ISBN: 978-88-98533-01-5
DOI: 10.13133/978-88-98533-01-5

Secoa

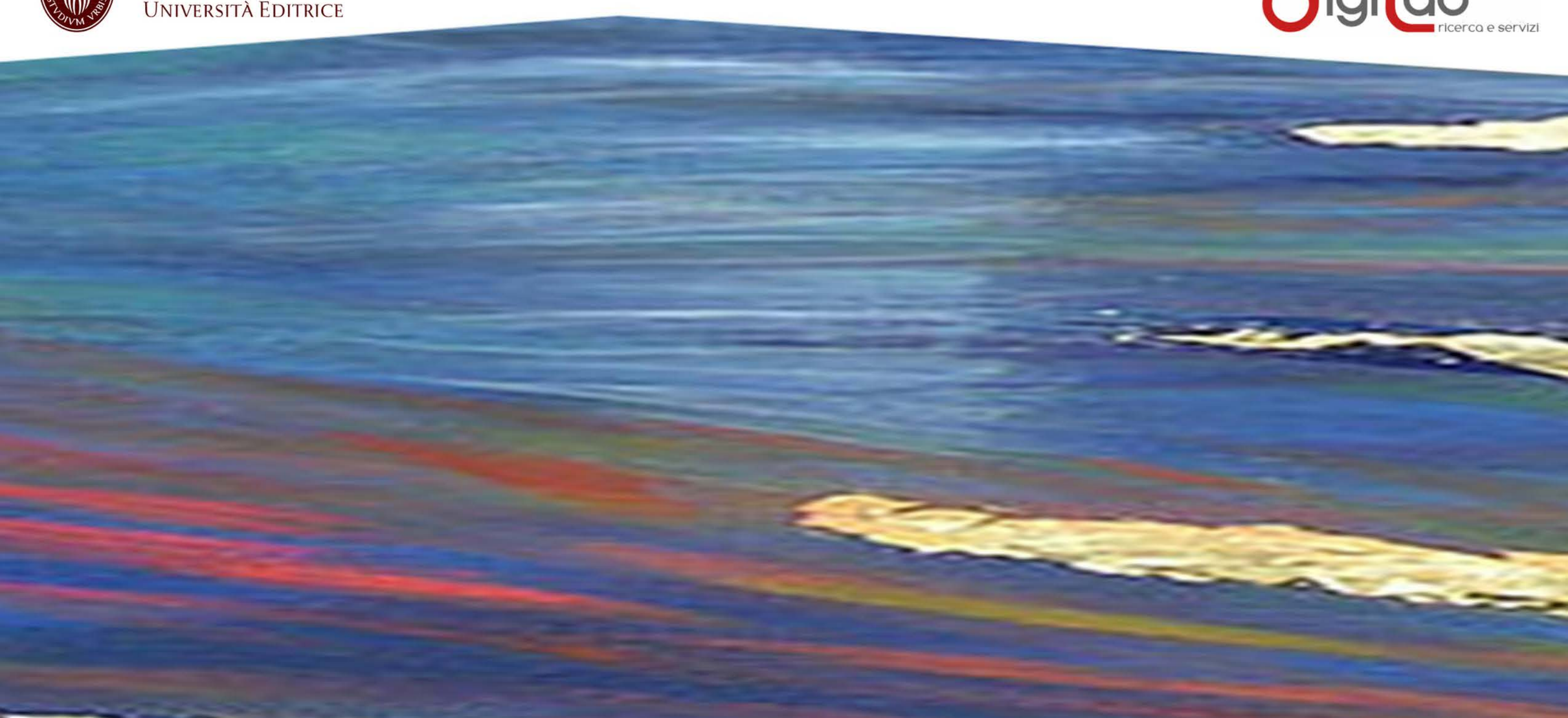
URBAN COASTAL AREA CONFLICTS ANALYSIS METHODOLOGY

Human mobility, climate change and local sustainable development

edited by
Armando Montanari

Sapienza Università Editrice

digiLab ricerca e servizi



Collezione di "Testi e Studi Umanistici"

Serie

Letterature, linguaggi e culture

SECOA, 5

The research leading to these results has received funding from the European Community's Seventh Framework Program FP7/2007-2013 under grant agreement n° 244251

URBAN COASTAL AREA CONFLICTS ANALYSIS METHODOLOGY

HUMAN MOBILITY, CLIMATE CHANGE AND LOCAL SUSTAINABLE DEVELOPMENT

Edited by
Armando Montanari

SAPIENZA
UNIVERSITÀ EDITRICE

digilab
ricerca e servizi

2013

SECOA FP7 Research Project is published by **Casa Editrice Università La Sapienza**

Published online in **2013**

The copyediting is by **Digital Publishing Division of DigiLab**

(Centro interdipartimentale di ricerca e servizi) – La Sapienza Università di Roma

Managing editor: **Gianfranco Crupi**

Copyediting and Layout editor: **Ivan Macculi**

Graphics of book cover: **Mariella Scelsi**

DOI: http://dx.doi.org/10.13133/978-88-98533-01-5

Table of Contents

Editor's Preface to the Series

This volume is part of a new series on cross-national comparative research in the fields of global climate change, coastal areas, sustainable urban development and human mobility. These factors, which arise at both the local and global level, are confronted with many conflicts of interest in every possible combination between the local and the global. The volumes being published in this series attempt to provide a contribution to resolving these conflicts. This multi-national and multi-disciplinary network was set up in 2009 on the occasion of the European Commission's call for proposals for a Seventh Framework Programme (FP7) project. The research project Solutions for Environmental Contrasts in Coastal Areas (SECOA), Global Change, Human Mobility and Sustainable Urban Development won the bid and began work in December 2009 (http://www.projectsecoa.eu/), coordinated by Sapienza Innovazione (Riccardo Carelli) with scientific coordination by Sapienza Rome University (Armando Montanari).

Global changes affect both the environment and socio-economic conditions: first the economic crisis of the 1970s and then the financial crisis of the first decade of the new millennium have had a profound impact on environmental and socio-economic conditions. SECOA examines the effects of human mobility on the growth and restructuring of urban settlements in coastal areas, where: a) the environment is particularly fragile and space is limited, b) every phenomenon is far more concentrated and c) the effects on natural and cultural resources and the environment are more acute. Being aware of these effects can be extremely useful for governments and companies – particularly in the building sector, but also in tourism – in planning their future growth. Awareness of the environmental status of the coast and the local population's usage preferences can help to plan the development of homes, retail and leisure facilities. The problems have multiplied as a result of climate change and its influence on environmental parameters such as the sea level, sparking an increased risk of flooding, the spread of pollution and the displacement of a large number of inhabitants. The control and reduction of undesirable consequences is leading to increased conflict among stakeholders. An integrated approach to the ecosystem incorporating the social, economic and natural sciences is essential to understand the complex and dynamic problems typical of coastal towns, as the figure illustrates. The complexity of the problems and the heterogeneousness of the data required to document very diverse phenomena are being managed using Geographic Information Systems (GIS). SECOA aims to: 1) identify conflicts, 2) analyse their quantitative and qualitative effects on the environment, 3) create models to synthesise the various social, economic and environmental systems and 4) compare the priorities of each type of coastal town using a taxonomic tool. Coastal areas have traditionally been considered difficult to manage because of the problem of the weather, the tides and the seasons and the overlapping of the specificities features of physical geography and hydrography, as well as overlapping jurisdictions and remits of individual government bodies and the competing needs of various civil society stakeholders. Local, regional and national administrations are often responsible for similar aspects of the same physical area and the uses of coastal zones, such as fisheries, environment, agriculture, transport (inland and marine), urban planning, the land registry and the national cartographic and hydrographic services. Many people are able to intuitively recognise a coastline, although they find it harder to determine its precise landward or seaward extent and vertical growth. For this reason, and considering the diversity of the stakeholders, managing authorities and administrative structures, there are inevitable conflicts between users of coastal zones, developers and the rest of society. Similarly, there is a conflict between human society and natural resources. Because of the complexity of the problems involved, the spatial component of data has also been taken into account through the use of GIS, which offer enhanced possibilities of contributing to coastal zone management for a number of reasons: (i) their ability to manage large databanks and integrate data relating to quite heterogeneous criteria; (ii) their inherent tendency to harmonise data from different sources and thereby contribute to the exchange of information between governing bodies and research institutes; (iii) the possibility they offer of using shared data banks; (iv) their inherent aptitude for modelling and simulation that allows for alternative scenarios to be built before being implemented. The basic function of information that can appropriately inform decision-makers is the ability to produce online geographical maps to illustrate the location of problems, the densification and concentration of shortcomings, the density, the content, what happens in the environs, and changes.

Together with the problems created by climate change, the SECOA project examines the spread of human mobility – an area that principally involves the social science disciplines, each with its own research framework, levels of analysis, dominant theories and hypotheses of application. The social science fields can be considered according to the dependent and independent variables they use. For example, anthropology, demography and sociology consider behaviour a dependent variable; for economics, it is microeconomic flows and impacts; for geography, it is decision-making ability; for history, it is experience; for law, it is treatment and for political science, the dependent variables are management policies and their results. Examples are always hard to agree on, but in this case they are being used to emphasise the differences that exist even between related sectors, and the obvious multiplication of variables when the ones proposed by the social sciences must include geomorphological variables (the way the coast physically changes) and environmental and cultural resources (their availability and the way they are consumed). The SECOA project has attempted to tackle this problem by also measuring types of individual mobility and the attractiveness of the territory. For previously mentioned reasons, these data are not generally registered, so it was decided to use the GIS tool to add space and time values. Space in coastal metropolitan areas is characterised by the differences among the various spatial components, and it is not always easy to identify the coastal stretch used as the element of comparison. Time, on the other hand, is defined in terms of recurring daily, temporary and permanent mobility, with a further variant of mobility that is either production-led (blue-collar, white-collar, managers, regular and irregular workers) or consumption-led (including mobility for reasons of tourism, leisure and retirement). The prediction models, on the other hand, are an instrument to connect the past to the future, and hence to integrate the natural and cultural heritage and contribute to building prediction scenarios.

The SECOA Series Editor would like to thank all the colleagues who contributed their work, advice and arguments to this volume. The book covers a logical phase during which a substantial amount of data – some of them apparently unrelated – turn into an evolved form of knowledge able to sum up the real world. It is for this reason that all the SECOA national research groups, whose coordinators are on the editorial board for SECOA publications, have contributed to putting together this volume. The quantitative and qualitative data were collected on the basis of operational programmes drawn up by the coordinators of the Work Packages (WP) n. 1, 2, 3 and 4, whose work has been documented in previous issues of the SECOA Publication Series. A simplified interpretation of the context was provided by putting together a process concerning specific aspects of the territory: its environmental problems, the characteristics and habits of civil society, and finally the interpretation of the environmental conflicts identified in urban coastal areas. WP5 arranged and catalogued the data in accordance with the INSPIRE Directive.

This volume is the result of interpretation and synthesis by a group of researchers at Rome Sapienza University's Department of European, American and Intercultural Studies, composed of scientists, social scientists and humanists: Luca Deravignone, Alessandro Londei, Marco Ramazzotti, Barbara Staniscia and Irene Viaggiu. Like Charles Percy Snow, who is mentioned in the third chapter of this book, Ramazzotti combines a humanities background with natural science studies. Similarly, Londei combines natural sciences research with producing and directing theatrical texts. Viaggiu puts on her archaeological hat to study productivity in human mobility. And Deravignone and Staniscia are skilled at photography, music and composing poetry, besides the professional competencies they demonstrate in this book.

The group coordinated by the undersigned met on a regular basis over a period of 18 months with the focus, drive and sheer enthusiasm that are typical – or should be typical – of all researchers involved in a project. Considering the amount of work that was put in, the resulting book should be perfect. But it is with a lingering worry that it may appear imperfect or incomplete that we entrust it to all our potential readers, to whom we would be grateful for pointing out any errors or omissions.

Armando Montanari

Rome, May 2012

CHAPTER 0.

Introduction

Armando MONTANARI

Department of European, American and Intercultural Studies,
Sapienza University of Rome, Italy

1. Introduction

For research projects investigating complex issues, a framework is prepared by deciding on the issues to be studied and the disciplines that will be involved. Once a framework in line with the proposed objectives of the project has been developed, the research work is divided into work packages (WP).

Figure 0.1. The SECOA PERT diagramme.

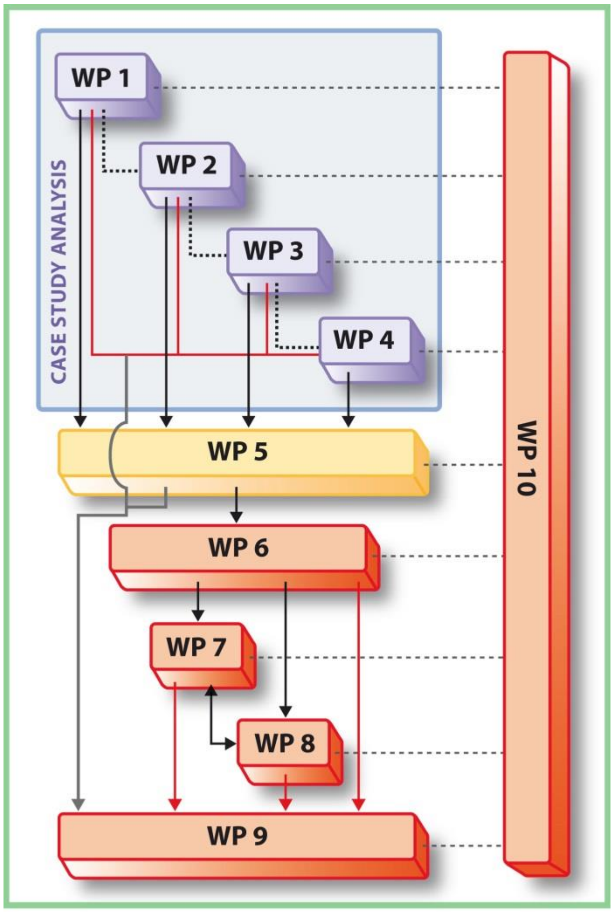


Figure 0.1 shows the SECOA project's PERT chart, with the following sequence:

- WP 1 – Climate change and natural hazards;
- WP 2 – Environmental stresses and resources uses for sustainable development;
- WP 3 – Social, economic, cultural dimensions of urban development;
- WP 4 – Analysis of conflicts of uses of coastal resources amongst users and sectors;

- WP 5 – Data integration;
- WP 6 – Taxonomy and modelling;
- WP 7 – Methodology and tools for appropriate policies;
- WP 8 – Scenario building;
- WP 9 – Dissemination of results;
- WP 10 – Project Management.

SECOA's logical structure is based on the DIKW hierarchy: (i) Data, (ii) Information, (iii) Knowledge, (iv) Wisdom, (v) Understanding. WP6, the subject of this volume, fits into the knowledge category of the DIKW hierarchy. Any major multi-disciplinary, cross-national research project should be a benchmark for knowledge, but the issue is far from obvious. There is nothing obvious about the move from information to knowledge. The SECOA researchers put in a great deal of time and effort, and used their recognised competence in the various disciplines involved in the project, to acquire information. But information is made up of many components that will not produce knowledge unless they can be made to communicate. And knowledge cannot be a concise interpretive summary based on only part of the information collected – that would be unthinkable. At first glance, it would seem that only some of the information – which communicates because of analogous references, or communicates more easily with those seeking to produce knowledge – is comparable. However, in this way there would be a serious shortage, which would result in inexistent or incomplete multi-disciplinary comparison. Indeed, if we refrained from using even part of the information produced by some disciplines to draw up a synthesis, multi-disciplinary research would remain a formal concept devoid of content, and the very essence of scientific knowledge would be lacking. SECOA is not the first multi-disciplinary project, as the researchers who put the project proposal together were aware from the outset. In drawing up the proposal, they took previous experience in this sector into account. Neither is SECOA the first project to study the problems of coastal areas; hence the framework and findings of similar projects had to be appraised.

This volume examines the problems of cross-national research and embarks on a historical analysis to evaluate the difficulties that these cognitive processes have encountered over the past half-century. The problems currently encountered by researchers working on European projects are no different to those of their counterparts many decades ago. Going back a few decades, comparative research projects were of interest to a limited number of researchers – an elite few who had to tackle the political differences and issues of the time. These researchers had a great deal of scientific curiosity, but the funding available to them was not necessarily commensurate with the level of commitment required. Although large-scale international funding was not available, the main prize to which researchers aspired was being able to overcome the considerable hardships they faced. Things began to change perceptibly in Europe in the 1990s, after the collapse of the barriers created by the cold war, and even more so when, immediately afterwards, many Central and Eastern European countries became members of the EU. The author did not participate in the past half-century's international research projects, but has had the honour of participating in the debate on these issues for at least the past thirty years, and meeting and talking to players in comparative research projects in previous decades. There are thousands of EU-funded projects today. Initiating and funding research has gone from being a national responsibility to an EU responsibility. While there has been substantial growth in this field over the past few decades, the changes have been in style rather than substance. The work of Niessen and Peschar (1982) is a significant benchmark, as it provides an overview and reflections on the theory, methodology and organisation of international comparative research in Eastern and Western Europe. Their book is the result of research and a number of validation meetings carried out on behalf of the European Coordination Centre for Research and Documentation in Social Sciences (Vienna Centre). Their work followed other Vienna Centre initiatives aimed at examining the problems of international projects in the previous decade. The author was a member of the research staff at the Vienna Centre from 1979 to 1983, the very period during which Niessen and Peschar (1982) were carrying out their research, and took part in the various formal meetings during official events – or rather informal meetings, held as they were in one Viennese *beisel* (tavern) or another! All the main inventions and innovations in this sector had taken place by the end of the 1980s. With the disintegration of the Eastern European regimes, cross-comparative analysis ended up becoming a political issue; there was no longer any need for research to be used to maintain the social and cultural connection between different socio-economic systems. In those years, alongside international research there were national or bilateral research streams, which were far more numerous, and better funded. Gradually, interest shifted from the political aspect to the importance of international comparative research as a key tool for contributing to innovation and competitiveness in European society as a whole. As a result, the EU increasingly took over leadership in deciding on strategic research sectors and funding. Being European, the projects it funded had to involve research teams from many different countries. The enlargement of the EU and the entry of new member states increased the number of participating teams as well as the differences between participating countries. With the EU tending to play a greater role overall, some calls for proposals began to specify that non-EU researchers could also participate. SECOA submitted its application in response to a call stipulating that research groups from India and at least one Asian country on the Pacific Rim would have to be involved. This is why the SECOA project proposal was constructed around problems encountered in the EU and Israel as well as in India and Vietnam. Alongside its cross-national approach, SECOA requested the collaboration of colleagues working in the natural sciences, social sciences and humanities in order to collect information that would also be useful for developing adequate knowledge.

Once the framework had been drawn up, a research methodology had to be identified. SECOA examined global problems, of a general nature, and local problems, which are justified and explained through studying the territory. The approach used was a bottom-up projection, starting from the incoming data from the territory to reach an analysis of global phenomena. This method made it necessary to set aside pre-conceived models and algorithms to which to apply the information produced by the different work packages. The researchers therefore sought a method that would make it possible to draw up an interpretive model starting from the data. Many past international comparative research projects have concluded with data analysis. Like many other current European projects, SECOA uses data analysis as a means to propose policy and build scenarios. This volume deals with data processing and analysis through taxonomy and interpretive modelling. Any researcher who has been involved in a cross-national project is aware of the difficulties to be tackled, and knows that even though the findings of a single project will hardly bring about great progress in the field, a contribution, however limited, to the spread of knowledge remains a duty.

SECOA developed on the basis of a collaboration involving three cultures – the natural sciences, humanities and social sciences. The findings presented in this volume have also benefited from the participation of and close collaboration among researchers representing the three cultures: Londei for the natural sciences, Ramazzotti and Viaggiu for the humanities and Deravignone, Montanari and Staniscia for the social sciences.

The first part of the volume reviews the main stages of cross-national and multi-disciplinary research. The results of the analysis are used to verify SECOA's knowledge structure and justify the next set of methodological choices. The SECOA project essentially studied conflicts relevant to achieving sustainable development in urban coastal areas. The project had to take the reference territory into consideration whilst managing the economic, social and environmental variables. It was within this territory, composed of urban and metropolitan areas, that the various components of society that are both affected and active at the local and global levels were studied. Conflict derives from the inability of the various elements of the system to achieve representation in a way that tallies with the importance of each theme, and not merely on the basis of the varying ability of interest or pressure groups to influence or act.

The first chapter employs the instruments of analytical archaeology to introduce the way in which artificial adaptive systems are used; these include Learning Systems, Artificial Neural Networks (ANNs) and Evolutive Systems. The reference to archaeology is significant in this context, as it is traditionally considered to be a humanities discipline. In fact, for a few years now, some scholars have been identifying similarities between archaeology, a science of the territory, and geography, a social science. In this regard, Barcelò (2009) points out that the only difference is that archaeology studies completed social actions, while in the social sciences these actions are either underway or have just begun. So while the spatiality and temporality of actions can be observed in the social sciences, in the humanities they are inferred from indirect evidence. In these disciplines, the social players in action cannot be studied; everything must be interpreted on the basis of observations, not of the actions, but of the consequences of previous actions. The collaborative model established between analytical archaeology and quantitative geography that is described in this volume can be represented by a butterfly diagram in which on the one hand archaeology collects and summarises the cognitive foundations of the other humanities, while geography's role in collaboration is to convey, through the use of georeferenced data, the cognitive mass that comes from its being a social science, with an enhanced ability to familiarise itself with economics, sociology, statistics and urban planning. Pacinelli (2008) refers to the "cone of plausibility", through which he demonstrates that the trajectories of the exploratory scenario move from the present towards the future, while those of the regulatory scenario go from the future towards the present. In this theory, the butterfly takes on the shape of a double cone; one projects itself from the present, via the trajectories, towards the future, on an exploratory scenario, while the other moves in the opposite direction, projecting itself through the trajectories from the present towards the past, on a predictive scenario (Montanari, 2010). Ramazzotti cites natural computing, in which each phase of artificial generation is a model comparable to a natural cultural process. ANNs are algorithms for drawing up information that makes it possible to rebuild the rules connecting a certain number of explanatory data of the problem being studied.

The second chapter looks at projects and literature in which the application of ANNs intersects with the capacity of Geographic Information Systems to manage geographical data.

The third chapter is devoted to international comparative analysis: the progress of international scientific comparison, the levels of complexity it has currently reached, and a reflection on the need to apply new interpretive models so as to achieve forms of knowledge increasingly capable of interpreting this complexity. While the third chapter remains an element of transition between reflections on methodology and SECOA's fieldwork, the fourth chapter leads us into the problems of the 17 urban coastal settlements and the 27 conflicts that were studied. The fifth chapter examines the specificities of SECOA and explains the ANNs used to manage the taxonomy: self-organising maps (SOMs); nonlinear modeling; a feedforward neural network (FFNN). The sixth chapter looks at the role of the GISystem in regularising, homogenising and comparing data. It also examines the integration of the SOMs into the GISystem to contribute to achieving a level with greater added value, the Geographic Information Science (GIScience). Chapter Seven presents and discusses the application of ANNs to SECOA so as to supply a response to two main tasks: the taxonomy of the case studies, and determining the relations among the variables that are the main components of sustainable development in order to define SECOA's descriptive model. In the eighth chapter, the outputs derived from the application of the SOMs and FFNNs are interpreted from a geographical standpoint. In this way, five types of territory and five types of conflict were identified.

2. References

Barselò A.J., 2009, Computational Intelligence in Archeology, Oxford: IGI Global.

Montanari A., 2010, Turismo e sistemi di informazione geografica, Milano: Bruno Mondadori.

Niessen M. and Peschar J. (eds.), 1982, International Comparative Research. Problems of theory, methodology and organisation in Eastern and Western Europe, Oxford: Pergamon Press.

Pacinelli A., 2008, Metodi per la ricerca sociale partecipata, Mlano: Franco Angeli.

CHAPTER 1.

Logic and Semantics of Computational Models for the Analysis of Complex Phenomena. Analytical Archaeology of Artificial Adaptive Systems.

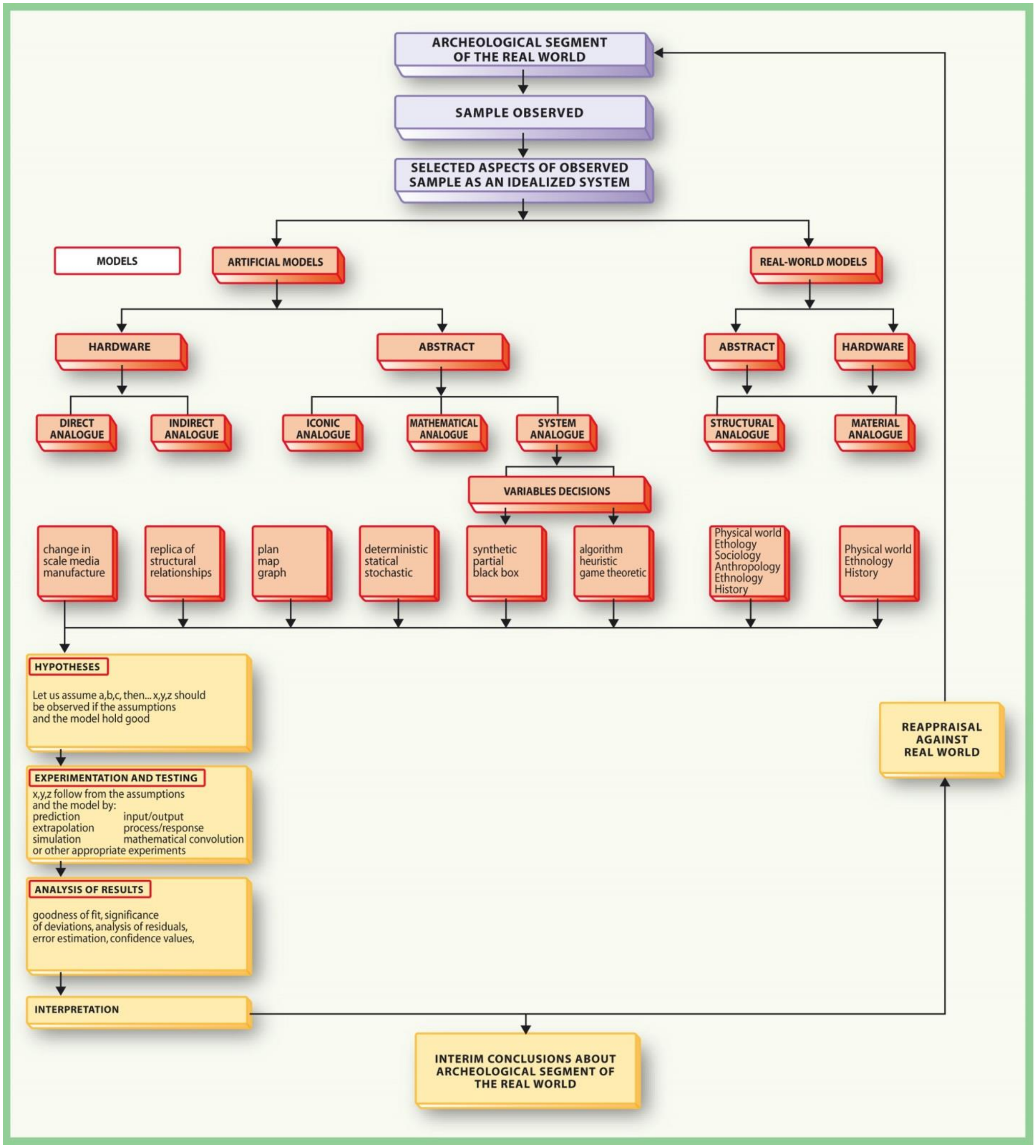
Marco RAMAZZOTTI

Dipartimento di Scienze dell'Antichità, Sapienza University of Rome, Italy

1. Introduction. The Encoding of Cognitive Principles and Spatial Complexity

Analyses and methods which use deductive inference tend "to *predict* the Result of a (true) Law through a Case" and return that result (model) which represents a projection of the historical meaning ascribable to the data, in other words its prediction. Pessa noted that "no automated reasoning program can be universal, in the sense of deciding, for any set of inference rules and axioms, whether or not a given symbolic expression is a theorem of the theory in question and, if it is, supplying an effective deduction procedure" (Pessa 1992: 83). Within the humanities, the observations of the mathematical, statistical, economic and geographical relationships processed for a given body of data are represented with tables, matrices, histograms and dendrograms which perform the dual purpose of spatialising and structuring the values, the percentages, the trends and the intersections between a limited number of variables. These graphs are therefore already models which summarise the repeated observation across multiple cases as a result expressed through frequencies whose different variation and intensity always constitutes a degree of (cultural) intentionality (fig. 1.1).

Figure 1.1. Models classifications after D. L. Clarke 1972.



Cultural intentionality in realising a given production of artefacts presupposes the concept of "type" as a principle, a finite, projectual entity, expressed by the intentional correlation of different attributes and resulting axioms for which each hidden organisation of the attributes will define the characteristics of a type, multiple types the characteristics of a class, and a class the "intentional" product of a culture. Given the technical clarifications on the topic, the analyses which were used first by Analytical Archaeology, and today by many other of the so-called Social Sciences to classify attributes, types and classes of a given culture have been manifold, extremely varied and more sophisticated the greater the variability of the systems observed (Ramazzotti 2010: 88-126). Analysis of the metric frequencies of the attributes was carried out to select trends, distributions and correlations in structuring the artefacts, and the first histograms represented their formal characteristics in terms of modes and frequencies; in the same manner, analysis of nominal frequencies performed through contingency tables and *Chi-squared tests* allowed recovery of association matrices of two or more classes and verification of whether a given decoration on the surface of a type of container was random or not. The necessarily accelerating increase in homogeneity of the classes and the presence of documents with strongly variable attributes (many of which shared by other artefacts, but none necessary or sufficient to distinguish or characterise them) was incorporated into the concept of "polythetic groups", which is key because it gave rise to specific research on the most suitable tools for highlighting the *similarities* and *differences* which could structure composite and/or highly specialised production. Recognition of these qualities (analogies and differences) in the material culture follows the psychological research intended to apply those methods to isolate such essential functions of the cognitive process. In the first cumulative analyses, which studied the growth of the level of technology in the same manner as the evolutionary process, the percentages of artefact types are even associated with cranial capacities to explain the presumed symmetry between the growth in functional complexity of a given implement and man's evolutionary growth, essentially understood as adaptive growth, in other words due to the necessary acquisition of technological experiences (Leroi-Ghouran 1977). In the same way, the methods of multivariate analysis, factor analysis, automatic classification and Principal Component Analysis intended to show the *structural* nature of that variability present in the class, both to make future comparative exploration of it in a more precise way and to draw its unique and irreducible associative root. This attempt to trace the origin of the class in order to redraw its relational structure was, on the other hand, equivalent to the first experiments which were performed in analytical psychology to outline the human ability to structure reality into 'similar' and 'different' (Sternberg 1987: 19-27), and it is indeed the very first studies applying differential logic to understand intelligence that gave rise to the suggestion to use techniques such as Correspondence Analysis to reduce the high level of variability of cultural traits recognised in a limited and more controllable number of factors. Each culture economic behaviour presupposes the site *principle* as an entity defined by a group of geomorphological, stratigraphic and morphometric values which can be measured. Accepting this principle, many territorial analyses have used statistical methods to document the existence of valid laws which could explain different economic behaviours. The first were performed to verify how the settlement structures defined in graphical models of spatial structure could be traced onto consistent logical networks which would automatically transform with changes to the geometric measurement of their respective areas of influence. Changes to their reciprocal linear distances, changes to the 'weight' applied to the group of nearby settlements and, finally, changes in the relationships between distances and sizes of the centres present in a defined system. These automatic changes, which evidently transform every organisation of the territory into a different settlement model, were then clarified by the application of ever-more sophisticated rules which all aimed to define the structural characteristics of the transformations, considering them determined by the interaction of two variables, those of the size occupied by the site in the hierarchy and those of its distance from the other nearby centres, or of three variables, size, distance, influence on the region[1]. The automatic change in the transformation of

[1] Among the most suitable tools to activate this process are the so-called Thiessen Polygons, which have had a strong impact since the early eighties. The theory, (originating with Descartes and then developed mathematically in the 19th century by Dirichlet and Voronoi), provides for a grid of polygons to be drawn around the set of distributed points – each of the polygon's sides is formed by the perpendicular line passing between the minimum distance of two points. The calculation of the area of each polygon is understood to be indicative of the macroscopic differences in the distribution of the points in that space; in any case, since this grid does not take into account the high variability of factors of three-dimensional space (e.g. geomorphological obstacles) and their temporal differentiation (e.g. the dynamics of attendance of the territory), using it mechanically can produce questionable results. The measurement characteristics of the polygons have indeed recently been redefined, and a 'weighting' system has been proposed based on moving from the perpendicular to the distance between centres, therefore no longer passing between the midpoint, but fluctuating proportionally to the difference in size between sites. In the so-called *X-Tent* model, therefore, calculation of the 'weight' (in other words the influence) of every site on its surroundings. Application of these models has become typical in research on a territorial scale, but they have also been used to simulate and compare landscapes of power on a much larger scale. Vide Level - Renfrew 1979: 145-167; Renfrew 1984: 86-155.

a settlement structure did not, however, offer the chance to trace which were the functions activated in the structure able to transform the settlement landscape. From this limit, a more ambitious attempt emerged to further summarise the recognised *frames* in order to ascertain which economic rules system was the origin of the change. Once the syntax of the main relationships between the sites of a structure had been systematised in a theoretical list of associative constraints, exploration of the *rules* held to be at the base of the change could begin, rules which – initially – would make use of the principles of urban economics, first of all that founded on the interpretation of the relationships between production and transport cost, then between production, cost and geometry of transportation. The latter (which, let us remember, was a true geographical theory of optimisation in early 1930s Germany) is still widely applied in research into urban economics, economic geography and territorial archaeology[2]. Analysis of the economic relationships between settlement sites did not, in any case, have to be limited, or reduced to the presentation of those optimal operating rules towards which the different systems would have aimed and which would have been behind their diversity[3]. The limits of considering the complexity of the relationship behind the morphometric parameters of the sites and

[2] Central Place Theory presupposes that an organised distribution across a territory is based on at least two macro-categories of settlements – one composed of the most important in terms of size, population, availability of services, commercial structure etc., and one composed of the less important ones. The second group will tend to arrange themselves around the first depending on the ease of access or administrative control, until they form a homogeneous hexagonal lattice. According to this economic theory devised by the German geographer Christaller (1933), distribution of the centres in a given territory is regulated by "Principles", a logic by which optimising industrial production, reducing the difficulty of travel and transport and minimising production times and methods is inevitable. In this sense, CPT can be considered a deductive model founded on a series of postulates and axioms, and in this same sense it has been widely applied in the area of European and US territorial archaeology since the early nineteen seventies. Vide Johnson 1972: 769-785; Clarke 1977: 1-32.

[3] The problem which was nevertheless met more than once was that of poor consideration of the geomorphological aspects of the territory, which *de facto* involve a constant deformation of the theoretical model (hexagonal lattice) into more or less rhomboidal or trapezoidal shapes, making the causal explanation of the structural change highly questionable. Moreover, as far back as 1972, Clarke's *Gravity Model* simulated the intensity of the economic relationship between multiple settlements by defining an expression proportional to the product of the activities performed and inversely proportional to the cost of transport, seeking to reinterpret, in this manner, the meaning of the economic relationship in relation to the transport costs and direct productivity variables, present in the classic CPT formulation, but subordinate to its rigid spatial geometry. Cf. Clarke 1972: 7.

landscapes were soon felt; the first ecological models can, indeed, be understood as an attempt to shift the observation towards greater spatio-dimensional formalisation of the (theoretical) structure which could support more refined analysis of the (adaptive) mechanisms at the origin of the locational choices and the same socio-cultural transformations. As such, these mechanisms will later always be presented more as highly complex and non-linear phenomena which can be simulated on an ethnoarchaeological level by 'comparing' the spatial action typical of living cultures in modifying, changing and structuring the landscape; on an anthropological level by 'identifying' osmotic and/or reciprocal relationships between groups, or in other words a spatiality no longer just diffusive and sequential, but able to alter the geometry itself of the occupation; on a social level by 'predicting' the settlement developments by applying distribution curves, e.g. normal, originally employed for correcting geometric measurements and, finally, on a demographic level by 'interpreting' the influence of different conditions on the theoretical and regular growth trends[4].

4 The more culture is studied in biological terms, the less it is reduced to an automatic phenomenon and the reasons for its specific spatial distribution are pursued with more sophisticated simulations which tie the occupational process to the action of specific algorithms. Vide Allen 1997: 48-49; Foster 1989: 40-50; Chippindale 1992: 251-276. Moreover, some of the more recent research, aiming to experiment with connectionist computational models, employs the formal and conceptual elements peculiar to Cellular Automatons to simulate the dynamic complexity of the locational choices using *Bottom-up* logic. The emphasis in these cases is not placed on the evolution of the settlement system so much as on the learning capacities of the automata; their ecology is wholly artificial, but the choice of their location does not just depend on the environmental Input so much as on a complex relationship which connects sites, environment and experience. We must remember, however, that the rules of transition which drive the location choice and organisation of the territory are a hotly debated topic; in isolated cases, indeed, on employing integration between Cellular Automata and Neural Networks, the use of predictive functions is preferred, such as the normal (or Gaussian) distribution curve which provides a prediction of the trend as the mean and standard deviation change, using the limit theorem to minimise systematic and/or accidental errors. In archaeology, it has been widely and indiscriminately used both as a correction tool for geometric and topographic measurements and as an aid in verifying hypotheses on the economic and social behaviour of cultures. Vide Rogers 1962; Renfrew ed. 1984: 390-417.

2. Building Rules for the Analysis of Spatial Behaviour

Analyses and methods which use inductive inference tend "to *generate* Rules from the repeated observation of a Case", providing a formalisation (model) of this which identifies and selects them and allows others to be hypothesised. In *experimental* archaeology, these rules are stated as mathematical operations (equations, functions, algorithms) which offer reasonable hypotheses of "which causes" are behind the relationships between variables and which can generate other relationships with significance. The fact that each cultural context leaves the traces of a series of actions produced by ancient man on the territory, and that this evidence constitutes only the *trail* left so that we can trace those same actions, has given strong support in humanities research to the adoption of the circumstantial paradigm, better known as the hypothetico-deductive method. Analyses, procedures and models starting with this have always tended towards the formulation of a hypothesis or a series of hypotheses which could reveal (or *justify*) the events. While the first models nevertheless purported to formulate hypotheses based on the comparative observation of the cases, or by comparing 'the case' recovered with the living one in the cultures, over time a true method was refined which aimed to transform information into *evidence* and *evidence* into the apex of a network of semantic associations. This inference is particularly exploited today when the intention is to present 'reasonable' hypotheses of the spatial and temporal structure of the data – in the first case, the models constructed aim to supply a possible view of the physical causes causing distribution of the materials, while in the second the models generate a structural framework, generally phylogenetic, which observes the constraints imposed by the spatial structure. The search for informative distribution rules therefore led to a long debate in spatial geography, territorial archaeology, ethnogeography and ethnoarchaeology, but it currently seems to have been reduced to the suggestion to use models which are able to select which 'physical' conditions are behind the formation of the deposits and which 'theoretical' constraints are behind the adaptation. Nowadays, informative distribution models can be understood to mean any processing which can facilitate widespread understanding of the information by indicating where it is significantly lacking in intensity and predicting some of its structural typologies. These models converge on the selection of transverse (physical and theoretical) rules which allude to the existence of a 'natural' behaviour of the material culture and categorise those peculiar 'adaptive' behaviours of the group (or groups) which use them. Despite the fact that this commitment has produced enlightening hypotheses over the years on the associations which could connect the structure of spatial data to group adaptation, currently there is a tendency not to go too far in terms of the 'transversality' of the system, rather examining the complex physical, mechanical and natural causes. In the micro-space, on the other hand, the search for spatial articulation rules has, over the last decade, made use of generative models which trace (or assimilate) the occupation and, in particular, the construction of intra-site architectural spaces, to some processes, direct and inverse, in which 'nuclear' elements are added or removed by following certain constraints (or rules). These constraints (or rules) differ depending on the case, but all aim to make the complexity of a class of spatialised attributes as if they were 'generated' by a oriented relational process; progressive or inverse as they may be, these relationships always replicate a linear evolutionary trend. Given their simple behavioural mechanism, generative models find widespread use in archaeological and geographical research and have been used in the past both as tools for the automation of archaeological hypotheses at a given level of complexity (Doran 1972: 525-425), and as technical tools which identify the steps of each specific evolution or regression of the typologies (Chippindale 1992: 251-276). On rare occasions, in any case, the interpretation in studies thus designed comes from a shared linearity of the process for investigating the action of variables on logotechnics[5]. The network of semantic associations drawn from the physical relationships of the geological stratigraphy is also behind those models which aim to select the most suitable rules for the relative temporal placement of the document. That the position of the document in its stratigraphic level is the best evidence for defining it in terms of relative chronology is, on the other hand, evident if only we observe the many experiments which are performed presently, both to formalise the stratigraphical logic and to use IT techniques to refine the procedures for referencing the elements conserved in the archaeological deposits. The desire to save all evidence from the destructive mechanics of the archaeological dig therefore also favoured the introduction of analyses parallel to the usual ones which could provide integrated management of the position of the document in space and time and, in this manner, feed the chain of hypotheses and deductions necessary to perform any historical interpretation. The highly variable nature of the features present in some classes of artefacts posed the problem of tracing a classification for them which could subclassify them into other groups with relevance in the relative spatial and temporal placement of the artefacts themselves. For these reasons, the first 'combinatorial models' were implemented which, through

[5] It is nevertheless interesting to observe the growth of applicative research which integrates Geographic Information Systems and RNA, see: Black 1995: 159-166; Openshaw & Openshaw 1997; Fischer - Reismann 2002: 1-23; Zubrow 2004.

formalisation of the observation into decomposition rules, founded initially on the presence/absence relationship, would offer a structured description of them in calculation algorithms. But although this decomposition could have been sufficient for a reasoned 'deconstruction' and 'reclassification' of the highly specialised artefacts, it would not have added anything new to their relationship with other variables of their original archaeological context. This possibility was offered to experimental research precisely when it was decided to manage the contextual issues in an integrated manner through a more general theory. Therefore, it was allowed to the action to be applied by a variable on the system of connections tying it to further important characteristics of the context to be understood, these variables being similar to those of a constructed mechanical system or akin to those of a known natural one. Precisely, to trace the functions fulfilled by each given archaeological and/or geographical variable to a specific context, systematically structured, three paths were available to the analytical approach: to interpret the entire mechanism as if it were that of a theoretical machine, always functioning, in which the relationships between the various elements would be known; to hypothesise that each cultural and/or natural system had a biological life characterised by complex moments of growth, withdrawal and collapse; and to find the natural probability of the connections between the various elements of the system. The first case would offer an optimal and theoretical presentation of the syntax regulating the various parts of the system. A presentation by which the same variability of the aspects of a context, in order for it to be able to function 'mechanically', would be defined *a priori* (Systems Theory); in the second case, the system of variables of a given context would be superimposed on the functional cycles of the biological systems and therefore undergo their same growth rules up to the point of collapse (Catastrophe Theory); in the third case, the probability of the associations between variables of the cultural system would be codified, and their behaviour simulated with probabilistic networks (such as, to give one example from the many, *Bayesian networks*)[6]. Contemporary archaeological research does not limit itself to demonstrating the principles of complexity in cultural, economic and social systems, slowly beginning to work alongside radically different disciplines from the early sixties, for instance cybernetics which, by their own theoretical admission, live and develop for analysis of and experimentation on the rules of complexity (Ashby 1964; Hall 1989: 39-120; Gehlen 2003: 46; Rogers - McClelland 2004). This approach to other fields prompted the construction of a wide

---

[6] On the relationships between Computational Archaeology and Systems Theory, see: Gardin 1987: 27-42, Clarke 1994: 44-47; Bintliff 1997; Ramazzotti 2010: 171-198.

range of models which select the causes of complexity in systems and, therefore, attempt to represent their origins themselves. It should be noted that Archaeology and *κυβερνητική* (cybernetics) working together was neither sudden nor revolutionary (nor even linked to the advent of the American New Archaeology). it was rather the first, enlightening attempts to compare archaeological contexts with dynamic systems which promoted the inevitable 'intersection' of the two different disciplines. As in mechanical systems, the relationship between variables of a given context could therefore be described not only in recognisable geometric forms then redefined into theoretical maps which caused a distinctive 'activity', but from which also those Rules able to transform their internal connections and their organisation could be selected. Feedback, attractor, and dynamic equilibrium, which had already been studied by "Mechanical Intelligence" even before Cybernetics, would, on the one hand clarify the limit of automatic operation of the systems and, on the other, led towards the replacement of the mechanical (and linear) nature of the relationships with the physical-biological (non-linear) one of the 'connections'. The 'feedback' rule was selected, reviewed and discussed in many models which aimed to represent the complex formation of the State; the recognition of 'attractors' in as many models which aimed to recognise the homoeostasis of social and economic organisation in a chronologically and spatially localised state, and, moreover, the 'dynamic equilibria' in many others which emphasised the anomalies of a given territorial structure (Waldrop 1993; Page 2010).

3. Models for the Comparison of Spatial Behaviour

Models which perform analogical inference, "which *form* Hypotheses based on the comparison between Cases", are founded on, amongst others, analogies or similitudes, through which the intention is to grasp the logic of a particular phenomenon in relation to the operation of that which is better known or directly observable. In this sense, working them out requires a "symbolic capacity", in other words the ability to grasp the whole from the allusive evidence of the part (Gehlen 1983: 207-208). The models which use analogies have dealt with the relationship between the physical scattering of the materials in their contexts and the operation of the cultural systems which had produced them; the relationship between the behavioural variables of the cultural systems and the formal ones of the major economic theories; the relationship between the perceptions of the objects and their environment and those of the operation of memory and perception. The models which establish a comparison between Culture and Environment aim to present the transformation of cultural phenomena as being dependant on physical and biological laws which can be checked scientifically; but while most of this processing reaches that which, in theoretical archaeology is known as environmental determinism, a term which, indeed, highlights how a group of natural causes can 'condition' the structural change of the same social organisations, the confidence that these transformations could instead relate to man's action on the resources is owed to the department of prehistory at Cambridge, known as *Palaeoeconomy*. For the founder G. Clark, indeed, the comparison between the physical-biological laws which control characterisation of the environment and the cultural-economic laws which govern the social structures would lend itself to the construction of an integrated and global transformation model in which they would all interact; the discipline would therefore be responsible for the job of reorganising them in other different theories (local and contextual) of the individual processes. This model which, starting from a common base of principles, would allow hypothetical theories of the individual economic developments to be established, would be represented in a spatialised and structured system in which social organisation would occupy the centre of a network of connections (direct and inverse), with most variables depending on it or which could simply characterise it; a system nevertheless constrained at its base by the 'reciprocal' relationship between *Habitat* and *Biome*. The model generated by keen comparison between physical-biological laws of natural transformation and cultural-economic laws of social transformation, and spatialised as a complete system tied to interdependent relationships would therefore offer a precise and general theory of the economic and cultural operation regulated by systemic principles of structural equilibrium, or of homoeostasis, as explicitly stated by the archaeologist (Clark 1992: 162). As we have observed, the analogy between Culture and Environment, decoded into a systemic model by Clark, had a very large following in archaeological research aimed at reconstructing economic processes, but it was also relevant for introducing the *concrete* possibility of also comparing the principles which regulated cultural transformation with those which controlled natural transformation, an analogy which we could conventionally define as being second level, and whose analytical potential was noticed rather early by D. L. Clarke. From the comparison between the hypothetical operational rules of organic systems and those of cultural systems, the author of *Analytical Archaeology* sets out to undertake the demanding and brilliant conversion of cultural complexity into physical-biological complexity, the subject therefore of research out of analytical and, fundamentally, logical-mathematical requirements. This is the indelible mark left by his greatest work, which also led to the desire for a radical transformation of the cognitive morphology of the discipline. When we talk of conversion, we must not, however, simply mean an attempt to summarise cultural complexity, which Clarke was aware was, in a certain sense, fruitless. We should rather highlight how the models that he created always aimed to specify every assessable segment of cultures understood, conventionally, as Complex Systems. This measurability of relationships mapped out with the then pioneering aid of Cybernetics was *de facto* the first, and still unequalled attempt to offer an explanation for the operation which regulated the interaction between the parts of a system, an explanation which was able to make decisive use of the contribution of mathematics and the nascent new mathematics. Cybernetics, which Clarke explored as the direct expression of Systems Theory, would offer him an appropriate language to consolidate that second level analogy between Cultures and Organism and, in the early sixties, Cybernetics was essentially intended, and was actually born from the desire to conceive of the organic-biological function as the mechanical operation of interconnected parts; driven by an Input, these parts would be able to report the whole procedure which caused the alteration of equilibria and transformations and these would not be very dissimilar from those observable in the so-called Cultural Systems[7]. This mechanisation of cultural complexity, like its transfer onto the level of the mathematical discussion of the action performed by one or more factors on its entropy (Edelman 2004: 131), then inspired almost half a century of experimental archaeology and, above all in the United States, became a standard practice in archaeological research which was continually updated. Measurement of the difference between these two *operations*, nevertheless, over time, became ever more the study of cultural complexity through its reduction to groups of calculable parameters, and Clarke's central idea was, in a certain sense set aside. Insertion of cultural variability in the more refined and contemporary Systems Theories and Expert Systems took away from the search for other possible analogies which could be involved to solve highly complex problems and, above all, radicalised a single meaning of complexity itself, in other words its expression external to man, independent of human cognitive nature, a product existing in and of itself and the specific topic of the

[7] The computational models which today use analogue operators have been divided into three types: 1. Symbolic: defined as such since they refer to the paradigms of artificial intelligence and allow analogue codification of information to be examined by constructing more extended generic classes 2. Connectionist: since they process relationships between differentiated objects and classes of objects, allowing the degrees similarity to be measured 3. Hybrid: since they constitute architectures which integrate the functions of the first and second groups French 2002: 200-205.

research[8]. But at the end of the nineteen eighties, numerous studies resurfaced in the attempt to understand complexity no longer as external to man and subject of our predominantly applicative research, but rather as a living expression of our *intelligence*, our mnemonic, perceptive and creative capacity. In this sense, complexity was almost removed from the undisputed supremacy of external interpretation, able to be analysed through mechanical and linear systems, and became the subject of specific research which aimed to trace man's mental capacity to create it. The analogy between cultural complexity and the complexity of intelligence then gave rise to a new system of theoretical knowledge, methods and applications linking archaeological research to the New Artificial Intelligence (See now: Buscema – Tustle eds. 2013). Theories, methods and applications which are already in use and identify a whole new world of archaeology, which is not a paradigm of it, as cognitive archaeology aims to be, but a (contemporary) way to undertake the same historical reconstruction[9]. Indeed, while we certainly cannot debate the possibility to artificially recreate intelligence, it is equally evident that many models emulate and quite clearly come close to some segments of the cognitive process – memorisation, classification, perception, creation and reflection. Segments which, although they act in parallel and for which today architectures which are able to make them operated in an integrated manner are being studied, only allow the rules which control memory, classification, perception and reflection to be explained, rules which are no longer tracked down in the linear mechanisms of automatic operation, but in the networks which connect the known physical units of the brain, neurons[10]. Transferred to the level of the necessary logical-mathematical identity, these biological entities are defined as nodes, and the synapses which regulate their dynamic functions are called connections. The terms imply another important *conversion*, that of the biological-cognitive complexity of the world of intelligence into physical-cognitive complexity of the system of intelligence which, in this manner, favours the processes of analysis,

---

8 For these issues opened up by Computational Archaeology, see: Clarke 1962; Binford 1965; Clarke 1968; Gardin 1970; Clarke 1972; Level – Renfrew 1979: 145-167; Bintliff 1997; Barceló 2008; Ramazzotti 2010: 171-198; Ramazzotti 2012: 6-10.

9 For the issues opened up by Cognitive Archaeology, see: Zubrow 1994: 107-118; Doran 1996: 1231-1245; Gardin 1996: 1221-1230; Djindjian 2003: 178-199; Zubrow 2004; Malafouris - Renfrew 2010: 1-12; Ramazzotti 2010: 128-198.

10 On the operation of Artificial Neural networks, see: Minsky 1954; Minsky – Papert 1968; Amari – Arbib 1977; Grossberg 1982; Hopfield 1982: 2554-2558; Grossberg 1988; Anderson – Rosenfield 1988; Kosko 1992; McCulloch - Pitts 1993: 115-133; Arbib 1995; Kohonen 1995; Bishop 1995; Szczpaniak ed. 1999; Smolensky - Legendre 2006; Ehsani 2007; Nunes de Castro 2007: 1-36.

experimentation and simulation. Today, there is unceasing talk of *Computer Semiotics* as a discipline which aims to establish the function of the logical operators of programming on the basis of structured and complex semantic units, but the semiotic analyses centred on redefining the analytical object are also one of the main trends in *Computer Science* and, in particular, that sector interested in constructing nodes or cells which make up many of the artificial models which fit into the class of Artificial Adaptive Systems (Beckerman 1997; Miller – Page 2007; Ramazzotti 2012: 6-10), whether they be synthetic representations of the observed reality which must undergo interrogation processes (Expert Systems, Cellular Automata, Logical Networks) or the most advanced analytical tools for learning and modelling complex configurations (Artificial Neural Networks, Contractive Mapping, Genetic Algorithms). Given these elementary co-ordinates, it seems clear that simulating the behaviour (dynamic and complex) of the high variability of the cultural 'factors' in networks thus conceived equals tracking down, selecting and recreating (separately) a wide variety of functions which associate variables, a wide variety of inferences which control their semantic structure and an equally wide variety of causes which produce their transformation (Zubrow 2004; Bintliff 2005: 125-131; Barceló 2008: 154-184; Renfrew 2008: 2014-2047; Malafouris - Renfrew 2010: 1-12). This perception of functions, inferences and causes which multiply and generate the complex phenomena demands an archaeology concerned with interpreting the past by debating the history itself of its different perceptions and which, at the same time, attends to tracing the complexity of the culture by thwarting the classical and dualistic models to display all its extraordinary variability and richness. In this specific sense, the application of artificial intelligence models to archaeological problems has *value*: it recreates a possible world of other associations of meaning from the body devoid of sources and dispersed information, it exhibits the nuances and complex interrelations and, furthermore, it helps the interpretant codify others which were unforeseen (or hidden). In a certain sense, this it is a sort of metaphor with which we understand that the complexity of intelligence has to do with that of culture, a metaphor which we could keep on at forever without answers which, however, it is impossible not to 'feel'.

4. Cognitive and Spatial Behavioural Logic

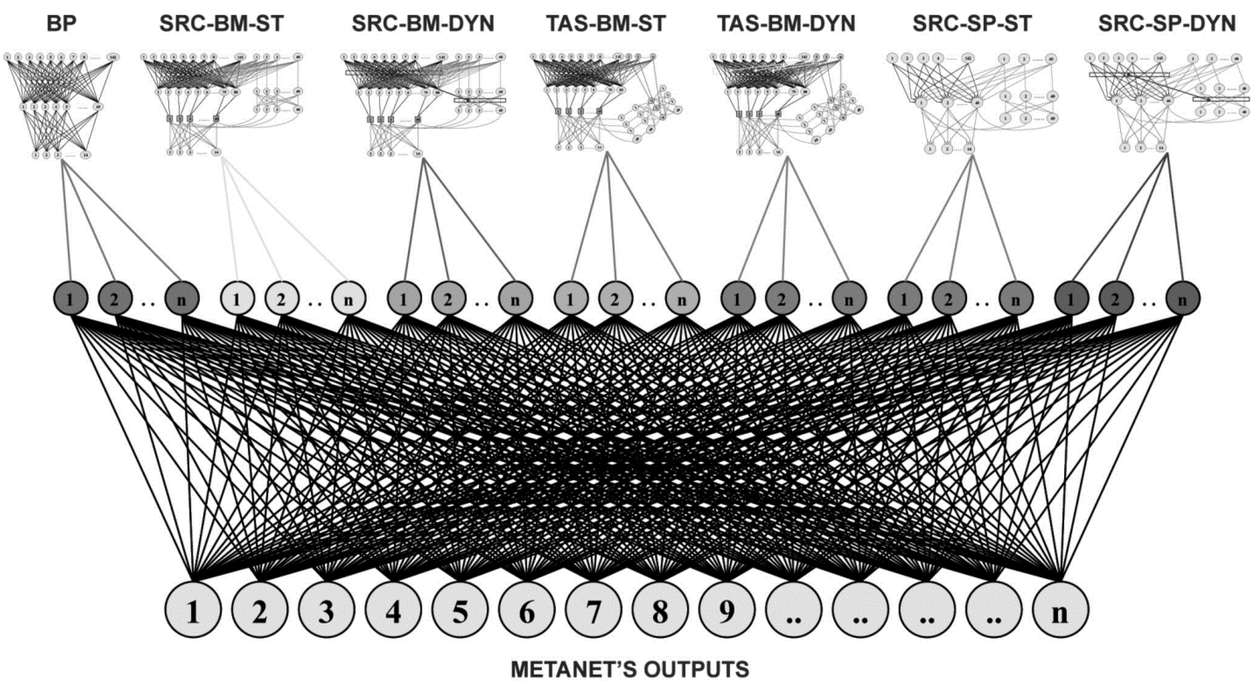
Analyses and methods which follow abductive inference generate theories on these cases from the possible results and offer all those representations (models) which outline a theory of the cultural, social, political, economic and cognitive function, looking for their foundations outside the perimeter of 'strictly' archaeological analyses, methods and techniques. These theories can be displayed as "closed systems" tied to a precise logical structure, or as "open systems" characterised by a dynamic combination of connections, but in both cases they always express a global and integrated interpretation of the processes, events and facts; they exhibit the set of relationships that these three 'elements' of the story have with the world of the present and the past; they change their morphological structure in relationship to the quantitative and qualitative growth of the documentation. From the results obtained in the biological area on the study of the rules of operation of the selective process, some archaeological research is starting to update the classical body of tools of linear and multi-linear neo-evolutionist theories and heading resolutely to set up a new semantics of cultural function. The attempt to deconstruct the functional complexity of archaeological documentation is indeed encouraging some experiments on 'generative grammar' which can be related to it, which is understood as the *codes* of a given spatial and temporal structure, necessarily constrained and therefore subject to the same rules of the selection process ingrained in the theory of evolutionary biology. Other, even more specific, studies examine the processes of cultural diffusion, aggregation and classification through the models of *coevolution*, thus identifying how they fulfil themselves by the interaction between 'genetic evolution' of species and the 'effect of human action' in influencing the exploitation of resources. Going beyond the limits of analysing archaeological cultures as concrete expressions of their documents and the transformation processes with the mechanical logic of human action on the environment, or conversely of the conditioning the environment might have on it, the possibility also stands out to define cultural diversity on a genetic basis, to interpret its slow transformations by making an equivalence between the concept of 'population' and that of 'culture' and establishing a spatial circumscription for it which is not just geographical, ethnic or linguistic. This diversity, which has many different contours and nuances to that only apparently recognised on ethno-linguistic foundations, necessarily subject to the possible manipulation of old and new ideologies, foreshadows a map of the ancient human *genome*, predominantly useful for medicine, but also sufficient to make the spatio-temporal relationships which have always focused, in one way or another, on the critical distance taken in relationship to some principles of equivalence, such as the classic one between 'culture and typology', more complex. The results obtained concerning the social complexity of the mechanisms which regulate cultural relationships not only produced a reduction of the classical application of Systems Theory to archaeology, but at the same time a trend of the discipline developed which, by updating that same theory with the aid of the new mathematics, aims to define some characters which predate cultural complexity itself and which, in the same way, can be examined again in light of that renewed theory. As already highlighted, beginning in the nineteen eighties, there was an attempt to dissolve the intrinsic rigidity of the regulatory mechanisms of the Systems by replacing the linear *Input-Output* function with some rules of anthropological relationships, considered more flexible in documenting the structural transformation of cultures (*Peer-Polity-Interaction*), or by clarifying the natural and biological direction followed by each system (Catastrophe Theory). But since the early nineties, that 'critical' path aiming to detail these limits and, therefore identify the mathematical rules which predict the ancient functional use of the objects was undertaken (*Rough Set Theory*). Some studies examine the dynamic oscillations of the processes of territorial organisation, employing models which organise themselves and which, as other cases present in nature show, tend to reach a stationary state of equilibrium, beyond which they transform themselves. Other, simulative ones trace the problem of location choices, traditionally dealt with using the principles of urban economics (agglomeration, accessibility, interaction, and hierarchy), back to the multiplicity of choices and the constraints determining this. The self-organisation of highly complex structures and the multifactorial nature of the choices which influence the territorial form of the systems are all recent phenomena which inevitably act to construct a new semantics of *cultural transformation*. Indeed, these 'phenomena' require overcoming the obstacle of an interpretation which reads the purely 'human' construction of ever-more complex systems to understand the cultural transformation and advance the illusion (or the ambition) of paying more attention to those, present in nature, which do not require external Inputs to act dynamically and transform themselves; in the same manner, they require the limit of the 'mechanical' construction of organisational complexity to be crossed, in which an 'Instruction' (or a few fundamental relationships) would be sufficient to transform the system in order to observe, classify, formalise and organise other rationalities of the locational choices. From the continuous results that contemporary Semiotics, Semantics and Logic obtain in shaping, deconstructing and reconstructing the world of meanings of structures, signs and symbols, their internal structure and the rules governing their perception and communication, some of the most recent archaeological studies have taken their cue to further theorise other information analysis procedures, propose other forms of their communication and generate other theories of their combination. These are also expressed in models, more structured than the ones we have already dealt with, because new research *epistemologies* are summarised and conserved within them. As far as analysis of the data is concerned, as we have already seen, Clarke, with the definition and elaboration of polythetic entities, had already supplied categories for learning, translating and transferring technical experience, seeing them in the grammar of the first semantic-perceptive models which were still, in the late seventies, largely structuralist. This first step towards a logic of sign transmission – which has been improperly left out of critical-archaeological literature and storiography – is presently right at the roots of that search for techniques and models suitable to transform the documentary archaeological situation into structures, codes and messages, both written and visual which can, in a certain sense, distinguish it. Gardin's constant attempts to prepare the logic of archaeology's theoretical-inferential reasoning are, indeed, operations destined to yield an anatomy of the perceptive mechanisms which are behind historical knowledge and its inevitable reworking. An anatomy of human historical reasoning which wishes to display all the strength of the subject, the arbitrariness and the conditioning of meaning and which, for this reason, has always offered itself to elaboration and simulation (Gardin 1970; 1980; 1996: 185-209; 1996: 1221-1230; 1987: 27-42).

Labelled as a speculative and sometimes ineffective attempt, this trend has, on the contrary, painted a picture of archaeology as logical-scientific research, always open to enquiry, aiming to build a 'metalanguage' which expresses its most profound nature. As far as the modalities of scientific communication are concerned, on the other hand, many archaeologists draw inspiration from the semiotic theories of communication to design and exhibit other forms of research. Thus, for example, Renfrew attempts to overcome the purely descriptive and analytical obstacles of investigation to relate the complexity of archaeological reasoning through the transmission of aesthetic experiences which organise new suggestive and effectual descriptions of the discipline, without, however, denying its internal order, its syntax, the organisation of work and the costs which, he presumes, are useless, boring and uninteresting details for the public (for specialists, in any case). More structured, on the other hand, is the network model proposed by Hodder, for whom horizontal circulation of archaeological information, established from a non-hierarchical, co-operative and collaborative work set-up combines perfectly both with the workable metaphor of IT networks, which without a centre and with many nodes are able to transfer the complexity of information, its frequency and its relativity, like a metalanguage, and with the literary metaphor of archaeology as narration, able to exhibit the plurality and the same antimony as the evaluations concerning the given object, as if to write the book on it, never unique nor conclusive. But above all, the results obtained, together, in the disciplines of Semantics, Semiotics, Logic and Neurobiology, gave rise to different theories of the so-called 'cognitive function' which were gathered first by Analytical Psychology and, only later, by Experimental Psychology. The contact between Archaeology and Cybernetics, which Clarke had already identified as central for archaeology to abandon its aura of innocence, is not therefore a purely generational fact, but the product of a *wonderful* intuition, through which the English archaeologist intended to reconnect the abyss which had formed, essentially due to the effect of the historical-cultural approach, between the Humanities and the Sciences, and in this way offer a new category of meaning to the definition itself of archaeology as 'human science' (and it is not by chance that today we group it with the social sciences). In the late sixties, as we have notice, Cybernetics was, in any case, simply understood as an extension of Systems Theory and Artificial Intelligence, then nascent, effectively representing the pioneering technological illusion of reducing the function of the logic of knowledge to a formalism which could be managed with that theory. Nevertheless, while initial research in Experimental Philosophy used guinea pigs as if they were automatic machines and applied multifactorial analysis to trace the differentiality of intelligence, beginning in the nineteen fifties neuroscience began to disown the paradigm of a brain regulated only by the modularity of electrical impulses (Oliverio 2004: 22-23). The division of the observed world into classes and forms (*Gestalt*) then began to be gathered into more complex models which attempted to integrate the awareness that some perceptive functions were not simply calculable, predicted or predictable actions or functions with the necessary linearity of automation and robotisation in 'naturalising' the cognitive function (innervating the experience in the elaborative process of the synapses between neurons). But the New Artificial Intelligence inspired, instead, by 'theoretical' neuroscience research and by 'practical' neurobiology research, in other words inspired by the interpretation of the cognitive function as a certainly more complex expression of a relation between the physical and irreducible elements of the brain, appeared only in the early nineties. Today, its foundations and origins are still being discussed, but it is at least unanimously recognised that its advancements in recreating some segments of 'knowledge' trace back to that area of the 'neurobiology of memory' which, having risen in radical opposition to the most ancient and traditional behaviourist school, was then defined as 'connectionism'[11]. Since this is, therefore, the border that any analysis of the principles which regulate memory, learning and classification (that is to say the fundamentals of all Theoretical and Experimental Archaeology) rests on today, back in the late nineties a series of studies was proposed, in the archaeological field, which had sensed the epistemological relevance of the connectionist models (Ramazzotti 1997; 1999; 1999: 261-269). Currently, not even a decade later, we can already distinguish at least two different directions – the first aims to explore the high level of complexity of archaeological processes, structures and systems, supporting the '*semantic* instrumental' value of the New Artificial Intelligence in rewriting a General Theory of Archaeology; the second instead shows and emphasises all the 'statistical instrumental' potential of the same models which are in reality replacing the more traditional mathematics of classification. Most research relating to both directions makes use of the principles by which every cultural expression is the reflection of conceptual and cognitive human processing (in the real, or not random sense), and that according to which human conceptual and cognitive expression can be reproduced through the construction and simulation of systemic and mathematical rules (in empirical, not dogmatic terms). Even though they leave open the whole problem of how non-calculable factors (such as emotivity) can alter any state of equilibrium, the first Networks, which are physical architectures of the complex relationships between the irreducible 'elements' of the brain, historically represent the point of reference of the two analogies – they are inspired by the biological model of 'knowledge' (hence 'neural') and

[11] "Connectionism" is understood as a theory born inside neurobiological research on natural intelligence and the reproduction of an artificial intelligence; for it, the brain is not ascribable just to a system of rules and symbols, but is composed of the operation of simple and non-intelligent elements known as "neurons", whose connections (synapses) express properties of coherence. The connectionist approach represents, in this sense, a reaction to the "behaviourist" and "representationalist" hypotheses which did not tackle the study of artificial intelligence based on the dynamic and connective relationship between the neurons, but by interpreting the logic of its operation in the mechanical Input-Output flow (vide. Minsky 1986; Rumelhart - McClelland 1986: 1-20; Ackley 1987; Touretzkey 1988; Fodor – Pylyshyn 1988: 3-81; Fodor 1975; 1985; Feldman – Ballard 1989; Marcus 2001; Clark – Eliasmith 2002: 886-888; McClelland et al. 2010: 348-356). In international archaeology, it was chiefly Doran's research which supported the applicability of this logic for studying social systems; of particular importance was the insertion of *Multiple Agents System Theory* (MAS) and *Distributed Artificial Intelligence Theory* (DAI) as the foundations for building dynamic socio-cultural models. Vide Doran 1970: 57-69; 1996: 381-393; 1997: 283-297). In particular, the second theory (DAI) aimed to conceptualise the structure of artificial scenarios to combine the representations of the individual cognitions of every agent and verify the qualitative characteristics of the birth of social hierarchies in the French Palaeolithic communities. Vide Doran - Palmer 1995: 103-125.

represent, in a structured manner, some of their aspects (or segments), which can be empirically checked, through a manifold variety of techniques and models (hence 'artificial'). In the psychological-cognitive area it has by now been shown that the organisation of reality by each individual does not just depend on the action of the environment on their formation, nor (vice versa) only on the action of their choices on the environment, but rationality is understood as functions of experience. Among these functions of experience, neuroscience studies memory, first in its evolutionary dynamic, then as a biological model and finally as the dependent nature of genetics. Observed as a biological model, memory is a product of the nervous circuits subjected to an experience, and in the area of dynamic mathematics it can be simulated in regulated (or self-regulated) systems. Associative Memories are, in effect, among the first processing mechanisms which learn by taking rules from complex systems and, as such, show themselves capable of finding possible solutions to non-linear problems. An archaeologist studying the complexity of archaeological processes through these models investigates the relational dynamics between classified variables of a given context 'like' a psychobiologist studies the complexity of the perceptive-analytical process of an individual (or a group of individuals) and a neurobiologist explores it on an empirical level. Rather than calibrate 'suitable' and 'expert' tools which repeat the relationship between variables like that present in connected parts of a machine, that archaeologist will therefore aim to trace the possible, nuanced and non-linear rules of their operation and will never obtain a single result, but rather 'analytical surfaces' (*hypersurfaces*) which will lend themselves (with humble, controllable repeatability) to historical-archaeological interpretation. In the psycho-cognitive field it is well known that the perception and organisation of reality by an individual increases and changes in relation to the quantitative and qualitative growth of the information and messages they exchange and receive. Phenomena such as *orientation*, which in psychobiology are also considered adaptive phenomena (dependent on learning and memory), can be simplified into highly complex and structured systems which change dynamically in relation with the increase in information. Some experiments performed with Cellular Automata demonstrate how these can identify the hazy rules which connect different groups of data and organise their rules by transforming them in relation to that given quantitative and qualitative (but nevertheless dynamic) growth or regression of the information (or the constraints). Their first applications to archaeological spatiality are therefore to be understood as simulations destined to explore, recognise, classify and typify the different spatial forms that a behaviour takes on in relation to the increase and decrease of the information (instructions) it receives or are subtracted from it, and in this sense is today driving the application of neural models for topology analysis of the forms of dynamic adaptation (Buscema – Breda et al. 2012; Buscema – Sacco et al. 2012). The archaeologist studying the complexity of archaeological processes through such models therefore intends to investigate the forms of the adaptation as a psychologist would investigate the perceptive-analytical reasons driving the organisation of a space (individual and collective) and a psychobiologist the physiological models which are at the origin of orientation; rather than therefore building closed memorisation systems, which gather all possible information on a given context, that archaeologist will design systems which are as 'open' as possible to receive the natural growth of information. Since this will not only become another data typology, but will also be able to be processed with other Artificial Intelligence models, its historical-archaeological interpretation will inevitably be subject to continuous updates, to the extent necessary. In the experimental and cognitive psychology area, an individual's process of classifying reality occurs by effect of their capacity owing to the experience they have acquired to operate 'summaries' and 'generalisations'; as such, the organisation of forms into categories of meaning is considered a complex phenomenon which provides for mnemonic and learning skills. The Artificial Neural Networks which have been trained to trace those rules which structure a given complex system can also perform corrected generalisations on the System, redefining them into other relational classes. Their first applications in the seriation of archaeological data, a classification which necessarily increases in complexity in relation to the increase in information, are therefore also to be understood as attempts to apply a given individual's adaptive skills and experience to organising the reality surrounding them. In the area of experimental and cognitive psychology, an individual's predictive abilities depend on the dynamic behaviour and structure of the variables they have at their disposal to take on the solution to a given problem; this behaviour therefore focuses into a response which transforms the input information into highly structured surfaces which represent form and substance of the observed reality. Artificial Neural Networks which have been trained to trace the nuanced rules of a given complex system can construct a surface geometry of their learning, which changes in relationship to the qualitative and quantitative characteristics of the other stimuli received (Buscema – Breda – Lodwick 2013). Those trained to define the rules of a given complex system therefore offer a possible generalisation which lends itself, at a later date, to being interrogated using 'prototypical questions' in order to delineate a given object of investigation in quantum terms, both by modifying the number and intensity of the inputs (simple interrogation) and by adding other possible (complex) ones and observing how that representation (surface) changes structure with each response. Their first applications to predictive archaeological problems are therefore to be understood as attempts to apply the diagnostic abilities of an individual to transform what they observe into a possible prediction of its function. The archaeologist who simulates the dynamic behaviour of a Complex System through these New Artificial Intelligence models intends to explore the configuration of the data (which has been learned) as an *analytical surface*, in a similar way to that of a psychologist investigating the perceptive-analytical processes of an individual's predictive potential. Rather than therefore describing the purely systemic complexity of a given context, they will aim to work on how that context was learned in order to interrogate it in a diversified manner and trace every possible combination of it, thus providing a wide cross-section of historical-archaeological predictions of its 'state' (fig. 1.2.).

Figure 1.4. METANET general topology. © Semeion Research Center. Da Buscema et al. 2013: 144, Figure 5.1.



5. Conclusions: Analytical Archaeology of Complex Phenomena

The study of complex archaeological systems which can make use of the *philosophy* of New Artificial Intelligence is, ultimately, a research project which *evaluates* the historical meaning of the relationships between archaeological documents as an essentially human construction which repeats, in this, a *strong* position of Analytical Archaeology, but updates it on the basis of the progress which Cognitive Science, Neuroscience and Cybernetics have made in simulating the principles which regulate memory, orientation, classification and interpretation of reality. It is important to highlight that these models, unlike others, must make use of a precise *encoding* of the documents and take on an important role in the research only when the results which they produce become the *hyper*-surface to continue, update, refine or open the analysis itself. Some considerations, after this brief, subjective and critical introduction to the history and use of the models in archaeology are necessary before concluding this paper. On the other hand, these models, as the reader will have understood, almost always introduce new problems, even only for the fact that they are subject to constant re-examination. In this sense they irremediably prevent the closure of the research. Even though today those researching theories seem to be the ones most in use in experimental and applicative studies, it should not be considered that they form a new 'paradigm' as they respond, like others, to specific questions which the past asks of man. If this were the case, if they really did represent a paradigm, we would indeed have to think that those questions had diversified, developed, become more complex when, on the contrary, they always fit into the all too human categories of enquiry. That which is changing, however, and decidedly so, is man's relationship with the technique which is now transforming natural reality from the inside, moving it onto a first artificial, then virtual plane, where everything, within the (desirable) limits imposed by ethics, is *apparently* possible. Our action on this new world is a field still to be explored, on a historical, anthropological and archaeological level, but already – by observing our models – we seem to perceive at least the formation of different research areas into communication. The deductive models are headed towards a check or an *anticipation* of the possible context and they, known to the discipline right from its creation, today show the desire to orient themselves in the world which is created and replicated. They spread out from underlying convictions, as old as the postulates and their axioms, they communicate a certainty that the referent can grasp in all their lucidity and rationality. The inductive models which continue to produce rules from observation, on the other hand, arrange themselves (given their specific nature) in a more chaotic manner; they do not have those certainties, but they always live in the *experiment*, communicating the state of a new discipline, or one always in the course of renewal. Those which inflect analogy, on the other hand, resist and grow in this reality codified in signs and symbols precisely because they require "symbolic capacity". When it would seem that the *hiatus* which every analogy entails might fade away, here the 'metaphor', the 'similitude', the 'allusion' appears. In the *technical age* it is even too predictable that the last perceptible limit is still that of the relationship (metaphorical, nuanced or allusive) between 'mind and machine'. Besides, in this age, it is almost instinctive to replicate the function of knowledge, to retrieve its origin and to hypothesise a backstory for it. The models which, on the other hand, have searched for a place in the discipline by drawing their inspiration from other far away things and at the same time from the theories which emerged and tried to explain cognitive function, would, in the technical age, be absorbed by the recreation, even though minimal or "impossible", of intelligences, first Ancient and then the New Artificial Intelligence. The other they would be inspired by was *reason as a tool* and, in the technical age, this becomes, today, the condition for interpreting and communicating man's historical, archaeological and anthropological complexity as well.

6. References

Ackley, D. H. 1987, Connectionist Machine for Genetic Hill Climbing, Boston: Kluwer Academic Publishers.

Allen, J. 1999, Spatial Assemblages of Power: from Domination to Empowerment, in Massey, D., Allen, J., Sarre, P. (eds.), Human Geography Today, Cambridge: Wiley, pp, 194-218.

Amari, S., Arbib, M. A. 1977, Competition and Cooperation in Neural Nets, in System Neuroscience, in Metzler, J. (ed.), New York: Academic Press, pp. 119-165.

Arbib M. A. 1995, The Handbook of Brain Theory and Neural Networks, Cambridge Ma.: The MIT Press.

Ashby W. R. 1964, An Introduction to Cybernetics, London: Methuen.

Anderson J. A., Rosenfeld, E. (eds.) 1988, Neurocomputing Foundations of Research, Cambridge Ma.: The MIT Press.

Barceló J. A. 2008, Computational Intelligence in Archaeology. Investigations at the Interface between Theory, Technique and Technology in Anthropology, History and the GeoSciences, London: IGI Global.

Beckerman M. 1997, Adaptive Cooperative Systems, New York: Wiley-Interscience.

Binford L. R. 1965 Archaeological Systematics and the Study of Culture Process, *American Antiquity* 31/2, pp. 203-210.

Bintliff J. 1997, Catastrophe, Chaos and Complexity: the Death, Decay and Rebirth of Towns from Antiquity to Today, *Journal of European Archaeology* 5, pp. 67-90.

Bintliff J. 2005, Being in the (Past) World: Vermeer, Neural Networks and Archaeological Theory, in T. L. Kienlin (ed.), Die Dinge als Zeichen: Culturelles Wissen und materielle Kultur, Bonn: Verlag Dr Rudolf Habelt, pp. 125-131.

Bishop C. M. 1995, Neural Networks for Pattern Recognition, Oxford: Oxford University Press.

Black W. R. 1995, Spatial Interaction Modelling using Artificial Neural Networks, *Journal of Transport Geography* 3/3, pp. 159-166.

Buscema P. M., Tastle W. J. 2013, Intelligent Data Mining in Law Enforcement Analytics: New Neural Networks Applied to Real Problems, Heidelberg - New York: Springer.

Buscema P. M., Breda M., Lodwick W. 2013, Training with Input Selection and Testing (TWIST) Algorithm: A Significant Advance in Pattern Recognition Performance of Machine Learning, *Journal of Intelligent Learning Systems and Applications* 5, pp. 29-38

Buscema P. M., Breda M., Grossi E., Catzola L., Sacco P.L. 2012, Semantics of Point Spaces through the Topological Weighted Centroid and Other Mathematical Quantities-Theory & Applications. In Data Mining Applications Using Artificial Adaptive Systems; Tastle, W., Ed.; Springer: New York, NY, USA, 2012.

Buscema P. M., Sacco P. L., Grossi E., Lodwick W. 2012, Spatiotemporal Mining: A Systematic Approach to Discrete Diffusion Models for Time and Space Extrapolation. In Data Mining Applications Using Artificial Adaptive Systems; Tastle, W., Ed.; Springer: New York, NY, USA, 2012.

Buscema P. M., Tastle W. J., Terzi, S. 2013, Meta Net: A New Meta-Classifier Family, in Buscema P.M., Tastle W. J. 2013, Intelligent Data Mining in Law Enforcement Analytics: New Neural Networks Applied to Real Problems, Heidelberg - New York: Springer, pp. 141-182.

Chippindale, C. 1992, Grammars of Archaeological Design: A Generative and Geometrical Approach to the Form Artifacts, in Gardin, J.-C. e Peebles C. S. (eds.), Representations in Archaeology, Bloomington: Indiana University Press, pp. 251-76.

Clark A., Eliasmith C. 2002, Philosophical Issues in Brain Theory and Connectionism, in Arbib M. (ed.), Handbook of Brain Theory and Neural Networks (2nd ed.). Cambridge Ma.: The MIT Press., pp. 886-888.

Clarke D. L. 1962, Matrix Analysis and Archaeology with Particular Reference to British Beaker Pottery, *Proceedings of the Prehistoric Society* 28, pp. 371-383.

Clarke D. L. 1968, Analytical Archaeology, London: Methuen.

Clarke D. L. 1972, Models and Paradigms in Contemporary Archaeology, in D. L. Clarke (ed.), Models in Archaeology, London: Methuen, pp. 1-57.

Clarke, D. L., 1973, Archaeology: the Loss of Innocence, *Antiquity* 47, pp. 6-18.

Clarke D. L. 1994, Culture as a System with Subsystem, in Pearce, M. S. (ed.), Interpreting Objects and Collections, New York – London: Routledge, pp. 44-47.

Djindjian F. 2003, Modéles «Cognitifs» et modéles «paradigmatiques» en archéologie, in AA. VV. (eds.), I modelli nella ricerca archeologica. Il ruolo dell'Informatica (Roma, 23-24 novembre 2000), Accademia Nazionale dei Lincei, Anno CCC – 2003, Contributi del Centro Interdisciplinare "Beniamino Segre", n. 107, pp. 178-199.

Doran J. E. 1970, Archaeological Reasoning and Machine Reasoning, in Gardin J.-C. (ed.), Archéologie et Calculateurs. Problèmes sémiologiques et mathématiques, Paris: Ed. du CNRS, pp. 57-69.

Doran J. E. 1972, Computer Models as Tools for Archaeological Hypothesis Formation, in Clarke, D. L. (ed.), Models in Archaeology, London: Methuen, pp. 425-451.

Doran, J. E. 1996 Simulating Societies using Distributed Artificial Intelligence, in Troitzsch, K. G, Mueller, U., Gilbert, G. N., Doran, J. E., (eds.), Microsimulation, Berlin, pp. 381-393.

Doran, J. E. 1996, Artificial Societies and Cognitive Archaeology, *Archeologia e Calcolatori* 7, pp. 1231-1245.

Doran, J. E. 1997, Distributed Artificial Intelligence and Emergent Social Complexity, in van de Leeuw, S. E. McGlade, J. (eds.). Time, Process and Structured Transformation in Archaeology, London: Routledge, pp. 283-297.

Doran, J. E., Palmer, M. 1995, The EOS Project: Integrating two Models of Palaeolithic social Change, in Gilbert, N., Conte, R. (eds.), Artificial Societies: the Computer Simulation of Social Life, London, pp. 103-125.

