to be 0.23 km by 0.18 km. Prediction results (probability of exceeding the level 70 Ci/km^2 and predicted classes with a threshold probability of 0.5) are presented in Figures 15 and 16. The percentages of misclassified data are shown in Table 3.

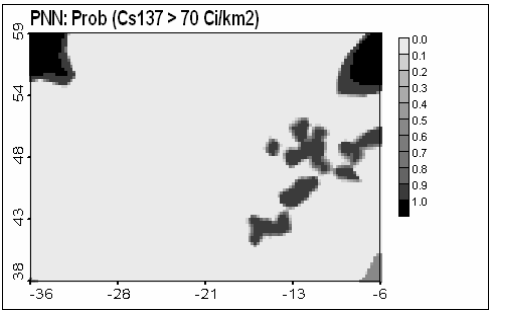


Fig. 15. Prediction mapping using PNN model: probability of exceeding level 70 Ci/km^2

In this case study anisotropic PNN model significantly outperforms the k-NN algorithm. Both training and (more important) validation errors are smaller. The quality of mapping can be examined even visually: in PNN prediction there is a well-defined area with high probability of values exceeding studied level 70 Ci/km^2 in the right/top corner of the map. In k-NN result there is no such area. But keeping in mind the distribution of dataset (training and validation) and results of the regression mapping it seems that high values in this area are more probable than small ones.

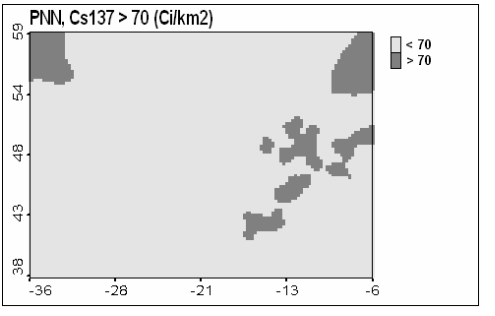


Fig. 16. Prediction mapping using PNN model: predicted classes with threshold probability 0.5

Table 3. Error statistics of the predictions for all models (classification task)

Model	Error, %%	
	Training	Validation
5-NN	3.6	8.1
Anisotropic PNN	1.4	7.5

4 Conclusions

Environmental automatic decision-oriented treatment of data – from exploratory data analysis to mapping and classification, is an important decision-oriented task. In general, the problem is to develop nonlinear, robust, adaptive and data driven methods which can detect spatial patterns in noisy data and are able to produce prediction and probabilistic/risk maps which can be used in a real decision making process. Other important properties of such methods should be their ability to model multi-scale variability in data and to detect hot-spots (e.g. accidental releases).

In the present study data driven nonlinear modelling methodologies based on General Regression Neural Networks and Probabilistic Neural Networks were considered in detail.

Both models can be used at different stages of environmental data analysis: from exploratory data analysis to the spatial prediction tasks. Depending on the

number of data and the dimension of the space either cross-validation or an n-fold cross-validation techniques can be used.

The results were compared with a simple and easy to use k-NN model which can be considered as a benchmark model in machine learning.

The future adaptations of these models for the automatic decision-oriented mapping deal with an application of different types of kernels, adaptation to the clustered monitoring networks and kernels varying in space, extensions to multivariate data, direct multi-scale analysis, and better description and visualisation of the corresponding uncertainties.

Detailed description of traditional and advanced models of geostatistics, machine learning and Bayesian maximum entropy approaches and their applications for advanced mapping of geospatial data can be found in Kanevski et al. (2009) and Kanevski (2008).

Acknowledgments. The work was supported in part by Swiss National Science Foundation projects "GeoKernels." (№200021-113944, №200020-121835) and "ClusterVille" (№ 100012-113506).

References

Aha, D.W. (ed.): Lazy Learning. Kluwer Academic, Dordrecht (1997)

Chiles, J.-P., Delfiner, P.: Geostatistics. Modelling Spatial Uncertainty. Wiley, Chichester (1999)

Dubois, D.: Automatic mapping algorithms for routine and emergency data. European Commission, JRC Ispra, EUR 21595 (2005)

Fan, J., Gijbels, I.: Applied Local Polynomial Modelling and Its Applications. In: Monographs on Statistics and Applied Probability 66. Chapman and Hall, London (1997)

Hardle, W.: Applied Nonparametric Regression. Cambridge University Press, Cambridge (1989)

Haykin, S.: Networks: A Comprehensive Foundation. Prentice-Hall, Englewood Cliffs (1998)

Kanevski, M. (ed.): Advanced Mapping of Environmental Data. Geostatistics, Machine Learning and Bayesian Maximum Entropy. iSTE and Wiley, London (2008)

Kanevski, M., Pozdnoukhov, A., Timonin, V.: Machine learning algorithms for spatial data. In: Theory, applications, and software tools. EPFL Press, Lausanne (2009)

Kanevski, M., Timonin, V., Pozdnoukhov, A.: Automatic Decision-Oriented Mapping of Pollution Data. In: Gervasi, O., Murgante, B., Laganà, A., Taniar, D., Mun, Y., Gavrilova, M.L. (eds.) ICCSA 2008, Part I. LNCS, vol. 5072, pp. 678–691. Springer, Heidelberg (2008)

Kanevski, M., Maignan, M.: Analysis and Modelling of Spatial Environmental Data. EPFL Press, Lausanne (2004)

Kanevski, M., Arutyunyan, R., Bolshov, L., Demyanov, V., Maignan, M.: Artificial neural networks and spatial estimations of Chernobyl fallout. Geoinformatics 7(1-2), 5–11 (1996)

Kanevski, M.: Spatial Predictions of Soil Contamination Using General Regression Neural Networks. Systems Research and Information Systems 8(4), 241–256 (1999)

Nadaraya, E.A.: On estimating regression. Theory of Probability and its Applications 9, 141–142 (1964)

Parkin, R., Kanevski, M., Saveleva, E., Pichugina, I., Yatsalo, B.: Implementation of Neural Networks for Assessment of Surface Density Contamination with 90Sr. Nuclear Power Engineering (2), 63–69 (2002)

Parzen, E.: On estimation of a probability density function and mode. Annals of Mathematical Statistics 33, 1065–1076 (1962)

Rosenblatt, M.: Remarks on some nonparametric estimates of a density function. Annals of Mathematical Statistics 27, 832–837 (1956)

Specht, D.E.: A General Regression Neural Network. IEEE Transactions on Neural Networks 2, 568–576 (1991)

Specht, D.E.: Probabilistic Neural Networks. Probabilistic Neural Networks 3(1), 109–118 (1990)

Timonin, V., Savelieva, E.: Spatial Prediction of Radioactivity Using General Regression Neural Network. Applied GIS 1(2), 1901–1914 (2005), doi:10.2104/ag050019

Watson, G.S.: Smooth regression analysis. Sankhya: The Indian Journal of Statistics, Series A 26, 359–372 (1964)

Detecting Landforms Using Quantitative Radar Roughness Characterization and Spectral Mixing Analysis

Andrea Taramelli

ISPRA Institute for Environmental Protection and Research,
via di Casalotti, 300, Rome, Italy
Office: +39-06-615-70546, +39-06-5007-4096,
Fax: +39-06-615-70543
e-mail: andrea.taramelli@isprambiente.it
www.isprambiente.it

Abstract. In Geomorphology the landforms delineation and delimitation are based on traditional techniques which is usually achieved by two different approaches: semantic and geometric. The semantic approach involves interpreting aerial photographs as well as using field knowledge to map the possible landforms extent. The geometric approach involves inferring the probability of extension on related properties through observations of topographic attributes like slope, elevation and curvature. This report, based on a similarity geometric model, uses quantitative roughness characterization, Spectral Mixing Analysis and fuzzy logic to map alluvial fans. The research is applied in the Alashan region in China because the timing of alluvial deposition is tied to land surface instabilities caused by regional climate changes. The main aim of the research is to understand where they form and where they extent in an effort to develop a new approach using the physical response within the SMA, the backscatter roughness parameters and primary attributes (elevation and curvature) derived from the Shuttle Radar Topography Mission (SRTM) Digital Elevation Model (DEM). So that, this study helps to provide a benchmark against which future alluvial fans detection using roughness, SMA and fuzzy logic analysis can be evaluated, meaning that sophisticated coupling of geomorphic and remote sensing processes can be attempted in order to test for feedbacks between geomorphic processes and topography.

Keywords: Alluvial Fan, SMA, Roughness, Fuzzy Logic, Curvature, Elevation.

1 Introduction

One of the most important overall objectives of remote sensing activities in topography and surface change is to locate and characterize physical properties of topography. This can be accomplished with multitemporal, moderate resolution reflectance data from the Landsat ETM+ sensors, as well as topographic elevation and radar backscatter data obtained from the 2000 Shuttle Radar Topography Mission (Rabus et al. 2003; Kobrick 2006). An important consideration for

B. Murgante et al. (Eds.): Geocomputation, Sustainability & Environ. Planning, SCI 348, pp. 225–249.
springerlink.com

monitoring surface is the nominal temporal frequency of remote sensor data acquisitions required to adequately characterize specific events. Topography specific parameters are an important consideration for determining the required frequency of data collections to minimize change omission errors. The broadband multispectral imagery provided by Landsat can be used to classify and map the extent of specific topographic parameters characteristic of some landforms and distinguish them from soil and vegetation classes. The characteristic spatial scale and the spectral variability of steep ridge areas cause serious problems for traditional image classification algorithms that discriminate the aforesaid classes on the assumption of spectral homogeneity (Settle and Drake 1993).

In areas where the reflectance spectra varies appreciably at scales comparable to or smaller than the Instantaneous Field Of View (IFOV), the spectral reflectance of an individual pixel will generally not resemble the reflectance of a single class (dark surface, soil or vegetations), but rather a mixture of the reflectances of two or more classes present within the sensor's IFOV (Singer and McCord 1979). Classification algorithms that accommodate spectral heterogeneity have been shown to discriminate heterogeneous "built" environments from more homogeneous natural environments. Spectral Mixture Analysis (SMA) instead classifies individual mixed pixels according to the distribution of spectrally pure end member fractions and provides a tool for discrimination and classification of surface topography (Adams et al. 1986; Taramelli and Melelli 2009). In addition, SMA can be used to discriminate among different types of natural environments on the basis of end member fraction distributions. Spectral characteristics of the surface topography show significant differences between the spectral reflectance of different landform surfaces (Small 2001; 2004).

The SRTM topography and backscatter data are particularly useful for augmenting the Landsat land surface classification as well as for providing input into the landform identification (Allievi et al. 2003; Tarmelli and Melelli 2008a). Landforms in geomorphology are features on the earth's surface produced in the past or in the present time. Topography, soil type and climate conditions are the most important control factors on specific landform processes. Climate conditions change in different areas and in a wide range of times and can differ in periods than times evolution of geomorphologic landforms. In this context the topographic variables (as slope, aspect, curvature or surface roughness) are the important factors that can control the boundary conditions for actual geomorphic processes. Thus landform delineation is usually achieved using the remote sensing interpreting techniques applied to topography and soil type that are often adequate to provide unique results. Soil type is most efficiently mapped using optical reflectance data (using Landsat, as discussed above), whereas surface roughness can be obtained from synthetic aperture radar (SAR) backscatter measurements in addition to other radar (Antonello et al. 2004; Taramelli and Melelli 2008b) and optical reflectance techniques (Imhoff et al. 1987). At present, both optical and radar-based methods for surface roughness topography can yield only coarse (0.25º grids). To provide a sound empirical evaluation of the delineation of landforms using remote

sensed analysis, a strict definition of the landforms being investigated is needed to minimize the error matrix and consequent delineation errors in mapping transposition. In this paper, a fuzzy logic computer-based algorithm that uses the Landsat SMA, the Shuttle Radar Topography Mission (SRTM) DEM and a quantitative analysis of the SRTM radar backscatter data for roughness estimation is employed to investigate the relationship between alluvial fans and the distribution of boundaries including independent variables like elevation and surface curvature.

2 Landforms: Semantic and Geometric Approaches

The goal of defining landforms in geomorphology is usually achieved using two main approaches that are widespread and well known in the scientific literature (Wood et al. 1960; Fairbridge 1968; Dehn et al. 2001): the semantic and the geometric one. According to the semantic approach, a landform is the result of a classification that simplifies the real world. The classification of the landforms is dependent on the scientists' background and research context. When we look at a landscape which is made up of the surface of Earth continuously varying in elevation, together with natural and anthropogenic phenomena superimposed on it, the strategy for classification is first based on different conditions (i.e. structural) and that can lead to some errors in mapping transposition. These errors increase if the landform is defined by different scientific approaches (hydrology, geology, sedimentology or others) because of the heterogeneity of names and consequent low consistency. The geometric approach (including of topological considerations and semantic definitions) highlights the geometric characteristics of a feature related to the topographic surface properties. Some landforms show, independently from the environmental conditions and the scientific disciplines, the same geometric response expressible in terms of angles (slope, curvature), and distance ratio along the principal directions (height, width, thickness).

The attempt to analyze the landscape using a geometric approach based on the semantic one, has given rise to an area of research which predicts the shape of the surface by delineating a set of rules within SAR data (Sainsbury 1989; Pike 2000; Robinson 1988; 2002). The advent GIS software and of spatial data in the form of SAR data has made possible the task of generating descriptive statistics of the shape of the surface and locations in the landscape to an exhaustive set of features based on the local form of the land surface.

Among the simpler, geometric, and therefore computable, set of forms is the assignment of a location of different morphometric parameters to primary or secondary (compound) attributes (Speight 1984). Thus, if we consider a regular grid of elevations stored as a backscatter value, a set of grid cells (or pixels) in the data can be assigned to a specific morphometric set of features which people recognize in the landscape. In this context the algorithms traditionally included in most raster processing systems use neighborhood operations to calculate slope, aspect, shaded

relief and points of inflection. To overcome possible limitations, software has been developed that uses interactive spatial techniques as well as a neighborhood operation that can best be visualized as region-growing procedures. They provide an analyst with the ability to extract information from roughness on morphologic features and properties, specifically topographic depression and curvature.

That knowledge recalls the one underlined about the landscape concept that is characterized by a degree of uncertainty, especially in its spatial extent. The ambiguity of the above attributes is probably best matched to the model of vagueness (Sainsbury 1989). Despite a precise definition of the different morphometric classes, any location can be allocated to a specific class, but the class to which a location is assigned by this precise process varies because of the scale of measurement giving rise to ambiguity as to the correct classification and so vagueness. Recently, several researchers have introduced the idea that the vagueness in geomorphology may be appropriate for analysis by fuzzy sets (Zhu, 1997a; 1997b; MacMillan et al. 2000; Argialas et al. 2003; Petry et al. 2003; Robinson 2003). Algorithms, used to identify such features, are now standard tools within GIS. They can provide a standardized approach to identify landforms (Dikau, 1989; Dubois and Prade 1996; Cross et al. 2000; Cobb et al. 2003; Zadeh 2002). Two methods for defining the membership values of the fuzzy sets have been developed (Ahamed, 2000; Fisher 2000a; 2000b; Ladner et al. 2003): the semantic import model based on the a priori knowledge that assigns a value of fuzzy membership to a landscape feature with a particular metric property, such as height (Usery 1996); the similarity relation model that uses surface derivatives, such as slope and curvature, as input to a multivariate fuzzy classification which yields the membership values (Irvin et al., 1997). The research in based on the second one, taking into account that the method may only be applied at the data resolution.

In this method the similarity representation of different classes in the parameter domain is based on fuzzy logic applied in GIS (Zhu 1997). Under fuzzy logic, a class at a given pixel can be assigned to more than one geometric class with varying degrees of class assignment (Burrough et al. 1992; 1997; McBratney et al. 1992; 1997). These class degrees assignment are referred to as “fuzzy memberships”. This fuzzy representation allows a class at each pixel to bear a partial memberships in each of the prescribed classes. Then each fuzzy membership is regarded as similarity measure between the local class and the typical case of the given class. By coupling this similarity representation with a raster backscatter data model, allow alluvial fans in different area to take property values of the prescribed classes (Zhu et al. 1996; 1997).

3 Methodology

The major problem with accurate definitions of landforms is the complexity of the boundaries, which are the result of the interplay of many factors, some of which are known and mappable while others are known but cannot be effectively expressed digitally. Transposing a feature on a map could show very different results

and depends on various variables: the scale of the feature (meso- or macro-landform), the map scale corresponding to the degree of accuracy, human error introduced by subjectivity and the errors from insight transposing mapping techniques. Landform delineations using traditional surveyor techniques are sometimes inadequate to provide unequivocal results. The quiescent or inactive boundaries of a landform may change over a very short time scale due to non-conservative lithotypes, even thought change may be expected only over a geomorphologic time scale. Moreover, specific land use and land cover (for example, in areas with much human activity) could mask the features. The landforms representation on thematic maps is thus the first fundamental step to visualize the features and assess associated processes. In this context the research studied the alluvial fans, depositional features of water-transported material (alluvium) with a longitudinal section cone geometry. They typically form at the base of slopes or at the junction of a tributary stream with a main one, and are characterized by a distinct break in slope at the junction between mountain areas and floodplain (Bull 1977). They appear as a cone segment radiating away from a point source (fan apex) with the coarser sediment in the upper fan but gradually diminishing toward the plain. Deposits of the alluvial fan may result from debris flows deposit (Melelli and Taramelli 2004) and/or be water-laid (Nemec and Steel 1988). The morphologic and morphometric characteristics of the fan are strictly related to the lithotypes and the shape of the catchment area (Leeder 1997).

3.1 Broadband Mixture Analysis

A mixing space can be thought of as a coordinate system in which reflectance spectra are represented as linear mixtures of spectrally pure endmember spectra. Spectrally mixed pixels can therefore be represented as combinations of the pure endmember spectra corresponding to spectrally distinct features within the area of the pixel (Small 2004; Li et al. 2005; Bach et al. 2007; Tarmelli and Melelli 2009). The location of a mixed pixel in the mixing space specifies the relative areal abundance of spectral endmembers that contribute to the mixed pixel reflectance. Linear combinations of spectral endmembers can therefore be used to describe spectra that occur within a convex hull prescribed by the distribution of mixed pixels and endmembers (Boardman 1993). The concept of the mixing space was developed to facilitate analysis of mixed reflectance with linear mixture models. Using the concepts of convex geometry allows the linear mixture problem to be treated using linear inverse theory (Boardman 1989, 1993). Linear mixture models are based on the observation that, in many situations, radiances from surfaces with different "endmember" reflectances mix linearly in proportion to area within the IFOV (Nash and Conel 1974; Johnson et al. 1983). This observation has made possible the development of a systematic methodology for Spectral Mixture Analysis (Adams et al. 1986; 1993; 1995; Smith et al. 1985; Gillespie et al. 1990)

that has proven successful for a variety of quantitative applications with multispectral imagery (e.g. Smith et al. 1990; Elmore et al. 2000; Roberts et al. 1998b; Small 2004; Taramelli and Melelli 2009). If a limited number of distinct spectral endmembers can be found, then it is possible to define a mixing space where mixed pixels can be described as linear mixtures of these endmembers. Given sufficient spectral resolution, a system of linear mixing equations can be defined and solved for the best fitting combination of endmember fractions for each of the observed reflectance spectra. The solution to the linear mixing problem can be cast as a linear inverse problem in which the system of mixing equations is inverted to yield estimates of the endmember fractions that best fit the observed mixed reflectances (Boardman 1993; 1994). The strength of the SMA approach is that it explicitly takes into account the physical processes responsible for the observed radiances and therefore accommodates the existence of mixed pixels (Lu and Weng 2006).

The first task, described here, is to determine the extent to which Landsat can distinguish important land cover and the land surface properties that influence the mass and energy fluxes to be modeled. Once this has been determined and validated, it is possible to develop a strategy for mapping and monitoring the spatiotemporal variations in the land surface.

The primary requirements for this landform analysis are spatial, spectral and temporal resolution, accurate calibration and a suitable temporal baseline. Of the available data sources, Landsat imagery is best suited to this task because it provides a calibrated 30 year baseline and both the spatial and temporal resolution necessary to distinguish among important land cover types and to characterize their temporal variability on seasonal to internannual scales. Our strategy is to 1) use spectral mixture analysis to represent land surface reflectance as continuous fields of biophysical land surface properties, 2) use field validation to verify the accuracy of the mixture model and 3) use decision tree classification to divide the spectral mixing space into discrete land cover types. The Alashan mosaic can be represented as a tetrahedral mixing space bounded by four distinct spectral endmembers. The endmembers correspond to green vegetation, non-reflective dark surface and two rock/soil substrates. The non-reflective dark surfaces represent both transmissive (e.g. clear water), absorptive (e.g. Fe rich rocks) and non-reflective (e.g. deep shadow) targets. The rock/soil substrates represent SWIR bright sands and a higher albedo, more spectrally flat reflectance corresponding to both evaporates and mud/silt lithologies. Following the procedure described in detail by Small (2001; 2004), the research generate a mixing space for each individual scene in the Alashan mosaic and a composite mixing space for the entire mosaic. As would be expected, the endmembers derived from the composite mixing space bound those found for each individual scene. However, the selection of 3 or 4 endmembers per individual scene resulted in 20 endmember suites (16 scenes + double redundancy for 4 hotspot scenes) that were generally consistent with the endmembers derived from the regional composite – as well as generic endmembers found in the global analysis described by Small (2004). The mixing space and endmembers are shown in Figure 1.

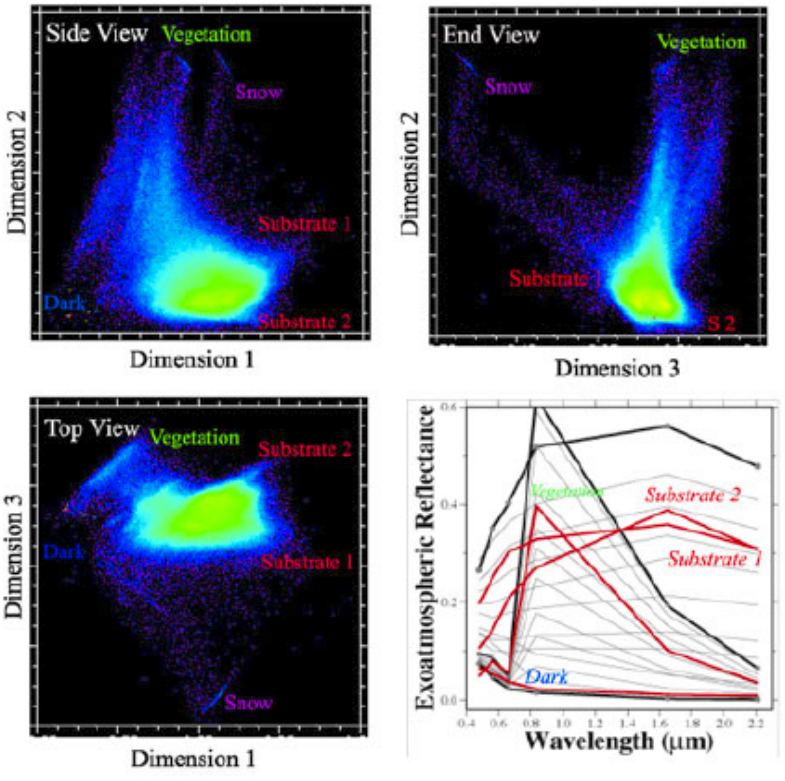


Fig. 1. Spectral mixing space for Alashan ETM+ mosaic of the Alashan district (Pasqui et al., 2008) that correspond to a shallow lacustrine depositional environment. Color density shaded (warmer = more pixels) scatterplots show orthogonal projections of the 3D mixing space spanned by the 3 low order principal components (PCs) containing>90% of image variance. Dimension 1 corresponds to rock/soil substrate fractions while dimension 2 corresponds to vegetation fraction and dimension 3 accommodates the divergence of substrate 1 (mud/siltstone) and Substrate 2 (sand). Snow and ice form a separate mixing continuum not contained within the primary tetrahedral mixing space bounded by other four endmembers.

Endmember fraction composites illustrate primary land cover distributions and spatial variations in the areal abundance of the primary biophysical properties of the land surface. Fraction composites are generated by assigning three endmembers to the red, green and blue layers of an image. For Alashan, using one substrate (red), vegetation (green) and dark surface (blue) highlights the landform areas and the contrast between high and low albedo rock/soil substrates (Figure 4). Using both mud/silt (red) and sand (green) as well as dark (blue) highlights lithologic differences between the low albedo Fe rich crystalline rocks, high albedo mud/silt deposits and evaporates and intermediate albedo sands (Figure 2).

The research conducted a preliminary interpretation of alluvial deposits and lacustrine depositional environment features for the Alashan area where the alluvium meets the crystalline basement.

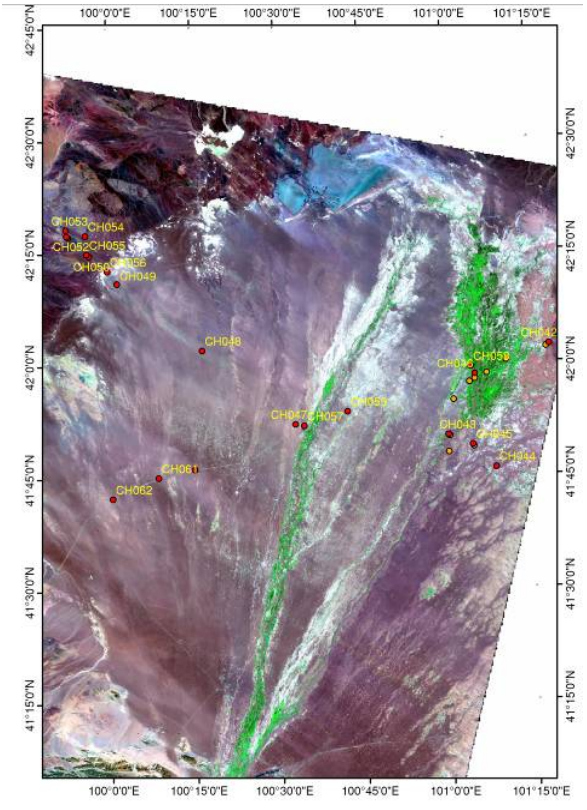


Fig. 2. On the ground, the white stripes appear as alternating of surfaces covered with coarse-grained gravel a few centimeters in diameter and the intervening surfaces mantled by finegrained gravel less than 1 cm in diameter

3.2 Estimating SubPixel Surface Roughness Using the C-Band SAR Backscatter from SRTM

In order to obtain a quantitative surface characterization, parameters have to be selected that will allow a clear discrimination between different types of landforms. The most obvious parameters are the magnitude and horizontal length scale of surface height variations. If the surface can be modelled as a stationary random Gaussian process, mean and variance of the elevation, and the autocorrelation function (which is related to the horizontal length scale of height variation) provide a complete description of the statistical surface properties. In theoretical models of rough surface scattering, it is often assumed that the surface is stationary with a Gaussian height distribution (Ulaby 1982). In addition, it is assumed that the mean elevation of the surface is subtracted from all height data. In this case, the autocorrelation function is identical to the autocovariance function, and

the square root of the height variance is the standard deviation around zero mean. Work in the 1980s, 1990s and 2000s (Greeley et al. 1995; 1997; Blumberg 2007) showed that the logarithm of roughness length zo depends linearly on radar backscatter power, as measured by normalized radar cross-sections σ^o expressed on a decibel scale (Figure 3). More recently, radar based estimates of roughness length have been made globally from C-band ERS1/2 radar scatterometer measurements by a group based in the CNRS/Universités Paris VII-XII (Pringent et al. 1997). Although the spatial resolution is somewhat coarse, the main advantage of ERS1/2 scatterometry lies in its full coverage of the globe every ~4 days with 50 km resolution, so that seasonal and interannual variations in z_o can be taken into account. Prigent group (Pringetnt et al., 1997) show that the scatterometer estimates of z_o are fully consistent with those of the ASU group's results from the airborne and space Shuttle SAR data. After correcting for variable look angle effects, they find a linear relation between the logarithm of z_o and radar cross section σ^o (Figure 3).

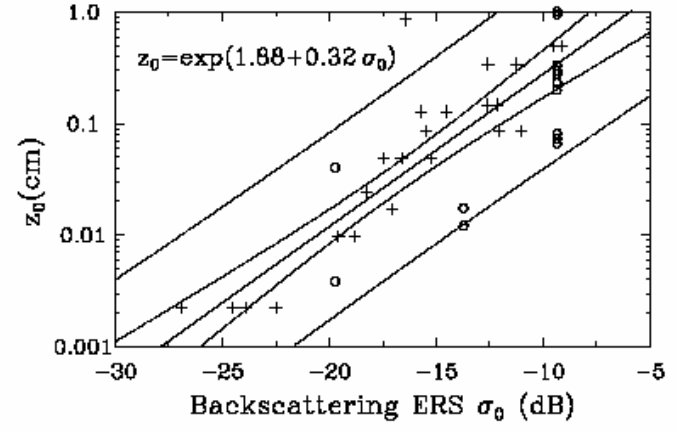


Fig. 3. Scatter plot of z_o estimates from previous studies (various symbols) versus the ERS backscatter coefficient σ^o (Pringent et al., 1997). The log linear regression is indicated by the solid line, and the dashed lines represent the 95% confidence interval and the 2 standard deviation interval (close to, and farther from the linear regression, respectively)

More recently, a method for estimating roughness length z_o based on bi-directional reflectance distribution functions (BDRF) was developed, obtained by the POLDER-1 instrument flown aboard ADEOS 1 platform from October, 1996 through June, 1997 (Laurent et al., 2005). POLDER-1 was a broad swath instrument that operated in the VNIR, and was capable of a range of zenith viewing angles up to 60°, allowing multiple viewing angles for a given site. Averaging over a 30-day period allowed determination of the BDRF with a nadir spatial resolution of 6 km x 7 km. Small scale roughness elements cast shadows which affect the BDRF. The BDRF data are used to determine a dimensionless quantity called the "protrusion coefficient" PC, and the aerodynamic roughness length z_o is founded through an empirically determined relation to the PC parameter:

$$zo = a * \exp(PC/b) \quad (1)$$

where a is a constant with units of length, and b is a dimensionless constant. Earlier works found that the PC parameter and normalized radar cross-sections σ^o (in db) are linearly correlated (Figure 4). Furthermore, both parameters are linearly related to the logarithm of the aerodynamic roughness length zo.