Ehsani 2007: A. H. Ehsani, Artificial Neural Networks: Application in Morphometric and Landscape Features Analysis, Stockholm: Byggvetenskap, Kungliga Tekniska högskolan.

Feldman J. A., Ballard D. H. 1989, Connectionist representation of concepts. In R. Pfeifer, Z. Schreter, F. Fogelman-Soulié e L. Steels (eds.), Connectionism in Perspective, Amsterdam: North-Holland, pp.

Fischer M. M., Reismann M. 2002, A Methodology for Neural Spatial Interaction Modeling, *Geographical Analysis* 34/3, pp. 1-23.

Fodor J. 1987, Psychosemantics: The Problem of Meaning in the Philosophy of Mind, Cambridge Ma.: The MIT Press.

Fodor J., Pylyshyn Z. 1988, Connectionism and Cognitive Architecture: A Critical analysis, *Cognition* 28, pp. 3-71.

French R. M. 2002, The Computational Modeling of Analogy-making, *TRENDS in Cognitive Sciences* 6/5, pp. 200-205.

Gardin J-C. 1970: Archéologie et calcolateurs: problèmes mathématiques et sémiologiques, Paris: Éditions du Centre national de la recherche scientifique.

Gardin J-C. 1980, Archaeological Constructs. An Aspect of Theoretical Archaeology, Cambridge: Cambridge University Press.

Gardin J.-C. 1987, De l'analyse logiciste aux systèmes experts, in J.-C. Gardin, O. Guillame et al (eds.), Systèmes experts et sciences humaines: le cas de l'archéologie, Paris : Eyrolles, pp. 27-42.

Gardin J-C. 1996, Formalisation et simulation des raisonnements, Revel, J., e Wachtel, N. (a cura di), Une école pour les sciences sociales, ehess, Paris: Cerf, pp. 185-208.

Gardin J.-C. 1996, La révolution cognitive et l'archéologie, *Archeologia e Calcolatori* 7/2, pp. 1221-1230

Geheln A. 1940, L'uomo, la natura e il suo posto nel mondo, Milano: Feltrinelli.

Geheln A. 1957, L'uomo nell'era della tecnica, Roma: Armando.

Grossberg S. 1982, Studies of Mind and Brain: Neural Principles of Learning, Perception, Development, cognition, and motor control (Boston Studies in the Philosophy of Science, Vol. 70), Dordrecht: Reidel

Grossberg S. 1988, Neural Networks and Natural Intelligence, Cambridge Ma.: The MIT Press.

Hall R. 1989, Computational Approaches to Analogical Reasoning: a Comparative analysis, *Artificial Intelligence* 39, pp. 39-120.

Hopfield J. J. 1982, Neural Networks and Physical Systems with Emergent Collective Computational Abilities, *Proceedings. Natl. Acad. Sci. USA*. 79/8, pp. 2554-2558.

Johnson, G.A. 1972, A Test of Utility of Central Place Theory in Archaeology, in Ucko, P. J. et al. (edd.), Man, Settlement and Urbanism, London: Duckworth, pp. 769-785.

Kohonen, T. 1995, Self-Organizing Maps, Berlin-Heidelberg: Springer-Verlag.

Kosko B. 1992, Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence, NJ.: Prentice Hall, Englewood Cliffs.

Leroi - Gourhan, A. 1977, Il gesto e la parola. La memoria e i ritmi, Torino: Einaudi.

Level, E. V., Renfrew, C. 1979, Exploring Dominance: Predicting Polities from Centers, in Renfrew, C., Cooke, K. (eds.), Transformation: Mathematical Approaches to Culture Change, New York: Academic Press, pp. 145-167.

Malafouris L. & Renfrew C. 2010, The Cognitive Life of Things: Archaeology, Material Engagement and the Extended Mind. In L. Malafouris & C. Renfrew (eds.), The Cognitive Life of Things: Recasting the Boundaries of the Mind., Cambridge: McDonald Institute Monographs, pp. 1-12.

Marcus G. F. 2001, The Algebraic Mind. Integrating Connectionism and Cognitive Science, Cambridge Ma.: The MIT Press.

McClelland J. L., Rumelhart D. E. (eds.), 1986, Parallel Distributed Processing. Explorations in the Microstructure of Cognition, Vol. I, Foundations, Vol. II. Psychological and Biological Models, Cambridge: The MIT Press.

McClelland J. L., Botvinick M. M., Noelle D. C., Plaut D. C., Rogers T. T. Seidenberg M. S. and Smith L. B., 2010, Approaches to Cognitive Modeling. Letting Structure Emerge: Connectionist and Dynamical Systems Approaches to Cognition. Cognitive Sciences» 14, pp. 348-356.

McCulloch W. S, Pitts W. H. A 1993, A Logical Calculus of the Ideas Immanent in Nervous Activity, *Bull. Math. Biophys* 5, pp. 115-133.

Miller J. H., Page E. S. 2007, Complex Adaptive Systems. An Introduction to Computational Models of Social Life (Princeton Studies in Complexity), Princeton: Princeton University Press.

Minsky M. 1954, Neural Nets and the Brain-Model Problem, Doctoral Dissertation, Princeton: Princeton University Press.

Minsky M 1986, The Society of Mind, New York: Simon & Schuster (Trad. Ital. La società della mente, Milano: Adelphi 1989).

Minsky M., Papert, S. 1968, Perceptrons, Cambridge Ma.: MIT Press (Expanded Edition 1988).

Nunes de Castro L. 2007, Fundamentals of Natural Computing: an Overview, *Physics of Life Reviews* 4, pp. 1-36.

Openshaw S. & Openshaw C. 1997, Artificial Intelligence in Geography, Chichester: John Wiley.

Page, S, E. 2010, Diversity and Complexity (Primers in Complex Systems), Princeton: Princeton University Press.

Pessa E. 1992, Intelligenza artificiale. Teorie e sistemi, Torino: Bollati Boringhieri.

Ramazzotti M. 1997, La fase 'Middle Uruk': studio tramite Reti Neurali Artificiali su un orizzonte latente nella protostoria della Bassa Mesopotamia. In Matthiae P. (ed.), Studi in memoria di Henri Frankfort (1897-1954) presentati dalla scuola romana di Archeologia Orientale, Roma: Sapienza Ed., pp. 495-522.

Ramazzotti M. 1999a, La Bassa Mesopotamia come laboratorio storico in età protostorica. Le Reti Neurali Artificiali come strumento di ausilio alle ricerche di archeologia territoriale (= Contributi e Materiali di Archeologia Orientale VIII), Roma: Sapienza Ed.

Ramazzotti M. 1999b, Analisi qualitativa dei depositi archeologici come indice guida delle ricerche a scala territoriale, in Reti Neurali Artificiali e sistemi sociali complessi. Teoria - Metodi - Applicazioni. Edited by M. Buscema, Voll. II, Milano: Franco Angeli, pp. 261-269.

Ramazzotti M. 2010, Archeologia e Semiotica. Linguaggi, codici, logiche e modelli, Torino: Bollati Boringhieri.

Ramazzotti M. 2012, ARCHEOSEMA. Un modello archeo-logico per la ricerca teorica, analitica e sperimentale dei fenomeni complessi, *Archeomatica* 2, pp. 6-10.

Renfrew C. 1984, Trade as Action at Distance, in Renfrew, C. (ed.), Approaches to Social Archaeology, Cambridge Ma.: The MIT Press, pp. 86-155.

Renfrew C. 2008, Neuroscience, Evolution and the Sapient Paradox: the Factuality of Value and of the Sacred, *Phil. Trans. R. Soc. B* 363, pp. 2041-2047.

Rogers T. T., McClelland J. L. 2004, Semantic Cognition: A Parallel Distributed Processing Approach, Cambridge Ma.: The MIT Press.

Rosenblatt F. 1958, The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain, *Psychological Review* 65/6, pp. 386-408.

Rothman 1987: M. S. Rothman, Graph Theory and the Interpretation of Regional Survey Data, *Paléorient* 13/2, pp. 73-91.

Singh J. 1966, Great Ideas in Information Theory, Language and Cybernetics, New York: Dover Publications Inc.

Smolensky P. - Legendre G. 2006, The Harmonic Mind: From Neural Computation to Optimality-theoretic Grammar, Vol. 1: Cognitive Architecture, Vol. 2: Linguistic and Philosophical Implications, Cambridge Ma.: The MIT Press..

Sternberg, R. J. 1985, Teorie dell'Intelligenza. Una teoria tripolare dell'intelligenza umana, Milano: Bompiani.

Szczpaniak P. S. (ed.) 1999, Computational Intelligence and Applications (= Studies in Fuzziness and Soft Computing, Vol. 33), Heidelberg - New York: Springer.

Waldrop M. M. 1993, Complexity, The Emerging of Science at the Edge of Order and Chaos, London: Simon & Schuster.

Zubrow E. B. W. 1994, Knowledge Representation and Archaeology: a Cognitive example using GIS, in Renfrew C., Zubrow E. B. W. (eds.), The Ancient Mind: Elements of Cognitive Archaeology, Cambridge: Cambridge University Press, pp. 107-118.

Zipf, G. K. 1949, Human Behaviour and the Principle of Least Effort, Cambridge: Hafner Publishing Co Ltd.

Zubrow E. B. W. 2004, The Archaeologist, the Neural Network, and the Random Pattern: Problems in Spatial and Cultural Cognition. In Forte M., Williams P. R. (eds.), The Reconstruction of Archaeological Landscapes through Digital Technologies Italy-United States Workshop , (BAR S1151), Oxford: Archaeopress.

Marco RAMAZZOTTI, Dipartimento di Scienze dell'Antichità, Sapienza University of Rome, Italy.

email: marco.ramazzotti@uniroma1.it

Ramazzotti, Marco, "Logic and Semantics of Computational Models for the Analysis of Complex Phenomena. Analytical Archaeology of Artificial Adaptive Systems". In: Armando Montanari (ed.), *Urban Coastal Area Conflicts Analysis Methodology: Human Mobility, Climate Change and Local Sustainable Development*. Cap. 1. *SECOA FP7 Research Project*, Vol. 5. Rome: Sapienza Università Editrice, 2013. *http://digilab-epub.uniroma1.it*. DOI: 10.13133/978-88-98533-01-5. Web. 30 September 2013.

ABSTRACT: This paper is a further attempt to apply Artificial Adaptive Systems to the analysis of complex natural and cultural phenomena through the lens of Analytical Archaeology. These phenomena are essentially understood to be the product of cognitive behaviour, in other words aspects, models and ideal types which represent it and can be analysed on a formal logical level. This introductory exploration leads to a strong syntactic diversification of logical inferences and a progressive human attempt to trace them back to the simulation of cognitive complexity. Artificial Adaptive Systems, as mathematical tools which express these emulative properties, are historiographically animated in the connectionist reaction to behaviourism and therefore effectively shape the social sciences' attempts to ascribe the complexities developed by our brains to advanced, non-linear and dynamic computational models.

KEYWORDS: Logic, Semantics, Computational Models, Complex Phenomena, Analytical Archaeology, Adaptive Artificial Systems, Artificial Intelligence.

Submitted:	6th June 2013
Accepted:	18th July 2013
Final Revision:	23th July 2013
Published:	30th September 2013

Annex I.

Annex I.

CHAPTER 2.

GIS and SOM Application in Different Scientific Areas

Irene Viaggiu

Department of European, American and Intercultural Studies, Sapienza University of Rome, Italy

1. An Introduction to GIS and SOM

The huge amount of spatial data generated by software GIS development, the increasing number of geographic informatic applications available, the computerization of a large amount of information sources, and the availability of digital maps has increased the opportunity and need for the utilization of methods for spatial classification, for both research and applied purposes (Dardala, Reveiu 2011).

Artificial Neural Networks (ANNs) can be used to develop a classification procedure which blends traditional statistical methods with a machine learning approach, allowing the system to iterate over a collection of datasets until patterns can be learned and realized (Coleman 2009).

This concept introduces the notion of "adaptive learning" and is especially effective for evaluation of the data that continuously evolves in time and space, such as meteorology, vehicle traffic, the spread of disease, and landuse dynamics.

"The Self-Organizing Map (SOM) is an unsupervised ANN that projects and maps high-dimensional, complex, linear, or nonlinear data into iteratively organized clusters in a topology-preserving geometric structure for the creation of a low-dimensional discrete data space" (Coleman 2009).

As many architectures of neural networks, the SOMs have biological basis that inspire the model. A similarity between a biologic system and an artificial system, such as of the SOMs is the macroscopic level of functioning of the brain. In the human brain, during the development stages, a subdivision in specialised areas for input *stimuli* occurs. In a similar manner the SOMs, during the learning process, there is the creation of a map in the layer of Kohonen (the Finnish academician father of this computational method), constituted by Processing Element (PE) disposed spatially in ordered way.

We can have one-dimensional, bi-dimensional, three-dimensional layers of Kohonen and also more than three dimensions. The typical dimensionality for the layer of Kohonen is of two dimensions. This PE layer evolves during the learning, specializing the positions of the single PEs as indicators of the statistical characteristics relevant for the input stimulus. This process of spatial organisation of the characteristics of input data is called also Feature Mapping. The SOMs realize the feature mapping with a technique of non-supervised learning, from which the name that indicates the self-organization. Thus the SOM is also known in literature as Kohonen Feature Map.

"The added value of the SOM is its ability to uncover hidden data patterns, structures, and relationships in multivariate datasets. Because the SOM classifies data in an unsupervised form, no training data are presented to the network; thus no *a priori* knowledge about the data distributions or placement of data into discrete output space is incorporated". (Coleman 2009).

The SOM structure is obtained by iteratively presenting the same input data signals to the network and adjusting the network weights to create "meaningful order, as if some feature coordinate system were defined over the network" (Kohonen 2001).

SOM is also a well-developed neural network technique for data clustering and visualization (Jiang, Harrie 2004). That is, data samples that are close to each other in the input space are also close to each other on the low dimensional space. In this sense SOM is like a geographic map on the distribution of phenomena, in particular regard to First Law of Geography: everything is related to everything else but near things are more related to each other (Tobler, WR 1970).

An integration of Self-Organizing Maps (SOM) into a Geographic Information Systems (GIS) data model is an area that has received particular attention, and is critical for information summarization. SOM provides a versatile platform upon which various data exploration methods can be built efficiently. The training file (SOM prototype data) has potential use in spatially-oriented databases.

The list of papers relating to the Self Organizing Maps (SOM) is reported by Saski, Kangast, Kohonen (1998) for the period 1981-97, by Oja, Kaski and Kohonen (2002) for the period 1998-2001, and by Pöllä, Honkela, Kohonen (2009) for the period 2002-2005 for a total of 7768 titles in total. Of this large amount of reference material only papers of Aggarwal, Lin, Kim (2004), Hatzichristos (2004), Lin, Aggarwal, Kim (2004), and Haidong Yang, Yueming Hu, Fei Qi Deng, Xian Tian, and Baorong Li (2004) explicitly refer to the application of GIS to the SOM.

The GIS data models are computerized encodings of abstracted forms of geographic space based on a simple graph. Although geographic space is modelled using objects and fields, the formalization of these models in a computer is obtained in a different way.

Formalization of continuous space is encoded by approximations based on tessellations (Samet, H 1995). Object-like geographic objects are encoded with appropriate vector data structures.

GIS data model encodes interactions between spatial and attribute information based on topological relationship directional, and metric.

The total of the GIS data model has certain key features and supports vector and raster data structures, geometric properties, algorithms, and database functions-data storage and retrieval.

2. Applications

Technology has created a shift to an extremely data rich environment with significant increases in the resolution and warehouse of space-time attribute data. This challenges traditional analysis with exponential demands on computational resources. Data mining and knowledge construction address this issue by reducing the volume and complexity of data.

However, these methods do not indicate the significance of their results and thus require the intervention of an expert. This means that solutions to the growth of data require both information system and human intelligence. These models have some spatial qualities, allowing them to be combined with geographic information systems (GIS) to provide practical interactive solutions for the analysis and exploration of large high-dimensional datasets.

That's why the use of GIS and SOM are well suited to the studies of different scientific areas and of course to the more strictly geographic and cartographic. Here are just some of the fields in which they are now used this formidable combination:

1. **Archaeology**: "Most archaeological information is spatial in nature, because it deals with the placement of archaeological finds, contiguity and neighbourhood relationships between archaeological entities" (Barceló, Pallares 1996).
2. **Geosciences** (Bação, F., Lobo, V., Painho, M. 2005)
3. **Market and finance** (Carlson, E. 2000).
4. **Medical research**: Community health is a non-linear function of environmental and socioeconomic effects that are not normally distributed among communities. The integration of multivariate data sets representing social, economic, and physical environmental factors allows us to evaluate the hypothesis that communities with similar environmental characteristics exhibit similar distributions of disease (Basara, Yuan 2008).
5. **Social area** (Yan Li 2007).

The use of GIS maps and self-organization is not limited to "simple" data analysis, but according to numerous studies published in recent years can be used as reliable predictive system for urban development models.

The evolution of urban systems is characterized by complex non-linear relationships between socio-economic attributes of spatial interaction and land use.

Traditional urban models have had many limitations in the simulation these urban dynamics. Nevertheless the last decade has seen a rise in a new class of models, which aim to address the shortcomings of their predecessors.

The system incorporates SOM algorithms within an existing GIS application to function as a design and decision support system. Urban data is simulated in the neural net and correlated, in varying degrees, with the data obtained from case studies relative to other regions. This gives the user a chance to see and understand the impacts of the proposal of research, which is otherwise difficult to imagine because of its complexity.

The ever increasing volume of studies and related outcomes of this research field demonstrate the importance of this methodology. Also the calculations based on SOM and GIS seem to be able to perceive and reliably predict the spatial logic for future urbanization and development processes can be a guiding tool to identify appropriate policies.

3. Modelling: A General Overview

Models are of fundamental importance in both scientific and economical contexts. The great importance of models such as the Bohr model, the MIT bag model of the nucleon, the Lorenz model of the atmosphere, the double helix model of DNA, the models based on evolutionary agents in the social sciences, and models general equilibrium of markets in their respective fields mark the rhythms of scientific developments of the last decades. The models are undoubtedly one of the main tools of modern science (Frigg-Hartmann 2012).

Philosophers are recognizing the importance of models with increasing attention and are investigating the roles they play in scientific practice. The result has been an incredible proliferation of type-models in the literature, citing only a few: phenomenological models, computational models, development models, explanatory models, test models, idealized models, theoretical models, scale models, heuristic models, teaching models mathematical models, models substitute iconic models, formal models, analogue models and instrumental models are some of the categorizations that are used to classify models.

A scientific model generally deals with phenomenon. With the term phenomenon we can identify its representation in terms of observable quantities. For example, empiricists like van Fraassen (1980) only allow for observables to qualify as such. This is the common approach used in physics and particularly in quantum mechanics. It can be fixed in mind that the observables must not be confused with the phenomenon itself. It is only for practical purpose that we need of observable for a better description of the phenomenon. For practical purposes, we need of a certain number of observables that are sufficient to characterize completely a given phenomenon or a given system. The time evolution of such variables must 'capture' the main features of the evolution of the system under consideration.

The first problem is to explain in virtue of what a model describes a phenomenon or a real situation. If we understand models as descriptions, the above question would be reduced to the problem of how language relates to real world. However, in all other cases, we are faced with the question of what it is for an object to represent a phenomenon.

Such problems are present in the recent literature (Bailer-Jones 2003, Contessa 2007, Elgin 2010, Frigg 2006, 2010c, Knuuttila 2009, Morrison 2009, Giere 2004, Suárez 2003, 2004, 2009, Suárez and Solé 2006, Thomson-Jones 2010, Toon 2010, 2011, 2012, van Fraassen 2004).

The second problem is concerned with the choice of representation, i.e. the fact that the same phenomenon can be represented in different ways.

A typical example is provided by quantum mechanics where the same phenomenon can be represented with different but equivalent set of observables. To address such problems see for example (van Fraassen 1980; Suppes 2002), (Da Costa and French 2003). In theory of systems representations is concerned with the ways in which a system can be described in mathematical terms. Also, it may happens that a given problem is best understood from a theorical point of view in a given representation, but for numerical purposes it is more useful to use another representation. For furthers information on this issue see J.C. Willems.

A typical issue that is often present in modelling a certain phenomenon is the one of considering a down-sized (Black 1962) version of the system itself. An important example is given by the study of systems with many degree of freedom like gas or liquid or also social dynamics with many agents. In this case the goal is to reduce the degree of freedom of the system and then study the equivalent down sized system. This is typical for example in mechanical statistics. The most important mathematical tools in this frame are certainly provided by the renormalization group approach (see Wilson). With this powerful technique we can study a given system by reducing the degrees of freedom and as a consequence we can "see" the original system at different distance scales. By passing from a scale to a smaller one the system is equivalent physically, but with less parameters to be described, i.e. less degrees of freedom.

4. Idealization Procedure

Often, to describe a certain phenomenon or system one is faced to the problem of 'idealize' the same. By idealization we mean a simplification of a 'complex' object to making it more tractable. Frictionless motion, point masses, infinite velocities, isolated systems, omniscient agents, and markets in perfect equilibrium are but some well-known examples. Generally there exist two general kinds of idealizations: so-called Aristotelian and Galilean idealizations. By Aristotelian idealization we mean to 'eliminate' all properties from a certain object that we believe are not relevant to the problem at hand. An example is a classical mechanics model of the planetary system, describing the planets as objects only having shape and mass, disregarding all other properties. To this purpose see Cartwright 1989, Ch. 5, Musgrave 1981 and Mäki 1994.

Galilean idealizations are ones that involve certain ad hoc assumptions. Physicists work with point masses moving on frictionless planes; economists assume that agents are omniscient; biologists study isolated populations, and so on. To address the questions related to the meaning of such approach see Laymon (1991) which understands idealizations as ideal limits.

In any case, the Galilean and Aristotelian approach are not mutually exclusive but they often come together. Consider again the mechanical model of the planetary system: the model only takes into account a narrow set of properties and distorts these, for instance by describing planets as ideal spheres with a rotation-symmetric mass distribution. In fact in both approaches one attempts to isolate a small number of essential properties of a system and transformate them into an extreme case.

5. Approximation and Similarity

Another fundamental concept in modelling systems is the approximation. First of all, it should be stressed that approximations arise at mathematical level, i.e. after a Galileon or Aristotelian approach is chosen. In practice, the approximations come into action to simplify calculations in a given scheme of modelling. Even though a theory may give a rigorous mathematical description of a given phenomena, the mathematical difficulties might be so great to solve that it is just not feasible to solve a problem exactly. One approximation can be done by ruling out a part of the calculations or to use an average rate rather than an exact mathematical description. The other commonly used method of approximation is variations, perturbations, use of simplified functions and fitting parameters to reproduce the result of a given experiment.

Often in modelling nature one can describe different objects that are related by similarity relations in the same way. As an example, consider the earth and the moon. They are different objects with different chemical composition. But they can be considered spherical. So, if we are interested in the cinematics properties of such bodies, what interests is the mass and shape and not the chemical composition. For a classification of different kind of similarity see Hesse (1963). In this sense similarity can be used with the meaning of 'universality'. This implies that in modelling a certain system of phenomenon we must choose the essential minimal set of features that are common to a large set of physical objects. As an example, the motion of all test particles of a certain mass in a gravitational field is the same independently on their chemical composition or internal structure.

Another interesting class of models is provided by the so called 'Phenomenological models'. A common used definition of these models is that they only represent observable and measurable quantities. Such models are only the outcome of empirical facts. In practice such models express mathematically the results of observed phenomena without care to their interpretation or significance. In first instance, such models can be viewed as a rule that must been satisfied by the empirical data. A further step for phenomenological models (and also generally) is the attempt to insert them in a given theory, from which the empirical laws have a sound interpretation.

6. Models of Data

An important class of models is given by the 'models of data' (Suppes 1962).

In this case, the empirical data must be the fundamental ingredient to test the validity of a given theory. Stated in another way, the data must be explained and reproduced in a given model. Obviously, one must firstly eliminate the possible errors that are always present. The construction of a data model is generally rather cumbersome. It requires complicated statistical techniques and raises serious methodological as well as philosophical questions as for example, the way to remove points on the data set or decide the best curve fitting the data (see for instance Galison 1997 and Staley 2004). A typical example is provided by the cosmology in fitting the distances in terms of the redshift of distant supernovae. A general problems arising in this context is the one of degeneracy.

In fact, in the interpretation of a given data set, more than one model could fit the data and then we are face to the problem to choose the best model that can reproduce the data. Generally the preferred model is the one that can reproduce more independents data set. Moreover, if exists more models explaining a given data set at the same level, the one that predicts new phenomena experimentally verifiable or explain better another set of independent data must be preferred. Further discussions of data models can be found in Chinn and Brewer (1994), Harris (2003), Laymon (1982) and Mayo (1996).

7. Mathematical Models

Another group of models are the ones often termed 'mathematical models'. As important examples, see the Black-Scholes model of the stock market or the Mundell-Fleming model of an open economy. Obviously all models must be confirmed by experiment, the natural arena for any reasonable theory. A theory can be perfectly expressed in terms of a coherent and nice mathematical construction, but nevertheless the model can have nothing to do with reality. Cartwright (1983), for instance, presents several case studies illustrating that good models are often false. In this context, an important property of a model is the possibility to contradict itself by means of practical experiment: a model that cannot be falsifiable does not represent a good model but rather a philosophical construction. It should be stressed that the Maxwell theory of electromagnetism is the first example in the history of science of a pure mathematical construction, built without any reference to real data set. Different authors (e.g. Brown 1991, Gendler 2000, Norton 1991, Reiss 2003, Sorensen 1992) have studied this line of argument with different conclusions within. In any case, validation data is an essential step to assume a model as a valid one. But validation is not always available. One possible reason is that many models predict into the future, and so we must wait until the validation data is at our disposal. In this case, the golden rule is to give a model that is good given all the information available. In absence of any validation data, one way to select the best model is using Akaike Information Criterion (AIC). AIC will choose a model that fits the data well without involving too many parameters.

Also in the context of pure mathematical models, often the equation involved is quite complicated and generally explicit solutions are lacking. In this case the solution relies on making a computer simulation. Consider for example equation for the dynamics of the universe or of the so called many bodies problem (Hegselmann et al. 1996, Skyrms 1996). The aim of a simulation is to solve the equations of motion of a system where many objects are involved. Hence a simulation approximates a real process by another process (Hartmann 1996, Humphreys 2004).

For a discussion of this technique of research see Humphreys 2004, 2009, Rohrlich 1991, Winsberg 2001 and 2003, and various contributions to Sismondo and Gissis 1999. In any case, independently on the meaning of this approach, the simulations are essential tools in modern science. It is essential to note the discrete nature of the calculations carried out on a digital computer. The problem can be mitigated by increasing the power of modern computers. However there exist powerful technique for integrating differential equations (Runge gutta methods) from which it is possible to control the error and then obtain a fixed degree of precision.

8. Choice of the Model

A further interesting question is the link between models representing the same phenomenon. Quite generally, different models interpreting the same data set or more data set are not necessarily in opposition. As an example, Newtonian theory well describes the motion of celestial body in the approximation of low speed of the body with respect to the speed of the light. If we consider bodies with velocities comparable with the speed of the light or which are very massive, Einstein's General relativity is more appropriate to deal with the astrophysical data. But if we restrict the data to the range of validity of Newtonian equation, both General relativity and Newtonian equation are practically equivalent with only very small differences not yet detectable within the actual technology.

This obviously does not mean that Newtonian mechanics is wrong, but only that it is appropriate in a certain context, within certain hypothesis. When such hypothesis is relaxed, a more general model comes in action. In fact, general relativity contains Newtonian mechanics as a limit of low velocities. As a consequence, to modelize a certain phenomenon, first of all one must specify some restrictions, within which the model 'works', i.e. it can explain experimental facts. Stated in other words, two models can as well explain the same empirical facts and diverge if the data set is augmented. The fact that a modelling can be obtained as a limited case of a more general model is rather common in modern science. As a consequence, to formulate a sound model, the limits of validity of this must be always taken in consideration. It is within these constraints that a given model can be successfully used to reproduce or explain a given phenomenon. Often in modern science an old model can be replaced by a new model that incorporate all the 'good' features of the old one but in a new more general context.

It should be stressed that the experimental data are always affected by errors, called 'systematics' that cannot be avoided or simply eliminated. This obviously takes more complicated the operation of finding the best model representing a certain data set. More, for a certain class of practical situations, the same data set available are not sufficient to build a model or are not easily available. Also in this case, the modern computers represent very useful tools. As an example, the computers can generate a casual distribution of initial data (Monte Carlo simulations) set an then simulate the relative dynamics of complex systems containing for example a large number of degree of freedom as, for example, the dynamics of a self gravitating cloud of particle. The outcome of such simulations must then be compared with the observations of such systems. The comparison with the collective behaviour of such systems available in the nature can give confirmation of the hypothesis given to perform simulations.

9. Formulas Versus Reality

The more important problem dealing with modelling is represented by the relation between formulas and the real world. As well known, Galileo said that the nature is written in mathematical language. In modern science this point of view is outdated. The mathematics is perhaps the most powerful instrument to investigate that nature, but there exist some fundamental process and phenomena that cannot be at our disposal in terms of formulas. A typical example is provided Biology or by Human Sciences, where the incredible amount of degree of freedom seems to be a serious obstacle to a complete mathematization of the nature. Perhaps may be that in a near future the next generation of quantum computer can alleviate such formidable task. In any case, the Goedel incompleteness theorem seems to pose problems to the mathematical ''knowability'', although this issue is a theme of an interesting dispute (see Hilary Putnam 1960, Avi Wigderson (2010)). However, it is important to stress that, within certain idealizations that are an important ingredient for any modelling procedure, the use of mathematics can works to describe the essential properties of a given phenomena or system. More generally, the most powerful technique at our disposal to describe complex systems with a large number of degree of freedom is provided by statistical methods and computer simulations.

10. Models and Theories

What is the relationship between models and theories? This is a rather complex question.

Generally a theory explains the 'why' of a model. Further, we can say that a theory represents an empirical knowledge only when validation from data is available. The validation is obtained by means of predictive mathematical models built within the framework of the theory. A fundamental property of models is that they are theory wannabes. Moreover, 'theory' operates at a more general level. Hence, theory concerns the essential properties and their relation between each other, while models deals with real objects and how such objects can be represented to best describe the empirical data.

Another important difference is that a theory can be used to describe many different type of models. As an example, general relativity can be used to build a model of the

Universe (cosmology) or to describe the gravitational field of a single star. As a consequence, a theory does represent nothing existing in the real world, it represents a complex language. The complex language provided by theory, when applied to models permit us to describe, also in an idealized form, the reality. To this purpose see Cartwright. A model thus gives an idealized representation of the world, but this representation is neither true of false. For example a chart of our solar system is not obviously a true representation of the solar system. Nevertheless, a chart of the solar system can give us information about the distance of the planet with respect the sun, i.e. true informations on the real world. Thus, once again models must give practical and verifiable informations among real objects. With a sentence, models represent objects and processes, while theories give us the language to 'speak' of such objects and processes.

In modern science there exist some remarkable examples of well posed theory explaining the real world built only from a theorical point of view, i.e. as a pure mathematical formulation. Examples are provided by Maxwell theory of electromagnetism, mechanics statistics, general relativity. In the majority of cases, the empirical data are a fundamental ingredient that guides us to build models and then, as a final step, help us to represent the model within a sound scientific theory. This is the final scope of any modelling procedure, since the scientist is always devoted to discover the rational relation between isolated experimental facts and not only to reproduce them in a simple model.

11. Selected Bibliography

Abrantes, P., Pimentel D. Tenedório J.A. (2010) *Metropolitan Dynamics Typology of the Portuguese Urban System*. The Open Urban Studies Journal, 2010, 3, 68-77.

Aggarwal, R.K, Lin T., Kim C.H. (2004) *A novel approach to diagnosis of defective equipments in GIS using self organizing map*. IEEE Power Engineering Society-General-Meeting IEEE Vol.1, 2004.

Aksoy, E. (2006) *Clustering With GIS: An Attempt to Classify Turkish District Data*. XXIII International FIG Congress- INTERGEO, Shaping The Change, 8-13 October, Munich, Germany.

Al Douri, A. (2010) *The impact of 3D modeling function usage on the design content of urban design plans in US cities. Environment and Planning B*. Planning and Design 2010, volume 37, pp 75-98.

Bação, F., V. Lobo, and M. Painho (2004) *Geo-Self-Organizing Map (Geo-SOM) for Building and Exploring Homogeneous Regions. Geographic Information Science*. Third International Conference GIScience 2004. M. Egenhofer, C. Freksa and H.J. Miller. Berlin, Springer.3234.

Bação, F., Lobo V., Painho M. (2005*)* The Self-Organizing Map, the Geo-SOM, and relevant variants for geosciences. Computers and Geosciences, Vol. 31, Elsevier, pp 155-163. http://www.isegi.unl.pt/ensino/docentes/fbacao/papers/The%20self-organizing%20ma%20the%20Geo-SOM.pdf#_blank.

Bação, F., Lobo V., & Painho M. (2005) *Self-organizing maps as substitutes for k-means clustering. Computational Science.* Iccs 2005, Pt 3, 3516, 476-483.

Bação, F., Lobo V., & Painho M. (2008) *Applications of different self-organising map variants to geographical information science problems.* In P. Agarwal & A. Skupin (Eds.), Self Organising Maps: Applications in Geographic Information Science (pp.21-44) Chichester, England:John Wiley & Sons, Ltd.

Bailer-Jones, (2003) *When Scientific Models Represent.* International Studies in the Philosophy of Science *17: 59–74.*

Barceló, J. A, Pallarés, M. (1996) *A Critique of G.I.S. in Archeology.From Visual Seduction to Spatial Analysis.* Archeologia e calcolatori No. 6.

Barceló, J. A. (2008) *Computational intelligence in archaeology*. The automatic production of knowledge. Paper presented at the ARCHAIA. Training Seminars on Research Planning, Characterisation, Conservation and Management in Archaeological Sites. Bologna (Italy), 15-17 May, 2008.

Basara, H.G., Yuan M. (2008) *Community health assessment using self-organizing maps and geographic information systems.* International journal of health geographics.;7: 67.

Bel, D.A., Nijkamp P., Scholten H. (2011) *Multidimensional urban sprawl in Europe: A self-organizing map approach.* Computers, Environment and Urban Systems Volume 35, Issue 4, July 2011, Pages 263–275. http://www.sciencedirect.com/science/journal/01989715.

Black, M. (1962) *Models and Metaphors. Studies in Language and Philosophy*. Ithaca, New York: Cornell University Press.

Brewer, W. F. and C. A. Chinn (1994) *Scientists' Responses to Anomalous Data: Evidence from Psychology, History, and Philosophy of Science.* In: Proceedings of the 1994 Biennial Meeting of the Philosophy of Science Association, Volume 1: Symposia and Invited Papers, 304–313.

Brown, J. (1991*) The Laboratory of the Mind: Thought Experiments in the Natural Sciences.* London: Routledge.

Carlson, E. (2000*) Kohonen Maps, Real Estate Market and the Effect of Individual Components.* In: H. Hyötyniemi (ed), Proceedings of the Finnish Artificial Intelligence Symposium SteP 2000, (Espoo, 28-31 August 2000) Finnish Artificial Intelligence Society, Vol. 3, pp. 227-234.

Cartwright, N. (1983*) How the Laws of Physics Lie*. Oxford: Oxford University Press.

Cartwright, N. (1989) *Nature's Capacities and their Measurement.* Oxford: Oxford University Press.

Chao, D.K., Chowdhury S., Lan H., Morgan A.J., Froese C.R. (2009*) Implementation of a Web-GIS Application with ArcGIS Server 9.2 for the Peace River Landslide Project*. Alberta. ERCB/AGS Open File Report 2009-08.

Clarke, K. C., Hoppen S., Gaydos L. (1997) *A Self-modifying Cellular Automaton Model of Historical Urbanisation in the San Francisco Bay Area.* In «Environment and Planning B: Planning and Design», 24, pp. 247-261.

Coleman, A.M. (2009) *Adaptive Classification Using Self Organisation Maps.* Sensor and system 18 October 2009. 5870.

Contessa, G., (2007) *Scientific Representation, Interpretation and Surrogative Reasoning.* Philosophy of Science 74(1): 48–68.

Da Costa, N., French, S. (2003) *Science and Partial Truth: A Unitary Approach to Models and Scientific Reasoning.* Oxford: Oxford University Press.

Dai E., Wu S., Shi W., Cheung C.K., Shaker A. (2005) *Modeling change-pattern- value dynamics on land use: An integrated GIS and artificial neural networks approach*. Environmental Management 36(4):576-591.

Dardala, M., Reveiu, A. (2011) *Geographical Information System Tecnologies for Spatial Visualization of Statistical Data.* MPRA Paper No 31090.

Elgin, C. (2010) *Telling Instances. In: Roman Frigg and Matthew Hunter (eds.):* Beyond Mimesis and Nominalism: Representation in Art and Science *Berlin and Ney York*. Springer, 1–17.

Ferentinou, M., Karymbalis E., Charou E., Sakellariou M. (2011) Using Self Organising Maps in Applied Geomorphology, Self Organizing Maps. Applications and Novel Algorithm Design. Josphat Igadwa Mwasiagi (Ed.). http://hua.academia.edu/EfthimiosKarymbalis/Papers/442218/Ferentinou_Maria_Karymbalis_Efthimios_Charou_Eleni_and_Sakellariou_Michael_2011_._Using_Self_Organising_Maps_in_Applied_Geomorphology_Self_Organizing_Maps_-_Applications_and_Novel_Algorithm_Design_Josphat_Igadwa_Mwasiagi_Ed._ISBN_978-953-307-546-4_InTech_Available_from_http_www.intechopen.com_articles_show_title_using-self-organising-maps-in-applied-geomorphology.

Fraser S. J., Dickson B. L. (2007) *A New Method for Data Integration and Integrate Data Interpretation: Self-Organising Maps.* In "Proceedings of Exploration 07: Fifth Decennial International Conference on Mineral Exploration" edited by B. Milkereit, 2007, p. 907-910.

Frigg, R. (2006) *Scientific Representation and the Semantic View of Theories.* Theoria 55: 37–53.

Frigg, R.,Hartmann, S., (2012).*Models in Science*. The Stanford Encyclopedia of Philosophy (Fall 2012 Edition), Edward N. Zalta (ed.).

Galison, P. (1997) Image and Logic. A Material Culture of Microphysics. Chicago: Chicago University of Chicago Press.

Gendler, T. (2000) *Thought Experiment: On the Powers and Limits of Imaginary Cases*. New York and London: Garland.

Giere, R. (2004) *How Models Are Used to Represent Reality*. Philosophy of Science 71, Supplement, S742–752.

Haidong Y., Yueming Hu, Fei qi Deng, Xian Tian, Baorong Li. (2004) *Fuzzy SOFM-GIS space cluster model and its application analysis.* In 2004-8th International Conference on Control, Automation, Robotics and Vision-Icarcv 1.

Harris, T. (2003) *Data Models and the Acquisition and Manipulation of Data*. Philosophy of Science 70*:* 1508–1517.

Hartmann, S. (1996) *The World as a Process. Simulations in the Natural and Social Sciences.* In Hegselmann et al. 1996, 77–100.

Hatzichristos T.(2004) *Delineation of demographic regions with GIS and computational intelligence.* Environment and Planning B-Planning & Design, 31(1):39–49, January 2004.

Hegselmann R., Müller, U. and Troitzsch, K. (eds.) (1996) *Modelling and Simulation in the Social Sciences from the Philosophy of Science Point of View.* Theory and Decision Library. Dordrecht: Kluwer.

Henriques R., Bação F., Lobo V. (2011) *Exploratory geospatial data analysis using the GeoSOM suite.* Computers, Environment and Urban Systems.

Hesse, M. (1963*)* Models and Analogies in Science. London: Sheed and Ward.

Howard T.E., Mendenhall M.J., Peterson G.L. (2009) *Abstracting GIS layers from hyperspectral imagery*. In proceeding of: Hyperspectral Image and Signal Processing: Evolution in Remote Sensing, 2009. WHISPERS '09. First Workshop on. http://www.researchgate.net/publication/224604951_Abstracting_GIS_layers_from_hyperspectral_imagery?ev=auth_pub.

Humphreys, P. (2004) *Extending Ourselves: Computational Science, Empiricism, and Scientific Method*. Oxford: Oxford University Press.

Humphreys, P. (2009) The Philosophical Novelty of Computer Simulation Methods. Synthese 169: 615–626.

Kamal M., Passmore P. J., Shepherd I.D.H. (2010) *Integration of geographic information system and RADARSAT synthetic aperture radar data using a self-organizing map network as compensation for realtime ground data in automatic image classification.* Journal of Applied Remote Sensing, Vol. 4, 043534.

Ji, C.Y. (2000) *Land-Use Classification of Remotely Sensed Data Using Kohonen Self-organizing Feature Map Neural Networks. Photogrammetric Engineering & Remote Sensing* Vol. 66, No. 12, December 2000, pp. 1451-1460.

Jiang, B.,Harrie, L. (2003) *Cartographic Selection Using Self-Organizing Maps*. In Fifth Workshop on Progress in Automated Map Generalization.

Jiang, B.,Harrie, L. (2004). *Selection of Street from a Network Using Self Organisation Maps.* Transaction in Gis, 8(3): 335-350.

Knuuttila, T., (2009*) Representation, Idealisation and Fiction in Economics: From the Assumptions Issue to the Epistemology of Modelling.* In: Mauricio Suárez (ed.):*Fictions* in Science. Philosophical Essays on Modelling and Idealisation, London: Routledge, 205–233.

Kohonen, T. (1982) *Self-organized formation of topologically correct feature maps*. Biological Cybernetics, 43(1), 59-69.

Kohonen, T., Hynninen J., Kangas J., Laaksonen J. (1995) SOM PAK: *The selforganizing map program package (version 3.1)* Rakentajanaukio 2 C SF-02150 Espoo, Finland: Helsinki University of Technology Laboratory of Computer and Information Science.

Kohonen, T. (2001) *Self-Organizing Maps.* Springer-Verlag, Berlin, Germany, pp. 501.

Kvamme, K.L. (1990) *The fundamental principles and practice of predictive archaeological modeling.* In A. VOORRIPS (ed.), Mathematics and Information Science in Archaeology: A Flexible Framework, Studies in Modern Archaeology, 3, Bonn, Holos-Verlag, 257-295.

Laymon, R. (1982*) Scientific Realism and the Hierarchical Counterfactual Path from Data to Theory.* Proceedings of the Biennial Meeting of the Philosophy of Science Association, Volume 1, 107–121.

Laymon, R. (1991) *Thought Experiments by Stevin, Mach and Gouy: Thought Experiments as Ideal Limits and Semantic Domains.* In Horowitz and Massey 1991, 167–91.

Li, B. (1998) *Exploring spatial patterns with self-organising maps.* Paper presented at the Proceedings of the Geographic Information Systems and Land Information Systems Consortium.

Li, Y. (2007) *Social Area Analysis Using SOM and GIS: A Preliminary Research*. RCAPS Working Paper No.07-3.

Lin, T., Aggarwal R.K., Kim C.H. (2004) *Identification of the defective equipments in GIS using the self organising map.* IEE Proceedings- Generation Transmission and Distribution, 151(5), September 2004.

Lobo, V., Cabral P., Bação F. (2007*) Self organizing maps for urban modelling*. In Proc. 9th Int. Conf. on geocomputation.

López E.O. (2009) *High Resolution Images: Segmenting, Extracting Information and GIS Integration.* World Academy of science ,Engineering and Technology 54.

Mayo, D. (1996) *Error and the Growth of Experimental Knowledge.* Chicago: University of Chicago Press.

Morrison, M., (2009) *Fictions, Representations and Reality*. In: Mauricio Suárez (ed.): Fictions in Science. Philosophical Essays on Modelling and Idealisation. London: Routledge, 110–135.

Nykänen, V. (2011) *Prospectivity mapping in GIS: integrate geochemistry data with geophysics and geology.* Workshop in the 25th International Applied Geochemistry Symposium 2011 22-26 August 2011 Rovaniemi, Finland. Vuorimiesyhdistys, B92-5, 88 pages.

MacEachren, A. M., Wachowicz M., Edsall R., Haug D., Masters R. (1999)*Constructing knowledge from multivariate spatiotemporal data: Integrating geographical visualization with knowledge discovery in database methods.* International Journal of Geographical Information Science, 13(4), 311-344.

Mäki, U. (1994) *Isolation, Idealization and Truth in Economics.* In Bert Hamminga and Neil B. De Marchi (eds.), Idealization VI: Idealization in Economics. Poznan Studies in the Philosophy of the Sciences and the Humanities, Vol. 38: 147–168. Amsterdam: Rodopi.

Musgrave, A. (1981) *Unreal Assumptions in Economic Theory*. The F-Twist Untwisted. Kyklos 34: 377–387.

Norton, J. (1991) *Thought Experiments in Einstein's Work*. In Horowitz and Massey 1991, 129–148.

Penn, B.S. (2004) *Using self-organizing maps to visualize high-dimensional data*. Computers & Geosciences 31 (2005) 531–544.

Pöllä, M., Honkela T., Kohonen T. (2009) *Bibliography of Self-Organizing Map (SOM*) Papers: 2002-2005, addendum, TKK Reports in Information and Computing Science, 1-236.

Putnam, H. (1960) *Minds and Machines*. In Sidney Hook, ed. Dimensions of Mind: A Symposium. New York University Press. Reprinted in Anderson, A. R., ed., 1964. *Minds and Machines*. Prentice-Hall: 77.

Oja, M., Kaski S., Kohonen T. (2002) *Bibliography of Self-Organizing Map (SOM)* Papers 1998-2001, addendum, Neural Computing Survey 3, 1-156.

Oyana, T. J., Achenie L. E. K., Heo J. (2012) *The New and Computationally Efficient MIL-SOM Algorithm: Potential Benefits for Visualization and Analysis of a Large-Scale High-Dimensional Clinically Acquired Geographic Data.* Hindawi Publishing Corporation Computational and Mathematical Methods in Medicine Volume 2012, Article ID 683265, 14 pages.

Pijanowski, B.C. (2002) *Using neural networks and GIS to forecast land use changes:a Land Transformation Model*. Computers, Environment and Urban Systems, 553–575.

Reiss, J. (2003) *Causal Inference in the Abstract or Seven Myths about Thought Experiments*. In Causality: Metaphysics and Methods Research Project.Technical Report 03/02. LSE.

Rohrlich, F. (1991) *Computer Simulations in the Physical Sciences*. In Proceedings of the Philosophy of Science Association, Vol. 2, edited by Arthur Fine et al., 507–518. East Lansing: The Philosophy of Science Association.

Samet, H. (1995) *Spatial data structures*. Reading: Addison-Wesley/ACM.

Saski, S., Kangast J., Kohonen T. (1998) *Bibliography of Self-Organizing Map (SOM)* Papers: 1981-1997, Neural Computing Survey 1, 102-350.

Sen, A. (2012) *Clustering approaches for hydrographic generalization*. GIS Ostrava 2012 - Surface models for geosciences.

Sismondo, S. and Gissis, S. (eds.) (1999) *Modeling and Simulation*. Special Issue of Science in Context 12.

Skyrms, B. (1996) *Evolution of the Social Contract.* Cambridge: Cambridge University Press.

Skupin, A., & Hagelman R. (2003) *Attribute space visualization of demographic change. GIS '03*. Proceedings of the 11th ACM international symposium on Advances in geographic information systems. Pages 56-62.

Skupin, A., & Hagelman R. (2005) *Visualizing demographic trajectories with selforganizing maps.* Geoinformatica, 9(2), 159-179.

Skupin, A., & Agarwal P. (2008) *Introduction: What is a self-organizing map?* In A. Skupin & P. Agarwal (Eds.), Self-organising maps: Applications in geographic information science (pp. 1-20) Chichester, England: John Wiley & Sons, Ltd.

Skupin, A., & Fabrikant S. I. (2008) *Spatialization*. In J. P. Wilson & A. S. Fotheringham (Eds.), The handbook of geographic information science (pp. 61-79) Malden, MA: Blackwell Publishing.

Sorensen, R. (1992) *Thought Experiments*. New York: Oxford University Press.

Spielman E.S.,Thill J.C. (2008) Social area analysis, datamining and GIS. Computers, Environment and Urban Systems Volume 32, Issue 2, March 2008, Pages 110–122. http://www.sciencedirect.com/science/journal/01989715.

Staley, K. W. (2004) *The Evidence for the Top Quark: Objectivity and Bias in Collaborative Experimentation.* Cambridge. Cambridge University Press.

Suárez, M., (2003) Scientific Representation: Against Similarity and Isomorphism. International Studies in the Philosophy of Science. 17: 225–244.

Suárez, M., (2004) *An Inferential Conception of Scientific Representation.* Philosophy of Science 71, Supplement, S767–779.

Suárez, M., Solé, A. (2006) *On the Analogy between Cognitive Representation and Truth.* Theoria 55, 27–36.

Suárez, M., (2009) *Scientific Fictions as Rules of Inference*. In: Mauricio Suárez (ed*.):* Fictions in Science. Philosophical Essays on Modelling and Idealisation. Routledge: London, 158–178.

Suppes, P. (1962) *Models of Data*. In Ernest Nagel, Patrick Suppes and Alfred Tarski (eds*.),* Logic, Methodology and Philosophy of Science: Proceedings of the 1960 International Congress. Stanford: Stanford University Press, 252–261. Reprinted in Patrick Suppes: Studies in the Methodology and Foundations of Science. Selected Papers from 1951 to 1969. Dordrecht: Reidel 1969, 24–35.

Suppes, P. (2002) *Representation and Invariance of Scientific Structures*. Stanford: CSLI Publications.

Thomson, J. M., (2010) *Missing Systems and the Face Value Practice*. Synthese 172 (2): 283–299.

Tobler WR (1970) *A Computer Movie Simulating Urban Growth in Detroit Region*. Economic Geography,46,pp 234-240.

Toon, A. (2010*) Models as Make-Believe*. In: Frigg, R and Hunter, M. (eds.*):* Beyond Mimesis and Convention: Representation in Art and Science. Boston Studies in the Philosophy of Science: Springer, 71–96.

Toon, A. (2010) *The Ontology of Theoretical Modelling: Models as Make-Believe.* Synthese 172: 301–315.

Toon, A. (2011) Playing with Molecules. Studies in History and Philosophy of Science 42: 580–589.

Toon, A. (2012) *Models as Make-Believe: Imagination, Fiction and Scientific Representation*. Palgrave Macmillan.

Van Fraassen, C. (1980) *The Scientific Image.* Oxford: Oxford University Press.

Van Fraassen, C. (2004) *Science as Representation: Flouting the Criteria*. Philosophy of Science 71, Supplement, S794–804.

Villmann, T., Wieland H., & Michael G. (2000) *Data mining and knowledge discovery in medical applications using self-organizing maps.* Paper presented at the Proceedings of the First International Symposium on Medical Data Analysis.

Wang, Y.C., Feng C.C. (2008) *Visualizing Research of Land Use Land Cover Change*. Geoinformatics 2008 and Joint Conference on GIS and Built Environment: The Built Environment and Its Dynamics, Lin Liu, Xia Li, Kai Liu, Xinchang Zhang, Xinhao Wang, Eds., Proc. of SPIE Vol. 7144, 71440J.

Wigderson, A. (2010) *The Gödel Phenomena in Mathematics: A modern view*.

Willems, J.C. (1979) *System Theoretic Models for the Analysis of Physical Systems*. Ricerche di Automatica, Vol. 10, No. 2, pp. 71–106.

Willems, J.C. (1980) *Topological classification and structural stability of linear systems.* J. Differ. Equa. 35, (3), p. 306-318.

Willems, J.C. (1980) *Almost A (mod B.)- invariant subspace*. Asterisque 75-75, 239-248.

Willems, J.C. (1981) *Almost invariant subspaces: An approach to high gain feedback design.* Part I: Almost controlled invariant subspaces. IEEE Trans. Automat. Control AC-26, 1, 235–252

Willems, J.C. (1983) *Input-output and state space representation of finite-dimensional linear time invariant system.* Lyn. Alg. Appl. 50, 581-608.

Willems, J.C. (1986) *)* From Time Series to Dynamical Systems. Part I: finite- dimensional linear time invariant system. Automatica 22, 561-580.

Willems, J.C. (1986) *From Time Series to Dynamical Systems.* Part II: Exact Modelling, Automatica, Vol. 22, No. 6, pp. 675–694.

Willems, J.C. (1986*) Deducing the input/output and the input/state/output structure from the external behaviour.* In Proc. 25th IEEE Conf. Dec. Contr. (Athens, Greece, Dec. 1986) IEEE Press, New York. 1936-1937.

Willems, J.C. (1987) From time series to linear system - Part III. Approximate modelling Automatica Vol. 23, 87-115**.** http://homes.esat.kuleuven.be/~jwillems/Articles/JournalArticles/1986.3.pdf.

Wilson, K.G. (1975*) The renormalization group: critical phenomena and the Kondo problem*. Rev. Mod. Phys. 47, 4, 773.

Winsberg, E. (2001) *Simulations, Models and Theories: Complex Physical Systems and their Representations.* Philosophy of Science 68 *(*Proceedings): 442–454.

Winsberg, E. (2003) *Simulated Experiments: Methodology for a Virtual World*. Philosophy of Science 70: 105–125.

Xu, Z. (2011) Application of System Dynamics model and GIS in sustainability assessment of urban residential development. In: Proceedings of the 29th International Conference of the System Dynamics Society, Washington DC (USA), July 24 – 28 2011. http://porto.polito.it/2460435/.

Yan, Li (2007) *Social Area Analysis Using SOM and GIS: A Preliminary Research*. Ritsumeikan Center for Asia Pacific Studies (RCAPS), Ritsumeikan Asia Pacific University. Working Paper No.07-3.

Yuan, M., Buttenfield B. P., Gahegan M., & Miller H. (2001) *Geospatial data mining and knowledge discovery*. Washington, DC: University Consortium for Geographic Information Science.

Zondac, B., Geurs K. (2011) *Coupling a Detailed Land-Use Model and a Land-Use and Transport Interaction Model*. In Land Use Modelling in Planning Practice, Geojurnal Library, 2011, Volume 101, Part 2, 79-95.

Irene VIAGGIU, Department of European, American and Intercultural Studies, Sapienza University of Rome, Italy.

email: ireneviaggiu@gmail.com

Viaggiu; Irene, "GIS and SOM Application in Different Scientific Area". In: Armando Montanari (ed.), *Urban Coastal Area Conflicts Analysis Methodology: Human Mobility, Climate Change and Local Sustainable Development*. Cap. 2. *SECOA FP7 Research Project*, Vol. 5. Rome: Sapienza Università Editrice, 2013. *http://digilab-epub.uniroma1.it*. DOI: DOI: 10.13133/978-88-98533-01-5. Web. 30 September 2013.

ABSTRACT: In this chapter, we briefly review the relevant literature in the field of the Gis technology used to obtain a large amount of data and informations to be analyzed and classified. In particular, we review the main features of the technique named 'Artificial Neural Networks' that mixes statistical methods together with a 'machine learning approach' that allow-us to generate dataset collections up to a workable level. Also the most quoted bibliography on the 'Self Organizing Map' is presented and commented, by stressing in particular the present day capability to reduce high-dimensional data to a less dimensional one in the form of clusters preserving the initial topology. Moreover, we analyze the main issues raised in the current literature regarding the concept of modelling in different scientific disciplines. In particular, we explore the relevant existing literature on the fundamental concept of idealization of the real world and its relation with the modeling procedure. Finally, the relation between models and theories is reviewed in light of the current ideas.

KEYWORDS: GIS, SOM, MODELLING.

Submitted: 6th June 2013

Accepted: 18th July 2013

Final Revision: 23th July 2013

Published: 30th September 2013

Annex I.

CHAPTER 3.

Interpreting Local Issues in a Global Context: International Comparative Analysis and Coastal Urban Areas

Armando MONTANARI

Department of European, American and Intercultural Studies,

Sapienza University of Rome, Italy

1. Introduction

The following chapter examines the concept of knowledge in terms of awareness of knowledge. Cross-national comparative research has become an increasingly important part of social sciences research over the past half-century. Other fields had independently carried out this task in the early years of the century, and possibly come up against fewer difficulties than the social sciences. The need for such research has become increasingly evident over the past twenty years, with the consequences of the cumbersome presence of human beings on Earth forcing all the sciences to revisit environmental issues using multidisciplinary methods.

This chapter is divided into two parts. The first part reviews the efforts made to carry out genuinely (i.e. concrete, not merely theoretical) multidisciplinary research capable of offering ways to interpret phenomena, and possibly also some solutions. The subject was urban development in coastal areas, viewed through the prism of a sequence of experiences as well as the socio-economic situation in Europe, which has gone through a profound change over the past fifty years in terms of type of development as well as the ability and willingness to cooperate. Cooperation is essential in order to resolve problems that, by their very nature, are both global and local. Even the variables connecting these two dimensions are interconnected. To try and reconstruct the key elements of the history of international scientific cooperation, we consulted scientific literature as well as unpublished documents produced in some areas of research, where the researchers had put so much effort into resolving incidental problems that, despite their best intentions, they had not managed to publish an account of their experiences. The author of this paper learned both scientifically and culturally from these research experiences, either by taking part in them or by studying them. International comparative research is like an archaeological dig reversed in time. If we do not report and document what we discover as we go along, the "discovery" does not exist in absolute terms. If we fail to provide documentation, and an interpretation of the stones, we will never know anything about the wall those stones were part of, or about the city and its buildings made up of those stones. On the basis of what we know, it is possible to reconstruct how research went from monothematic or monodisciplinary sectors to wider compendia capable of using the cognitive approaches of scientific fields that, although different, co-operate with each other. The findings of these broader – and inevitably more complex – approaches have often led to debate on the advantages and disadvantages of multidisciplinary and cross-national research. Researchers often find that the results of such complex activity are disappointing. So the question of why comparative analysis is required must always be asked in advance. Is it worth it to increase complexity in sectors or situations in which added value cannot be guaranteed? And finally, what is to be compared? This chapter highlights the key elements of interest to SECOA ("Solutions for environmental contrasts in coastal areas. Global change, human mobility and sustainable urban development"): coastal areas, urban settlements, the environmental problems in these settlements and their continued expansion in fragile coastal areas. The progress of cross-national research is subsequently interpreted in two different ways deriving from two thought processes that can be traced back to the Stockholm Conference (1972). Firstly, there was the need to bring development back within the limits imposed by environmental themes. Secondly, there was the opening in 1963 of the Vienna Centre by the ISSC (which had been set up by UNESCO ten years previously) to meet the need for cooperation among European scientists in order to improve the quality of results while reducing the differences between European countries. The paper examines the characteristics and findings of some research projects preceding and following the end-1980s from these two perspectives. The end-1980s marked the collapse of governments presiding over planned economies, and the concept of sustainable growth was introduced at the Rio de Janeiro conference in 1992.

In the second part of the chapter, the methodological references and complex situations described previously are compared with the set-up and management of the SECOA research project. SECOA's set-up was based on the need to identify and resolve conflicts arising in coastal areas, and manage the complex balance required for sustainable development. At the local level, sustainable development is the result of a compromise – not always an entirely acceptable one – among the numerous competing stakeholders in contemporary society's decision-making process. These are quite different types of stakeholders, who manage to impose their own interests, depending on their ability to comprehensively state their case. But the local aspect is only one component of the elements transferred to the territory as a result of what is decided and carried out at the local level. SECOA identified two key, and growing, components of this action and reaction pairing: human mobility and global climate change. The book, and hence also this chapter, focuses mainly on the local dimension and reference areas. The second part of the chapter aims to explain why an interpretive model based on the use of two instruments, geographical information and the neural network, was made. This combination is not totally original, but after examining the experiences of other projects, SECOA decided to test an independent path. In its methodological approach, different types of neural networks worked in different ways with the geographical information within a framework that favoured greater scientific complexity over the technological instrument. This method made it possible, by using taxonomy and then interpretive modelling, to connect the phase of collecting data and organising information to that of drawing up policies to reduce conflict and understand phenomena, so as to gain a more thorough grasp of problems and thereby contribute to resolving them.

2. Cross-national Comparative Research. Methodological Issues, Theories, Organisation and Management

2.1. From single cultures to a multidisciplinary approach

The SECOA project was based on collaboration among different cultures – the natural sciences, social sciences and humanities. This reflected not so much a decision of the researchers as the need to work within the reference framework, i.e. the concept of sustainable development and its practical application in urban coastal areas. It was absolutely indispensable for the SECOA project be organised in this way, even though it made it more complicated to process and compare data from such heterogeneous approaches and experiences.

SECOA was devised between 2008 and 2009, around fifty years after the 1959 lecture in which Charles Percy Snow lamented the breakdown in communication between scientists and literary intellectuals, and their inability to work together. Scientific and technological research is important for the social development of a community, while public life tends to be dominated by the humanities. Snow, who was both a physicist and a man of letters, describes the discomfort of the intellectual who must acknowledge the sharp dichotomy between scientific research and the humanities, and the political, environmental and cultural problems that this chasm creates for society as a whole. He believes that both cultures must be represented in the social and political spheres in order to provide greater depth of perspective. Snow's work (1959) remained the focal point for discussions on this theme for many decades.

In 2009, some leading scientific journals reviewed the thinking of the previous half-century. Writing in the *New Scientist*, Collini (2009) noted that Snow had identified the issue of hyper-specialisation in disciplines, which on the one hand makes it harder to spread awareness of scientific results and, on the other, does not allow non-scientists to reflect on and write about scientific subjects. The editorial of an issue of the magazine *Nature* (2009) also takes up the subject of dissemination, and extends it to the North-South relationship between developed and developing countries, and between the rich and the poor.

John Brockman (1995) cites Snow's prediction in the second edition of his book (1963) of "a third culture" which, as Brockman says, "would emerge and close the communications gap between the literary intellectuals and the scientists". Brockman continues: "In Snow's third culture, the literary intellectuals would be on speaking terms with the scientists. Although I borrow Snow's phrase, it does not describe the third culture he predicted." Brockman brought together some of the leading researchers of the time to see if it was possible to identify the third culture. In the introduction to his book, Brockman complains that Snow did not really describe the third culture he mentioned. "Literary intellectuals are not communicating with scientists," Brockman continues. "Scientists are communicating directly with the general public. Traditional intellectual media played a vertical game: journalists wrote up and professors wrote down. Nowadays, third culture thinkers tend to avoid the middleman and endeavour to express their deepest thoughts in a manner accessible to the intelligent reading public." For Brockman (1995), the third culture is a process, not a discipline, embodied by scientists and other thinkers who, through their work and ability to express their ideas, have replaced traditional intellectuals in "rendering visible the deeper meaning of our lives, redefining who and what we are". So the third culture is an instrument that looks towards the future of science and human beings.

The disclosure of the findings of scientific research has become a crucial point in recent years, and is being put into practice in the most advanced countries. In the summer of 2012, the European Commission outlined measures to improve access to scientific information produced as a result of publicly funded research. The open access philosophy will make it easier for researchers and businesses to build on these findings, thereby giving a better return on the €87 billion annual investment in R&D and "boosting knowledge and competitiveness in Europe", to quote Máire Geoghegan-Quinn, European Commissioner for Research and Innovation. On 22 February 2013, U.S. President Barack Obama's Executive Office published a memorandum stating that: "The Administration is committed to ensuring that, to the greatest extent and with the fewest constraints possible and consistent with law and the objectives set out below, the direct results of federally funded scientific research are made available to and useful for the public, industry, and the scientific community. Such results include peer-reviewed publications and digital data." In the EU, the open-access publication policy rolled out in 2011 will become compulsory for all member states from 2014, through Horizon 2020, the research funding programme for 2014-2020. Open access will be made operational in two ways: "Gold open access", with articles immediately being made accessible online by the publisher, and "Green open access", with articles archived in a repository and subsequently made available to all after an embargo period. In the first case, the costs will be transferred from the reader to the public body financing the research, and in the second case the costs are shifted from the reader to the producer of the article.

Of course open access can only reduce communication problems between cultures, particularly between businesses and society at large, not resolve them entirely. Kagan (2009) uses Snow's premise of two cultures – natural science on the one hand, the arts and humanities on the other – as his starting point. He notes that a third culture – that of the social sciences, comprising sociology, anthropology, political science, economics, psychology and geography – has emerged over the past fifty years. Kagan describes the assumptions, vocabulary and contributions of natural science, the social sciences and the humanities. He argues that the social sciences and humanities have contributed greatly to our understanding of human nature, and questions the popular belief that biological processes are the main determinant of variation in human behaviour.

Kagan divides up the sciences using quite pragmatic parameters such as the amount of funding required, finding sources, research management methods and the way data are collected and handled. In the case of the humanities, he mentions "philosophers, scholars of literature and historians… most work alone, are not highly dependent on grant support from government agencies and rely primarily on semantic text as a source of evidence". He cites a triangle of powers, knowledge and interests, at whose apex are natural science, the social sciences and the humanities. In some phases of human evolution, the triangle has been modified to the extent of almost becoming a straight line, with the system of checks and balances petering out. Kagan's message can be summed up into two concepts. The physical and cultural environment is much more important than genetic structure. The correct study of human behaviour is based on analysing culture and symbols. The key sources for the social sciences are utterances, observed and analysed behaviour, and physiological measurements. The social sciences could successfully work together. Any explanation of human behaviour depends – at the very least – on all the different ways to ask and answer the questions that science requires. The "big science" that has developed over the past fifty years has taken away the excitement of the individual discovery. The large-scale funding required by "big science" has benefited the social sciences, where flashes of intuition and theoretical innovation are still possible with reasonable funding.

So we need to reflect on the essence of the social sciences. All the fields that study society and human beings – in the sense of their participation in society – are called social sciences. Anthropology, archaeology, criminology, economics, education, geography, history, international relations, law, linguistics, political sciences, social psychology, sociology and social work are generally considered to be social sciences. The point is not to catalogue these fields at any price, but it is undoubtedly true that some fields have changed over time, and definitions that were valid in the post-World War II period need to be reviewed and reinterpreted today. A case in point is geography, which comprises physical geography (a natural science) and economic geography (a social science). Physical and economic geographers follow the same university syllabus and belong to the same professional organisations, and are all members of the International Geographical Union (IGU), which in its turn is a member of the International Social Science Council (ISSC) while also being a part of the International Council for Science (ICSU), which is made up of associations representing the fields of anthropology and psychology, but not the other social sciences. So geography alone represents at least two cultures, and can therefore naturally contribute to overcoming intercultural divisions. Geography studies the spatial distribution of phenomena through the use of the most up-to-date systems of geographical information, and uses place to compare the way different events and cultures interact.

Relying on funding from the public sector and foundations has often led to slavish dependency on them, at the expense of the kind of learning that comes from writers and historians. The reason this happens is that natural and social scientists are expected to provide the funds to keep the books of academic institutions in the red, as a result of which there are often unpleasant disagreements within the institution.

Researchers working in three different cultures – natural science, the social sciences and the humanities – participated in the SECOA project. The choice of the three cultures was necessary at a time when global researchers are trying to understand climate change (natural science) and human mobility (social sciences). At the local level, conflicts in urban coastal areas have required the skills of natural scientists (sea level rise and water management), the social sciences (socioeconomic variables and conflict types) and the humanities (history and cultural heritage sites). The involvement of people from different fields did not create insurmountable problems for cooperation among individual researchers, in part because the research network was selected on the basis of tried and tested forms of collaboration. Problems did, however, crop up when it became necessary to compare the data collected by researchers in the different fields. In this book, we mention not only the difficulty due to the multidisciplinary nature of the researchers, but also their disparate geographical origins, which meant different legal systems and totally dissimilar cultural habits and administrative policies. The convergence of disciplines in research programmes had become a necessity from the 1990s, when the concept of sustainable development was introduced. The three cultures have to meet to analyse phenomena through the lens of sustainable development, or programme such development. So although the differences persist, the gap will inevitably narrow.

2.2 Why compare? The problems of cross-national and multicultural analysis in the natural and social sciences

The main stakeholders are naturally interested in undertaking comparative analyses to gain a better understanding of what is happening in their community. And, if they are responsible enough, they try to give a broader perspective to the decisions that they make and will have to be accountable for. According to Bradshaw and Hatland (2006), the reason social scientists are interested in carrying out comparative analyses is that comparative research is the bedrock of methodology in their field. Comparing is like setting up an experiment to understand the origins and nature of a social and economic phenomenon more thoroughly than would be possible by studying a single country, or with a single case study. Natural science is applied to sectors of research which, by their very nature, are compatible. The geological structure of the Earth, climate change and even the human body have constant characteristics. So scientific analysis sums up the findings of research carried out in different places, which can however be combined to supply more information and carry out further checks. In a natural science research project, when one team has to quit for whatever reason, the project as a whole is only partially impacted. The amount of evidence gathered will be smaller, but the exercise is still worth completing, and at any rate it can always carry on with an alternative team. The social sciences are based on comparability. Human beings all over the world have exactly the same physical make-up, but are totally different when it comes to their behaviour, culture and history. Comparative research in the social sciences exists because of these differences. If a team drops out of a comparative research project, in most cases the entire project would collapse. So we can affirm that comparative analysis has a particular scientific value, a specific added value, especially for the social sciences. Lisle (1985) writes that in the natural sciences, the languages and cultures of the researchers participating in an international research project are practically irrelevant, and hardly ever an obstacle to the progress of knowledge. In the social sciences, on the other hand, langue and culture are not only instruments of communication, they are also the subject of scientific study.

Comparative analysis began after the Second World War, around the same time as it began to develop and be used in the social sciences. The physical and moral destruction caused by the war had to be studied attentively to ensure that the mistakes of the past would not be repeated. American academics had got a head start, and were carrying out research using the funding available to them and the cognitive tools they had developed (which, however, were not always suited to the other countries they were studying). Meanwhile, the world had begun dividing into countries with free market economies and countries with planned economies, and the early impact of this new, divided world order was becoming perceptible. To overcome potential discord, in 1951 UNESCO set up the International Social Science Council (ISSC) to examine whether greater international cooperation in the social sciences was possible, and subsequently open an institute where researchers from different fields, countries and cultures could meet and work together (Montanari, 2012b). After a number of preliminary meetings, in 1963 the European Centre for the Coordination of Research and Documentation in the Social Sciences (Vienna Centre) was set up, with the specific goal of organising research projects involving researchers from countries with opposing cultural systems – a constant risk for peace and stability, i.e. countries with planned economies and countries with free market economies, mainly Eastern European countries and the West.

The early years of ISSC activity were devoted to promoting infrastructure for use by social scientists, such as data banks, libraries and data processing software. According to Szalai, Petrella et al., (1977) processing systems were applied to already available data. Following the example of research by other organisations in preceding decades, such as that of the Royal Anthropological Society at the end of the 19th century, problems of data collection and comparability were disregarded, as it was not easy to influence the many national authorities in charge of data collection. The availability of comparable data has been a significant problem ever since comparative analysis began. It is a problem the SECOA project too has come up against, and one that will probably hinder comparative research in the future too. The efforts made by the US Administration and the EU in this respect are proof that it is a genuine problem, and that in any case, even if the data were totally comparable, the issue of their total availability would still have to be tackled.

The need to resolve the deep-rooted conflicts that had led to the outbreak of World War II spurred UNESCO to start funding comparative analysis as early as 1948 (Montanari 2012b). The variable studied most frequently in those years was the concept of the "nation". We can refer to the findings of a well known comparative analysis of ten countries by Ornauer, Wiberg, Sicinski, and Galtung (1976), who state, among other things, that "when it comes to the great issue of space, peace and war, our data indicate that the nation will probably continue to be the salient actor for a long time to come… (p.574)". In a later work, Galtung (1982) specifies at least three meanings of the word "nation". Meanings should be defined for each variable as a matter of course, so as not to unwittingly make the error of comparing "words" rather than meanings. Galtung (1982) makes a distinction between (i) "nation" in the sense of a country, a political entity in territorial space (a country is also often called a state); (ii) "nation" in the sense of ethnic group, a socio-cultural entity in non-territorial space; and finally (iii) "nation" in the sense of a nation-state, meaning a state populated almost exclusively by members of the same ethnic group. But the "nation" must also be considered part of a more complex and general system and, simultaneously, itself be considered as a complex entity made up of subsystems. When we delve into the problems of comparative analysis, we should always consider that each entity has at least more than one horizontal dimension and, similarly, more than a single vertical dimension. Galtung (1967) proposed the following classification of variables for the term "nation" (tab. 3.2.1).

Table 3.2.1. Classification of variables for countries/states. Source: Galtung (1967).

classification	**variables**	**elements**
Type 1. Nation by itself	Absolute	Size, population, continental belongingness, or to other groups
Type 2. Nation as part of a super-system	Relative	Any kind of variable on which nations may be ranked
Type 3. Nation as part of a super-system	Relational	Interaction of nations in pairs, dyads, bilaterally
Type 4. Nation as part of a super-system	Structural	Interaction of nations in n-tuples, n-ads, multilaterally
Type 5. Nation as having subsystems	Inside	Same four types

Galtung (1967) explains the variables in detail, focusing in particular on the relational variables, which he believes to be the most significant, as they analyse interaction between nations rather than the nations per se. We can generalise Galtung's theory and say that the principal element of comparative analysis is the ability to represent the elements of interaction between variables, and therefore go beyond evaluations such as "greater than" or "less than", which are relative variables. So how do we evaluate interaction processes? Galtung introduces an evaluation of interaction based on levels of "symbiosis" and "exploitation". Symbiotic interaction occurs when the advantages or disadvantages of a variable immediately impact all the variables in an equal manner. In exploitative interaction, the advantages of a variable damage all the others, so there is no convergence of interests.

2.3. The pros and cons of cross-national, multidisciplinary comparative research

According to Hantrais (2005), one of the main themes of the social sciences is the complexity of research across national and cultural borders. The advantages of comparative research have been appreciated right from the end of World War II, as mutual understanding among researchers encourages international scientific cooperation and the spread of results. Neither should we underestimate the importance of mutual understanding and collaboration as an embryonic form of international cooperation during the years when the world was divided into two political, economic and social blocs. Galtung (1982) acknowledges that there were political implications that would have contributed to discussions on common problems of development and the future and encouraging cooperation between universities and businesses, as part of what was known as the "spirit of Helsinki". The term refers to the Final Act of the Conference on Security and Co-operation in Europe, approved in Helsinki on 1 August 1975. One section of Chapter 4 (Co-operation and Exchange in the Field of Education) of the Final Act is devoted to "Science". Scientific research is described as: i) exact and natural science; (ii) medicine; and (iii) the humanities and social sciences. The goal of research is "to develop in the field of scientific research, on a bilateral or multilateral basis, the co-ordination of programmes carried out in the participating States and the organisation of joint programmes …. which may involve the combined efforts of scientists and in certain cases the use of costly or unique equipment". In the past decade, following the collapse of the Berlin Wall and the overcoming of the differences of the preceding decades, the number of international projects has increased substantially, undoubtedly as a result of the internationalisation of the economy and business and the increase in human mobility, but also because EU funding has focused on cross-national projects. These are now given preference over national projects because they encourage an international exchange of ideas, information and scientific culture, and therefore contribute to innovation. In the contemporary world, innovation is key to competing in international markets, and hence achieving economic growth and more jobs. The Lisbon Strategy (2000) ratified the need for greater scientific co-operation among EU countries. EU proposals to overcome the financial crisis that began in 2007 include "investing in knowledge" to make it possible for knowledge to circulate freely, and thereby contribute to creating a European research area with improved conditions for scientific innovation.

Since research findings depend mainly on the availability of comparable data, we cannot rule out the risk that the scientific quality of a cross-national research project may be inferior to that of a single-nation study. The less comparable the data, the smaller the possibility of complex extrapolation and the greater the risk of insignificant results. Hantrais (2005) divides researchers involved in comparative analysis into two categories: those who use a quantitative approach and those who use a qualitative approach. The former adopt a systematic and reliable scientific method, imitating the way natural scientists work, and are therefore able to outline possible scenarios and ways to generalise trends of convergence or divergence. In the latter case, the focus is on the differences between and within states, and the complexity of the elements involved. In this second case, the instrument used is the case study, which has many advantages, but is difficult to generalise. However, the evolution of research projects and the complexity of the phenomena researchers come up against have reduced the quantitative-qualitative divide; the two approaches are now considered mere tools to use, sometimes simultaneously – methodological pluralism based on the needs and objectives of the research project. Methodological pluralism refers to a mix of methods, not just the quantitative or qualitative approach, as Hantrais (2005) points out in her analysis of the methods used and the results obtained in a cluster of European Commission-funded FP4 and FP5 projects.

From the method we come to project management. The project co-ordinator and work package co-ordinators play a very important role, as they have to select – in advance – the variables and methods for collecting variables that will work with the overall objectives of the project. This co-ordination is essential at every stage – before the start of the project, to devise the research framework; during the data collection phase, to fine-tune methods and tools, and finally when interpreting findings (this phase must be consistent with what has been specified in previous phases). The findings of an international research project cannot be systematically compared if comparison is not set as the project's objective from the outset, let alone during the data collection phase. The characteristics of data vary substantially from one country to another, in Europe as well as outside Europe, as data collection methods and classification criteria used by public administrations are variable, and can also change over time. Official statistics are often aggregate data for an entire country, or for macro areas, and are not necessarily collected either continuously or at regular intervals. In some countries, data are collected locally in a specific way, depending on local planning requirements. Many researchers point out that a lot of time and energy therefore goes into "reducing" the classification of shared and shareable databases. The term "reduce" is used because some researchers working on comparative studies say that by its very nature, this kind of research involves a higher level of compromise than a single-country study. Debate of this nature was cited at the early stages of comparative research (Montanari, 2012b). Gusdorf (1977) held that interdisciplinarity and interculturality did not mean a search for the lowest common denominator or the highest common factor, thereby making them reductionist; rather, they should be seen exclusively as an enterprising process involving the investigation of unknown sectors. Therefore, no tools can be suggested *a priori*; we must trust entirely in management project managers capable of recognising, understanding and giving value to the cultural, economic and social contexts in which they have developed data that may be compared, even though it is apparently not comparable. So the problem is not to compare the data as such, but rather to try and understand whether a variable indicates an international or an intra-national difference. Alternatively, national data have to be grouped into clusters, and in some cases the clusters will be compared to make it easier to understand the more general aspects of a specific phenomenon. The greatest risk is of having an interpretive model before, or regardless of, the variables, thereby implying that the data must adapt to the model, not the other way around. Successful comparative research has almost always overturned this approach, as it is the research topic that determines the selection of data, and the evaluation method can only be determined once the data are available and comparable (Montanari, 2012a). Peschar (1982) reviews the positions of Frey (1970) and subsequently Elder (1976) to remind us that there are four possible strategies for selecting countries: (i) administrative convenience; (ii) maximise similarity; (iii) maximise diversity; (iv) factorial matrix based on relevant dimensions. While the first strategy is the most widely used, we cannot overlook a pragmatic choice that also takes into account the presence of researchers or institutions with an established reputation. Peschar also mentions obligatory frameworks imposed by the institutes that organise and fund research. For instance, for the Vienna Centre, country teams from both Western and Eastern Europe had to be involved while, in the case of EU-funded projects, every call for proposals details what categories of countries can, or should, be involved.

To lower the risk of failure, the selection of countries to be analysed is planned in advance, so as to proceed in line with the programmed objectives right from the outset. Social scientists in particular have seen the need to identify in advance the basic differences between countries, or at least developed ones, so as to provide a valid basis for synthesis when selecting the assortment of cases to study, and to better evaluate the context of the variables collected. This division was mainly made on the basis of welfare policies, but their validity has been demonstrated for other sectors too. Parameters such as pension systems, the tools that regulate the public and private sector relationship, and the attention paid to social stratification are important, not least the way that relations between public administrations and civil society are managed. In the last two decades, the right to a pollution-free environment and the introduction of sustainable development policies have reflected previously identified differences in the way social policies are applied. Esping-Andersen (1990) analysed 18 OECD countries with social welfare systems. The parameters considered were: (i) decommodification (the extent to which an individual's welfare is reliant upon the market, particularly in terms of pensions, unemployment benefit and sickness insurance); (ii) social stratification (the role of welfare states in maintaining or breaking down social stratification) and the private-public mix (the role of the state, the family, the voluntary sector and the market in welfare provision). Using these parameters, Esping-Andersen (1990) identified three regime types: Liberal (Australia, Canada, Ireland, New Zealand, UK, USA), Conservative (Finland, France, Germany, Japan, Italy, Switzerland) and Social-Democratic (Austria, Belgium, The Netherlands, Denmark, Norway, Sweden). Esping-Andersen's was long considered the reference typology, but he was also criticised for basing it solely on social transfers and totally ignoring the role of healthcare, education and social services. Bambra (2007) examines the key elements of Esping-Andersen's work, the debate it sparked and the main scientific works that followed in the next fifteen years using other variables. Bambra explains that his work is intended to be a cultural bridge between social science researchers and public health researchers and epidemiologists, "… so that our research can reflect and benefit from the more contemporaneous insights on offer from the 'welfare modelling business'."

The debate about the pros and cons of international comparative research has, however, died down over the past decade. By launching Framework Programme 4 (FP4) and subsequently the FP5, FP6 and FP7 projects, the European Commission has considerably increased opportunities for applying comparative research to large-scale research projects.

2.4. Coastal areas

Coastal zones are one of the most significant geographical and cultural boundary areas on the planet. Three totally different elements – land, sea and air – meet and clash along a line that is constantly shifting, in both time and space. Human beings, who, by nature, are "terrestrial" creatures, do not appreciate a totally liquid and unfamiliar element, the sea; it is uncontrollable, full of uncertainty and surprises. Bodei (2008) talks of inhospitable and hostile places, such as the ocean, that most humans have avoided for millennia because they evoke death; they humiliate us with their vastness, threaten us with their power and make us aware of our precarious existence. Man's pleasure in challenging a tempestuous sea is first recorded in the eighteenth century. From then on, these *loci horridi* were intentionally frequented and perceived as sublime, endowed with a more "...intense and enthralling beauty. This radical inversion of tastes…implies a new way to forge and consolidate individuality by challenging the magnitude and the supremacy of nature…" (Bodei, 2008, p7). According to Hegel (1770-1831), the sea emboldens and invites man to conquest, and is therefore an element unifying mankind, beyond the physical barrier (Hegel, 1975). This perception varies depending on individual cultural attitudes: there are also those who fear the sea as a dangerous and unknown element, a solid boundary that can be identified with *finis terræ.* Hegel (1991) provides a glimpse of mankind's renewed attraction for coastal settlements. People are attracted by the sea because it is perceived as an element of communication and trade, and hence also of growth. Carter (1989) defines the coastal zone as that space in which the terrestrial environment influences the marine environment and vice versa. On the basis of this definition, in specific physical and cultural conditions, entire cities, regions and states can be defined as coastal zones. This definition gives us the idea of an evolved category of people who settle along the coast and derive advantages and further cultural development from their proximity to the sea. Based on these considerations, the whole of Tokyo is a coastal settlement, but so is Rome, even though its historical centre is now located a few dozen kilometres away from the sea. Irrespective of the predominant geographical situation of its districts, a coastal city's problems and evolution emerge in the parts closest to the water. "...Tokyo's coastal zone, with its rows of factories and warehouses, found itself suddenly in the spotlight as the setting for this development. Vast tracts of landfill, long slumbering in disuse, took centre stage. A plethora of proposals to erect a glittering city of the future along the shoreline appeared, some of which have now been brought to completion..." (Jinnai, 1995, p219).