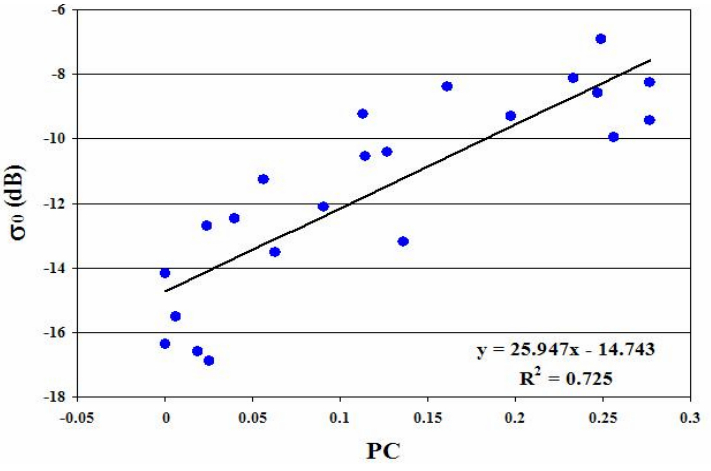


Fig. 4. Relation between normalized radar cross section σ^o derived from SAR data and protrusion coefficient PC obtained from the POLDER-1 bi-directional reflectance distribution function (from B. Marticorena, Universités Paris VII-XII). Both σ^o and PC relate to surface roughness.

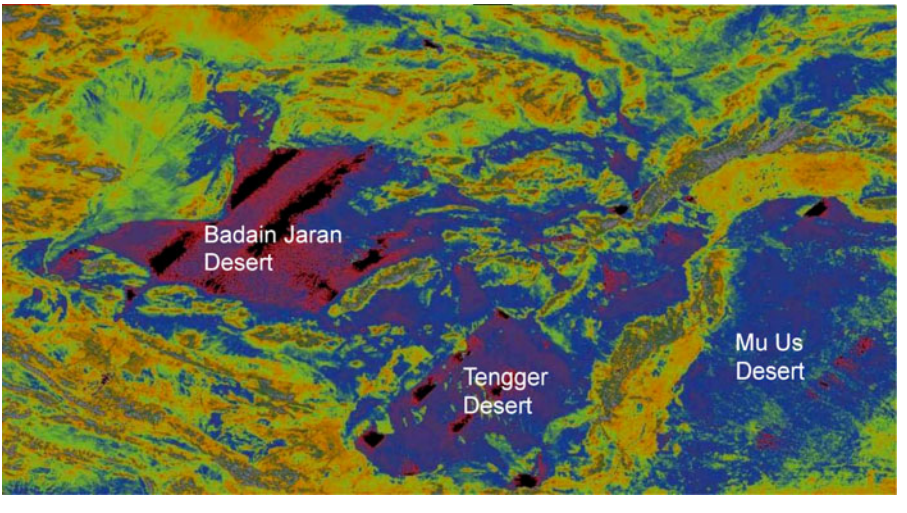


Fig. 5. C-band backscatter power from the February 2000 SRTM mission. Purple -> blue is low; orange -> gray is high. Data gaps are black.

Based on the aforesaid research an higher resolution (~25 m) estimates of zo can be derived for the entire world from the C-band SAR backscatter from the February 2000 SRTM. An image of the study area of SRTM C-band backscatter is shown in Figure 5 (from http://onearth.jpl.nasa.gov/). The most striking characteristic of SAR data is the low power in the radar returns over arid areas like deserts. Figure 5 shows an image of C-band backscatter power obtained during the February 2000 Shuttle Radar Topography Mission (SRTM). The deserts of the Alashan region stand out by their very low backscatter (dark blue and purple in Figure 5). Such low radar returns contrasts the higher backscatter (orange to gray) observed from rock outcrop and vegetated areas.

The research analyzed C-band SRTM backscatter data in the form of 1-degree geographic tiles for the Umbria region. Detailed topography, at selected elevation ranges that matches alluvial flood plain and junction areas typical of the intermontane basin, needs to be recognized using useful parameters based on SRTM backscatter data (Blumber, 2005; Laurent et al. 2005); to be able to detect transient surface events that can occur over widely separated geographic locations. The SRTM backscatter analysis leads to the detection of the roughness signature that represents the small-scale variation in the relief of a terrain surface (Benallegue et al. 1995; Dierking 1999); based on the statement that the amount of backscatter is proportional to the roughness of a surface. As a result, the total amount of the increases in the backscatter signal within the alluvial fans area where the other smooth surfaces, like flood plain, reflect most of the microwave energy away from the radar and produce a smaller backscattered portion. The following data products were developed from the SIR-C data takes over Alashan:

1) RGB false-color composites that display C-Band $\sigma^{o}hh$ in the red and blue channels, and L-Band $\sigma^{o}hh$ in the green channel. Thus, areas where the shorter wavelength C-Band backscatter dominates are purple and areas where L-band backscatter is important show as green.
2) For the SLC and MLD data, two additional data products were found to be very useful:
   a. The sum of C-Band $\sigma^{o}hh$ plus L-Band $\sigma^{o}hh$ expressed on a blue to red color scale, where blue is low total backscatter power, and red is high backscatter power.
   b. The ratio L-Band $\sigma^{o}hh$ to the sum of C-Band $\sigma^{o}hh$ plus L-Band $\sigma^{o}hh$. This expresses the relative importance (or fraction) of the longer wavelength L-band backscatter in the total backscatter power from both bands at each pixel. For arid to semi-arid areas like the Alashan, C-Band backscatter is stronger that L-band for most surface types, as shown in Figure 6.

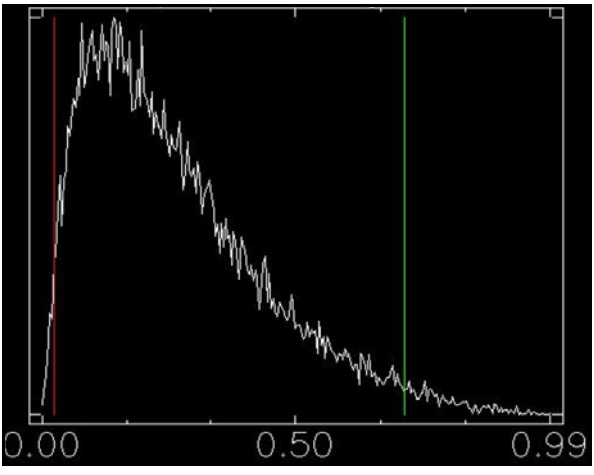


Fig. 6. Fraction of L-Band in the total backscatter power at L- and C-Bands for each pixel in data takes 11440/1

4 Analyses

The main objective in this section is to illustrate the general characteristics of the SIR-C radar data, focusing on its ability to delineate surfaces composed of dry, unconsolidated materials like sand, and loose soil, and in some case to "see" through a thin (< a few meters) cover of dust/sand to the bedrock below.

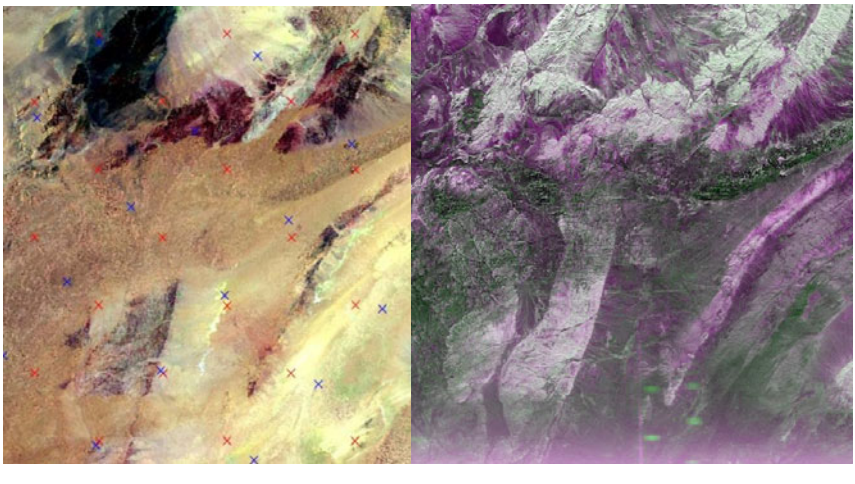


Fig. 7. a. Portion of spring 2004 Landsat TM mosaic corresponding to SIR-C data in Figure 7b. North is towards upper left. **b.** RGB composite of C- band, L-band, C-band σ^{o}hh for data takes 41833/4. Center latitude is 40.94°N; center longitude is 103.92°E. Radar illumination is from the top, and the swath is about 90km wide.

The research also compare the coverage SIR-C data to the 2004 Landsat TM which covers the entire Alashan region. In Figures 7a and b is shown the comparison of the Landsat with L- and C-band SIR-C data takes 13046/7 for an area in the northeast part of the Badain Jaran Desert. In the Landsat image (Figure 7a), bedrock outcrop is dark (black or purple), the sand and dust covered areas are brown and beige, and alluviated hill sides at the top are light and medium gray. In the C- and L-band SIR-C image (Figure 7b), bedrock outcop is bright grey or purple hued, the latter areas are dominated by C-band (5.8 cm wavelength) backscatter, whreas in green hued areas L-band (23.5 cm wavelength) backscatter is important (but might not be dominant – see Figure 6).

There are three main points to take from Figures 7a and b. First, backscatter power from unconsolidated sandy surfaces is low compared to rock outcrop. Second, L-band backscatter is important from these dry, unconsolidated materials (green-ish areas Figure 7b). And third, the longer, meter-scale wavelength energy used by radars (compared to the micro-meter scales sensed by optical instruments) allows microwave energy to penetrate a few meters into dry, unconsolidated materials, scattering from buried surfaces and other subsurface targets. This is demonstrated by comparing Figures 7a and b, noting that the SIR-C radars, particularly the L-band, "see" bedrock surfaces buried beneath thin sand sheets, whereas these are not visible in the Landsat data. Examples of SIR-C L- and C-band data products, defined above, are compared with the 2004 Landsat TM data in Figures 8a-c. Figure 8a gives the distribution of summed L-band plus C-band $\sigma^{o}hh$ backscatter power for data takes 13046/7. Comparing with the Landsat TM (Figure 8c), is possible to see again that that backscatter from the sand streams and sand sheets is low (dark blue), whereas backscatter from rock outcrop is high (red). A minor exception to this is the very strong backscatter from the sand dune facets facing the radar (Figure 8a, upper left), as mentioned earlier. Figure 8b demonstrates that L-band backscatter is relatively important for the dry, unconsolidated sandy materials (light tones). However, recall that C-band is generally stronger than L-band for radar backscatter over arid and semi-arid areas (Figure 6). The results presented in Figs. 8a and b suggest the following classification criteria for mapping areas of alluvial, unconsolidated sandy materials using dual-frequency (C- and L-band) SAR: surfaces composed of dry, unconsolidated sandy materials might be found where the backscatter power in both bands is very low, but where the backscatter in L-band is relatively important. The exception to this rule will occur for alluvial fields as noted above. However, the high backscatter from alluvial facets is very local in character, commensurate with the size and spacing of fans that are favorably oriented relative to the radar.

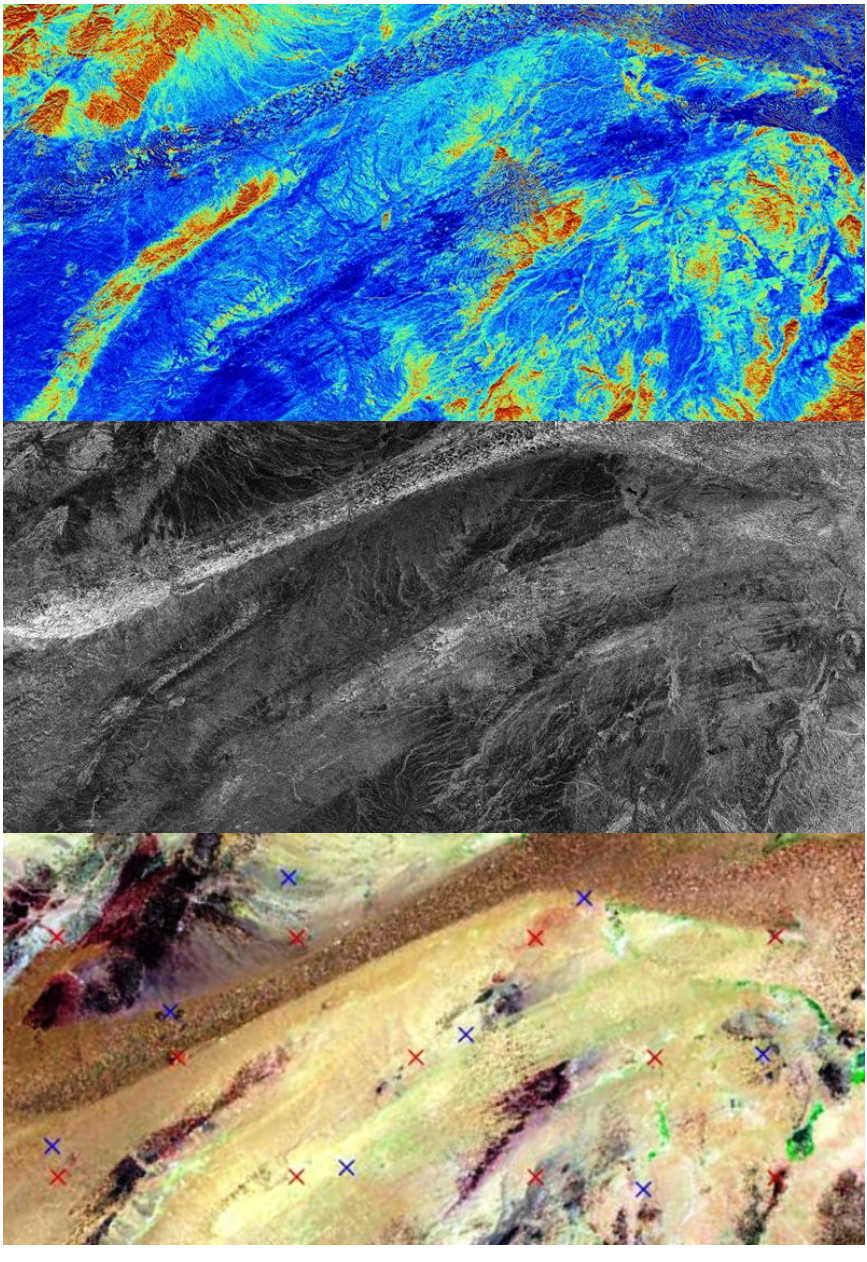


Fig. 8. (a) False color image of total σ°hh power in both C- and L-bands from data takes 13046/7. Blue is low, red is high. Center latitude is 40.52°N; center longitude is 104.43°E. Overlaps Figure 7b to southeast. Swath width about 45 km. Radar illumination from top. **(b)** Grayscale image of the ratio L-Band σ°hh to the sum of C-Band σ°hh plus L-Band σ°hh from data takes 13046/7 (Figure 8a). **(c)** Portion of spring 2004 Landsat TM mosaic corresponding to SIR-C data in Figs. 8a and b. North is towards upper left.