Coastal zones are hard to manage because of temporal issues such as tides and seasons and the overlapping of physical geography and hydrography, as well as overlapping jurisdictions and remits of individual government bodies and the competing needs of various civil society stakeholders. Longhorn (2005) points out that local, regional and national administrations are responsible for similar aspects of the same physical area and the uses of coastal zones. He gives the examples of fisheries, environment, agriculture, transport (inland and marine), urban planning, the land registry and the national cartographic and hydrographic services. Bartlett (2000) notes that many people are able to intuitively recognise a coastline, though they have more difficulty determining precisely where it extends landwards or seawards and its vertical growth. For this reason, and considering the diversity of the stakeholders, managing authorities and administrative structures, there are inevitable conflicts between users of coastal zones, developers and the rest of society. Urban coastal zones therefore become a prime vantage point from which to survey the way the global and local aspects interact, as Hegel intuited (1991, p268): "...in order to appreciate what an educational asset is present in the link with the sea, one should compare the relationship to the sea of those nations in which creativity has flourished whit those which have shunned and which…have stagnated internally and sunk into the most appalling and miserable superstition; one should likewise note how all great and enterprising nations put their way to the sea". Coastal zones are once again becoming an element of risk, terror and destruction, bearing out mankind's inability to manage natural disasters. Scientific research has provided several insights into ways to manage the conflicts sparked by the innumerable and discordant interests that mankind has generated along the coast, and to take a coordinated approach to tackling emergencies and catastrophes.

2.5 Urban areas: Scientific discussions in the 1970s and 1980s

In the 1970s and 1980s, a new phenomenon was noted, first in the US and subsequently in several European countries, that gave fresh impulse to research into urban areas: a slowdown and then a sharp decline in the population of urban areas in the US. These urban areas had begun to sprawl as a result of the increased availability of personal and public transportation. With changes in the organisation and location of manufacturing and business activities, people had begun moving from the city centre to the suburbs. Meanwhile, city centres, which essentially had cultural and historical value in Europe, were either being abandoned or occupied by marginal classes of society and immigrants, or radically transformed to make them more suitable to host new residential functions, emerging social classes and new tertiary activities. Two groups emerged in the course of these changes: those who wanted new architecture to create space for new activities, possibly saving a few particularly valuable historical buildings, and those who considered the value of the urban fabric as a whole, not just the value of individual monuments.

In his studies on urban morphology, Gianfranco Caniggia (1933-1987) conceptualised the city as an organic result of a dynamic procedural typology, which interprets the political, social, economic and cultural aspects to identify the forces that have given shape to the historical landscape. Caniggia (1976, p9) affirmed that the term 'typology', while maintaining the original term of 'classification' for similarity, could no longer be understood as an *a posteriori* recognition of such similarity. Caniggia acknowledged that his intellectual guide Saverio Muratori (1910-1973) had the great merit of attributing the concept of work of art and organism to a collective work such as the city (Muratori, 1960). In so doing, he had overturned the approach of the philosopher Benedetto Croce (1866-1952), according to whom reality had to be conceived as history. The Roman school of thought led by Muratori and Caniggia was shared, albeit not openly, by the British school that had developed around the work of M.R.G. Conzen (1907-2000), a German geographer who had been forced to emigrate to the UK in 1933 because he was being persecuted as a member of the social-democrat movement (Conzen M.P, 2004). Conzen was part of the German tradition of which Otto L. K. Schlüter (1872-1959) was a precursor, recognising the economic, social and cultural importance of the physical shape of cities and introducing the concept of the morphology of the cultural landscape as a counterpart to geomorphology in cultural geography. Conzen (1960) developed the technique of the "town-plan analysis" based on the study of "the town plan (streets, plots, lots, buildings and their aggregation)", "pattern of building forms" and "pattern of land use". He was the first to introduce methodological principles to define the analysis of urban morphology; he adopted an evolutionary type of thought approach in a geography text in English; he recognised the individual elements of a city as fundamental units of analysis; he used detailed cartographic analysis, and finally he drew up a conceptual system to analyse the evolution of the landscape.

The latter concept gave rise many years later to the "Urban Morphological Zone (UMZ)" classification in the Coordination of Information on the Environment (CORINE) programme begun in 1985 by the EC, with the priority subject being the identification of land cover. UMZs are defined as "a set of urban areas lying less than 200m apart" and comprise, as core classes: continuous urban fabric, discontinuous urban fabric, industrial or commercial units and green urban areas. In a text published after his death by his son Michael (Conzen M.P, 2004, p39-40), M.R.G. Conzen writes: "...during the life of any society in fixed geographical location its past and present experiences, thoughts, actions, behaviour patterns and aspirations accumulate to form the distinctive heritage of its spiritual possessions. It influences its actions as a particular society in a particular place or region. Such a coherent heritage, augmented by current experiences and handed down as a tradition from generation to generation constitutes an important historical factor. It accounts for the continuity of action of a local society and establishes its personality...". The term 'urban morphology' is used international literature to cover various types of research on the subject. Whitehand (2001) explains that while all these studies focus on the physical form of cites, the methods and tools are quite different for each disciplinary approach. He does however acknowledge that the British school of urban morphology is primarily geographical: "...it is primarily about how things fit together on the ground. It is hard to envisage ideas that are more geographical than the fringe belt concept and the morphological region. They are about how the urban parts of the earth's surface have been configured and reconfigured…" (Whitehand, 2001, p108). Two schools of analysis can be identified in architecture and town planning: typomorphology and space syntax. In geography there are those who follow the tradition of Conzen and others who interpret it as spatial analysis. Until now, there has been little interaction between architecture, geography and the other disciplines involved in the study of urban morphology. Sima and Zhang (2009) have identified the four fields most involved in this type of research: architecture, geography, science and philosophy. They cite G. Caniggia as the reference in architecture and, in geography, M.R.G. Conzen and his school. Their reference for science is B. Hillier (1999), who introduced the concept of space syntax, which explains the functional impact of spatial relations on human behaviour. In the view of French sociologist and philosopher Henri Lefebvre (1901-1991), space is not a neutral entity but a social existence. His theory (Lefebvre, 1991) suggests new ways to understand the processes of urbanisation. The social conditions of everyday life are connected, through the urban scale, to global flows of people, capital, information, ideas and culture. Thus the city becomes an active element of knowledge and an instrument of innovation in itself.

In the 1970s, the circulation of ideas and a different cultural climate encouraged the spread of the principles of preserving the environment and cultural heritage beyond these fields. The 1975 Amsterdam Congress following the European Year of Architectural Heritage organised by the Council of Europe approved the European Charter of the Architectural Heritage (Amsterdam Charter). Architectural heritage was defined as being an integral part of world culture and an instrument to encourage cultural cooperation and exchanges in a continent then divided into two opposing areas of influence. Although the Soviet Union was hostile to the Council of Europe, the

Amsterdam Charter was considered to be in line with the Helsinki Final Act (1975) of the Conference on Security and Co-operation in Europe, held only a few months previously.

A few of the ten points of the Amsterdam Charter are relevant to the issues discussed in this paper. Article 3 identifies architectural heritage as "a capital of irreplaceable spiritual, cultural, social and economic value". Article 4 considers all historical buildings to be assets capable of encouraging social balance by offering an environment suited to the development of a vast range of activities, which have contributed in the past to preventing social segregation. Article 7 calls for an integrated conservation of these assets and points out that the restoration of historical buildings must be carried out in a spirit of social justice and should not be accompanied by the exodus of residents of modest means. Finally, the charter expresses the hope that the conservation of architectural heritage will be tackled comprehensively and include all buildings with cultural value, from the most prestigious to the humblest, as well as the environment in which they are situated.

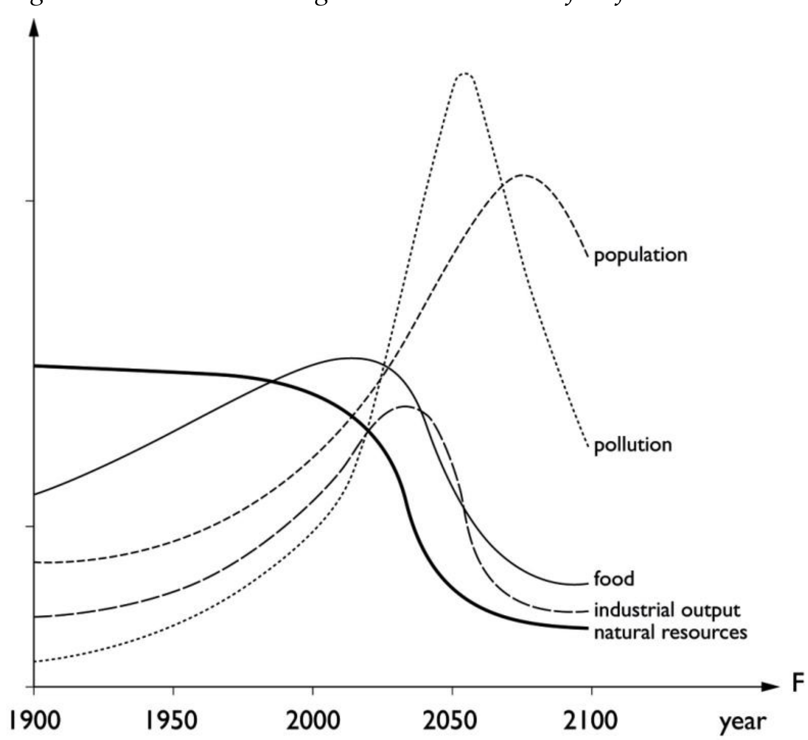
2.6. Urban areas at the heart of the debate on environmental resources

In 1972, the Stockholm Conference expressed concern about excessive atmospheric pollution and marked a turning point in the development of international environmental policies. Several major international conferences on the deterioration of the environment were held from the late 1970s on. The number of such events went up significantly in the early 1980s and warned about the emergence of global environmental risks. Significant events include the setting up of the Club of Rome in 1968, UNESCO's Man and Biosphere (MAB) Programme in 1971 and the United Nations Conference on the Human Environment, UNCHE, in Stockholm in 1972.

The Club of Rome was set up, among other things, to "understand why it is that today, when we have so much knowledge, information and power, we have got ourselves into a situation from which we are unable to extricate ourselves", to quote its founder Aurelio Peccei (1908-1984). Solutions for the present and the future can only be found if mankind as a whole is taken into consideration, within the "problem" of our own environment. Peccei complained of the reductive view until then of "man as an individual economic unit, a producer and consumer, or biological entity". In 1984, the year he died, he wrote (Peccei, Ikeda and Cage, 1985) that our model of civilisation not only exalts man's mastery over the world, if not the entire universe; it condones practically anything he does to assert his primacy, and justifies any means he chooses to use to this end. While theoretically we are all aware of the noble principles that should guide our actions, our civilisation is so anthropocentric that we feel free to ignore these principles whenever it suits us and whenever they interfere with our own interests or the flattery of our own ego.

In those years, many academics held the view that analysing the present was a means to anticipate future problems. Peccei, on the other hand, was convinced that unless we try to get an idea of where we are going, we could end up moving in undesirable directions; he was aware that it is only by identifying future scenarios that we can make an impact on the present. One of the contributions of the Club of Rome that can truly be described as a "future shock" was a report titled "The Limits to Growth" (Meadows et al., 1972), in which international growth trends for five variables were examined for the period 1900-2100: world population, per capita industrial production, available food production per capita, natural resources and pollution (fig.3.2.1).

Figure 3.2.1.The limits to growth. Source: modified from Meadows et al. (1972).



An exponential growth in consumption would lead to available resources being depleted by the end of the 21st century, with tragic consequences for mankind, as per capita food and industrial production would collapse, leading to a diminution in the size of the population. In Peccei's view, the problem was not resources, but rather the use that was made of them. Even oil could still be considered abundant at the time, but if it were incorrectly assumed to be infinitely available, there would soon be too little of it. The report was translated into thirty languages and ten million copies were distributed. It sparked considerable debate and a great deal of criticism of its methodology, but it undoubtedly had the merit of drawing public attention to the issue of the scarcity of resources. The fact that raw materials are available in finite quantities was brought home to all the next year, in 1973, when people in many industrialised countries were forced to walk or cycle to work and switch off shop signs and street lamps to cut down on oil consumption. The finite supply of available resources also brought up the issue of the need to reduce consumption, either directly or by recovering the by-products of industrial activity that were abandoned in the environment. The work of American biologist Barry Commoner (1971) was particularly well known; he pointed out the limits of "economic" societies that, by withdrawing resources from nature and returning to it by-products in the form of waste, function in an "open cycle", unlike "ecological" societies, which operate in a "closed cycle" as they do not envisage the production of waste.

UNESCO organised the MAB Programme to enhance the relationship between people and their environment in the whole range of bioclimatic and geographic situations in the biosphere, and thereby contribute to resolving problems connected to the use and maintenance of natural resources. Scientists, intellectuals, public opinion and eventually governments began be concerned about atmospheric and water pollution and the decrease in arable land. The 1972 UNCHE conference in Stockholm clearly acknowledged these concerns. Following this landmark event, the problem of the environment and the need to protect it gradually began to be included in all policies and planning by international organisations as well as national governments. The Stockholm Conference started off a period lasting around twenty years (1972-92) during which international, national and local government bodies, political parties, businesspeople, unionists and individual citizens became aware that natural resources are finite, and that they must be preserved (Montanari, 2009).

In 1992, twenty years after the Stockholm Conference, the Earth Summit in Rio de Janeiro once again addressed environment concerns and paved the way for a new cultural process: sustainable development.

2.7. The process of urbanisation: The market economy and central planning

Another phenomenon that was studied in depth in the 1970s was that of urbanisation and the growing number of people residing in urban areas. The ratio of the urban population to the total population was about one to two in the 1950s; it grew to two-thirds in the 1970s and is expected to reach five-sixths by 2030. Tisdale (1942, p.311) wrote: "... Urbanisation is a process of population concentration. It implies a movement from a state of less concentration to a state of more concentration..." From the 1970s on, data concerning the changing urbanised population started a fresh debate on a phenomenon defined as counter-urbanisation, to indicate the diminution in the size, density and heterogeneity of urban areas. Adapting Tisdale's definition, Brian J. L. Berry (1976) defined counter-urbanisation as "... a process of population deconcentration; it implies a movement from a state of more concentration to a state of less concentration..."

The reasons for this phenomenon can be traced back to the lifestyle changes of the 1960s – the economic boom and the high birth rate in European countries, especially northern Europe – that made personal transportation available at affordable prices. Urbanised people could now achieve their desire to move to residential areas on the outskirts of large cities areas where they could have more spacious homes and districts with large green spaces, and where the distribution of homes was more democratic than in the historic centres, where it had always been hierarchically structured. The process was first identified in the US by Berry (1976) and then studied in Europe with the CURB (1971-1982) or "Costs of Urban Growth" project sponsored by the Vienna Centre (ISSC). CURB was one of the first European international comparative research projects in the field of urban geography, and Berry's work was its starting point. Berry took a global view of the phenomenon of urbanisation, and tested the theories discussed by the International Geographical Union's Commission on Processes and Patterns of Urbanisation, which he had applied to the US, in Western countries – both free market economies and centrally planned economies – and developing countries. In an earlier work, Berry (1973) had outlined the consequences of ever-increasing urbanisation on people, and how the various socio-political forms had a significant influence on the morphology and dimension of contemporary urbanisation. His interest in the contrasts offered by the different situations then appearing in the world had led Berry to put together a group of scientific colleagues in various parts of the world.

Some of them, particularly the Europeans Drewett and Goddard, went on to form the nucleus of researchers working on the CURB project. The main results of CURB, particularly the theory of urban cycles, were published in v.d. Berg et al (1982). The theory of urban cycles (fig.3.2.2 – fig.3.2.3) illustrates the increase and decrease of the population of the CORE, the RING, and the Functional Urban Region (FUR) as a result of the various phases of urban development: "...Urbanisation is in force when the growth of the core dominates that of the ring, while the FUR as a whole is growing. Suburbanisation is in force when the growth of the ring dominates that of the core, while the FUR is still growing. Desurbanisation takes place when the decline of the core population leads to a decline of the population of the whole FUR. Reurbanisation takes place when the share of the core population in the total population of the FUR is increasing again..." (v.d.Berg et al., 1982, p36).

Figure 3.2.2. The theory of urban cycles. Source: modified from Berg v.d. et al. (1982).

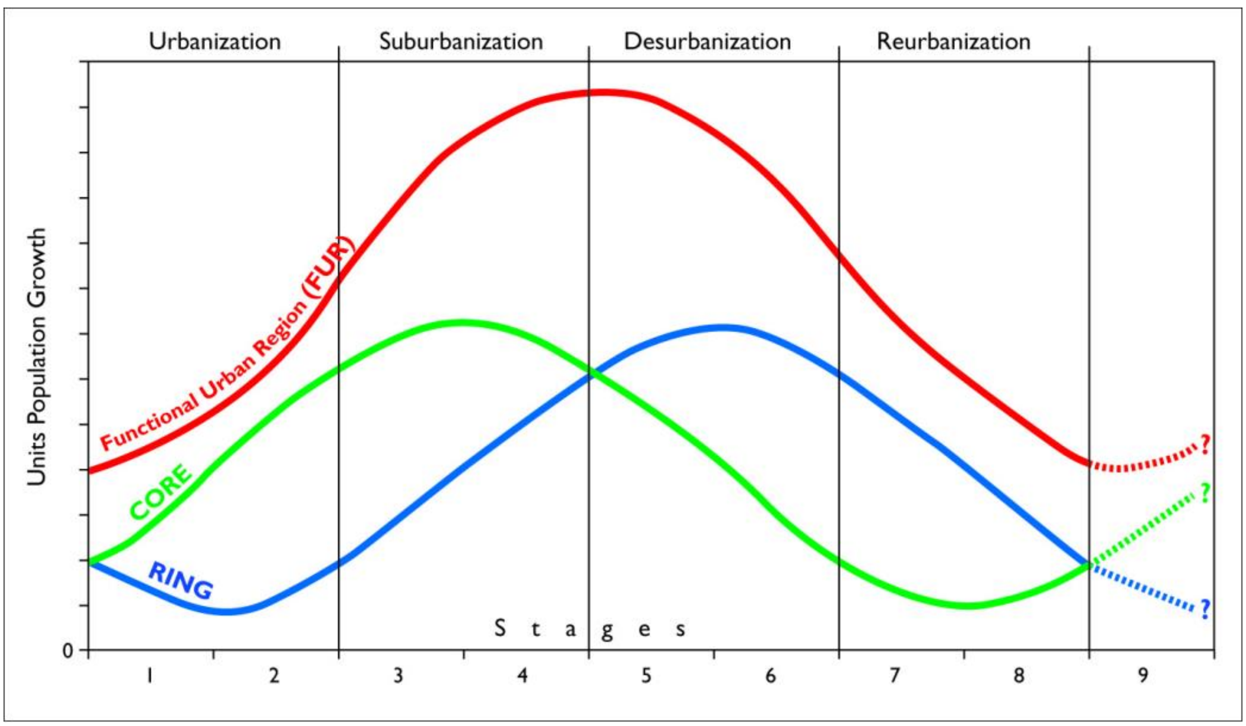
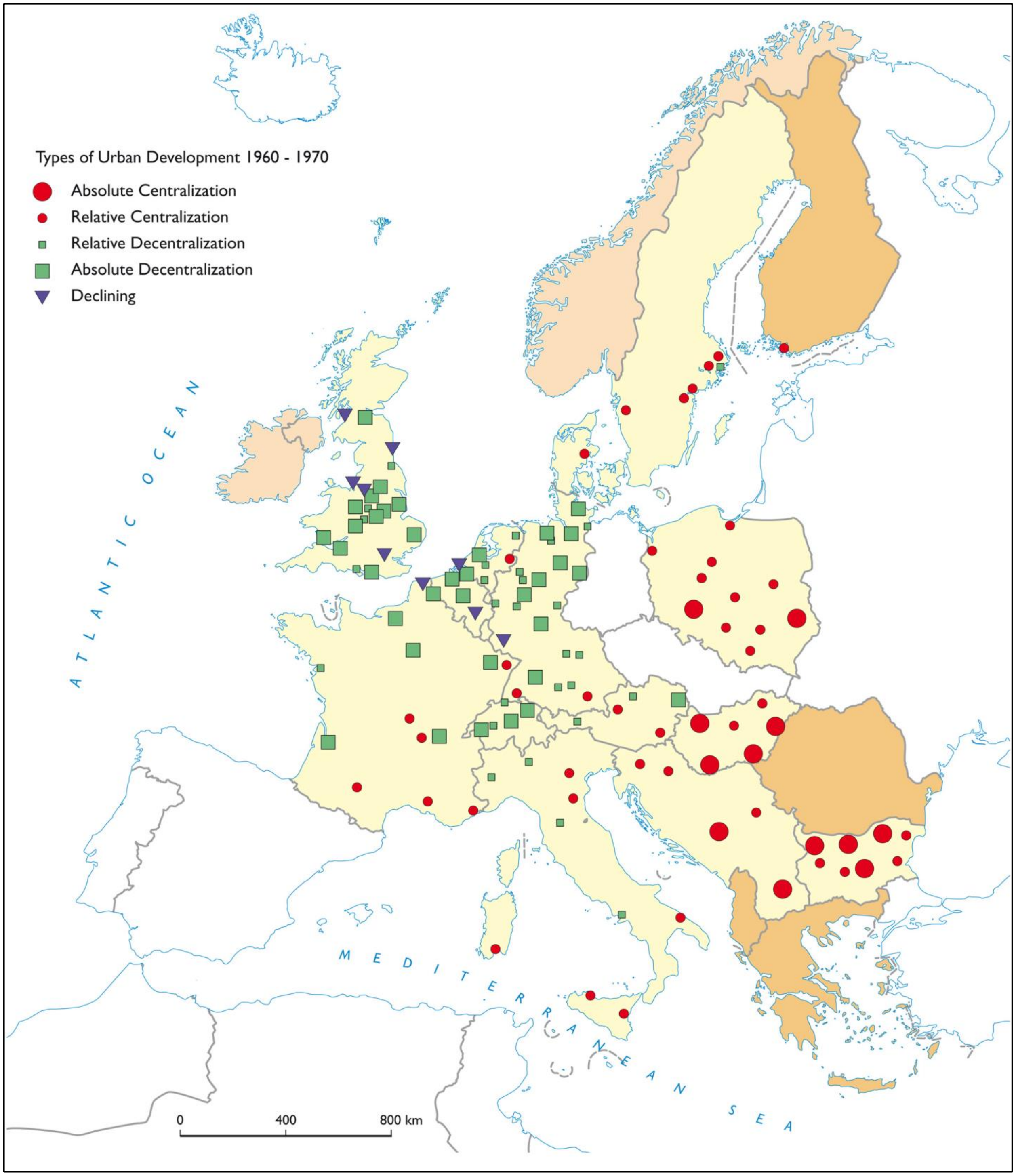


Figure 3.2.3. Types of urban development. Source: modified from Berg v.d. et al. (1982).



Concerning this model and the idea that there could be a reference model for urban development, it emerged, within CURB as previously within the IGU, that the researchers from countries with centrally planned economies had different views to the researchers from countries with free market economies. The former denied the existence of such a model, because the Eastern European countries had concentrated their efforts on developing manufacturing activity and settlements within an "urban network" of the rural type. This is why v.d. Berg et al. (1982, p5) clearly point out that the form, scope and methods of development of an urban settlement are the result of socio-economic development, geographical conditions, institutional characteristics and finally people's own choice. As these conditions are very different in countries with centrally planned economies, the results could also be expected to be different. Enyedi (1992), who as a member of the Hungarian Academy of Sciences had participated in the CURB project, later retraced the urban history of Eastern European cities after 1988. The picture that emerges shows quite a difference between "ideological" theory and its implementation: "...in political declarations, they praised the Soviet example, but planners were looking for Western patterns in physical planning management and in developing technical civilisation in cities…informal responses have had a much more important effect on urbanisation than earlier supposed…" (Enyedi, 1992, p. 878).

CURB is important for the simultaneous management of problems related to international comparative research, not least because it helps to understand how information must be managed in order to become an element of knowledge. The first project proposal was presented by Professors G. Gaudard and J. Valarché of the Istitut de sciences économiques et sociales at the University of Fribourg, Switzerland, to the board of directors of the Vienna Centre on 23-24 April 1971. The project was based on these premises:

(i) Urban growth is too costly, for free market as well as planned economies;

(ii) Unlike geographers, economists have not undertaken studies of urban areas, and therefore have not considered the problem of increased costs;

(iii) Sociologists have identified a system of social imbalances in cities without, however, attempting to quantify them.

Urban growth was considered from three standpoints:

(i) Demographic growth;

(ii) Economic growth and the product of the urban economy;

(iii) The spatial growth and spatial extent of urbanised areas.

This wide range of growth typologies gave rise to the need for a study involving all the social sciences in order to do the following: (i) analyse costs using statistical data, field surveys and monographs. The costs to be identified were monetary, non-monetary, public and private; (ii) quantify the costs of urban growth; (iii) the policy of urban growth. Next, parameters were drawn up to choose the case study cities, where the following had to be analysed:

(i) The operating costs needed for the city to maintain a constant quality of urban services for all users – both producers and consumers;

(ii) The adaptation costs required to improve the quality of urban services provided to a stable number of users;

(iii) The costs of the growth required to ensure the quality of services provided to a growing number of users;

(iv) The development costs required to improve the quality of services provided to a growing number of users.

To the abovementioned costs must be added costs in areas over which the city has a certain influence; the size of these areas was also studied and analysed. This is why the concept of the functional urban region gradually came into being and began to be developed as the research project progressed. This definition was then expanded, on the basis of statistical characteristics, into core areas and outlying areas of influence. Together with other connected areas, all these areas form a bigger settlement system. Using these parameters, CURB demonstrated that it is possible to describe the cost elements for a particular city, or for several cities within the same country. It is considerably harder to construct a cross-national comparison, and not only because of the generic problem of public and private costs, monetary and non-monetary costs. The greatest difficulty is the differences in the value systems, administrative bodies, levels of centralisation or decentralisation, cultural heritage in the broadest sense and, above all, the way that the variety of social and economic policies implemented by individual cities can influence settlements. These interpretive difficulties led the CURB participants to research additional factors that are key for determining the urban and regional structure, for generating the distinctive aspects of social, economic and spatial changes, and for the political implications of the processes that highlight these changes. They collected documents, statistical information, economic data and econometric analyses that made it possible to:

(i) Analyse the theories that have most probably influenced urban planning policies in Europe on the basis of the knowledge acquired in the sphere of the connections between the economic growth and urban transformation phases;

(ii) Systematically summarise existing research and the trends of specialised literature on the themes of urban growth and policies. This activity was particularly appreciated in countries where language problems make knowledge harder to acquire;

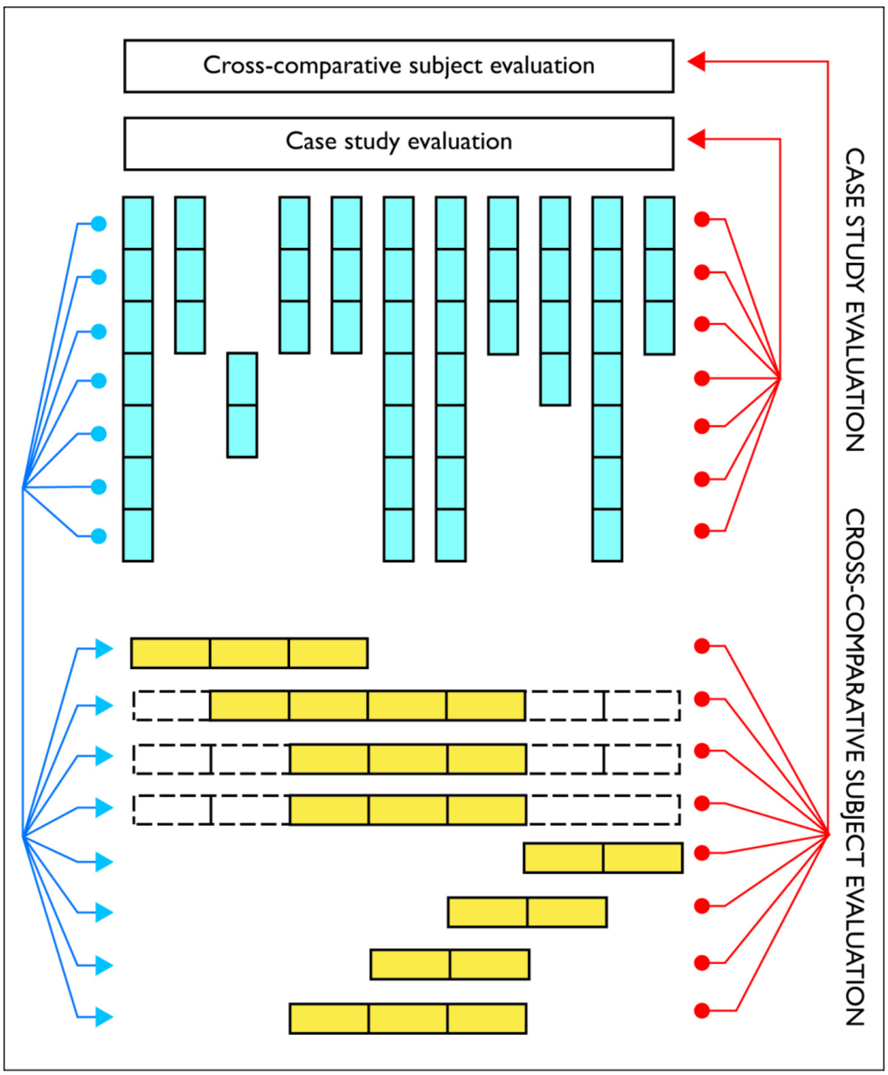
(iii) Examining urban policy themes so as to identify specific characteristics, major methodological changes and inadequacies in both the analytical approach and policy implementation phases.

The CURB project findings demonstrated that it is not the type of economy or administrative system that causes major differences in urban growth, contrary to a number of theories held until then, which made a superficial connection between urban growth phenomena and free market economies or highly centralised administrative systems. These situations exist in all European regions, albeit with specific characteristics. CURB formed a basis and provided an opportunity for an intense scientific debate that lasted a few years and produced around thirty working documents. Although special attention was paid to the unique aspects of the comparativeness of Eastern and Western European countries, very soon the main concern became to give the research project a purpose. So the objective was to give anyone in charge of managing the project advance knowledge of development trends and the ability to anticipate problems in a timely fashion. This is why CURB was extended to the largest possible number of European countries. Meanwhile, CURB's findings were presented at various international conferences to be compared with the findings of other national and international projects. One of the first opportunities to share the findings was the VII Conference of Mayors of the World's Major Cities, held in Milan and Turin in 1978, which had the costs of urban growth and all the related problems as its theme. This event involving the mayors of the world's major cities was an opportunity for a heated debate with the people who tackled urban management issues on a daily basis, and therefore for a very critical examination of the transferability of methodological approaches. Conversely, CURB's presentation at the IIASA's conference on "The Analysis of Human Settlement Systems" in Laxenburg, Austria, in 1978 focused mainly on the theoretical aspects. Through other events in a number of European countries, CURB researchers drew up a scientific itinerary that would allow them to check the development of the territory from a micro and macroeconomic standpoint as well as at the international, national and regional levels. It was on the elaboration of this research material, and these scientific debates, that the most original part of the project – the theory of growth and decline in urban areas in Europe – was constructed.

2.8. Urban models as an instrument of a sustainable future

In the 1980s the quantitative dimension of the urban phenomenon began to stabilise, and research began to focus mainly on the qualitative aspects of the way processes had unfolded. The URBINNO (1987-89) project ("Innovation and urban development: the role of technological and social change") used innovation parameters to identify urban landscape dynamics. Researchers analysed a hundred-year period from the end of the 19th century onwards to document the principal changes wrought by industrial and post-industrial development. For each of the five historical periods identified, seven levels of analysis were considered (fig.3.2.4) following the above-mentioned parameters of urban morphology (Montanari, 1990).

Figure 3.2.4. From case-study to thematic approach. Source: modified from Montanari et al. (1993).



Innovations were then studied in further detail in the transport sector and the urban form, the spatial effects of new technology, the spatial organisation of the city at the level of the smallest urban units, urban centres, lifestyle changes in relation to the different urban forms, the development of public spaces as a basic element transforming different lifestyles, and the productive and functional transformation of the waterfronts (Montanari, Curdes and Forsyth, 1993). Figures 3.2.5, 3.2.6, 3.2.7, 3.2.8 following the theory of urban cycles illustrates the changes in lifestyle and urban form.

Figure 3.2.5 & 3.2.6. Changing life styles versus urban built form: stage of urbanization. Source: modified from Dunin-Woyseth (1993).

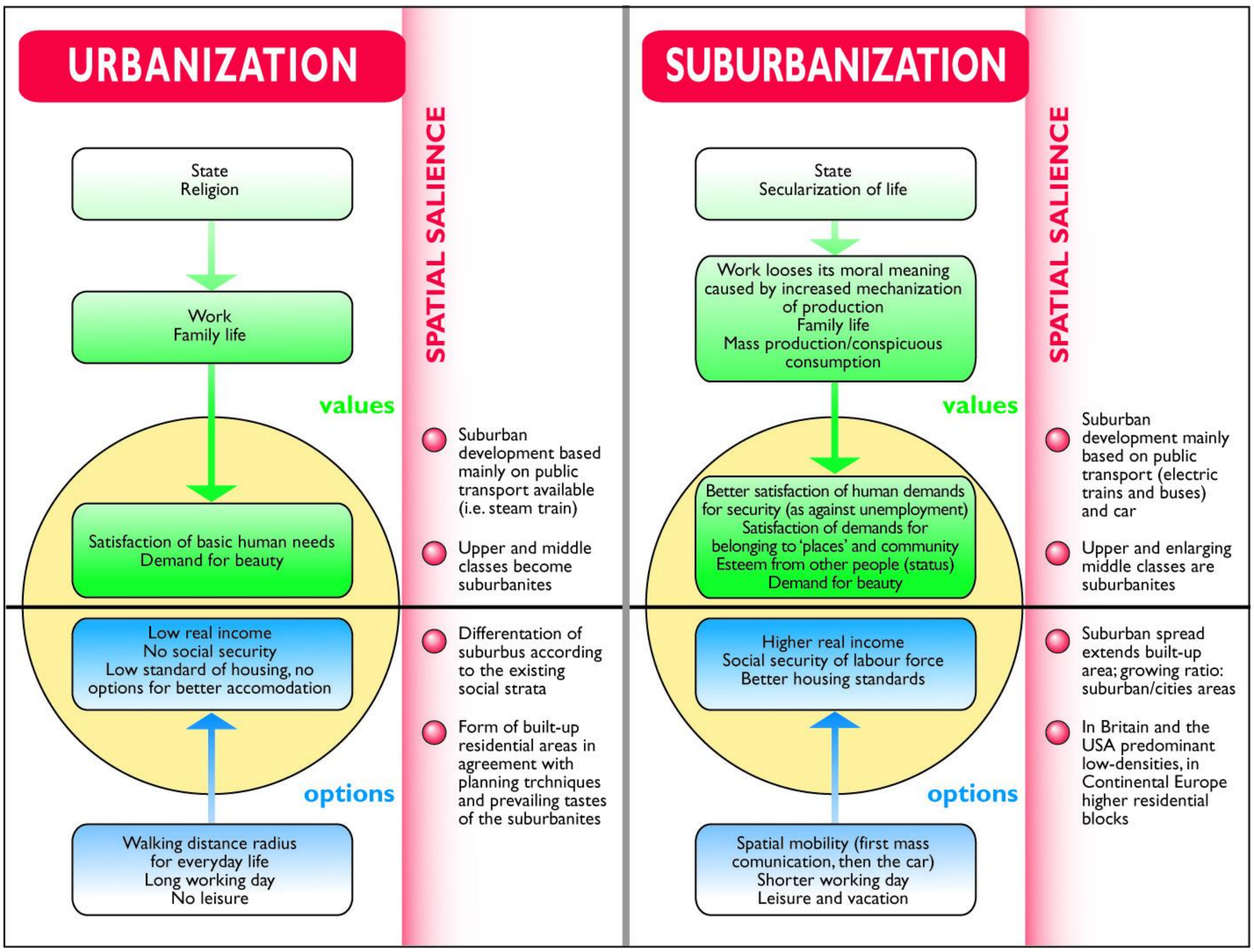
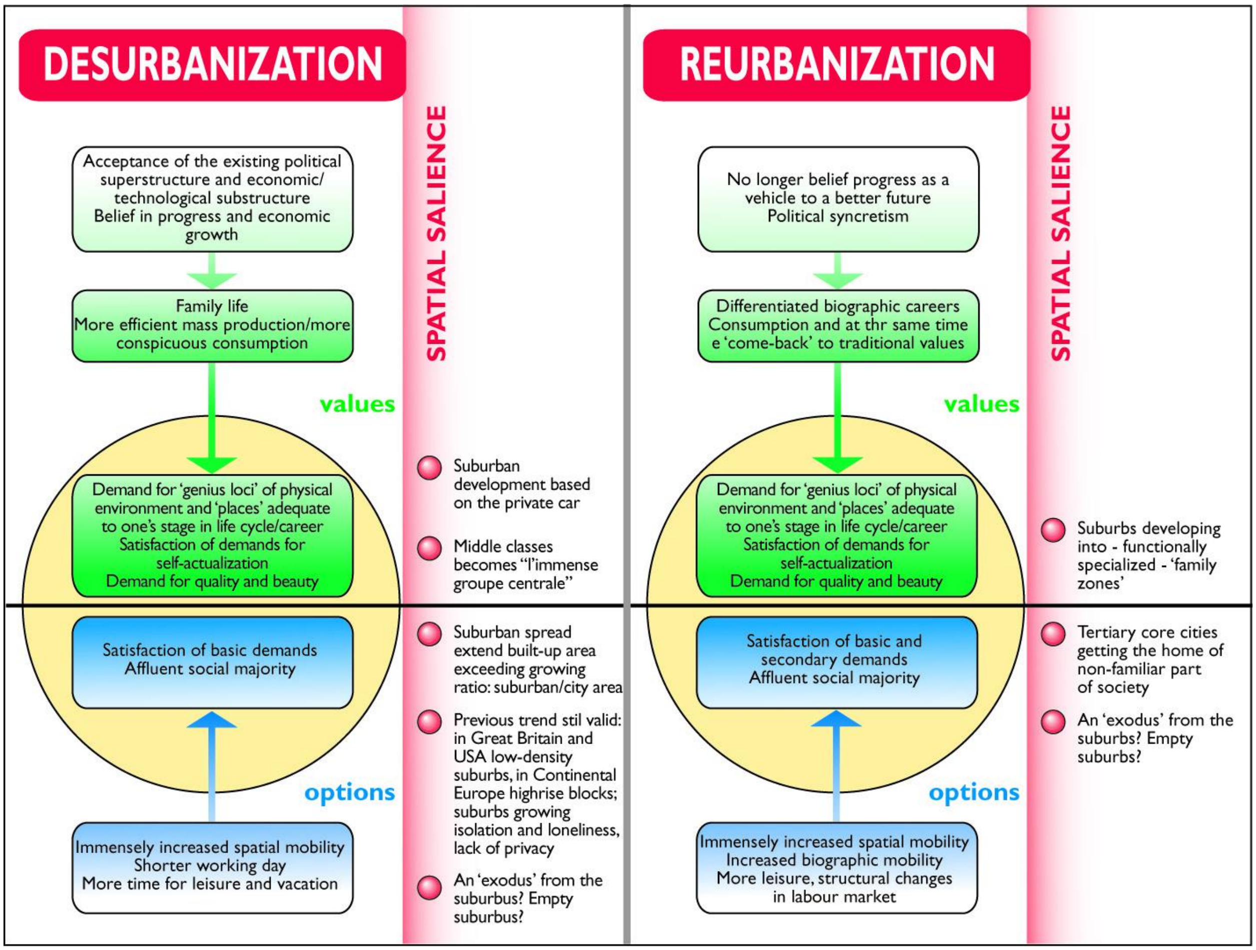
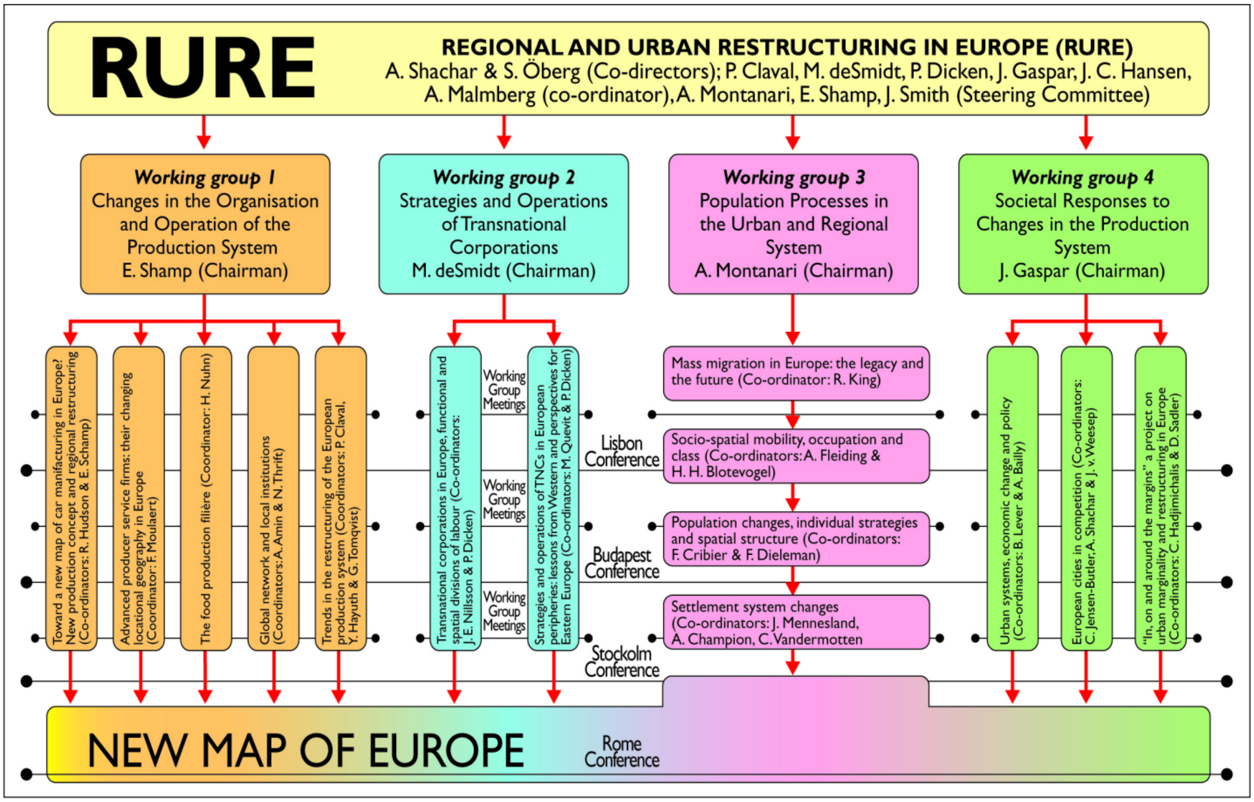


Figure 3.2.7 & 3.2.8. Changing life styles versus urban built form: stage of desurbanization. Source: modified from Dunin-Woyseth (1993).



The consequences of the internationalisation of the economy and the social, economic and political processes that changed Europe in the 1980s and 1990s were the basic elements examined by the RURE ("Regional and Urban Restructuring in Europe") project (1988-1993). The SELMA ("Spatial Deconcentration of Economic Land Use and Quality of Life in European Metropolitan Areas") project (2002-2006), on the other hand, examined the phenomenon of deconcentration of production activities and its effects on the environment and the quality of life.

Figure 3.2.9. RURE Project. The thematic approach.



The RURE (fig.3.2.9) project was undertaken during a period of great global change: production changes following the growing introduction of ICT and the internationalisation of the economy. In Europe, the period from the end-1980s to the early 1990s marked the collapse of the Soviet Union and the fall of the Berlin Wall. These major transformations brought about a shift in Europe's political and economic barycentre, altered relations between the USA, Europe and Russia (Malmberg, 1996) and created a significant change in migratory flows (fig 3.2.10), with four different types of mobility emerging in Europe: mobility due to a demand for unskilled labour, mobility for family unification purposes, mobility for hybrid migration and leisure and finally migratory flows from countries with post-colonial links (fig.3.2.11).

Figure 3.2.10. Mass migration in the context of regional and urban restructuring in Europe.
Source: modified from Fielding (1993).

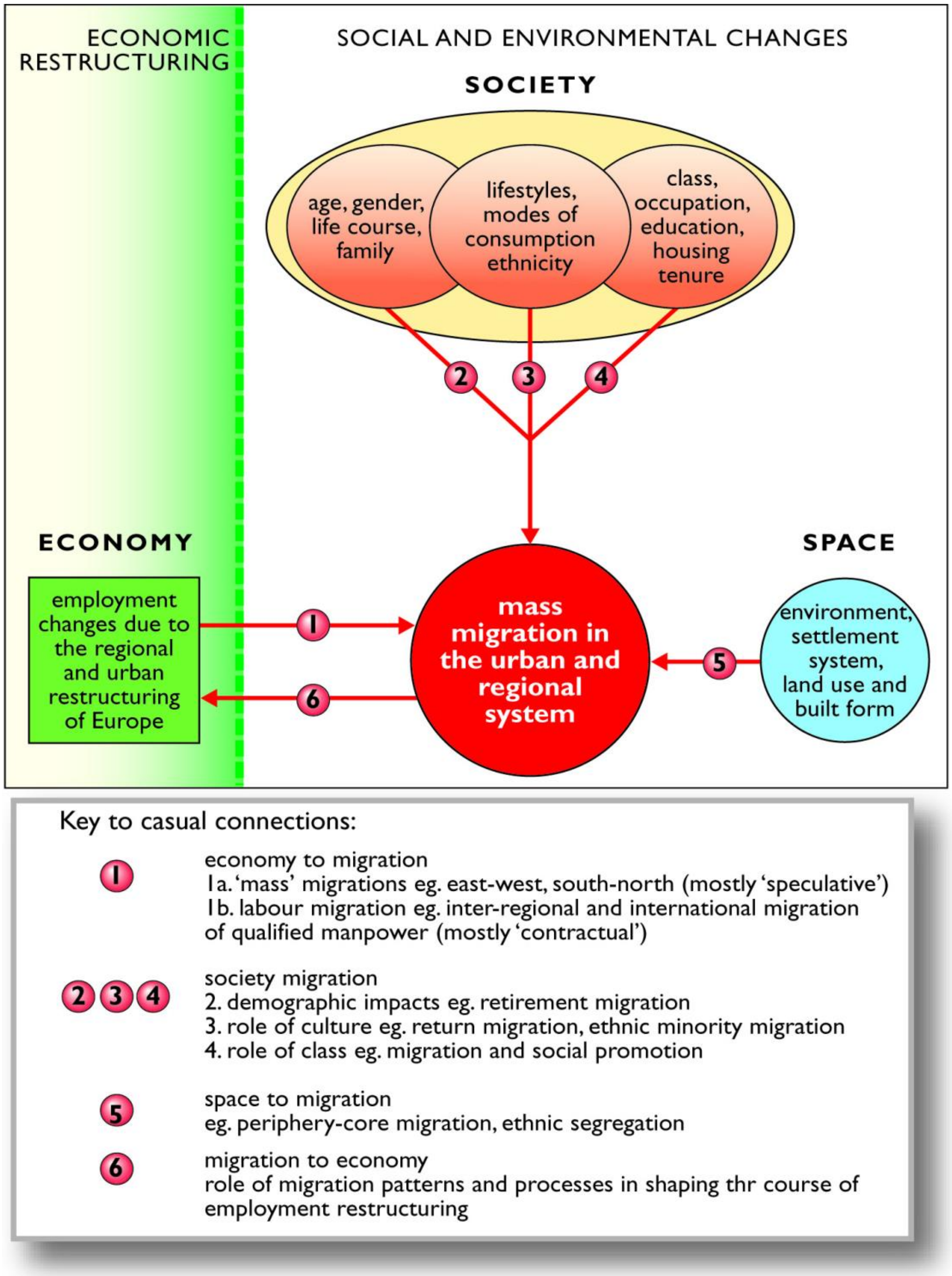
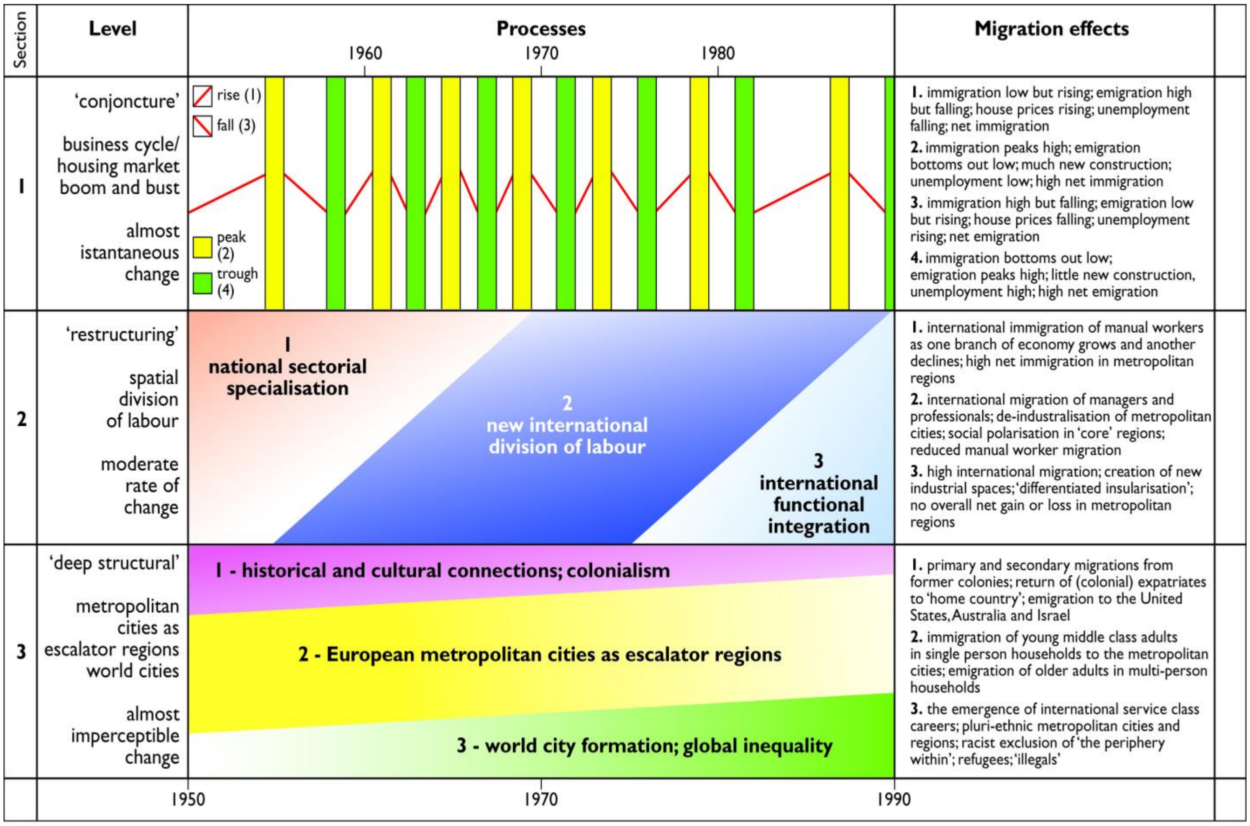


Figure 3.2.11. The effects of economic restructuring on mass migrations to and from Western Europe's metropolitan cities and regions. Source: modified from Fielding (1993).



From preliminary studies on migratory and tourist flows (King, 1993; Montanari and Williams, 1995) RURE began to identify one of the key themes for social scientists worldwide: human mobility, which takes on various forms, including migration and tourism, and is an element of scientific study because of the relationships and overlapping that these flows imply. Tourism is interdependent with particular forms of migration; it is a form of mobility, of variable duration, that in turn engenders further, distinct forms of migration such as those activated by the demand for tourist services or prevalently consumerist forms of varying duration and motivations, such as holiday homes, seasonal migration, migration related to lifestyle changes and the migration of old people and pensioners.

An IGU Commission on Global Change and Human Mobility, GLOBILITY (2000-2012) was set up on the subject; the information and scientific results it has collated are available on the website: http://w3.uniroma1.it/globility/.

The findings of the empirical research presented by members of the "Globility" commission and published in studies by Montanari (2002; 2006), Ishikawa and Montanari (2003), Montanari and Salvá Tomás (2005), Montanari (2012a) provide a mosaic which, however, remains incomplete, as it lacks the findings of targeted research from which we can infer the existence of a qualitative model able to help us to understand how the various forms of human mobility are formed, and overlap, at different stages of the economic growth of urban areas. The model was intended to check how the various types of mobility were interconnected – in such a way that each form of mobility leads to others, in either simultaneous or successive phases. The model was debated, enhanced and finally approved after discussions with a significant number of special observers – the academics from around one hundred economically and culturally diverse countries who participated in "Globility".

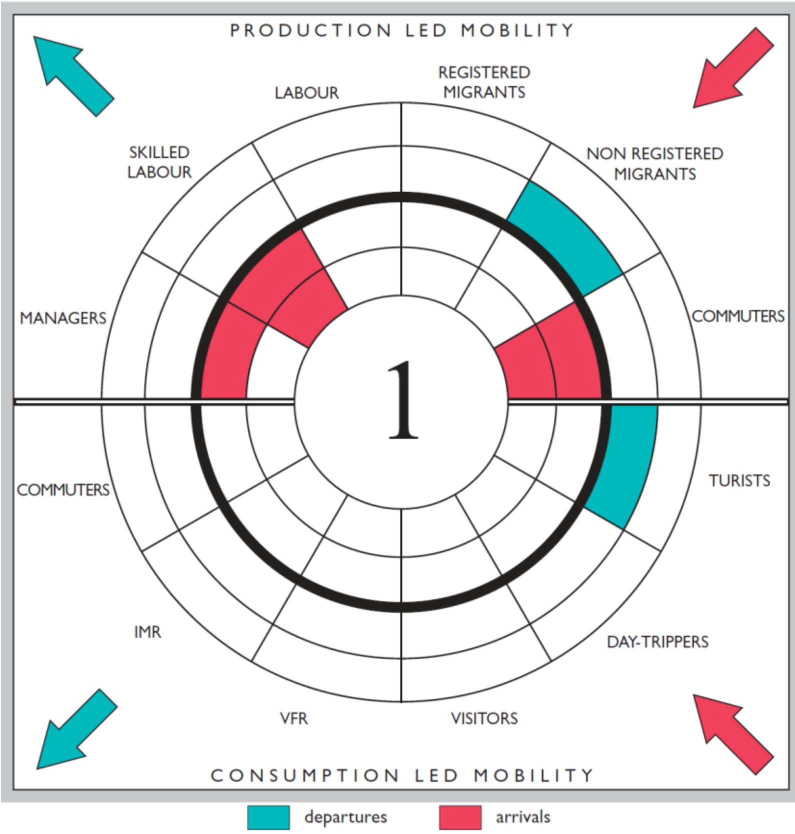
The first two variables taken into consideration were human mobility and commuting, as it was thought to be important for the territory to detect whether an individual manages to return to his place of residence the same day or has to spend the night elsewhere (Montanari, 2005) – also keeping in mind the progress of transport, which now allows for shorter travel times than in the past, regardless of the distance. The overnight stay is considered a strong element of the relationship with the territory, as it implies the availability and use of accommodation and, therefore, the activation of a different type of housing and time use. The second group of variables is temporary and permanent mobility, the difference being the extent of the minimum stay, which can be considered in terms of weeks, months or years. In fact, rather than the minimum or maximum stay, it would be important to refer to the lifestyle category of the person who decides to move from one place to another. So the meaning of "temporary" can be highly variable. "Temporary" means that there is no intention of "putting down roots" in the territory – on the contrary, a return to the place of usual residence is envisaged, even if decisions that appear to contradict this hypothesis are taken, such as buying a house and transferring one's family; hence an inferior category of relationship with the territory. On the other hand, when we look at Australia – where individuals move house twelve times in their lifetime on average – we realise that in that case it would probably be hard to identify a place of usual residence, and that each move would be considered "temporary" if referring exclusively to the length of stay. Taylor and Bell (1996) point out that in Australia, an increasing number of the people who migrate temporarily have a "network of places" rather than a place of usual residence. The place where they are going to stay is no longer one fixed point in the territory so much as a network of interconnected places. This is not a surprising consideration, because in fact manufacturing is organized into a network.

Other variables taken into consideration are local and regional flows and national and international flows, in relation to local development within a network system. In this context, it is important to consider the international flow as the manifestation of a border crossing, whether in terms of bringing in capital, participating in the completion of the project, or drawing benefits from the development that will possibly come about as a result of it. In the illustrative diagram here, the difference is mainly between two kinds of mobility: production-led and consumption-led. It is merely an example demonstrating a priority activity that will certainly be followed by a number of superimposed layers of secondary mobility activated by different production and consumption requirements. Another basic difference is proposed, between incoming and outgoing flows. In production-led mobility, the main job – that of the managers – is distinct from that of specialised technicians and workers. It is possible to differentiate between the many hierarchical levels in production during the various growth phases. Then there are migrant flows for which the only identifiable factor is their documented status; they have presumably been in the territory for a long time, and have therefore integrated more easily. There are other, more recently acquired migrants – by definition, temporary migrants – who, besides the uncertainties of their temporary situation, do not have a precise regulatory and administrative standing. In consumption-led mobility, tourist flows were taken into consideration, as well as day trippers and local visitors. The latter two are not counted as tourists because they do not book an overnight stay, and therefore do not fall within the conventional international definition. The VFR (visiting friends and relatives) category is also made up of tourists, who have not so far been considered to be tourists because they do not book hotel accommodation. Another important category of consumption-led mobility is the migratory flow of pensioners, like the VFR, they are better known by their acronym, IRM (international retirement migration). This category includes different lengths of stay. IRM often includes a production-led component, as, once they reach their holiday location, people who have moved there might become managers of real estate agencies, cafés, restaurants, hotels and other businesses, mainly targeting their own countrymen. Commuting – recurring, repeated mobility – qualifies as both production-led and consumption-led mobility. In the case of production-led commuting, this flow indicates a lack of manpower in the urban settlement that is being studied, but it can also point to a significant lack of housing, whether in absolute or qualitative terms, thereby forcing managers, technicians or workers to live in surrounding areas and undertake a daily commute. A settlement undergoing restructuring and development cannot always manage to offer new arrivals the structures and services they require. In the early stages of development, therefore, there is an outgoing commute in search of structures for sport, leisure and shopping as well as healthcare, training and education. In the later stages of development the inverse can happen, i.e. it is the settlement that offers surrounding areas more and better-quality services and infrastructure.

The model theorises six stages of development of an urban settlement, hypothesising the different ways in which people behave depending on the reason for their movement – work, leisure, family life, migration or escape. Larsen, Urry and Axhausen (2006) identify five independent types of mobility that shape and reshape the geographies of networks and travel in the contemporary world. These can be the physical journeys of people, goods and things, journeys of the imagination based essentially on memory, virtual journeys on the Internet, and finally journeys involving direct communication between individuals by post, telephone, fax, videoconference or other means. Figures 3.2.12, 3.2.13, 3.2.14, 3.2.15, 3.2.16, 3.2.17 show – for incoming as well as outgoing flows – three levels of qualitative analysis: inexistent phenomenon, low and high flow intensity. Urban tourism is the result of the interconnection between production stimuli and the consumption phenomena related to them. There is a preliminary stage that we can define "Stage 0", in which the settlement reproduces itself in a state of self-sufficient balance. We must assume that there are no significant endogenous, let alone exogenous, development phenomena. Hence the territory is not affected by any form of dynamic mobility; however, over the years it has constantly lost population because of outward-bound flows towards either the rest of the country or the rest of the world, with the partial compensation of a possible natural population growth.

Stage 1 – the activation of the process – is a phase of preparation, planning and checking the feasibility of a project, for example the organisation of an event through forms of exogenous development devised by an international body or a multinational, or a production-related investment made by returning migrants. International managers come to the territory – for short periods, and therefore temporarily – to roll out the first stages of the project. For longer periods – which we can define as semi-permanent – and with larger flows, it is the managers who will carry out the project who arrive, followed by specialized technicians. As yet, there is no housing market and no hotel accommodation, and so these employees find themselves makeshift housing options within and outside the concerned area – contributing, therefore, to an incoming commuter flow that also comprises generic manpower from the surrounding areas. At this stage the area is still poorly equipped in services and infrastructure compared to the areas in its immediate vicinity, as a result of which outgoing consumption-led commuter flows are created to satisfy the need for leisure, sport, and better-quality shopping (fig. 3.2.12).

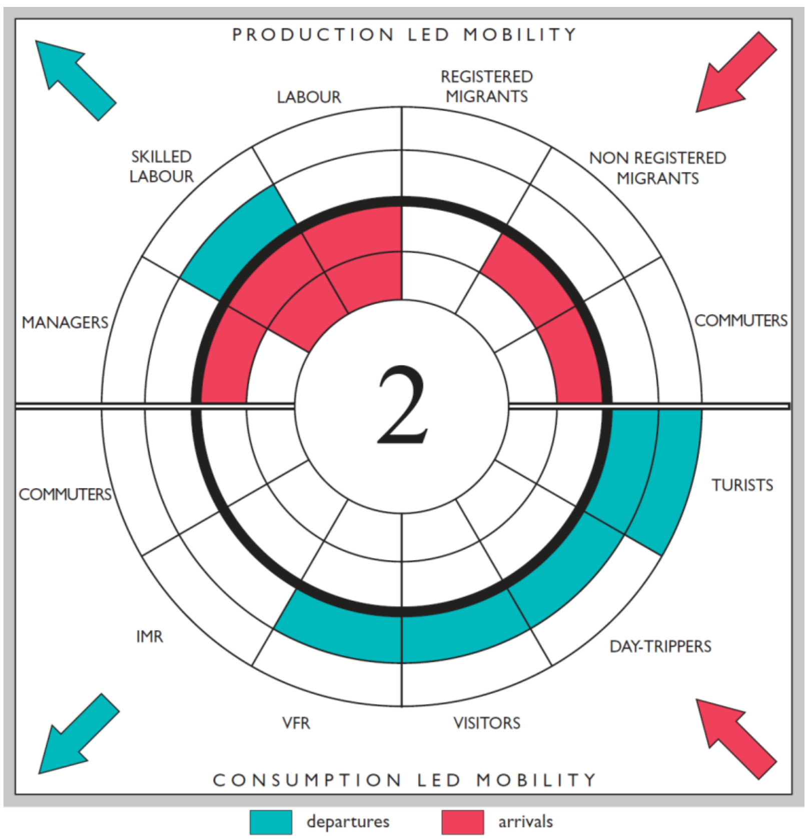
Figure 3.2.12. Stage 1, the activation of the process.



Stage 2 – the operational stage of the project – marks the start of the roll-out phase, stimulating, strengthening and pulling expectations in its wake. There is a need for manpower to manufacture goods and build infrastructure. While this is generally development of the exogenous type, the public body behind the settlement weighs in to ensure that local manpower is given priority, even if this involves a period of training or refreshing skills. These are workers from the agricultural sector with no specific experience. They are poorly suited and often not particularly keen on the kind of jobs on offer in the new businesses and services, in sectors where a lot of manual work is involved. In addition to these local workers, there are workers from a larger area within the country. The creation of specialised jobs sets in motion a process of return of workers who had previously emigrated, either abroad or elsewhere in the country. Meanwhile, another process begins, of replacing the manpower that was sent into the area in the early stages of the project. Various forms of temporary mobility then begin to get going alongside these forms of permanent mobility. Seasonal workers are employed to tackle specific production needs within and outside the company. Moreover, a flow of middle managers and managers from other branch

offices of the same corporations, whether in the country or abroad, begins. To satisfy these new needs, new housing and hotels are built, or existing ones are refurbished. A higher income is assured for employees of the company as well as its suppliers, and the introduction of new lifestyles and habits creates, among other things, the prerequisites for a lucid use of leisure time and holiday periods – outgoing tourist flows, therefore – either immediately outside the area or, in more general terms, in the country or abroad. The area is not yet sufficiently equipped, and therefore it is possible that commuter flows take place towards neighbouring areas with better schools, health services and shops (fig. 3.2.13).

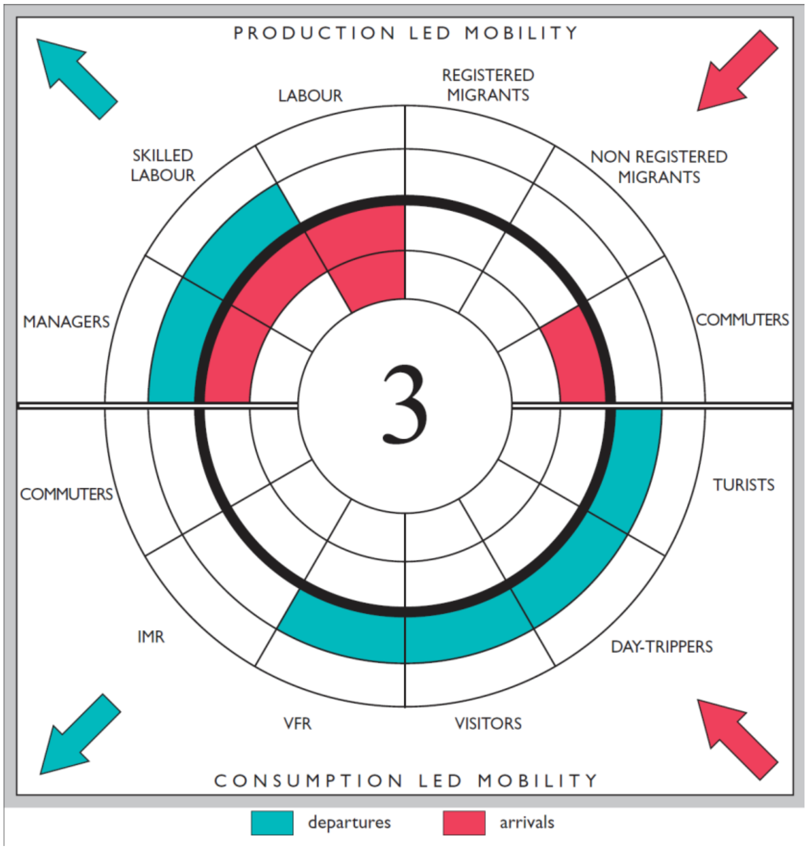
Figure 3.2.13. Stage 2, the operational stage of the process.



Stage 3 – the development of the project – considers a period in which there are new production phases and, therefore, the substitution of the specialised manpower and managers who had transferred to the area for the long term as well as the short term. This period also marks the emergence of the first cases of regulation by local communities demanding greater involvement in making the decisions that will have the greatest impact on the territory. There are stronger demands to hire local manpower, which is now unemployed because it had abandoned

traditional sectors of activity as a result of the labour market restructuring in the previous stage of development. Meanwhile, the process of providing suitable services and infrastructure has started, and there is therefore less and less need to use those of neighbouring areas (fig. 3.2.14).

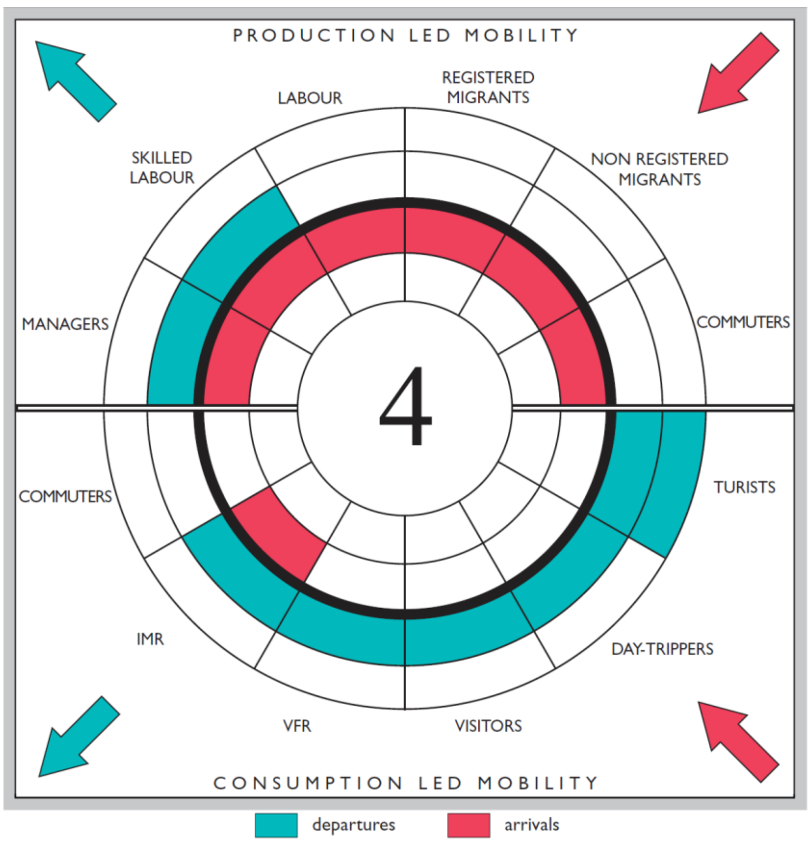
Figure 3.2.14. Stage 3, the development of the project.



Stage 4 – the reinforcement of the project – envisages the strengthening of the production sector and, simultaneously, the launch of a mediation policy with the local authorities for joint development. The exogenous development of some manufacturing situations is followed by a phase of adapting to the needs of local development and, therefore, the social, cultural and economic growth of the entire territory, with strategies and instruments typical of the bottom-up approach. There is no longer any possibility of finding the required manpower locally, and so recourse is had to specialised personnel, recruited both nationally and internationally. In addition, the first flows of workers from developing countries begin to arrive, both documented immigrants – to be used at the head offices – and undocumented, to be employed mainly in satellite companies, services and the black market. Undocumented migrants can fit into the "permanent" category when the work they have moved to the area for continues over time, so

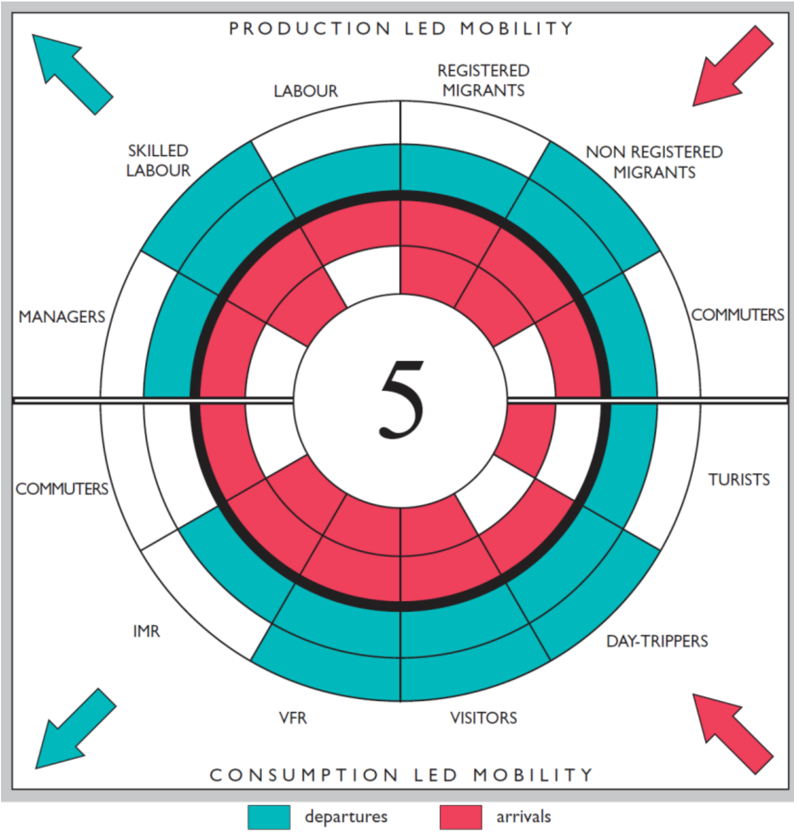
that they expect to benefit from regularisation and then family reunification. However, there can also be temporary flows: a short-term transfer, usually for seasonal work, with the objective of either returning home once they have accumulated sufficient funds or, more frequently, of drawing on a network of connections to undertake onward mobility to other areas or countries. By now, the area has improved the quality of its hotels, housing and services. It has become a benchmark for residents of neighbouring towns, and the direction of the commuting, which was mainly outgoing until the previous phase, is now reversed. The area also becomes a tourist attraction in its own right, particularly if it can provide a range of high quality natural and cultural assets. But even if these are lacking, an initial flow of VFR tourists begins to arrive (fig. 3.2.15).

Figure 3.2.15. Stage 4, the reinforcement of the project.



Stage 5 – restructuring and relaunching the project – hypothesises the overcoming of a period of stagnation and crisis due to local factors as well as an international recession. The recurring objectives of this phase are a policy of partnership, decentralising production, going from manufacturing to creative planning, introducing innovative processes and polices, reducing environmental impact, introducing policies for environmental restoration and cultural resources, enhancing the quality of life and introducing policies to help SMEs to develop. To achieve these objectives, innovative policies have been put in place to restructure production; these policies are beginning to take the characteristics of society and the territory into account. Moreover, pride in the territory is being promoted through a close collaboration between public administrations and private sector players (Cooke and Morgan, 1998). On the basis of a new report founded on primarily endogenous development, as defined by Garofoli (1992), the territory is promoted not only as a manufacturing capital but also, through tourism, as a consumer good. In the restructuring phase, there is a probable loss of jobs that will, however, mainly affect poorly qualified manpower. Workers of this type will presumably be expelled towards neighbouring areas. For the rest, there is large-scale replacement of middle management and managers, with flows that are therefore both incoming and outgoing. As a result of specific sector-wise policies, such as the improvement of the territory's infrastructure and image, national and international tourist flows are on the rise – not to be confused, however, with those registered in the previous stage – connected mainly to work activity and short or medium-term consultancy work. The area also affirms itself at the local level and therefore becomes a centre of attraction for commuter flows that come here to carry out a number of different activities. It is generally considered that widespread wellbeing encourages a consistent flow of tourists towards national as well as international locations. This flow can also be local, as a component of the commuting towards less well-equipped and built-up areas which, for this very reason, are more sought after for recreational activities (fig. 3.2.16).

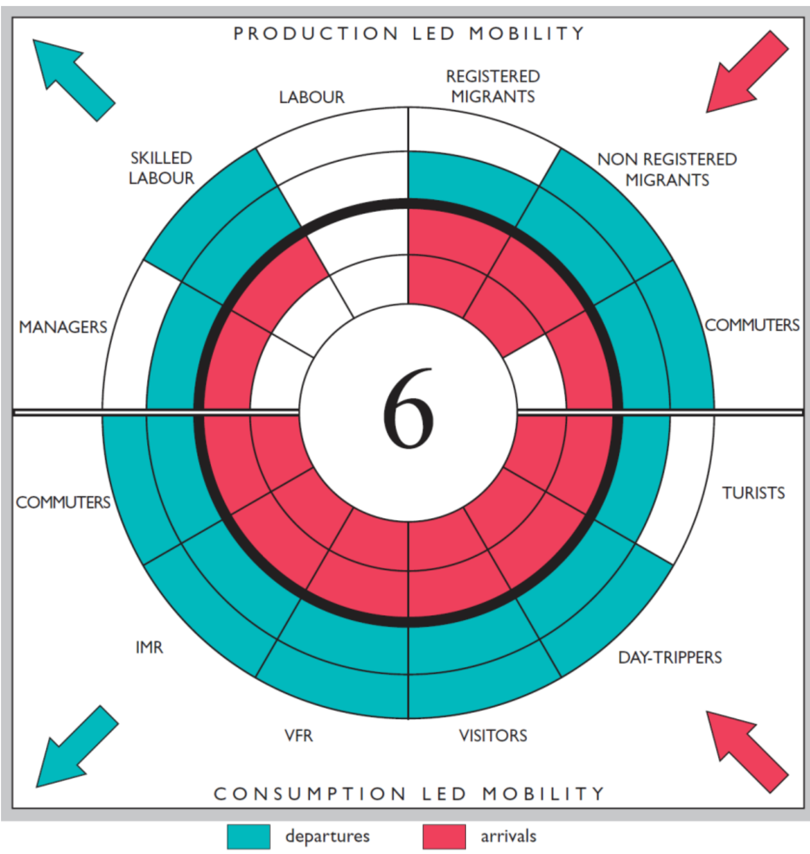
Figure 3.2.16. Stage 5, restructuring and relaunching of the project.



By **Stage 6** – technological innovation – the area is a well-established production locale, and is therefore dominated by high technology and professional service providers. There has been substantial investment in research and development but, to reach these levels, a wide-ranging process has been started to involve the main players and social groups through governance policies backed up by intensive information, training and continuing education initiatives. Overall quality is therefore not applied exclusively to companies, but also to the territory, according to the synopsis drawn up by Hunt (1995). In these conditions, the area is interesting and attractive in its own right for new companies, new activities and new flows of goods and services. In terms of mobility, the trend of temporary flows for the purposes of learning and exchanging information remains high, and extends quite naturally into longer-duration stays. The place's specific cultural offer is firmly in place, as are other forms of attraction that create of niche tourist flows. In this phase the flows increasingly resemble magma; their direction cannot be easily foreseen or understood, and their motivations are apparently irrational. Taylor

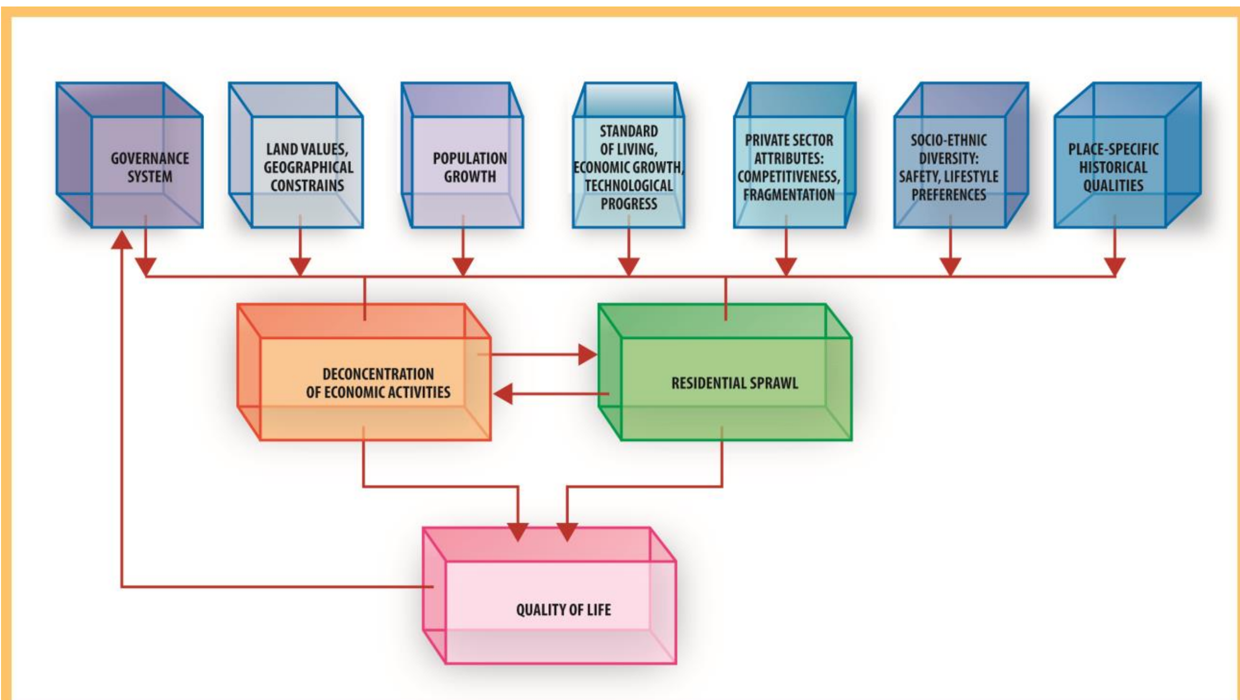
(1997) indicates that the premise for this phenomenon is the Americanization, modernization and globalization of society – hence, a phenomenon that should contribute to creating a totally homogenized world. In reality, this process – homogeneous and even ordinary in some cases at the global level – becomes differentiated, and develops peculiarities, at the local level. In this sense, global and local are not opposing factors; they form part of a single system in which the local attempts to gain an effective dialectical position for itself. Urry (2000) proposes the development of a "sociology of fluids" that contrasts with the network to create "*heterogeneous, uneven and unpredictable mobility*" (fig. 3.2.17).

Figure 3.2.17. Stage 6, technological innovation.



The European FP5 project, "Spatial deconcentration of economic land use and quality of life in European metropolitan areas", SELMA (2002-2006), is similar to CURB in many ways, although the similarities are not immediately obvious because of the time gap between the two projects, and events that appear to have changed the context completely. The first point concerns the fact that urban processes always seem to get underway in North America first. So researchers always pay attention to events in the U.S. to try and anticipate, and occasionally pre-empt, the problems that are sometimes an integral – though not inevitable – part of new urban phenomena. In CURB, decentralisation concerned the urban American population, and the American scholar of reference was B. Berry. SELMA, too, examined issues affecting the non-residential sprawl by analysing phenomena that had been previously observed in American cities (Ewing, 1997), which researchers had already attempted to quantify (Galster, Hanson, Ratcliffe, Wolman and Freihage, 2001) and which had already been analysed by non-American researchers (Razin and Rosentraub, 2000). SELMA was conceived at a time when environmental problems and the issue of the quality of life were beginning to be taken more seriously. So decentralisation was examined not only in the light of the economic costs of the process, but also in terms of the dignity and protection of citizens (fig. 3.2.18).

Figure 3.2.18. A framework for evaluating deconcentration of economic activities in metropolitan areas. Source: modified from Razin (2007).



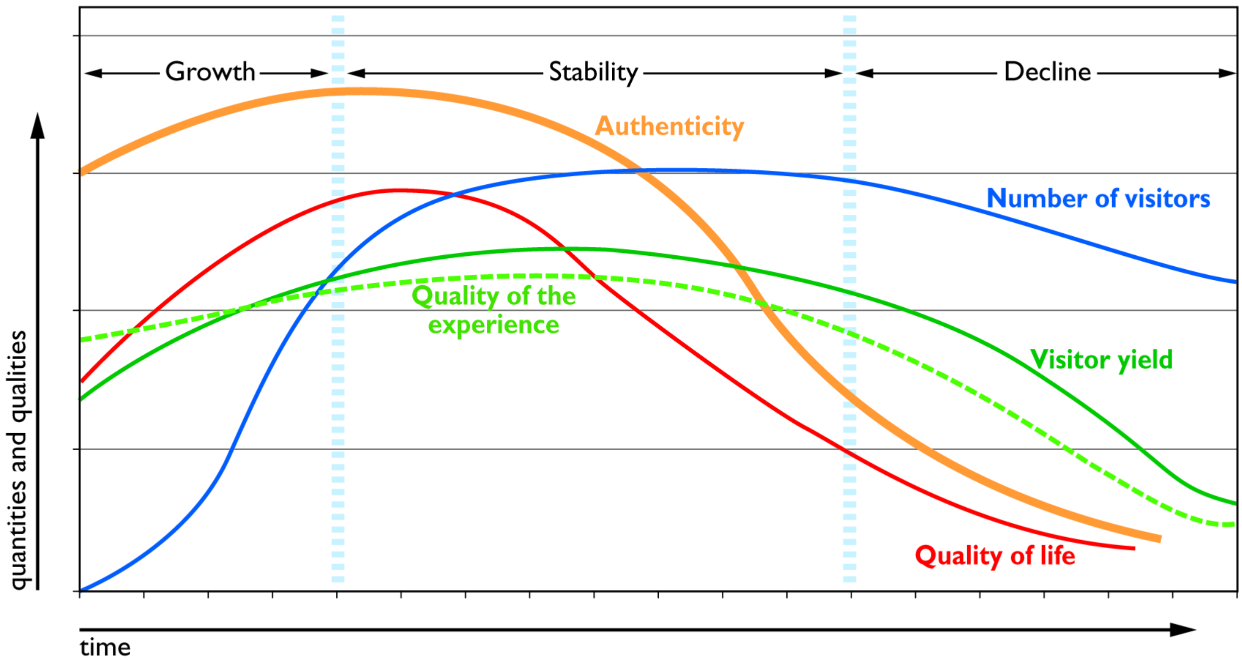
When the project was conceived around the year 2000, Europe was no longer clearly divided into countries with free market and planned economies. But in the first decade of the new millennium, what researchers wanted to check was whether the process of economic decentralisation was developing in the same way, and at the same pace, in countries that had had planned economies until only a few years previously, like the Czech Republic. The project studied fourteen metropolitan areas in the Czech Republic, Denmark, Israel, Italy, the Netherlands, Spain and the UK. A qualitative taxonomy was carried out on the four main attributes of decentralisation (Montanari and Staniscia, 2006): (i) magnitude; (ii) physical form; (iii) sectoral composition; (iv) government-policy context (tab. 3.2.2).

Table 3.2.2. Characteristics of deconcentration in the 14 SELMA metropolitan areas investigated. Source: rielaborated from Montanari and Staniscia (2006).

Urban areas	**Profile**	**Characteristics of deconcentration process**			
Valladolid	0001	Small-scale	concentrated	industry	Market-led
Copenhagen	0010	Small-scale	concentrated	services	Planning-led
Aarhus	0010	Small-scale	concentrated	services	Planning-led
Breda	0010	Small-scale	concentrated	services	Planning-led
Prague	0021	Small-scale	concentrated	retail	Market-led
Brno	0021	Small-scale	concentrated	retail	Market-led
Chieti-Pescara	0021	Small-scale	concentrated	retail	Market-led
Rome	0101	Small-scale	scattered	industry	Market-led
Beer-Sheva	1000	Large-scale	concentrated	industry	Planning-led
North Wing of Randstad	1010	Large-scale	concentrated	services	Planning-led
Tel Aviv	1021	Large-scale	concentrated	retail	Market-led
Bristol	1111		scattered	services	Market-led
Southampton	1111	Large-scale	scattered	services	Market-led
Madrid	1111	Large-scale	scattered	services	Market-led

The European Culture Project "Preserving places. Managing urban tourism, urban conservation and quality of life in historic centres", PLACE (2008-2010), took a fresh look at the theme of old districts in European countries to examine whether physically preserved historic districts are now suffering from excessive tourism consumption that makes it problematic to conserve their identity (Montanari, 2010a). The Rome case focused on three variables (Fig. 3.2.19): the number of tourists, the wellbeing of residents and the satisfaction of tourists (Montanari and Staniscia, 2010; Montanari, 2010b).

Figure 3.2.19. Visitors, absolute number vs. qualitative approach Source: modified from Plog (2001) in Montanari (2010).



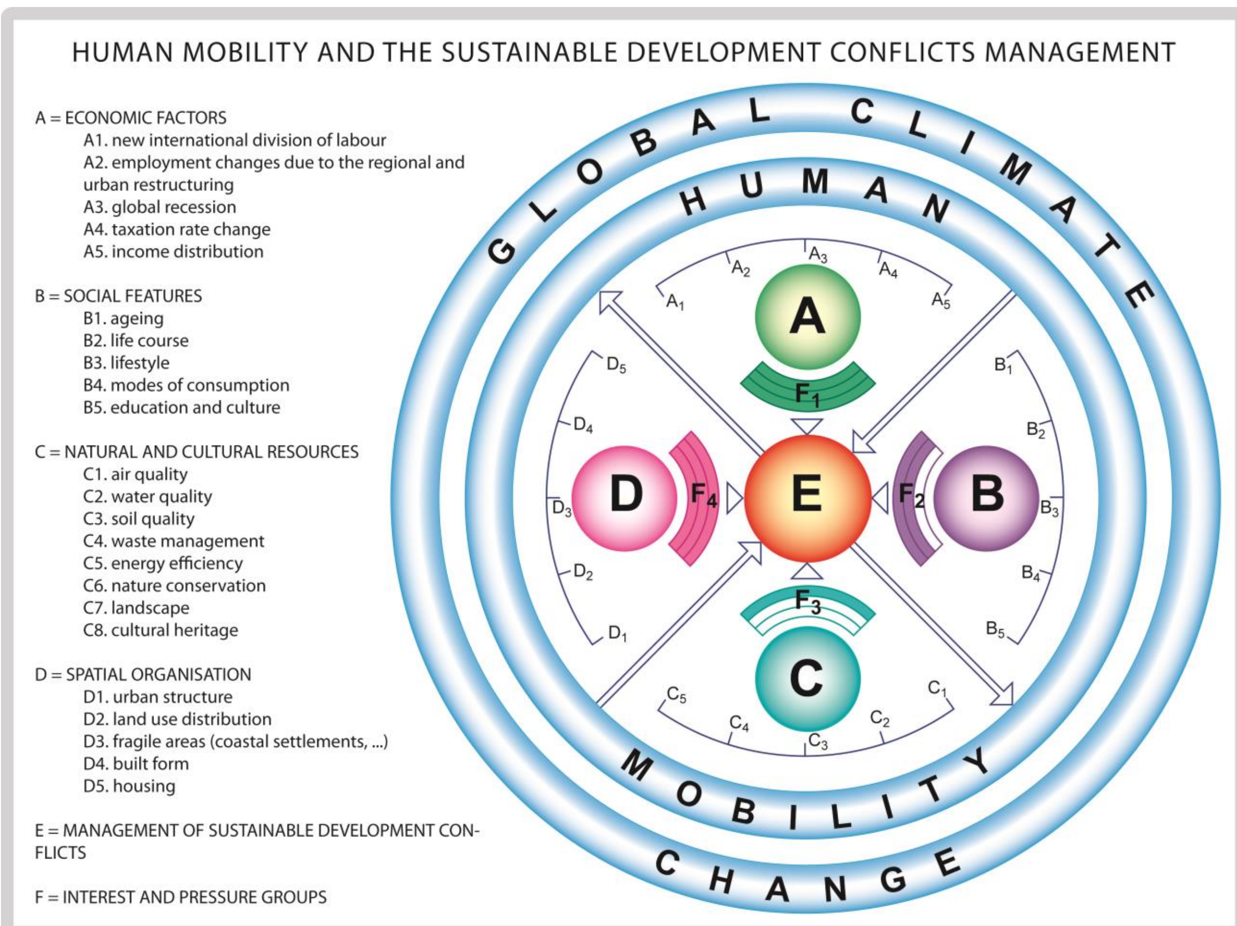
3. Cross-national Comparative Research: The SECOA Project. Data Analysis and Interpretation for Taxonomy and Modelling

3.1 SECOA: The basis for comparative analysis

Taxonomy and modelling (Work Package 6 in the SECOA Project) is the node through which the collected data were processed into information and knowledge for the later stages of the SECOA project. SECOA's data management methodology drew from the outset on the theories of Ackoff (1989) and other scholars who studied (Bellinger, Durval & Mills, 2004; Bernstein, 2009) and further developed his scientific work. These theories are also known as the DIKW (data, information, knowledge, wisdom) hierarchy. The initial programme elements were checked during the actual management of SECOA data, and adjusted as necessary on the basis of the data collected during the first phase of the project. The practice of data collection and processing in what is known as the "Internet age" has been expanded and elaborated within research on data mining and geographic knowledge (Giannotti & Pedreschi, 2007) and the subject of file sharing (Aigrain, 2012). Berry and Linoff (1997) define data mining as the process of exploring and analysing large quantities of data so as to identify "meaningful patterns and rules". The objective of the process is therefore to identify and make visible hidden associations contained within data. Alongside data mining, it is worth mentioning Business Intelligence (BI) processes: the term defines a series of techniques for collecting and analysing strategic information, the technology used to carry out these techniques, and the information collected as a result of using them. The philosophies used are data aggregation, data warehouses and the discovery of patterns within data, data mining. SISTER, one of the end users of the SECOA project, organised a seminar for project participants at La Sapienza University in Rome on 13 January 2011. On this occasion, SISTER presented the work it has carried out since the 1990s, setting up data warehouses for government bodies (public administrations) as well as the private sector (banks, insurance companies and supermarkets) and creating its own technology, such as the ST StatPortal® and Graph Data Browsing, which combine traditional BI tools and new uses such as analysis based respectively on the territory and graph theory. The similarity between data mining and statistical tools is based on the fact that both approaches focus on a data-based learning process, and function because the data can be converted into information. SECOA used both approaches, with two operational components working in a parallel, convergent and synergistic manner. As regards the specific case of this chapter, we will not go into general assessments of a comparative analysis of the advantages and disadvantages of each approach. Lovleen and Mehra (2008) believe that, generally speaking, for statisticians the term "data mining" has a pejorative meaning because the connotation of searching for data to fit preconceived ideas is more important than the value of the data itself. The work of these two scholars can be referred to with regard to considerations about the advantages and disadvantages of the two techniques, and how it is possible to reduce the gap between statisticians and data miners.

The problem of collecting comparable data in the social sciences began to be tackled in the 1970s, when UNESCO set up the ISSC (International Social Science Council) to examine this issue, among other things. The issue of the advantages and disadvantages of comparative analysis in relation to the quantity and quality of available data has been studied from the outset (Montanari, 2012a). The natural sciences have the same problems, but the methods used do not envisage total comparability so much as repeated control tests using the same procedures, and possibly the same techniques (Montanari, 2012b). The SECOA project took into account data originating from the experiences of both the natural and social sciences; as a result, data collection and processing was even more complex. Complexity is a characteristic feature of our society and to explain social phenomena, it is becoming increasingly necessary to analyse all available data, regardless of their provenance. Any existing phenomena have some relevance in the formation of reality. They must therefore be identified and measured using devices that subsequently make it possible to integrate them into the remaining body of traditionally known data. SECOA took two levels of phenomena into account: global phenomena such as climate change and human mobility, and local-level phenomena such as the emergence of conflicts, the dynamics of urban development and the natural, environmental, social and economic factors at the local level. The local level is dependent on the global one in that it is the positive or negative consequence of decisions and policies applied at a higher level. The local level is a concentration of consequences while contributing, through its errors and positive behaviours, to the global level (fig. 3.3.1).

Figure 3.3.1. The Secoa Project (local – global – local).



The SECOA project (2009-2013) looks at urban settlements that undertook a process of regional and urban reconstruction following the economic crisis of the 1970s to give themselves a new international image. Renewed economic success has brought fresh flows of human mobility – permanent, semi-permanent, temporary and daily – for reasons of consumption (leisure and tourism) and production (economic migration). Global competition between metropolitan areas has emphasised the importance of natural and cultural resources. SECOA evaluates the effects of human mobility on the growth of urban settlements and restructuring in coastal zones where: a) the environment is more fragile and the available space finite b) each phenomenon is more concentrated and c) the natural and cultural effects on the environment are more acute. The knowledge of such effects can greatly help government bodies and businesses – especially in the construction sector but also in tourism – to plan future growth. Awareness of the environmental status of the coast and the population's preferences of use are useful for the planning and development of new residences, shops and leisure facilities. The problems have increased because the climate changes that influence environmental parameters – such as the sea level – have caused an increased risk of flooding, the spread of pollution and the dislocation of a large number of inhabitants. Controlling and reducing such undesirable consequences leads to greater conflict between stakeholders. An integrated approach to the ecosystem that incorporates the social, economic and natural sciences is therefore essential to understand and resolve the complex problems and dynamics of coastal cities. The operational phase of the project aims to: 1) identify conflicts 2) analyse their quantitative and qualitative effects on the environment 3) create models to summarise the complexity of different social, economic and environmental systems and 4) tackle the priorities of each type of coastal city by drawing up a taxonomy of each. The SECOA project model is targeted at managing sustainable development conflicts and draws on the analysis of economic factors, social characteristics, natural and cultural resources and spatial organisation within the larger context of sustainable development objectives. These local problems are then considered in the context of two factors of globalisation typical of coastal cities: global climate change and human mobility. The values F1, F2, F3, F4 (Fig. 3.3.1) refer to the level of ability of each pillar of sustainability in a coastal urban area to influence to its own advantage the decisions regarding sustainable development on the basis of the interests expressed by each stakeholder. The conflicts and differences within each community influence the receptiveness of government bodies and the political power of the various pressure groups and interest groups.

An authoritative contribution to this debate by the OECD's Global Science Forum (2011) supports the theory that many cities are currently unable to manage the problems they create, and are therefore in search of better management systems. International problems (capital flows, human mobility, climate change) that need to be tackled at the local level move faster than the speed at which local government bodies are able to operate: "...modelling of urban systems is one of the research and support tools available for decision making, and provides understanding of complex interrelationships and interactions in the urban domain, acting as a guide to urban policy and practice..." (OECD, Global Science Forum, 2011, p.4).

Right from the planning stage, the SECOA project mooted the possibility of solving these problems by localising data in space and hence using Geographic Information Systems (GIS).

GIS, considered as "tools", are the result of progress that began in the 1960s:

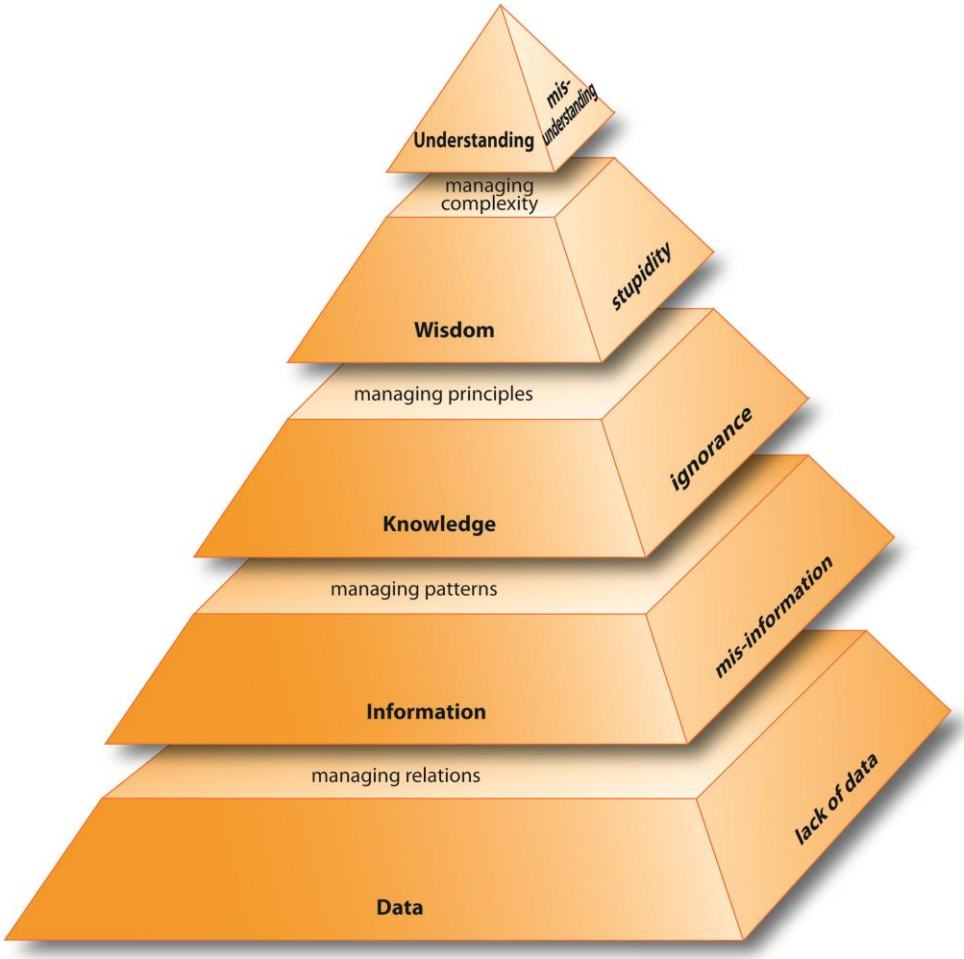
- The Geographic Information Science (GIScience) concept was introduced for the first time 20 years ago by Mr Goodchild, but the methods were developed long before GIS.
- GIScience studies the fundamental issues arising from geographic information.

GIS have evolved over the past 20 years from Systems to Science (Montanari, 2010b). GIScience can also be defined as a body of knowledge; data that GISystems are capable of managing, enhancing and using (Longley, Goodchild, Maguire, and Rhind, 2005). A georeferenced piece of data is characterised by its position in geographical space, and has a set of attributes describing its non-spatial dimension. GIS contributes to collecting, processing and visualising georeferenced data, thereby facilitating the researcher's work of interpretation.

3.2 The information chain

The SECOA project used the DIKW hierarchy theory to draw up its logical path, within which the meaning of each level of information and its opposite (e.g. existence of data and lack of data) could be made to converge. Since empirical reality often differs from theoretical reality, individual realities must be considered. The levels of information can therefore achieve the objectives of the individual SECOA Work Packages (WPs), and comprehension and interpretation can be used to go from one level to the next.

Figure 3.3.2. The La ACKOFF Pyramid rivisited Source: derived from Ackoff, 1989; Bellinger, Durval & Mills, 2004.



The pyramid (fig.3.3.2) was used in the SECOA project through the following phases. The data collected represented within SECOA by:

- WP1 – climate change and natural hazards.
- WP2 – environmental stresses and resources use for sustainable development.
- WP3 – Social, economic, cultural dimensions of urban development.
- WP4 – analysis of conflicts of uses of coastal resources among users and sectors are the raw material (tab. 3.3.1).

Table 3.3.1. SECOA WPs 1 – 4, topic, issue, dimension and data selection.

WP	topic	Issue	dimension	Data selection
1	Natural hazards	Climate change	Global-local	objective
2	Environmental stress	Sustainable development	Global-local Local-global	objective
3	Socio economic features and Human mobility	Urban development	Local global	objective
4	Coastal resources	Conflicts among users and sectors	Local global	subjective

For the specific feature of SECOA, the data measured both natural and environmental as well as socioeconomic phenomena, the objective being to interpret them. At this stage, however, the data have a meaning than does not transcend their existence. In view of subsequent stages, and on the basis of the information set out in SECOA project description, the data were collected in the form of characters, titles and numbers, and georeferenced. In WP4, case studies were selected using a subjective method. The selected conflicts had to involve at least one of the following themes:

- Economic development (industrial development, tourist industry, harbour restructuring, marina construction) vs. environmental protection (creation of protected areas);
- Preservation of natural sites and biodiversity;
- Conflicts over the use of resources between residents and newcomers as a result of human mobility (from economic migration to tourism).

The second level of the Ackoff pyramid shows the information process corresponding to the work of WP5 and WP6. The need to describe such heterogeneous phenomena in such different natural, social and economic situations was certainly a key presupposition in drawing up parameters intended to supply answers to questions such as "Who?", What?", "Where?" and "When?". This approach, in line with the rationale of SECOA, involved organising data using two synergistic approaches, global and local, so as to:

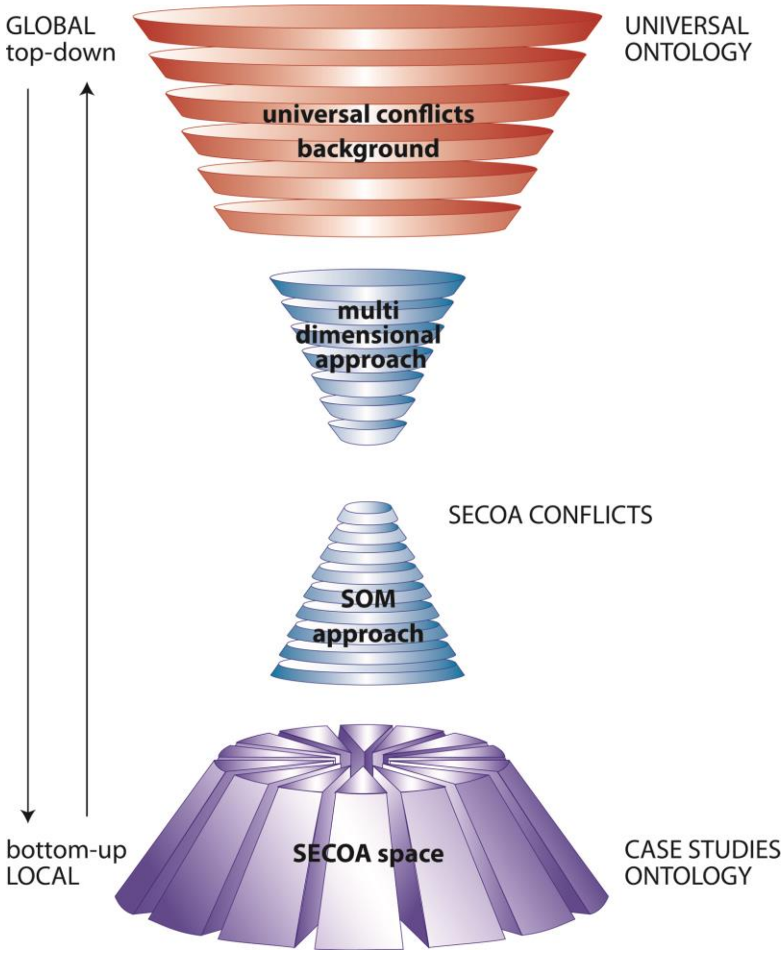
- Use all the data collected by national research groups and all the WPs, regardless of their comparability;
- The methodology involved analysis of the horizontal type for the global approach and of the longitudinal type along each case study for the local approach;
- The conflicts were therefore considered within the context of the variables of each case study as well as in the context of the global variables.

In the information the data, explained by WP5 (data integration) and ordered into clusters in WP 6, took on the form of a logical relation that can be interpreted in relation to their consultation, which can be simple or basic or even complex.

WP6 is the knowledge level at which SECOA acquired forms of knowledge through the taxonomy tool, which made it possible to appropriately catalogue information intended to be used at the subsequent stages. Knowledge presupposes the conscious use of data and information, and responds to the questions of "how" and "why" these are organised together. Wisdom (in WP7, methodology and tools for appropriate policies, of SECOA) is the stage of discerning and judging what can and cannot be done and what is right and wrong, based on the previous stages. The knowledge stage is typical of human beings and requires the use of forms of reasoning (based on the presence of the soul, heart and mind) that a machine can generally never have, for "a soul is something machines will never possess" (Bellinger, Durval & Mills, 2004). Understanding (in WP8, scenario building, of SECOA) completes the process with a stage that is simultaneously cognitive and analytical; this stage is an evolution of the previous level, knowledge. A community that has a correct understanding of environmental and socioeconomic phenomena is capable of taking correct initiatives because it is capable of synthesising new knowledge – which can also be based on information that may be new, since it is built on the further elaboration of previously noted and described information and phenomena.

The SECOA project uses its knowledge at both the global and local levels, considering their relation of cause and effect. This is due to the characteristics of coastal areas, which are an optimal point at which local and global phenomena intersect, and where their natural and socioeconomic cause and effect can therefore be observed. The taxonomy is the result of an approach reflecting both the global and local dimensions (fig. 3.3.3).

Figure 3.3.3. The SECOA taxonomy cones: global - local vs. local – global.



The core of taxonomy is an examination of conflicts that generally originate from global phenomena and are reflected at the local level, and vice versa. The data relating to and characteristics of these conflicts are the result of the work of WP4. The data were then processed and interpreted in relation to the need to describe global or local phenomena and causes (Khan, Le, Canters, Corijn, 2013; Khan, Le, Corijn, Canters, 2013) .

As regards this theme too, it is worth mentioning the work of Ackoff (2010) who, in a posthumous work on the concept of multiple meanings, refers to the binomial of ethics and morality, and that of conflict and competition. Identifying differences in meaning is essential, as it makes it possible to differentiate between concepts with multiple meanings, often to the point of being considered synonymous with each other. Ackoff's theory examines an individual who diminishes another's objectives in order to achieve his own, thereby creating a state of conflict. When, on the other hand, the first individual's objectives are in synergy with the others, the result is cooperation. The "zero-sum game" situation, in which the victories of one individual are equivalent to the defeats of others, is an obvious basis for a state of conflict. The conflict-cooperation equation can be explained relatively easily; the conflict-competition equation less so. Competition can be described as conflict in accordance with pre-defined rules. But the rules alone do not turn a conflict into a competition. A war between countries that do not obey any rules is an obvious state of conflict. But even when the Geneva Convention is respected, the local populations still suffer the consequences of a war, which is still a conflict. In theory, the maximum level of conflict is reached in terrorism, which by definition does not obey any rules. In Wp6s taxonomy "global" approach (Alleva, 2013), the conflicts examined within the SECOA case studies are compared with the results of a databank (ICOW) that examines a series of more wide-ranging conflicts in terms of reasons, time and space. The WP6s taxonomy "local" approach undertakes analysis and taxonomy of the conflicts studied by SECOA, as well as the analyses of metropolitan or urban areas studied by SECOA research where conflicts emerge. Conflicts are considered territorial because they emerge and take place in a spatially defined rather than an abstract area. By region we mean here a network of places. The place is a space of living, consuming and producing; it is the space of social relations and mobility. The place, so defined, is a social construction (Staniscia, 2003).

3.3. The application of GIS and SOM in different scientific fields

The results of the bibliographical research mentioned in the previous chapter (Viaggiu, ibidem) reveal that until the first decade of the 21st century, there was little work done on the use of the SOM method with GIS. Skupin and Agarwal (2008) attempt to justify the situation by recalling, that the SOM method was not developed by GIScience. In fact the main tests on GISystems, defined in literature as "tools" (Montanari, 2010b), tend to introduce variously heterogeneous approaches and therefore do not go into detail on any specific application. Skupin and Agarwal's book (2008) opens up a new phase in the use of the SOM technique and how it can interact with GIScience. SOM, especially in relation to the Principal Component Analysis (PCA) method, has often been seen as an instrument to transform and reduce variables. The objectives of European research projects (FP7) like SECOA include the need to communicate the results of fairly complex computational procedures to end-users, stakeholders and the general public, and it is therefore particularly useful to interpret SOM outputs using GIS.

Goodchild (2008), the academic known for his study of GIS and for introducing the concept of GIScience two decades ago (Goodchild, 1992), notes that the SOM algorithm describes objects that are not necessarily localized in geographical space, and does not make any reference to their location. GIScience, on the other hand, makes the basic assumption that all objects are georeferenced. This is why SOM and GIScience researchers have so far had few shared interests, and hence few contacts. Bação, Lobo and Painho (2008) acknowledge the need to transform large-scale georeferenced databases into "knowledge" tools. For this reason, many researchers have attempted to contribute to making an industrial product. But it has not yet been made, and therefore there is a need for continued research so as to try and adapt the SOM algorithm to the needs and functions of GIScience. In many scientific fields, the classification of objects is based on the similarity of attributes. In GIScience the concept of similarity takes into consideration both the "attribute space" parameter as well as geographical space, where the concept of similarity therefore takes on the value of proximity.

Bação, Lobo and Painho (2008) propose embedding the First Law of Geography (Tobler, 1970) into the SOM method to see if there is a possible balance between geographical proximity and the proximity of the values of attributes. The work of Tobler (1970), also known as the Tobler First Law (TFL), has been published in a number of geography and GIS manuals, and is based on the concept that "everything is related to everything else, but near things are more related than distant things". Some people have criticised TFL because it appears to be a simplistic interpretation of geography in a situation that seems rather more complex than can be explained merely by the relationships established between nearby objects. Specifically, it would seem impossible to explain complex ecological and socioeconomic systems merely by looking at the spatial relationships established between different entities. Notwithstanding these observations, TFL is a benchmark for many geography scholars and academics working in the social sciences in general, who consider it useful and satisfactory for their research. Generally speaking, TFL makes a distinction between geography and other disciplines, as it assigns specific and substantial relevance to the spatial component. Certainly TFL on its own cannot be considered to be an exhaustive concept, but on the basis of this principle it is possible to extrapolate, through the use of sophisticated techniques, systems of geospatial interpretation to better understand the present and anticipate future events. Longley, Goodchild, Maguire and Rhind (2005) maintain that it is impossible to imagine that two maps of the same area do not reveal some kind of similarity, even if they illustrate different phenomena. They therefore conclude that the affirmation, made on the basis of numerous statistical inference methods, that a spatial autocorrelation is equal to zero is contradictory to what seems obvious on the basis of TFL. The complex adaptive systems (CAS) theory suggests that the parameter of proximity could suffice.

This happens because interactions at the local level, although simple, can contribute to the formation of global behaviours, often rather more complex, that cannot be entirely foreseen or controlled. Montello, Fabrikant, Ruocco and Middleton (2003) introduce the concept of visualizing the semantic content of non-spatial information, such as library documents. The principle of spatialisation is based on the "distance-similarity" metaphor. On the basis of TFL, it is distance that determines similarities. The metaphor is based on analogous presuppositions to TFL, but it reaches the opposite conclusion: it is similarity that determines distance in spatialisation. This principle has been defined the First Law of Cognitive Geography (FLCG). The FLCG is based on empirical research in which participants evaluated the similarities between "documents depicted as points in spatialised displays". Further inspiration on the use of SOM in GIScience comes from the work of Tobler and Wineberg (1971), which refers to the use of cuneiform tablets to identify the location of ancient urban settlements in Cappadocia. The analysis was based on the number of tablets found in one settlement that were similar to those in other settlements. The authors expressed their conviction that mathematical models based on contemporary urban structures have a "temporally and geographically invariant" structure, and could therefore have been used in Cappadocia. The study did not make reference to the paths and itineraries that might have existed at the time, only to the distance between the settlements. So it was then possible to identify the relative position exclusively on the basis of distance, and thereby hypothesise the geographical coordinates of the settlements to be identified. Tobler (2004) went back to observations on and criticism of his work over the years and explained that in the case of Cappadocia, the spatial gravity model was turned upside down in order identify vanished settlements, and that spatial autocorrelation need not necessarily be positive, as can be seen from the effects of NIMBY in contemporary society.

Starting from the basic assumption that one of the objectives of spatial research is the construction of uniform, homogeneous regions, F. Bação, O. Bação Lobo and Painho (2004), consider the idea of drawing up a Geo-Self-Organizing Map (Geo-SOM). The Geo-SOM is a tool to identify homogeneous regions for which predictive analysis can be done using tools that make the visualisation of positive and negative correlations possible. Bação, Lobo and Painho (2008) warn that in many scientific fields, classification is aimed at grouping different entities based exclusively on the idea of similar attributes. In GIScience, classification is a compromise between similarity as "attribute science" and that which identifies the "proximity" of geographical space. One example of this theory is the application of Geo-SOM to the 2011 census data in Portugal, composed of 70 attributes for 250 counties, indexed by the number of inhabitants. Kropp and

Schellnhuber (2008), on the other hand, used SOM to examine the relationship between climate, soil type and the overall distribution of vegetation. In their case the SOM were used as a method to analyse a large-scale, multi-dimensional data base using an algorithm that supplies quantitative measurements for topological ordering (SOMTOP model). When the analysis subjects are subsequently georeferenced, it also becomes possible to draw up maps indicating the classes. A paper by Skupin and Hagelman (2005) uses a demographic data set with socio-economic variables for all 254 county units in Texas for 1980, 1990 and 2000. They propose a method to spatialise multi-temporal and multi-dimensional trajectories using SOM and GIS to represent them. The experiment supplied good results, although there is still the difficulty of integrating the related software, which the authors describe as a "loose coupling of SOM and GIS components". Fincke, Lobo and Bação (2008) explain the function of the SOM algorithm to present a number of frequently used visualisation techniques. They also demonstrate a method to import data processed with SOM into a GIS environment, so that operations intended for spatial analysis can be applied to data that are originally non-spatial. Yi-Chen Wang and Chen-Chieh Feng (2011) refer to research on around 700 articles published between 1987 and 2007 in seven specialised magazines on the subject of "land-use/cover change (LUCC)". The papers were analysed using SOM, and the results presented in visual form to contribute to a better understanding of LUCC research activity. Yan Li and Shanmuganathan (2007) apply a "social area analysis" model to Beppu City, a small Japanese town. The authors point out that this kind of analysis, which was quite frequent in earlier decades, has not been used much in recent times for two main reasons: (i) the excessive simplification of the complexity of residential urban structures, particularly through the use of linear correlation methods that have often supplied banal results; (ii) the substantial amount of work required to visualise the spatial characteristics of phenomena. The combined use of SOM and GIS made it possible to see how, in Japan, socio-spatial divisions mainly occur in relation to age, or rather different periods in the life of each individual. It emerges that Beppu City is divided into two macro-areas, each of which is further divided into 6 and 2 sub-areas, for a total of 8 social areas.

Finally, Goodchild (2008) believes that the SOM is not merely an ordinary complement to the spatial analytic toolbox, but a new paradigm for Exploratory Spatial Data Analysis (ESDA) and spatial data mining. Increasingly intensified use of SOM and the good results this method makes it possible to obtain poses significant questions for GIScience in general and its research agenda in particular. It is also a signal to the GIS software industry to consider the need to create a stand-alone toolbox rather than merely integrate a part of GIS functions.

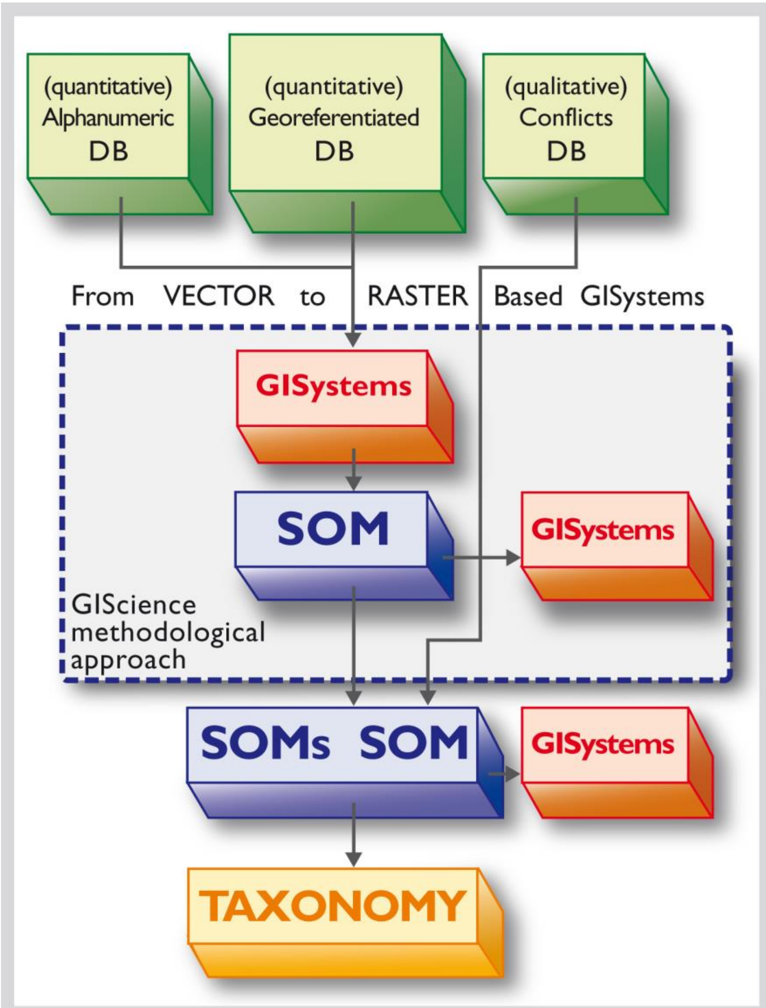
3.4. The taxonomy for the SECOA space and conflicts: methodology

The general introduction explains how, for better management of the data collected by WP1-WP2-WP3, the decision was made to operate at two parallel and synergistic levels: (i) the global approach (Alleva, 2013) and (ii) the local approach and hence the available data and methods used. As mentioned previously, SECOA DB is heterogeneous because of the nature itself of the data: environmental and socio-economic; qualitative and quantitative. The environmental data were collected during WP1 and WP2 (Lan, Olsson, Alpokay, 2013), and the socio-economic data during WP3. The environmental data use the same units of measurement worldwide (or can in any case be converted from one system to another), the same measuring systems, and an analogous method for collecting and cataloguing data (Montanari, 2012b). The socio-economic data measure the same phenomena using different collection techniques and tools. In the first case, there is no substantial difference between measurements taken within EU countries and countries outside the EU. In the case of the socio-economic data, there are differences between EU countries too. Often the same phenomena are identified in a different, even opposite, way because they represent different historical traditions and geographical features, and a different phase of economic progress (Montanari, 2012a). The data collected by WP4 were gathered directly by individual SECOA national teams on the basis of a common design drawn up by the WP4 leader, and this is why it has been described as being of the qualitative type. This type of data are heterogeneous but have a common point, their geographical location, either because they concern a specific territory or because, on the basis of Tobler's Law and its previously mentioned interpretations and applications, the georeferencing of each datum makes it possible to identify further interpretational analogies and synergies. The method used was therefore to line up geographical analogies using GIS, and logical analogies using SOM. The simultaneous use of GISystem and SOM is part of a characteristic GIScience methodological approach.

The urban areas to be considered as case studies had already been identified at the preparatory stage of SECOA, based on the principle that they must be coastal areas of two categories: (i) metropolitan areas of international/global importance; (ii) metropolitan areas of regional/national importance. The fact that the SECOA case study areas were located both within and outside the European Union undoubtedly enriched the meaning of the research, but made the data comparability aspect more complicated. The identification of conflicts was also qualitative, as the data were collected on the basis of SECOA's choice. The selected conflicts had to cover at least one of the following themes: (i) Economic development (industrial development, tourist industry, harbour restructuring, marina construction) vs. environmental protection (creation of protected areas); (ii) Preservation of natural sites and biodiversity; (iii) Conflicts over the use of resources between residents and newcomers as a result of human mobility. So, besides being located in a coastal area, the conflicts had to represent SECOA's need for comparability rather than the need to statistically illustrate the representative requirements of the entire metropolitan area in which they take place. This data comparison complexity had been foreseen on the basis of previous experiences, and therefore the data georeferencing and the GIS software to use had already been agreed on before SECOA went into operation. With regard to this approach, the need for processing on a local scale had also already been envisaged, with the following aims: (i) Using all the data collected during the preliminary stage of SECOA, regardless of whether they were perfectly comparable (ii) In any case the social and economic differences between the territories are based on values that do not make it possible to directly compare the data generated by the different case studies; (iii) Hence the comparison was not made between individual data from all the metropolitan areas studied, but between the metropolitan areas, each one interpreted according to a specific typology of their wholes. This is why, in Figure 3.3.3 (the SECOA taxonomy cones, global-local vs local-global) of the general introduction the cone at the bottom is shown with vertical divisions; (iv) The conflicts, too, were not considered in the context of the variables of each metropolitan area, but subjected to a preliminary evaluation with a specific SOM. Next, they were geographically localised, together with their problems, within each metropolitan area; (v) In this way, the conflicts were studied in the territory pertinent to them, in the place where they started and continue, hence in a concrete, not an abstract space; (vi) Each metropolitan area is therefore considered a network of places. The place is a space of living, consuming and producing; it is the space of social relations and mobility. The place, so defined, is a social construction. (Staniscia, 2003; 2006).

In the EU countries, it was possible to have the GIS database for all the urban territories; in the non-EU countries, this was not possible with single cell data. In the EU countries, considering the availability of geographical bases and information, the data were converted from vector to raster based GIS. The vector data represent geographical components as points, lines and polygons. The spatial data model defines space as a sequence of cells of the same size, organised into columns and lines and composed of single or multiple bands. Each cell contains an attribute value and the location coordinates. Unlike the vector structure, the raster coordinates are contained within the matrix. Groups of cells with the same values represent the same type of geographical characteristics. On the basis of tests carried out on the available data, the decision was made to choose squared cells, 1 km per side. Each metropolitan area was represented by a variable number of cells, from a few thousand to several thousand. On the basis of experience acquired during previous projects, it was deemed more convenient to use the raster data base to construct the SOMs.

Figure. 3.3.4. Logic model of the taxonomy: the local approach.



As shown in Fig. 3.3.4, the organisation of data into the GIS database will allow for faster, more substantial data transfer to a SOM. This will make it possible to organise and structure metropolitan areas according to a sequence of qualitative parameters grouped together in different ways. A further construction, a SOM of the SOMs, makes it possible to identify some "territory types" in situations concerning the metropolitan areas used as case studies., For reasons of safety, correctness, right of use etc., data for the non-EU countries on the other hand was only supplied grouped together in macro areas: core, ring, coastal area. The simplified processing of these data used the "territory types" mentioned earlier as a reference. The data on conflicts from WP4 were on the other hand processed with an independent SOM, and gave rise to the identification of "conflict types". As can be seen in Fig. 3.3.4, the SOM and the SOM of the SOMs were further processed using GIS to make them easier to understand for end users and stakeholders.

3.5. Modelling methodology: A few preliminary concepts

The search for the results of operational projects and general literature in this field started with the OECD Global Science Forum (GSF) on the Effective Modelling of Urban Systems to address the challenges of climate change and sustainability (OECD GSF, 2011), notably with the Melbourne workshop (30 January to 1 February 2011). SECOA's invited representative at the Melbourne workshop was José Manuel Simões, the co-ordinator of the Portuguese research team and chairman of SECOA WP8. The European Union DG RTD, Unit 1.4 (Environmental Technology and Pollution Prevention) was represented by Astrid Kaemena, also a member of the steering committee of the OECD GSF, and EU Commissioner for SECOA until end-2012. SECOA had not yet elaborated its considerations on modelling in early 2011, so it was unable to provide any specific contribution to the workshop apart from the competence of Simões. Its considerations were developed from 2011 onwards in WP6, WP7 and WP8. However, right from the start SECOA took a great interest in the conclusions of the OECD forum, particularly regarding the "local" approach detailed in this volume, and consulted literature and methodological approaches that were not part of the OECD GSF.

The OECD GSF also provided an opportunity to take a closer look at around thirty international research and exchange programmes. A further opportunity for information on the modelling used in some European research programmes and the work of a number of European Commission offices was provided by the "Meeting of co-ordinators: urban issues, soil biodiversity, desertification and land degradation" (Brussels, 22-23 October 2012) organised by the European Commission, at which the author of this paper was present.

While the initial hypothesis of the OECD GSF is not the same as SECOA's, a few elements could be the result of shared considerations on how to create and manage modelling. The first problem is data availability, as data are the key component of modelling. The main element of the OECD GSF's initial hypothesis was the unprecedented need to adapt research into urban settlements to the issues of global climate change and sustainable development. SECOA, with its bottom-up approach, did not choose a model and then apply the data; on the contrary, there was the problem of creating a method, a model, capable of interpreting what was available. On the basis of the research carried out by the OECD GSF, it emerges that the problem of data availability is still given little importance.

The problem of the lack of data is further complicated by the need to compare quite different phenomena, collected using the tools available to disciplines that are not used to working together. Many data are either unavailable, or not available in the desired form, or inaccessible for commercial, legal or governance-related reasons. Historically, the traditions of the various scientific disciplines have produced data in different forms. In the social sciences, these data are generally ready to be inserted into a model, but this is not always the case with other disciplines (OECD GSF, 2011). "Likewise, data from one given city, urban region or country may differ from others in its form, despite increasing expectations that urban performance can be compared across and between cities and urban regions" (OECD GSF, 2011, pag.9).

Based on the experience of collecting data and the problems SECOA encountered in this field, it was decided to use a differentiated approach for selecting data before inserting them into the modelling process. In the case of EU countries and Israel, thanks in no small part to the skills and availability of the researchers involved, it was possible to use quantitative data, even though the data came from heterogeneous sources and had been processed by researchers from different disciplinary cultures. The problem emerged when data from this group of countries were compared with data from India and Vietnam. To overcome the issue of the non-comparability of data, the researchers proceeded to work on qualitative data, rather than give up all the information collected. So the quantitative data from the EU countries and Israel were mechanically transformed into qualitative data. The data from India and Vietnam were integrated and transformed into qualitative data with the help of the key observers. The modelling referred to in this volume attempts to answer the question "what is?", interpretative modelling, leaving SECOA WP8 to define "possible alternative futures through scenarios based on "what if?", predictive modelling (OECD GSF, 2011, p.7).

New trends emerged in the early years of the 21st century because of improved calculation capacity, a bottom-up epistemological approach, and increasingly widespread use of the GIS instrument. As a result, the focus of urban modelling began to shift from macro to micro (Openshaw and Abrahart, 2000; Batty, 2005). The theory of Multi-Agent modelling was also taken into account; it is based (Diappi, 2004) on:

- Computing models (neural networks, cellular automata and multi-agent systems) deriving from artificial intelligence;
- A system capable of extracting rules from a database and reproducing macro-phenomena using a bottom-up process;
- Processing adapted to a description of the system on the micro-urban scale. The rules of individual behaviour are defined according to the way the background contexts interact.

The simulation of real processes carried out with the Multi-Agent Systems (MAS) entails building a complex system composed of individual entities that are at least partially independent and can interact with each other according to pre-established rules. In this way, according to Ferber (1995), these entities act together to integrate their individual ability. Diappi, Buscema and Ottanà (2004) point to the complexity of studying the interaction of social systems, economic systems and the environment. For this reason, they consider it opportune to use new methods of scientific analysis involving neural networks. While the traditional modelling approach is based on probabilistic rules, in neural networks the rules are factually defined on the basis of a learning process of distributed "unit processing" architecture (Diappi, Buscema, Ottanà, 1998).

Torrens (2010) points out that the social sciences have been obliged to use IT tools to explore the mechanisms that guide economic and social systems. Geography has the ability to connect the phenomena studied in the natural and social sciences. Advances in computational science have further enhanced geography's role, as many of the elements making up the complexity are geographical by nature.

Complexity is not a new phenomenon; it is inherent to all multidisciplinary, cross-national research (Montanari, 2012a). The situation has certainly become more complex since the introduction in the 1990s of the concept of sustainable development, which has obliged social scientists and natural scientists to compare their methods and findings (Montanari, 2012b). The question is, why model? It is considered necessary to examine the opportunity, and/or need, to interpret observations regarding human behaviour using a model. Also, what can one hope to achieve with a model? Van der Leeuw (2004) answers these questions by explaining that models make it possible for researchers to describe a wide range of relationships far more precisely than with words. He points out that models are particularly useful when it comes to multidisciplinary research. Each discipline has its own vocabulary and approach. The advantage of modelling is that it provides a form of expression that is acceptable to all the disciplines involved while being free from the connotations of any of them. "It is thus a major asset of models in that they can be used to express phenomena and ideas in ways that can be understood in the same rigorous manner by practitioners of different disciplines" (Van der Leeuw, 2004). Diappi, Bolchi and Buscema (2004) point out a further characteristic of neural networks – their ability to enhance our knowledge of urban planning by multiplying the information capacity of GIS and offering a new approach to territorial modelling.

The OECG GSF (2011) report identifies three categories of models that have been developed over time. These include a family of models for Land Use and Transport Interaction (LUTI), most notably UrbanSim, a "micro-simulation" type model that simulates land use. The model's findings are exported using a GIS, which makes them easy to understand and make available.

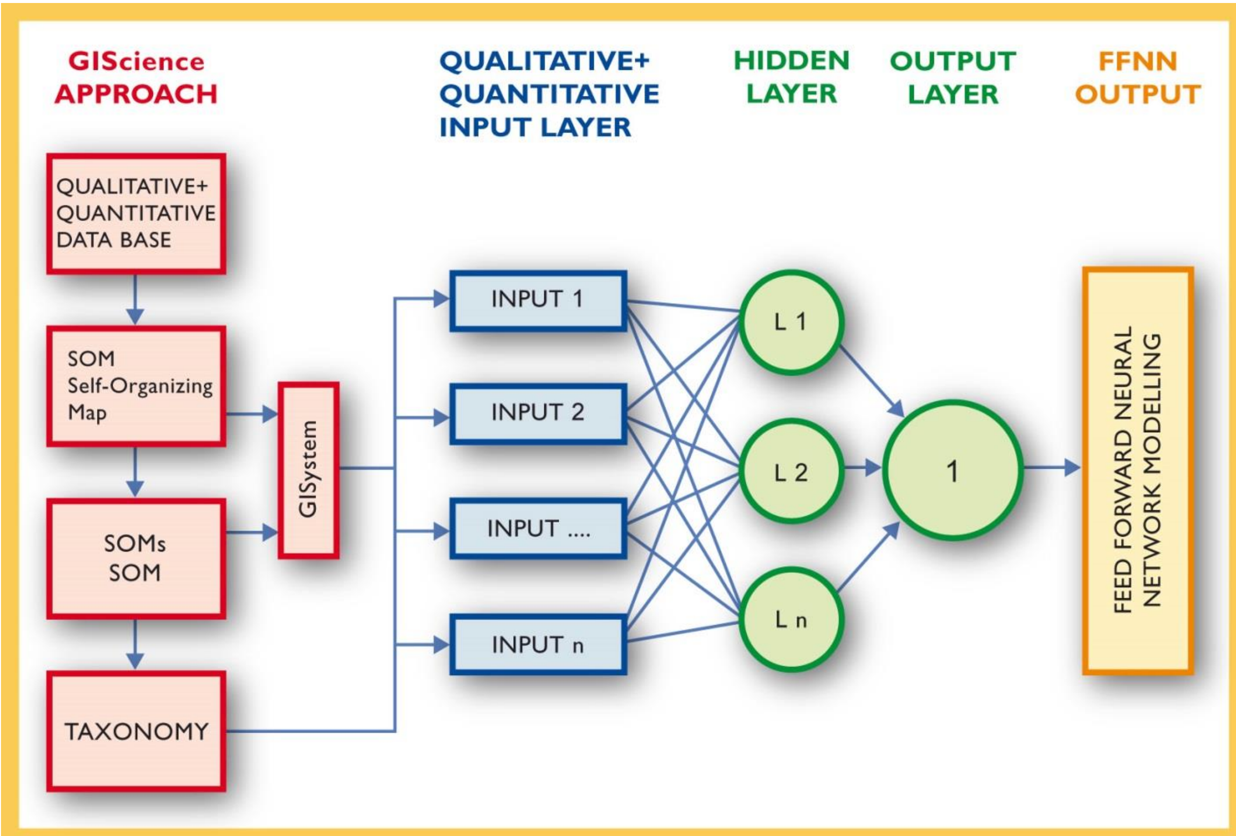
SECOA gave UrbanSim careful consideration for a number of reasons. Two SECOA research groups tested UrbanSim as part of the European project SELMA (FP5), with a specific research and operational project co-ordinated by the University of Jerusalem. Subsequently, a research group was set up to examine the various ways in which UrbanSim was adapted to the different situations of European countries. The findings were published in a special issue of the *Journal of Transport and Land Use* (Felsenstein, Axhausen and Waddell, 2010). The Israeli group published an article on developer behaviour and land prices (Felsenstein and Ashbel, 2010), and the Italian group published the findings of running an UrbanSim application in the city of Rome (Di Zio, Montanari, and Staniscia, 2010). It was decided not to use UrbanSim in this part of the SECOA project because the variables used by the model are different from the variables chosen by the project, and moreover one of the main characteristics of the model is that it is largely predictive. The possibility of using UrbanSim at least partially for WP8 is being considered, as well as the possibility of adapting it to the requirements of SECOA and integrating it with the neural networks approach of the GIScience, especially considering the advantage of exporting the results with a GIS. The possibility of using GIS to export data processed using necessarily complex tools is particularly useful in order to connect the "model maker" (generally an expert) and a "model user" (usually not an expert). The OECD GSF (2011) report considers this a significant connection, particularly because of the need to involve stakeholders in cognitive processes and, increasingly, in mitigation policies. Most research focuses on research methodology, leaving little room for studying interaction between model makers and model users. "While the interplay between socio-cultural practices, and mitigation and adaptation, in response to climate change remains poorly understood, emerging agent-based models (including those developed as 'serious games') present possibilities for informed decision making (OECD GSF, 2011, p.9-10)". In this phase of the research, SECOA has drawn up forms of communication represented by the "fingerprints" of each of the case studies presented later on in this volume. The term "fingerprints" to describe the individual characteristics of each case study was thought to be better suited to SECOA's requirements than "codebooks", the earlier definition used by Diappi, Buscema, Ottanà (1998, p.41). These instruments also serve to provide continuity with WP7 and WP8, which will conclude with an SDSS, Spatial Decision Support System, to be considered as a possible example of the "serious game". The "fingerprints" provide a brief description of each case study – the physical and socioeconomic context in which an environmental conflict has emerged. These "fingerprints" interpret the situation using a descriptive method that collects both spatial and temporal variables. The time variables will be processed further to make it possible to construct predictive scenarios and models.

3.6. The method used

The method used was to proceed with continuity of data and processing to switch from the process of data classification (taxonomy) to the narration system (paradigm) of the relationships in data (descriptive modelling). The first tool was a preliminary to the second, and both tools were a preliminary to the construction of scenarios and predictive modelling. Knowledge discovery in databases (KDD) was used to discover significant regularities, features and concepts in the data collected. Descriptive modelling was the mathematical process that made it possible to provide a concise description of the conflicts in coastal urban areas in relation to the social, economic and cultural factors that cause them. It was used to contribute to the procedure of identifying scenarios. The taxonomy took quantitative and/or qualitative data and processed them with GIS. The GIS data were made geographically homogeneous by transforming both vector and raster data to a raster dataset on a 1kmx1km grid, with a characteristic GIScience procedure in which the SOM, Self Organised Map, was integrated with the GIS to be subsequently classified in the taxonomy (fig.3.3.4).

Figure 3.3.5 confirms the linear process from the input layers to the outputs via the hidden layers, and the close connection between the taxonomy process, the results obtained in this phase, and the production of modelling through the FFNN.

Figure 3.3.5.The linear process of the Feed Forward Neural Networking Modelling.



For each case study – hence for each urban area taken into consideration – a hierarchy of involvement of each geographical area in the problems of each conflict was drawn up. The graph on the left in Figure 3.3.6 shows three geographical areas:

- the conflict's direct area (100%);
- the conflict's closest area (20%);
- the conflict's metropolitan area (5%).

For each area, the importance of the variables in understanding the conflict was defined, in percentage terms. For each of the conflicts in the 17 case studies, a FFNN was built with 52 output layers, 30 hidden layers and finally output, i.e. modelling.

Figure 3.3.6. The Feed Forward Neural Network Modelling applied to a SECOA conflict. The case of Civitavecchia, Rome Metropolitan Area.

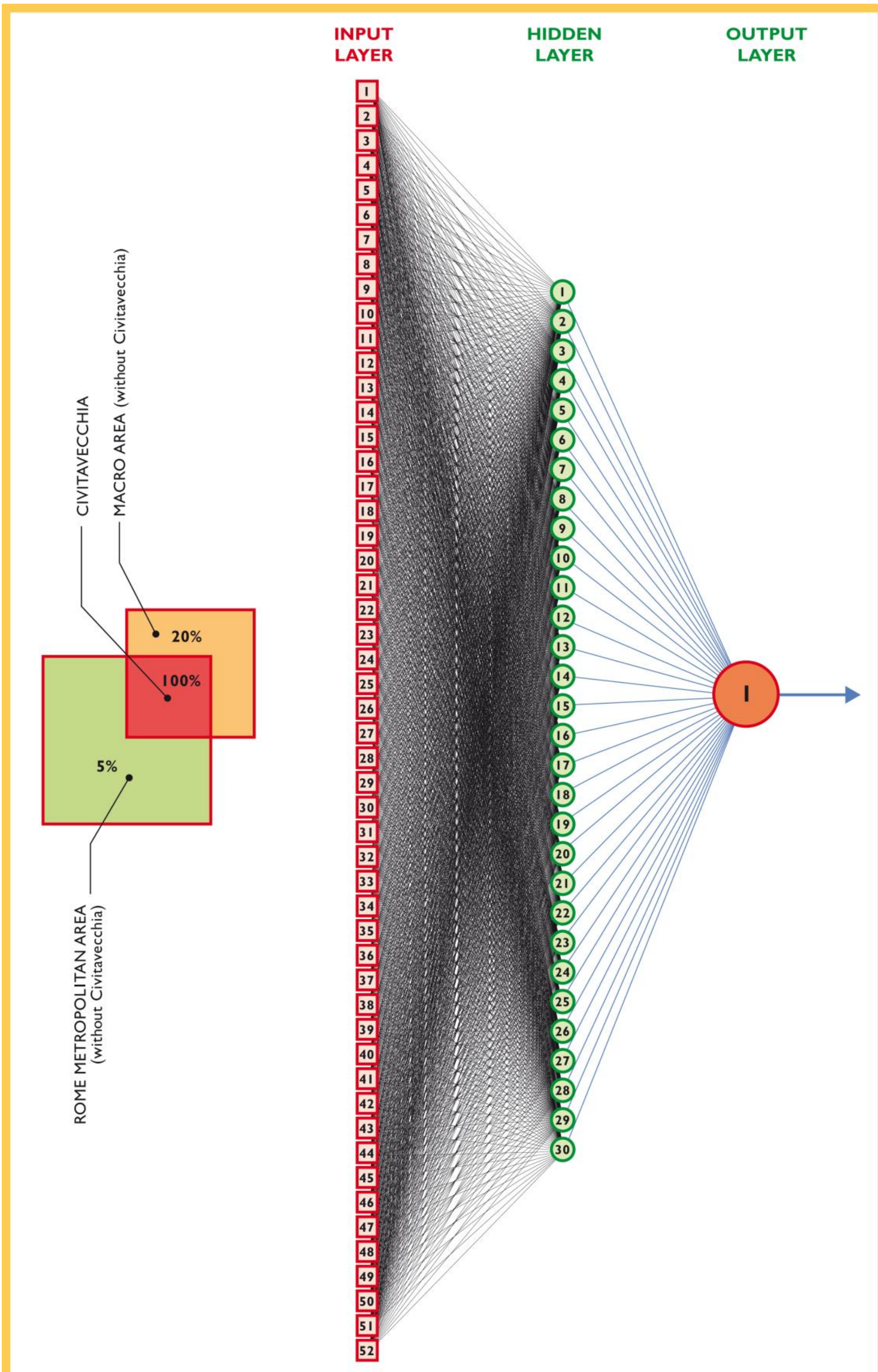


Figure 3.3.6 provides an example: the Civitavecchia conflict in the Rome metropolitan area. The backdrop of the illustration shows how the variables interacted to indicate each model of interpretation.

Figure 3.3.7. The output layer, the models.

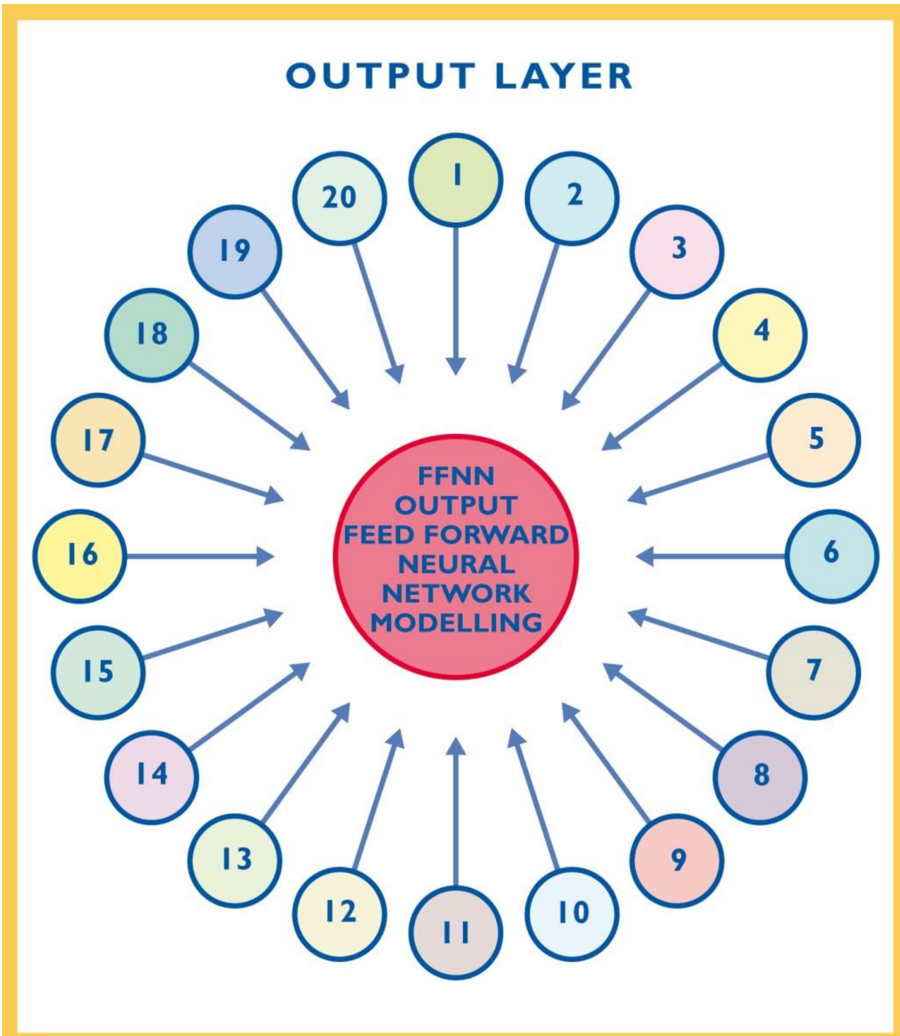


Figure 3.3.7 shows an example of the 20 output layers that, when juxtaposed, comprise the modelling, hence the sequence of all the fingerprints discussed in the following chapters. The fingerprints identify the way in which the variables tend to move over time in relation to the subsets of problems represented. Hence:

1. category of uses (n. 4);
2. economic development vs. environmental protection (n. 10);

3. preservation of natural sites and biodiversity (n. 4);
4. human mobility and contrasts for uses of resources (n. 4);
5. dynamic/manifestation over time (n. 4);
6. underlying causes/substances (n. 5);
7. scale (n. 5);
8. stage (n. 5);
9. final ranking (n. 3).

A check was carried out to evaluate the accuracy of the system's interpretive capacity (Figure 3.2.7), and one case study was subtracted. The neural network was started up again without the subtracted case study, and then the subtracted case study was tested with a new norm, the interpretive model. The network's interpretive capacity was judged to be sufficiently accurate, as the following chapters will show. Working on the fingerprints of each case study, the researchers were able to evaluate the system's capacity to modify its components by altering the value of the variables. The correlation between the variables and each conflict will be used to construct scenarios, and contribute to building an SDSS (Spatial Decision Support System). This tool will make it possible to resolve semi-structured spatial problems. It will comprise a Decision Support System linked to the Geographic Information System so as to simulate dynamic spatial processes and evaluate the effects of different future scenarios in advance.

3.7. Conclusions

Cross-national and multidisciplinary research projects are characterised by a large quantity of heterogeneous data, whether because of the nature of phenomena or the way in which these phenomena are measured and evaluated in the cultural systems of participating countries. This chapter has explained how comparative research has become increasingly complex over the years. To tackle environmental problems, the competence of the disciplines involved must be enhanced. Also, the increasing globalisation of the phenomena being studied presupposes the involvement of quite different countries. The complexity must be tackled using suitable tools. The tools used in this phase of the research have made it possible to grasp the relation between the variables of sustainability (environmental, sociological and economic) and the qualitative characteristics of the conflicts. We have also been able to gauge the influence of these variables on the conflicts and the intensity of the variables, and identify which of the variables are most important in the emergence of conflicts.

4. References

Ackoff, R. L., 1989, From data to wisdom. *Journal of Applied Systems Analysis* 15, pp. 3-9.

Ackoff, R.L., forward by Charles Handy, 2010, Differences that make a difference: an annotated glossary of distinction important in management, Devon, UK: Triarchy Press.

Aigrain P. , 2012, Sharing. Culture and the economy in the Internet age. Amsterdam: Amsterdam University Press

Alleva G. (ed.) 2013, *Taxonomy and modelling of conflicts in SECOA case studies,* Roma: Sapienza Università Editrice, forthcoming.

Bação F., Bação O., Lobo V., Painho M., 2004, Geo-Self-Organizing Map (Geo-SOM) for Building and Exploring Homogeneous Regions, in M.Egenhofer, C. Freska, H.Miller (eds.), Lecture notes in computer science, Berlin: Springer-Verlag, pp.22-37.

Bação F., Lobo V., Painho M., 2008, Application of different Self-Organising Map variants to Geographic Information Science problems, Self-Organising Maps. Application in Geographic Information Science, in P.Agarwal, A.Skupin (eds.), Self-Organising Maps. Applications in Geographic Information Science, Chichester: John Wiley & Sons, pag.21-44.

Bambra C., 2007, Going beyond the three worlds of welfare capitalism: regime theory and public health research, *Journal of epidemiology and community health*, 61 (12), pp.1098-1102.

Barlett D.J., 2000, Working on the frontiers of science: Apllyng GIS to the coastal zone, D.Wright and D.Barlett (eds.), Marine and coastal Geographical Information Systems, Philadelphia, PA: Taylor & Francis, pp. 11 – 24.

Batty M. (2005), Approaches to modeling in GIS: spatial representation and temporal dynamics, in D.I. Maguire, M. Batty and M.F. Goodchild (eds), GIS, spatial analysis and modelling, Redlands CA: ESRI Press, pp. 41 - 61

Bellinger G., Castro D., Mills A., 2004, Data, information, knowledge & wisdom, http://www.systems-thinking.org/dikw/dikw.htm

Berg v.d. L., Drewett J.R., Klaassen L.H., Rossi A. and Vijverberg C., 1982, *Urban Europe: a study of growth and decline*, Oxford: Pergamon Press.

Bernstein J., 2009, The data-information-knowledge-wisdom hierarchy and its antithesis, *Journal of information science,* 35(2), 68-75.

Berry B. J. L., 1973,The human consequences of urbanisation, Basingstoke: Macmillan.

Berry B. J. L. (ed.), 1976, Urbanisation and counterurbanisati*on*, London: Sage Publications.

Berry M.G., Gordon S.Linoff, 1997, Data mining techniques: for marketing, sales, and customer support, New York: John Wiley & Sons.

Bodei R., 2008, Paesaggi sublimi. Gli uomini davanti alla natura selvaggia (Sublime landscapes. The men in front of the wilderness), Milano: Bompiani.

Bradshaw J. And Hatland A., 2006, Introduction, in J. Bradshaw and A. Hatland (eds.) Social policy, employment and family change in comparative perspective, Northampton, MA: Edward Elgar, pp.1-9.

Brockman J (ed.), 1995, The third culture: beyond the scientific revolution, New York: Simon and Schuster.

Caniggia G., 1963, Lettura di una città*: Como,* Roma: Centro Studi di Storia Urbanistica.

Caniggia G., 1976, Strutture dello spazio antropico. Studi e note, Firenze: Uniedit.

Collini S., 2009, Science and art: still two cultures divided? *New Scientist*, 06 May

Commoner B., 1971, *The Closing Circle: Nature, Man, and Technology*. New York: Knopf.

Conzen M.P. (ed.), *2004, Thinking about urban form. Papers on urban morphology, 1932-1998. M.R.G.Conzen*, Bern: Peter Lang.

Conzen M.R.G., 1960, Alnwick, Northumberland: a study in town-plan analysis, London: Institute of British Geographers Publication n. 27.

Di Zio, S., Montanari A., and Staniscia B., 2010, Simulation of urban development in the City of Rome: Framework, methodology, and problem solving. *The Journal of Transport and Land Use,* 3(2):85–105.

Diappi L., Bolchi P. and Buscema M. , 2004, Improving understanding of urban sprawl using neural networks, in J.P. van Leeuwen and H.J.P. Timmermans, Recent advantages in design and decision support systems in architecture and urban planning, Dordrecht: Kluwer, pp. 33-52.

Diappi L., Buscema M. and Ottanà M., 1998, A neural network investigation of the crucia facets of urban sustainability, *Substance Use and Misuse*, 33(3), pp.783-817.

Diappi L., Buscema M., and Ottanà M., 2004, Complexity in sustainability: an investigation of the Italian Urban System through Self-Reflexive Neural Network, in L.Diappi (ed.) Evolving Cities. Geocomputationin territorial planning, Aldershot: Ashgate, pp. 46-84.

ElderJ.W., 1976, Comparative cross-national methodology, *Annual Review of Sociology*, 2, pp. 209 – 230.

Enyedi G., 1992, Urbanisation in East Central Europe: social processes and societal responses in the State Socialist Systems, *Urban Studies*, 29(6), pp. 869-880.

Collini S., 2009, Science and art: still two cultures divided?, *New Scientist,* issue 2706, 06 May 2009, pp.26-27.

Esping-Andersen G., 1990, The three worlds of welfare capitalism, London:Polity.

Ewing R. ,1997, Is Los Angeles –style sprawl desiderable? *Journal of American Planners Association*, 63, pp. 107-126.

Felsenstein D. and Ashbel E. , 2010, Simultaneous modeling of developer behavior and land prices in UrbanSim, *the Journal of Transport and Land Use*, 3(2), pp. 117-127.

Felsenstein D., Axhausen K. and Waddell P., 2010, Land use-transportation modeling with UrbanSim: Experiences and progress, *the Journal of Transport and Land Use*, 3(2), pp.1-3.

Ferber J. 1995, Les Systèmes Multi-Agents: Vers une intelligence collective, Paris: InterEditions.

Fielding A. , 1993, Mass migration and economic restructuring in R. King, Mass migration in Europe. The legacy and the future, London: Belhaven Press, pp.7 – 18.

Fincke T., Lobo V., Bação F., 2008, Visualising Self-Organizing Maps with GIS. Proceedings of GI Days, Interoperability and spatial processing in GI applications, University of Münster, 16-17 June.

Frey F.W., 1970, Cross-cultural survey research in political science, in R.T. Holt and J.E. Turner (eds.) The methodology of comparative research, New York: Free Press, pp. 173-294.

Galster G., Hanson R., Ratcliffe MR, Wolman H, Coleman S., Freihage J (2001) Wrestling sprawl to the ground: defining and measuring an elusive concept, *Housing Policy Debate,* 12: 681-717

Galtung J., 1967, Theory and methods of social research, New York: Columbia University Press.

Galtung J., 1982, On the meaning of "Nation as a variable", in M.Niessen and J.Peschar, International comparative research. Problems of theory, methodology and organisation in Eastern and Western Europe, Oxford: Pergamon Press.

Giannotti F., Pedreschi D. ,2007, Mobility, data mining and privacy. Geographic knowledge discovery, Berlin: Springer.

Goodchild M.F., 1992, Geographical information science. *International Journal of Geographical Information Systems* 6(1): 31-45.

Goodchild M.F., 2008, Epilogue: intelligent systems for GIScience, where next? in P. Agarwal , A. Skupin A. (eds.), 2008, Self-Organising Maps. Applications in Geographic Information Science, Chichester: John Wiley & Sons, pp. 196-198.

Gusdorf, G., 1977. Past, present and future in interdisciplinary research. *International Social Science Journal*, 29, pp. 580 - 600.

Hantrais L., 2005, Combining methods: a key to understanding complexity in European societies?, *European Societies*, 7 (3), pp.399-421

Hegel G.W.F., 1975 (originally published in 1837), Lectures on the philosophy of world history, Cambridge: Cambridge University Press.

Hegel G.W.F., 1991 (originally published in 1833) , Elements of the Philosophy of Right or Nature Law and Political Science in outline, Cambridge: Cambridge University Press.

Hillier B., 1999, Space is the Machine: A Configurational Theory of Architecture, Cambridge: Cambridge University Press.

Jinnai H.,1995, Tokyo. A spatial anthropology, Berkeley: University of California Press.

Kagan S., 2009, Kagan cooperative learning, San Clemente: Kagan Publishing.

Khan A.Z., Le X.Q., Canters F., Corijn E. , Canters F. (eds.), 2013, *Sustainability in Coastal Urban Environment: Assessing conflicts of uses,* Roma: Sapienza Università Editrice.

Khan A.Z., Le X.Q., Canters F., Corijn E.(eds.), 2013, *Sustainability in the coastal urban environment. Thematic profiles of resources and their users*, Roma: Sapienza Università Editrice,

Kropp J.P., Schellnhuber H.J., 2008, Prototyping broad-scale climate and ecosystem classes by means of self-organizing maps, in P. Agarwal , A. Skupin A. (eds.), 2008, Self-Organising Maps. Applications in Geographic Information Science, Chichester: John Wiley & Sons, pp. 155-75.

Leeuw v.d. ,2004, Why Model?, Cybernetics and Systems, *An International journal*, 35, pp. 117-128.

Lefebvre H., 1991 (originally published in French 1974), The production of space, Oxford: Blackwell Publishing.

Lisle, E. ,1985, Validation in the social sciences by international comparison, *Cross-National Research Papers*, 1(1): 11–28.

Longhorn R.A., 2005, Coastal Spatial Data Infrastructure, in D. Bartlett and J. Smith (eds.), GIS for Coastal Zone Management, Boca Raton, Florida: CRC Press, , pp. 1-15.

Longley P.A., Goodchild M.F., Maguire D.J. , Rhind D.W., 2005, Geographic Information Systems and Science, Chichester: John Wiley and Sons, Chichester.

Lovleen K.G. and Mehra R. , 2008, The lure of statistics in data mining, *Journal of Statistics Education*, 16,1, pp.1-8.

Meadows D, Meadows D.H. and Randers J.,1972, *The limits to growth*, Universe Book, New York.

Montanari A., 1990, Review of the volume "Changing face of cities" by J.W.R.Whitehand, *Town Planning Review*, 61(2), p.210-211.

Montanari A., 2009, Ecoturismo (Ecotourism). Principi, metodi e pratiche, Milano: B.Mondadori.

Montanari A., 2010a, Urban tourism in the context of the global economic crisis: the city of Rome, *Rivista di scienze del turismo*, 2 (1), pp. 381 – 399

Montanari A., 2010b, Turismo e sistemi di informazione geografica (Tourism and GIS), Mlano: B.Mondadori.

Montanari A., 2012a, Cross-national co-operation and human mobility: an introduction, *International Review of Sociology: Revue Internationale de Sociologie*, 22(2),pp.175-190.

Montanari A., 2012b, Social sciences and comparative research in Europe: cross-national and multi-disciplinary projects for urban development. The role of geography, *BELGEO*, 1, pp. 1-9

Montanari A., Curdes G. and Forsyth L. (eds.), 1993, Urban landscape dynamics. A multi-level innovation process, Avebuty: Aldershot.

Montanari A. and Staniscia B. (2006), Types of economic deconcentration in European urban space. Magnitude, physical form, sectoral composition and governance context, *Die Erde,* 137(1-2): 135-153.

Montanari A. e Staniscia B., 2010, Rome: a difficult path between tourist pressure and sustainable development, *Rivista di Scienze del Turismo,* 2/2010, pp.301-316.

Montello D.R., Fabrikant S.I., Ruocco M, Middleton R.S., 2003, The distance-similarity metaphor in network-display spatializations, Testing the first law of cognitive geography on point-display spatialization, in W.Kuhn, M.Worboys, S. Timpf (eds.) Spatial information theory: foundations of Geographic Information Science, Berlin: Springher-Verlag, pp.316-31.

Muratori S., 1960, Studi per una operante storia urbana di Venezia, Roma: Istituto Poligrafico dello Stato.

Nature Editorials, 2009, Doing good, 50 years on, *Nature,* vol 459, 7 may 2009, p.10

Niessen M. and Peschar J. (eds.), 1982, International Comparative Research. Problems of theory, methodology and organisation in Eastern and Western Europe, Oxford: Pergamon Press.

OECD GSF, Global Science Forum, 2011, Effective modelling of urban systems to address the challenges of climate change and sustainability, Paris: OECD.

Openshaw S. and Abrahart R.J. , 2000, Geocomputation, London: Taylor & Francis.

Ornauer H., Viberg H., Sicinski A., Galtung J. (eds.), 1976, Images of the world in the year 2000, Mouton: The Hague

Peschar J., 1982, Quantitative aspects in cross-national comparative research: problems and issues, in M. Niessen and J. Peschar J. (eds.), International Comparative Research. Problems of theory, methodology and organisation in Eastern and Western Europe, Oxford: Pergamon Press, pp. 57-81.

Peccei A., Ikeda D. and Cage R.L., 1985, Before it is too late, New York: Kodansha America.

Plog S.C., 2001, Why destination area rise and fall in popularity: un update of Cornell Quarterly Classic, *Cornell hotel and restaurant administration quarterly*, 42 (13), pp.13-24.

Razin E. (2007), Introduction: deconcentration of economic activities within metropolitan regions: a qualitative framework for cross-national comparison in E.Razin, M. Dijst, C. Vázquez (eds.) Employment deconcentration in European metropolitan areas. Market forces versus planning regulations, Dordrecth: Springer, pp.1-27.

Razin E., Rosentraub M., 2000, Are fragmentation and sprawl interlinked? North American evidence, *Urban Affair Review*, 35: 864-879

Sima Y. And Zhang D., 2009, "Comparative precedents on the study of urban morphology" in D. Koch, L.Marcus and J. Steen (eds), Proceedings of the 7th International Space Syntax Symposium, Stockholm: KTH, pp. 103.1-103.8.

Skupin A., Agarwal P., 2008, Introduction: what is a Self-Organizing Map? in P. Agarwal , A. Skupin A. , eds., 2008, Self-Organising Maps. Applications in Geographic Information Science, Chichester: John Wiley & Sons, pp. 1-20.

Skupin A., Hagelman R. (2005), Visualising demographic trajectories with Self-Organising Maps, *Geoinformatica*, 9(2), pp.159-79.

Snow C.P., 1959, The two cultures, Cambridge, England: Cambridge University Press.

Snow C.P., 1963, The two cultures: and a second look, New York: The New American Library.

Staniscia B., 2003, L'Europa dello sviluppo locale. I patti territoriali per l'occupazione in una prospettiva comparata, Roma:Donzelli.

Staniscia B., 2006, Sviluppo locale e turismo. Un percorso di ricerca, in Buzzetti L. e Montanari A. (eds.), Nuovi scenari turistici per le aree montane, Trento: Valentina Trentini Editore, pp. 51-80.

Szalai, A. and Petrella, R., eds, 1977. Cross-national comparative survey research: theory and practice. Oxford: Pergamon Press.

Tran T.D., Olsson G.A., Alpokay S. (eds.), 2013, Environmental stresses and resource use in coastal urban and peri-urban regions, Roma: Sapienza Università Editrice, forthcoming.

Tisdale H., 1942, The process of urbanisation, *Social Forces*, 20, pp.311-316.

Tobler W., 1970, A computer movie simulating urban growth in the Detroit region. *Economic Geography*, 46(2): 234-240.

Tobler W., 2004, On the First Law of Geography: A Reply, *Annals of the Association of American Geographers*, 94(2), pp. 304–310.

Tobler, , Wineberg S., 1971, A Cappadocian speculation, *Nature*, 231, 5297, pp. 39-42.

Torrens, P.M., 2010, Geography and computational social science, *GeoJournal*, 75, pp. 133-148.

Whitehand J.W.R., 2001, British urban morphology: the Canzanian tradition, *Urban Morphology*, 5(2), pp.103-109.

Yan Li and Shanmuganathan S. , 2007, Social area analysis using SOM and GIS: a preliminary research, Ritsumeikan Asia Pacific Studies (RCAPS), Ritsumeikan Asia Pacific University, Working Paper n. 07-3. August.

Yi-Chen Wang, Chen-Chieh Feng , 2011, Patterns and trends in land-use cover change research explored using self-organised maps, *International Journal of Remote Sensing*, 32, 13, pp.3765 – 90.

Armando MONTANARI, Department of European, American and Intercultural Studies, Sapienza University of Rome, Italy.

email: armando.montanari@uniroma1.it

Montanari, Armando, "Interpreting Local Issues in a Global Context: International Comparative Analysis and Coastal Urban Areas". In: Armando Montanari (ed.), *Urban Coastal Area Conflicts Analysis Methodology: Human Mobility, Climate Change and Local Sustainable Development*. Cap. 3. *SECOA FP7 Research Project*, Vol. 5. Rome: Sapienza Università Editrice, 2013. *http://digilab-epub.uniroma1.it*. DOI: 10.13133/978-88-98533-01-5. Web. 30 September 2013.

ABSTRACT: This text looks at urban problems in coastal areas which have undergone international comparative analysis. The first section concerns the emergence of comparative analysis to study urban development, and how it evolved after World War II, when an international effort was made to encourage scientific cooperation in order to overcome the cultural and ideological barriers and hostility that had been partly responsible for the world wars of the 20th century. The experiences of a half-century are then classified according to the political developments and the international agreements that resulted from changes in cultural and scientific thought. The aspects of this wide-ranging review of cognitive processes that are particularly significant for the SECOA project and the topics covered in this book are highlighted. In all comparative analysis projects, there is a phase of transition from analysing empirical data to drawing conclusions in the knowledge phase. The second part of the chapter outlines the problems the SECOA project encountered, and the methods that were used to get to knowledge.

KEYWORDS: Multinational, Comparative, Social Sciences, Humanities, Natural Sciences, Taxonomy, Interpretive Modelling.

Submitted:	6th June 2013
Accepted:	18th July 2013
Final Revision:	23th July 2013
Published:	30th September 2013

Annex I.

CHAPTER 4.