5 Populating the Similarity Model: Automatic Geometric Model from SAR Backscatter under Fuzzy Logic

Fuzzy sets and operators are the subjects and verbs of fuzzy logic. These if-then rule statements are used to formulate the conditional statements that comprise fuzzy logic. A single fuzzy if-then rule assumes the form:

IF x IS A THEN y IS B

where A and B are linguistic values defined by fuzzy sets on the ranges X and Y, respectively. The if-part of the rule "x is A" is called the antecedent or premise, while the then-part of the rule "y is B" is called the consequent or conclusion. An example of such a rule might be: *IF slope IS inclined THEN area IS suitable.*

The input to an if-then rule is the current value for the input variable (slope) and the output is an entire fuzzy set (suitable). This set will be later defuzzified, assigning one value to the output. Interpreting an if-then rule involves distinct parts: first evaluating the antecedent (which involves fuzzifying the input and applying any necessary fuzzy operators) and second applying that result to the consequent (known as implication). In the case of two-valued or binary logic, if-then rules don't present much difficulty. If the premise is true, then the conclusion is true. If the antecedent is true to some degree of membership, then the consequent is also true to that same degree. In this context the research first semantic define alluvial fans based on literatures. Then to complete the analysis the research had to create a geometric definition of alluvial fans. In this analysis the morphometric class is the geographic scale of measurement. By scale, the research mean a combination of spatial extent and spatial detail or resolution of the variation in the extent over which the feature is defined as the basis of the fuzzy membership (Band and Moore; 1995; Wang, 2000). Thus geomorphometric measures can indicate the location of a landform by assigning a location in a landscape to an exhaustive set of classes that could be run by an algorithm process. These different classes are illustrated below (Taramelli and Melelli 2008a; 2008b - Figure 9):

1) Altitude. The highest value of altitude is equal to the top of the apex of the highest fan while the lowest one is equal to the toes of the lowest alluvial fan. Because the range of values of altitudes differ in large areas, the break values to be considered have to be selected for each basin. The alluvial fans must be grouped in a well defined range of threshold altitude for each basin.

2) Convex contour. The shape of the alluvial fan shows typical boundaries where it is possible to define successive breaks of slope on a two-dimensional profile. A convex class referred to slope is recognizable in all the types of the alluvial fan and can be emphasized by contour lines. The same geometric shape is not recognizable in other debris accumulations (i.e. talus heaps) where the convexity attribute is random.

3) Increase in arc circumference. In plan view, the alluvial fans show a cone geometry where the boundary value may approximate natural division. In this context only two parameters, the values for the lower and upper boundaries are needed: the arc circumference increases from the apex (upper fan) to the toe (lower fan). The algorithm uses a linear relationship: while the altitude decreases, the arc circumference increases.

4) Convex-concave radial. Along the longitudinal profile an alluvial fan shows a convex – concave radial shape (from the top to the bottom) because the sediment grain size decreases toward lower altitudes values. Whereas gravels in the upper fans are associated with a convex profile, sands and clays are related to a concave profile in the bottom area. The algorithm first delineates the convex profile within the higher values of the altitudes and then the concave profile within the lower ones.

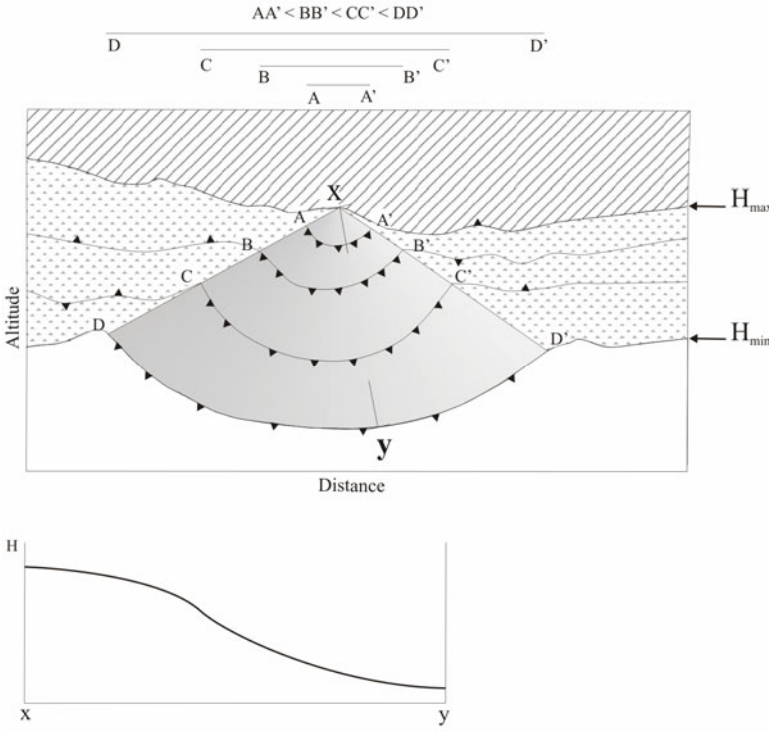


Fig. 9. Scheme of the geometric classes of alluvial fans (Tarameli and Melelli, 2008a; 2008b). The altitude class has a range between Hmax and Hmin. The convex contour class is highlighted by the direction of the triangles. The cone geometry is evident from the increase in the arc circumference from the segment AA' to DD'. The longitudinal profile (xy) shows a convex-concave radial shape from the top (Hmax) to the bottom (Hmin).

5.1 Delineation of the Alluvial Fans: Uses of the Similarity Model

Once a set of reliable signatures was created and evaluated, the next step was to perform a classification of the backscatter data with "Parametric/No-parametric" classifier to detect the range of altitude and curvature within the roughness signature using a probability function:

$$Pi = (2\pi)^{-1/2n} \mid Ci \mid^{-1/2} e^{-(1/2)[(X-Mi)'C^{-1}(X-Mi)]} \Pr i \tag{2}$$

Where:

Pi is Maximum likelihood probability of attribution to the class.
n Number of measurement variables.
Ci Covariance matrix of the class considered.
Mi Mean vector of the class considered.
X Pixel vector.
Pri Prior probability of the class considered defined from the frequency histograms of the training sets.

$$\Pr i = Fr / Frt \tag{3}$$

Where:

Fr is the pixel count of the class under examination.
Frt Is the sum of counts of all the classes.

In fuzzy theory, the algorithm has to define the class which exactly matches the core parameter set to assign a class membership of 1. The membership is assigned by a decreasing real number for classes as they are increasingly dissimilar from that core parameter until they have no similarity at all to the class. At that point the membership is assigned a value 0. The created algorithm has a specific sequence to assign the boundary value of class sets (Table 1).

The first selected parameter is the range of altitude values. Within this range of values the second assignment chooses only the convex contour shape. As a third boundary the algorithm selects only convex contours with an arc circumference that increases toward lower altitude. Finally, as a fourth boundary, convex-concave radial slope values are chosen. Ideally this approach maximize internal homogeneity and between-unit heterogeneity, and is characterized by unique groups of morphometric parameters. In that way the two studied areas were partitioned into topographic sections, adopting a semi-quantitative approach that combined an unsupervised eight-class cluster-analysis of derivatives. Finally, error matrices (Confusion Matrix) and the usual coefficient (Commission Error, Omission Error, Kappa Coefficient) were evaluated to quantify the different sets of ranges detected:

$$K = \frac{N\sum Xkk - \sum Xk_{+}X_{+}k}{N^{2} - \sum Xk_{+}X_{+}k} \tag{4}$$

Table 1. The heuristic rule base for converting initial terrain derivatives into fuzzy landform attributes

No.	Input terrain derivative	Output fuzzy landform attribute	Description of fuzzy landform attribute	Standard Index	Dispersion index
1	Elevation	Near_max	Relatively near maximum elevation	90.0	15.0
2	Elevation	Near_min	Relatively near minimum elevation	10.0	15.0
3	Profile	Concave_D	Relatively convex profile (down)	10.0	5.0
4	Profile	Convex_D	Relatively concave in profile (down)	-10.0	5.0
5	Profile	Planar_D	Relatively planar in profile (down)	0.0	5.0
6	Planar	Convex_A	Relatively convex in plan (across)	10.0	5.0
7	Planar	Concave_A	Relatively concave in profile (across)	-10.0	5.0
8	Planar	Planar_A	Relatively planar in profile (across)	0.0	5.0

6 Results and Conclusion

By analyzing data at different polarizations and look angles for backscatter, potentially derivable parameters were highlighted. While the elevation range is the basic surface pattern to detect roughness, the planar and radial curvature values are the primary attributes to detect alluvial fans surface change. Computation for the fuzzy k-means classification of the stratified random sample cells for the test area indicated that sensitivity analysis applied to the area illustrates a significant increase of object suitable for classification when roughness from the landform sources (alluvial fans) are decreased within the lower limit while the planar backscatter is increased.

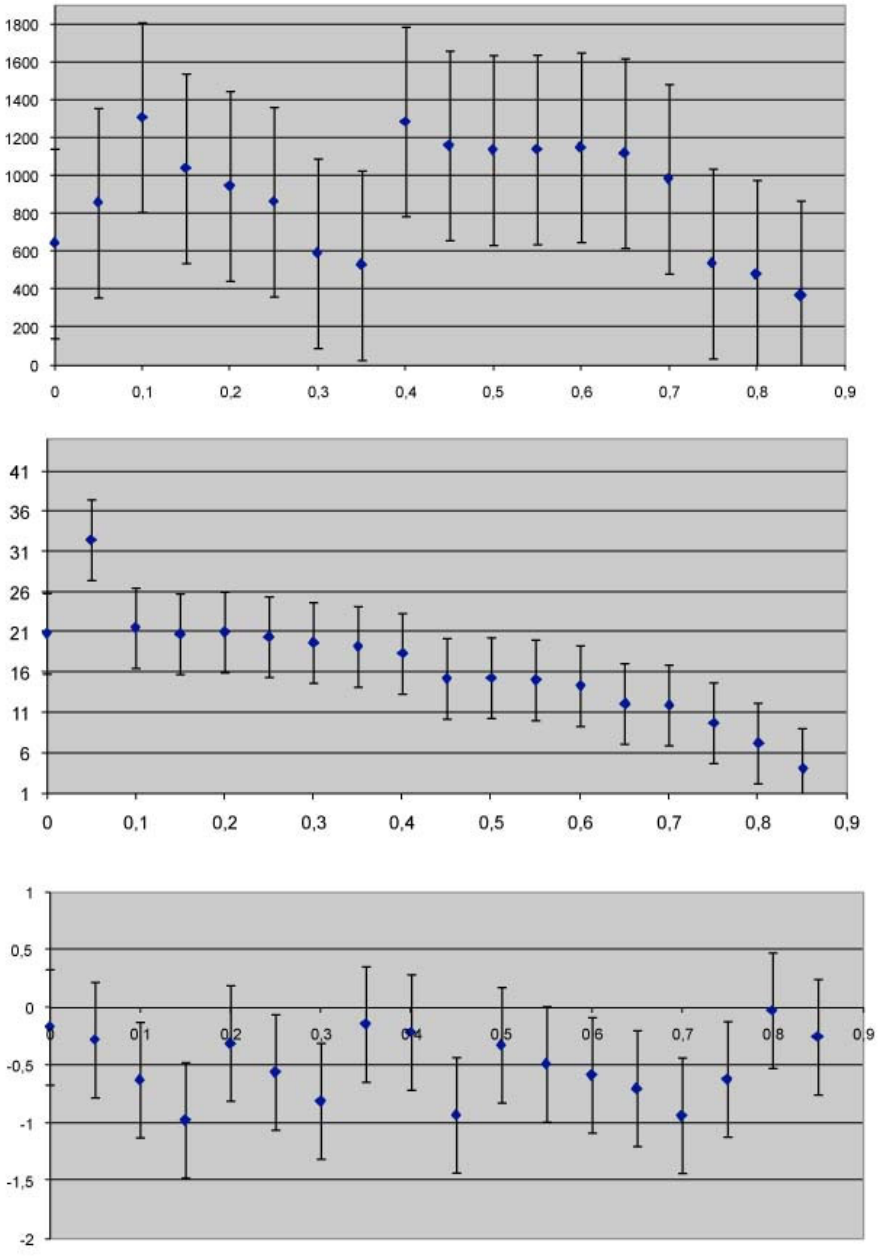


Fig. 10. The graphs shows the real value interval related to the classes related to the roughness memberships function of: Elevation (A), planar curvature (B) e radial curvature (C)

The alluvial features were first identified on Landsat satellite using SMAand SAR imagery. Topographic attributes can be divided into primary and secondary (compound) attributes. Primary attributes are directly calculated from elevation data and include variables such as elevation and slope. Compound attribute involve combinations of the primary attributes and are indices that describe or characterize the spatial variability of specific process occurring in the landscape such as soil water content distribution or the potential dust erosion. These attributes can be derived empirically but it is preferable to develop them through the application and simplification of the underlying physics of the processes. The algorithm used here is based on three topographic parameters related to roughness (altitude and two curvature) and shows the identification of alluvial areas characterized by a specific range of altitudes and curvature values (Figure 11).

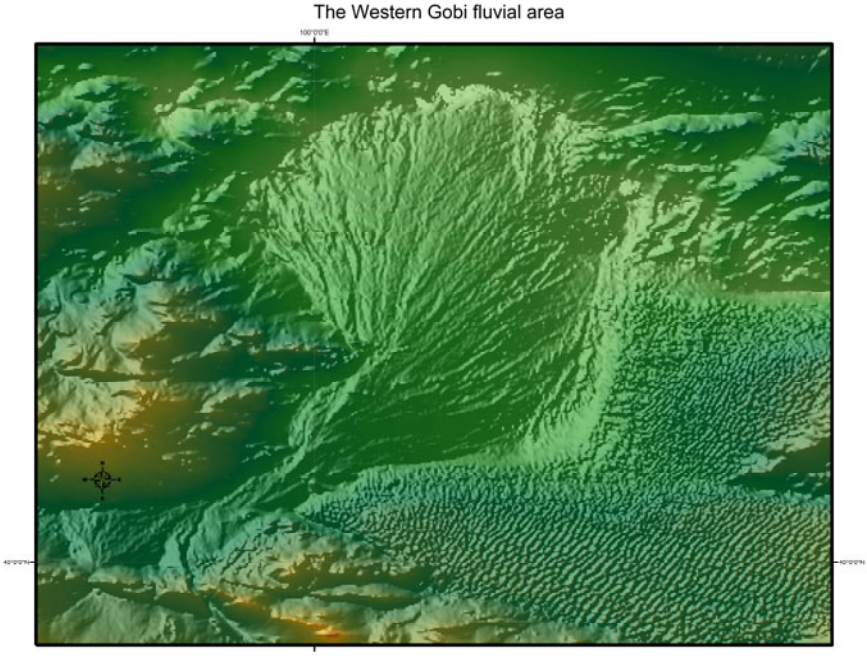


Fig. 11. Western Gobi region. In the central area is possible to see the huge alluvial area individuated with the algorithm.

This paper presents a new approach to answer the fundamental question of spatial information processing in geomorphology: can the simple knowledge of the spatial extent of a semantic defined alluvial fan with indistinct geographical locations, be improved with fuzzy memberships? This research has shown how alluvial fans phenomena can be defined using a novel fuzzy computer - based algorithm to processes the Sar backscatter. The geometric - morphometric analysis does not directly map alluvial fans, but highlights, the primary attributes (roughness, elevation and curvature) of an alluvial fan. Delineation of alluvial fans is then identified within an approximate spatial extent together with fuzzy

memberships. This paper has confirmed that fuzzy k-means classification of alluvial fans is then possible using data derived from SAR. In particular demonstrates that the procedure based on sampling to obtain training set and a classification scheme could be used with areas with low resolution calls. This analysis provides new insights and statements that can be made about to what extent type of fans and the morphology are dependent on the physical environment.