SECOA Territories and Conflicts: Data Dilemmas

Barbara STANISCIA
Department of European, American and Intercultural Studies,
Sapienza University of Rome, Italy

1. Introduction

The implementation of the principles of GIScience is strongly influenced by the definition of the phenomenon and the territory to be explored, the explanatory variables, and the availability of data. This chapter will deal with those issues: how the problems were placed in SECOA and which solutions have been adopted. The themes and topics characterizing SECOA are defined, and the variables chosen for analyzing, classifying, and modelling them are described.

The themes characterizing SECOA are those connected to *places* and to *conflicts*. We first discuss place and territory followed by conflict and contrast. Then, SECOA places and conflicts are defined, and SECOA variables and data are presented.

2. Places, Regions, and Territories

The notion of *place* has been the focus of intense scientific debates in geography. Cresswell (2004, p. 51) identifies three approaches in the conceptual definition of *place*: descriptive, social constructionist, and phenomenological. The first considers place in its specificity and peculiarity. It is typical of regional, cultural and human geography and serves as the basis for research on regions. Regions and places are considered as unique objects to be studied for their specific characteristics. The second approach is followed by Marxist, feminist, and post-structuralist geographers. It considers both the peculiarities of a place (per se) and the social process that leads to the construction of place. The third approach is taken by humanistic geographers, neo-humanists, and phenomenological philosophers. It considers human existence as determinant and essential in the definition of place. These perspectives can be considered as alternative or as complementary. They can also be jointly considered to obtain a holistic definition of place.

In the liquid modernity (Bauman, 2000), mobility becomes an important element in defining places. Hannam, Sheller and Urry (2006, p. 1) remind us that "the concept of mobilities encompasses both the large-scale movements of people, objects, capital and information across the world, as well as the more local processes of daily transportation, movement through public space and the travel of material things within everyday life". Thrift (1994, p. 212) wonders "what is place in this 'in-between' world? The short answer is compromised: permanently in a state of enunciation, between addresses, always deferred. Places are 'stages of intensity'. Traces of movement, speed and circulation. One might read this depiction of 'almost places' (...)". Hannam, Sheller and Urry (2006, p. 2) observe that "the global order is increasingly criss-crossed by tourists, workers, terrorists, students, migrants, asylum-seekers, scientists/scholars, family members, business people, soldiers, guest workers and so on. Such multiple and intersecting mobilities seem to produce a more 'networked' patterning of economic and social life, even for those who have not moved."

Place, therefore, is considered in various dimensions: the objective dimension views place as a geographical space, naturalistically and environmentally determined, existing irrespective of the existence of humans; the subjective dimension views place as a geographical space defined by the presence of the subject and from the perception of the subject; the social dimension views place as a geographical space defined by the action of society or the community and by the interaction of its members; the last dimension considers place as a manifestation of a given economic and social paradigm. The place in SECOA was considered as the result of the interaction of objective, subjective, social, and economic dimensions.

In this context, place is considered as the basic unit of a region. Place is a space of living, consuming, and producing; it is the space of social relations and mobility. Place is a social construction. There is no opposition to the global dimension, and there is no definition of scale (regional, provincial, local). There is a reference to the place. Place is understood in its complexity as a spatial-physical, social, and economic entity. A place, therefore, is given by the interrelationship of a human community, settled in a given physical space in which it carries on its business and social activities. According to Sforzi (2000, p. 186), "the place is a piece of territory to which a human group gives an individuality that comes from the individual functions and the overall role it plays in the system of spatial structures of society. And since the subject, as well as the recipient of such functions is the population, whether it lives and acts in the place or visits it from time to time (Nice, 1987, pp. 117-118), here it is that the place results as a social construction". A variety of places linked together through networks gives life to the region, which is a complex and derived entity. The importance of a region can be considered a recent acquisition. In the Fordist model, the region does not appear. The factors of production — declined in different ways and with different roles in the different schools of thought — remain: land, capital and labour. In the post-Fordist models, however, land is replaced with the region understood as ground-terrain-soil, physical space, local community-social capital, local players, processes of production and consumption. The notion of region as a factor of production is much more complex than that of land. It includes intangible factors such as synergies among local players, the interactions of the components of local community and the active participation of the institutions. Sforzi (2000, p. 187) points out that "the economic, social and political organization of the region is expressed in a system of places, and there is the distinct possibility of interpreting the society and the economy of a nation-state through a configuration of local systems".

The consideration of place in this way has led to the need to involve different disciplines, to find a common language among them, and a common way to select, analyze, synthesize, and compare the data. This will be discussed in detail in the fourth section of this chapter.

3. Conflicts and Contrasts

According to the Oxford English Dictionary, a conflict is "a serious disagreement or argument, typically a protracted one." Contrast is defined as "the state of being strikingly different from something else in juxtaposition or close association." In literature, several kinds of conflicts and contrasts are analyzed. We are interested in the conflicts and contrasts linked to the alternative uses of the natural environment and resources by different users.

The definition of conflict in the environmental field is part of the recent scientific debate; there is not a unique interpretation of what conflict means. In early publications, conflict involves violence or armed fight and refers to clashes between territorial units (e.g. neighbouring states). Homer-Dixon (1991, p. 77) was one of the first authors to introduce the concept of environmental conflict; in presenting his research, he states that "environmental change may play a variety of roles as a cause of conflict, but I bound my analysis by focusing on acute national and international conflict, which I define as conflict involving a substantial probability of violence". Libiszewski (1992, p. 14), referring to violent environmental conflicts lists three characteristics: "environmental conflicts are characterized by the principal importance of degradation in one or more of the following fields: overuse of renewable resources; overstrain of the environment's sink capacity (pollution); impoverishment of the space of living". This definition of conflict was behind the ENCOP project – (Environment and Conflicts Project), of which the author was a member. Presenting the results of the Swedish Research Programme SUCOZOMA - Sustainable Coastal Zone Management, Bruckmeier (2005, p. 66) offers a comprehensive and exhaustive classification of conflicts starting from the definition of Tillitt (1999): "a conflict arises when two (or more) people (or groups) perceive that their values or needs are incompatible – whether or not they propose, at present or in the future, to take any action on the basis of those values or needs. Thus, while a problem or dispute relates to a specific action or situation …, a conflict can exist without such a specific focus. Two parties can be in conflict because of what each believes, regardless of whether any action has been or is being taken on the basis of the belief." Bruckmeier (2005) emphasizes that research on conflicts crosses disciplinary fields, due to the intrinsic characteristics of the object of analysis. The author laments that "instead of studying conflicts for the purpose of improved conflict resolution, of mitigation and cooperation between stakeholders they are neglected or only dealt with indirectly in research as well as in management." He summarizes the way in which conflicts have been considered in the literature into four types:

(i) Conflicts defined as dichotomies, using juxtapositions such conflicts manifest-latent, constructive and destructive, real and virtual;

(ii) Conflicts defined from the various spheres of human activity and the various scales involved, using categories such as: (a) intra-personal conflicts (psychic conflicts); (b) inter-personal conflicts (personal relations, small groups); (c) social conflicts: intra-societal conflicts (national- and sub-national levels) between groups (political, religious, economic, social, ethnic, race, gender conflicts); conflicts related to information, communication, knowledge (access or exclusion, distribution); technology related conflicts (e.g. use of nuclear energy); environmental conflicts (conflicts in environmental policy; "livelihood conflicts"); violent conflicts (civil war, criminality, terrorism); (d) international and global conflicts: power-based conflicts (for example trade wars); war; competition between socio-political systems and worldviews (East-West conflict/cold war); global conflicts about resources and distribution of resources;

(iii) Conflicts as defined by the different attitudes and behaviour of the parties involved (Rapoport, 1970): (a) fights (unwillingness to compromise), (b) games (with pre-established rules for conflict solution), and (c) debates (based on contradicting values, interests or world views); and

(iv) Conflicts as defined by the objects of contention (Schmidtz, 2000): (a) conflicts in use, (b) conflicts in values, (c) conflicts in priorities or needs.

The way in which conflicts were defined and studied in SECOA is presented in the next section.

4. Places and Conflicts in SECOA

SECOA focused on places where conflicts were occurring. Table 4.4.1 provides an overview of the regions and conflicts studied by SECOA.

Table 4.4.1 SECOA regions and conflicts.

Countries	Metropolitan/Urban Regions/Areas	Conflicts
BE	Ostend	Ostend airport
BE	Zeebrugge	Schipdonk canal; Zeebrugge harbour
IL	Haifa	Haifa Port
IL	Tel Aviv	Palmachim beach; Netanya sandstone cliffs
IN	Chennai	Pallikaranai Marshland
IN	Mumbai	Sanjay Gandhi National Park; Mangrove forest
IT	Chieti-Pescara	Costa Teatina National Park
IT	Rome	Civitavecchia; Ostia water-use & management
PT	Algarve	Barrier islands (Ria Formosa Natural Park)
PT	Funchal	Funchal bay (Madeira Island)
PT	Lisbon	Trafaria and Costa da Caparica
SE	Gothenburg	Torsviken; Kungsbacka
SE	Malmö	Managing urban sprawl; Falsterbo-Peninsula: Vellinge municipality
UK	Portsmouth	Langstone Harbour / Farlington Marshes; Tipner Regeneration
UK	Thames Gateway	Barking Riverside; Lower Thames Crossing
VN	Haiphong	Haiphong port; Industrial zone; Cat Ba
VN	Nha Trang	Nha Trang

The first problem addressed was the operational definition of places. The problem was how to define places that could be studied in a comparative way. The solution was to give a definition of place that made sense for each country and then use methods of taxonomy and modelling to make a comparison.

Places are located in coastal areas. Places are metropolitan and urban regions defined according to case-based methodologies, formed of different functional zones (Williams, 2012). Metropolitan and urban regions are considered in both their static and dynamic aspects, namely, their structures (natural, environmental, social, economic, and urban) and the flows by which they are crossed (of persons; temporary and permanent; consumption-oriented and production-oriented). Williams (2012) identifies the following models for the definition of metropolitan areas and city-regions: (i) based on local labour market, journey to work or travel to work definitions, (ii) based on housing-market definitions, (iii) economic activity-based definitions, (iv) based on service-district definitions, (v) based on administrative area definitions, and (vi) based on transport data (as a proxy for journey to work) definitions.

The local context has been an essential element in the definition of the territories. Metropolitan and urban areas have been identified with reference to what "metropolitan area" and "urban area" mean in the different national and regional contexts. The meaning of a metropolitan area cannot be the same in India and Portugal, for instance. Accordingly, metropolitan areas were defined using a combination of different criteria and variables. They are summarized in Table 4.4.2.

Table 4.4.2. Criteria and variables used for the definition of metropolitan zones in the case studies.

Variables/Case studies	**BE**	**IL**	**IN**	**IT**		**PT**		**SE**	**UK**	**VN**
				Rome MA	**Chieti-Pescara UA**	**Lisbon MA**	**Eastern Algarve and Funchal regions**			
Administrative boundaries	X	X	X	X	X	X	X	X	X	X
Population density and/or dynamics	X			X	X		X			
Migration	X					X				
Housing (first and/or second homes)	X			X			X			
Income	X									
Employment	X				X		X		X	
Journeys to work and/or travelling times	X	X		X		X		X	X	
Flows related to the use of services (journeys to schools, hospitals)				X						
Transport systems and/or costs				X	X					

Functional zones were identified in each metropolitan/urban area based on urban structures and data availability.

Table 4.4.3. Functional zones in the case studies.

Functional zones/Case studies	BE	IL	IN	IT	PT	SE	UK	VN
Core	X	X	X	X	X	X	X	X
Ring	X	X	X	X	X	X	X	X
Further subdivisions								
Inner ring		X		X		X		
Middle ring		X						
Outer ring		X		X		X		

The definition of metropolitan/urban areas and functional zones was particularly significant for some socio-economic phenomena, in particular for the study of migration and tourism. In each region, coastal areas were identified based on criteria adapted to national situations, data availability, and themes addressed. The definition of coastal area was not relevant for the processes of creating taxonomy and modelling, since the use of GIS made different information and definition compatible and comparable.

Figure 4.4.1. Ostend region and its functional zones.

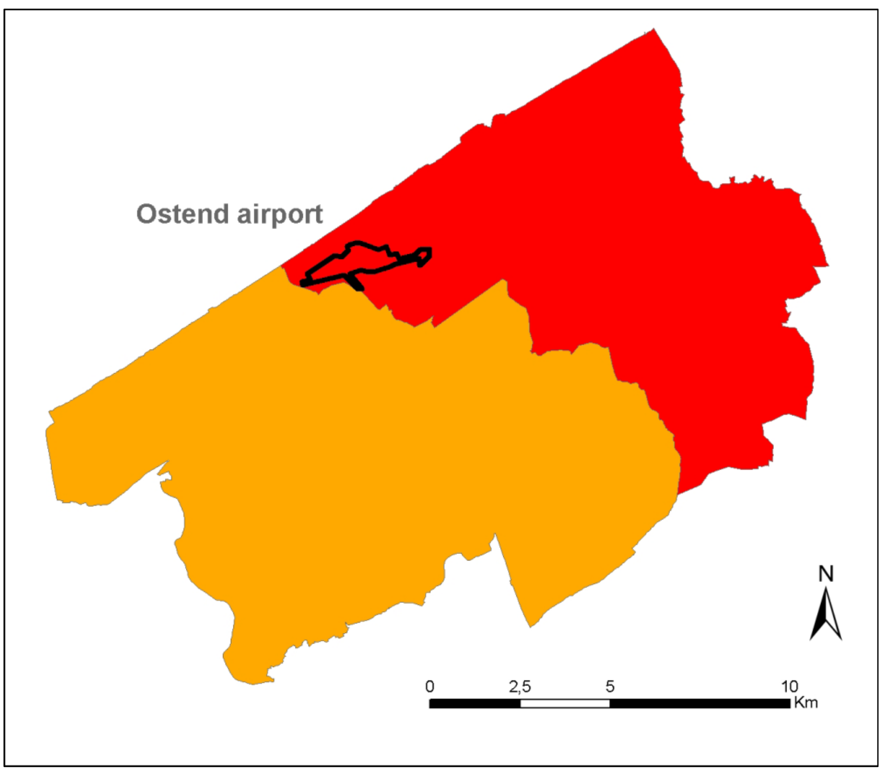


Figure 4.4.2. Zeebrugge region and its functional zones.

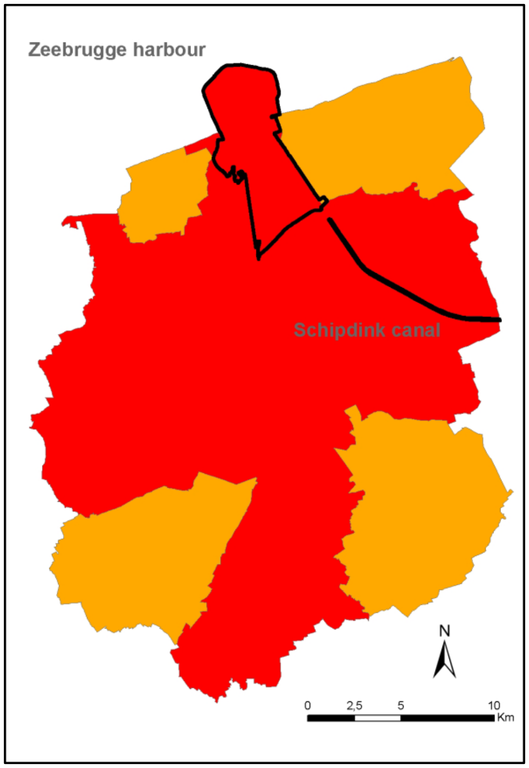


Figure 4.4.3. Haifa region and its functional zones.

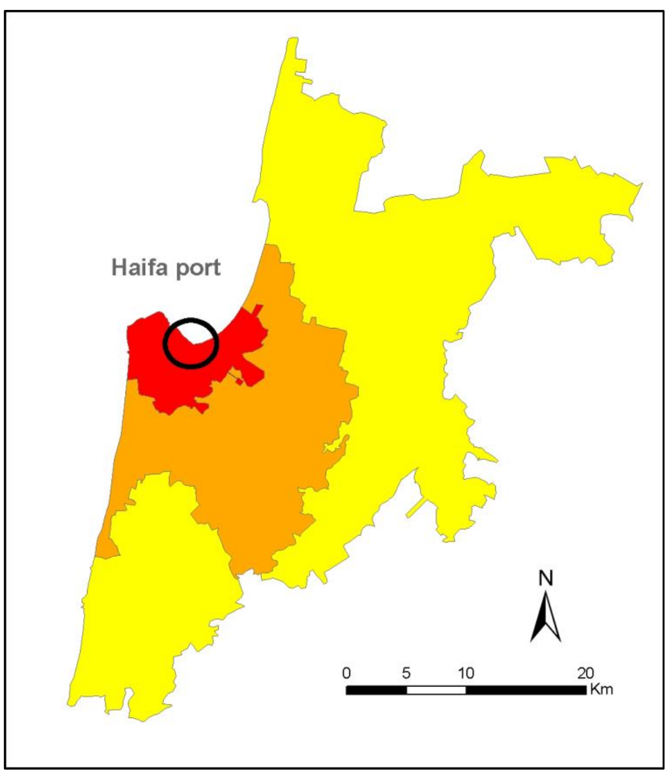


Figure 4.4.4. Tel Aviv region and its functional zones.

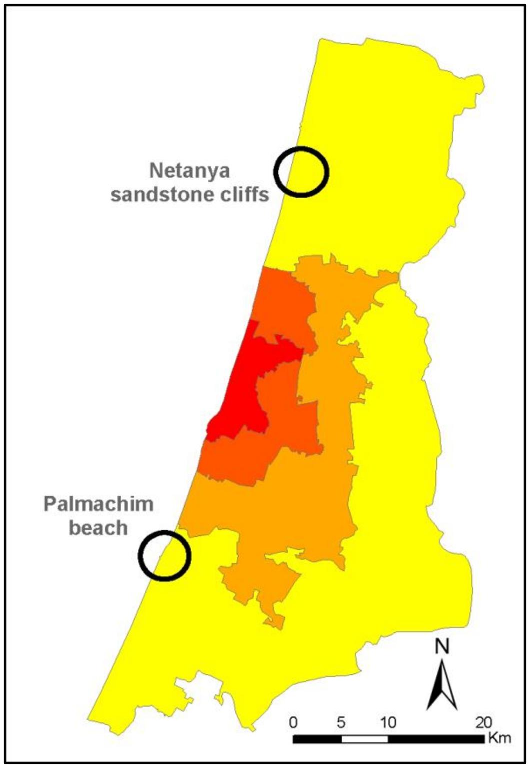


Figure 4.4.5. Chennai region and its functional zones.

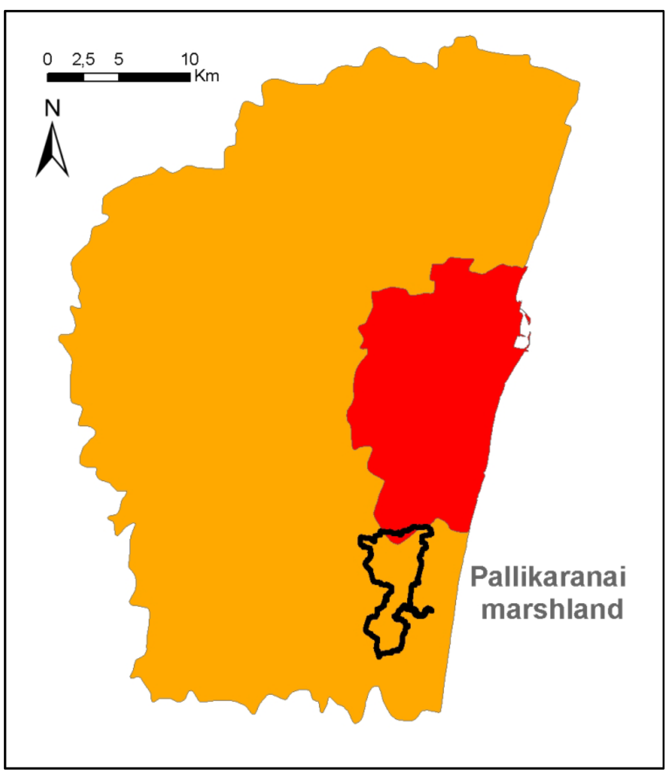


Figure 4.4.6. Mumbai region and its functional zones.

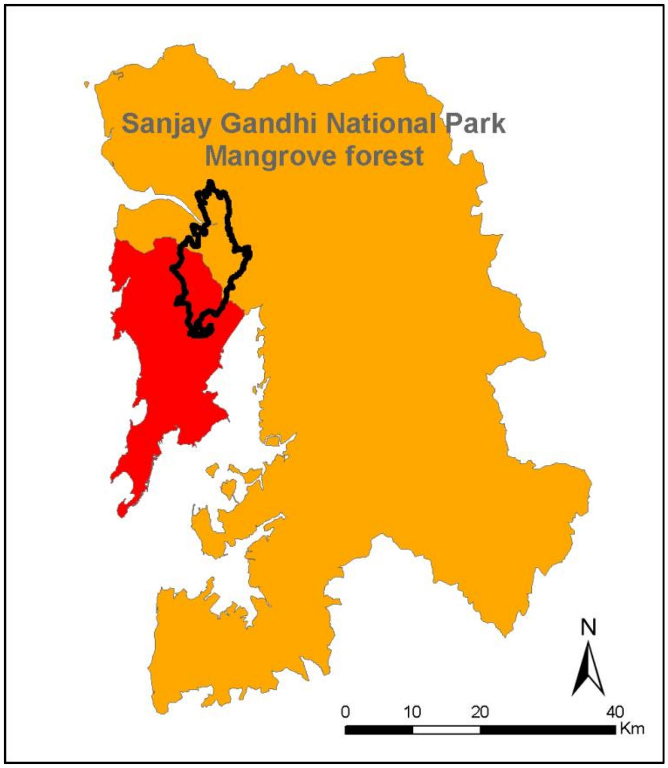


Figure 4.4.7. Chieti-Pescara region and its functional zones.

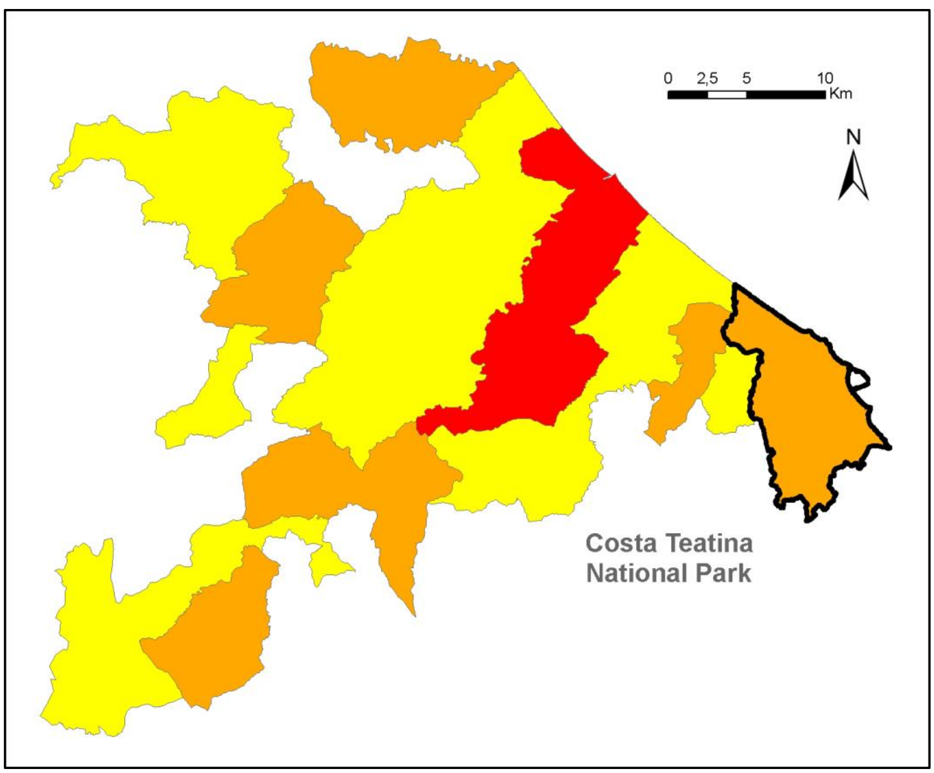


Figure 4.4.8. Rome region and its functional zones.

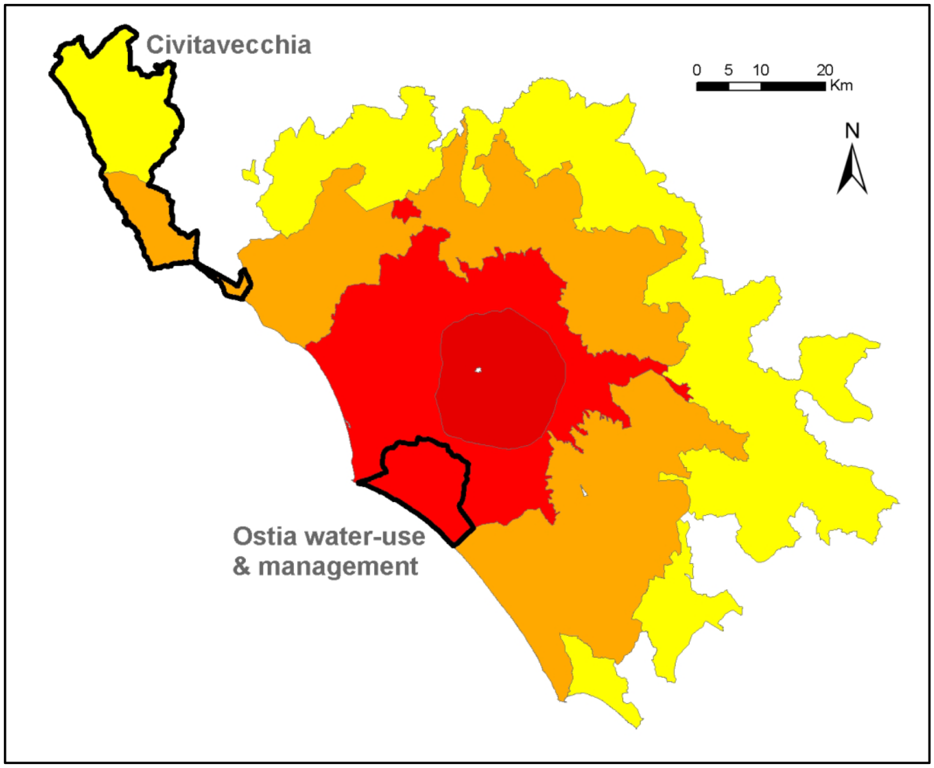


Figure 4.4.9. Algarve region and its functional zones.

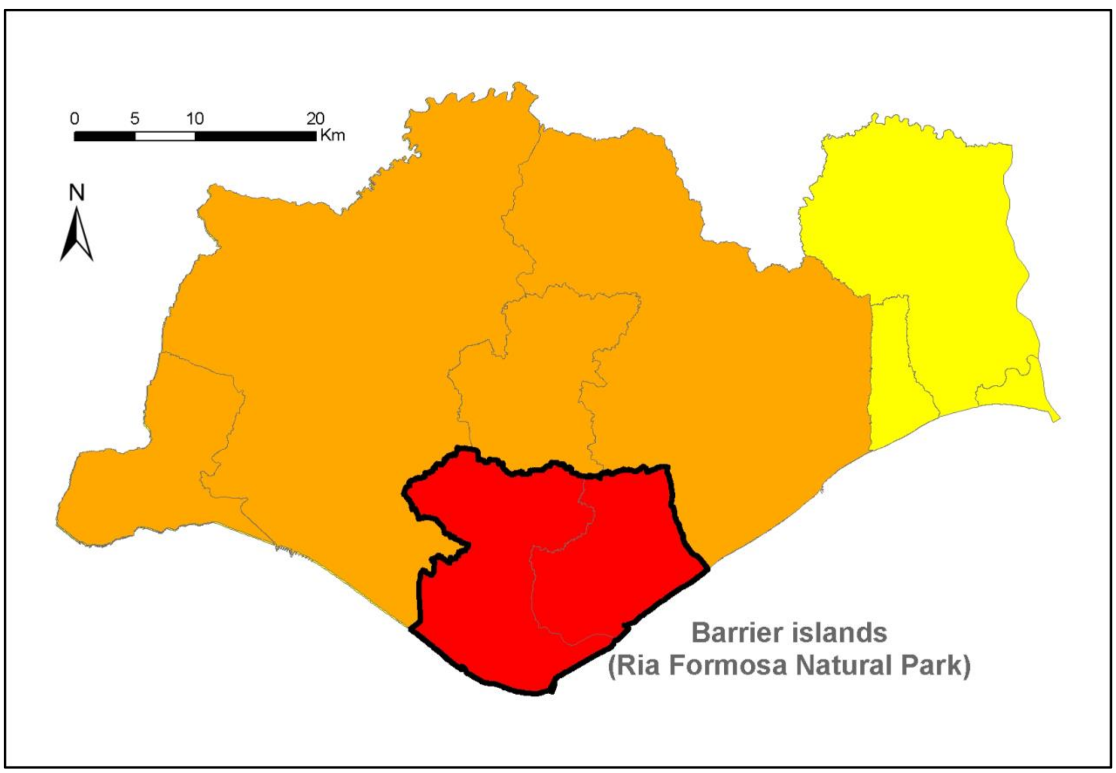


Figure 4.4.10. Funchal region and its functional zones.

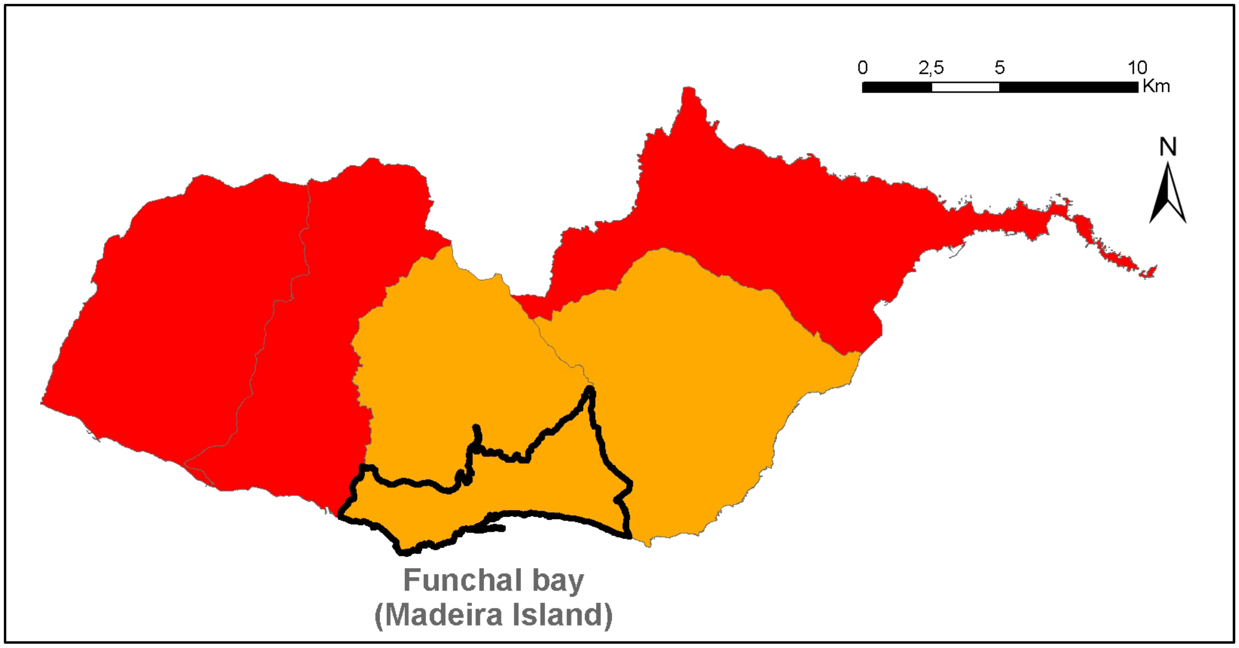


Figure 4.4.11. Lisbon region and its functional zones.

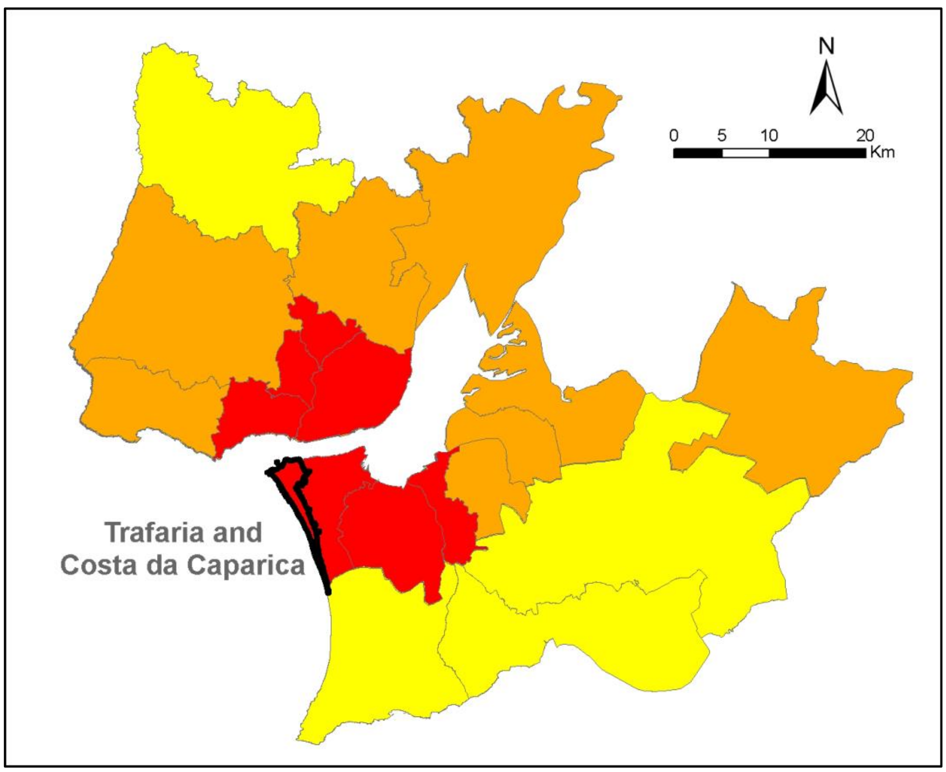


Figure 4.4.12. Gothenburg region and its functional zones.

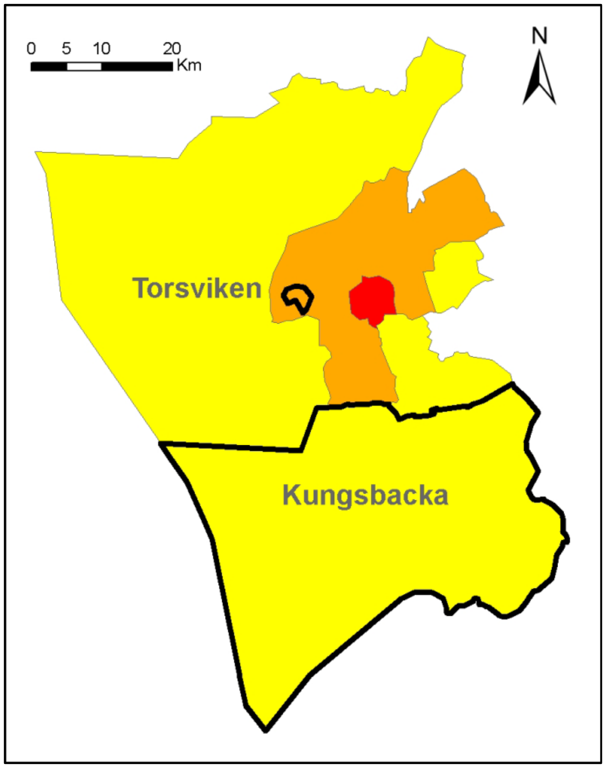


Figure 4.4.13. Malmö region and its functional zones.

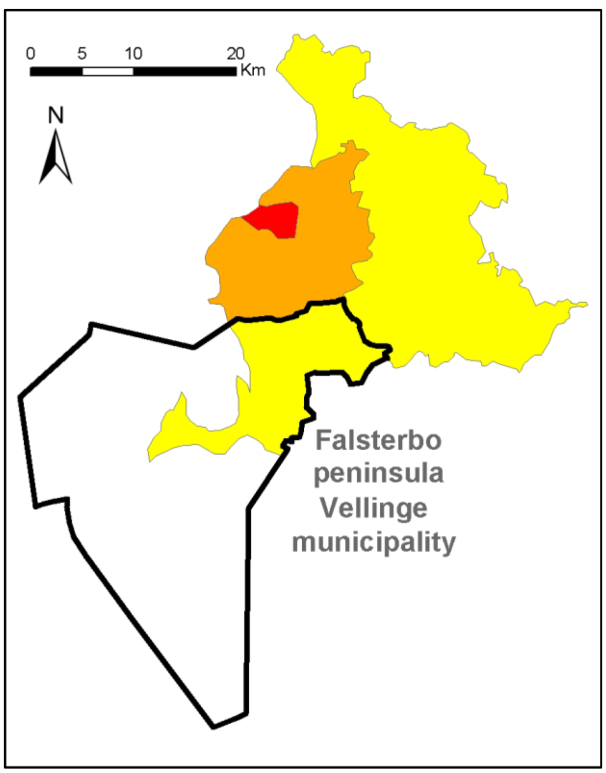


Figure 4.4.14. Portsmouth region and its functional zones.

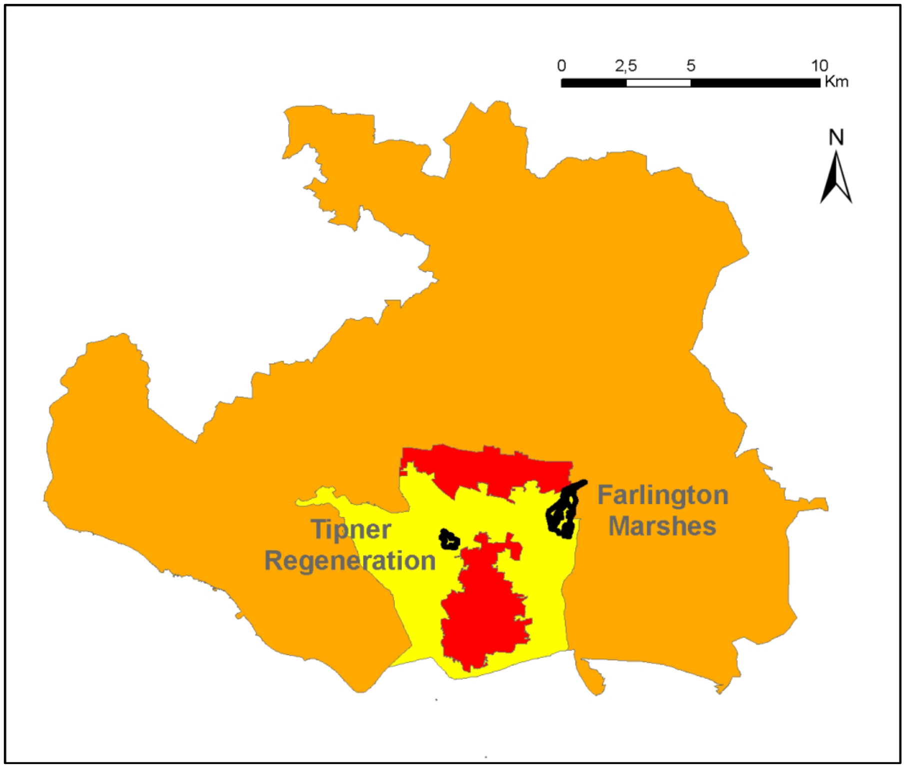


Figure 4.4.15. Thames Gateway region and its functional zones.

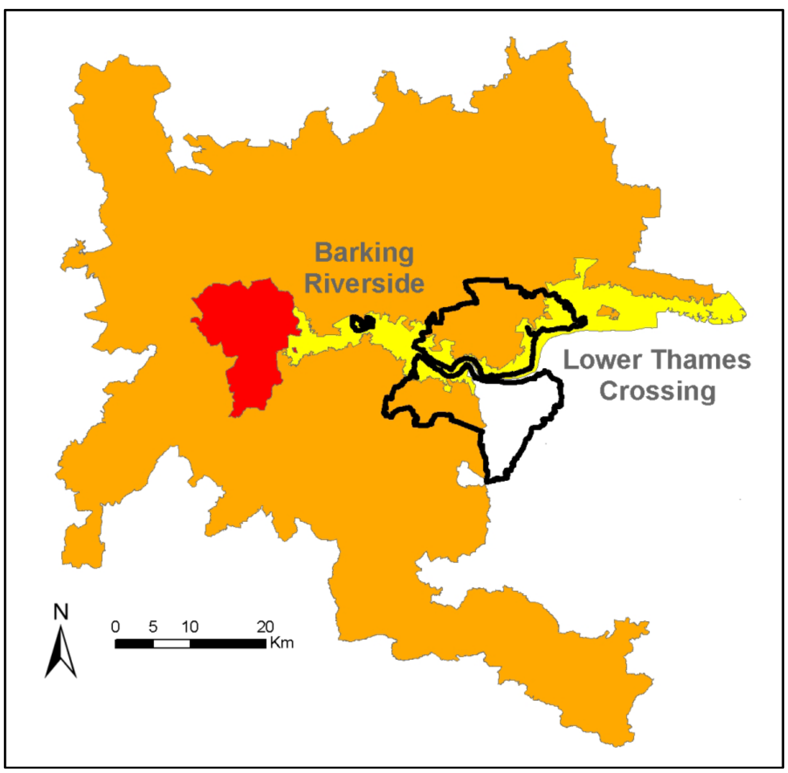


Figure 4.4.16. Haiphong region and its functional zones.

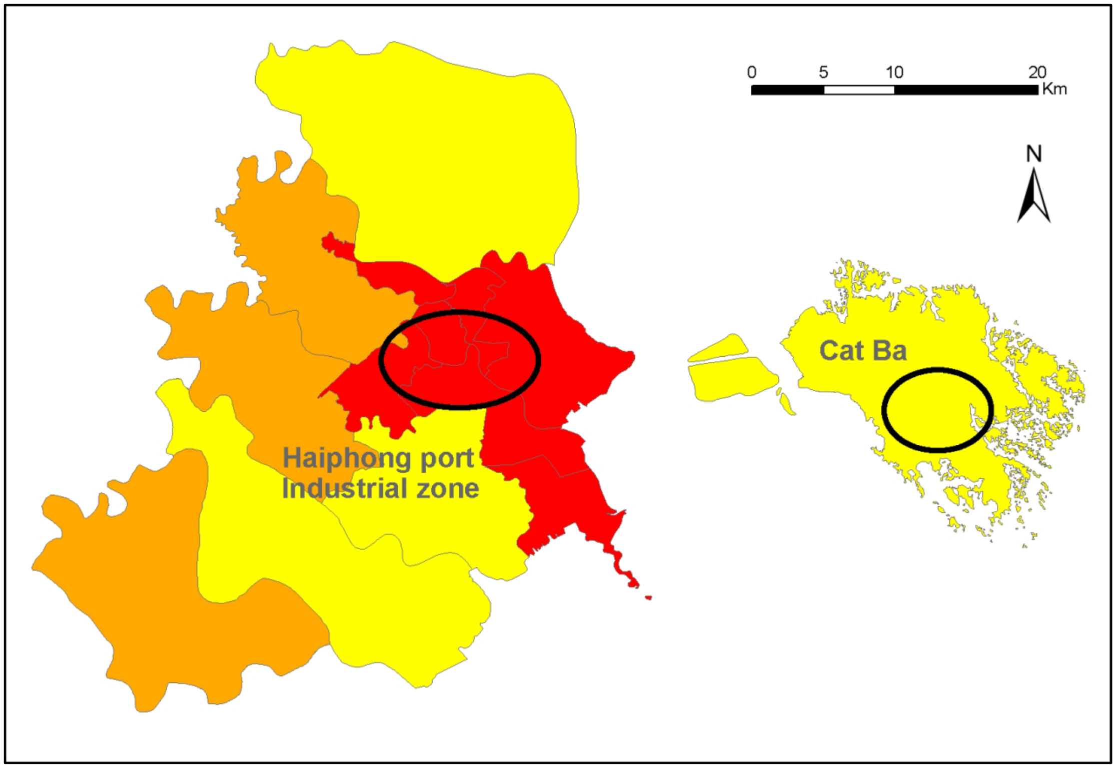
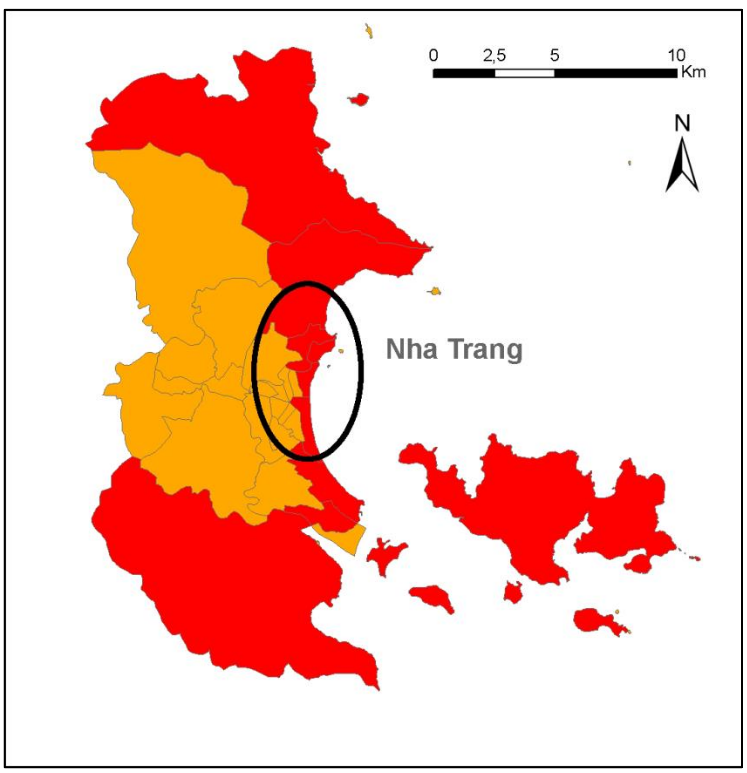


Figure 4.4.17. Nha Trang region and its functional zones.



The second problem addressed was the operational definition of conflict. What is a conflict? What are the characteristics identifying conflicts, and what is the object of the conflict? The perspective of environmental and urban policies is the one with which the theme has been addressed in SECOA. Conflicts were selected and analyzed considering that the ultimate goal was to seek innovative policies for different scenarios. It was, therefore, a policy-oriented research (Montanari, 2012). Conflicts were considered as "complex societal problems" (Khan *et al.*, 2013); conflicts in SECOA, thus, are a social construct as well as places.

Conflicts were linked to environmental issues. They were selected and defined following the Compram method as defined by De Tombe (2001). They were case-based and involved at least one of the following issues (Khan et al., 2013): (i) economic development (industrial development, tourist industry, harbour/port restructuring, marina re-construction) vs. environmental protection (creation, preservation and conservation of environmentally and ecologically sensitive, valuable and protected areas); (ii) preservation of natural sites and biodiversity; and (iii) contrasts for the use of resources between residents and newcomers for processes of human mobility. Conflicts were considered in a time span of 10 to 20 years.

Each conflict has been analyzed with respect to five aspects (Khan et al., 2013): (i) nature of the conflict, the context in which the conflict takes place and its causes; (ii) stakeholders involved in the conflict, with attention paid to their interests, goals, positions, capacities, relationships, salience; (iii) typologies of conflicts according to their manifestation, underlying cause, stage and scale; (iv) state of mediation or resolution of the conflict; and (v) ranking based on criticality, urgency and duration.

5. Descriptive and Explanatory Variables in SECOA

The transition from the idea of territory and conflict to the description of them required the definition of descriptive and explanatory variables. The choice of variables was based on two main criteria: significance, namely the ability to describe and explain the analyzed phenomena, and feasibility, namely the availability of data for all case studies. The main difficulty is posed by the second criterion. Data availability significantly diversified the case studies. Another difficulty was the difference in the languages specific to different disciplines involved in the analysis. The description of the territories required multidisciplinary analyses and the need to standardize several different languages.

Data used for the taxonomy and the model were both quantitative and geo-referenced and qualitative. The first concern natural, environmental, socio-economic characteristics of the metropolitan and urban areas, while the second concern the analyzed conflicts. The availability of quantitative data, their homogeneity, the degree of detail, and the spatial scale are strongly country-dependent: the EU countries' data were relatively more abundant, homogeneous and spatially detailed than data for Asian countries. All data were available for two time periods of ten years each; with some differences among countries, data were available for 1991-2000 and 2001-2011.

For the physical description of the territories, variables related to land use and flood risk were used. For the description of the social and economic aspects relating to individuals, variables related to the stable part of the population (residents) and the mobile part (migrants and tourists) were used. For a description of the economic aspects of enterprises, variables related to employees and local units in the main sectors of economic activities were used. The political dimension of governance has been analyzed considering the different systems of government and the different ICZM mechanisms in the case studies. The variables considered are certainly representative of the territories but cannot be considered as exhaustive. However, they were able to explain the spatial dynamics.

Table 4.5.1. List of the variables used for the description of the territories in the SECOA taxonomy and model.

Variables	**Variables' specifications**	**Used for taxonomy (only aggregated data)**	**Used for taxonomy (both geo-referenced and aggregated data)**	**Used for model**
land use types:				
	agriculture		X	X
	industrial		X	X
	industrial-commercial		X	X
	natural		X	X
	open space		X	X
	residential mixed		X	X
	government	X		X
road network length		X		
areas at risk of:				
	flood hazard		X	X
	inundation due to sea level rise		X	X
	extreme rainfall	X		
sustainability index		X		

number of residents			X	X
number of temporary residents present		X		
number of residents by social status		X		
number of persons living in slums		X		
number of households			X	X
number of dwellings		X		
number of second homes		X		
number of temporary dwellings		X		
number of commuters		X		
unemployment rate			X	X
average income			X	X
motorization rate or number of registered vehicles			X	X
net migrants			X	X
number of tourists			X	X
peak number of passengers		X		
number of higher education students		X		
employees/local units:				
	in agriculture		X	X
	in the industrial sector		X	X
	in the tertiary sector		X	X
mechanisms of ICZM:				
	environmental impact assessment			X
	planning hierarchy			X
	setback lines			X
	marine spatial planning			X
	regulatory commission			X
political systems:				
	centralized			X
	decentralized			X
	unitary			X
	federal			X

Conflicts have been described through several characteristics and associated variables. The first group of variables concerns the object of conflict; three main themes have been identified: the conflict between economic growth and environmental protection *tout court*, the need to preserve specific natural sites and biodiversity, and the contrast for the use of resources resulting by the presence of newcomers in the area. The second, third, and fourth group of variables relate to categories of use of resources subject to a conflict, the underlying causes of the conflict (as defined by Chandrasekharan, 1996), and the size of the conflict (as defined by Warner, 2000 and Bruckmeier, 2002), respectively. The fifth and sixth group of variables concern the development of the conflict over time (as defined by Cadoret, 2009) and the stage at which the conflict is currently (as defined by Rupesinghe, 1995), respectively. The seventh group concerns a comprehensive assessment of the conflict (obtained through Delphi or alike methods).

Table 4.5.2. List of the variables used for the description of the conflicts in the SECOA taxonomy and model.

Variables	Variables' specifications	Variables' contents
Theme	Economic Development vs. environmental protection	Port/harbour restructuring, expansion and infrastructure; tourism lead infrastructure and urban development; industrial zones and airports expansion; waterfront and brown fields regeneration; energy/power generation; pollution (air, soil, water) associated; landscape/nature conservation; wildlife habitat protection; parks/beaches /protected areas; waste water management
	Preservation of natural sites and biodiversity	Preservation of natural sites/ islands /marshlands; protection of biodiversity habitats; national parks; cultural heritage and landscapes
	Human mobility and contrast for use of resources	Physical infrastructures (transport, utilities and waste water); social infrastructure (migrants/new comers, social exclusion/segregation, slums); tourism (housing, recreation and second homes); commuting (job related, daily, occasionally)
Category of uses		Ports and harbours related uses; urban growth and development in terms of specific urban functions including tourism; energy generation/production; natural environment and habitat including national parks and protected areas
Underlying cause		Change in resource quality and availability; legal/policy reasons; infringements over access; authority over resource; conflicts that are value based

Scale		Micro-macro conflicts; inter micro-micro conflicts; intra micro-micro conflicts; hybrid; between local and territorial/regional scales
Dynamics and manifestation over time		Chronic, anticipation, hybrid, hushed or deferred
Stage		Endurance, management, formation, transformation, manifestation
Final Ranking		Criticality; duration; urgency

6. Conclusion

17 coastal urban areas and 27 conflicts in SECOA have been analyzed, classified, and modelled using variables expressed through quantitative and measurable data and qualitative information. The selected variables had to meet two requirements: to be able to represent the phenomena and to be comparable through different case studies. This second requirement led to a strict selection of the variables related to quantitative data. Territories have been studied through variables related to their natural-environmental, socio-economic, and policy-related characteristics. Conflicts have been studied through variables related to their objects, their causes, their manifestation, their duration, and their scales. National, cultural, historical differences among the case studies may have affected the processes of data collection and information creation. Taxonomy and modelling aim to reducing the background noise and highlight similarities and differences.

7. References

Bauman, Z. (2000) *Liquid Modernity*, Cambridge: Polity Press and Oxford: Blackwell Publishers.

Bruckmeier, K. (2002) Conflict Mitigation through Local Resource Management, *HERS-SUCOZOMA-Report* 8, pp. 1-25.

Bruckmeier, K. (2005) Interdisciplinary Conflict Analysis and Conflict Mitigation in Local Resource Management, *Ambio* 34(2), pp. 65-73, http://www.ambio.kva.se.

Cadoret, A. (2009) Conflict dynamics in coastal zones: a perspective using the example of Languedoc-Rousillon (France), *Journal of Coastal Conservation* 13, pp. 151-163.

Chandrasekharan, D. (1996) Addressing Natural Resource Conflicts through Community Forestry: The Asian Perspective, *Proceedings of electronic conference on Addressing Natural Resource Conflict Through Community Forestry,* Rome: Food and Agricultural Organization of the United Nations.

Cresswell, T. (2004) *Place. A Short Introduction*, Oxford: Blackwell Publishing.

DeTombe, D. J. (2001) Compram, a method for handling complex societal problems, *European Journal of Operational Research* 128, pp. 266-281.

Hannam, K., Sheller, M., and Urry, J. (2006) Mobilities, Immobilities and Moorings, *Mobilities*, 1(1), pp. 1-22.

Homer-Dixon, T.F. (1991) On the threshold: environmental changes as causes of acute conflict, *International Security*, 16(2), pp. 76-116.

Khan, A., Le Xuan, Q., Corijn, E., and Canters, F. (2013) Comparative Analysis and Assessment of Environmental Conflicts: A Synoptic Overview of Methodological Developments, in: Khan, A., Le Xuan, Q., Canters, F. and Corijn, E. (Eds) (2013) *Environmental Conflicts in Coastal Urban Areas. Towards a Strategic Assessment Framework for Sustainable Development*, Rome: Sapienza University Press on line, pp. 35-97.

Libiszewscki, S. (1992) What is an environmental conflict?, *ENCOP Occasional papers*, Zurich: Center for Security Studies-ETH, pp. 1-16.

Montanari, A. (2012) Social sciences and comparative research in Europe: cross-national and multi-disciplinary projects for urban development. The role of geography, *Belgeo* (on line) 1-2, http://belgeo.revues.org/6085.

Nice, B. (1987) *Introduzione alla geografia economica*, Firenze: Cusl.

Rapoport, A. (1970) *Games, Fights and Debates*, Ann Arbor: The University of Michigan Press.

Rupesinghe, K. (1995) Multi-Track Diplomacy and the Sustainable Route to Conflict Resolution, *Cultural Survival Quarterly* 19(3), http://www.culturalsurvival.org.

Schmidtz, D. (2000) Natural Enemies: An Anatomy of Environmental Conflict, *Environmental*

Ethics 22, pp. 397-408.

Sforzi, F. (2000) Il sistema locale come unità d'analisi integrata del territorio, in: Gori, E., Giovannini, E. e Batic, N. (a cura di) *Verso i Censimenti del 2000. Atti del Convegno della Società Italiana di Statistica* – *Udine - 7/9 giugno 1999*, Udine: Forum, pp. 185-192.

Thrift, N. (1994) Inhuman geographies: landscapes of speed, light and power, in: Cloke, P. (Ed) *Writing the Rural: Five Cultural Geographies*, London: Paul Chapman, pp. 191-250.

Tillitt, G. (1999) *Resolving Conflict. A Practical Approach*, Oxford: Oxford University

Press (2nd ed.)

Warner, M. (2000) Conflict management in community-based natural resource projects: experiences from Fiji and Papua New Guinea, *Working Paper 135*, London: ODI.

Williams, A. (2012) Human mobilities and urbanization impacts: a conceptual framework, in: Williams A. (Ed) *Human Mobility in Coastal Regions: the Impact of Migration and Temporary Mobilities on Urbanization*, Rome: Sapienza University Press on line, pp. 19-34.

Barbara STANISCIA, Department of European, American and Intercultural Studies, Sapienza University of Rome, Italy.

email: barbara.staniscia@uniroma1.it

Staniscia, Barbara, "SECOA Territories and Conflicts: Data Dilemmas". In: Armando Montanari (ed.), *Urban Coastal Area Conflicts Analysis Methodology: Human Mobility, Climate Change and Local Sustainable Development*. Cap. 4. *SECOA FP7 Research Project*, Vol. 5. Rome: Sapienza Università Editrice, 2013. *http://digilab-epub.uniroma1.it*. DOI: 10.13133/978-88-98533-01-5. Web. 30 September 2013.

ABSTRACT: The present chapter addresses the definition of territories and conflicts in SECOA, the choice of variables, and related data. Territories and conflicts are considered as social constructs. Territories are studied through variables related to their natural-environmental, socio-economic, and policy-related characteristics. Conflicts are studied through variables related to their objects, their causes, their manifestations, their durations, and their scales. Variables are expressed through quantitative and measurable data and qualitative information. The involvement of multiple disciplines and very heterogeneous countries in SECOA has imposed choices of variables and data that were meaningful for the analysis of the phenomena. The 17 urban coastal areas and 27 conflicts in SECOA are presented here together with their descriptors.

KEYWORDS: Places, Territories, Conflicts, Variables, Data.

Submitted:	6th June 2013
Accepted:	18th July 2013
Final Revision:	23th July 2013
Published:	30th September 2013

Annex I.

CHAPTER 5.

Neural Networks for Multidisciplinary Approach Research

Alessandro LONDEI
Department of European, American and Intercultural Studies,
Sapienza University of Rome, Italy

1. Overview: Non-linear Systems and Connectionism

1.1. Introduction

Understanding the world around us is usually a difficult task. All dynamically evolving phenomena in the natural world are produced by a strong interaction among a great number of causes of which only few are visible or measurable. Moreover, the phenomena, like the weather evolution, may be so distributed over the space or time that only a small number of measurements can be done, making the understanding of the overall system difficult and approximated. As will be shown later, some characteristics of systems can produce a very strange behaviour, even when the elements constituting the system are a small number. All these elements and their mutual interaction can produce the so-called *complexity*.

In order to understand the approach a researcher may use in analysing a system, a very simple metaphor may be adopted: the iceberg. An iceberg is a floating ice mountain in the sea that shows only a small visible part above the waterline. If some specific tools are not used to improve our investigation of the iceberg, all we can describe is the movement of its visible part, the rate of melting, the colour, the transparency, and the like. Any other aspect that belongs to the submerged part is excluded by a direct measurement. Therefore, any hypothesis about the global behaviour of the iceberg can be proved by using only what we are allowed to see directly. Is all this incomplete amount of accessible information enough to fully describe the iceberg and its future evolution? This is a very difficult question to answer. All we could say is that the visible behaviour, in some sense, contains also the occulted information and everything that is out of our sight can be extracted by what is known. Even when no theories or hypotheses are allowable to create a reference framework, complex systems have the characteristic to show an evolution through the mixed actions or interactions of the variables.

In observing natural, social, economical, physical, biological systems, we basically deal with measured data that give us a partial knowledge of the "visible part" of the system. Therefore, data is required to re-build a mathematical or algorithmic framework that could be sufficiently detailed and powerful to describe the fundamental aspects of the system under study, its evolution over time, and the meaningful characteristics.

The next parts of chapter 5.1 focus on the definition of complex system and the techniques, some traditional and accepted, some more recent and sophisticated, that can be a useful tool for accomplishing this task.

1.2. Linear and Non-linear Systems

All natural phenomena that can be observed regularly in sciences, including physics, biology, sociology or geography, and in technological worlds, are governed by specific and well-determined rules. These rules are usually deterministic, but there can be some probabilistic element, such as noise, uncertainness, or a multitude of quasi-negligible influencing or uncontrolled causes, which are usually not included in the sketch of our phenomenon comprehension. From the mathematical point of view, such phenomena are considered as the outcomes of a particular entity called *system*.

A system is an organized, purposeful structure that consists of interrelated and interdependent elements (components, entities, factors, members, parts etc.). These elements continually influence one another (directly or indirectly) to maintain their activity and the existence of the system in order to achieve the goal of the system.

Although all systems have outputs, which are considered as observable variables that make it possible to have a measure of what the system is doing at a given time, they may also:

a) have inputs and feedback mechanisms;
b) maintain an internal steady-state (called homeostasis) despite a changing external environment;
c) display properties that are different than the whole (called emergent properties) but are not possessed by any of the individual elements;
d) have boundaries that are usually defined by the system observer.

Systems underlie every phenomenon and all are part of a larger system. Together, they allow understanding and interpretation of the universe as a meta-system of interlinked wholes and organize our thoughts about the world. If a system has no input variables, it is called autonomous; otherwise, if input variables can modify the outcomes of the system, it is called non-autonomous.

Although different types of systems (from a cell to the human body, soap bubbles to galaxies, ant colonies to nations) look very different on the surface, they have remarkable similarities. At the most basic level, systems are divided into two categories:

1) Closed systems: theoretical systems that do not interact with the environment and are not influenced by its surroundings. Only the components within the system are significant. Example: a sealed jar, nothing enters or exits the jar, but whatever is inside can interact.

2) Open systems: real-world systems the boundaries of which allow exchanges of energy, material, and information with the larger external environment or system in which they exist. Example: a company, even if there are separate departments in one organization, the workers share data and interact with each other on a daily basis.

Some other differences among systems can be found in terms of determinism. Before addressing this aspect in the world of systems, it is necessary to define the system *state*. In a system, the state describes the minimum set of inner variables that are able to uniquely describe any part of the system. When a system returns to a specific state or situation, which it already visited in the past, no differences can be found between the two situations. Therefore, two identical systems with the same cannot be distinguished. Of course, not all the systems have inner states. If a system has no inner states, it is called a *0*-order system, and the outputs depend only on the input values. An example of a system like that is, for instance, the motion of a particle in the space: if the particle velocity doubles, the time to cover a certain distance is halved, if the velocity is three times greater, the time is three times smaller and so on. No information about what the particle is doing is necessary to calculate the relationship between space and time. Otherwise, the presence of inner states in some way gives the system a sort of memory of the past: what happens now depends on the inputs and also on what the system did previously. These kinds of systems are called *N*-order systems, where *N* is, in some sense, the amount of memory the system beholds. Since a system's evolution over time depends on the inputs and the past, the future outcomes of the system should also be determined by these two elements. In deterministic systems, the past and the future evolution over time is determined uniquely for a specific input. This means that if the inner state of a system is known and the input sequence in time is given, every future evolution of the system will be known and defined. For instance, in thermodynamics, it is sufficient to know the volume, pressure and temperature of a perfect gas in order to define the average velocity of gas molecules and their global position or, in classical physics, it is required to know the position and velocity of a solar system planet as system state variables to determine every future position and velocity of the planet. From the mathematical point of view, a differential equation form represents a continuous time-deterministic system:

$$\frac{dx}{dt} = F(x; y(t)) \tag{5.1.2.1}$$

where x is a vector containing all state variables of the system and y is a vector describing inputs (explicitly depending on time). If y is zero, the system is autonomous. The same relation in discrete time is:

$$x_{n+1} = G(x_n; y_n) \tag{5.1.2.2}$$

F (or G in discrete time domain) is the operator linking the rate of variation of the system variables to the present state, and it can be either linear or non-linear. A linear operator L fulfils the characteristic:

$$L(ax + by) = aL(x) + bL(y) \tag{5.1.2.3}$$

while a non-linear operator does not (Birkhoff & Rota, 1978).

In general, the evolution in time of linear differential equations is completely determined and can be calculated by means of well-established mathematical techniques. Conversely, as will be shown later, non-linear differential equations do not have a general solution mechanism and in most of cases, do not admit analytical solutions. Anyway, several mathematical and geometrical techniques were developed to define the long-term evolution of this kind of equations and outline the global behaviour of the differential dynamical system.

An efficient way to graphically show the behaviour of such systems is to draw the time evolution on a graph, with axes defined by means of the state variables of the differential equation. In general, dynamical systems can behave in any of the ways depicted in Figure 5.1.2.1.

Figure 5.1.2.1a. converging states.

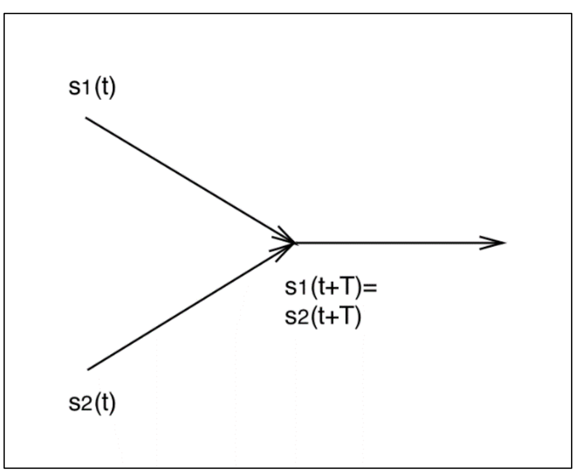


Figure 5.1.2.1b. independent states evolution.

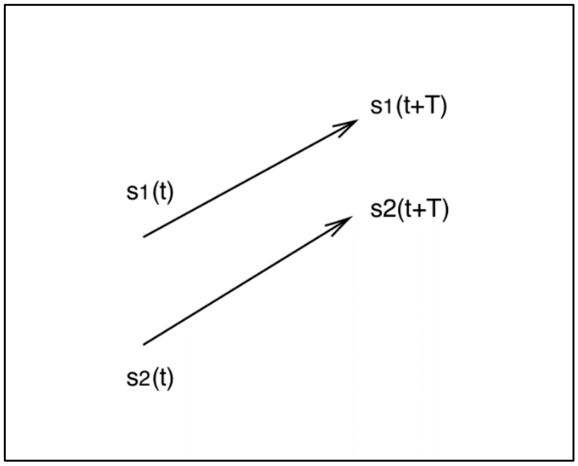
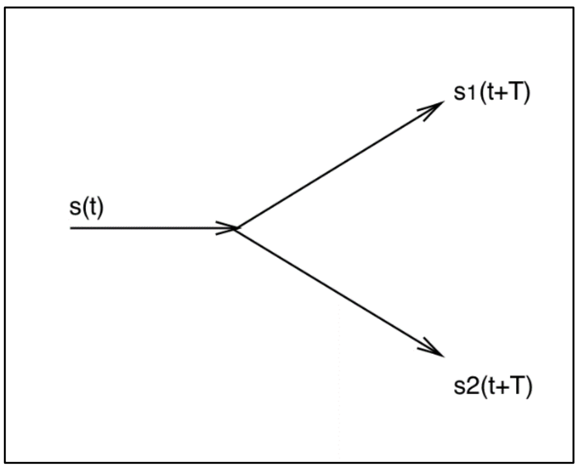


Figure 5.1.2.1c. diverging states.



The first and the third example in Figure 5.1.2.1 show two cases where the deterministic feature is lost in past evolution (a) and in future evolution (c). In fact, in (a) case that starts with two different states, s1 and s2, there will be no possibility to discriminate the original state (information lost in the past) after some time. In (c) case, conversely, a single state evolves into two different states, and there is no rule to determine the path that the evolution should follow. The only deterministic behaviour is described in (b) where two distinct states evolve over time with no common point.

Autonomous differential systems have steady states if there exists some combination of $\boldsymbol{x}$ variables where $\boldsymbol{F}(\boldsymbol{x})=0$. In these points, the variation of $\boldsymbol{x}$ is null and the system will keep this steady state until some perturbation is applied from the external environment (input). Evolution to (or from) a steady state can be described on two-dimensional state graph, as illustrated in Figure 5.1.2.2.

Figure 5.1.2.2a. Steady states in two dimensions. Real attractive (or repelling) fixed point.

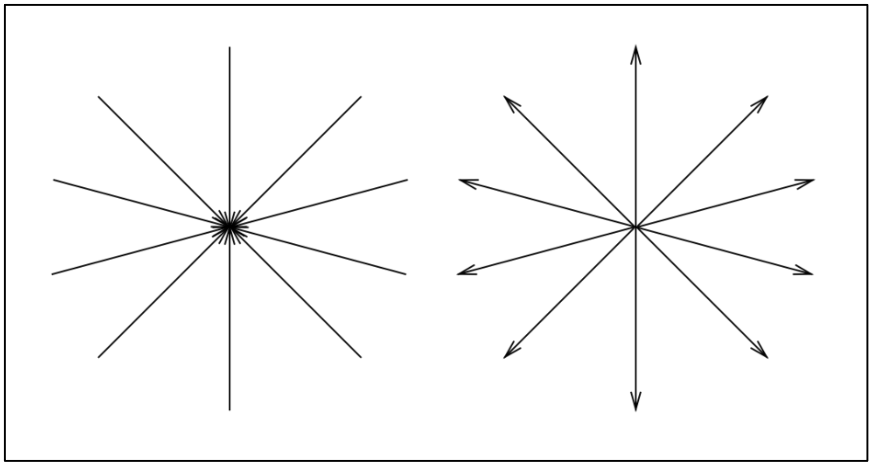


Figure 5.1.2.2b. Complex attractive (or repelling) fixed point.

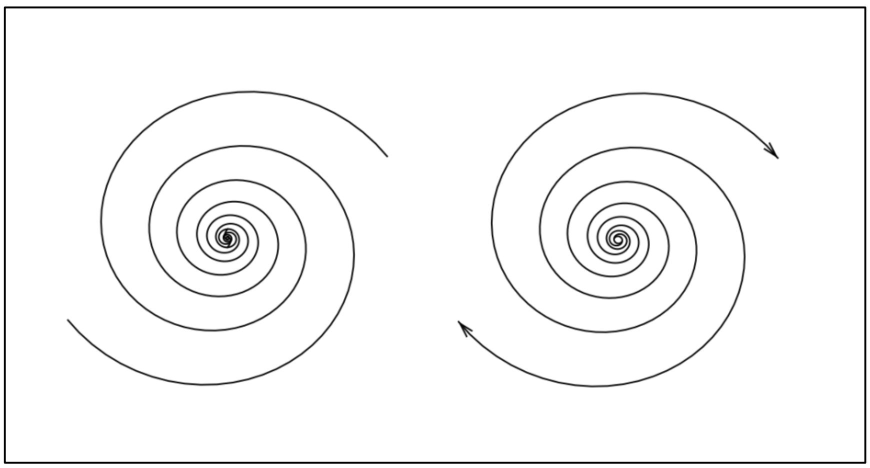
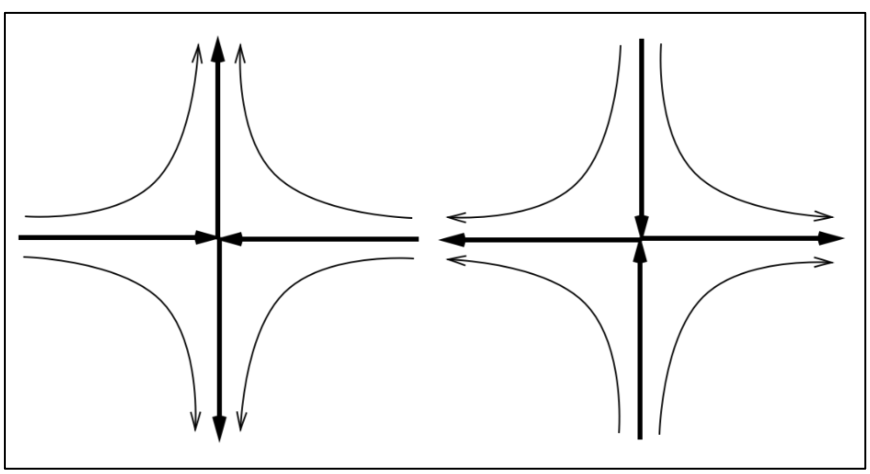
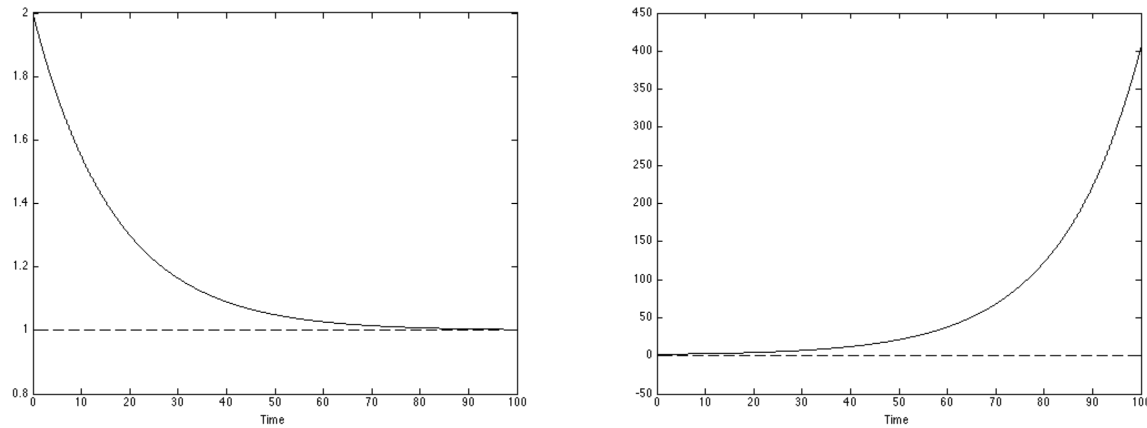


Figure 5.1.2.2c. Saddles.



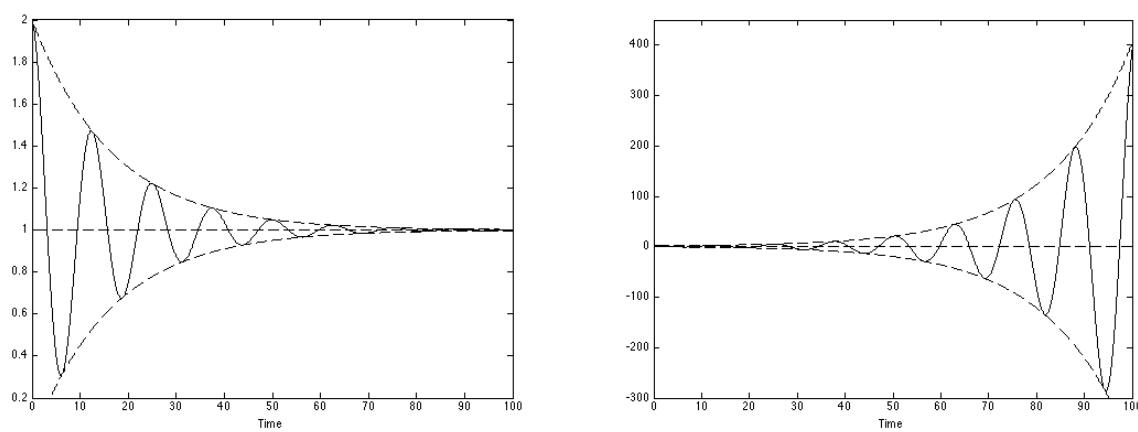
Steady states in two dimensions depicted in Figure 5.1.2.2a represent trajectories that converge or diverge toward the fixed point in the centre with a real exponential behaviour (Figure 5.1.2.3a).

Figure 5.1.2.3a. Time evolution of a system single variable: (a) real eigenvalues convergence or divergence (fixed point x=1).



In Figure 5.1.2.2b, a complex exponential behaviour rules the attraction (or repulsion) law; the case of attractive behaviour is usually called damped oscillation (Figure 5.1.2.3b).

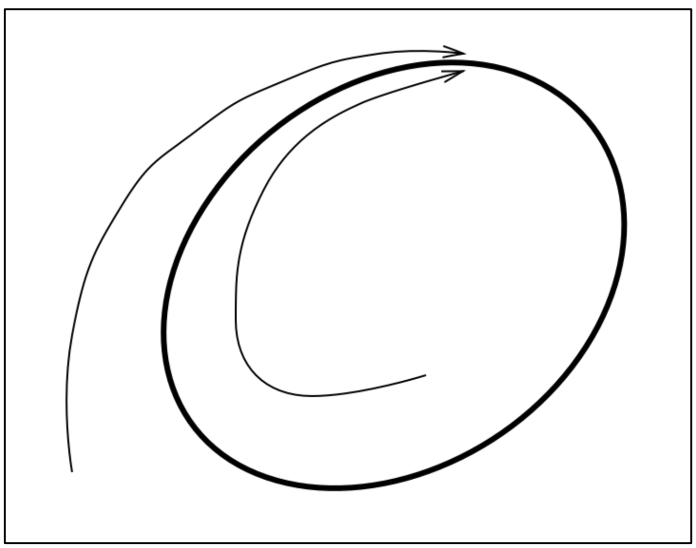
Figure 5.1.2.3b. Time evolution of a system single variable: Complex eigenvalues convergence or divergence.



The left side of Figure 5.1.2.2a and 5.1.2.2b represents the stable case: every initial state around the steady state will converge toward it. Moreover, the system would force any small perturbation applied to the steady state to move again toward the steady state. Conversely, the right side represents the unstable behaviour: any small perturbation would turn the future state away from the fixed point. The third case (Figure 5.1.2.2c) represents the last possible behaviour in two dimensions: the saddle case. In bold are presented the main contracting and diverging directions (eigenvectors) where the fixed point cannot be considered as stable since, again, any small perturbation would send the state far from the steady state. Therefore, the stability of the fixed points is described by the dynamical behaviour of the surrounding space. The local space can be studied by a linearization of the dynamic system, and the general behaviour of the system around the fixed point can be evaluated by means of the main directions of convergence or divergence (eigenvectors) and their associated eigenvalues (as in Figure 5.1.2.2c) (Robinson, 2004).

In two-dimensional space, another classical behaviour, the periodic orbit, can appear, as shown in Figure 5.1.2.4.

Figure 5.1.2.4. An attractive periodic orbit.

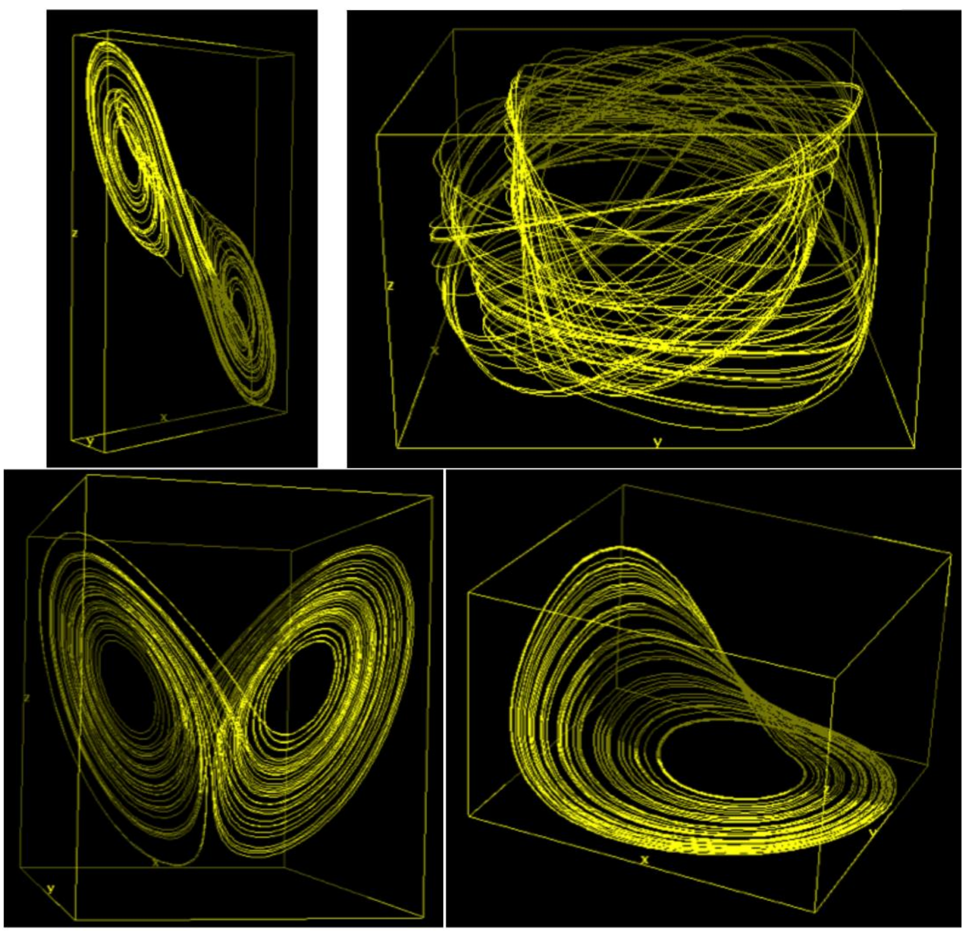


As in previous cases, the limit periodic cycle is attractive for any initial neighbouring state and after a transitory time, a periodic evolution is kept for any future time. A system like this is, for instance, the pendulum clock, where the 1 Hz period is reached independently of the initial state (angular position and velocity) of the pendulum.

Linear differential equations, as dynamic outcomes, can produce only fixed stable or unstable points and oscillations instead of more complex geometric objects (both in two or higher dimensions). Conversely, non-linear differential systems can show a greater amount of time evolutions, some of which are definitely more strange and difficult to deal with.

In three or more dimensions, all previous cases can appear but additional behaviours may be added to the geometric taxonomy of *attractors* (Guckenheimer & Holmes, 1983; Jordan & Smith 2007; Khalil, 2001). An attractor is a set of points in the phase space where all trajectories starting in a sufficiently close state will converge. The set of all points fulfilling this request is called basin of attraction. Therefore, the attractive fixed points and orbits shown in two-dimensional examples are attractors. As mentioned in the previous part, since the dynamic evolution is considered deterministic, two different trajectories cannot intersect each other to preserve the uniqueness of the future system evolution. Starting with this consideration, one may ask what kind of new attractors may emerge from a high-dimensional non-linear system. Around 1970, physicists and computer scientists encountered a special kind of attractors that, even if they were describing a deterministic system, they could not forecast the long-term evolution (or limit behaviour) unless considering a new geometrical object called fractal. This kind of time evolution of a system was named *chaos*. Some examples of chaotic attractors are shown in Figure 5.1.2.5.