Acknowledgments. The author thanks Chris Small for introducing me into the Spectral Mixing Analysis methodology and Laura Melelli for implementing with me the alluvial fan test algorithm within the intermontane basin in Umbria - Italy.

References

Adams, J.B., Smith, M.O., Johnson, P.E.: Spectral mixture modeling; a new analysis of rock and soil types at the Viking Lander 1 site. JGR 91, 8098–8122 (1986)

Adams, J.B., Sabol, D.E., Kapos, V., Filho, R.A., Roberts, D.A., Smith, M.O., Gillespie, A.R.: Classification of multispectral images based on fractions of endmembers: Application to land cover change in the Brazilian Amazon. RSE 52, 137–154 (1995)

Adams, J.B., Smith, M.O., Gillespie, A.R.: Imaging Spectroscopy: Interpretation based on spectral mixture analysis. In: Pieters, C.M., Englert, P. (eds.) Remote Geochemichal Analysis: Element and Mineralogical Composition, pp. 145–166. Cambridge University Press, New York (1993)

Allievi, J., Ambrosi, C., Cerini, M., Colesanti, C., Crosta, G.B., Ferretti, A., Fossati, D.: Monitoring slow mass movements with the Permanent Scatterers technique, pp. 07803–79306. IEEE, Los Alamitos (2003)

Ahamed, T.R., Gopal Rao, K., Murthy, J.S.R.: GIS-based fuzzy membership model for crop-land suitability analysis. Agric. Syst. 63, 75–95 (2000)

Antonello, G., Casagli, N., Farina, P., Leva, D., Nico, G., Sieber, A.J., Tarchi, D.: Ground-based SAR interferometry for monitoring mass movements. Lands. 1, 21–28 (2004)

Argialas, D.P., Tzotsos, A.: Geomorphological feature extraction from a digital elevation model trough fuzzy knowledge-based classification. In: Ehlers, M. (ed.) Proc. SPIE, Remote Sensing for environmental monitoring, gis application and geology, 2, vol. 4886, pp. 516–527 (2003)

Bach, D., Barbour, J., Macchiavello, G., Martinelli, M., Scalas, P., Small, C., Stark, C., Taramelli, A., Torriano, L., Weissel, J.: Integration of the advanced remote sensing technologies to investigate the dust storm areas. In: El-Beltagy, A., Saxena, M.C., Wang, T. (eds.) Proceedings of 8th ICCD Conference on Human and Nature – Working tighter for sustainable development of Drylands, Beijing, China, February 25-28, pp. 387–397. ICARDA, Aleppo (2007)

Band, L.E., Moore, I.D.: Scale: landscape attributes and Geographical Information Systems. Hydrol. Porces. 9, 401–422 (1995)

Benallegue, M., Taconet, O., Vidal-Madjar, D., Normand, M.: The Use of Radar Backscattering Signals for Measuring Soil Moisture and Surface Roughness. RSE 53(1), 61–68 (1995)

Blumberg, D.G.: Analysis of large Aeolian (wind-blown) bedforms using the SRTM digital elevation data. RSE 100, 179–189 (2006)

Boardman, J.W.: Inversion of Imaging Spectrometry Data Using Singular Value Decomposition. In: Porceedings of the 12th Canadian Symposium on Remote Sensing, pp. 2069–2072 (1989)

Boardman, J.W.: Automating spectral unmixing of AVIRIS data using convex geometry concepts. In: Green, R.O. (ed.) Fourth Airborne visible/infrared Imaging Spectrometer (Aviris) Airborne Geoscience Workshop, pp. 11–14. Jet Propulsion Laboratory, Pasadena (1993)

Boardman, J.W., Kruse, F.A.: Automated spectral analysis: a geologic example using AVIRIS data, north Grapevine mountains, Nevada. In: 10th Thematic Conference on Geologic Remote Sensing, pp. I407–I418. Environmental Research Institute of Michigan, Ann Arbor (1994)

Burrough, P.A., MacMillan, R.A., Van Deursen, W.: Fuzzy classification methods for determining land suitability from soil profile observations. J. Soil Science 43, 193–210 (1992)

Burrough, P.A., van Gaans, P., Hootsmans, R.: Continuous classification in soil survey: Spatial correlation, confusion and boundaries. Geoderma 77, 115–135 (1997)

Bull, W.B.: The alluvial fan environment. Progres. Physic. Geo. 1, 222–270 (1977)

Cobb, M., Petry, F., Robinson, V.: Special issue: uncertainty in geographic information systems and spatial data. Fuzzy Set and System 113(1), 1–159 (2003)

Cross, V., Firat, A.: Fuzzy objects for geographical information systems. Fuzzy Set and System 113(1), 19–36 (2000)

Dehn, M., Gärtner, H., Dikau, R.: Principles of Semantic Modeling of Landform Structures. Comp. & Geosc. 27, 1005–1010 (2001)

Dierking, W.: Quantitative Roughness Characterization of Geological Surfaces and Implications for Radar Signature Analysis. IEEE 7(5), 2397–2412 (1999)

Dikau, R.: The application of a digital relief model to landform analysis in geomorphology. In: Raper, J. (ed.) Three-Dimensional Applications in Geographical Information Systems, pp. 51–77. Taylor and Francis, London (1989)

Dubois, D., Prade, H.: What are fuzzy rules and how to use them. Fuzzy Set and System 84(2) (1996)

Elmore, A.J., Mustard, J.F., Manning, S.J., Lobell, D.B.: Quantifying vegetation change in semiarid environments: Precision and accuracy of spectral mixture analysis and the normalized difference vegetation index. RSE 73, 87–102 (2000)

Fairbridge, R.W.: The ecyclopedia of Geomorphology. Reinhold Book Corp., New York (1968)

Fisher, P.F.: Fuzzy modeling. In: Openshaw, S., Abrahart, R., Harris, T. (eds.) Geocomputing. pp. 161–186. Taylor and Francis, London (2000)

Fisher, P.F.: Sorties paradox and vague geographies. Fuzzy Set and System 113, 7–18 (2000)

Gillespie, A.R., Smith, M.O., Adams, J.B., Willis, S.C., Fischer III, A.F., Sabol, D.E.: Interpretation of residual images: Spectral mixture analysis of AVIRIS images. In: 2nd Airborne Visible/Infrared Imaging Spectrometer (Aviris) Airborne Geoscience Workshop, Owens Valley, California. NASA Jet Propulsion Laboratory, Pasadena (1990)

Greeley, R., Blumberg, D.G., McHone, J.F., Dobrovolkis, A., Iversen, J.D., Rasmussen, K.R., Wall, S.D., White, B.R.: Applications of spaceborne radar laboratory data to the study of aeolian processes. JGR 102(10), 971–983 (1997)

Greeley, R., Blumberg, D.G., Dobrovolskis, A.R., Gaddis, L.R., Iversen, J.D., Lancaster, N., Rasmussen, K.R., Saunders, R.S., Wall, S.D., White, B.R.: Potential transport of windblown sand: Influence of surface roughness and assessment with radar data. In: Tchakerian, V.P. (ed.) Desert Aeolian Processes, pp. 75–99. Chapman & Hall, London (1995)

Johnson, P.E., Smith, M.O., Taylor-George, S., Adams, J.B.: A semiempirical method for analysis of the reflectance spectra of binary mineral mixtures. JGR 88, 3557–3561 (1983)

Imhoff, M.L., Vermillion, C., Story, M.H., Choudhury, A.M., Gafoor, A., Polcyn, F.: Monsoon flood boundary delineation and damage assessment using spaceborne imaging radar and Landsat data. Photog. Eng. and rem. Sens. 53(4), 405–413 (1987)

Irvin, B.J., Ventura, S.J., Slater, B.K.: Fuzzy and isodata classification of landform elements from digital terrain data in Pleasant Valley, Wisconsin. Geoderma 77, 137–154 (1997)

Kobrick, M.: On the toes of Giants – How SRTM was born. Photog. Eng. and rem. Sens. 3, 206–210 (2006)

Laurent, B., Marticorena, B., Bergametti, G., Chazette, P., Maignan, F., Schmechtig, C.: Simulation of the mineral dust emission frequencies from desert areas of China and Mongolia using an aerodynamic roughness length map derived from the POLDER/ADEOS I surface products. Jour. Gephys. Res. 110, 21 (2005), doi:10.1029/2004JD005013

Ladner, R., Petry, F., Cobb, M.: Fuzzy set approaches to spatial data mining of association rules. Transaction in GIS 7(1), 123–138 (2000)

Leeder, M.R.: Sedimentary basins: tectonic recorders of sediment discharge from drainage catchments. Earth surf. Proces. Landf. 22, 229–237 (1977)

Li, L., Ustin, S.L., Lay, M.: Application of multiple endmember spectral mixture analysis (MESMA) to AVIRIS imagery for coastal salt marsh mapping: a case study in China Camp, CA, USA. Int. Journ. of Rem. Sens. 26(23), 5193–5207 (2005)

Lu, D., Weng, Q.: Spectral mixture analysis of ASTER images for examining the relationship between urban thermal features and biophysical descriptors in Indianapolis, Indiana, USA. RSE 104, 57–167 (2006)

Nemec, W., Steel, R.J.: What is a fan delta and how we recognize it? In: Nemec, W., Steel, R.J. (eds.) Fan Deltas: sedimentology and tectonic settings. pp. 3–13. Blackie and Son, Glasgow & London (1988)

McBratney, A.B., De Gruijter, J.J.: A continuum approach to soil classification by modified fuzzy k-mean with extragrades. J. Soil Scince 43, 159–175 (1992)

McBratney, A.B., Odeh, I.O.A.: Application of fuzzy sets in soil science: Fuzzy logic, fuzzy measurements and fuzzy decisions. Geoderma 77, 85–113 (1997)

MacMillan, R.A., Pettapiece, W.W., Nolan, S.C., Goddard, T.W.: A generic procedure for automatically segmenting landforms into landform elements using DEMs, heuristic rules and fuzzy logic. Fuzzy Set and System 113, 81–109 (2000)

Melelli, L., Taramelli, A.: An example of debris-flows hazard modeling using GIS. Nat. Haz. Erath Syst. Scien. 4, 347–358 (2004)

Nash, E.B., Conel, J.E.: Spectral reflectance systematics for mixtures of powdered hypersthene, labradorite and ilmenite. JGR 79, 1615–1621 (1974)

Pike, R.J.: Geomorphometry – diversity in quantitative surface analysis. Prog. In Phy. Geogr. 24, 1–20 (2000)

Prigent, C., Tegen, I., Aires, F., Marticorena, B., Zribi, M.: Estimation of aerodynamic roughness length in arid and semi-arid regions over the globe with the ERS scatterometer. JGR 110, 12 (2005), doi:10.1029/2004JD005370

Rabus, B., Eineder, M., Roth, A., Bamler, R.: The shuttle radar topography mission- a new class of digital elevation models acquired by spaceborne radar. Jour. of Photog. and Rem. Sens. 57, 241–262 (2003)

Roberts, D.A., Batista, G., Pereira, J., Waller, E., Nelson, B.: Change Identification using Multitemporal Spectral Mixture Analysis: Applications in Eastern Amazonia. In: Elvidge, C., Lunetta, R. (eds.) Remote Sensing Change Detection: Environmental Monitoring Applications and Methods, pp. 137–161. Ann Arbor Press, Ann Arbor (1998a)

Roberts, D.A., Gardner, M., Church, R., Ustin, S., Scheer, G., Green, R.O.: Mapping chaparral in the Santa Monica mountains using multiple endmember spectral mixture models. RSE 65, 267–279 (1998b)

Robinson, V.B.: A Perspective on Geographic Information Systems and Fuzzy Sets. IEEE, 1–6 (2002), 0-7803-74614102

Robinson, V.B.: Some implications of fuzzy set theory applied to geographic databases Computers. Environment and Urban Systems 12, 89–97 (1988)

Sainsbury, R.M.: What is a vague object? Analysis 49, 99–103 (1989)

Settle, J.J., Drake, N.A.: Linear mixing and the estimation of ground cover proportions. Int. Journ. of Rem. Sens. 14(6), 1159–1177 (1993)

Singer, R.B., McCord, T.B.: Mars: Large scale mixing of bright and dark surface materials and implications for analysis of spectral reflectance. In: Proceedings of the 10th Lunar and Planetary Science Conference, AGU, pp. 1835–1848 (1979)

Small, C.: Estimation of urban vegetation abundance by spectral mixture analysis. Int. Journ. of Rem. Sens. 22(7), 1305–1334 (2001)

Small, C.: The Landsat ETM+ Spectral Mixing Space. RSE 93, 1–17 (2004)

Smith, M.O., Johnson, P.E., Adams, J.B.: Quantitative determination of mineral types and abundances from reflectance spectra using principal component analysis. JGR 90, 792–804 (1985)

Smith, M.O., Ustin, S.L., Adams, J.B., Gillespie, A.R.: Vegetation in deserts: I. A regional measure of abundance from multispectral images. RSE 31, 1–26 (1990)

Speight, J.G.: The role of topography in controlling through-flow generation: a discussion. Earth Surf. Proces. Landf. 5, 187–191 (1984)

Taramelli, A., Melelli, L.: Contribution of the Landsat ETM+ Spectral Mixing Space and SRTM analysis to characterize Deep Seated Gravitational Slope Deformations in central Apennines (Italy). Int. Journ. of Rem. Sens. 30(2), 357–387 (2009)

Taramelli, A., Melelli, L.: Detecting alluvial fans using quantitative roughness characterization and fuzzy logic analysis. In: Gervasi, O., Murgante, B., Laganà, A., Taniar, D., Mun, Y., Gavrilova, M.L. (eds.) ICCSA 2008, Part I. LNCS, vol. 5072, pp. 1–15. Springer, Heidelberg (2008a)

Taramelli, A., Melelli, L.: Detecting alluvial fans using quantitative roughness characterization and fuzzy logic analysis using the SRTM data. IJCSST (2008b) (invited Paper) (in press)

Ulaby, F.T., Moore, R.K., Fung, A.K.: Microwave remote sensing, vol. 2. Addison-Wesley, Reading (1982)

Usery, E.L.: A conceptual framework and fuzzy set implementation for geographic features. In: Burrough, P.A., Frank, A. (eds.) Geographic Objects With Indeterminate Boundaries, pp. 87–94. Taylor & Francis, London (1996)

Wang, F.: A fuzzy grammar and possibility theory-based natural language user interface for spatial queries. Fuzzy Set and System 113(1), 147–159 (2000)

Wood, W.F., Snell, J.B.: A quantitative system for classifying landforms. U.S. Army Quartermaster Research and Engineering Center, Natick, MA, Tech. Rep. EP-124 (1960)

Zadeh, L.A.: Toward a perception-based theory of probabilistic reasoning with imprecise probabilities. Jour. Stat. Plann. and Inference 105(1), 105–119 (2002)

Zhu, A.X.: A similarity model for representing soil spatial information. Geoderma 77, 217–242 (1997a)
Zhu, A.X.: Measuring uncertainty in class assignment for natural resource maps using a similarity model. Photog. Eng. and Rem. Sens. 63, 1195–1202 (1997b)
Zhu, A.X., Band, L.E., Vertessy, R., Dutton, B.: Deriving soil property using a soil land inference model (SoLIM). Soil Sci. Soc. Am. J. 61, 523–533 (1997)
Zhu, A.X., Band, L.E., Dutton, B., Nimlos, T.: Automated soil inference under fuzzy logic Ecol. Modell 90, 123–145 (1996)

A Framework of Map Comparison Methods to Evaluate Geosimulation Models from a Geospatial Perspective

Alex Hagen-Zanker[1,2] and Pim Martens[3]

[1] Research Institute for Knowledge Systems, P.O. Box 463, 6200 AL Maastricht, The Netherlands
[2] Urban Planning Group, Technical University Eindhoven, PO Box 513, 5600 MB Eindhoven, The Netherlands
[3] International Centre for Integrated assessment and Sustainable development, Maastricht University, PO Box 616, 6200 MD Maastricht, The Netherlands
e-mail: ahh34@cam.ac.uk

Abstract. Geosimulation is a form of microsimulation that seeks to understand geographical patterns and dynamics as the outcome of micro-level geographical processes. Geosimulation has been applied to understand such diverse systems as lake ecology, traffic congestion and urban growth. A crucial task common to these applications is to express the agreement between model and reality and hence the confidence one can have in model results. Such evaluation requires a geospatial perspective; it is not sufficient if micro-level interactions are realistic. Importantly, interactions should be such that meso- and macro- level patterns emerging from the model are realistic. In recent years, a host of map comparison methods have been developed, which address different aspects of the agreement between model and reality. This paper places such methods in a framework to systematically assess breadth and width of model performance. The framework expresses agreement at the continuum of spatial scales ranging from local to whole landscape and separately addresses agreement in structure and presence. A common reference level makes different performance metrics mutually comparable and guides the interpretation of results. The framework is applied for the evaluation of a constrained cellular automata model of the Netherlands. The case demonstrates that a performance assessment lacking either a multi-criteria and multi-scale perspective or a reference level would result in an unbalanced account and ultimately false conclusions.

Keywords: geosimulation, calibration, validation, map comparison.

1 Introduction

Geosimulation is a field of geography that seeks to understand geographical patterns and dynamics as the consequence of the interactions between individual

B. Murgante et al. (Eds.): Geocomputation, Sustainability & Environ. Planning, SCI 348, pp. 251–269.
springerlink.com

entities, like tenants, land owners, car drivers, trees, etc. Geosimulation models only prescribe the behaviour of these entities and not the resulting large scale patterns, such as segregation, urban sprawl, road congestion, forest fire, etc. These emerge as a product of the interactions between entities. The field of geosimulation is spurred by advances in computing. As a consequence, the average desk computer can function as a virtual laboratory, where researchers can grow their own virtual cities, transport systems, rural societies, forests, etc. (Benenson and Torrens 2004).

Geosimulation models are increasingly finding region specific applications. Rather than using the models to grow geographical systems from scratch, the models are fed with a real initial situation. Models are then not just theoretical constructs, but have a more applied nature. A striking illustration is the contrast between the Schelling model (Schelling 1971) and the entity-based model of urban residential dynamics for the Yaffa area in Tel Aviv (Benenson et al. 2002). Both models are concerned with segregation and the underlying microdynamics, but the level of detail and the lessons that can be drawn differ considerably.

The constrained cellular automata land use model that is evaluated as a case in this paper is a good example of a model that has developed from a theoretical model to a practical tool. The first application of the model concerned an imaginary island with characteristics typical for Caribbean islands (Engelen et al. 1995). Later applications focused on the city of Cincinnati Metropolitan Area on a timescale of more than 100 years (White et al. 1997). Further developments have elaborated the use of GIS data, including road network data and the dynamic integration with socio-economic land use models at multiple scales (White and Engelen 2000). The model became part of Policy Support Systems, such as the Environment Explorer (Engelen et al. 2003; de Nijs et al. 2004). Currently the model is the cornerstone of several modelling frameworks of urban and regional growth, meaning that new regional applications can be setup within hours, including METRONAMICA (van Delden and Engelen 2006) and MOLAND (White et al. 2000; Barredo and Demicheli 2003) frameworks.

The problem that confronts the new generation of geosimulation modellers is to assess how well their virtual worlds correspond to reality. ‘Good modelling practice’ (Refsgaard and Henriksen 2004) prescribes different analytical steps. Of these, calibration and validation require an expression of goodness-of-fit of the model. Since results of the models are typically maps, it makes perfect sense to address this question by map comparison; however the nature of geosimulation models provides some particular challenges that need to be considered.

One problem is that the resolution at which the model is defined is not equal to the scale at which the results are interpreted. The interest in the models lies in the geographical structures that unroll as a consequence of the interactions of the individual entities. The concept of complexity is relevant here. Typically the elements in the model are mutually dependent, causing feedback processes and self-organization, but also making the models sensitive to small deviations to the extent that they are chaotic or unpredictable. The consequence is that even a geosimulation model that perfectly captures the dynamics of a geographical

system cannot be expected to produce result maps that correspond perfectly to reality. Models should therefore not be evaluated just at the location-to-location level, but in terms of the patterns that emerge. On the other hand; when applying a model for a particular region, one is not just interested in the global patterns, but also how the patterns are distributed in space.

A balance is sought between finding realistic patterns and finding them at the right location; the geosimulation model should create spatial configurations that are similar to reality and place them in approximately the right locations. Existing comparison methods do not strike such a balance. With few exceptions, they are either local and based on cell-by-cell overlap, or global and based on metrics summarizing the whole landscape in a single value. Despite the lack of formal methods, an expert can make this kind of balanced comparison by just looking at the map. It is therefore not surprising that in practice geosimulation models are often evaluated on the basis of such face validation. (Batty and Torrens 2005)

There are several problems associated to face validation, most pressingly the lack of objective reproducibility. A practical concern of face validation is that for some tasks, such as calibration, large numbers of consistent assessments are required. Depending on a human judge of map similarity may be too time-consuming, costly and prone to inconsistencies.

Another challenge is posed by the dynamical nature of the simulation models and the relatively small number of changes that may occur over a simulation period. It is, for instance, not uncommon for land use models to attain percentages of agreement between reality and model above 95%. This ‘good performance’ is then invariably due to the fact that land use patterns at the beginning of the simulation period are highly similar to those at the end (Pontius Jr. et al. 2008). This causes a real risk of misinterpretation and false confidence in the results of geosimulation models.

This paper introduces a framework to evaluate model performance. It applies a number of statistics that can be categorized according to two axes. The first axis is typically recognized in geographical information science and is based on the spatial unit of the analysis; it ranges from local, via focal to global. The second axis is more commonly applied in landscape ecological applications and discerns whether the presence of certain classes is considered or their spatial structure. The interpretation of the results is guided by a neutral model of landscape change which provides a common reference for all metrics (Hagen-Zanker and Lajoie 2008). This reference model is subject to the same initial situation and constraints but does not represent the processes that characterize the evaluated model. The difference in performance between the two models can thus be attributed to those processes.

The purpose of this framework is in first instance to provide an order in the large variety of performance metrics. Secondly, the framework can be a guide towards a comprehensive approach of performance assessment in the calibration and validation process of geosimulation models.

2 Method

2.1 The Axis: Local, Global, Focal

Geographical Information System (GIS) operations can be classified as being local, focal or global in nature (Takeyama and Couclelis 1997). Local statistics relate a particular location on the map, in the case of raster maps: a cell. Analysis based on local operations is often called overlay analysis. Focal statistics have the focus on one location, but also take location neighbourhood into account. Typical focal operations are spatial smoothing functions and density estimation. Focal operations are also called moving window, filter (typically for small neighbourhoods) or kernel operations (typically for large neighbourhoods). Global operations are based on aggregates over the whole map.

In the past, map comparison methods that are used as performance criteria for spatial models have been either local (e.g. Monserud and Leemans 1992; Pontius Jr. 2000) or global. Global analysis includes fractal dimension (Batty and Longley 1994; White 2006), cluster size distributions (Dungan 2006) and landscape metrics (Turner et al. 1989; Barredo and Demicheli 2003). The disadvantage of these methods is that either spatial structure (local) or spatial specificity (global) is ignored. The continuum from local to global has been investigated by multi-scale analysis on the basis of step-wise aggregation model results and data (Costanza 1989; Kok et al. 2001; Pontius Jr. 2002) and wavelet decomposition (Briggs and Levine 1997). Both aggregation and wavelet based approaches however, suffer under the rather arbitrary positioning of the coarse scale grid relative to the original grid.

Methods in this paper consider the whole spectrum of local to global operations. The emphasis, however, is on focal operations. Not just because these have received little attention in the past, but primarily because focal operations are the ideal means to simultaneously investigate similarity in structure and location-to-location correspondence.

The use of local or focal statistics does not imply that map comparison results are only presented at the local level, i.e. as a comparison map. Local agreement can be aggregated to global statistics. The crucial distinction is whether a metric compares local, focal or global attributes.

2.2 The Axis: Presence and Structure

The field of landscape ecology (Turner 2005) studies the relationship between landscape structure and ecological processes. One of the major considerations is that inference about spatial structure is only possible if it is objectively quantified. Consequently many metrics of landscape structure have been introduced and analyzed.

In particular two types of spatial structure are recognized: composition and configuration. Metrics of composition are based on the fraction of occurrence of land use/land cover classes at the global level. Examples of composition metrics

are diversity indicators. Metrics of configuration relate to the spatial positioning of land use classes relative to each other. Configuration metrics are often calculated at patch level. Patches are contiguous areas of a class. In a geographical context, patches are often called clusters. Examples of patch level structure metrics are size, perimeter, shape index and fractal dimension. Other configuration metrics which not based on patches include edge and contagion index. Configuration metrics are common in landscape ecology (McGarigal et al. 2002) and they are sporadically used as performance criteria for geosimulation models of urban dynamics (Turner et al. 1989; Barredo and Demicheli 2003).

In landscape ecological studies, metrics are typically applied and analyzed at landscape scale, i.e. global. When landscape metrics are applied at the focal level, the distinction between composition and configuration blurs, since focal composition depends on global configuration. For example, focal assessments of patch size will correlate strongly with focal assessments of entropy, depending on the size of the focal window, , since large patches lead to homogenous areas. Therefore, this paper re-emphasizes the distinction between configuration and composition to that between structure and presence. Similarity in structure is achieved when structure metrics (either of configuration or composition) describing two maps, focal windows or locations, are similar. Similarity in presence is location specific and is achieved when the compositions of two locations or focal windows are similar.

2.3 Comparison Methods

Considering the two axes, six classes of map comparison methods could possibly be identified. However it is not possible to recognize spatial structure locally, at the level of a single cell, . Otherwise, at the global level all notions of location are lost and therefore the notion of presence is obsolete. Thus, four classes of map comparison methods remain. Table 1 presents the comparison methods of this paper and places them in the framework of the two axes.

Table 1. Overview of comparison methods applied in this paper

	Local	Focal	Global
Presence	Kappa	Moving Window Euclidean	-
Structure	-	Moving Window Patch size	Cluster size distribution

2.3.1 Local Presence: Kappa Statistics

Cell-by-cell map comparison methods consider two compared maps as a number of paired observations. Each paired observation consists of the classes found at one cell in the compared maps. Apart from direct overlap, spatial structure is not considered and the full information available to cell-by-cell methods can therefore be tabulated in the contingency table. This matrix tabulates for each pair of classes how often it occurs. Table 2 gives the generic form. It is well established as a cornerstone of accuracy assessment (Foody 2002).

An obvious metric of map correspondence is the fraction of agreement, which is the fraction of all observed pairs where the first and second class are identical. The more often used cell-by-cell metric however is Kappa. This statistic corrects the fraction of agreement for the fraction of agreement that can be expected given the number of cells of each class. For instance, let us consider two maps which are both 80% forest and 20% desert. A random spatial distribution of these quantities over the map would be at least 60% identical and the expected agreement of these maps is $0.8^2 + 0.2^2 = 0.68$.

Table 2. Contingency table (t_{ij} is the number of cells of class i in map A and class j in map B. t_{i+} is the number of cells of class i in map A. t_{+i} is the number of cells of class j in map B. t_{++} is the total number of cells)

Map A \ Map B	1	2	⋯	c	Sum
1	t_{11}	t_{12}		t_{1c}	t_{1+}
2	t_{21}	t_{22}	⋯	t_{2c}	t_{2+}
⋮	⋮	⋮	⋱	⋮	⋮
c	t_{c1}	t_{c2}	⋯	t_{cc}	t_{c+}
Sum	t_{+1}	t_{+2}	⋯	t_{+c}	t_{++}

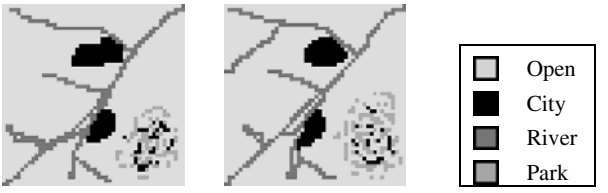
The following equations express how Kappa is calculated from the contingency table:

$$K = \frac{P(A) - P(E)}{1 - P(E)} \tag{1}$$

$$P(A) = \frac{1}{t_{++}} \sum_{i=1}^{c} t_{ii} \tag{2}$$

$$P(E) = \frac{1}{t_{++}} \sum_{i=1}^{c} t_{i+} t_{+i} \tag{3}$$

where K is Kappa, $P(A)$ is the fraction of agreement and $P(E)$ is the expected fraction of agreement. A K of 0 corresponds to the expected level of agreement, identical maps get value 1 and the lowest possible score is -1. The Kappa statistic as originally introduced (Cohen 1960) is not intended as a map comparison metric, but as a general method of evaluating agreement of paired observations. Interestingly the same method was already introduced decades earlier for the purpose of comparing weather forecast maps (Heidke 1926). The metric is therefore also known as the Heidke Skill Score.



a. Map 1 b. Map 2.

Fig. 1. Example pair of land use maps

Further developments have extended the use of Kappa metrics; Monserud and Leemans (1992) calculated Kappa values for individual classes by temporary reclassifying the maps to binary maps. Figure 1 shows two maps; the corresponding contingency table and Kappa statistics are given in tables 3 and 4. Kappa results indicate that the two maps are most similar in terms of the class city and least similar for the class park, despite the fact that the contribution to the fraction of agreement of that class is very small. The fact that all cells of the class park are contained within the same (bottom right) region of the map is not recognized by the Kappa statistic, but will be by the focal statistics discussed in the following sections.