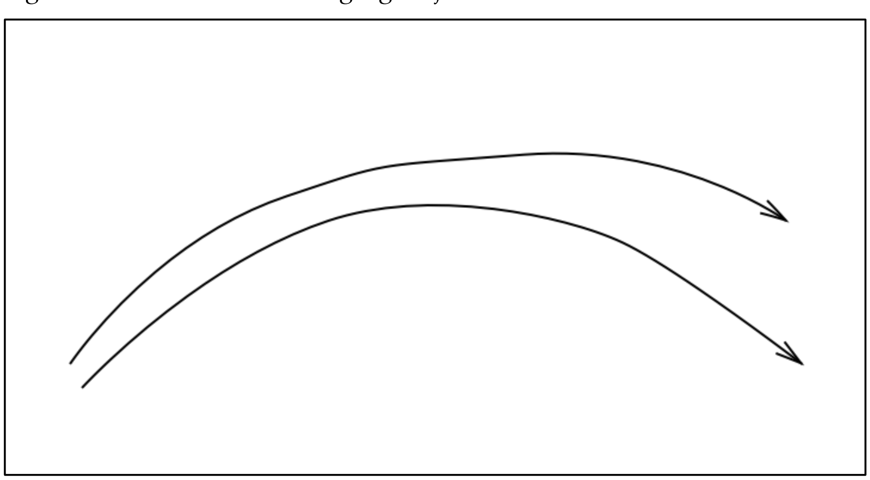
Figure 5.1.2.5. Chaotic attractors. From top left to bottom right: Chua, Duffing, Lorenz, Rossler.



As can be noted from the pictures, a chaotic attractor shows a geometrical form similar to a ball of thread. Trajectories pass very close to each other but they never intersect, preserving the deterministic nature of the system. It can be proved that trajectories, belonging to the chaotic attractor, do not fill the space in which they are embedded in a uniform way. In previous cases, an attractive fixed point has a dimension equal to zero, an orbit has a dimension equal to one (length), surfaces are two-dimensional, volumes three-dimensional, and so on.

Chaotic attractors have a non-integer dimensionality, since they do not fill the space uniformly and densely. For instance, the Lorenz attractor has a geometrical (Hausdorff) dimension equal to 2.06. It means that the trajectory fills the space more than a 2-dimensional surface, but the density of points is not sufficient to fill the space as a dense volume. This is the reason why these attractors are called strange or fractal. Another feature characterizing the strange attractors is the local divergence of close trajectories. Because of the geometrical aspects of this kind of strange objects, two close initial states are expected to move away from each other with an exponential law of divergence. The rate of local divergence is measured by the so-called Lyapunov exponent (Barreira & Pesin, 2007). Therefore, even if the chaotic attractor geometrically describes the global behaviour of the system and the trajectory remains in that part of the space, when the system explores a state, which is close to another one visited in past, its evolution is expected to be very different after some time. The effect of diverging trajectories is called, by using a metaphor, the *Butterfly Effect*. This effect explains the dependence of the system evolution on small indetermination of the initial state. As a matter of fact, the calculation of a dynamic system time course requires the knowledge of the initial state with infinite precision. If either a small perturbation or simply a rounding operation were applied to the initial state, the future evolution of the trajectory would be expected to diverge from the predicted one. This effect is depicted in Figure 5.1.2.6.

Figure 5.1.2.6. Chaotic diverging trajectories.



If d(0) is the initial distance between the two initial states, after a time T, the distance is expected to increase exponentially:

$$\boldsymbol{d(T) \approx d(0)exp(\lambda T)} \tag{5.1.2.4}$$

As mentioned previously, the averaged increasing distance is put in relation with the Lyapunov exponent Λ, which is:

$$\boldsymbol{\Lambda = \lim_{t\to\infty} \lim_{\delta x_0 \to 0} \frac{1}{t} ln \frac{|\delta x(t)|}{|\delta x_0|}} \tag{5.1.2.5}$$

Therefore, the Butterfly Effect describes the fundamental importance of small perturbations of the knowledge of the initial states. The name of the effect, coined by Edward Lorenz, is derived from the theoretical example of a hurricane's formation being contingent on whether or not a distant butterfly had flapped its wings several weeks before (Lorenz, 1996).

Finally, another feature characterizing the chaotic attractors is that a chaotic evolution is neither periodic nor quasi-periodic (i.e., sum of several periodic evolutions the frequencies of which have irrational ratio). Therefore, chaotic evolutions are hardly distinguishable from random evolutions, and the time series coming from chaotic systems may be misinterpreted as unpredictable noise. The power spectrum of chaotic signals reveals continuous dense zones, similarly to noisy and weakly self-correlated systems.

According to the existing literature, non-linear dynamic systems are deterministic but manifest their time evolution in a way that is very difficult to describe, analyse, and predict. Long-term prediction is to be fully excluded, even if the deterministic machine gives the possibility to extract some useful and interesting parameters to identify the systems (Ruelle, 1989).

Complexity can therefore be summarized by mixing the following factors: high number of dimensions (or descriptive variables), non-linearity in description of differential equation systems, some noise, which may come naturally from environment, from exclusion of any marginal aspect of the system description, or from measurement errors. Complex systems are therefore characterized by strange, non-periodic, unpredictable time evolution, strong inter-relation among variables, sensitivity to initial condition, and difficult discrimination by noisy non-deterministic phenomena. One may ask the reason why it is so interesting to define, identify, analyse and understand complex systems. The answer lies in the fact that most natural systems are ruled by non-linear differential equations. When these systems are non-autonomous and admit inputs from external stimuli, a very complex evolution may be difficult to define: the amount of chaos may change over time and the understanding of these phenomena becomes difficult. The traditional tools as statistics or classical mathematical approaches can fail to give sufficient information about the nature of what was observed. It has been proved that weather prediction (Lorenz, 1963), socio-politic systems (Campbell & Mayer-Kress, 1991; Pere, Plaza, Font, & Régis, 2006), economic markets (Guégan, 2009), stocks (Levy, 1994), currency markets (Chorafas, 1994), biological and ecological natural systems (Stone & Ezrati, 1996), among others, are ruled by chaotic equations that, even with a small set of variables, can show complex and unpredictable evolution.

1.3. Predictable Systems

Predicting a system means, in general, tries to forecast or predict a state of system that is not described by measured data and observations. The operation of prediction may involve both the temporal and/or the spatial aspect. A temporal prediction is usually related to forecasting field while the spatial prediction is sometimes called static model. Naturally, in several cases, these two aspects may coexist and the algorithmic tool used in the analysis has to be sufficiently powerful to process both of these aspects. In some sense, a static model is an approach that does not need any memory about the past; its present state depends only on the input stimulation (even when the input and the output are in high dimensional space). Anyway, these two approaches can be connected when the system is non-autonomous. The dependence on the past may be reconstructed in a static model by means of the temporal windowing. The input of such a system is not only representing the present input state, but the input vector may increase in dimension by adding the past values of the input variables. The amount of past or delayed values strongly depends on the geometry of the system under analysis and on the amount of memory needed to describe the system. Therefore, in order to describe dynamical systems, it is possible to build models with inner dynamics by conjecturing the suitable internal dynamical structure of the modelling system or static non-linear models with delayed inputs.

The concept of predictability is strongly connected to determinism. If a system is deterministic, the description of a unique temporal evolution has to fulfil some linear or non-linear differential equation system. A specific state of the evolution cannot lie on different trajectories in phase space, that is, it cannot fulfil the solutions of the equations that describe two or more trajectories.

A particular aspect that differentiates linear and non-linear systems often regards the influence of small perturbations on the actual state. As viewed in the previous section, chaotic systems are strongly affected even by microscopic deviation of the state under analysis. The natural behaviour of a chaotic system is to exponentially separate the close states while they evolve in time. This separation is intended to result from a small perturbation in case of light noise or the inability to detect or measure the initial state with an infinite amount of precision. Such repulsion among close trajectories is limited to a local behaviour of the system, since the same attractor belonging to evolution for sufficient time gives the possibility of new rapprochements. Conversely, in linear and stable cases, a small deviation in detecting the initial state will retain the same order of magnitude, that is, close trajectories will remain sufficiently close to each other while the time increases.

Therefore, the ability to predict the future evolution of a system strongly depends on the linear or non-linear nature of the system itself. Traditional classic mathematical and statistical approaches often fail to give a global representation of the system, although they may give a sufficient precise description for small time evolution or space volumes (linearization approach). As a matter of fact, any deterministic evolution, whether linear or not, can be locally represented by means of a linear tangent space representing the state and its neighbourhood, which were originally linear. An approach like that describes the system sufficiently well only if the representation of the system is limited to the so-called small variations. These approaches are therefore based on linear techniques that, according to their nature, can put in relevance the linear aspects of the system under analysis: global aspects in case of linear systems, local aspects for non-linear systems.

Some very popular algorithms belonging to this class are the following:

- **ARMA method**: This Auto Regressive Moving-Average method is a type of linear mathematical model that provides instant by instant an output value based on the previous incoming and outgoing values. The time evolution is described by the estimation of the linear parameters that are used in the model repeatedly. This approach can be considered useful when pursuing the forecasting of time series, that is, when forecasting the new outcome of the system when the present and the past states are known and the hypothesis of weak variability of inner parameters can be assumed. A global description of the system is out of the ability of ARMA method (Box, Jenkins, & Reinsel, 2008).
- **Kalman filter**: Also known as Linear Quadratic Estimation (LQE) filter, Kalman filter algorithm uses a series of measurements that contain noise (random variations) and other inaccuracies observed over time and produces estimates of unknown variables that tend to be more precise compared to those based on a single measurement alone. More

formally, the Kalman filter operates recursively on streams of noisy input data to produce a statistically optimal estimate of the underlying system state. From a theoretical standpoint, the main assumption of the Kalman filter is that the underlying system is a linear dynamical and that all error terms and measurements have a Gaussian distribution (Hamilton, 1994).

In addition to the linear approaches, which can be partially useful for the description of an undefined system, some other method classes can be mentioned as probabilistic and non-linear.

Probabilistic forecasting is based on the description of the statistical distribution of single variables involved in the system description and on their mutual statistical relation. In this framework, if any variable (or the variables with principal effects on the system evolution) can be measured and its statistical distribution may be estimated, a network of relations among the variables can be built by evaluating conditional probability connecting different variables (Bayesian belief networks). A request for this kind of approach reflects *stationarity*. A stationary process is a stochastic process of which the joint probability distribution does not change when shifted in time or space. Consequently, parameters such as the mean and variance, if they exist, also do not change over time or position. The probabilistic approach has several pros and cons. The main positive aspect of this approach is that it can provide a global description of the system behaviour since the distribution of the variables has to consider a sufficiently wide manifestation of each variable in order to be built. The whole spectrum of possibilities has to be investigated and described (even by a theoretical point of view) in statistical way. Moreover, the probabilistic approach allows including all unknown causes into a probabilistic framework and avoiding a huge amount of negligible connections in some cases. Conversely, the drawbacks of this approach lie in the necessity to formulate a priori hypotheses about the choice of descriptive variables and their deterministic relations (one-way statistical dependence), the great amount of data required for a suitable probabilistic description, and the implicit lack of precision for the state prediction.

A more specific probabilistic approach is given by the analysis of Markov processes. A Markov chain is a special sort of belief network used to represent sequences of values, such as the sequence of states in a dynamic system or the sequence of words in a sentence. A Markov chain is stationary if the transition probabilities are the same for each time point. Stationary Markov chains are of interest because they provide a simple model that is easy to specify. The assumption of stationarity often involves the natural model, because the dynamics of the world typically does not change in time. If the dynamics change over time, it is usually because some other feature could also be modelled. The network can extend indefinitely. Specifying a small number of parameters can give an infinite network. You can ask queries or make observations about any arbitrary points in the future or the past (Durrett, 2010).

Finally, the last class of methods presented in this chapter refers to non-linear algorithms and models. A very important aspect of non-linear approaches regards the possibility to build a base of non-linear functions that can approximate the target function when combined together. Note that linear systems can be considered a particular case of non-linearity. Therefore, the non-linear approach is considered a more general class of techniques in comparison with the traditional ones because, in general, the non-linear set of mathematical tools contains the linear ones as well.

In modelling and forecasting, the most popular non-linear approaches are the neural networks and, more in general, the application of artificial intelligence algorithms. Both classes of techniques contain several approaches and structures with a plethora of modern algorithms, some of which are still under testing.

1.4. Connectionism and Complexity

Complex systems represent a new approach, which studies how relationships between parts give rise to the collective behaviours of a system and how the system interacts and forms relationships with its environment. The equations from which complex system models are developed generally derived from statistical physics, information theory, and non-linear dynamics, and represent organised but unpredictable behaviours of systems of nature that are considered fundamentally complex. The physical manifestations of such systems cannot be defined; thus, the usual choice is to refer to "the system" as the mathematical information model without referring to the undefined physical subject that the model represents.

The key problems of complex systems are difficulties with their formal modelling and simulation. From such a perspective, in different research contexts, complex systems are defined based on their different attributes. Since all complex systems have many interconnected components, the science of networks and network theory are important aspects of the study of complex systems. A consensus regarding a single universal definition of complex system does not yet exist.

For systems that are less usefully represented with equations, various kinds of narratives and methods are used to identify, explore, design and interact with complex systems.

Some definitions of complexity focus on the question of the probability of encountering a given condition of a system once characteristics of the system are specified. Complexity of a particular system is the degree of difficulty in predicting the properties of the system, given the properties of the system's parts (Weaver, 1948). In Weaver's view, complexity comes in two forms: disorganized complexity and organized complexity. Disorganized complexity results from the particular system having a very large number of parts, say millions of parts, or many more. Although the interactions of the parts in a disorganized complexity situation can be seen as largely random, the properties of the system as a whole can be understood by using probability and statistical methods. Organized complexity, on the other hand, resides in nothing else than the non-random, or correlated, interaction between the parts. These correlated relationships create a differentiated structure that can, as a system, interact with other systems. The coordinated system manifests properties not carried or dictated by individual parts. The organized aspect of this form of complexity can be said to "emerge" without any "guiding hand". The number of parts does not have to be very large for a particular system to have emergent properties. The properties of a system of organized complexity may be understood through modelling and simulation conducted particularly with computers.

A very important aspect of complexity can be found in the field of *connectionism*. Connectionism comprises a set of approaches in artificial cognition modelling that models mental or behavioural phenomena as emergent processes of interconnected networks of simple units. The key word linking complexity and connectionism is "emergence" because the strange and complex phenomena that may arise from non-linear world are, in some sense, unexpected from the point of view of classical system analysis. For instance, the complex behaviour emerged in Lorenz model of weather was so unexpected that the author himself was convinced that it was an error in the implementation of the algorithm. The non-linear relationships between weather single elements and between the neural cells in the brain have in common the possibility of the emergence of unexpected and extremely interesting behaviour. The interesting part of complexity in brain structures is well known, as it involves the emergence of efficient approaches to solve difficult tasks that traditional algorithmic techniques fail to describe even the simplest cases. In the last years, several problems have been addressed using techniques inspired by natural connectionism: face recognition (Le, 2011), language recognition (Cole 1989), automatic robot guidance (Gowdy, Pomerleau, & Thorpe, 1991), pattern recognition (Ripley, 1996), economic prediction (White, 1988), and many others.

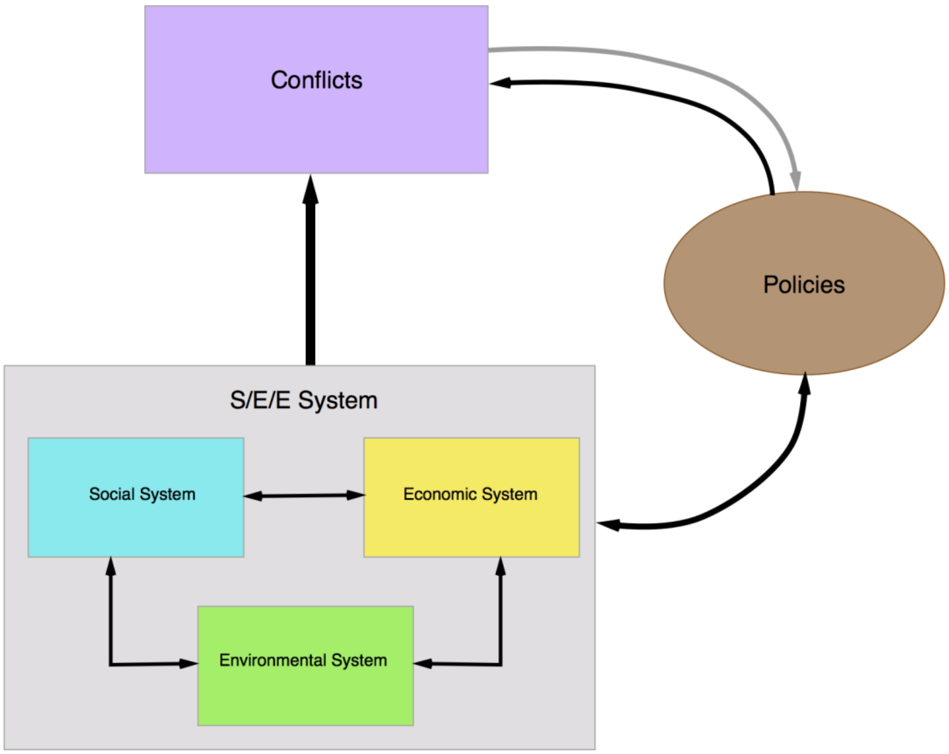
Another aspect of connectionism related to complexity is the network of interconnected simple units. Any interconnected structure of dialoguing elements that influence the future is related to the behaviour of some set of neighbour elements of the same kind and is likely to show complex behaviour in its time evolution. Once again, such a complex behaviour is given either by the eventual non-linear relationships among elements, by their inner non-linear dynamics, or by the great amount of elements synchronically evolving in time. In brain, for example, the complex dynamics can be measured in several cognitive states but, at the same time, some sort of cooperative coherence can be relevant depending on the task that the specific cortex area is performing. Different kind of coherence and different kind of chaotic evolution can relate to different kind of cognitive states and perceptions.

According to what previously described, complexity is an attribute of connectionist systems, as the spatial distribution complexity increases complexity emerges more easily. Therefore, simple non-linear processing units connected to each other according to some defined rule can be considered as the fundamental elements for building a complex system the behaviour of which may reflect the complexity of a target system under the investigation (Sporns, Tononi, & Edelman, 2000).

1.5. The SECOA System

The SECOA project concerns the analysis of coastal conflicts in relation to socio-economic-environmental (S/E/E) variables using geo-referenced data. The degree of conflict aspects on the one hand and the spatiotemporal distribution of S/E/E variables on the other hand represent the manifestation of two well-defined non-autonomous and potentially non-linear dynamical systems. In SECOA, conflicts, intended as the manifestation of some specific social and political features, are hypothesized to be produced and characterized by the territorial S/E/E conditions of the surrounding area. Up to now, the connection between conflict and territorial S/E/E aspects has not yet been explored by means of the dynamical system analysis although several disciplines, such as sociology, geography and economy, explored the dynamics of social and economical aspects of territories using classical mathematical and algorithmic tools (Montello *et al.*, 2003; Tobler, 2004; Yan Li, & Shanmuganathan, 2007). As mentioned in previous sections, economic and social evolutions are connected with non-linear world, making it possible to build a conceptual scheme representing the relations among the elements involved in the global SECOA system, as shown in Figure 5.1.5.1.

Figure 5.1.5.1. Conceptual Scheme of Dynamic SECOA System.



Social, economic and environmental sub-systems are tightly connected with each other to reflect the specific territorial developmental dynamics. Solid black double-arrow lines describe the connections among them. The set of S/E/E variables and their relationships form the S/E/E system. The S/E/E system is without any doubt a dynamic system, and some non-linear aspect in their inner evolution and in their mutual relationship may be assumed in accordance with the bibliography and past studies.

Conflicts, on the other hand, are supposed to be generated by several causes but, as mentioned before, it is expected that every specific form a conflict assumes in the associated territory is basically ruled by the S/E/E conditions and the adopted policies to address and solve the conflict itself. Moreover, policies are strictly connected with the S/E/E territory structure that can be influenced by long-term policies and can influence the development of appropriate policies for the specific territory. Therefore, it is more accurate to consider a dynamic system whose evolution is determined by the combined and interconnected effect of S/E/E and policies domains (curved solid black double-arrows lines between S/E/E and policies systems in Figure

5.1.5.1). Finally, in conflict temporal evolution, some sort of feedback from conflict features and adopted policies should be considered, as the progressive modification of conflict due to applied policies may affect and calibrate different approaches by local institutions (grey solid line in Figure 5.1.5.1). Anyway, this last aspect will not be considered in SECOA since the temporal aspect of sampled data did not allow for a suitable description of such phenomena.

In synthesis, the previous SECOA scheme can be depicted as in Figure 5.1.5.2.

Figure 5.1.5.2. Schematic SECOA system.

In Figure 5.1.5.2, S/E/E and policies systems were combined together to form a single interacting system, the outcomes of which drive the conflict aspects. The black spiral into the S/E/E + Policies box symbolizes the non-linear dynamical nature of the system.

Considering the non-linear aspect of such systems and the complexity they may manifest in their spatial and temporal evolution, the algorithmic approach to identifying and modelling the whole SECOA system was based on the adoption of a connectionist approach to describe the specific cases studied in the SECOA project and their mutual relationships (taxonomy) as well as to model the causal relationships between S/E/E + policies features (independent variables) and the conflict characteristics.

2. Reductionist Approach: The Neuron Model

2.1. Introduction

As described in the previous chapter, a simple unit is the basic element used in a connectionist system. The simplicity of this unit is intended to be at functional level. In nature, brain is the most complex connectionist system, and the units that compose it are a particular kind of biological cells called *neurons*. As it will be described in the next section, a neuron is a small device that can perform a very simple processing of input stimulations and, in turn, can transmit its internal state to other neurons. The dynamics performed by the network of neurons, even if the single processing can be considered simple and may show a very complex, rich and amazing emergent global behaviour. The cooperation or competition among neurons may reveal an extraordinary organized complexity that, under some circumstances, may reflect the complex dynamics of a target system. This makes a network of neurons a strong and powerful tool for modelling a system by observing its behaviour and for building a suitable neural network of which spatio-temporal evolution should be as similar as possible to the system under analysis.

A neural network derives its computing power through its massively parallel distributed structure and its ability to learn and therefore to generalize. Generalization refers to the neural network producing reasonable outputs for input that were not encountered during the training phase. These capabilities make neural networks able to solve complex and large-scale problems that are currently intractable. The use of neural networks offers the following useful properties and capabilities:

- *Nonlinearity*. A neural network composed of an interconnection of non-linear units is itself nonlinear. Moreover, the nonlinearity is distributed throughout the network. Nonlinearity, as expressed in previous chapter, is a highly important property; particularly, if the underlying mechanism that is responsible for generating input signal is inherently non-linear.
- *Input-Output Mapping*. As will be shown later, a popular paradigm of learning, called *supervised learning*, involves modification of the synaptic weights of a neural network by applying a set of *training samples*. Each sample consists of a unique input signal and a corresponding desired response. The synaptic weights of the network are modified to minimize the difference between the desired and the actual response. The training of the network is repeated for many sample/response couples until the network reaches a steady

state where there are no further significant changes in the synaptic weights. Thus, the network learns from the examples by constructing an *input-output mapping* for the problem. Note that no prior assumptions are made for the input data of the statistical model. This approach is very similar to nonparametric statistical inference methods, from the conceptual point of view.

- *Adaptivity*. Neural networks have a built-in capability to adapt their synaptic weights (or internal free parameters) to changes in the surrounding environment. In particular, a neural network trained to operate in a specific environment can be easily retrained to deal with minor changes in the operating environmental conditions. Moreover, in the presence of a non-stationary environment (where statistic changes with time), a neural network can be designed to change its synaptic weights in real time. To realize the full benefits of adaptivity, the principal time constants of the system should be long enough to ignore spurious disturbances and yet short enough to respond to meaningful changes in the environment. The problem described is referred to as the *stability-plasticity dilemma* (Grossberg, 1988).
- *Evidential Response*. In pattern classification, a neural network can be designed to provide information not only about which particular pattern to select, but also about the confidence in the decision. The latter information may be used to reject ambiguous patterns and thereby improve the classification performance.
- *Contextual Information*. In a neural network, knowledge is represented by the structure and the activation state of the network. Every neuron of the network is potentially affected by the global activity of all the other neurons. Consequently, a neural network deals with contextual information naturally.
- *Fault Tolerance*. A neural network has the potential to be inherently fault tolerant or capable of robust computation. For example, if a neuron or its connecting links are damaged or removed, the quality of the recall of stored information is impaired. Due to the distributed nature of information stored in the network, the damage has to be extensive before the overall response of the network is degraded seriously.
- *Neurobiological Analogy*. The design of a neural network is motivated by analogy with the brain, which is a living proof that parallel processing is not only physically possible but also fast and powerful. Neurobiologists look to neural networks as a research tool for the interpretation of neurobiological phenomena. On the other hand, physicists look to

neurobiology for new ideas to solve problems that are more complex compared to those based on conventional techniques.

The next sections address the description of the functionality of neurons and the functions that a neural network may perform. Several algorithms designed to build a network able to perform specific tasks will be described along with the topological structures and the definition of internal parameters of a neural network.

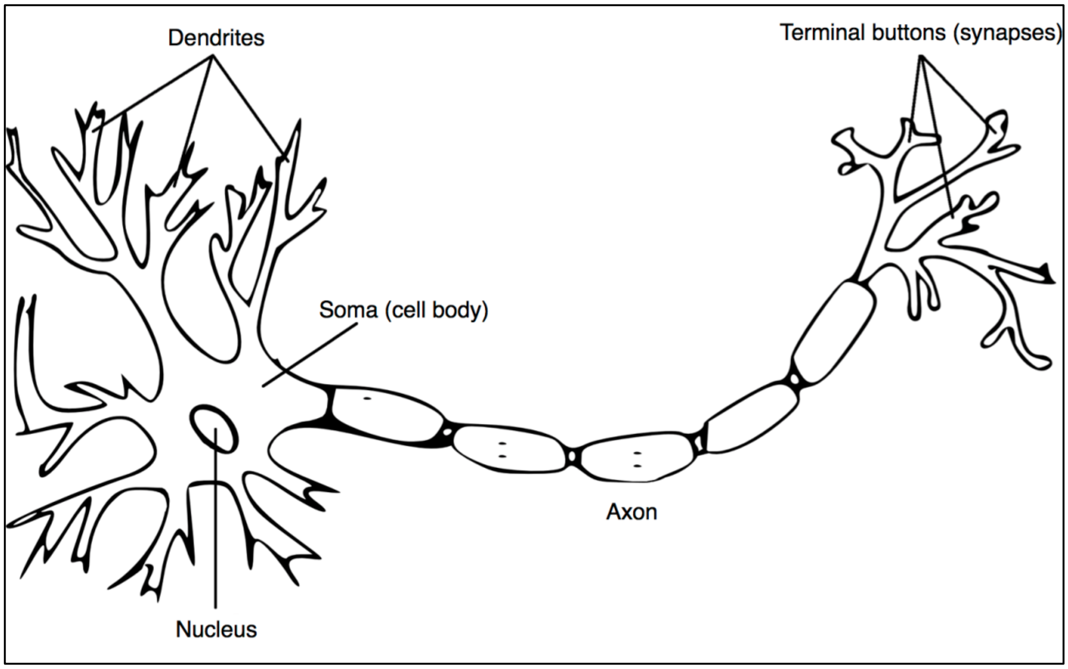
2.2. The Biological Neuron

The simple units that comprise a neural network are called artificial neurons, whose behaviour is based on the biological neurons by means of the functions performed by the latter operating in their natural environment. What we know about biological neurons is due, among the others, to the pioneering work of Ramón y Cajál (Ramón y Cajál, 1911) who introduced the idea of neurons as structural constituents of the brain. Typically, neurons are rather slower than silicon logic gates, but the brain makes up for the relatively slow rate of operation of a neuron by having a truly staggering number of neurons with massive interconnections between them. It is estimated that there are approximately 10 billion neurons in the human cortex and 60 trillion synapses or connections. The result is that the brain is an enormously efficient structure.

Synapses are elementary structural and functional units that mediate the interactions between neurons. The most common kind of synapse is the chemical synapse. When a presynaptic process liberates a transmitter substance (neurotransmitter), it diffuses across the synaptic junction between neurons and then acts on a postsynaptic process. Therefore, a synapse converts a presynaptic electrical signal into a chemical signal and then back into a postsynaptic electrical signal. In terms of physics language, a synapse operates as a one-directional gate in which information or signals may flow in only one direction. A synapse can have excitatory or inhibitory function on the receptive neuron but not both.

Modification of synaptic configuration is called *plasticity* in neurobiology. Plasticity permits the developing nervous system to adapt to its surrounding environment. In an adult brain, plasticity can operate by means of two mechanisms: the creation of new synaptic connections between neurons and the modification of existing synapses. The former part will be implemented in the phase of building the structure of an artificial neural network while the second part will be used in the training phase of a neural system. Bioelectrical signals reach the synaptic zones, flowing into a special transmission line called axon. Axon is the unique output of a neuron, and the signal flowing into it is supported without leakage by the axonal transmitting system until it reaches the synaptic terminals. As mentioned before, a given amount of neurotransmitters is released and by diffusion, the neurotransmitter molecules reach the receptive sites of the postsynaptic neurons in specific neural structures called *dendrites* (Figure 5.2.2.1).

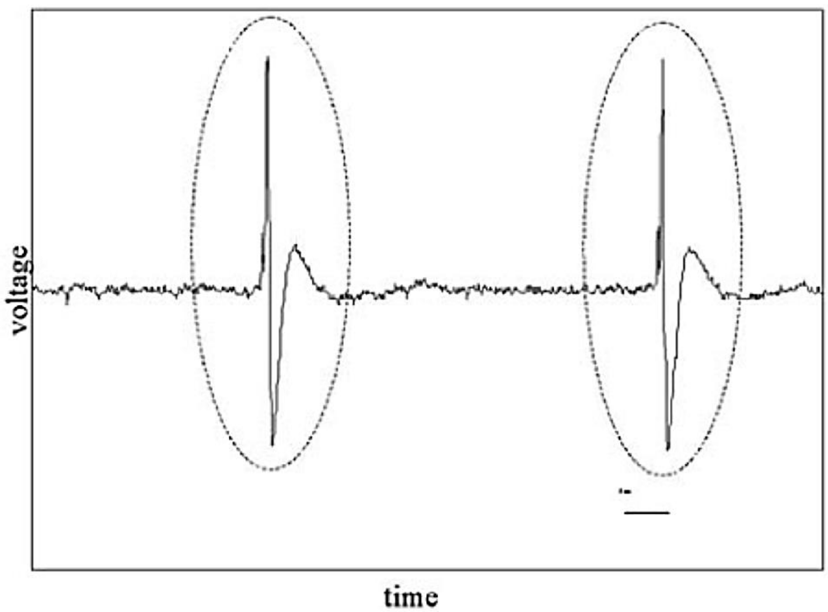
Figure 5.2.2.1. A neuronal Cell.



To give an order of magnitude, an average neuron cell in the brain cortex receives about 10,000 or more synaptic contacts and can project onto thousands of target cells.

A neuron is a single biological cell with a cell body and a nucleus where all the genetic information about the organism is stored; it is filled with fluid and cellular organelles and surrounded by a cell membrane. However, the neuron is unique in having very extended finger-like processes of dendrites and axon and a real electrical mechanism for processing of incoming and outgoing signals. To enable different neurons to communicate with each other, there are little openings in the cell membrane called channels. The basic mechanisms of information processing (i.e., for integrating inputs, thresholding and communicating outputs) in a neuron are based on the movement of charged atoms (ions) in and out these channels and within the neuron itself. A key element in the electrical model of the neuron is the difference in electrical charge (voltage) of the neuron relative to its external environment. This difference is known as the *membrane potential*, because the cell membrane separates the inside and outside of the neuron. The difference in electrical charge exists across this membrane. As ions flow in and out of the neuron through channels, this changes the membrane potential. Once the cell body transitions into the axon, the membrane potential determines if the neuron will fire. The thresholded property of this firing process is due to the sensitivity of a set of special channels to the membrane potential; these channels open up only when the membrane potential is sufficiently elevated (voltage-gated channels). Inputs to a neuron are transmitted when the neurotransmitters open particular types of channels on the receiving neurons, which allow specific types of ions to flow. These trigger the electrical propagation and integration in the receiving neuron. The action of neural firing is called *spiking*, or triggering an *action potential* (Figure 5.2.2.2).

Figure 5.2.2.2. Two spikes (action potential) on a neuron axon. The negative wave after the positive front of the spike is the refractory period.



The mechanism governing the spike emission on the axon to be propagated to other neurons is ruled by two kinds of voltage-gated channels, distributed rather densely near to the axon hillcock that only become activated when the membrane potential reaches a specific threshold value. When one type of voltage-gated channel at the hillcock opens, the channels cause the neuron to become even more excited (depolarization). This further excitation causes the other type of channels to open, and these channels act to inhibit the neuron. This results in a spike of excitation followed by inhibition. When the membrane potential is brought back down to the inhibitory channels, it tends to overshoot the basic resting potential (-70 mV) slightly. This causes a refractory period following a spike, where it is unable to fire another spike until the membrane potential climbs back up to the threshold level again. This refractory period determines the fixed maximum rate at which a neuron can fire spikes.

The Hodgkin–Huxley model is an analytic model that describes how action potentials in neurons are initiated and propagated (Hodgkin & Huxley, 1952). The two researchers who introduced the model won the Nobel Prize in 1963 for investigating the functionality of a squid neuron and defining a precise and accurate mathematical description of the neural activity. In particular, the differential equation describing the time evolution of a neuron is:

$$I = C_m \frac{dV_m}{dt} + g_k(V_m - V_k) + g_{Na}(V_m - V_{Na}) + g_l(V_m - V_l) \tag{5.2.2.1}$$

where V_m is the membrane potential, I is the total membrane current per unit area, C_m is the membrane capacitance per unit area, g_k and g_{Na} are the potassium and sodium conductances per unit area, V_k and V_{Na} are the potassium and sodium reversal potential, and g_l and V_l are the leak conductance and leak reverse potential, respectively. Sodium and potassium are the main chemical substances. Their ion density inside and outside the neuron defines the membrane potential and the electrical currents entering in and out the cell through the synaptic gates. In voltage-gated ion channels, the channel conductances g_k, g_{Na} and g_k are non-linear function of both time and membrane potential while the leak conductance g_l is constant.

The associated electrical circuit of the neuron is depicted in Figure 5.2.2.3.

Figure 5.2.2.3. Electrical model of a neuron based on Hodgkin-Huxley equations.

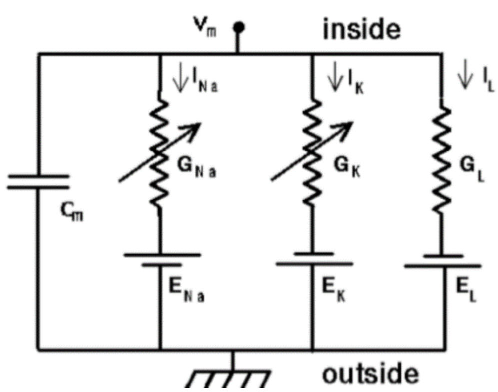
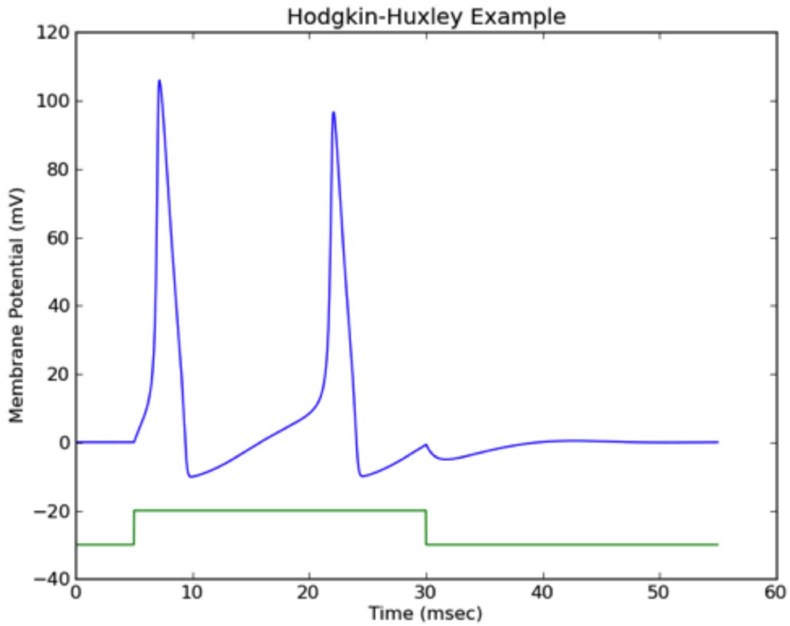


Figure 5.2.2.4 shows the time evolution of the membrane potential when a current stimulus is injected into the neuron: the Hodgkin-Huxley model put in evidence the formation of two well-defined spikes with a refractory period of about 10 msec between them.

Figure 5.2.2.4. Time evolution of membrane potential (in blue) when an activating 10 μA current is applied for 25 msec (in green).



Therefore, the basic mechanisms underlying the functioning of a neuron can be summarized as follows:

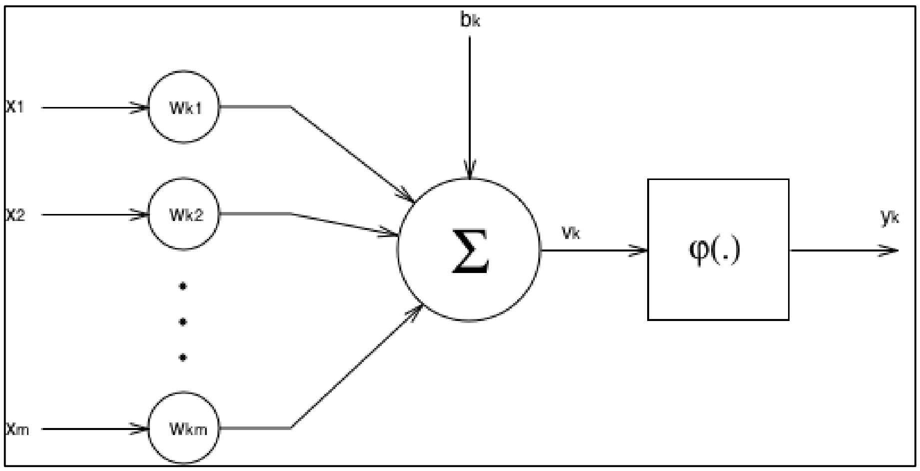
1. The external stimuli reach the neuron inputs by means of synaptic transmission. The efficiency and the nature of every synaptic site determine the amplitude of the signal read by the neuron cell.
2. All the inputs are integrated to define the internal membrane potential.
3. If the membrane potential is greater than is a reference threshold potential, an action potential is generated as a sequence of spikes that is transmitted along the axon (output channel).
4. The action potential reaches the terminations where the phenomenon of neuro-transmitters diffusion is repeated and the synaptic sites of the post-synaptic neurons can again read the neuronal stimulus at their inputs.

In the next section, a simple functional model of the neuron will be described in terms of the expressed mechanisms. Once defined, this model will be the fundamental element used to form a network of neurons with architecture that will allow the emergence of interesting and useful algorithmic approaches to non-conventional problems and modelling.

2.3. Activation Function, Bias and Synaptic Weights

A neuron is an information-processing unit that is fundamental to the operation of a neural network. The block diagram of Figure 5.2.3.1 shows the model of a neuron, which forms the basis for designing artificial neural networks.

Figure 5.2.3.1. Block model of a neuron.



Here, we identify three basic elements of the neuronal model:

1. A set of synapses, or connecting links, each of which is characterized by a weight or connection strength. Specifically, a signal xj at the input of synapse j connected to neuron k is multiplied by the weight wkj. The first subscript refers to the neuron in question and the second subscript refers to the input end of the synapse to which the weight refers. Unlike a synapse in the brain, the synapse weight of an artificial neuron may lie in a range that includes negative as well as positive values.

2. An adder (S) for summing (or integrating) the input signals weighted by the respective synapses of the neuron. The operation described here constitutes a linear combiner.

3. An activation function (j) for limiting the amplitude of the output of a neuron. The activation function is also referred to as a squashing function, since it squashes (limits) the permissible amplitude range of the output signal to some finite value. Typically, the normalized amplitude range of the output of a neuron is written as the closed interval [0,1] or alternatively [-1,1].

The neuronal model in Figure 5.2.3.1 also includes an externally applied bias denoted by b_k. The bias has the effect of increasing or lowering the net input of the activation function, depending on whether it is positive or negative, respectively. The bias is strictly related to the concept of thresholding expressed in the previous section, i.e., mathematically, the bias is simply the inverse value of the threshold ($b_k = -T_k$). In other words, if the sum of the weighted inputs is greater than the threshold, the neuron turns on and the action potential is transmitted through the axon, otherwise the neuron is silent (off).

In mathematical terms, we may describe a neuron *k* by writing the following pair of equations:

$$\boldsymbol{u_k = \sum_{j=1}^{m} w_{kj} x_j} \tag{5.2.3.1}$$

and

$$\boldsymbol{y_k = \varphi(u_k + b_k)} \tag{5.2.3.2}$$

where $x_1, x_2, \ldots, x_m$ are the input signals; $w_1, w_2, \ldots, w_m$ are the synaptic weights of neuron *k*; u_k is the linear combined output due to the input signals; b_k is the bias; j(.) is the activation function; and y_k is the output signal of the neuron. The activation potential v_k is given by:

$$\boldsymbol{v_k = u_k + b_k} \tag{5.2.3.3}$$

The bias b_k is an external parameter of artificial neuron *k*. Therefore, equivalently, we may combine equations 2 and 4 as follows:

$$\boldsymbol{v_k = \sum_{j=0}^{m} w_{kj} x_j} \tag{5.2.3.4}$$

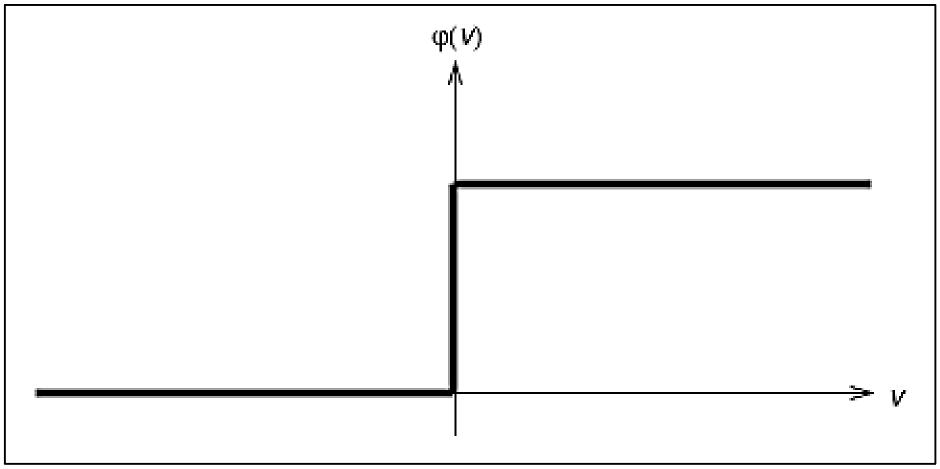
and

$$\boldsymbol{y_k = \varphi(v_k)} \tag{5.2.3.5}$$

In equation 5.2.3.4, we have added a new synapse. Its input is $x_0 = +1$ and its weight is $w_{k0} = b_k$.

The activation function, denoted by j(v), defines the output of a neuron in terms of the induced local field v. As mentioned before, a neuron can be on or off depending on whether the potential u_k given by the weighted integration of inputs is greater than the threshold. Therefore, the simplest activation function that can be imagined about an artificial neuron is depicted in Figure 5.2.3.2.

Figure 5.2.3.2. Threshold activation function.



For this type of activation function, called threshold function or Heaviside function, we have:

$$\varphi(v) = \begin{cases} 1, & if v \geq 0 \\ 0, & if v < 0 \end{cases} \quad (5.2.3.6)$$

Correspondingly, the output of neuron k employing such a threshold function is expressed as

$$y_k = \begin{cases} 1, & if v \geq 0 \\ 0, & if v < 0 \end{cases} \quad (5.2.3.7)$$

where v_k is the induced local field of the neuron; that is:

$$v_k = \sum_{j=1}^{m} w_{kj} x_j + b_k \quad (5.2.3.8)$$

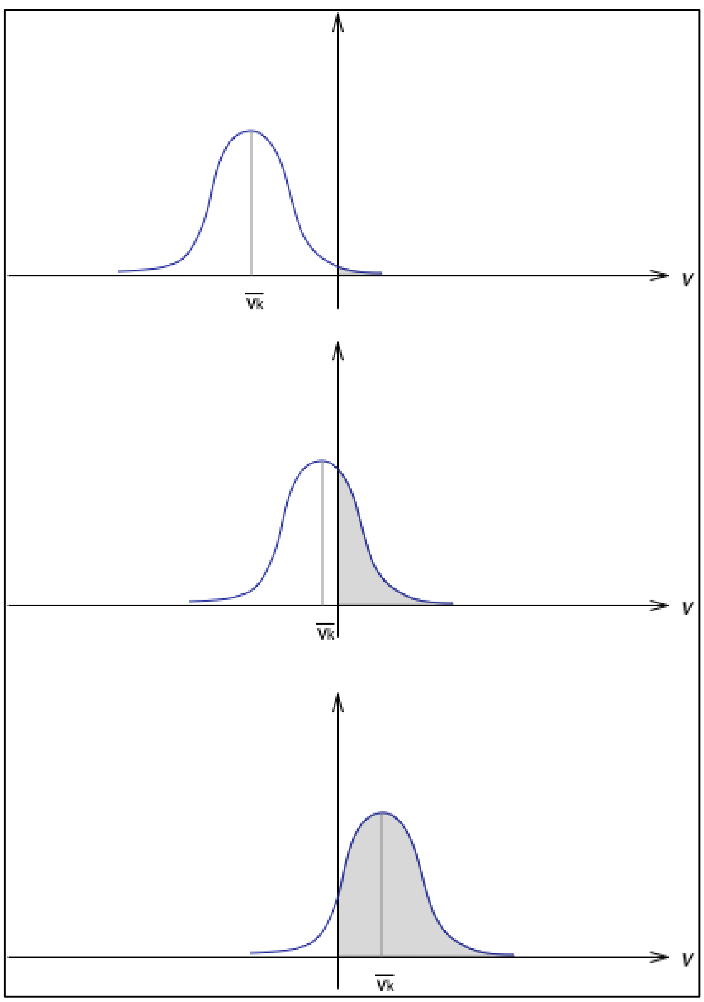
Such a neuron is referred to in literature as the McCulloch-Pitts model in recognition of the pioneering work done by McCulloch and Pitts (McCulloch & Pitts, 1943). In this model, the output of a neuron takes on the value of 1 if the induced local field of that neuron is nonnegative and 0 otherwise. This statement describes the *all-or-none property* of the McCulloch-Pitts model.

As stated in Amit (Amit, 1992), some unexpected perturbations may influence the output of a neuron. Basically, several sources of incoherent mechanisms may be identified in the field of biological processes of neurons. These perturbations may be due to small fluctuations of neurotransmitter densities in synaptic vesicles, by the quantized aspect of neurotransmitter molecules, and by unpredictable fluctuations of biological elements, as for instance hormones, in the area where the neuron is functioning. The total influence of these unpredictable causes of noise follows a Gaussian statistical distribution. The effect on the neuron functionality of such perturbations makes activation possible, at least to some extent, even if the neuron internal potential is lower than is the threshold. More specifically, the probability that a neuron is active is given by:

$$\boldsymbol{P(on) = P(v_k > 0) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_0^{\infty} e^{-\frac{(x-\overline{v_k})^2}{2\sigma^2}} dx} \qquad (5.2.3.9)$$

Equation 10 calculates the probability based on the Gaussian distribution in that, given a reference value of the local field $\overline{v_k}$ (mean of the distribution), the effective local field generated by the reference value and the effect of noise is greater than zero. Figure 5.2.3.3 shows some examples of this mechanism, taking into account that graphically, the probability is given by the area under the distribution curve.

Figure 5.2.3.3. Influence of Gaussian noise on neural activation: small, medium and high probability.

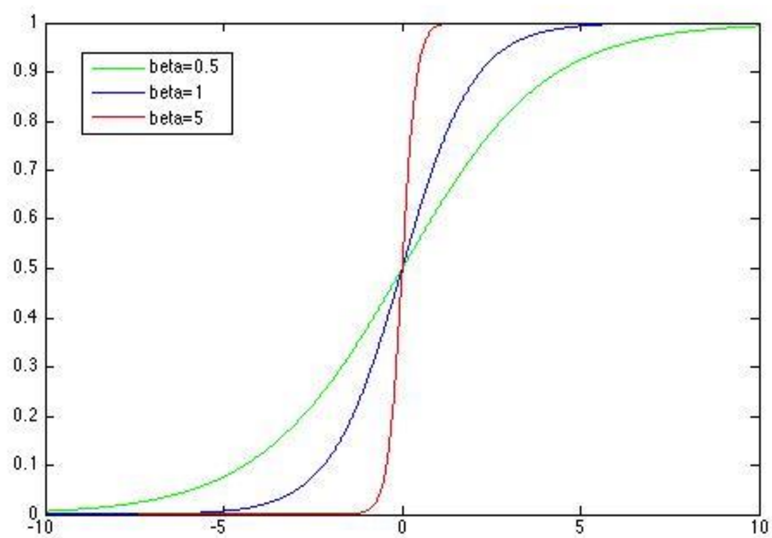


Since the amount of activity of a neuron is given by the frequency of spiking pulses, we can say that the number of spikes in the unit of time is proportional to the probability of activation described in equation 5.2.3.9. A spike is transmitted if the activation potential is greater than the threshold; therefore, the activity of a neuron can be formulated in terms of probability depending on the local field $\overline{v_k}$. The mathematical relation linking the activation probability and the local field is given by the equation 5.2.3.9, which can be accurately approximated by the relation:

$$\boldsymbol{P(on) = \frac{1}{1 + exp(-\beta \overline{v_k})}} \qquad (5.2.3.10)$$

where β is a parameter proportional to the standard deviation σ. The graphical description of equation 5.2.3.10 is shown on Figure 5.2.3.4.

Figure 5.2.3.4. Sigmoid activation function with different values of parameter β.



This curve is usually called sigmoid function or logistic function.

As noted in Figure 5.2.3.4, the greater the β values, the more the sigmoid function is similar to a threshold function; otherwise the function slope decreases and the aspect of the curve is more similar to the letter S (sigmoid comes from sigma – the letter S in Greek alphabet).

The sigmoid function is by far the most common form of activation function used in the construction of artificial neural networks. It is defined as a strictly increasing function that exhibits a graceful balance between linear and non-linear behaviour. Whereas a threshold function assumes the value of 0 or 1, a sigmoid function assumes a continuous range of values from 0 to 1. Note also that the sigmoid function is differentiable while the threshold function is not. Differentiability is an important feature of neural network theory, as described in the next chapter about the feed-forward neural networks.

2.4. Static and Dynamic Neurons

The neuron model depicted in Figure 5.2.3.1 is a rather general scheme that does not describe the type of input stimuli that are processed by the neuron or the temporal relations between inputs and outputs. According to the theory of neural networks, neurons usually work and cooperate in a neural structure in a synchronous mode that is in synchrony with an external clock defining the precise instants when the neuron reads the inputs and produces an output. From such an approach, the input signals are considered as constant for the whole period between a time tick and the next one. In general, and more specifically in the natural network of neurons forming the brain cortex, inputs reach a specific neuron continuously in the form of a train of pulses, as described in the Hodgkin-Huxley model. Therefore, the implementation of a neuron of the first kind, or static neuron, even if it can be a fully functional model for performing specific cognitive tasks, is a rough approach for describing a likely neural dynamics in terms of computational processes. Using biophysically accurate Hodgkin–Huxley-type models is computationally prohibitive, since we can simulate only a handful of neurons in real time.

Izhikevich (2003) recently proposed a very interesting and powerful model based on spiking neurons. This model is described by a simple system of differential equations computationally more convenient compared to the classical approach of Hodgkin and Huxley. The Izhikevich's dynamical model is described by the following 2-dimensional system of differential equations:

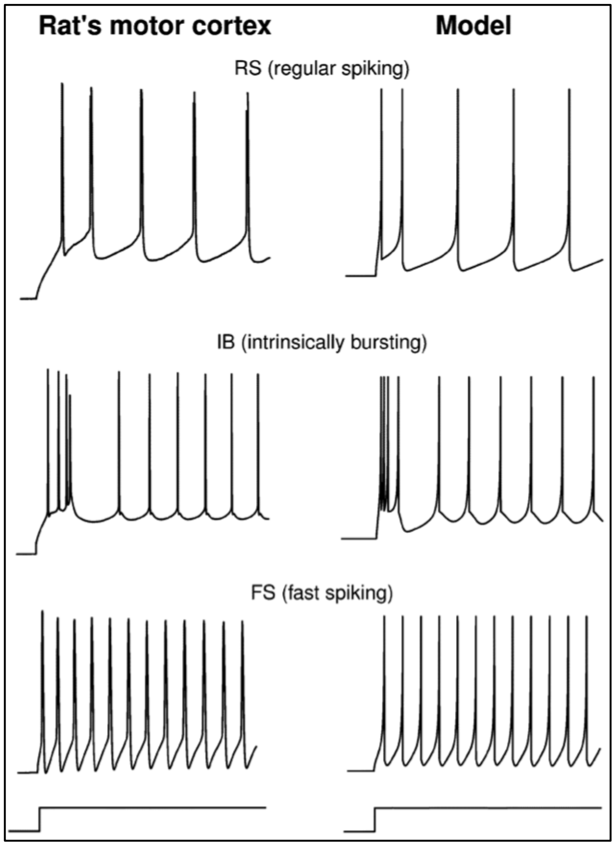
$$\begin{gathered} \dot{v} = 0.04v^2 + 5v + 140 - u + I \\ \dot{u} = a(bv - u) \end{gathered} \tag{5.2.4.1}$$

with the auxiliary after-spike resetting

$$if\, v \geq 30\, mV\, then \begin{cases} v \leftarrow c \\ u \leftarrow u + d \end{cases} \tag{5.2.4.2}$$

Where *u* and *v* are dimensionless variables, and *a*, *b*, *c* and *d* are dimensionless parameters. The Izhikevich model allows performing very similar behaviours with respect to the cortex dynamics, simulating several neuronal evolutions manifested into the mammalian cortex, as shown in Figure 5.2.4.1.

Figure 5.2.4.1. Izhikevich model of a dynamic neuron: comparison with a rat's motor cortex.



Recently, the Izhikevich model has been implemented to artificially describe a large part of the mammalian cortex called *Thalamocortical System* composed of about 10^{11} neurons (one hundred billions) and almost 10^{15} synapses (one quadrillion) by using a supercomputer that simulates the Izhikevich equations with its circuitry (Izhikevich & Edelman, 2008). This approach yields more answers about the functionality of brain dynamics and the comprehension of high-level cognitive processes involved in non-conventional information storage, non-linear classification, uncertain and incomplete information treatment, as well as generalization and correction of corrupted information.

3. Putting Neurons Together: The Networks

3.1. The Beginnings: Perceptron

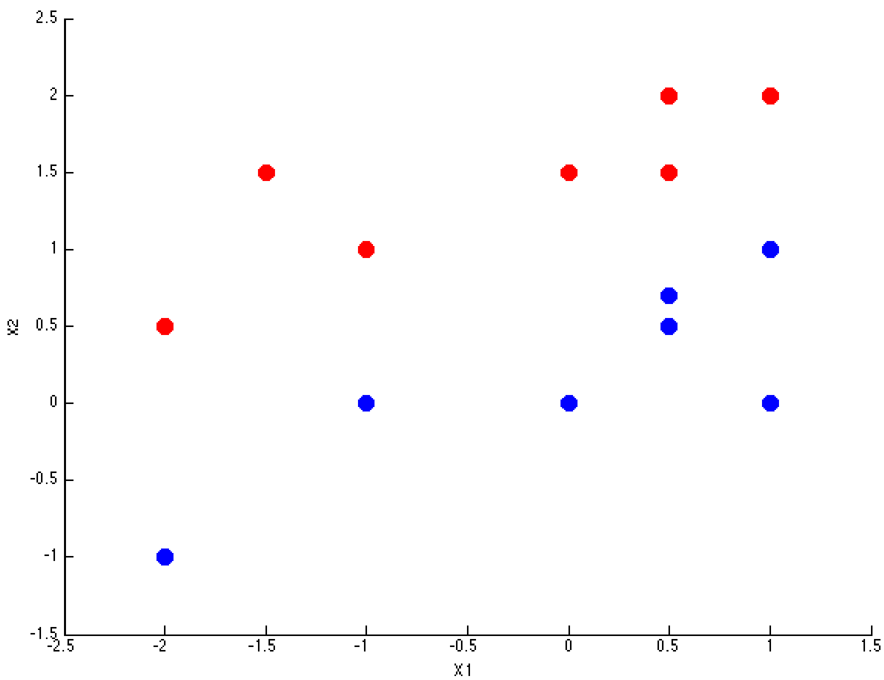
In 1952, Frank Rosenblatt, a psychologist and researcher at the Cornell University, invented an algorithm to perform a simple learning by an artificial neural network (Rosenblatt, 1958). Since the Rosenblatt attempted to model a sensory system of the brain, this typology of neural network was called *Perceptron*. The basic idea was that human beings learn to enter information and concepts by using common senses (mainly sight and hearing) and store the information in some kind of memory, such that when specific information is recalled, it has to be equal to the original one. If the recalled information were learnt incorrectly, it would be necessary to learn such information again so that the new recall operation higher probability to be correct compared to before. This approach can be repeated until all input information is correctly stored and classified, if possible. A neural network that is able to process such information should have a suitable number of inputs for reading the proposed information and an appropriate number of outputs for describing the class to which information belongs. Table 5.3.1.1 describes a very simple example of information classification problem.

Table 5.3.1.1. An example of information classification.

INPUTS		OUTPUT
X1	**X2**	**T**
1	2	1
1	0	0
0	1,5	1
-1	0	0
-1	1	1
0	0	0
1	1	0
-2	-1	0
-1,5	1,5	1
0,5	0,7	0
0,5	2	1
-2	0,5	1
0,5	0,5	0
0,5	1,5	1

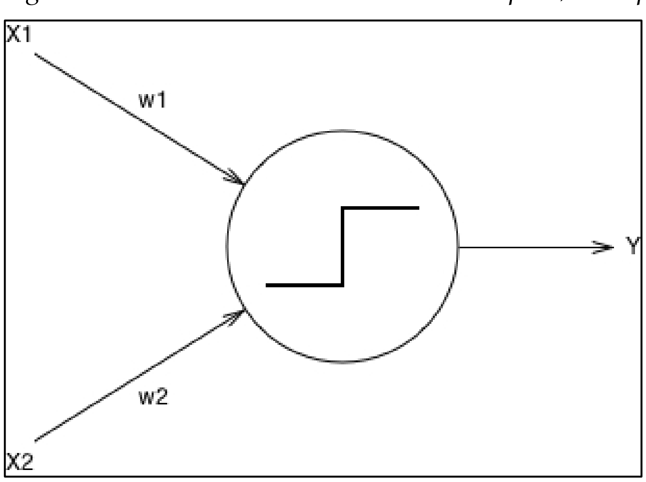
In this table, two input variables X1 and X2 are classified according to the values of the output target variable: 1 if X's belong to the class, 0 otherwise. A sketch of the variables distribution is shown in Figure 5.3.1.1.

Figure 5.3.1.1. Variables distribution for classification example: blue = class 0, red = class 1.



Therefore, suitable neural network architecture to perform the proposed classification task presents two inputs for the variables X1, X2 and one single output for the class label T. The simplest neural network that can be implemented for the proposed task comprises a single neuron and two synaptic weights (w1, w2) that describe the connections between inputs and the neuron. Moreover, since the output target is binary, it can be used as a threshold neuron, as described in the previous chapter (Figure 5.3.1.2).

Figure 5.3.1.2. 1 neuron network with 2 inputs, 1 output and a threshold neuron.



The inner potential of the neuron can be calculated as:

$$\boldsymbol{v = w_1 x_1 + w_2 x_2 + b} \tag{5.3.1.1}$$

that is, the weighted sum of inputs added to the internal bias b.

As said in the previous chapter, when v is positive then the neuron is active, otherwise it is silent. The edge and discriminating point of potential v is zero, since trespassing this value, the neuron changes its state:

$$\boldsymbol{w_1 x_1 + w_2 x_2 + b = 0} \tag{5.3.1.2}$$

Equation 5.3.1.2 has the same form of a line in a 2-dimensional space. By substituting the letters, the equation of a line is:

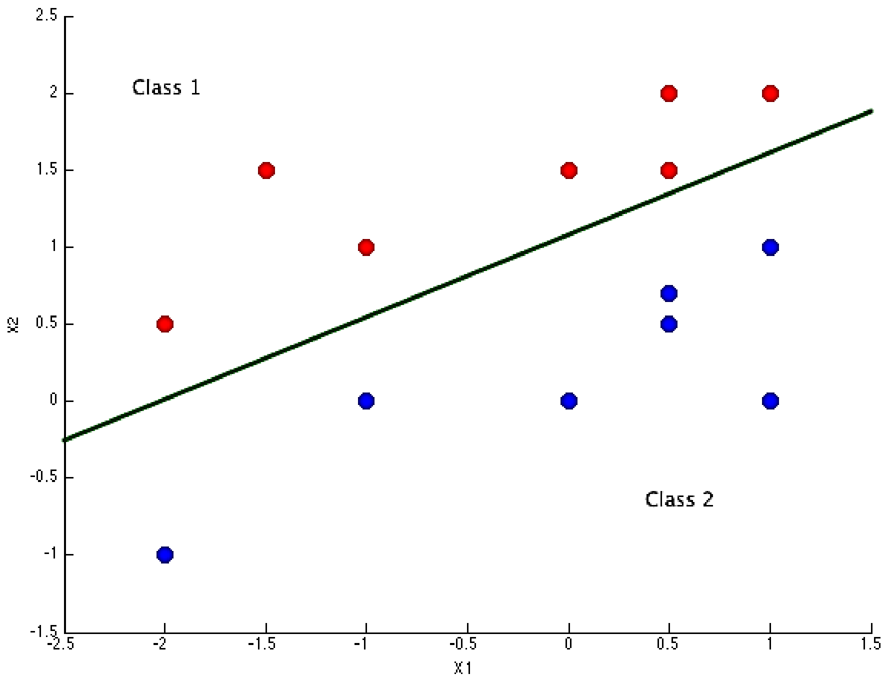
$$\boldsymbol{ax + by + c = 0} \tag{5.3.1.3}$$

where a, b, c are parameters defining the position and the slope of the line in the 2-dimensional space defined by the x-y axes. Note that, if $a > 0$, a point lying in the half-plane over the line changes the equal sign in equation 3 to inequality sign '>', that is, by substituting the coordinates of the point on *x* and *y-axis,* the final sum becomes positive while if the point is in the half-plane under the line, the final sum is negative. Therefore, the line in the Cartesian 2-dimensional space discriminates between two zones separated by the line itself.

According to these last considerations, the neural network in Figure 5.3.1.2 performs a discriminating operation by using a linear rule to distinguish the classes.

The example shown in Figure 5.3.1.1 can then be approached by identifying a line that divides the plane into two zones characterizing the classes of points (Figure 5.3.1.3).

Figure 5.3.1.3. Linear classification applied to the example of Figure 5.3.1.1.

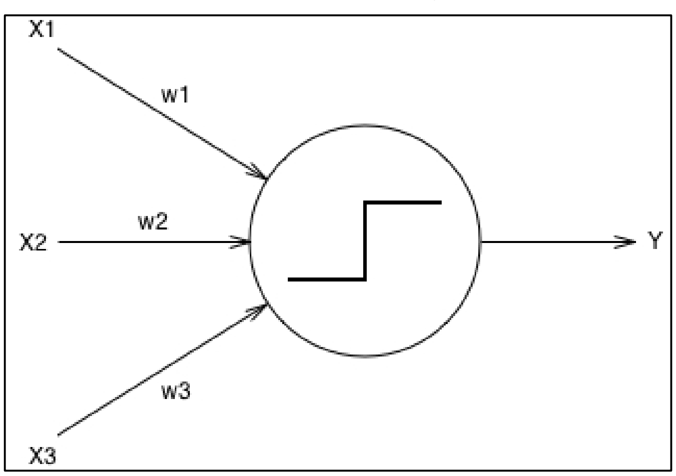


A line fulfilling this request has the equation $x - 2y + 2 = 0$, that applied to the Figure 5.3.1.2 correctly divides the two classes.

Thus, the line parameters can be easily transformed into network parameter by recalling equations 5.3.1.2 and 5.3.1.3, so that $w_1 = 1$, $w_2 = -2$ and $b = 2$.

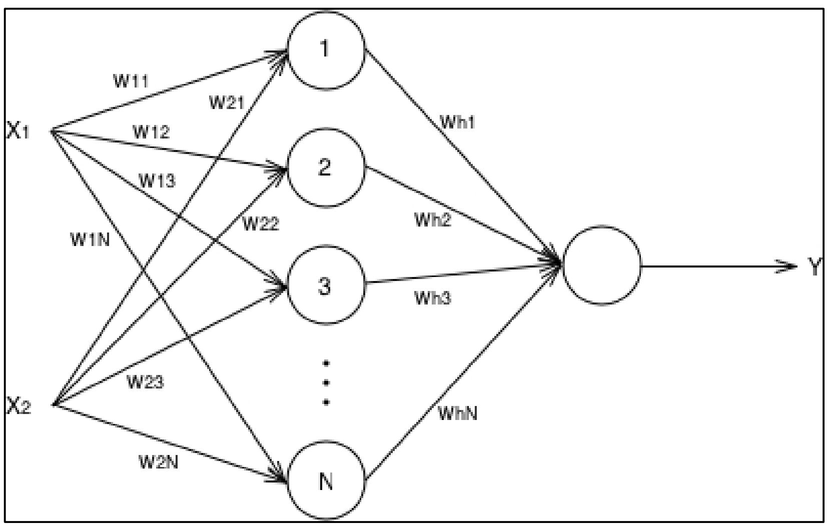
The proposed problem can be extended to a greater output dimensionality by increasing the number of inputs, as in Figure 5.3.1.4.

Figure 5.3.1.4. A 3-dimensional 1-layer Perceptron.



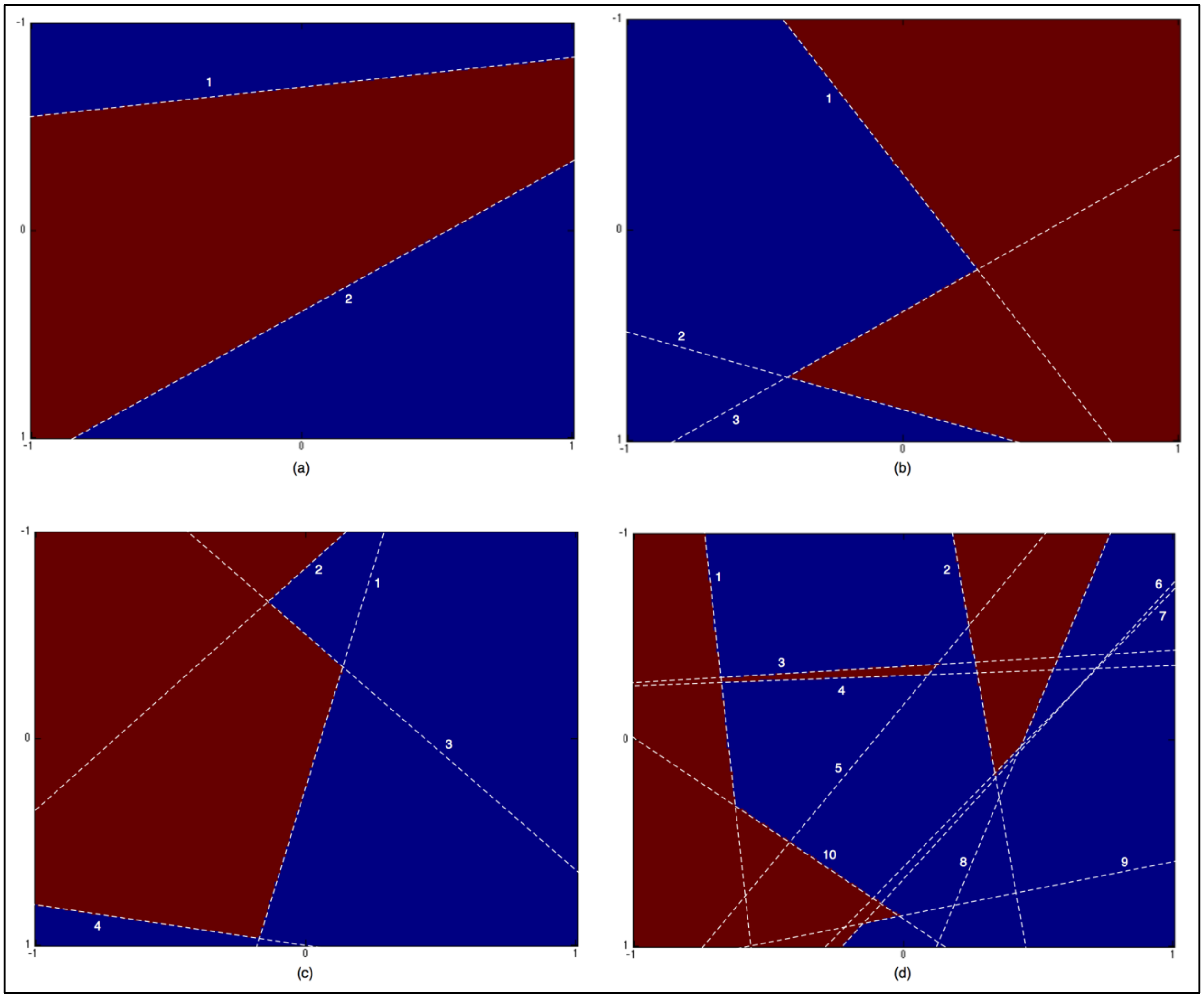
In this case, three inputs are implemented for the acquisition of input information, making it possible to divide the 3-dimensional space into two regions divided by a linear surface (plane). By iterating this concept, if the neuron in Figure 5.3.1.2/4 has *N* inputs, the data are described in a *N*-dimensional space and the Perceptron can divide the space into two regions with a boundary being a (N-1)-dimensional linear surface (hyper-plane). This is the main limitation of a Perceptron with a single layer of neurons: the space can be divided only into two regions by means of a linear hyper-plane. This drawback can be solved by adding more layers of neurons between the input and output layers, as shown in Figure 5.3.1.5.

Figure 5.3.1.5. A 1-hidden layer Perceptron with N hidden neurons.



To show how a Perceptron can work when adding a hidden layer, some different configurations are proposed as examples. Four Perceptrons with one hidden layer and a progressively increasing amount of hidden neurons were used. Synaptic weights and biases were chosen randomly by a normal distribution with 0-mean and standard deviation equal to 2.

Figure 5.3.1.6. Perceptron binary classifications with: (a) 2 hidden neurons; (b) 3 hidden neurons; (c) 4 hidden neurons; (d) 10 hidden neurons. Blue region describes the output value 0, while the red region describes the output value 1.



In Figure 5.3.1.6, a Perceptron with 2, 3, 4 and 10 hidden neurons divides the plane into two classes. The x and y-axes represent the two input variables with values running continuously from -1 to +1.

Note that in every proposed network structure, the boundary between the class 0 and 1 is described by one or more piecewise linear curves. The amount of straight lines used to build the piecewise linear boundaries is always lower of equal to the amount of hidden neurons.

Therefore, we can say that a Perceptron with a single layer can only classify a linearly separable set of points (with a line in 2 dimensions, a plane in 3 dimensions and, in general, a hyper-plane in N dimensions) while when adding more hidden layers, the classification may approach more complex distribution of points.

A crucial aspect of connectionist models is their ability to learn by experience. Even in the case of Perceptron, Rosenblatt (Rosenblatt, 1958) proposed an algorithm named Delta Rule to define the suitable set of synaptic weights and biases for correctly classifying a set of input-output relations. If the response of an output unit is incorrect, the network can change to produce the correct response the next time that the stimulus is presented. The activity of a neuron is determined by the sum of inputs leading to it and, recalling equations 2 and 3 presented in the previous chapter, each input is given by the product of the activity of a presynaptic unit multiplied by the weight of the connection between them. This means that any change in connection weights will change the activity level of units in the next layer. Thus, an output unit with activity that is too low can be corrected by increasing the weights of connections from units in the previous layer that provide a positive input to it and by decreasing the weights of connections that provide a negative input. Output units with an activity that is too high can be corrected by the opposite procedure. Therefore, in a single layer Perceptron having M input units and N output units, the modification of synaptic weight between the *j*-th input and the *i*-th output will be given by:

$$\boldsymbol{\Delta w_{ij} = \delta(t_i - y_i)x_j} \tag{5.3.1.4}$$

where t_i is the output desired target value, y_i is the actual output value, x_j is the input value and δ is a parameter defining the modification amount of the synaptic weight.

To understand the Delta Rule, consider the three terms in equation 5.3.1.4:

- $(t_i - y_i)$ is the difference between the desired activity level of unit i (i.e., the correct activity) and the level that was actually produced by the flow of excitation or inhibition going to it. Including this term in the weight change, equation does two things. First, if the obtained activity is too low, the difference $(t_i - y_i)$ becomes positive and the weight of connection increases. Consequently, the activity of unit i will be higher if stimulus is presented again. If the activity of i is too high, $(t_i - y_i)$ becomes negative and the weight is reduced, leading to the opposite effect. Second, $(t_i - y_i)$ ensures that the size of any weight change is proportional to the size of the error. When the obtained value of y_i is very inaccurate, there will be very large changes in the weights of the connections to unit i; thus, there will be large changes in its activity level the next time the stimulus is presented. However, as y_i approaches the desired value, these corrections and the consequent changes in y_i become smaller. If y_i reaches the correct value, $(t_i - y_i) = 0$ and the changes to the weight of the connection will cease. The crucial term in this rule, $(t_i - y_i)$ is called the Delta Rule because it represents a difference and Delta character is usually associated with a difference.

- x_j is the activity of unit j. To understand its role in the Delta Rule, consider two units, which both provide input to unit i. From the model of a neuron, we know that the product of their activity and the weight of their connection to i determines the contribution, which they each make to the activity of i. Other things being equal, the input unit with the higher activity will have the greater effect on the activity level of i. This learning rule recognizes this and concentrates the weight change on the connections from units in a high state of activity because these have the most effect on the state of unit with the incorrect activity level.
- δ, the learning rate parameter, is a constant. It determines how large the changes to the weights will be on any learning trial.

Equation 5.3.1.4 describes the modification of synaptic weights but does not give any explicit indication about the modification of biases. Actually, as it can be recalled from the previous chapter about the neuron modelling, the bias can be considered as a particular connection weight with an input always equal to +1. Therefore, the Delta Rule extended to the bias case can be formulated as:

$$\boldsymbol{\Delta b_i = \delta(t_i - y_i)} \qquad (5.3.1.5)$$

The fundamental aspect of the Delta Rule is that, in the case of binary units, it cannot be applied to multilayer networks. To understand this limitation, it is sufficient to note that the output of a unit in equation 4 is always compared with the desired one. In a multilayer network, desired output of hidden units is unknown information since we want to train the network based on the final output values, which are the values of the last layer outputs. Therefore, Delta Rule can be applied only to a single layer Perceptron for linearly separable tasks. Novikoff (1962) proved that the perceptron Delta Rule algorithm converges after a finite number of iterations if the data set is linearly separable.

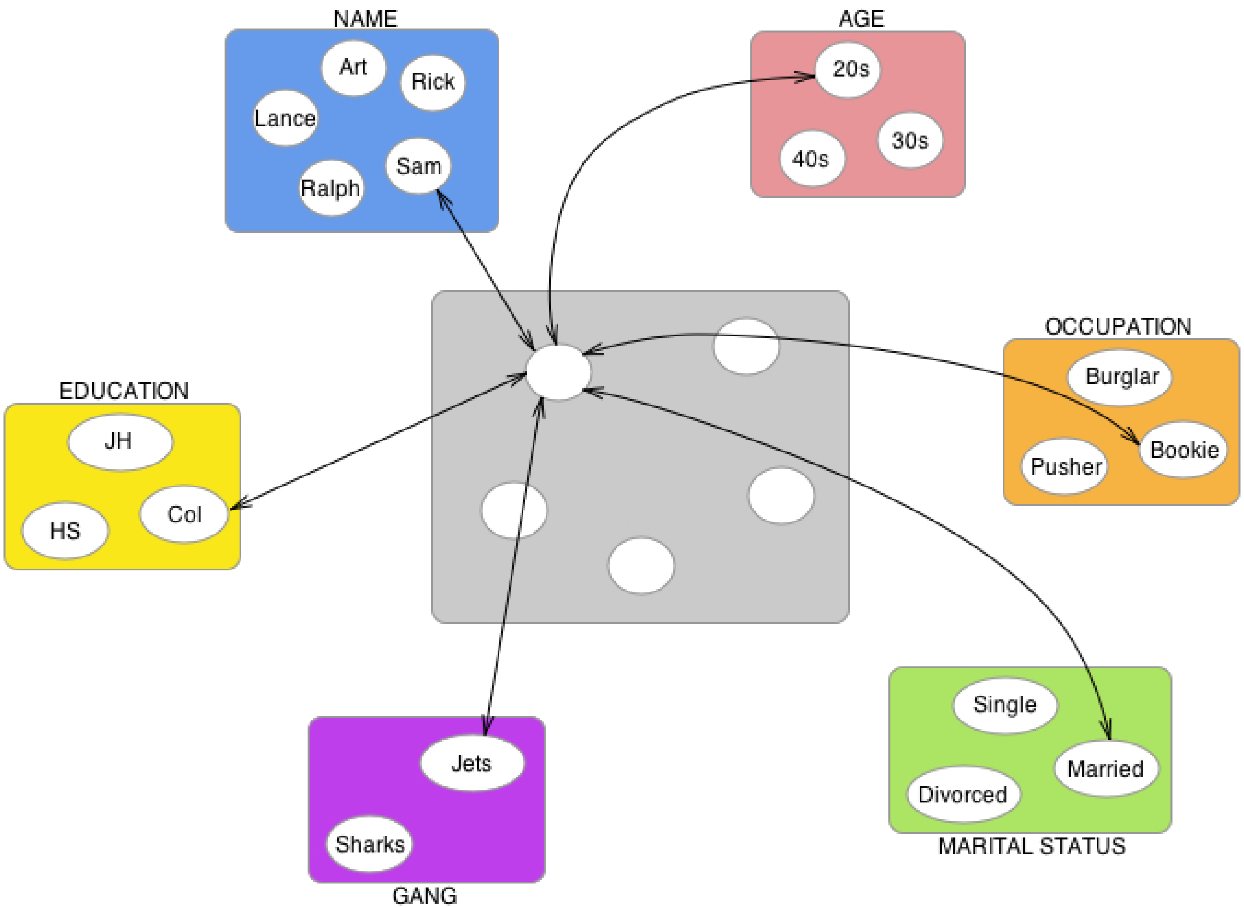
3.2. Storing and retrieving information.

Human memory is content-addressable. That is, you can access a memory by using some part of the information contained in the memory as a retrieval cue. This is unlike retrieval from familiar forms of information storage, such as dictionaries, telephone directories or computer disks. In these, the place where the information is stored has an address, which is needed to retrieve the information. Unlike a dictionary, human memory allows access to the stored content via any part of the information that forms the memory. One of the reasons why connectionist models are attractive is that content addressability is a natural consequence of their distributed structure.

To see how a distributed system with parallel processing works in practice, we will look at retrieval from a simple connectionist memory described by McClelland (McLelland, 1981). This memory demonstrates content addressability and fault tolerance. It also shows typicality effects in retrieval; if asked to retrieve a random member of a category, it will produce a typical member.

McClelland used an information database based on the criminal gangs' population appearing in the popular Broadway musical West Side Story. These gangs comprise several people, each of whom is characterized by an occupation (burglar, pusher or bookie), gang membership (Jets or Sharks), marital status (single, married or divorced), age (20s, 30s or 40s) and education level (College, Junior School or High School). For instance, in the McClelland database, Art is a Jet, 40 years old, is JS single and a pusher, while Rick is a Sharks, 30 years old, is HS, divorced and a burglar. McClelland used 27 elements, each of which is characterized by a record of the proposed features. The architecture of his system is shown in the upper part of Figure 5.3.2.1.

Figure 5.3.2.1. The McClelland retrieval information neural network: excitatory connections.



To store the information related to Jets and Sharks gang members, we need to present the facts about name, gang membership, age, education, marital status and occupation. Within each of these areas of knowledge, there is a node corresponding to the possible values that someone could have. A memory is formed by setting up a link between two nodes. Figure 5.3.2.1 presents a single case of Sam. Since he is a Bookie, a positive connection between the Sam node in the centre of the figure and the corresponding nodes in name and occupation categories is set up. The person nodes in the central part of the figure (grey area) represent the reference units for each member of the distributed database. Since only one feature can be associated with a specific person, nodes into each characteristic area have to be connected by a negative link in order to inhibit the activity of the other units while the reference unit is active.

Figure 5.3.2.2. The McClelland retrieval information neural network: excitatory connections in black (only one reference node is shown), and inhibitory connections in red.