Table 3. Example of contingency table

Map 1 \Map 2	Open	River	City	Park	Total	Map 1
Open	1767	61	119	48	1995	
City	32	92	18	8	150	
River	154	1	96	0	251	
Park	74	11	0	19	104	
Total Map 2	2027	165	233	75	2500	

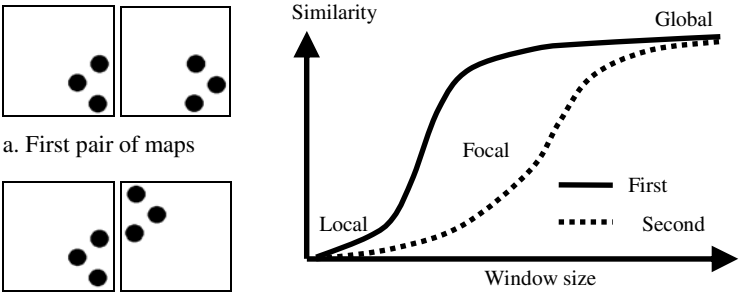
Table 4. Example of Kappa statistics

	Overall	Open	City	River	Park
P(A)	0.79	0.71	0.037	0.038	0.0076
P(E)	0.66	0.65	0.0040	0.0094	0.0012
Kappa	0.94	0.94	0.97	0.96	0.72

2.3.2 Focal Presence: Moving Window Euclidean Distance

As the previous section explains, cell-by-cell comparison methods do not consider spatial structure except for direct overlap. Many geosimulation models however, including the model evaluated in this paper, are not expected to achieve such precision in their predictions. If the model places land use classes approximately in the right location, it can already be considered to perform well. With the purpose of considering such near agreement, a moving window approach is followed comparing window composition in both maps.

The idea of this method is counting differences which are mitigated in the close neighbourhood only as small errors. In other words, an over-prediction of a class at one location can partially compensate the under-prediction of that same class at a close-by location. A larger window to detect such mitigating errors will lead to smaller errors. Figure 2 shows two pairs of maps that in a cell-by-cell approach would be considered fully distinct (Kappa = -1), but with increasing window sizes more of the similarity between the two maps is recognized. When the moving window covers the whole map, the two maps are considered identical. The crucial point is that with increasing window size, the similarity of the first map is recognized earlier (or is stronger) than that of the second map. It is apparent that focal statistics provide an insight in the nature of the agreement of both pairs that neither local nor global statistics can provide.



a. First pair of maps

b. Second pair of maps

c. Similarity as function of window size

Fig. 2. Two pairs of maps which are both considered fully distinct from the local perspective and identical from the global one. Balanced analysis should find the first pair more similar (source: (Hagen-Zanker 2006)).

The moving window Euclidean distance metric is calculated as the mean Euclidean distance on the basis of the proportions of each land use class over all cells in the map, as follows:

$$\bar{\mathrm{E}} = \frac{1}{\mathrm{n}}\sum_{\mathrm{i=1}}^{\mathrm{n}} \mathrm{E_i} = \frac{1}{\mathrm{n}}\sum_{\mathrm{i=1}}^{\mathrm{n}} \sqrt{\sum_{\mathrm{j=1}}^{\mathrm{c}} \left(\mathrm{p}_{\mathrm{i,j}}^{\mathrm{A}} - \mathrm{p}_{\mathrm{i,j}}^{\mathrm{B}} \right)^2} \tag{4}$$

where i iterates over all n cells on the map and j iterates over all c classes in the legend. E_i is the Euclidean distance of a moving window centered on cell i. $\bar{E}$ is the mean Euclidean distance. $p^A_{i,j}$ is the fraction of cells within a circular window around cell i that is in class j in map A (likewise for map B).

E_i values can be visualized as a map, and like the Kappa metrics presented in the previous section these metrics can be calculated for individual classes as well as for all classes combined. Figure 3 and table 5 present the results for the example introduced in figure 1. The results illustrate how errors decrease with increasing window sizes . Note that for class park the error decreases most strongly with increasing radius, because that class is dominated by small errors.

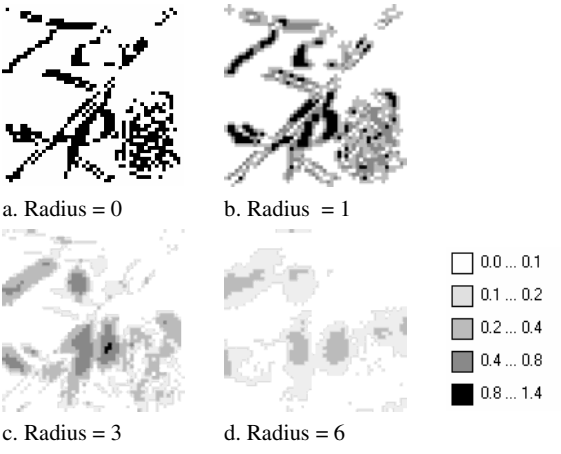


Fig. 3. Visualization of Euclidean distance metric for the example maps. Euclidean distance decreases and images blur more with increasing radius.

Table 5. Example of Euclidean distance statistics

Radius	Overall	Open	City	River	Park
0	0.30	0.28	0.074	0.16	0.080
1	0.21	0.19	0.065	0.12	0.049
3	0.13	0.12	0.050	0.074	0.028
6	0.081	0.072	0.035	0.044	0.020

2.3.3 Focal Structure: Moving Window Patch Size

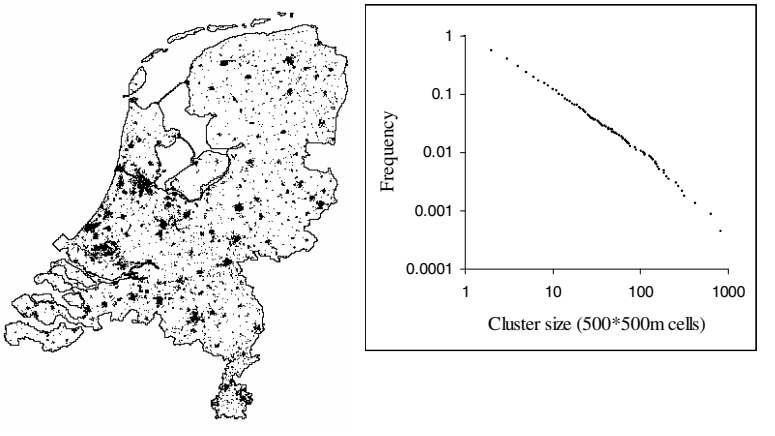
Land use models are not only expected to place land use classes at (approximately) the right location, but also to arrange classes in the right structure. As the strength of these models is capturing the processes underlying spatial morphology, they may even be expected to perform better in terms of structure than presence. There are many indicators of spatial structure. A commonly used metric, because of its straightforward interpretation, is patch size. A patch is also called a cluster and consists of all contiguous raster cells of one and the same class.

The focal comparison of spatial structure which is applied in this paper is based on two additional spatial (raster) layers derived from the original categorical maps. The value of a cell in the first layer is the size of the cluster that the cell belongs to. The second layer contains the weight for each cell, weights are chosen such that only one category is considered at a time (i.e. cells of other categories have a weight of zero) and the total weight of all cells in a cluster is 1. The focal comparison of spatial structure is made by comparing the weighted average cluster size of the focal window of both maps, according to the following equations:

$$\bar{D} = \sqrt{\frac{\sum_{i=1}^{n} (D_i)^2}{n}} = \sqrt{\frac{\sum_{i=1}^{n} (S_i^A - S_i^B)^2}{n}} \tag{5}$$

$$S_i^A = \frac{\sum_{j=1}^{s} w_{i,j}^A s_{i,j}^A}{\sum_{j=1}^{s} w_{i,j}^A} \tag{6}$$

where i iterates over all cells in the map and j iterates over all cells in the window. $s^A_{i,j}$ is the size of the cluster including the j-th cell in the window centered around cell i in map A. $w^A_{i,j}$ is the weight associated to that cell, which is the inverse of the cluster size or zero if the cell is not taken in by the class for which the comparison is performed. D_i is the difference in cluster sizes at location i and $\bar{D}$ is the root mean squared difference over the whole map. D_i can be used to visualize the spatial distribution of differences over the map.



a. Urban areas in the Netherlands b. Cluster size distribution

Fig. 4. Cluster size frequency distribution of urban clusters in the Netherlands. The straight line in the log-log plot indicates a power law distribution.

2.3.4 Global Structure: Cluster Size Frequency Distribution

At the global level many characteristics can describe spatial structure. For urban systems fractal metrics have been considered adequate (Batty and Longley 1994; de Keersmaecker et al. 2003). One particular aspect of spatial structure in which a fractal ordering becomes clear is the cluster size frequency distribution. In many urban systems a power law can be observed between size of clusters and frequency of occurrence of clusters that size or smaller. This regularity has been

utilized to describe urban patterns (Benguigui et al. 2006), simulate urban change patterns (Schweitzer and Steinbink 1997) and compare maps (Dungan 2006; White 2006). The global structure comparison applied in this paper compares cluster size frequency distributions of both the simulated and the real map. Figure 4 illustrates the occurrence of a power law distribution for the case of urban clusters in the Netherlands.

Notwithstanding this example, cluster sizes of land use classes do not always adhere to power law distributions. The Kolmogorov-Smirnov (KS) distance which can be used to express the similarity between series of sampled data without assuming any distribution is, therefore, of great practical use. The KS distance is the maximum difference in cumulative frequency between two sampled distributions:

$$D_{A,B} = \sup_{x} \left| F_A(X) - F_B(X) \right| \tag{7}$$

where $D_{A,B}$ is the KS distance of the two sampled distributions A and B of variable x. In other words, the KS distance is the maximal vertical distance between two cumulative frequency distributions.

2.4 Reference Levels

The methods described in the previous sections present a comprehensive overview of similarity between model and reality. It is not sufficient however to identify strengths and weaknesses of the model on the basis of these results, since metrics are expressed at various scales. Moreover, not all of the registered similarity is a consequence of the performance of the model. In practice it seems that much of the similarity must be attributed to boundary conditions and constraints which are exogenously imposed on the model (Hagen 2003; Pontius Jr. et al. 2004; Hagen-Zanker and Lajoie 2008), most strikingly this is the case for models with an exogenously determined initial situation.

These two problems of interpretation are mitigated by introducing a reference model. This model is subject to the same constraints and boundary conditions as the tested model, but otherwise represents as little process as possible. The difference in performance between the tested model and the reference model, can then be attributed to the processes present in the tested, but not in the reference model. The reference model provides a common standard and individual results become mutually comparable.

The boundary conditions and constraints in the current case are formed by the initial situation and the total area constraints, i.e. the total area of each land use class is an input to the model. The reference model to observe these conditions and meet the constraints is the ‘random constraint match’ model.

The random constraint match model compares the number of cells on the initial map to the area required by the constraints. Of those classes which are

over-presented in the initial map, the surplus cells are selected randomly to change to another class. The under-represented classes are randomly distributed over the cells selected in the previous step.

Let us consider the earlier example whereby map A is now used as the initial situation and map B for the total area constraints. In map A there are 2027 cells of open class, in map B there are 1995 ones. Therefore the class is over-represented in the initial situation by 2027 – 1995 = 32 cells. Likewise the city class is over-represented by 15 cells. The river and park classes are under-represented by 18 and 29 cells, respectively. The random constraint match model initializes with map A. It then randomly selects 32 cells of open class and 15 cells of city class. These 47 selected cells are then, in random order, replaced by 18 cells of river class and 29 of park class. Figure 5 shows the two maps and the derived results of the random constraint match model.

Map 1 Map 2 Random constraint match

Fig. 5. Application of the random constraint match model where total area constraints are posed by map 2 and the initial situation is that of map 1

3 Application and Results

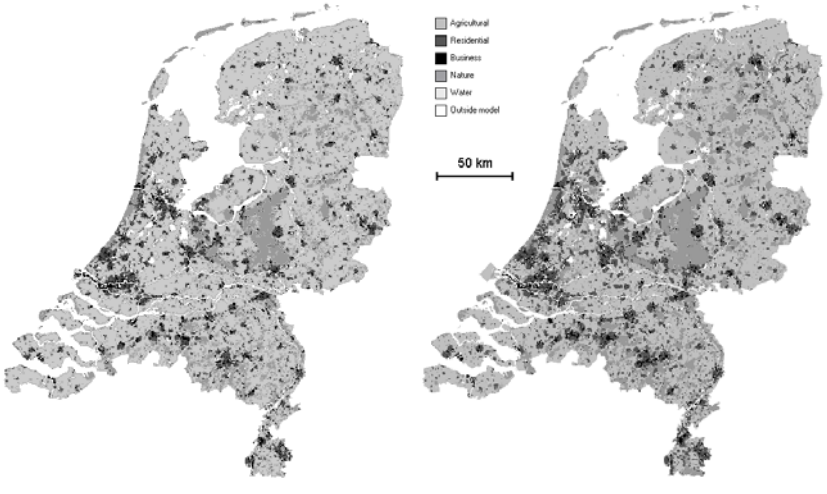
The evaluated model is the Constrained Cellular Automata (CCA) land use model (White et al. 1997) as it is applied in the Environment Explorer spatial planning support system (Engelen et al. 2003). The system incorporates several other models besides the CCA, but the spatial distribution of land use classes is the responsibility of the CCA and that aspect is evaluated here.

The CCA model simulates land use change year-by-year allocating land use classes to those cells for which they have the highest potential. The allocation process is constrained such that the exogenous area demand for each land use class is met. The potential of each location for different land use types is dynamic and changes over time as a function of land use classes found in the neighbourhood of each location. Hereby a reciprocal relation comes about, where the potential layers are determined by the spatial distribution of land use classes and at the same time the spatial distribution of land use classes is determined by the potential layers. This mutual relation causes the complex and self-organizing behaviour of cellular automata models and is hypothesised to underlie the formation of urban morphology.

The potential layers are not only a function of the dynamic neighbourhood effect, but a composite measure of several layers. Besides the neighbourhood effect, these are accessibility of road network, zoning status of cells, and physical suitability of land. Finally, a stochastic perturbation is included.

For the calibration and validation tasks only limited land use data were available. Land use maps were available for 1989, 1993, 1996 and 2000. In an early stage it was decided that the period of calibration was 1989-1996 and that of validation was 1996-2000. The consideration was that calibration affects model performance, whereas validation only measures the performance. Better results can therefore be expected if a longer period is used for calibration. The intended period of application is 2000-2030, therefore a considerable gap remains between the duration over which the model is calibrated/validated and applied.

The model identifies 15 land use classes. For the sake of model evaluation these are reduced to 5 main classes: agricultural, residential, business, water and nature. The model runs in steps of 1 year and the cell size is 500m. Figure 6 shows the initial land use map and the final model result.



a. Land use 2000 b. Model result 2030

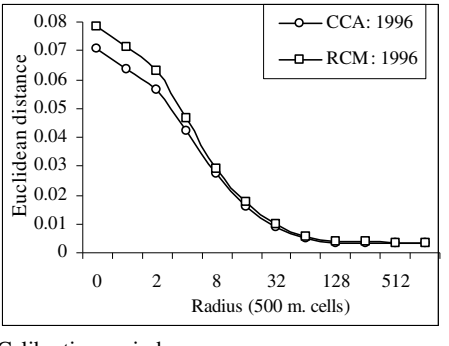
Fig. 6. Land use maps of the Netherlands; model and real

At the local level, cell-by-cell agreement is expressed by the Kappa statistics for the whole map and for individual land use classes (Table 6). Over the calibration period the agreement of the CCA model is better than that of the reference model for agriculture, residential, urban and nature classes. Note that the remaining two classes, water and foreign, are assumed constant in time by the model. Therefore the distribution of these classes is identical in both the CCA land use model and the RCM model. Over the validation period the CCA model does not outperform the reference model, except for the nature class, where the difference is minimal.