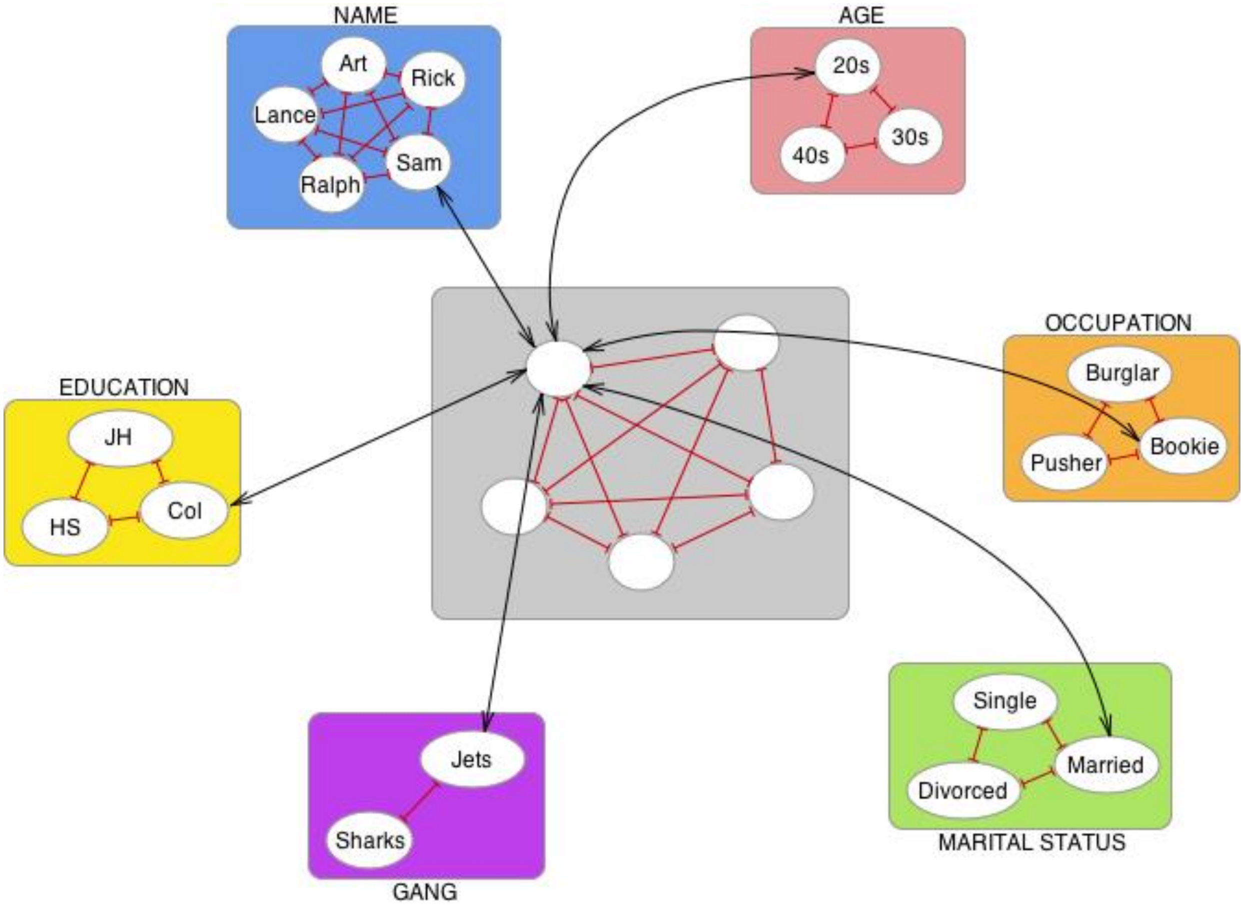


Figure 5.3.2.2 shows the connections between various nodes that are building up as information about five of the people is added to the system. The mutually inhibitory connections are built up between every couple of nodes in the same knowledge area and red lines represent them. These connections indicate, for example, that if someone is in his 20s, he/she cannot be in 30s or 40s.

To test the memory performance of this system, a question like "Can you remember the name of a pusher" can be asked of the system. This is done by activating the Pusher node in occupation area and determining, which unit becomes active in the name area as activity passes round the network. The activity of all nodes starts at a level of -0.1. Activity of the Pusher node increases the activity level of all nodes to which it has positive connections and decreases the activity of all those to which it has a negative connection. Once the activity of a node rises above 0 it excites all the nodes to which it has a positive connection and inhibits all those which it has a negative connection. Eventually, the system reaches a steady state in which the activity level of each node is constant either because it has reached its maximum or minimum permitted value or because its negative and positive inputs are exactly balanced.

The following relation defines the state of a unit in the network:

$$\boldsymbol{y_i} = \sum_{\boldsymbol{j \neq i}} \boldsymbol{w_{ij} x_j} \tag{5.3.2.1}$$

where y_i is the complete input for unit i and is calculated by a weighted sum of the synaptic weights of the connections to the neuron i multiplied by the output of the pre-synaptic neurons x_j. Once y_i is calculated for all units, activations are updated according to the following equations:

If $y_i > 0$,

$$\Delta \boldsymbol{x_i} = (\boldsymbol{max} - \boldsymbol{x_i})\boldsymbol{y_i} - \boldsymbol{decay}(\boldsymbol{x_i} - \boldsymbol{rest})$$

otherwise, (5.3.2.2)

$$\Delta \boldsymbol{x_i} = (\boldsymbol{x_i} - \boldsymbol{min})\boldsymbol{y_i} - \boldsymbol{decay}(\boldsymbol{x_i} - \boldsymbol{rest})$$

where Δx_i is the amount of modification of x_i. The parameters *max* and *min* are constants and define the upper and lower limits of the unit activity x_i. The *decay* parameter is used to force the activation to return to the *rest* value in the absence of external input. If the activation of a unit is equal to *max,* then the net supports the hypothesis completely. If it is equal to *min,* then the net rejects the hypothesis completely. The *rest* corresponds to an "I don't know state". Typically, the choice of these parameters is: *max* > 0 >= *rest* >= *min*. By default, *max* is set to 1, *min* is set to -0.2, *rest* is set to -0.1 and *decay* is 0.1.

To see whether this memory system allows access by content, the question "Do you know the name of a pusher" can be asked. To do this, the Pusher node is activated and left on for entire network evolution.

Figure 5.3.2.3. Name units activity when the Pusher node is activated.

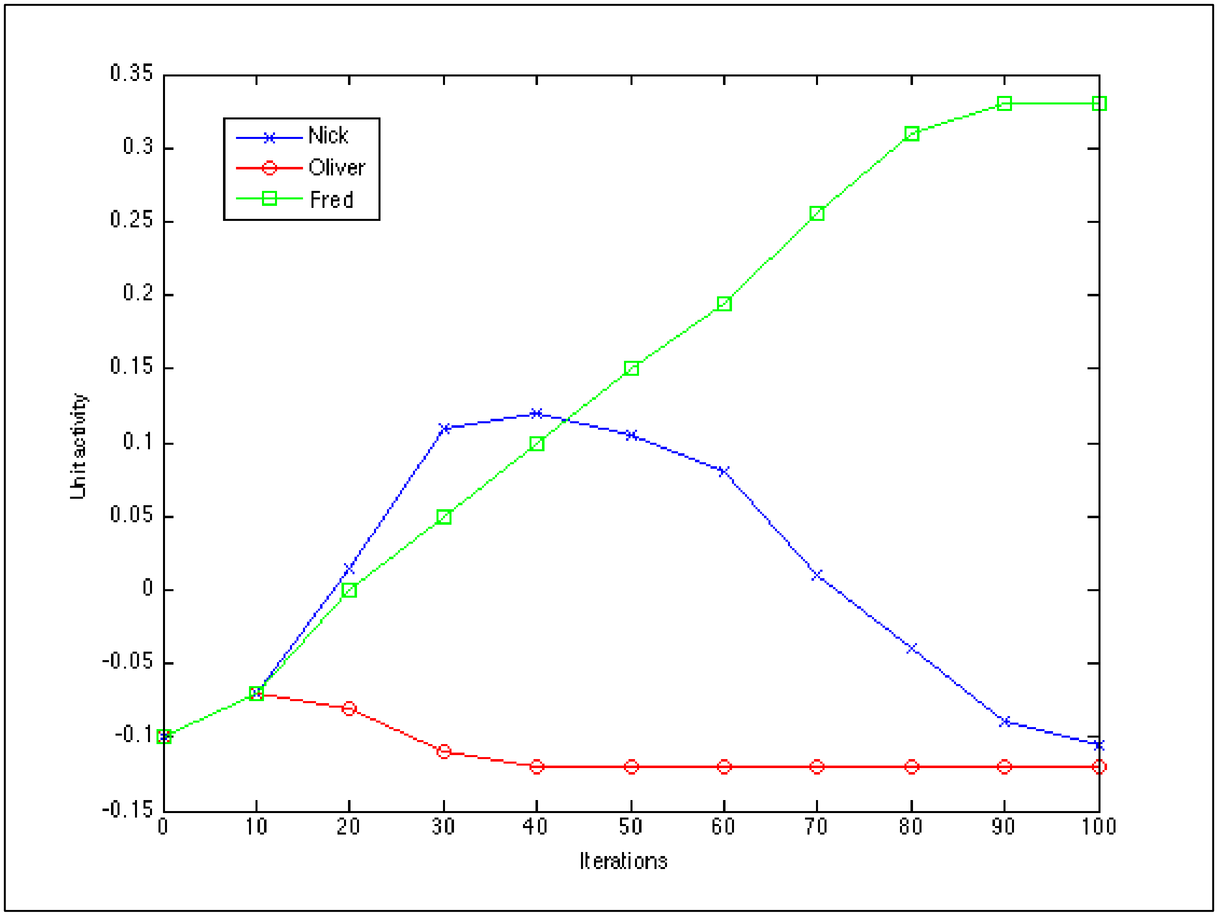


Figure 5.3.2.3 shows the activity of three of the name nodes as a function of the number of processing cycles for which the system has been allowed to run. In the original database proposed by McClelland, Nick, Oliver and Fred are pushers. All the pusher names become activated initially. Most of them, like Oliver, quickly return to their resting level. However, Fred and Nick both become increasingly activated. After about 50 cycles, Fred starts to dominate and soon the system enters a stable state, with Fred activated and all other names returning to their resting level.

The system answers the question with the reply: "Err… Fred"! Thus, this system allows information to be retrieved when it has been accessed by content rather than address.

It should be clear now that a connectionist memory system is totally unlike a conventional memory, such as a computer filing system. In a conventional system, independent pieces of information are stored separately. In a distributed connectionist system, an attempt to extract any information from the system leads to a flow of excitation and inhibition throughout the system to everything that has any relation to this information.

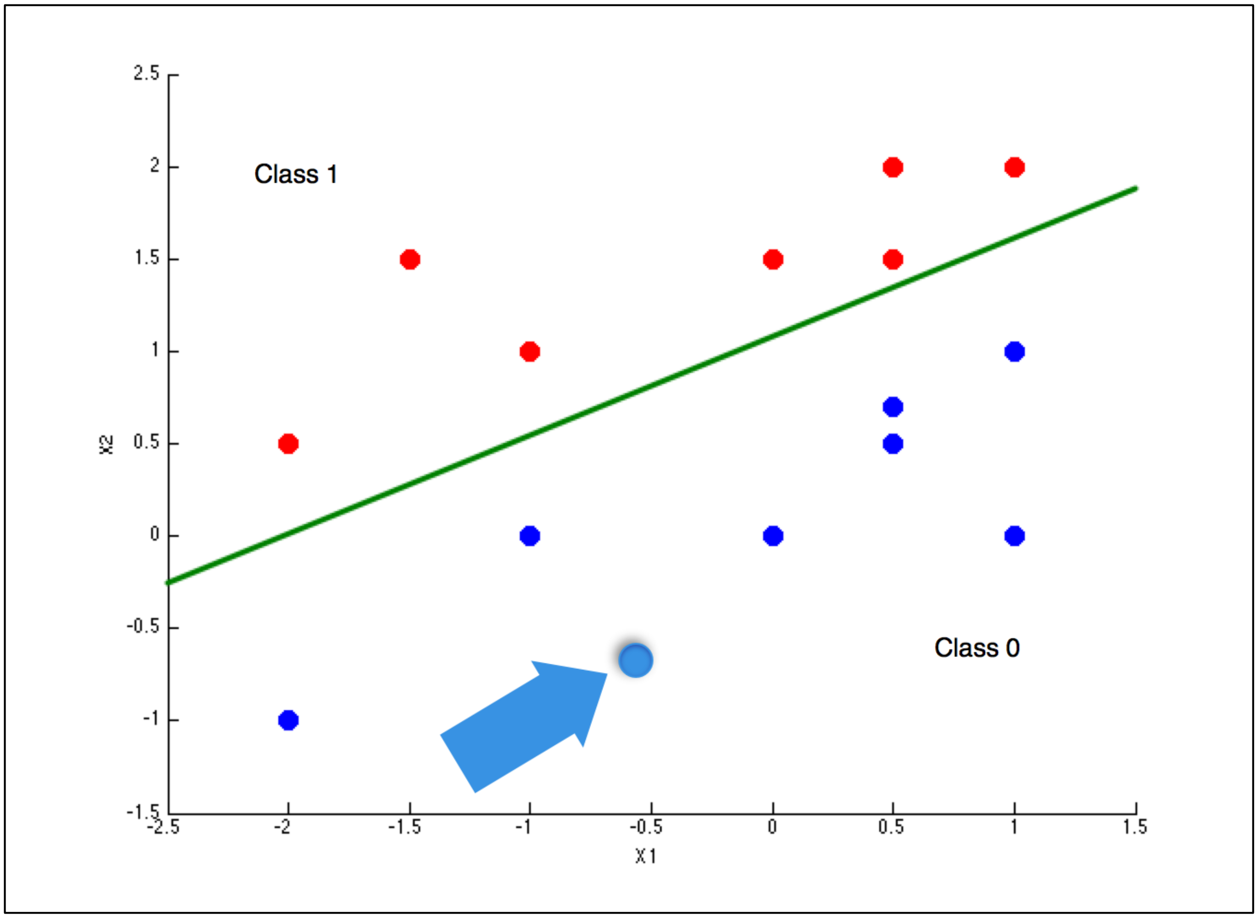
One general point to note about distributed representations is that they blur the distinction between memory and processing. All the information that the network has in its memory is stored in the weights of the connections between units. All the processing that the network can do is determined by the same set of weights.

3.3. Classification and Generalization.

In the first section of this chapter, the Perceptron was presented as a connectionist system able to divide the input space into two different classes and attribute a label to each class defined by the network output code. The classification ability of an artificial system is one of the most important skills a neural network has to process the complex information coming from the external world. As will be seen later, different network structures and specific training algorithms can lead the artificial connectionist system to perform the classification task. The Perceptron case belongs to the class of feed-forward neural network trained by means of a supervised learning algorithm. The supervision aspect regards the fact that to modify the weights of the network, a comparison between the desired and the actual outputs had to be performed. Thus, the training algorithm is based on some mathematical rules that allow the modification of weights based on the output error. An external supervising system should be available to observe the misclassified input data and to operate on the network in order to attain a correct attribution for every processed pattern. The limitation of the Delta Rule revealed in the analysis of Perceptron training, will be overcome for Multilayer Neural Networks using the sigmoid neural activation function and the more complex training algorithm Back-Propagation. With this different approach, the class of allowable classification problems will be extended to curve boundaries. The network complexity will represent internally the complexity of classifications.

Returning to the classification problem shown in Figure 5.3.1.3, after a suitable training, the Perceptron was able to distinguish the two classes by inserting a line in the middle section of the input space or some piecewise linear curves, specifically in the more complex cases shown in Figure 5.3.1.6. It has to be noted that the Perceptron does not just describe the individual points involved in the classification problem, but it also associates the class label (0 or 1) with the entire surface where the points are distributed. Every new input data point the coordinates of which are included in one of the plane portions attributed to a specific class will be associated with the same class, even if that specific point was not trained in the training phase (Figure 5.3.3.1).

Figure 5.3.3.1. New input untrained data point attributed to Class 0.



Since the neural system did not learn the new data point presented at the input layer, the neural network makes an inference about the correct classification of the new data. The attribution of untrained data to a class defined through the training algorithm is called Generalization. This ability is naturally performed by a cognitive system in the brain. For example, it is not necessary to learn all even numbers to define whether a particular number is even or odd. In a way, after the training process, the system extracts the overall rule, defining the right classification. No more training phases are required to assign a class to new incoming input data. The generalization aspect depends strongly on the data patterns used in the training process. Several empirical rules are applied to perform reliable training aimed at correcting generalizations. More precise information about these aspects will be treated in the following chapter on Multilayer Feed-Forward Neural Networks.

3.4. Types of Neural Networks

Neural networks can be classified based on their topological structure, training algorithms and internal dynamics. This last section describes a list of the most popular neural network architectures and dynamic neural systems.

- **Multilayer Feed-Forward Neural Network**. It is an artificial neural network where the information flows only in one direction, from input to output. The network structure is organized in progressive layers and the output of every neuron can be connected to the entire neuron or only a part of it belonging to the successive layer. No auto-connections, back-connections or connections toward neurons of the same layer are allowable. There are no cycles or loops in the network. Thus, this class of networks does not have any internal dynamics. When an input pattern is presented after a propagation delay of the information throughout the network, the output values appear on the output units and persist until the input is removed or changed. Perceptron, the first network architecture seen in the present chapter is a network of feed-forward class. This kind of networks can be trained by several training algorithm that usually depend on the task of the specific neural structure. Anyway, all training algorithms are based on the difference between the expected output value and the actual value when specific input patterns are presented to the input layer. The most popular training algorithms are: Back-Propagation (Rumelhart, Hinton, & Williams, 1986), the Conjugate-Gradient Method (Hestenes & Stiefel, 1952), the Quasi-Newton Mehod (Davidson, 1959), the Levenberg-Marquardt algorithm (Levenberg, 1944; Marquardt, 1963), and the Resilient Back-Propagation (RProp) algorithm (Riedmiller, Braun, 1993).
- **Recurrent Neural Network**. It is a class of neural networks where connections between units form one or more cycles. This creates an internal state of the network, which allows it to exhibit a dynamic temporal behaviour. Two of the most popular networks belonging to this class are the Helman network and the Jordan network (Cruse, 2009). The Helman network is very similar to a Multilayer Feed-Forward Neural Network, except for the presence of some context units forming a supplementary layer. At their inputs, the context

units receive the outcomes of specific hidden layer units and send their outcomes to the inputs of the same hidden units again. These connections form a cycle of information that flows into the network while the network outputs change in time and the input information remains constant. The network is then characterized by an internal dynamics and the output depends not only on the information presented at the input gates, but also on the internal network state. The outcomes of the Helman network are considered the real network responses when the relaxation process converges to an internal steady state. The Jordan network is similar to Elman networks. The context units are however fed from the output layer instead of the hidden layer. The recurrent neural networks take advantage of the presence of a memory of the recent past (internal recursion) to address complex problems in which the temporal aspect is crucial for achieving excellent results, such as handwriting recognition or spoken recognition.

- **Hopfield Network**. This network was the first attempt to realize an associative memory by means of a connectionist model. Hopfield proposed a network where each binary unit is connected to all the remaining units in the network (fully connected network), although no auto-connection are allowed (Hopfield, 1982). Then, for N neurons, the amount of required connections is $N(N\text{-}1)$. The connection weights are defined according to the Hebb's law (Hebb, 1949). The Hebb's law is a rule according to which the connection weight between a pre- and a post-synaptic neuron tends to grow when the activity of the neurons is coherent (that is, they are operating in the same state) while it decreases when there is incoherency between the units state. Starting with this aspect, Hopfield stored in this network some input patterns by determining the weights change according to the coherence/incoherence between all units that had to represent the specific "pixel" (or bit) of all the considered patterns. After defining the network according to its connection weights, when the neurons are initialized to an initial configuration and successively allowed to evolve freely, the network tends to converge its internal dynamics to one of the memorized patterns after a relaxation time. This network was largely used for pattern correction and associative storing of information. Several networks were proposed, starting with the Hopfield original idea to overcome some particular drawbacks of the Hopfield network, such as the poor storing capacity, the patterns of orthogonality constrain, and the presence of undesired memorized phantom patterns.

- **Self-Organizing Maps**. It is a neural network that is trained using unsupervised learning to produce a low dimensional (typically two-dimensional), discretized representation of the input space of the training samples called a map. As opposed to the training algorithms used for feed-forward networks, the unsupervised learning does not require an external observing system to define the distance between the correct and the actual output pattern to modify the connection weights; instead, it is based on implicit internal rules usually aimed at performing a suitable representation of the significant features of information. Kohonen (Kohonen, 1982) proposed the most popular and widely diffused Self-Organizing Map (SOM) connectionist model. This network is based on the Winner-Takes-All (WTA) mechanism and the local training of the units spatially closed to the winner neuron. As it will be described in section 5.4, these approaches can be very powerful in defining a strategy for finding the closeness between patterns and representing it in a low dimensional space. This characteristic is similar to the result of application of statistical Principal Component Analysis (PCA) but unlike it, the neural process is typically non-linear and the description power of a SOM can be considered as more reliable and robust with respect to the traditional approaches.

In this section, the most popular and used neural network structures have been described without the ambition of being exhaustive about the whole panorama of neural network structures, techniques and algorithms. It is mandatory, anyway, to cite the Cellular Neural Networks, the Boltzmann machines, the Brain-in-a-Box, the Neural-Gas networks, among others, because of their relevance to the field of theory and application of neural networks.

In sections 5.4 and 5.5, the Multilayer Feed-Forward neural networks and Self-Organized Maps will be further detailed because of their intense use in the SECOA data analysis.

4. Self-Organizing Maps: A New Method to Extract Similarities

4.1. Topological Maps

The main difference between a map and a topological map in geography is that the former represents a territory by preserving the mutual scaled distance between every couple of points while the latter is based on the representation of specific mutual information between points in order to assign a succession or ordering rule. The typical example is the comparison of a geographical map of the superposition of a town and its subway train lines with the sequence of subway stations in the same town (Figure 5.4.1.1).

Figure 5.4.1.1. A map of Rome and its subways, and the representation of the subway stations.

Thus, in some sense, a topological map preserves the skeleton of the relations between points. This aspect comes from the branch of geometry called topology, where the geometrical forms are defined in terms of invariance of deformations. Coming back to the sequence of subway stations, the picture on the right in Figure 5.4.1.1 can overlap the map on the left by stretching and deforming the sequence of stations without modifying the order of stations.

Some more general types of maps can consider the representation of some particular form of distance different from the classical concept of spatial Euclidean distance. For instance, a town map may represent the minimum time needed to cover the distance between two points instead of the spatial distance between them. The concept of distance can be very general and involve any possible numerical characteristic that define the relationship between two entities, which can be represented as points in an appropriate space. Nevertheless, the distance has to fulfill three fundamental requests:

1. The distance $d(X,Y)$ between a point and itself is always zero: $d(A,A) = 0$.
2. The distance between two points does not depend on the order of the points: $d(A,B) = d(B,A)$.
3. The distance between two points *A, B* can never be greater than the sum of the distances between the first point *A* and a reference third point *C* and between the third point *C* and the second point *B*: $d(A,B) \leq d(A,C) + d(C,B)$. This rule is called *triangular inequality*.

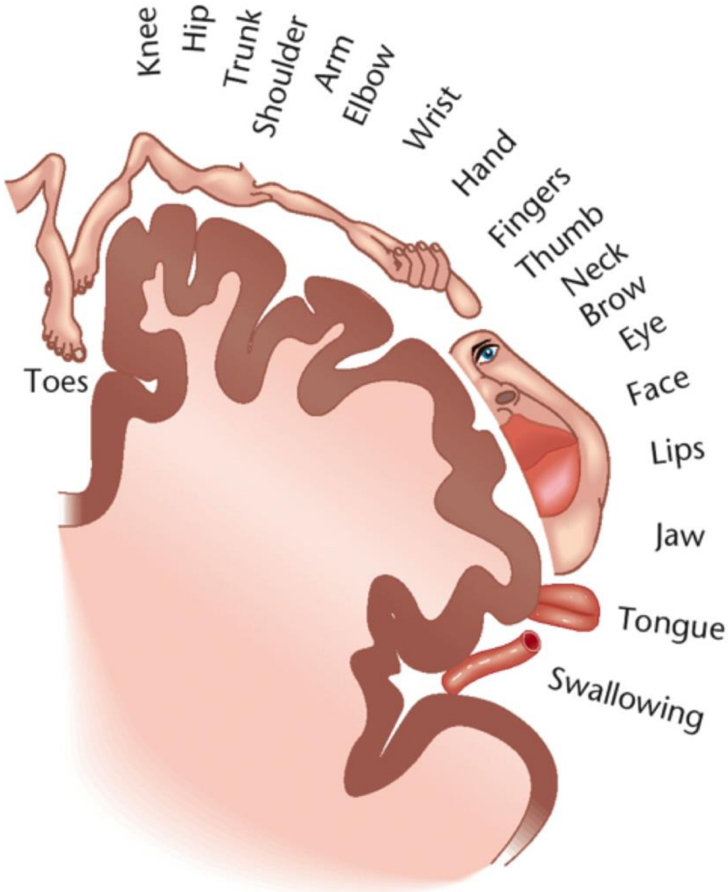
The most common distance is the Euclidean distance, based on the Pythagoras Theorem. In a generic number of dimensions, the Euclidean distance is:

$$\boldsymbol{d}(\boldsymbol{A},\boldsymbol{B}) = \sqrt{(\boldsymbol{A_1} - \boldsymbol{B_1})^2 + (\boldsymbol{A_2} - \boldsymbol{B_2})^2 + \cdots + (\boldsymbol{A_N} - \boldsymbol{B_N})^2} \qquad (5.4.1.1)$$

where *N* is the number of dimensions of the space in which the points are represented. In general, many definitions of distance can be accepted, as for example the *p*-norm distance or the Manhattan distance, but the three fundamental characteristics must always be met.

In neuro-functional description of the brain, a big importance is placed on a particular mapping called topographic mapping. A topographic map is the ordered projection of a sensory surface, like the retina or the skin, or an effector system, like the musculature, onto one or more structures of the central nervous system. This definition means that particular structures of the body are distributed across specific cortex regions by maintaining the representation of the mutual relations between the structures. A typical example of neural topographic map is the somatosensory homunculus depicted in Figure 5.4.1.2.

Figure 5.4.1.2. The somatosensory homunculus.



This figure shows the cortex devolved to the sensory representation of touch, temperature, proprioception, and nociception senses. This means that when a specific part of the human body touches something, the associated area of the somatosensory cortex becomes active in terms of neuronal activity. Therefore, all parts of skin surface, epithelia, internal organs, skeletal muscles, bones, and joints are placed along this cortical region. As can be noted by observing the homunculus in Figure 5.4.1.2, the somatosensory cortex is a topographic map. The cutaneous receptors of the skin project to the spinal cord in an orderly fashion and from there, via different afferent pathways, to the thalamus and the primary somatosensory cortex. Again, adjacent areas on the skin are represented by adjacent neurons in all aforementioned structures. This projection pattern has been termed somatotopy. An interesting aspect of somatosensory homunculus is that some representations are represented more widely compared to others. For example, lips, tongue, and the thumbs project their representations on a wider area than, for instance, elbows or knees. This aspect can be easily experienced on our body by touching these mentioned parts while not looking and trying to detect various sensations we can experience in the different cases.

Human body is intensely represented on the cortex even for other systems, including motor system, auditory system (tonotopic representation), olfactory and gustatory systems, and visual system (retinotopic representation).

As will be shown later, these kinds of natural systems inspire the Self-Organizing Maps and take advantage of the particular structure of these parts of the brain to distribute information over an artificial neural surface, which can conserve the representation of distances.

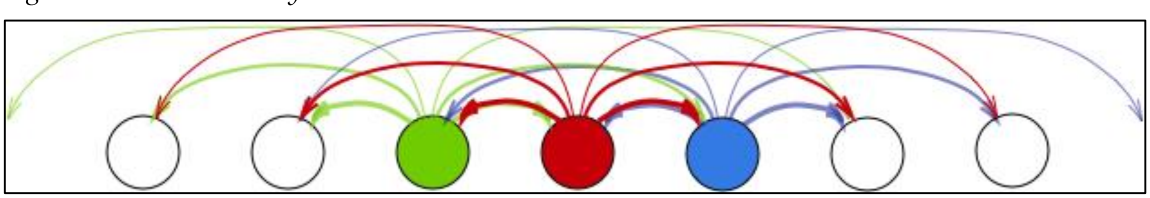
4.2. The Winner-Takes-All Mechanism

An artificial neural network able to simulate the organization of a topographic map based on information features would first require a mechanism that essentially distributes the same or strongly correlated signals to the neurons of a certain region. The longest distances reachable by common or highly correlated input as well as the ranges of the lateral interactions would then determine the spatial diameters of the maps. Second, we would need a mechanism that would select the winner neurons, that is, the center around which adaptation will take place. Third, the restriction of learning to the neighborhood of the winner might simply result from the spreading of triggering activity, but one can also assume that some type of learning factor that controls modification locally is being diffused from the active neurons.

The best self-organizing results are obtained if the following two partial processes are implemented: 1. Decoding the unit that has the best matching with the input information. 2. Improving the match in the neighborhood of neurons centered on the winner.

The former operation can be realized using a Winner-Takes-All (WTA) function. This mechanism is implemented by connecting each unit of the network by lateral connections with the units belonging to its neighborhood to inhibit their activity. Thus, the network has a structure as in Figure 5.4.2.1.

Figure 5.4.2.1. Inhibitory lateral connections.



In this picture, the red unit is laterally connected to the units belonging to a neighborhood of dimension 3. It sends its inhibitory action, progressively decreasing with the distance. Then, the blue and the green units receive a strong inhibition from the red unit, while the next units on the left and right receive a weaker inhibition action since the distance is greater. Figure 5.4.2.1 shows the inhibitory connections also for the green and blue units to illustrate that all units in the network are connected in the same way as the represented ones.

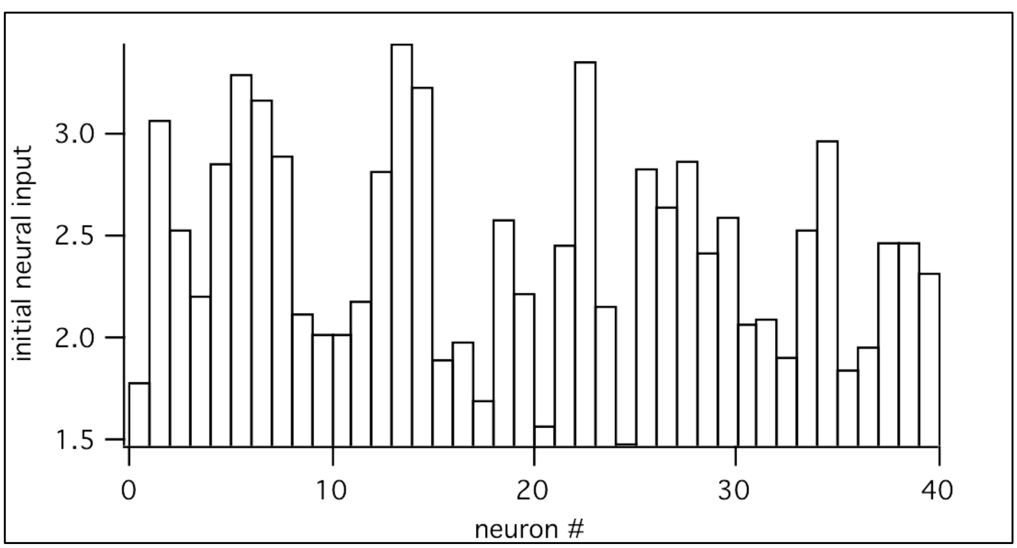
The following relation gives the temporal dynamics of the generic unit in the network:

$$\frac{dy_k}{dt} = I_k^e + I_k^f - f(y_k) \tag{5.4.2.1}$$

where I_k^e is the signal injected by the network input in the k-th unit, I_k^f is the inhibitory lateral contribution of the neighborhood units, and $f(y_k)$ is a generic convex function that depends on the k-th unit state y_k simulating a leakage effect. Since the equation 2 is a differential equation, it describes a temporal evolution of the unit. In particular, the unit state y_k tends to grow if an external stimulus coming from input is present (I_k^e), but the increasing process is opposed by the action of the inhibitory lateral effect (I_k^f). The leakage function ensures that the unit approaches zero if there is no input excitation.

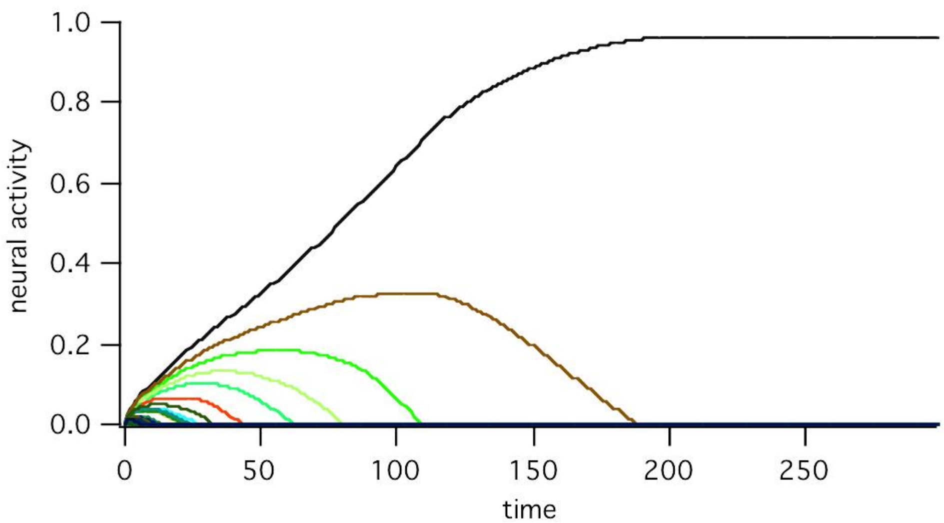
This unit description associated with the network structure and composed of lateral inhibitory connections allows producing the selection of a winner unit when the network is free to evolve. It can be shown in the following example. A 40-unit linear network is initialized with the states depicted in Figure 5.4.2.2.

Figure 5.4.2.2. Initial states of the WTA network.



The initial states can be thought to be produced by the injection of a signal from the input layer of the network. After the initialization phase, the network is allowed to evolve freely so that each unit receives an inhibitory action from its neighboring units, given by the weighted sum of inhibitions, and in turn sends its inhibitory influence. The temporal evolution of the units state is shown in Figure 5.4.2.3.

Figure 5.4.2.3. Temporal evolution of a WTA network.



After an initial transient, all units receive an inhibitory affect sufficient to reduce progressively their internal state, as represented by a solid black curve. The law governing this evolution is based on the proposition that the most activated unit at the initial time can inhibit other units more than they can be inhibited, so that at the end of the process, only the most activated unit wins the competition. In Figure 5.4.2.3, the black curve corresponds to the unit #14 in Figure 5.4.2.2, which has the greatest initial activity, as mentioned before.

The next section shows that the selection of the most activated unit will be associated with a training algorithm so that sufficiently correlated input signals will activate close units on the network surface. In this way, a topographic map based on a neural structure will be realized and a new definition of information distance may be defined.

4.3. The Kohonen map

The main goal of the Self-Organizing Map (SOM) is to transform an incoming signal pattern of arbitrary dimension into a one- or two-dimensional discrete map and to perform this operation adaptively in a topologically ordered fashion.

Figure 5.4.3.1. Two dimensional lattice of neurons.

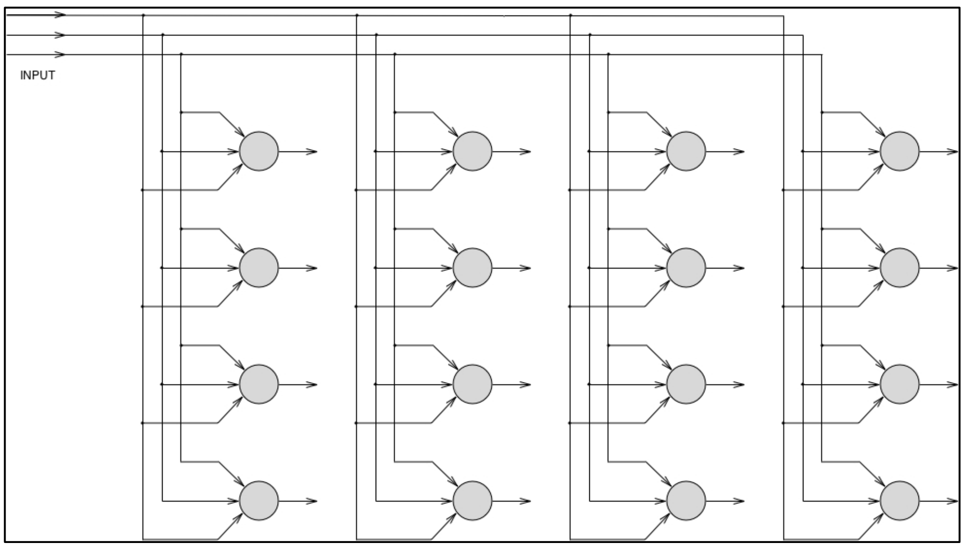


Figure 5.4.3.1 shows the schematic diagram of a two-dimensional lattice of neurons commonly used as a discrete map.

Each neuron in the lattice is fully connected to all the source nodes in the input layer. This network represents a feed-forward structure with a single computational layer consisting of neurons arranged in rows and columns. A one-dimensional lattice is a special case of the configuration depicted in Figure 5.4.3.1. In this special case, the computational layer consists simply of a single column or row of neurons.

Each input pattern presented to the network typically consists of a localized region or *spot* of activity against a quiet background. The location and nature of such a spot usually varies from one realization of the input pattern to another. All neurons in the network should therefore be exposed to a sufficient number of different realizations of the input pattern to ensure that the self-organization process has a chance to mature properly.

The algorithm responsible for the formation of the self-organizing map proceeds first by initializing the synaptic weights in the network. This can be done by assigning them small values picked from a random number generator. In doing so, no prior order is imposed on the feature map. Once the network has been properly initialized, three essential processes are involved in the formation of the self-organizing map:

1. *Competition*. For each pattern, the neurons in the network compute values of their discriminant functions. This discriminant function provides the basis for competition among the neurons. The unit with the largest value of discriminant function is declared winner of the competition. The physiological model of Winner-Takes-All mechanism has been described in the previous section.
2. *Cooperation*: The winning unit determines the spatial location of a topological neighborhood of excited neurons, thereby providing the basis for cooperation among such neighboring units.
3. *Synaptic Adaptation*: This last mechanism enables the excited neurons to increase their individual values of the discriminant function in relation to the input pattern through suitable adjustments applied to their synaptic weights. The adjustments made enhance the response of the winning unit to the subsequent application of a similar input pattern.

The competitive process can be described in terms of algorithmic approach as follows. The unit with afferent synaptic weights that are the most similar to the presented input pattern wins the competition. By denoting with $= [x_1, x_2, \dots, x_m]^T$ the input pattern with dimension m, and with $\boldsymbol{w_j} = \left[w_{j1}, w_{j2}, \dots, w_{jm}\right]^T$ the synaptic weight vector of unit j, the rule to detect the unit that best matches the input pattern is:

$$\boldsymbol{i}(\boldsymbol{x}) = \boldsymbol{\underset{j}{argmin}} \left\| \boldsymbol{x} - \boldsymbol{w_j} \right\| \tag{5.4.3.1}$$

that is, the winning unit has weight vector that is more similar to the input pattern. $i(\boldsymbol{x})$ is the index associated with the winning unit, and $\left\|\boldsymbol{x} - \boldsymbol{w}_j\right\|$ is the Euclidean distance between the input pattern $\boldsymbol{x}$ and the weight vector $\boldsymbol{w}_j$.

The winning neuron locates the center of a topological neighborhood of cooperating neurons. Based on the neurobiological evidence, in case of lateral interaction among a set of excited neurons, a neuron that is firing tends to excite the neurons in its immediate neighborhood rather than those that are farther away. Therefore, let $h_{j,i}$ denote the topological neighborhood centered on winning neuron i and encompassing a set of excited and cooperating neurons, a typical one of which is denoted by j. Let $d_{j,i}$ denote the lateral distance between winning neuron

i and excited neuron j. Then, we may assume that the topological neighborhood $h_{j,i}$ satisfies two distinct requirements:

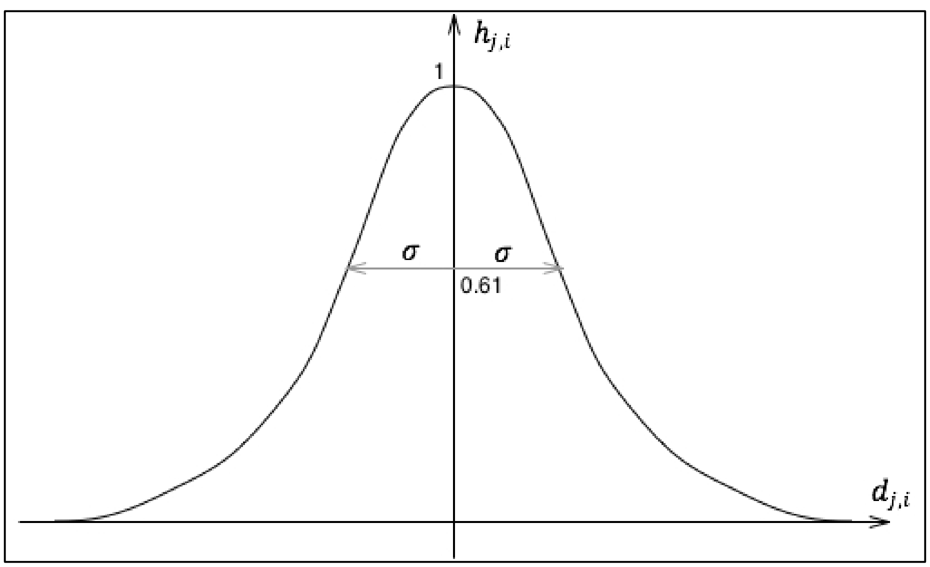
1. The topological neighborhood $h_{j,i}$ is symmetric about the maximum point defined by $d_{j,i} = 0$. In other words, it reaches its maximum value at the winning neuron i for which the distance $d_{j,i}$ is zero.
2. The amplitude of the topological neighborhood $h_{j,i}$ decreases with increasing lateral distance $d_{j,i}$, decaying to zero for $d_{j,i} \to \infty$.

   A typical choice of $h_{j,i}$ that satisfies these requirements is the Gaussian function

$$\boldsymbol{h_{j,i} = exp\left(-\frac{d_{j,i}^2}{2\sigma^2}\right)} \qquad (5.4.3.2)$$

which is translation invariant, that is, it is independent of the location of the winning neuron. The parameter σ represents the effective width of the topological neighborhood, as illustrated in Figure 5.4.3.2.

Figure 5.4.3.2. Gaussian neighborhood function.



It measures the degree to which excited neurons in the vicinity of the winning unit participate in the learning process. The lateral distance between winning neuron i and the excited neuron j is intended to be the simple Euclidean distance evaluated over the network surface:

$$\boldsymbol{d_{j,i}^2} = \left\|\boldsymbol{r_j} - \boldsymbol{r_i}\right\|^2 \tag{5.4.3.3}$$

where the discrete vector $\boldsymbol{r}_j$ defines the position of excited neuron *j*, and $\boldsymbol{r}_i$ defines the discrete position of winning neuron *i*, both of which are measured in the discrete output space.

Another unique feature of the SOM algorithm is that the size of the topological neighborhood shrinks with time. This requirement is satisfied by making the width σ of the topological neighborhood function $h_{j,i}$ decrease with time:

$$\boldsymbol{\sigma(n) = \sigma_0 exp\left(-\frac{n}{\tau_1}\right)} \tag{5.4.3.4}$$

where σ_0 is the value of σ at the initiation of the SOM algorithm, and τ_1 is a time constant. Correspondingly, the topological neighborhood assumes a time-varying form as:

$$\boldsymbol{h_{j,i}(n) = exp\left(-\frac{d_{j,i}^2}{2\sigma^2(n)}\right)} \tag{5.4.3.5}$$

Thus, as time *n* (i.e., the number of discrete iterations) increases, the width $\sigma(n)$ decreases at an exponential rate and the topological neighborhood $h_{j,i}(n)$ shrinks in a corresponding manner. The purpose of a wide $h_{j,i}(n)$ is essentially to correlate the directions of the weight updates of a large number of excited neurons in the lattice. As the width of $h_{j,i}(n)$ decreases, so is the number of neurons for which update directions are correlated.

The last process is the synaptic adaptive process in the self-organized formation of a feature map. For the network to be self-organizing, the synaptic weight vector $\boldsymbol{w}_j$ of neuron *j* in the network is required to change in relation to the input vector $\boldsymbol{x}$. Using discrete-time formalism, given the synaptic weight $\boldsymbol{w}_j(n)$ of neuron *j* at time *n*, the updated weight vector $\boldsymbol{w}_j(n+1)$ is defined by (Kohonen, 1982):

$$\boldsymbol{w_j(n+1) = w_j(n) + \eta(n) h_{j,i}(n) \left(x - w_j(n)\right)} \tag{5.4.3.6}$$

which is applied to all neurons in the lattice that lie inside the topological neighborhood of winning neuron *i*. $\eta(n)$ is the learning-rate parameter of the algorithm.

It seems obvious from equation 5.4.3.6 that the modification of the synaptic weight depends on the difference between the input pattern and the weight vector of the considered unit *j* while the strength of this modification depends on the distance from the winning unit by means of the neighborhood function $h_{j,i}(n)$. This equation has the effect of moving the synaptic weight vector $\boldsymbol{w}_i$ of the winning neuron *I* toward the input vector $\boldsymbol{x}$. Upon repeated presentations of the training data, the synaptic weight vectors tend to follow the distribution of the input vectors due to the neighborhood updating. The algorithm therefore leads to a topological ordering of the feature map in the input space in the sense that neurons that are adjacent in the lattice will tend to have similar synaptic weight vectors.

The learning-rate parameter $\eta(n)$ is time varying, starting at an initial value η_0 and then decreasing gradually with increasing time *n*. The dependence on time of the learning-rate parameter is similar to equation 5.4.3.4:

$$\boldsymbol{\eta(n) = \eta_0 exp\left(-\frac{n}{\tau_2}\right)} \tag{9}$$

where τ_2 is another time constant of the SOM algorithm.

The adaptation of synaptic weights can be decomposed in two distinct phases:

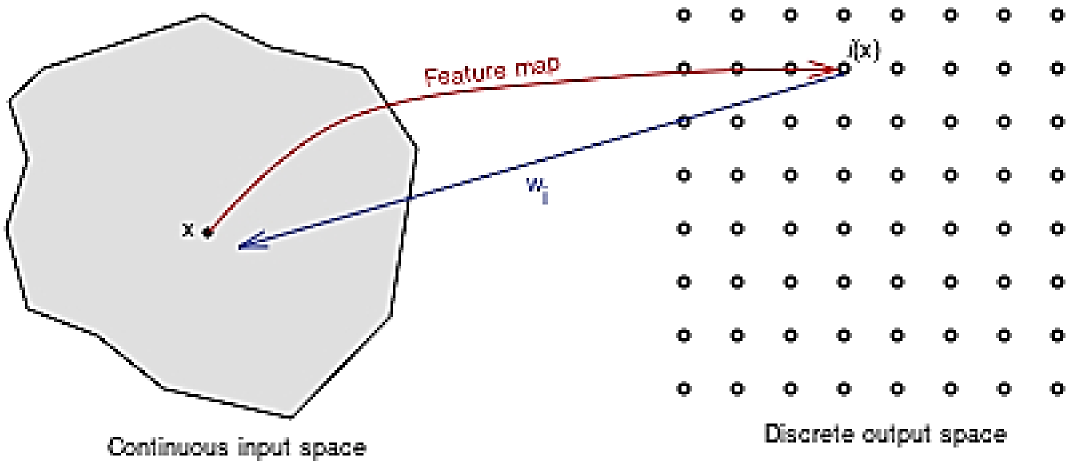
1. *Self-Organizing or Ordering Phase*. Adaptive process of the topological ordering in the weight vectors takes place during the first phase. The ordering phase may take as many as 1000 iterations of the SOM algorithm.

2. *Convergence phase*. The second phase of the adaptive process is needed to fine-tune the feature map and therefore provide an accurate statistical quantification of the input space. As a general rule, the number of iterations constituting the convergence phase must be at least 500 times the number of neurons in the network.

Once the Kohonen's SOM algorithm has converged, the feature map computed by the algorithm displays important statistical characteristics of the input space. By calling $\Phi: \mathcal{H} \rightarrow \mathcal{A}$ the feature map operator, it transforms a spatially continuous input data space $\mathcal{H}$ into a spatially discrete output space $\mathcal{A}$. The feature map Φ has the following properties:

1. *Approximation of the Input Space*. The feature map Φ, represented by the set of synaptic weight vectors $\{\boldsymbol{w_j}\}$ in the output space $\mathcal{A}$, provides a good approximation of the input space $\mathcal{H}$ (Figure 5.4.3.3).

Figure 5.4.3.3. Relationship between feature map and weight vector of winning neuron i.



2. *Topological Ordering*. The feature map Φ computed by the SOM algorithm is topologically ordered in the sense that the spatial location of a neuron in the lattice corresponds to a particular domain or feature of input patterns.

3. *Density Matching*. The feature map Φ reflects variations in the statistics of the input distribution. Thus, the region in the input space $\mathcal{H}$ from which sample vectors $\boldsymbol{x}$ are drawn with a high probability of occurrence are mapped onto larger domains of the output space $\mathcal{A}$, and therefore with better resolution compared to regions in $\mathcal{H}$ from which sample vectors $\boldsymbol{x}$ are drawn with a low probability of occurrence.

4. *Feature Selection*. Given the data from an input space with a nonlinear distribution, the self-organizing map is able to select a set of the best features for approximating the underlying distribution.

5. The Feed-Forward Networks: Strategies for Modelling

5.1. Training Algorithms

Chapter 3 described a simple type of Feed-Forward Neural Network made by a single output layer where the units were represented by a threshold binary activation function. Such network is able to divide the patterns presented at its input into classes and to find a general rule for a generalized classification. The algorithm used to find the best classification linear rule is the Delta Rule, which is based on the comparison between the desired output and the actual output when the associated input pattern is presented at the input layer of the network. The Delta Rule cannot be applied to networks having even just a single hidden layer of units, and the presence of hidden layers is crucial to solve more complex classification problems as classes the boundaries of which are piecewise linear or, more in general, curvilinear.

A more general approach can be pursued by considering the relationship between the overall errors of the network related to the patterns presented at the input. In accordance to the definitions given in chapter 5.3, the error in the network output when a single pattern p is presented, is:

$$\boldsymbol{E^{(p)} = \frac{1}{2}\sum_{i=1}^{N} \frac{\left(t_i^{(p)} - y_i^{(p)}\right)^2}{N}} \tag{5.5.1.1}$$

The sum runs over all N outputs of the network, and $t_i^{(p)}$ and $y_i^{(p)}$ are respectively the expected and the actual outcome of the i-th output of the network. The role of the factor ½ will be clearer afterwards. Basically, $E^{(p)}$ can be considered a sort of variance of the error distribution of the network output. When all P patterns are presented to the network, the overall error can be calculated as:

$$\boldsymbol{E = \sum_{p=1}^{P} E^{(p)} = \frac{1}{2NP}\sum_{p=1}^{P}\sum_{i=1}^{N}\left(t_i^{(p)} - y_i^{(p)}\right)^2} \tag{5.5.1.2}$$

Note that the error E is always positive unless the network outcomes are identically equal to the desired output, and in this case, the error is zero.

Since $y_i^{(p)}$ depends on the network synaptic weights, as well as on the presented pattern at input, the general error E may be modified by changing the synaptic inner parameters of the network. In particular, if a given modification of a synaptic weight Δw eliminates the output error when a certain pattern is presented, such modification may be considered as a useful contribution to the training. Thus, the error will be lower the next time the pattern is presented. In order to achieve a better comprehension of the pattern by the network, a positive increment of the weight should be associated with a negative error variation, that is, a decrease in the error. In mathematical terms:

$$\boldsymbol{\Delta w = -\eta \frac{\partial E}{\partial w}} \tag{5.5.1.3}$$

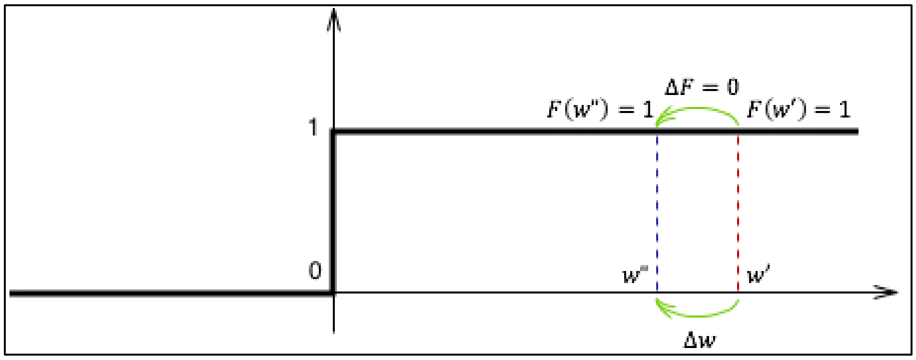
where η is a constant training parameter. This approach is called the Least-Mean-Squares procedure introduced by Widrow and Hoff (1960).

For a single layer Perceptron, after some substitution and the evaluation of the derivative, equation 3 is:

$$\boldsymbol{\Delta w_{ij} = \varepsilon \left(t_i^{(p)} - y_i^{(p)}\right) x_j^{(p)} F' \left(\sum_{k=1}^{M} w_{ik} x_k^{(p)} + b_i\right)} \tag{5.5.1.4}$$

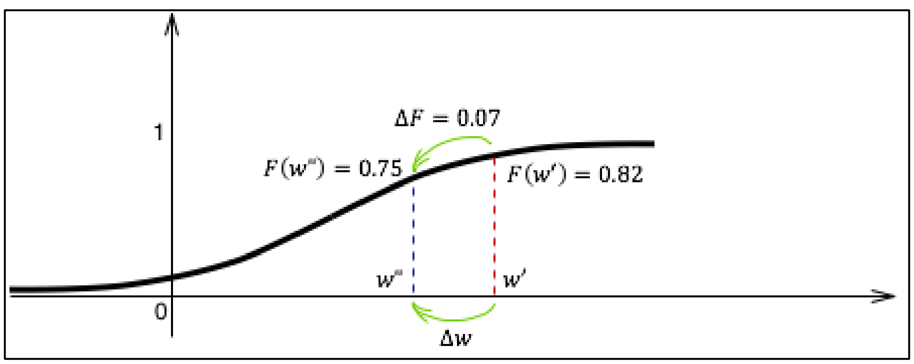
where $x_j^{(p)}$ is the *j*-th input of the pattern *p*, *M* is the total number of inputs, w_{ij} is the synaptic weight connecting the *j*-th input to the *i*-th output, and F' stands for the derivative of the activation function *F* with respect to the synaptic weight w_{ij}. The crucial part of equation 5.5.1.4 is the derivation of *F* with respect to w_{ij}, which reflects the rate of variation of a given output when one of the afferent synaptic weights is changed. It is easy to note that for a threshold binary activation function, a small change in w_{ij} leads, in most cases, to an unchanged output, since the step function is flat almost for every value of its argument. A drastic change in the output, which changes from 0 to 1 or from 1 to 0, occurs only when the change in the synaptic weight allows the argument to cross the threshold. The mathematical description of equation 3 includes the use of an infinitesimal variation of the synaptic weight; thus, the probability to cross the threshold value is usually zero (Figure 5.5.1.1).

Figure 5.5.1.1. Dependence of the network output on the synaptic weight modification – Threshold activation function.



The Widrow-Hoff procedure changes radically if it is applied to a continuously differentiable activation function as, for example, the sigmoid function shown in chapter 5.3. In this case, when a modification of a synaptic weight is applied, there is always a change in the value of the activation function such that the network error is different from zero (Figure 5.5.1.2).

Figure 5.5.1.2. Dependence of the network output on the synaptic weight modification – Sigmoid function.



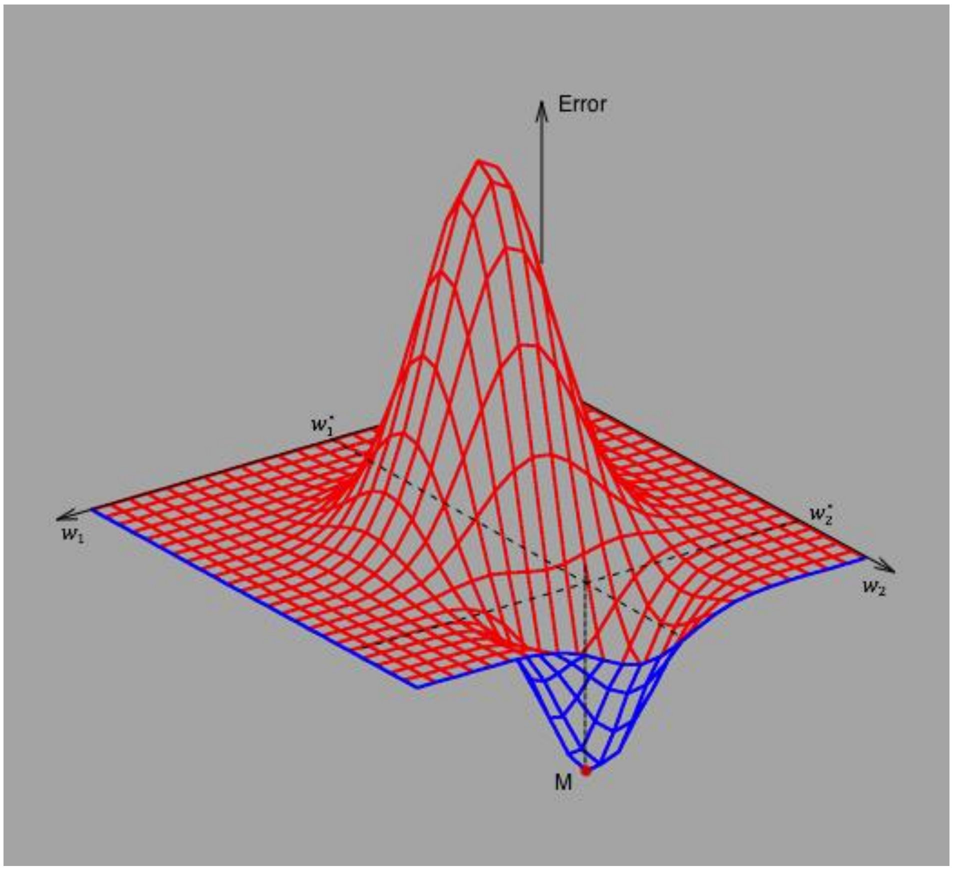
According to the presented information, the feed-forward can be trained with Widrow-Hoff procedure, providing that sigmoid activation function model describes the units. Anyway, in a neural network, the error reduction can be achieved only by taking into account the whole set of synaptic weights, unlike the Widrow-Hoff procedure that can treat only the single neural layer case. Generally, the error modification depends on the combined action of all the synaptic weights, and equation 5.5.1.3 is generalized as:

$$\overrightarrow{\Delta \boldsymbol{w}} = -\boldsymbol{\eta} \boldsymbol{\nabla}_{\boldsymbol{w}} \boldsymbol{E} \tag{5.5.1.4}$$

This new formulation of Widrow-Hoff procedure considers at once the modification of all synaptic weights described by the vector notation (the upper arrow over $\overrightarrow{\Delta w}$). The symbol $\nabla_w E$ is the gradient of the scalar field E, which is the vector of which elements are single derivatives of error function with respect to any single synaptic weight.

The graphical representation of a generic error function depending, for sake of simplicity, on two synaptic weights, is shown in Figure 5.5.1.3.

Figure 5.5.1.3. Error surface for a 2 synaptic weights neural network.



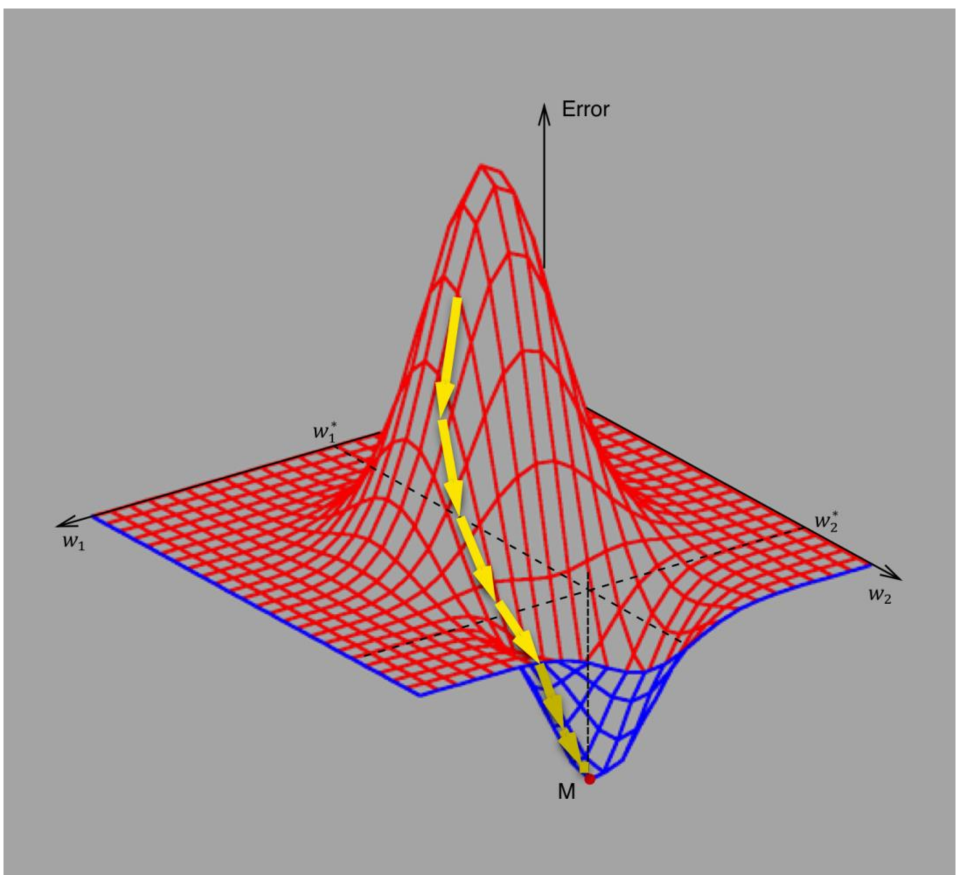
If the error is evaluated for any combination of w_1 and w_2, an error curve surface is produced. This means that by choosing a particular couple of values for w_1 and w_2, the intersection on the w_1-w_2 plane represents a point that projects on the error surface to identify the specific error connected to the synaptic weights choice.

The main task of a training algorithm is to find the minimum error value allowed by a particular neural network. This can be translated, in terms of graphical aspects, into finding the point on the error surface that has the smallest value among all the possible minima. This point is called absolute minimum.

In Figure 5.5.1.3, the best weights combination for the training is w_1^*-w_2^*, corresponding to the lowest minimum of the error surface *M*.

In general, the common use of Widrow-Hoff training algorithm involves random initialization of the synaptic weights of the network. Thus, at the very beginning of the training phase, the parameter state of the network is located somewhere in the weights space. As described previously, the main task of the training algorithm is to find a route to reach the minimum value of the error through successive small adjustments to the weights. Since the gradient vector determines the steepest direction for the error to decrease, where the direction is always tangent to the error surface, the sequence of training steps to reach the local minimum of the surface is depicted in Figure 5.5.1.4.

Figure 5.5.1.4. Training path of Widrow-Hoff procedure.



The main drawback of the Widrow-Hoff approach is that it allows for reaching a local minimum, which, even if it relates to the optimal path given by the gradient descent, does not represent the best training for the network. In the next sections, some heuristic approaches will be presented to make the network training as independent as possible from the shape of the error surface.

The Widrow-Hoff technique can be applied to a simple Perceptron with no hidden layers and sigmoid activation function. However, some problems arise when attempting to apply this method to a Multilayer Feed-Forward Neural Network. This is due to the specific form of equation 3, which binds the error gradient to the synaptic weights variation. As can be seen from equation 4, in the simple Perceptron, one of the factors is the difference between the expected value of the neuron output and the actual outcome. In this case, such difference can be easily computed since the expected value is known. When this approach is extended to a multi-layer neural network, it would be necessary to know the expected output value of all neurons, including those neurons belonging to the hidden layers. For these last neurons, there is no explicit indication about the value they should assume, since the training is based only on the input-output relationship and the comparison between expected and actual network outcome. This issue is avoided by the Back-Propagation training algorithm, which is based on the Widrow-Hoff approach and allows for estimating the expected values of the hidden neurons by reconfiguring the Widrow-Hoff algorithmic technique.

5.2. Feed-Forward neural networks: Structure and training

Figure 5.5.2.1. Graph of a multilayer Perceptron with two hidden layers.

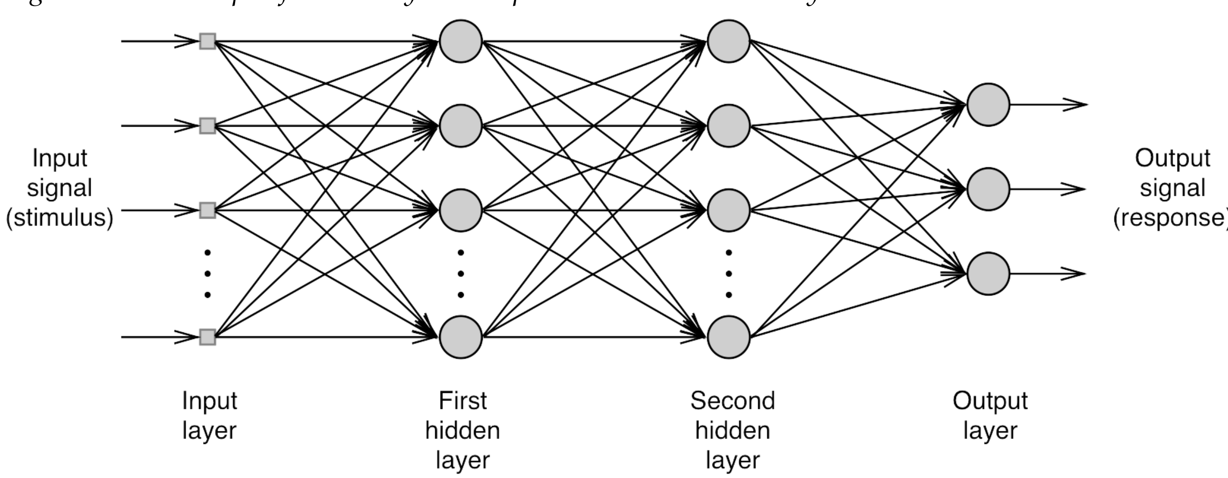


Figure 5.5.2.1 shows the architectural graph of a Multilayer Perceptron with two hidden layers and an output layer.

To set the stage for a description of the Multilayer Perceptron in its general form, the network layers must be fully connected, as shown in Figure 5.5.2.1. This means that a neuron in any layer of the network is connected to all neurons in the previous layer. Signal flow through the network progresses in a forward direction, from left to right, and on a layer-by-layer basis.

Figure 5.5.2.2. Directions of two basic signal flows in a multilayer Perceptron.

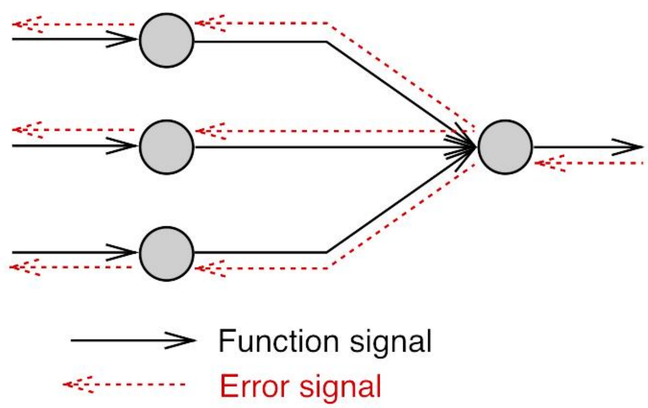


Figure 5.5.2.2 depicts a portion of the multilayer Perceptron. Two kinds of signals are identified in this network:

1. *Function Signals*. A function signal is an input signal (stimulus) that comes in at the input end of the network, propagates forward (neuron by neuron) through the network, and emerges at the output end of the network as an output signal. We refer to such a signal as a *function signal* for two reasons. First, it is presumed to perform a useful function at the output of the network. Second, at each neuron of the network through which a function signal passes, the signal is calculated as a function of the inputs and associated weights applied to that neuron. The function signal is also referred to as the input signal.
2. *Error signals*. An error signal originates at an output neuron of the network, and propagates backward (layer-by-layer) through the network. We refer to it as an *error signal* because its computation by every neuron of the network involves an error-dependent function in one form or another.

The output neurons constitute the output layer of the network. The remaining neurons constitute hidden layers of the network. Thus, the hidden units are not part of the output or input of the network; hence, their designation is *hidden*. The first hidden layer is fed from the input layer made up of sensory units (source nodes). The resulting output of the first hidden layer is in turn applied to the next hidden layer and so on for the rest of the network.

Each hidden or output neuron of a Multilayer Perceptron is designed to perform two computations:

1. The computation of the function signal appearing at the output of a neuron, which is expressed as a continuous sigmoid function of the input signal and synaptic weights associated with that neuron.

2. The computation of an estimate of the gradient vector, which is needed for the backward pass through the network.

The objective of the training process is to adjust the free parameters of the network to minimize the average error defined in equation 2. To do this minimization, the Back-Propagation algorithm, introduced by Rumelhart et al. (Rumelhart, Hinton, & Williams, 1986), can be used to exploit the backward error signals to overcome the limit of the Widrow-Hoff approach in case of multilayer networks, as expressed in the previous section. Specifically, we consider a simple method of training in which the weights are updated on a pattern-by-pattern basis until one epoch, that is, until one complete presentation of the entire training set has been dealt with. The weights are adjusted in accordance with the respective errors computed for each pattern presented to the network.

In a manner similar to the Widrow-Hoff algorithm, the Back-Propagation algorithm applies a correction Δw_{ij} to the synaptic weight w_{ij}, which is proportional to the partial derivative $\partial E/\partial w_{ij}$.

In general, after some mathematics, equation 4 written for a single synaptic weight can be described by:

$$\Delta \boldsymbol{w_{ij}} = \boldsymbol{\eta \delta_i x_j} \qquad (5.5.2.1)$$

where x_j is the *j*-th input to the *i*-th neuron, and δ_i is the so-called *local gradient* defined by:

$$\boldsymbol{\delta_i} = -\frac{\boldsymbol{\partial E}}{\boldsymbol{\partial v_i}} = \boldsymbol{e_i F'(v_i)} \qquad (5.5.2.2)$$

In equation 6, e_i is the error signal at the output of neuron *i* (that is the difference between the desired output for that neuron and the actual output), and $F'(v_i)$ is the derivative of the sigmoid function of the present activation potential of the neuron or local field (that is, the weighted sum of the input signals: $\sum_{k=1}^{M} w_{ik}x_k + b_i$). From equations 5 and 6, we note that the key factor involved in the calculation of the weight adjustment Δw_{ij} is the error signal e_i at the output of the neuron i. In this context, we may identify two distinct cases, depending on where in the network neuron i is located. In case 1, neuron i is an output node. This case is simple to handle because each output node of the network is supplied with a desired response of its own, leading to the same formulation obtained using the Widrow-Hoff algorithm (equation 4). In case 2, neuron i is a hidden node. Even though hidden neurons are not directly accessible, they share responsibility for any error made at the output of the network. The question is how to penalize or reward hidden neurons for their share of responsibility. This problem is solved in an elegant fashion by back-propagating the error signals through the network.

The case depicted in Figure 5.5.2.1 has an input layer, two hidden layers and an output layer. For the formal description used in Back-Propagation algorithm, the following symbols will be used:

- x_i is the i-th input ($i = 1,\dots,N_i$, that is the network has N_i inputs)
- $w_{ji}^{(1)}$ is the generic synaptic weight connecting the i-th input to the j-th neuron of the first hidden layer (the superscript index represents the layer where the post-synaptic unit is located). The index j runs from 0 to the number of neurons contained in the first hidden layer $N^{(1)}$. For the sake of simplicity, the bias effect is included in the synaptic weights symbolic representation, in this case $b_i^{(1)} = w_{0i}^{(1)}$.
- $w_{hj}^{(2)}$ is the generic synaptic weight connecting the j-th neuron of the first hidden layer to the h-th neuron of the second hidden layer ($h = 1, \dots, N^{(2)}$).
- w_{kh}^{o} is the generic synaptic weight connecting the h-th neuron of the second hidden layer to the k-th neuron of the output layer ($k = 1, \dots, N_o$).
- $F(\dots)$ is the sigmoid function and $F'(\dots)$ is its first derivative. Note that $F(\dots)$ does not depend on the layer where the neuron is located, that is, the activation function is the same all over the network.
- $v_j^{(1)}$, $v_h^{(2)}$ and v_k^{o} are the activation potentials or local fields of the j-th neuron of the first hidden layer, the h-th neuron of the second hidden layer and the k-th neuron of the output layer, respectively.

- $y_j^{(1)} = F\left(v_j^{(1)}\right)$, $y_h^{(2)} = F\left(v_h^{(2)}\right)$ and $y_k^o = F(v_k^o)$ are the output of the neurons of the first and second hidden layer and the output layer, respectively.
- $\delta_j^{(1)}$, $\delta_h^{(2)}$ and δ_k^o are the local gradients of the first hidden layer, the second hidden layer and the output layer, respectively, as defined in equation 6.

The following relations describe the activation potentials (for the first and second hidden layer, and the output layer), as seen in the previous chapters:

$$v_j^{(1)} = \sum_{i=0}^{N_i} w_{ji}^{(1)} x_i \quad \rightarrow \quad y_j^{(1)} = F\left(v_j^{(1)}\right)$$

$$\boldsymbol{v_h^{(2)} = \sum_{j=0}^{N^{(1)}} w_{hj}^{(2)} y_j^{(1)} \quad \rightarrow \quad y_h^{(2)} = F\left(v_h^{(2)}\right)} \tag{5.5.2.3}$$

$$v_k^o = \sum_{h=0}^{N^{(2)}} w_{kh}^o y_h^{(2)} \quad \rightarrow \quad y_k^o = F(v_k^o)$$

The Back-Propagation algorithm is used to determine the value of the local gradients of a neuron of a specific hidden layer according to the local gradients of the next layer on the right. For example, the output local gradient can be easily calculated using equation 6 because the error term is known:

$$\boldsymbol{\delta_k^o = e_k F'(v_k^o)} \tag{5.5.2.4}$$

and the modification of the synaptic weights to the output neurons is:

$$\Delta \boldsymbol{w_{kh}^o = \eta \delta_k^o y_h^{(2)}} \tag{5.5.2.5}$$

As described previously, the local gradients of the hidden neurons cannot be evaluated directly because the desired output of such neuron, and consequently the error term, is not known. The rule given by the Back-Propagation algorithm for the second hidden layer states that the local gradient $\delta_h^{(2)}$ depends on the output local gradient in the form:

$$\boldsymbol{\delta_h^{(2)} = F'\left(v_h^{(2)}\right) \sum_{k=1}^{N_o} \delta_k^o \, w_{kh}^o} \tag{5.5.2.6}$$

Thus, the local gradient of the second hidden layer can be calculated *backwards* starting with the knowledge of the local gradient of the output layer. Hence, from equation 5.5.2.1, the modification of the synaptic weights between the first and the second hidden layer is:

$$\Delta \boldsymbol{w}_{hj}^{(2)} = \boldsymbol{\eta} \boldsymbol{\delta}_h^{(2)} \boldsymbol{y}_j^{(1)} \tag{5.5.2.7}$$

Finally, the local gradient of the first hidden layer is:

$$\boldsymbol{\delta}_j^{(1)} = \boldsymbol{F}'\left(\boldsymbol{v}_j^{(1)}\right) \sum_{h=1}^{N^{(2)}} \boldsymbol{\delta}_h^{(2)} \boldsymbol{w}_{hj}^{(2)} \tag{5.5.2.8}$$

with the consequent modification of the synaptic weights from input to the first hidden layer:

$$\Delta \boldsymbol{w}_{ji}^{(1)} = \boldsymbol{\eta} \boldsymbol{\delta}_j^{(1)} \boldsymbol{x}_i \tag{5.5.2.9}$$

Figure 5.5.2.3. Signal-flow graph of a part of the system pertaining to Back-Propagation of error signals.

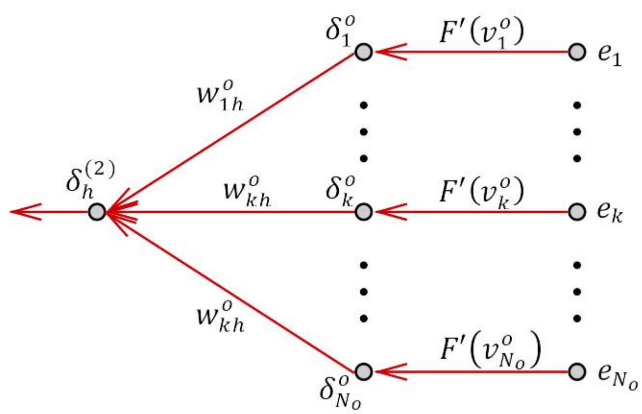


Figure 5.5.2.3 shows the signal-flow graph of error signals for the generic neuron of the second hidden layer.

We now summarize the relations derived from the Back-Propagation algorithm. First, the correction Δw_{ij} applied to the synaptic weight connecting neuron *i* to neuron *j* is defined by the delta rule:

(Weight correction Δw_{ij}) = (Learning rate parameter η) x (local gradient δ_j) x (input signal of neuron *j* y_i).

Second, the local gradient δ_j depends on whether neuron *j* is an output node or a hidden node:

1. If neuron *j* is an output node, δ_j equals the product of the derivative $F'(v_j)$ and the error signal e_j, both of which are associated with neuron *j*.
2. If neuron *j* is a hidden node, δ_j equals the product of the associated derivative $F'(v_j)$ and the weighted sum of the δs computed for the neurons in the next hidden layer or output layer that are connected to neuron *j*.

When applying Back-Propagation algorithm, two distinct passes of computation are distinguished. The first pass is referred to as the forward pass, and the second is referred to as the backward pass.

In the forward pass, the synaptic weights remain unaltered throughout the network, and the function signals of the network are computed on a neuron-by-neuron basis. The function signal appearing at the output of neuron *j* is computed as $y_j = F(v_j)$, as mentioned previously in this chapter and in the past chapters.

The backward pass, on the other hand, starts at the output layer by passing the error signals leftward through the network layer by layer and recursively computing the local gradient for each neuron. This recursive process permits the synaptic weights of the network to undergo changes in accordance with the delta rule.

In a practical application of the Back-Propagation algorithm, learning results from the many representations of a prescribed set of training examples to the Multilayer Perceptron. One complete presentation of the entire training set during the learning process is called an *epoch*. The learning process is maintained on an epoch-by-epoch basis until the synaptic weights and bias levels of the network stabilize and the average squared error over the entire training set converges to some minimum value. It is a good practice to randomize the order of presentation of training examples from one epoch to the next. This randomization tends to make the search in weight space stochastic over the learning cycles. For a given training set, Back-Propagation learning may thus proceed in one of two basic ways:

1. *Sequential Mode (or Epoch Mode)*. The sequential mode of Back-Propagation learning is also referred to as *on-line, pattern,* or *stochastic mode*. In this mode of operation, weight updating is performed after the presentation of *each* training example. To be specific, consider an

epoch consisting of N training examples (patterns) arranged in the order (**x**(1),**d**(1); …, **x**(N),**d**(N)). The first example pair (**x**(1),**d**(1)) in the epoch is presented to the network, and the sequence of forward and backward computations described previously is performed, resulting in certain adjustments to the synaptic weights and bias levels of the network. Then the second example pair (**x**(2),**d**(2)) in the epoch is presented, and so on. This process continues until the last example pair (**x**(N),**d**(N)) in the epoch is presented.

2. *Batch Mode*. In the batch mode of back-Propagation learning, weight updating is performed *after* presenting *all* training examples that constitute an epoch. For a particular epoch, the error (or cost) function is the average squared error described in equation 2.

From an *on-line* operational point of view, the sequential mode of training is preferred over the batch mode because it requires less local storage for each synaptic connection. Moreover, given that the patterns are presented to the network in a random manner, the use of pattern-by-pattern updating of weights makes the search in weight space stochastic in nature. This in turn makes it less likely for the Back-Propagation algorithm to be trapped in a local minimum. In the same way, the stochastic nature of the sequential mode makes it difficult to establish theoretical conditions for the convergence of the algorithm. In contrast, the use of batch mode of training provides an accurate estimate of the gradient vector.

In general, the Back-Propagation algorithm cannot be shown to converge, and there are no well-defined criteria to stop its operation. Instead, some reasonable criteria may be used to terminate the weight adjustments. Let the weight vector $\mathbf{w}^*$ denote a minimum, be it local or global. A necessary condition for $\mathbf{w}^*$ to be a minimum is that the gradient vector $\mathbf{g(w)}$ of the error surface with respect to the weight vector **w** needs to be zero at $\mathbf{w} = \mathbf{w}^*$. Thus, the Back-Propagation algorithm is considered to have converged when the Euclidean norm of the gradient vector reaches a sufficiently small gradient threshold. The drawback of this convergence criterion is that, for successful trials, learning times may be long. A useful modification of the stop criterion based on the evaluation of the norm of gradient vector is that the learning process is considered to have converged when the absolute rate of change in the average squared error per epoch is sufficiently small. In general, these last two stop criteria are combined to achieve the best operational learning results in Back-propagation algorithm.

In these recent years, the gradient descent approach, proposed by Widrow-Hoff procedure and applied via error Back-Propagation to a multilayer neural network, has undergone numerous improvements and variations aimed at a more reliable searching for the minimum. One of the most efficient modifications of Back-Propagation algorithm is the Levenberg-

Marquardt approach (Marquardt, 1963), also called Levenberg-Marquardt Back-Propagation (LMBP), which was applied to multilayer neural networks by Hagan and Menhaj (1994).

Like the quasi-Newton methods, the Levenberg-Marquardt algorithm was designed to approach second-order training speed without having to compute the Hessian matrix. When the performance function has the form of a sum of squares (as is typical in training feed-forward networks), then the Hessian matrix can be approximated as

$$\boldsymbol{H} = \boldsymbol{J}^{\boldsymbol{T}}\boldsymbol{J} \tag{5.5.2.10}$$

Where $\boldsymbol{H}$ is the Hessian matrix (that is, the matrix of which elements are the second derivative of the nonlinear error function with respect to all combinations of synaptic weight pairs: $H_{ij} = \partial^2 E/\partial w_i \partial w_i$).