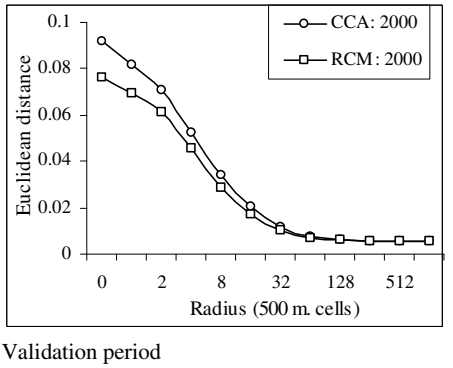
Table 6. Local presence comparison: Kappa statistics (CCA: Comparison between Constrained Cellular Automata model results and real data; RCM: Comparison between Random Constraint Match model results and real data)

	CCA: 1996	RCM: 1996	CCA: 2000	RCM: 2000
Overall	0.896	0.884	0.867	0.890
Agricultural	0.907	0.894	0.881	0.906
Residential	0.867	0.854	0.827	0.864
Business	0.769	0.766	0.731	0.809
Nature	0.924	0.908	0.907	0.904
Water	0.915	0.915	0.874	0.875
Foreign	0.960	0.960	0.949	0.949

Results of the focal presence comparison display (figure 7) the same pattern as Kappa results; over the calibration period the CCA land use model outperforms the RCM reference model, but it does not over the calibration period. With increasing radius of the moving window, errors of both CCA and RCM models decrease, but the relative position remains the same. Again, the difference in performance of both models is small.



a. Calibration period



b. Validation period

Fig. 7. Focal presence comparison: Euclidean distance

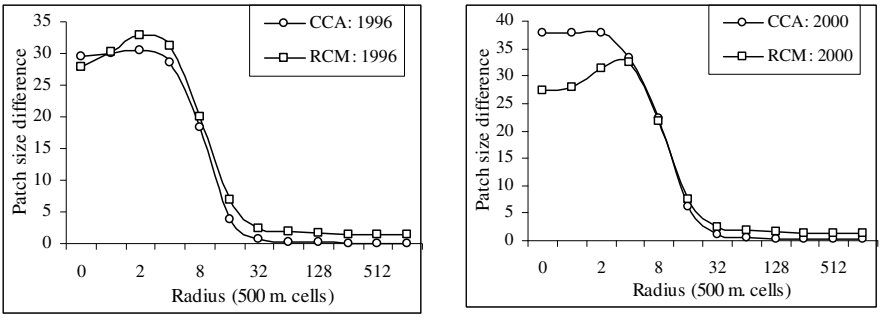
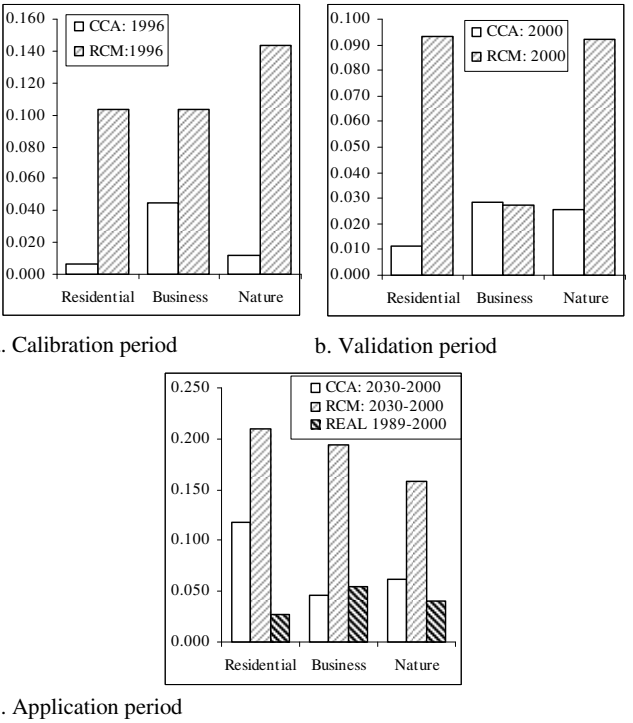


Fig. 8. Focal structure comparison: Patch size

Results for focal structure, measured on the basis of patch size are more diverse (figure 8). Like the earlier results, the CCA outperforms the reference model over the calibration period, but now also for the larger focal windows over the validation period.

From the perspective of global structure, the CCA model proofs to be more similar to reality than the reference model (figure 9). The difference in cluster size



a. Calibration period

b. Validation period

c. Application period

Fig. 9. Cluster size frequency distribution, compared by Kolmogorov-Smirnov distance

distribution between the CCA model results and reality are substantially smaller than those of the RCM model. This is not only the case for all classes over the calibration period, but also for residential and nature classes over the validation period. The business class shows little difference between the two models over the validation period. Agriculture class is not included in these graphs because patches of agriculture are so large that there are only a few of them on the map.

For the intended application period no data are available. The historical trend, however, is that the cluster size distribution is relatively stable in time. The CCA is better able to maintain a stable cluster size distribution.

4 Discussion and Conclusions

Data available for model calibration and validation are less than ideal. Although the model is to be used for explorations of approximately 30 years into the future, data are available for a period that spans over 11 years only. Moreover, this short period is split into separate calibration and validation periods, leaving 7 years for calibration and 4 years for validation. This is of particular significance since the land use model (like most geosimulation models) is a dynamic model, and rather than simulating land use patterns it simulates land use change patterns. The short calibration period means that relatively little change takes place and thus that the calibration routine has little information to pick up on. As a result, the risk of over-calibration is very real. By over-calibration is meant that too many parameters are fitted to too little observations and rather than fitting to the general trend, the parameters are fitted to eccentricities in the data. Over the period 1989-2000 about 6% of all cells changed status, it is unknown how many of these transitions are real changes and how many are data errors.

The comparison of the results between calibration and validation periods shows a decrease of performance over the validation period. This can always be expected and does not necessarily indicate over-calibration. At the local level results are better than the reference model over the calibration period, but not over the validation period. This is the strongest indication that over-calibration took place and the recommendation is made that future development of the model should aim to reduce the number of parameters and increase the amount of data by employing data available over longer time periods. There is a synergy between these two recommendations because the number of parameters can be reduced by considering less land use classes and if less land use classes are considered more data are available, over longer periods also. These recommendations have already been followed in a later project with good results (van Vliet 2006).

However the verdict of ‘over-calibrated’ must not be too rashly made. This conclusion is indeed suggested by the results at the local level and also at the focal level when presence is concerned. The focal results of spatial structure, as well as the global results offer a more positive interpretation. For these results too, the performance (relative to the reference model) over the validation period is less than that of the calibration period, but over both periods the land use model outperforms the reference model. Therefore we can conclude that the strength of

the model is in simulating urban structure at coarser scales rather than the precise (or approximate) land use class at particular locations. In summary, table 7 describes model performance according to the two axes along which the performance metrics have been organized.

Within the framework of the two axes many individual performance metrics can be applied. During this research we have applied more metrics than those reported here. In particular the Fuzzy Kappa metric (Hagen 2003) has been used to measure presence at the focal level. Structure at the focal level has also been measured in terms of edge density and fractal dimension (McGarigal et al. 2002). These results are not reported here, but do offer support for the summary presented in table 7.

Table 7. Overview of model performance according to the two axes of performance criteria (-- poor performance, - mediocre performance, + acceptable performance, ++ good performance)

	Local	Focal	Global
Presence	--	-	
Structure		+	++

All results, but most clearly, focal metrics show a strong correlation between the performance of the evaluated model and the reference model. Apparently, a large proportion of model performance must be attributed to factors exogenous to the model. This demonstrates the need of the reference model, since without this reference any analysis of performance criteria will be clouded by the impact of exogenous factors.

The contradiction in results between various performance metrics underlines the need for a multi-criteria analysis as applied in this paper. Focusing on a single metric bears the risk of over-confidence in the results (in the current case if only structure metrics would be considered) or under-evaluation of the model (in the current case if only presence metrics would be considered).

We generally recommend this approach for all studies involving calibration and validation of geosimulation models. Also exciting recent developments towards rigorous sensitivity analysis of large numbers of simulations (Jantz and Goetz 2005; Kocabas and Dragicevic 2006) can gain in scope and relevance if these recommendations are followed.

References

Barredo, J.I., Demicheli, L.: Urban sustainability in developing countries' megacities: modelling and predicting future urban growth in Lagos. Cities 20, 297–310 (2003)

Batty, M., Longley, P.: Fractal cities: a geometry of form and function. Academic Press Professional, Inc., San Diego (1994)

Batty, M., Torrens, P.M.: Modelling and prediction in a complex world. Futures 37, 745–766 (2005)

Benenson, I., Torrens, P.M.: Geosimulation: object-based modeling of urban phenomena. Comput Environ. Urban Syst. 28, 1–8 (2004)

Benenson, I., Omer, I., Hatna, E.: Entity-based modeling of urban residential dynamics: the case of Yaffo, Tel Aviv. Environ. Plan B: Plan Des. 29, 491–512 (2002)

Benguigui, L., Blumenfeld-Lieberthal, E., Czamanksi, D.: The dynamics of the Tel Aviv morphology. Environ. Plan B: Plan Des. 33, 269–284 (2006)

Briggs, W.M., Levine, R.A.: Wavelets and field forecast verification. Mon. Weather Rev. 125, 1329–1341 (1997)

Cohen, J.: Coefficient of agreement for nominal scales. Educ. Psychol. Meas. 20, 37–46 (1960)

Costanza, R.: Model goodness of fit: a multiple resolution procedure. Ecol. Model 47, 199–215 (1989)

de Keersmaecker, M.L., Frankhauser, P., Thomas, I.: Using fractal dimensions for characterizing intra-urban diversity: the example of Brussels. Geogr. Anal. 35, 310–329 (2003)

de Nijs, T.C.M., de Niet, R., Crommentuijn, L.: Constructing land-use maps of the Netherlands in 2030. J. Environ. Manag. 72, 35–42 (2004)

Dungan, J.L.: Focusing on feature-based differences in map comparison. J. Geogr. Syst. 8, 131–143 (2006)

Engelen, G., White, R., de Nijs, T.: Environment Explorer: spatial support system for the integrated assessment of socio-economic and environmental policies in the Netherlands. Integr. Assess 4, 97–105 (2003)

Engelen, G., White, R., Uljee, I., Drazan, P.: Using cellular automata for integrated modelling of socio-environmental systems. Environ. Monit Assess 34, 203–214 (1995)

Foody, G.M.: Status of land cover classification accuracy assessment. Remote Sens. Environ. 80, 185–201 (2002)

Hagen-Zanker, A.: Map comparison methods that simultaneously address overlap and structure. J. Geogr. Syst. 8, 165–185 (2006)

Hagen-Zanker, A., Lajoie, G.: Neutral models of landscape change as benchmarks in the assessment of model performance. Landsc Urban Plan 86, 284–296 (2008)

Hagen, A.: Fuzzy set approach to assessing similarity of categorical maps. Int. J. Geogr. Inf. Sci. 17, 235–249 (2003)

Heidke, P.: Berechnung des Erfolges und der Gute der Windstarkevorhersagen im Sturmwarnungsdienst. Geogr. Annal 8, 301–349 (1926)

Jantz, C.A., Goetz, S.J.: Analysis of scale dependencies in an urban land-use-change model. Int. J. Geogr. Inf. Sci. 19, 217–241 (2005)

Kocabas, V., Dragicevic, S.: Assessing cellular automata model behaviour using a sensitivity analysis approach. Comput. Environ. Urban Syst. 30, 921–953 (2006)

Kok, K., Farrow, A., Veldkamp, A., Verburg, P.H.: A method and application of multi-scale validation in spatial land use models. Agric. Ecosyst. Environ. 85, 223–238 (2001)

McGarigal, K., Cushman, S.A., Neel, M.C., Ene, R.: FRAGSTATS: Spatial Pattern Analysis Program for Categorical Maps. University of Massachusetts (2002), http://www.umass.edu/landeco/research/fragstats/fragstats.html (accessed October 14, 2008)

Monserud, R.A., Leemans, R.: Comparing global vegetation maps with the Kappa statistic. Ecol. Model 62, 275–293 (1992)

Pontius Jr., R.G.: Quantification error versus location error in comparison of categorical maps. Photogramm Eng. Remote Sens. 66, 1011–1016 (2000)

Pontius Jr., R.G.: Statistical methods to partition effects of quantity and location during comparison of categorical maps at multiple resolutions. Photogramm Eng. Remote Sens. 68, 1041–1049 (2002)

Pontius Jr., R.G., Huffaker, D., Denman, K.: Useful techniques of validation for spatially explicit land-change models. Ecol. Model 179, 445–461 (2004)

Pontius Jr., R.G., Boersma, W., Castella, J.-C., Clarke, K., de Nijs, T., Dietzel, C., Zengqiang, D., Fotsing, E., Goldstein, N., Kok, K., Koomen, E., Lippitt, C.D., McConnell, W., Pijanowski, B., Pithadia, S., Sood, A.M., Sweeney, S., Trung, T.N., Veldkamp, A.T., Verburg, P.H.: Comparing the input, output, and validation maps for several models of land change. Ann. Reg. Sci. 42, 11–37 (2008)

Refsgaard, J.C., Henriksen, H.J.: Modelling guidelines: terminology and guiding principles. Adv. Water Resour. 27, 71–82 (2004)

Schelling, T.C.: Dynamic models of segregation. J. Math. Sociology 1, 143–186 (1971)

Schweitzer, F., Steinbink, J.: Urban cluster growth: Analysis and computer simulations of urban aggregations. In: Schweitzer, F. (ed.) Self-organization of complex structures: From individual to collective dynamics, pp. 501–518. Gordon & Breach, London (1997)

Takeyama, M., Couclelis, H.: Map dynamics: integrating cellular automata and GIS through geo-algebra. Int. J. Geogr. Inf. Sci. 11, 73–91 (1997)

Turner, M.G.: Landscape ecology: What is the state of the science? Annu. Rev. Ecol. Evol. Syst. 36, 319–344 (2005)

Turner, M.G., Costanza, R., Sklar, F.H.: Methods to evaluate the performance of spatial simulation-models. Ecol. Model 48, 1–18 (1989)

van Delden, H., Engelen, G.: Combining participatory approaches and modelling: lessons from two practical cases of policy support. In: International Environmental Modelling and Software Society (2006),
http://www.iemss.org/iemss2006/sessions/all.html
(accessed October 14, 2008)

van Vliet, J.: Validation of land use change models: a case study on the Environment Explorer. Universiteit Wageningen (2006),
http://www.lumos.info/publications-en.php
(accessed October 14, 2008)

White, R.: Pattern based map comparisons. J. Geogr. Syst. 8, 145–164 (2006)

White, R., Engelen, G.: High-resolution integrated modelling of the spatial dynamics of urban and regional systems. Comput. Environ. Urban Syst. 24, 383–400 (2000)

White, R., Engelen, G., Uljee, I.: The use of constrained cellular automata for high-resolution modelling of urban land-use dynamics. Environ Plan B: Plan Des. 24, 323–343 (1997)

White, R., Engelen, G., Uljee, I., Lavalle, C., Ehrlich, D.: Developing an urban land use simulator for European cities. In: Fullerton, E. (ed.) Proceedings of the 5th EC-GIS Workshop held in Stresa, Italy, European Commission, Joint Research Centre Ispra, June 28-30, pp. 179–190 (2000)

Author Index