The gradient can be computed as

$$\boldsymbol{g} = \boldsymbol{J}^{\boldsymbol{T}}\boldsymbol{e} \tag{5.5.2.11}$$

where $\boldsymbol{J}$ is the Jacobian matrix that contains first derivatives of the network errors with respect to the weights and biases, and $\boldsymbol{e}$ is a vector of network errors. The Jacobian matrix can be computed using a standard back-propagation technique that is much less complex than computing the Hessian matrix. The Levenberg-Marquardt algorithm uses this approximation to the Hessian matrix in the following Newton-like update:

$$\boldsymbol{x_{k+1}} = \boldsymbol{x_k} - [\boldsymbol{J}^{\boldsymbol{T}}\boldsymbol{J} + \boldsymbol{\mu I}]^{-\boldsymbol{1}}\boldsymbol{J}^{\boldsymbol{T}}\boldsymbol{e} \tag{5.5.2.12}$$

When the scalar μ is zero, this relation describes just Newton's method, using the approximate Hessian matrix. When μ is large, the matrix becomes gradient descent with a small step size. Newton's method is faster and more accurate near an error minimum, so the aim is to shift toward Newton's method as quickly as possible. Thus, μ is decreased after each successful step (reduction in performance function) and increased only when a tentative step would increase the performance function. In this way, the performance function is always reduced at each iteration of the algorithm.

This algorithm appears to be the fastest method for training moderate-sized feed-forward neural networks (up to several hundred weights).

5.3. Operative Training Techniques and Analyses.

The Back-Propagation algorithm has emerged as the most popular algorithm for the supervised training of Multilayer Perceptrons. Basically, it is a gradient technique and not an optimization technique. Back-Propagation has two distinct properties:

- It is simple to compute locally.
- It performs stochastic gradient descent in weight space for pattern-by-pattern updating of synaptic weights.

These two properties of Back-Propagation learning in the context of a Multilayer Perceptron are responsible for its advantages and disadvantages.

- **Convergence**. The Back-Propagation algorithm uses an instantaneous estimate for the gradient of the error surface in weight space. The algorithm is therefore stochastic in nature, that is, it has a tendency to zigzag its way about the true direction to a minimum on the error surface. Indeed, Back-Propagation learning is an application of a statistical method known as stochastic approximation. Consequently, it tends to converge slowly. We may identify two fundamental causes of this property:

  1. The error surface is fairly flat along a weight dimension, which means that the derivative of the error surface with respect to that weight is small in magnitude. In such situation, the adjustment applied to the weight is small, and consequently many iterations of the algorithm may be required to produce a significant reduction in the error performance of the network. Alternatively, the error surface is highly curved along a weight dimension, in which case the derivative of the error surface with respect to the weight is large in magnitude. In this second situation, the adjustment applied to the weight is large, which may cause the algorithm to overshoot the minimum of the error surface.
  2. The direction of the negative gradient vector may point away from the minimum of the error surface; hence, the adjustments applied to the weights may induce the algorithm to move in the wrong direction.

  Consequently, the rate of convergence in Back-Propagation learning tends to be relatively slow, which may in turn make it computationally excruciating.

- **Local minima**. Another peculiarity of the error surface that influences the performance of the Back-Propagation algorithm is the presence of local minima, that is, isolated valleys, in addition to global minima. Since Back-Propagation learning is basically a hill climbing technique, it runs the risk of being trapped in a local minimum where every small change in synaptic weights increases the error function. However, another set of synaptic weights for which the error function is smaller than the local minimum in which the network is stuck exists somewhere else in the weight space. It is clearly undesirable to have the learning process terminate at a local minimum especially if it is located far above a global minimum.

Several heuristics were proposed in order to overcome the causes of the possible slow rate of the convergence of the Back-Propagation algorithm. These heuristics provide useful guidelines for thinking about how to accelerate the convergence of the Back-Propagation learning through learning rate adaptation.

- *Heuristic 1.* Every adjustable network parameter of the error function should have its own individual learning-rate parameter.
  Here we note that the Back-Propagation algorithm may be slow to converge because the use of the fixed-rate parameter may not suit all portions of the error surface. In other words, a learning rate parameter appropriate for the adjustment of one synaptic weight is not necessarily appropriate for the adjustment of other synaptic weights in the network. This heuristic recognizes this fact by assigning a different learning rate parameter to each adjustable synaptic weight in the network.

- *Heuristic 2.* Every learning rate parameter should be allowed to vary from one iteration to the next.
  The error surface typically behaves differently along different regions of a single weight dimension. In order to match this variation, these heuristic states that the learning rate parameter needs to vary from iteration to iteration.

- *Heuristic 3.* When the derivative of the error function with respect to a synaptic weight has the same algebraic sign for several consecutive iterations of the algorithm, the learning rate parameter for that particular weight should increase.
  The current operating point in weight space may lie on a relatively flat portion of the error surface along a particular weight dimension. This may in turn account for the derivative of the cost function (that is, the gradient of the error surface) with respect to that weight

maintaining the same algebraic sign; therefore, it points in the same direction for several consecutive iterations of the algorithm. This heuristic states that in such situation, the number of iterations required to move across the flat portion of the error surface may be reduced by appropriately increasing the learning rate parameter.

- *Heuristic 4.* When the algebraic sign of the derivative of the error function with respect to a particular synaptic weight alternates for several consecutive iterations of the algorithm, the learning rate parameter for that weight should decrease.

  When the current operating point in the weight space lies on a portion of the error surface along a weight dimension of interest that exhibits peaks and valleys (that is, the surface is highly curved), then it is possible for the derivative of the error function with respect to that weight to change its algebraic sign from one iteration to the next. To prevent the weight adjustment from oscillating, this heuristic claims that the learning rate parameter for that particular weight should decrease appropriately.

6. Conclusion

In Chapter 5 the biological and theoretical foundations of Artificial Neural Networks have been laid, taking into account the introductory and didactical scope of the present document. Basically, the path through the several neural network systems and architectures has been developed by starting from the functionality of a single biological neuron in algorithmic terms, to the information processing given by some basic neural architectures as feed-forward neural networks and self-organizing maps. However, some final conclusions are required in order to focus some specific aspects and special characteristics of this type of neural models and, more in general, of any modelling system.

- An artificial neural network is an algorithmic machine whose purpose is to exhibit a behaviour as close as possible to the system under analysis. This aspect requires that the complexity describing the observed system be somehow reproduced in the neural modelling system by defining a suitable architecture made of processing units and connections. Usually the definition of the neural architecture is a hard task for the user trying to implement a neural modelling system, mainly due to the partial or sometimes lack of knowledge about the system under analysis. It is then suggested to extract as much information as possible about the nature of the analysed system by means of a suitable mathematical or statistical analysis of the collected data and to use well-defined hypotheses to highlight the most significant aspects of the dynamics underlying the system.

- Although in nature is possible to deal with systems characterized by an evident and clear linear behaviour, such as for example some simple isolated mechanical systems, it is most likely to come across natural systems characterized by some kind of non-linearity whose presence may radically increase the complexity, as mentioned in paragraph 5.1. This makes the artificial neural networks excellent candidates as appropriate modelling systems for their non-linear characteristic and the ability to represent complexity by linking together a sufficient number of neural units.
- Finally, all the information about the complex nature of the observed system is contained in the connectionist architecture of the network. Therefore, the memory where the dynamical information is stored is distributed all over the neural network by means of the connection weights and the geometrical structure of the network.

7. References

Amit, Daniel (1992). *Modeling Brain Function: The World of Attractor Neural Networks*. Cambridge University Press.

Barreira, Luis and Pesin, Yakov (2007). *Nonuniform Hyperbolicity. Dynamics of Systems with Nonzero Lyapunov Exponents*. Cambridge University Press

Birkhoff, Garret and Rota, Gian-Carlo (1978), *Ordinary Differential Equations*, New York: John Wiley and Sons, Inc.

Box, George E.P.; Jenkins, Gwilym M.; Reinsel, Gregory C. (2008). *Time Series Analysis: Forecasting and Control*. Wiley

Campbell, David and Mayer-Kress, Gottfried (1991). *Chaos and Politics: Applications of Nonlinear Dynamics to Socio-Political Issues*. Santa Fe Institute, Paper #: 91-09-032

Chorafas, Ditmitris N. (1994). *Chaos Theory in the Financial Markets*. Mcgraw Hill Book Co

Cole, Ronald A. (1989). *Language identification with neural networks: a feasibility study.* IEEE Communications, Computers and Signal Processing, 1989.

Cruse, Holk (2009). *Neural Networks as Cybernetic Systems*. Brains, Minds & Media, 2nd and revised edition.

Davidson, W.C. (1959). *Variable metric method for minimisation*. A.E.C. Res. Develop. Rept. No. ANL-5996 (1959).

Durrett, Rick (2010). *Probability: Theory and Examples*. Cambridge University Press

Guckenheimer, John and Holmes, Philip (1983). *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*. Springer

Guégan D. (2009). *Chaos in Economics and Finance*. Annual Reviews in Control 33, 1 (2009) 89-93

Gowdy J., Pomerleau D.A., Thorpe C.E. (1991). *Combining artificial neural networks and symbolic processing for autonomous robot guidance*. Engineering Applications of Artificial Intelligence, 4(4):279-291

Grossberg S. (1988). *Competitive learning: From interactive activation to adaptive resonance*. Neural Networks and Natural Intelligence, MIT Press.

Hagan, M.T.; Menhaj, M.B. (1994). *Training feedforward networks with the Marquardt algorithm*. IEEE Transactions on Neural Networks, Vol. 5, Issue 6, 989-993.

Hamilton, J. (1994). *Time Series Analysis*. Princeton University Press

Hebb, D.O. (1949). *Organization of behavior*. New York: Wiley.

Hestenes, Magnus R.; Stiefel, Eduard (1952). *Methods of Conjugate Gradients for Solving Linear Systems*. Journal of Research of the National Bureau of Standards 49 (6).

Hodgkin, A. L.; Huxley, A. F. (1952). *A quantitative description of membrane current and its application to conduction and excitation in nerve*. The Journal of physiology 117 (4): 500–544.

Hopfield, J.J. (1982). *Neural networks and physical systems with emergent collective computational abilities*. Proceedings of the National Academy of Sciences of the USA, vol. 79 no. 8 pp. 2554–2558.

Izhikevich, Eugene M. (2003). *Simple Model of Spiking Neurons*. IEEE Transaction on Neural Networks, Vol.14, n.6, 1569-1572.

Izhikevich, Eugene M. and Edelman, Gerald M. (2008). *Large-scale model of mammalian thalamocortical systems*. PNAS, 105:3593-3598.

Jordan, D. W.; Smith, P. (2007). *Nonlinear Ordinary Differential Equations* (fourth ed.). Oxford University Press

Khalil, Hassan K. (2001). *Nonlinear Systems*. Prentice Hall

Kohonen, Teuvo (1982). *Self-Organized Formation of Topologically Correct Feature Maps*. Biological Cybernetics 43 (1): 59–69.

Le , Thai Hoang (2011). *Applying Artificial Neural Networks for Face Recognition*. Advances in Artificial Neural Systems, vol. 2011

Levenberg, Kenneth (1944). *A Method for the Solution of Certain Non-Linear Problems in Least Squares*. Quarterly of Applied Mathematics 2: 164–168.

Levy, David (1994).Chaos *Theory and Strategy: Theory, Application, and Managerial Implications*. Strategic Management Journal, Vol. 15, Special Issue: Strategy: Search for New Paradigms (Summer, 1994), pp. 167-178

Lorenz, Edward (1963). "Deterministic non-periodic flow". *Journal of the Atmospheric Sciences* **20** (2): 130–141

Lorenz, Edward (1996). *The Essence of Chaos*, University of Washington Press

Marquardt, Donald (1963). *An Algorithm for Least-Squares Estimation of Nonlinear Parameters*. SIAM Journal on Applied Mathematics 11 (2): 431–441.

McClelland, J.L. (1981). *Retrieving general and specific information from stored knowledge of specifics*. Proceedings of the Third Annual Meeting of the Cognitive Science Society, pp. 170-172.

McCulloch, W.S., Pitts, W. (1943). *A logical calculus of the ideas immanent in nervous activity*. Bulletin of Mathematical Biophysics, vol. 5, 115-133.

Montello D.R., Fabrikant S.I., Ruocco M, Middleton R.S. (2003) *The distance-similarity metaphor in network-display spatializations, Testing the first law of cognitive geography on point-display spatialization*, in W.Kuhn, M.Worboys, S. Timpf (eds.) Spatial information theory: foundations of Geographic Information Science, Springher-Verlag, Berlin, pp.316-31.

Novikoff, A. B. (1962). *On convergence proofs on perceptrons*. Symposium on the Mathematical Theory of Automata, 12, 615-622. Polytechnic Institute of Brooklyn

Pere Plaza i Font, Joan and Régis, Dandoy (2006). *Chaos Theory and its Application in Political Science*. IPSA – AISP Congress, Fukuoka, 9 – 13 July 2006

Ramón y Cajál (1911). *Histologie du Systémes Nerveux de l'homme et des vertébrés*. Paris, Maloine.

Riedmiller, Martin; Braun, Heinrich (1993). *A Direct Adaptive Method for Faster Backpropagation Learning: The RPROP Algorithm*. IEEE International Conference on Neural Networks.

Ripley, B.D. (1996). *Pattern Recognition and Neural Networks*. Cambridge University Press

Robinson, James C. (2004), *An Introduction to Ordinary Differential Equations*, Cambridge, UK.: Cambridge University Press

Rosenblatt, Frank (1958). *The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain*. Psychological Review, v65, No. 6, pp. 386–408.

Ruelle, David (1989). *Chaotic Evolution and Strange Attractors*, Cambridge University Press

Rumelhart, David E.; Hinton, Geoffrey E.; Williams, Ronald J. (1986). *Learning representations by back-propagating errors*. Nature 323 (6088): 533–536.

Sporns, O.; Tononi, G.; Edelman, G.M. (2000). *Connectivity and complexity: the relationship between neuroanatomy and brain dynamics*. Neural Networks (13):909-922

Stone, Lewi and Ezrati, Smadar (1996). *Chaos, Cycles and Spatiotemporal Dynamics in Plant Ecology*. Journal of Ecology, Vol. 84, No. 2 (Apr., 1996), pp. 279-291

Tobler W. (2004). *On the First Law of Geography: A Reply*. Annals of the Association of American Geographers, 94(2), pp. 304–310.

Weaver, Warren (1948). *Science and Complexity*. American Scientist 36 (4): 536–44

White, Halbert (1988). *Economic prediction using neural networks: the case of IBM daily stock returns*. IEEE International Conference on Neural Networks

Widrow, B.; Hoff, M.E. (1960). *Adaptive switching circuits*. In 1960 IRE WESCON Convention Record, part 4, pp. 96-104.

Yan Li and Shanmuganathan S. (2007). *Social area analysis using SOM and GIS: a preliminary research*. Ritsumeikan Centre for Asia Pacific Studies (RCAPS), Ritsumeikan Asia Pacific University, Working Paper n.07-3, August 2007.

Alessandro LONDEI, Department of European, American and Intercultural Studies, Sapienza University of Rome, Italy.

email: alessandro.londei@gmail.com

Londei, Alessandro, "Neural Networks for Multidisciplinary Approach Research". In: Armando Montanari (ed.), *Urban Coastal Area Conflicts Analysis Methodology: Human Mobility, Climate Change and Local Sustainable Development*. Cap. 5. *SECOA FP7 Research Project*, Vol. 5. Rome: Sapienza Università Editrice, 2013. *http://digilab-epub.uniroma1.it*. DOI: 10.13133/978-88-98533-01-5. Web. 30 September 2013.

ABSTRACT: The aim of this chapter is to give an overview of the modern world of neural networks to researchers involved in the field of geographical and territorial studies who utilize contemporary approaches that need to be connected to complex analysis of social, economical, environmental, and political data. Classical tools offered by statistical linear analysis reveal increasing weakness, mainly due to two fundamental reasons: the necessity to formulate some a priori hypotheses to determine the optimal tool to apply and the implicit inter-related connections among the variables evolving in a complex system, as an urban territory or a coastal area such as the SECOA case, for instance. For about 40 years, neural networks have been successfully implemented to deal with the complex natural system modelling due to nonlinear nature of connectionist systems and the possibility offered by neural networks to adapt to a huge class of problems, from nonlinear classification to time-series prediction or dimension reduction. The principal neural systems proposed in this chapter are the Feed-Forward Neural Network and the Self-Organized Maps. The first is useful for a nonlinear modelling of the SECOA system, and to the second is better fitted to di similarity analysis (Taxonomy), in order to describe the case studies of conflict territories.

KEYWORDS: Complex Systems, Nonlinear Systems, Neuron Model, Neural Networks, Feed-Forward Neural Networks, Back-Propagation, Self-Organizing Maps, Kohonen SOM.

Submitted:	6th June 2013
Accepted:	18th July 2013
Final Revision:	23th July 2013
Published:	30th September 2013

CHAPTER 6.

GIS Data and Territorial Management Approach

Luca DERAVIGNONE
Department of European, American and Intercultural Studies,
Sapienza University of Rome, Italy

1. Territorial Analyses and Multidisciplinary Research

"The map is not the territory", said Alfred Korzybski during a conference at the beginning of last century, applying to geography a wider and very long discussed philosophy concept. One of the strongest points of maps is their capacity of abstraction, letting us view the world from a different perspective, maybe simpler but understandable. Hypothetically, if we could create a map that really represents all aspects of the reality, the utility of the map would disappear because we would not be able to understand the complexity that surrounds us. At the same time, we can assume that if we could totally understand the world, there would be no need of maps.

The aim is always to understand at least some aspects of that complexity, and the only way to do it is by going from complexity to simplicity, with the big issue of not going simplistic. Moreover, more complex problem requires cooperation with other disciplines, such as geography, sociology, economy, and others, that is, it requires conducting multidisciplinary research. SECOA project is a typical example since, as already described in chapter 4, in a territorial environment we needed to manage several variables from very different fields.

Now, regardless of the fields of research involved, to understand the complex phenomenon, we have to use analytical methodologies involving quantitative and qualitative variables. GIScience was designed to study all aspects of geographical information and connections in depth and GIS tools help us perform several types of territorial analyses based on these kinds of data. GIS, in fact, supports gathering, describing, analyzing, and modeling the data; thus, an entire process of territorial management approaches field.

Usually, in a practical way, we can say that the fewer variables we consider, the more the analysis will differ from reality, but even if this separation could be perceived as a limit, it is the only way we can try to understand such complex phenomena that constitute the world we are living in. It depends on the path that the researcher chooses to follow in this process and on how far he/she decides to go, but these choices always depend also on practical issues, first of all data accuracy and availability.

This is especially true when we consider territorial data. Starting from their inner characteristics and considering them singularly, in fact, we can consider shape files and raster files as no more than schematic representations of territory: no one sees boundaries, lines or points while going around. The same "layers" concept, that is, the one based on GIS idea, lets us separate the features present in a territory, making an operation that tends to schematize reflect reality. In their digital form, these kinds of data have to be bonded to their resolution, as for satellite images and raster in general, or human perception, as in vector digitalization, just to give few clear examples. At the same time, one of the direct results, even if someone could call it solution, is that each data has to be measured using a specific scale.

That's why in SECOA project, we had to manage a wide range of matters related to data spatial nature, their heterogeneity and collecting scale, all related to very different territorial contexts going from Mediterranean area to Asia. The aim of the next paragraphs is to show the adopted workflow.

2. Managing Heterogeneous Datasets

Since the beginning, the purpose of using GIS technology was to compare all territories in terms of their spatial dimensions in their real context, considering the zones covered by the analyses represented by the whole SECOA metropolitan-urban areas so that conflicts could be considered in relation to the territory to which they belong, since they exist and develop in a real, not virtual, space.

The first step was to gather all information about our case studies, so all project partners were provided the GIS datasets relevant to their study areas, resulting in a huge amount of vector and raster dataset containing several variables collected from the previous Work Packages and related to the analyzed phenomena.

The main factors of differentiation between collected datasets were:

- Different source (census data, satellite imaging, data from public and private DBs)
- Different format (excel tables, vector shape files, raster/GRID data, geodatabases, CSV files)
- Different scale (census units, municipalities, provinces, zones)
- Coexistence of qualitative and quantitative data
- Presence of binary and continuous variables

Moreover, the variables were collected at two different time points; thus, the investigation considered time variable. In this way, we were also looking at changes, continuities, and discontinuities observable in the analyzed territories. (Figure 6.2.1; Figure 6.2.2).

Figure 6.2.1. Example of data viewed at metropolitan area scale (Rome, land use year 2006).

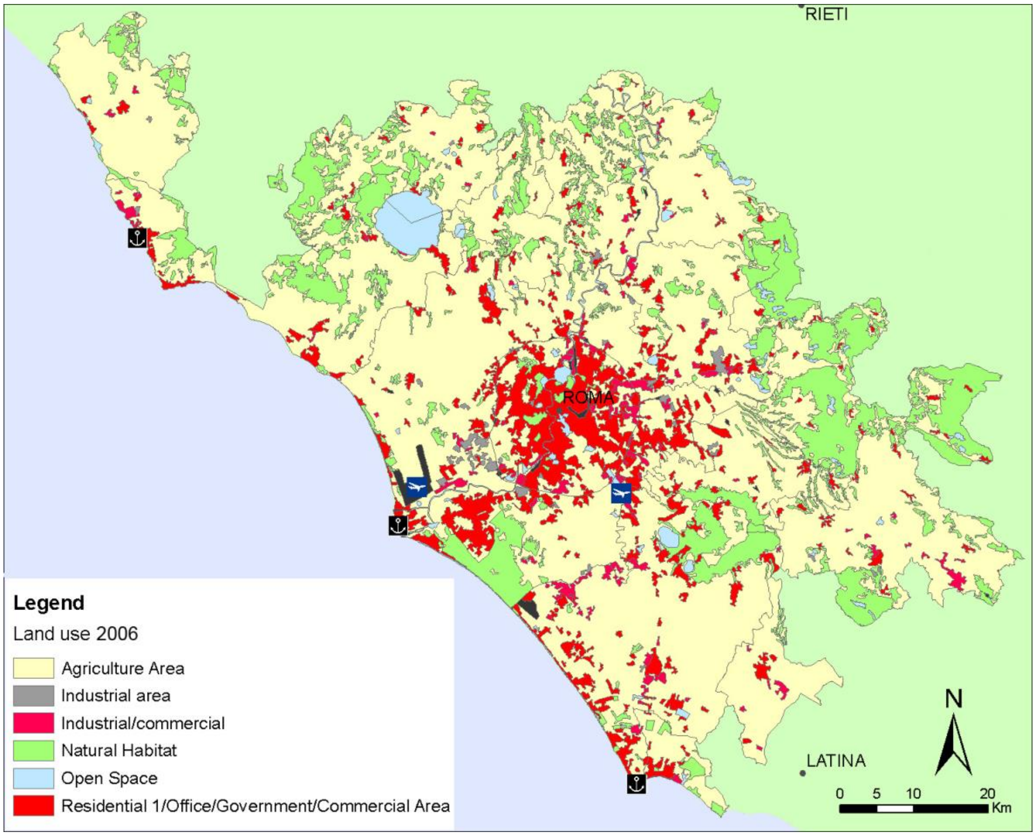
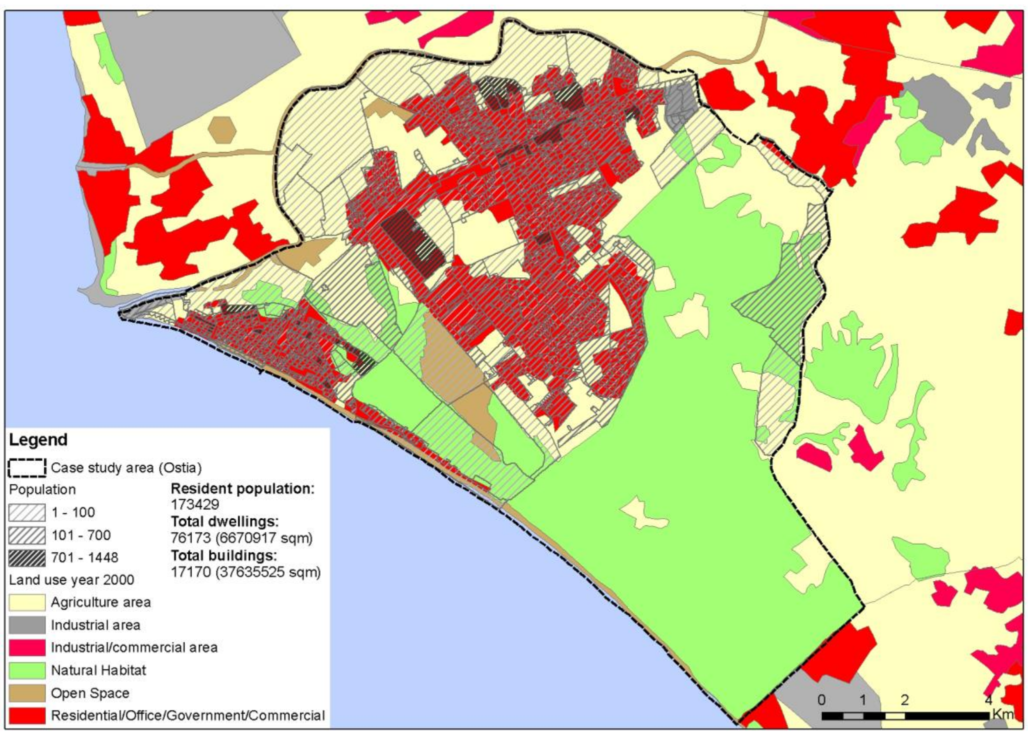


Figure 6.2.2. Example of data viewed at census scale (Rome, land use year 2006 of Ostia case study).



At this collecting phase, the use of metadata DB (provided along with data or available in the SECOA metadata website) was helpful, especially during the first phase of gathering datasets, in order to know how to treat datasets and analyze eventual discrepancies among data or at the end of the analysis.

For simplicity, we can organize the variables available for all case studies into five macro groups based on the treatment applied to data in GIS environment:

1. Socio-economic
2. Environmental
3. Natural hazards
4. Sustainability related
5. Existing policies

1) Since most socio-economic variables were in the form of Excel spreadsheets or .dbf tables, the first step was joining tables to their relative areas (polygons) constituted by municipalities or other areas with different grade of precision. The task was accomplished using an ID shared field (like municipality name/code) that act as a primary key for each attribute tables. In case of lack of data for one time period or for certain areas, special values were used in order to be "understandable" during the SOM processing, "telling the net" that the variable was missing.

2) European partners received the same information about land use because of the availability of Corine Land Cover photo interpretation made in several periods for the entire European Union area. Other partners provided interpretations of satellite images as well.

   The several land use categories (more than 40 in the case of Corine Land Cover) were then classified into these aggregations:

   - agriculture
   - natural habitat
   - residential mixed use
   - industrial
   - industrial/commercial
   - open space

For each of these categories, a "presence/absence" layer was created to be converted to raster.

3) The most of environmental and hazard data were in raster format. In these cases, the output had to be resized to match the 1km cell size. For some of them, like floods, landslides, and other kind of hazards, the variable is continuous and indicates, for example, the level of risk, while in others, like in case of sea level rise (SLR), a 1/0 approach indicated submerged areas.

4) Several different layers were included in this category, like sustainability indexes, water quality, air quality, presence of pollutants in water or air, flora and fauna at risk, presence of natural reserves.

5) Last category comprised binary values for Political System (Centralized/Decentralized, Unitary/Federal) and the presence/absence of Integrated Coastal Zone Management (ICZM), such as Environmental Impact Assessment (EIA), Planning Hierarchy (PH), Setback Lines (SL), Marine Spatial Planning (MSP), and Regulatory Commission (RC).

Although many data that compose these groups were the same for different countries, in certain cases it was necessary to use proxies in order to make a comparison between different countries and case studies: a clear example is the definition of "employed" variable, which ranged from "all person have been working at least one hour per week" to "who worked six months or more during the last one year" and so on, depending on the country of data acquisition. It is clear that also during this, it was necessary to look at the metadata in order to "weight" the inputs where necessary.

The question is, how to treat these datasets to analyze them in the best way?

All data analysis performed in a spatial context allow us to analyze events, recognize and interpret patterns, obtain results based on different scales, depending also on the scale of the original data (Haining 1994), so that, starting from the beginning, we can assert that the primary aim is to create new information, test hypotheses, and develop solutions.

Different aims and stages can be completed in 3 stages:

- Showing
- Analyzing
- Planning

The path that we see leads us directly toward Spatial Analysis and Spatial Modeling and, in this context, GIS assumes a role of not only a container, but also a tool that would allow us to add value to the original dataset after gathering all the data. The ability to analyze spatial data using different kind of spatial analyses provides this added value (Goodchild, 1988), so that, in general, we can say that spatial analysis has something to do with gaining new information from data using the spatial context of the problem of the data itself.

All this data can be managed in different ways, using several methodologies that often, but not necessarily, involve the use of GIS. Usually, we tend to divide these analyses into different categories, obviously considering that certain kinds of data allow only certain types of analyses. Typical differences in possibilities offered by vector and raster include:

- Simple queries on DB
- Measurements
- Transformations
- Interpolations
- Spatial queries
- Map algebra

Someone also stated (McDonnell, 1998) that depending on the conceptual model, we can divide all spatial operations into two different categories, "entities", which include attribute operations like distance/location operations or operations using built-in spatial topology, and "fields", like interpolations, terrain analyses, spatial filtering, and on the like. Nowadays, even if not officially, many of these terms and operations are codified in an almost standard language that can be found, with minor differences, in all most frequently used GIS software packages.

Beyond this, we must remember that besides all technical disquisitions, only the result is important.

3. From GIS to ANN and Back Again

We wanted to convert all datasets in a homogeneous format and then use them to link the GIS environment and the neural network simulator.

As mentioned before, the data, including variables and administrative boundaries, were in fact in different formats, mostly in geodatabases, but also in single shape files or grids and excel spreadsheets, while Vector data included polygons (boundaries), lines (roads, rivers), and points (air or water quality gathering stations) types.

After some considerations about the nature of the data and the analyses to be performed, we decided to apply a common approach using raster (GRID) data. The data available in this format were treated as described below, while in case of vector datasets, it was necessary to conduct a conversion operation. (Table 6.3.1; Figure 6.3.1).

Figure 6.3.1. Vector and raster.

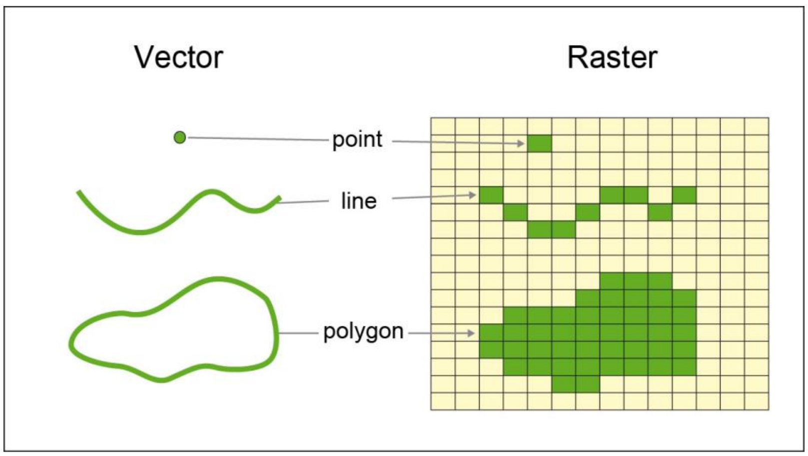


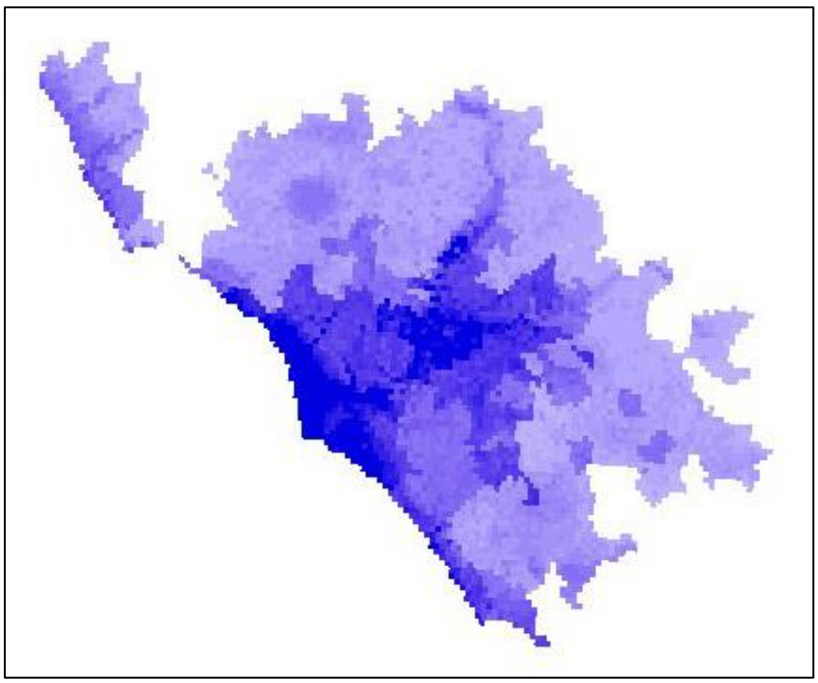
Table 6.3.1. Some characteristics of the vector and raster data (Longley et al. 2005).

Issue	Raster	Vector
Volume of data	depends on cell size	depends on vertices density
Sources of data	remote sensing, images, manipulation	social and environmental data
Applications	resources (enviromental)	social, economic, administrative
Software	raster GIS, image processing	vector GIS automated cartography
Resolution	fixed	variable

The choice of this treatment was based on several motivations explained before, i.e., mostly problems related to the different file formats, spatial and topological issues, and different scales. For example, since different institutions or teams collected the datasets of each case study, the boundaries were not exactly matching, causing some topological issues and leading us to deal with thousands of clipped polygons. In some cases, copyright or other kinds of legal restrictions issues applied from the data providers made it impossible to use "raw" variables at a detailed level, allowing us to use only their aggregations at zone level (Core and Rings).

Due to abovementioned issues, the analyses scale was fixed using square cells of 1x1 kilometers, which is a good representation of the territories without losing much information. Obviously, in the case of variables available only at a zone level, this procedure was not adopted because of the impossibility to obtain more information at that level (Figure 6.3.2).

Figure 6.3.2. Example of Rome metropolitan area at 1km resolution (Flood risk).



The study covered different areas in terms of extension too and the number of resulting cells after conversion varied from few hundreds to several thousands, as the table below shows (Table 6.3.2).

Table 6.3.2. The whole number of cells per case study.

Case study	Dimension	Pixel number
Roma	132 x 109	14388
Chieti-Pescara	54 x 44	2374
Lisboa	88 x 73	6424
Eastern Algarve	83 x 51	4233
London Thames Gateway	95 x 95	9025
Portsmouth	30 x 25	750
Malmö	55 x 67	3685
Gothenburg	75 x 98	7350
Brugge	26 x 35	910
Oostende	22 x 19	418

The entire process was done considering the final step of neural network analyses. Usually the problem of pre-packed software is the lack of control over processing, and that is why, from the beginning, we decided to use custom algorithms to process data since, by programming the single steps of the process, we had total control of the analysis flow, allowing us to treat and weight variables freely.

With this in mind, we had to choose a file format to "feed" the ANN system without problems. The requisites were mainly simplicity of import/read, compatibility with several software packages, ease of conversion, and the possibility of bringing the result back to GIS without losing spatial information. Considering these aspects, ASCII grids were adopted, in fact, every programming language can read this format easily. This is possible thanks to the very simple inner structure of this file type, as it is formed by a header containing information about georeferencing (X and Y coordinates of low left corner), extension (number of columns and rows), size of the pixels and no data value, and a body that contains all values. In practice, it does not differ so much from a text/tabulator delimited file, and at the same time, beyond this simple structure, there is no lack of information.

4. Masking the Areas of Interest

The final step involved isolating the areas of interest using the boundaries of each case study before transposing data to the ANN environment to perform the analyses. First, we need to point out that analysis mask is not analysis extent, this is often a common mistake, suggesting that each study area is like an island without anything around.

Since a raster or GRID data is formed by squares organized in rows and columns, the first step involved delimiting the total area of interest, in our case represented by the boundaries of the entire metropolitan areas. Using the given shape files relative to the zone organization (core and rings), we proceeded to create a total mask for each case study, so that all areas outside were directly excluded by inserting a NoData flag in the final ASCII grids. In this first stage, "masks" were essentially representing only the limits of the study areas with an aim to have a common boundary to be used during analyses. This is a typical case of analysis extent, since our data were not going beyond these boundaries, thus, it is not possible to know, beyond the boundaries, what the values of our variables are.

This step was necessary prior to performing the first SOM analyses, looking for clustering "similar" areas as explained in depth in chapter 7.

However, in the last stages of Feed Forward ANN analyses, there was also a different need, since a necessary step at that point was to consider the relationship between variables and conflicts.

All project partners were provided the GIS shape files relative to their conflict areas boundaries that were used to isolate the conflict itself. The support of experts who know the single territories in depth was given constantly at all stages of this project and was particularly useful during this stage.

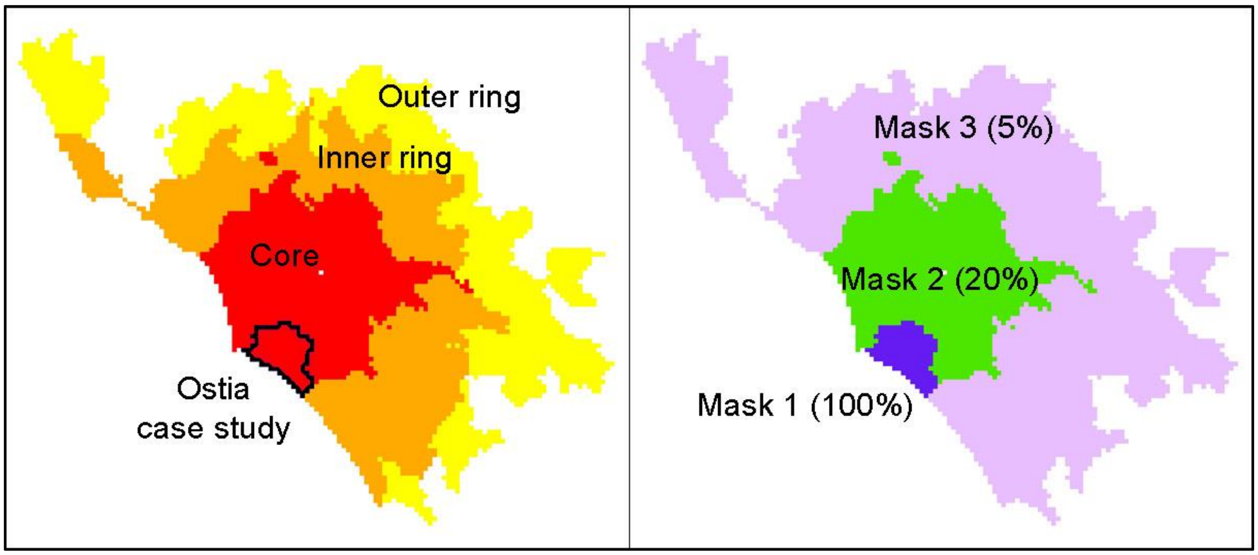
The last result was achieved using the conflict boundaries with an aim to divide and clip different zones, obtaining 3 different masks for each conflict:

- The actual conflict area
- The zone containing the conflict but without the conflict area
- The other zones

This choice was done because even if the area of conflict itself is directly related to its features, we wanted to compare them with surrounding areas. Moreover, even if it was necessary to indicate the conflict area for the purpose of the analysis, we can surely state that there are no real, physical barriers around the conflict. We decided to use these masks, in a certain way, as buffer zones useful for "mitigating" the input values and retaining their relationship with the surroundings.

The resulting masks were then used for ANN training in different ways in order to weight differently the relationship with single conflicts. After some tests, the data from the first mask was used at 100%, from the second only at 20%, and from the third at 5%, as explained more in depth in chapter 7 (Figure 6.4.1).

Figure 6.4.1. Example of masks for the Ostia case study.



Unfortunately, even if the main subject of comparison was the area of the conflict, it was not possible to reach that level of detail for all countries. The main factor was that for Asian countries, it was possible to georeference only few data, while the rest was available only in an aggregated way at zone level, making it impossible to use a more refined territorial approach while, at the same time, it was not correct to use values relative to the whole zone where the conflict is located.

This issue led us to the decision to perform 2 different kinds of analyses: one based entirely on geo-referred data that included all case studies except those from Asian countries, and another one that considered all case studies. The first kind of approach was adopted for Belgium, Israel, Italy, Portugal, Sweden, and the UK case studies while the other included also India and Vietnam.

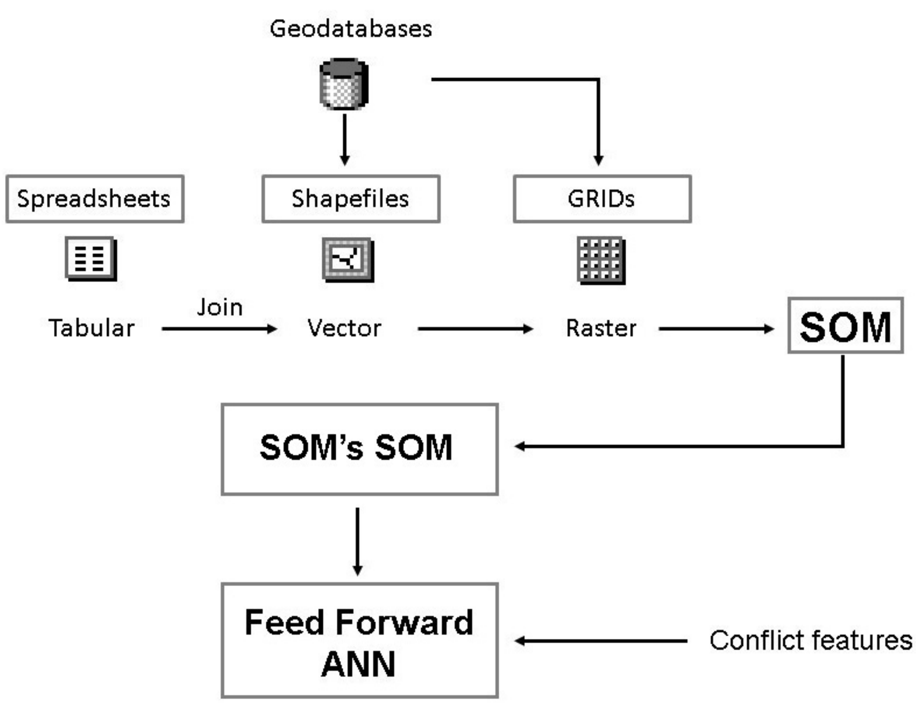
In both approaches, the variables relative to the areas of study were used as an input and the characteristics of the conflicts were used as an output. The only difference, since it was not possible to know the values of variables for India and Vietnam, was the use of a qualitative approach. The Asian countries conflicts experts were requested to complete a form that contained the same variables measured at two different time points as the other countries but with values ranging from 1 (minimum) to 5 (maximum). At the same time, values for all other countries were converted to the same form, considering maximum and minimum values for each country and qualitative values were converted to the quantitative ones, considering the minimum and maximum of each variable.

At that point, the idea was that since the conflict features are based essentially on the conflict areas themself, it was necessary to isolate them. The experts in different countries were asked to provide boundaries for every single conflict and the resulting polygons were applied to the GIS environment. In this case, we considered that conflict areas were part of the whole territory and that we had information about these surroundings.

5. Empirical Results and their Evaluation on Maps

Each single file was treated differently, as shown before, and the entire process from "raw" data to the chosen raster format and the final analysis, described in depth in chapter 7, can be reassumed in the following image (Figure 6.5.1).

Figure 6.5.1. Data processing.



Following the data treatment, the result was a huge dataset ready to be analyzed in a neural network environment, first by the unsupervised method of SOMs and then by supervised Feed Forward ANNs.

After performing the necessary analyses, the results always flow in maps representing clusters, aggregations and so on, depending on the analysis performed. The final map itself can be re-imported into every GIS environment. In this way, the spatial approach is never loose, as stated in the first phases of the processing workflow, and the results can also be used to perform new analyses.

Looking at the results on the maps allows us to see values and their representation directly on the territory: this is always much more informative than looking at the same results, for example, on a table or a graph. Moreover, by using the GIS technology, it is also possible to compare different territories in their spatial and time dimensions, looking at differences and common features in a synchronic and diachronic way.

In particular, the aim of the first SOM stage was to group similar areas based on variables that characterized themselves, leading us to create taxonomy of the investigated areas. Observing the results on a map, it is possible, at a first look, to conceive some connections between zones and clusters, since it is often possible to see a parallelism with the schema of core, inner, and outer ring.

Moreover, the Feed Forward supervised approach helped us identify the basics of a model, relating the variables of territories to the conflicts features. This gives us the ability of to stimulate, as variables change, the effect on the conflict, helping us define policies and program a sustainable management of the coastlines as part of the project purposes.

Luca DERAVIGNONE, Department of European, American and Intercultural Studies, Sapienza University of Rome, Italy.

email: luca.deravignone@uniroma1.it

Deravignone, Luca, "GIS Data and Territorial Management Approach". In: Armando Montanari (ed.), *Urban Coastal Area Conflicts Analysis Methodology: Human Mobility, Climate Change and Local Sustainable Development*. Cap. 6. *SECOA FP7 Research Project*, Vol. 5. Rome: Sapienza Università Editrice, 2013. *http://digilab-epub.uniroma1.it*. DOI: 10.13133/978-88-98533-01-5. Web. 30 September 2013.

ABSTRACT: SECOA is strongly territorial, which led us to adopt GIS technology since the very beginning of the project. The analyses were based on the data collected from several metropolitan-urban areas in Europe and Asia where the analyzed conflicts were considered in relationship to the territory to which they belong, since they exist and develop in a real rather than a virtual space. Almost entire dataset was territorial, and the results provided a huge amount of data organized in tables, vector and raster data. The necessity to manage such wide range of spatial data, considering also their heterogeneity and collecting scale, led us to conceive a certain workflow, converting all datasets into a homogeneous format and analyzing the resulting dataset with a Neural Network approach. In this sense, the idea was to create a sort of “bridge” between the GIS environment and the Neural Network simulator in order to bring the results back to GIS without losing spatial information. With the help of GIScience, we managed to analyze and compare all territories through their spatial dimension, in their real context.

KEYWORDS: GIS, GIScience, GISystem, Raster Analysis, Geodatabases, Neural Networks

---

Submitted:	6th June 2013
Accepted:	18th July 2013
Final Revision:	23th July 2013
Published:	30th September 2013

Annex I.

CHAPTER 7.
Application of Artificial Neural Networks to SECOA Project

Alessandro LONDEI
Department of European, American and Intercultural Studies,
Sapienza University of Rome, Italy

1. SECOA Taxonomy: Application of Local and Global SOM

1.1. Introduction

According to the material presented in Chapter 5.4, Self-Organizing Map (SOM) neural networks were chosen as a suitable tool for the analysis of SECOA in order to address the following characteristics and issues:

1. SOM can easily manage high-dimensional data by reducing the overall dimensionality and representing the data proximity characteristic on a two-dimensional surface.
2. Up to a reasonable level, SOM is able to deal with records containing missing data. This aspect is very important in the present analytical approach since some territories lack sensitive data.
3. SOM allows extracting specific prototypes closely related to subsets of data. The prototypes are then representative of the data and can take into account a more synthetic data description.

As discussed in chapter 6, SECOA data can be divided into three main classes:

1. Raster data from GIS maps describing socio-economic-environmental (S/E/E) aspects in every single pixel of the raster map. The territories related to this class are from European countries.
2. S/E/E data from macro-areas of Asian countries (with the exception of Funchal case study).
3. Data representing the conflict characteristics of all the cases involved in SECOA project.

Depending on the class, different aspects characterized the SOM approaches and analytical procedures. In general, data elaborations by SOM followed a common roadmap:

1. Definition of a data matrix comprised vectors (rows) of which elements were the variables involved in the specific analysis. This matrix may present some missing data.
2. Training of a suitably dimensioned SOM.
3. Cluster analysis (k-means) of the original data and representation of the clusters on the SOM neural surface.
4. Selection of prototypes (Best-Matching Units) belonging to every cluster.
5. Identification of cluster trend by means of cluster prototypes means and standard deviations.

This approach allowed both to reduce data dimensionality by distributing high-dimensional data on a 2-dimensional neural surface and to reduce the outcomes complexity by means of a simple description of significant variables defined by averaged prototypes.

1.2. SOM and GIS Raster Data

GIS raster data were associated with the following 10 metropolitan areas: Rome (Italy), Chieti-Pescara (Italy), Thames Gateway (United Kingdom), Portsmouth (United Kingdom), Gothenburg (Sweden), Malmö (Sweden), Zeebrugge (Belgium), Ostend (Belgium), Lisbon (Portugal), and Algarve (Portugal). Each GIS consists of a rectangular grid of pixels with dimensions of 1 Km by 1 Km. The 19 variables describing every territory are: average income, net migrants, motorization rate, number of households, number of residents, number of tourists, unemployment rate, employees in agriculture sector, employees in industrial sector, employees in services sector, agriculture land use, airports land use, industrial land use, industrial-commercial land use, natural land use, open-space land use, residential-mixed land use, flood hazard, and sea level rise. The whole set of variables consists of 38 elements since every variable is considered twice, one for the first period (up to 2000) and one for the second period (since 2001).

Table 7.1.2.1. Territories dimension (in pixels).

Territory	Dimensions	Total pixels	Data pixels	No-data pixels
Rome	132 x 109	14388	5303	9085
Chieti-Pescara	54 x 44	2376	907	1469
Thames Gateway	95 x 95	9025	3713	5312
Portsmouth	30 x 25	750	356	394
Gothenburg	75 x 98	7350	3905	3445
Malmö	55 x 67	3685	1465	2220
Zeebrugge	26 x 35	910	599	311
Ostend	22 x 19	418	194	224
Lisbon	88 x 73	6424	2982	3442
Algarve	83 x 51	4233	2175	2058

Table 7.1.2.1 shows the dimension of every territory (in pixels), the total amount of pixels, and the amount of data and no-data pixels (sea and unconsidered geographical zones).

Every territory was analysed by a specific SOM of which surface of neurons was determined by the amount of data pixels involved based on a heuristic formula. The hexagonal map grids used for SOMs are listed in Table 7.1.2.2.

Table 7.1.2.2. Territories SOM dimension (in units).

Territory	SOM Hexagonal Grid	Total neurons
Rome	30 x 12	360
Chieti-Pescara	17 x 9	153
Thames Gateway	24 x 13	312
Portsmouth	14 x 7	98
Gothenburg	28 x 11	308
Malmö	22 x 9	198
Zeebrugge	17 x 7	119
Ostend	12 x 6	72
Lisbon	21 x 13	273
Algarve	22 x 11	242

Every SOM was trained with a data matrix of which dimensions were given by number of data pixels X number of variables. For instance, Rome data matrix comprised 5303 rows (data pixels) and 38 columns (variables).

After the training phase, all data vectors were associated with a specific Best-Matching Unit (BMU) of the SOM, known as a Prototype. The connection weights considered as a vector connected to any BMU represent the best approximation of the data vectors associated with the BMU. The number of prototypes selected in every SOM is listed in Table 7.1.2.3.

Table 7.1.2.3. Territories SOM prototypes.

Territory	Number of Prototypes
Rome	211
Chieti-Pescara	49
Thames Gateway	178
Portsmouth	50
Gothenburg	25
Malmö	26
Zeebrugge	38
Ostend	30
Lisbon	113
Algarve	75

A cluster analysis based on the k-mean clustering algorithm was then performed on the original data. The best number of clusters for each territory was determined by the evaluation of the Davies-Bouldin index, a metric for evaluating clustering algorithms (Davies & Bouldin, 1979).

Table 7.1.2.4. Territories cluster population (pixels and prototypes).

Territory	**Clusters**	**N. of pixels in cluster**	**N. of prototypes in cluster**
Rome	8	21, 55, 18, 83, 165, 3078, 613, 1270	1, 3, 2, 7, 9, 118, 24, 47
Chieti-Pescara	6	117, 39, 29, 601, 69, 52	11, 4, 4, 25, 2, 3
Thames Gateway	8	7, 13, 1138, 1790, 95, 148, 180, 342	1, 6, 50, 59, 7, 24, 7, 24
Portsmouth	4	31, 173, 36, 116	8, 14, 11, 17
Gothenburg	7	976, 2566, 0, 3, 45, 1, 314	5, 11, 0, 1, 3, 1, 4
Malmö	6	1092, 40, 17, 310, 3, 3	16, 2, 1, 4, 1, 2
Zeebrugge	5	68, 84, 51, 261, 135	5, 8, 3, 14, 8
Ostend	5	15, 63, 24, 26, 66	3, 4, 5, 7, 11
Lisbon	8	660, 190, 379, 120, 41, 215, 1045, 332	18, 9, 24, 5, 4, 7, 26, 20
Algarve	7	212, 291, 65, 4, 98, 638, 867	12, 5, 4, 2, 5, 28, 19

Table 7.1.2.4 shows the amount of clusters for any territory and the number of pixels and prototypes per cluster.

To decrease the complexity of the S/E/E data, the analysis of the behaviour of each cluster was performed on prototypes rather than on pixels. Figures 7.1.2.1 to 7.1.2.10 in *Annex to Chapter 7* shows the distribution of clusters in each territory.

Distributions of clusters over the SOM surface are represented in Figure 7.1.2.11(a to j) in *Annex to Chapter 7*.

In order to support further analysis, every cluster was characterized by a series of parameters evaluated by variables. Variables in prototypes belonging to any cluster were averaged and the standard deviation was calculated. Hence, each cluster may be defined by means of averaged value of significant variables. Since variables were normalized in order to be evaluated by SOM, at any cluster was associated a vector composed of normalized variables averaged on all prototypes contained in that specific cluster. Values are distributed around zero (mean), and they can be positive or negative depending on their relations with the averaged value. Variables are not taken into account if the Student statistical test validates the null hypothesis.

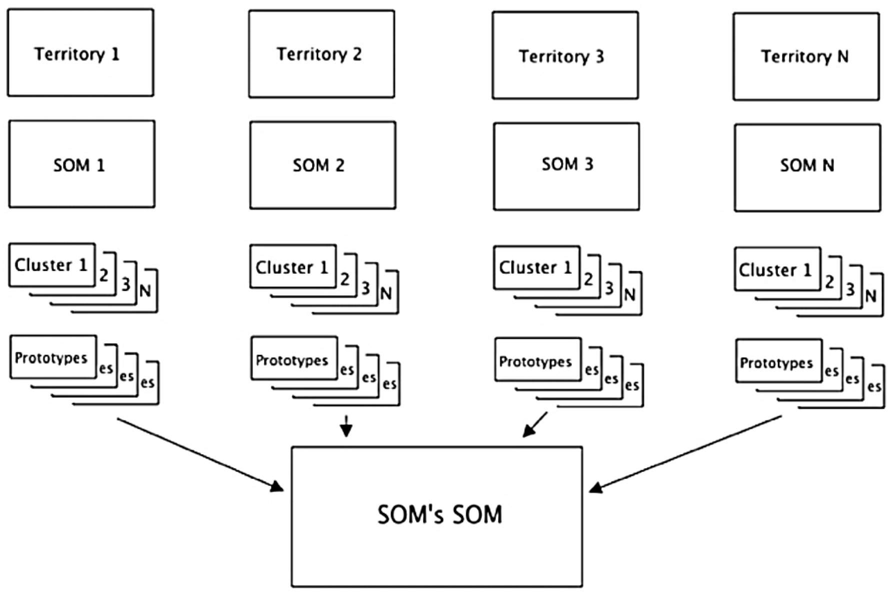
Figures from 7.1.2.12 to 7.1.2.21 and tables from 7.1.2.5 to 7.1.2.24 in *Annex to Chapter 7* represent the evolution of averaged variables of prototypes belonging to any cluster. Red dots represent significant non-null variables.

1.3. SOM's SOM and GIS

The global analysis of all territories as a whole was not allowed because of some differences among S/E/E data. For instance, employees' variables in some cases describe number of units instead of the real values of employees. Since each territory was described by either employees or units, it was decided to apply SOM algorithm on these kinds of variables as if they were uniform to each other, since in this specific analysis, the outcome of the SOM is a class of trends represented by the cluster prototypes.

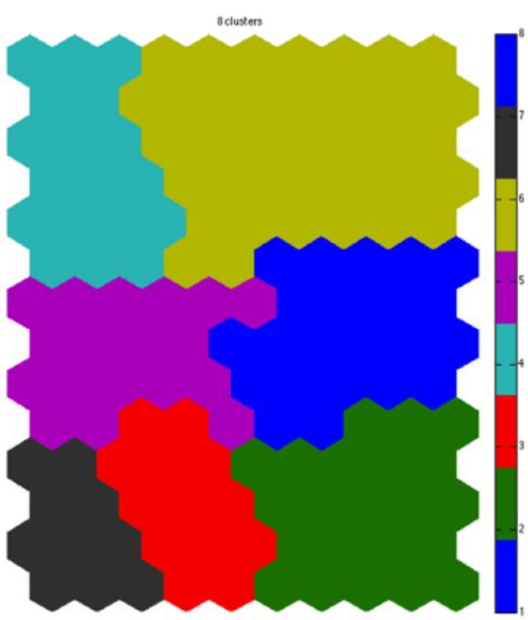
A different approach was applied to evaluate a global analysis. All prototypes belonging to a specific cluster from every territory were selected to create a new data set for a SOM analysis devolved to a global synthesis of the territory features. Every prototype was described by the connection weights of the BMU of the related SOM and was labelled by the name of the territory and the cluster number to which is belongs (Figure 7.1.3.1).

Figure 7.1.3.1. Blocks diagram of SOM's SOM procedure.



The SOM's SOM was then trained with a data matrix composed of 795 rows (prototypes) by 38 columns (variables). Note that a specific prototype was taken only once even if it appeared several times in the same cluster. The SOM surface was composed of 14 x 10 neurons distributed on a hexagonal grid. Cluster analysis applied to the prototypic data indicated 8 clusters distributed on the neural surface as shown in the following Figure 7.1.3.2.

Figure 7.1.3.2. 8 clusters SOM's SOM.



Therefore, every cluster contains prototypes related to the original prototypic data connected to some specific cluster territory.

Table 7.1.3.1. Sub-territories in SOM's SOM Clusters.

Cluster #1	**Cluster #2**	**Cluster #3**	**Cluster #4**	**Cluster #5**	**Cluster #6**	**Cluster #7**	**Cluster #8**
Rome 1	Rome 8	Rome 2	Rome 4	Rome 5	Rome 6	Thames Gateway 2	Rome 3
Chieti Pescara 1	Chieti Pescara 2	Chieti Pescara 1	Rome 7	Chieti Pescara 3	Chieti Pescara 1	Thames Gateway 6	Lisbon 3
Chieti Pescara 4	Chieti Pescara 3	Chieti Pescara 2	Chieti Pescara 5	Chieti-Pescara 4	Chieti Pescara 4		Lisbon 8
Lisbon 3	Chieti-Pescara 6	Chieti-Pescara 3	Lisbon 1	Lisbon 1	Lisbon 1		Algarve 7
Lisbon 5	Lisbon 6	Lisbon 2	Lisbon 3	Lisbon 3	Lisbon 7		Portsmouth 1
Lisbon 7	Lisbon 8	Lisbon 3	Lisbon 7	Lisbon 5	Algarve 1		Portsmouth 4
Lisbon 8	Algarve 5	Lisbon 4	Algarve 2	Algarve 1	Algarve 2		Thames Gateway 3
Algarve 1	Algarve 7	Algarve 4	Algarve 3	Algarve 7	Algarve 6		Gothenburg 4
Algarve 6	Portsmouth 1	Portsmouth 1	Algarve 5	Portsmouth 3	Portsmouth 2		Malmö 1
Algarve 7	Thames Gateway 3	Portsmouth 3	Algarve 6	Portsmouth 4	Portsmouth 4		
Portsmouth 2	Zeebrugge 5	Thames Gateway 5	Portsmouth 2	Thames Gateway 3	Thames Gateway 4		
Thames Gateway 1	Ostend 1	Zeebrugge 3	Portsmouth 4	Zeebrugge 2	Thames Gateway 7		
Thames Gateway 3	Ostend 3	Gothenburg 5	Thames Gateway 7	Ostend 4	Zeebrugge 1		
Thames Gateway 4	Gothenburg 1	Malmö 3	Thames Gateway 8	Ostend 5	Zeebrugge 3		
Thames Gateway 8	Malmö 4	Malmö 6	Zeebrugge 2	Gothenburg 2	Zeebrugge 4		
Gothenburg 1			Ostend 4	Malmö 1	Ostend 2		
Gothenburg 6			Gothenburg 2		Ostend 4		
Malmö 1			Malmö 1		Ostend 5		
Malmö 5			Malmö 2		Gothenburg 2		
					Gothenburg 7		
					Malmö 1		

Table 7.1.3.1 shows the original territories (sub-territories) related to any SOM's SOM cluster (the number denotes the original territory cluster).

Pictures in Figure 7.1.3.3 to 7.1.3.10 in *Annex to Chapter 7* show the original clusters territory related to any SOM's SOM cluster.

Even in the present analysis, every SOM's SOM cluster was described by the prototypes it contains. The prototype averaged values and the relative standard deviations are shown in the picture 7.1.3.11 in *Annex to Chapter 7*.

Variables significance of clusters is described in tables 7.1.3.2 and 7.1.3.3 in *Annex to Chapter 7*.

1.4. SOM: Asian and Funchal Macro-areas Data

The data gathered from Asian countries and from Funchal were very dissimilar to each other as well as compared to the European countries data for the following reasons:

1. Every Asian country and Funchal was described by different variables.
2. Asian countries and Funchal data represent only the S/E/E situation for macro-areas, which depending on the specific case, can be core, inner ring, middle ring, outer ring (or a generic ring) and coastal zone.

Because of these characteristics, SOM algorithm was applied to 5 different groups of data: Israel (Tel Aviv and Haifa), India (Mumbai and Chennai), Haiphong (in Vietnam), Nha Trang (in Vietnam) and Funchal (in Portugal).

The specific territories are described as following:

- Israel:

  7 macro-areas: Tel Aviv Core, Tel Aviv Inner Ring, Tel Aviv Middle Ring, Tel Aviv Outer Ring, Haifa Core, Haifa Inner Ring, Haifa Outer Ring.

  28 variables: total population, total number of in-migrants over the period, total number of out-migrants over the period, number of persons aged 0-17 who have changed address, number of persons aged >= 18 who have changed address, total residents of high social status (professional and managerial), total residents of middle social status (supervisory, other non manual, skilled manual), total residents in lower social status (unskilled and semi skilled manual), average household size (number of persons including children), number of persons commuting on daily basis for work related reasons in within the same sub-area, number of persons commuting on daily basis for work related reasons from one sub-area to another, peak number of passengers in within the same sub-area, peak number of passengers from one sub-area to another, employees in agriculture sector, employees in industrial sector, employees in private services sector, employees in public services sector, total area in sqmt, agriculture total area in sqmt, mixed industrial-commercial total area in sqmt, government total area in sqmt, natural habitat total area in sqmt, urban open space total area in sqmt, open space total area in sqmt, residential total area in sqmt, road network total length in km, extreme rainfall in mm (95 percentile), sea level rise.

2 time periods: up to 2001 and since 2002

- India:

  4 macro-areas: Mumbai Core, Mumbai Ring, Chennai Core, Haifa Ring

  23 variables: employees in agriculture, employees in industrial sector, employees in services, net migration, average number of tourists per day, max 1-day rainfall, slum population, total number of temporary dwellings, number of registered vehicles, total population, total in migration over the period, international in migrants over the period, national in migrants (beyond ring) over the period, average daily number of bus arrivals, average household size (number of persons including children), total area in sqmt, agriculture total area in sqmt, built up area in sqmt, natural habitat total area in sqmt, open space total area in sqmt, road network total length in km, sustainability index, sea level rise.

  2 time periods: up to 2001 and since 2002

- Haiphong:

  3 macro-areas: Haiphong Core, Haiphong Ring, Haiphong Coastal Zone

  22 variables: total population, international in migrants, national in migrants (beyond ring), in migrants from the metropolitan ring , in migrants from the metropolitan core, total out migrants over the period, average daily number of private motor vehicles on the roads, total land area (ha), total area of land allocated to economic activity (excluding agriculture or recreational sports areas), total area of land allocated to housing, total number of dwellings (including apartments within individual buildings and single family dwellings), total number of shanty or 'temporary' dwelling, number of households, total area in sqmt, agriculture total area in sqmt, industrial-commercial total area in sqmt, mixed total area in sqmt, residential total area in sqmt, extreme rainfall in mm (95 percentile), sustainability index, number of tourists, inundation.

2 time periods: up to 2001 and since 2002

- Nha Trang:

2 macro-areas: Nha Trang Core, Nha Trang Ring

12 variables: total population, total in migrants over the period, total out migrants over the period, average number of tourists per day, total land area (ha), average household size (number of persons including children), total area in sqmt, Extreme rainfall in mm (95 percentile), sustainability index, agriculture total area in sqmt, number of tourists, inundation.

2 time periods: up to 2001 and since 2002

- Funchal:

2 macro-areas: Funchal Core, Funchal Ring

35 variables: total population, total in migration, total out migration, international in migrants, total residents of high social status (professional and managerial), total residents of middle social status (supervisory, other non manual, skilled manual), total residents in lower social status (unskilled and semi skilled manual), total number of second homes, total number of temporary residents present (on census night), total number of higher education students from outside the metropolitan area, average number of tourists per day, average daily number of private motor vehicles on the roads, total land area (ha), total area of undeveloped land (inc agricultural land) (ha), total area of land allocated to economic activity (excluding agriculture or recreational sports areas), total area of land allocated to housing, total number of dwellings (including apartments within individual buildings and single family dwellings), total number of apartments, total number of shanty or 'temporary' dwellings, average household size (number of persons including children), number of persons commuting on daily basis for work related reasons in within the same sub-area, number of persons commuting on daily basis for work related reasons from one sub-area to another, employees in agriculture sector, employees in industrial sector, motorization rate per 100 inhabitants, number of households, rate of unemployment, total area in sqmt, agriculture total area in sqmt, mixed industrial-commercial total area in sqmt, natural habitat total area in sqmt, residential use: low density in sqmt, residential use: high density in sqmt, road network total length in km, Extreme rainfall in mm (95 percentile).

2 time periods: up to 2001 and since 2002

A preliminary selection of variables was performed before applying SOM algorithm to exclude the variables containing less than 50% of valid data. Moreover, the total number of variables listed above is doubled since the effect of time period was associated with every variable. After the variable selection, the territories were described as follows (the number next to any variable indicates the time point; a synthetic renaming of variables was adopted):

- Israel:

7 macro-areas: Tel Aviv Core, Tel Aviv Inner Ring, Tel Aviv Middle Ring, Tel Aviv Outer Ring, Haifa Core, Haifa Inner Ring, Haifa Outer Ring

42 variables: Population 1, Population 2, In-migrants 2, Out-migrants 2, Changed Address 0-17 1, Changed Address 0-17 2, Changed Address >=18 1, Changed Address >=18 2, Residents High Social 1, Residents High Social 2, Residents Middle Social 1, Residents Middle Social 2, Residents Lower Social 1, Residents Lower Social 2, Household size 1, Household size 2, Commuting In 1, Commuting In 2, Commuting Out 1, Commuting Out 2, Peak Passengers In 1, Peak Passengers Out 1, Employees Agriculture 1, Employees Agriculture 2, Employees Industrial 1, Employees Industrial 2, Employees Private Services 1, Employees Private Services 2, Employees Public Services 1, Employees Public Services 2, Total Area 1, Agriculture Area 1, Mixed Industrial-Commercial Area 1, Government Area 1, Natural Area 1, Urban Open Space Area 1, Open Space Area 1, Residential Area 1, Road Net Length 1, Extr. Rainfall 2, SLR 1, SLR 2

- India

4 macro-areas: Mumbai Core, Mumbai Ring, Chennai Core, Haifa Ring

39 variables: Employees in Agriculture 1, Employees in Agriculture 2, Employees in Industrial 1, Employees in Industrial 2, Employees in Services 1, Employees in Services 2, Net Migration 1, Net Migration 2, Rainfall 1, Rainfall 2, Slum Population 1, Slum Population 2, Temporary Dwellings 1, Temporary Dwellings 2, Vehicles 1, Vehicles 2, Population 1, Population 2, In migrants 2, International In migrants 2, National In migrants 2, Household Size 1, Household Size 2, Total Area 1, Total

Area 2, Agriculture Area 1, Agriculture Area 2, Built up Area 1, Built up Area 2, Natural Area 1, Natural Area 2, Open Space Area 1, Open Space Area 2, Road Net Length 1, Road Net Length 2, Sustainability Index 1, Sustainability Index 2, SLR 1, SLR 2

- Haiphong

  3 macro-areas: Haiphong Core, Haiphong Ring, Haiphong Coastal Zone

  24 variables: Population 1, International In Migrants 1, National In Migrants 1, In Migrants from Ring 2, In Migrants from Core 2, Land Area 2, Housing Area 2, Dwellings 1, Temporary Dwellings 1, Households 1, Total Area 2, Agriculture Area 1, Agriculture Area 2, Industrial-Commercial Area 1, Industrial-Commercial Area 2, Residential Area 1, Residential Area 2, Rainfall 2, Sustainability 1, Sustainability 2, Tourists 1, Tourists 2, Inundation 1, Inundation 2

- Nha Trang

  2 macro-areas: Nha Trang Core, Nha Trang Ring

  16 variables: Population 1, Population 2, In Migrants 2, Out Migrants 2, Land Area 1, Land Area 2, Households 1, Households 2, Total Area 2, Rainfall 2, Agriculture Area 1, Agriculture Area 2, Tourists 1, Tourists 2, Inundation 1, Inundation 2

- Funchal

  2 macro-areas: Funchal Core, Funchal Ring

  42 variables: Population 1, Population 2, In migrants 1, In migrants 2, Out migrants 1, Out migrants 2, International In migrants 1, International In migrants 2, High Social Residents 1, High Social Residents 2, Middle Social Residents 1, Middle Social Residents 2, Lower Social Residents 1, Lower Social Residents 2, Second Homes 1, Second Homes 2, Temporary Residents 1, Temporary Residents 2, Tourists 2, Vehicles 2, Land Area 2, Undeveloped Land Area 2, Economic Area 2, Housing Area 2, Dwellings 1, Dwellings 2, Apartments 1, Apartments 2, Temporary Dwellings 1, Temporary Dwellings 2, Household Size 1, Household Size 2, Commuting In 2, Commuting Out 2, Agriculture Employees 2, Industrial Employees 2, Motorization rate 1, Motorization rate 2, Households 1, Households 2, Rate of Unemployment 2, Extr Rainfall 2

The SOMs characteristics of every territory, as number of neurons and data dimension, are shown in Table 7.1.4.1.

Table 7.1.4.1. SOMs characteristics for Asian territories + Funchal.

Territory	SOM Hexagonal Grid	Total neurons	Data vectors (macro-areas)	Data variables
Israel	5 x 3	15	7	42
India	3 x 3	9	4	39
Haiphong	5 x 2	10	3	24
Nha Trang	3 x 3	9	2	16
Funchal	3 x 3	9	2	42

Table 7.1.4.2. SOMs clusters and prototypes for Asian territories + Funchal.

Territory	Number of clusters	Number of prototypes	Prototypes per cluster
Israel	3	7	2, 2, 3
India	3	4	1, 2, 1
Haiphong	3	3	1, 1, 1
Nha Trang	2	2	1, 1
Funchal	2	2	1, 1

Table 7.1.4.2 shows the number of clusters of territories and the number of prototypes.

The distribution of clusters on the SOM neural surface is shown in Figure 7.1.4.1 (a-e) in *Annex to Chapter 7*.

Figures from 7.1.4.2 to 7.1.4.6 and tables from 7.1.4.3 to 7.1.4.12 in *Annex to Chapter 7* represent the evolution of averaged variables of prototypes belonging to each cluster. Red dots represent significantly non-null variables.

1.5. Conflicts: Description of SOM Outcomes

The 27 SECOA conflicts were described by 44 variables. The conflicts type and associated zones are shown in Table 7.1.5.1.

Table 7.1.5.1. SECOA conflicts type and zones.

Country	Case	Labels	Conflict Name	Description of the conflict	Area
ITA	Rome	Rom1	Civitavecchia	Civitavecchia - air pollution	Inner Ring
ITA	Chieti-Pescara	ChPe	The Costa Teatina National Park	Costa Teatina National Park - protected area	Inner Ring
ITA	Rome	Rom2	Ostia water-use and management	Ostia - water management	Core
BEL	Ostend	Oste	Ostend airport	Privatization and anticipated expansion of the regional airport Ostend	Core
BEL	Zeebrugge	Bru1	Schipdonk canal	"Widening of the Schipdonk canal"	Core and Ring
BEL	Zeebrugge	Bru2	Zeebrugge harbour	"Expansion of the inland harbour of Zeebrugge"	Core and Ring
PRT	Lisbon	Lisb	Trafaria and Costa da Caparica	Trafaria/Costa da Caparica	Core
PRT	Algarve	Ealg	Barrier Islands - ria Formosa Natural Park	Barrier Islands	Core and Ring
PRT	Funchal	Func	Funchal Bay - Madeira Island	Funchal Bay	Core
UK	Thames Gateway	Tha1	Barking riverside	Barking Riverside urban regeneration	Ring
UK	Thames Gateway	Tha2	Lower Thames crossing	Lower Thames crossing proposal	Ring
UK	Portsmouth	Por1	langstone Harbour/ Farlington marshes	Farlington Marshes	Core
UK	Portsmouth	Por2	Tipner regeneration	Tipner Regeneration	Core
ISR	Tel Aviv	Tel1	Palmachim beach	Palmachim coastal area	Outer Ring
ISR	Tel Aviv	Tel2	Netanya sandstone cliffs	Nethania coastal cliff	Outer Ring
ISR	Haifa	Haif	Haifa port	"The new Haifa Port - coast sharing of sand extraction for land reclamation" and "The Haifa Bay: city-port-park-airport conflict resolution in the context of a transformation from hierarchical government to fragmented governance"	Core
SWE	Malmö	Mal1	Managing urban sprawl	"Managing urban sprawl in the Malmö area";	Core, Inner ring, Outer Ring
SWE	Malmö	Mal2	Falsterbo	"Falsterbo peninsula"	Outer Ring
SWE	Gothenburg	Got1	Torsviken	Torsviken	Inner Ring & Core
SWE	Gothenburg	Got2	Kungsbacka	Kungsbacka	Outer Ring
IND	Mumbai	Mum1	SGNP	Sanjay Gandhi National Park	Core
IND	Chennai	Chen	Pallikaranai marshland	Pallikaranai Marshland	Ring
IND	Mumbai	Mum2	Mangrove forest	Mangrove forest	Core and Ring
VNM	Haiphong	HaP1	Haiphong port	"Haiphong port"	Coastal Zone
VNM	Haiphong	HaP2	Industrial zone	"Industrial zone"	Core, Ring and Coastal Zone
VNM	Haiphong	HaP3	Cat Ba	"Cat Ba"	Coastal Zone
VNM	Nha Trang	NhaT	Nha Trang	"Nha Trang"	Coastal Zone (Core)

The variables defining conflicts are shown in Table 7.1.5.2.

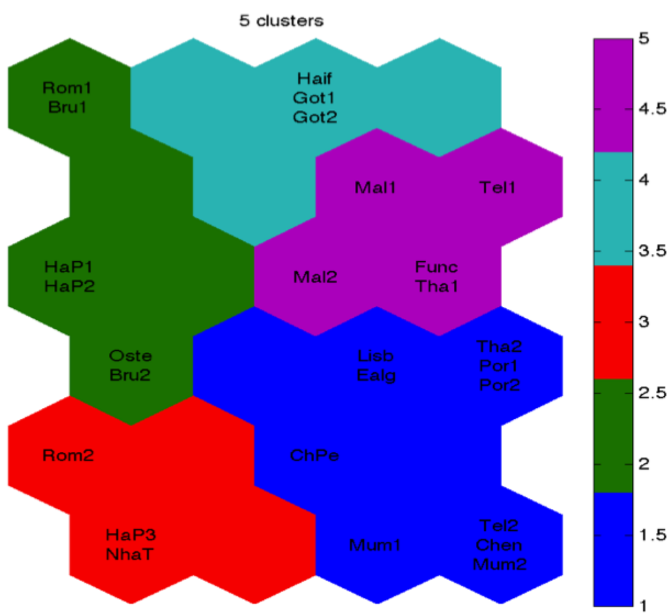
Table 7.1.5.2. SECOA conflicts variables.

Class	**Type**	**Description**	Labels
Category of uses	BINARY VARIABLE (0-1)	Ports and harbours related uses	CU/PH
		Urban Growth and Development in terms of specific urban functions including tourism	CU/UGD
		Energy generation/production	CU/EGP
		Natural environment and habitat including national parks and protected areas	CU/NEHP
Economic Development Vs. Environmental protection	RANKING 0 - 5 (Min.-Max.)	Port / harbour restructuring, expansion & infrastructure	EDEP/PHI
		Tourism lead infrastructure & urban development	EDEP/TIUD
		Industrial zones and airports expansion	EDEV/IZAE
		Waterfront & brown fields regeneration	EDEP/WBR
		Energy / power generation	EDEP/EPG
		Pollution [air, soil, water] associated	EDEP/PO
		Landscape / nature conservation	EDEP/NC
		Wildlife habitat protection	EDEP/HP
		Parks / Beaches / protected areas	EDEP/PA
		Waste water management	EDEP/WWM
Preservation of natural sites and biodiversity	RANKING 0 - 5 (Min. – Max.)	Preservation of natural sites / islands / marshlands	PNB/PNS
		Protection of biodiversity habitats	PNB/PBH
		National Parks	PNB/NP
		Cultural heritage & landscapes	PNB/CHL
Human mobility and contrast for use of resources	RANKING 0 - 5 (Min. – Max.)	Physical Infrastructure [transport, utilities & waste water]	HM/PIN
		Social infrastructure [Migrants / new comers, social exclusion / segregation, slums]	HM/SIN
		Tourism [housing, recreation & second homes]	HM/TOU
		Commuting (job related, daily, occasionally)	HM/COM
Dynamics/manifestation over time	BINARY VARIABLE (0-1)	Chronic	DY/CH
		Anticipation	DY/AN
		Hybrid	DY/HY
		Hushed or Deferred	DY/HUDE
Underlying cause/substance	BINARY VARIABLE (0-1)	Change in resource quality and availability	C/RES
		Legal/policy reasons	C/LP
		Infringements over access	C/INF
		Authority over resource	C/AUT
		Conflicts that are value based	C/VA
Scale	BINARY VARIABLE (0-1)	micro-macro conflicts	SC/MIMA
		inter micro-micro conflicts	SC/MIMI
		intra micro-micro conflicts	SC/IMIMI
		Hybrid	SC/HY
		between local and territorial/regional scales	SC/LR

Stage	BINARY VARIABLE (0-1)	Endurance	ST/EN
		Management	ST/MANA
		Formation	ST/FOR
		Transformation	ST/TRA
		Manifestation	ST/MANI
Final Ranking	(1 - 5) (Min.-Max.)	Criticality	FR/CRI
		Duration	FR/DUR
		Urgency	FR/URG

The SOM algorithm was applied to the conflict data by a hexagonal network grid composed of 24 neurons (6 x 4). Therefore, a cluster analysis (k-means) was performed on the conflicts data, producing 5 clusters, which is the best clusterisation based on the Davies-Bouldin index approach. The clusters and the associated conflicts are shown in Figure 7.1.5.1.

Figure 7.1.5.1. SECOA conflicts clusterised on a SOM surface.



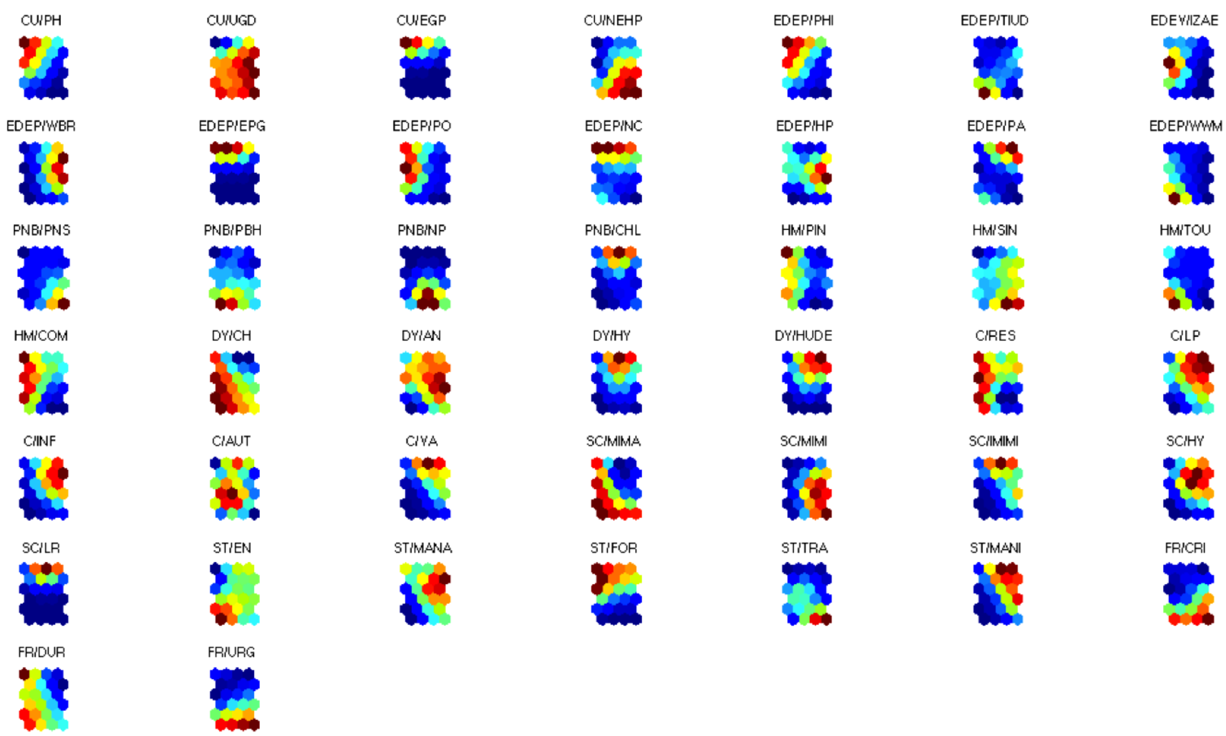
The conflicts zones described by clusters and the associated number of prototypes are shown in Table 7.1.5.3.

Table 7.1.5.3. Conflict clusters and number of prototypes.

	Zones	Prototypes
Cluster #1	Ch-Pe, Tha2, Por1, Por2, Lisb, Ealg, Tel2, Chen, Mum1, Mum2	5
Cluster #2	Rom1, Bru1, HaP1, HaP2, Oste, Bru2	3
Cluster #3	Rom2, HaP3, NhaT	2
Cluster #4	Haif, Got1, Got2	1
Cluster #5	Mal1, Mal2, Tel1, Func, Tha1	4

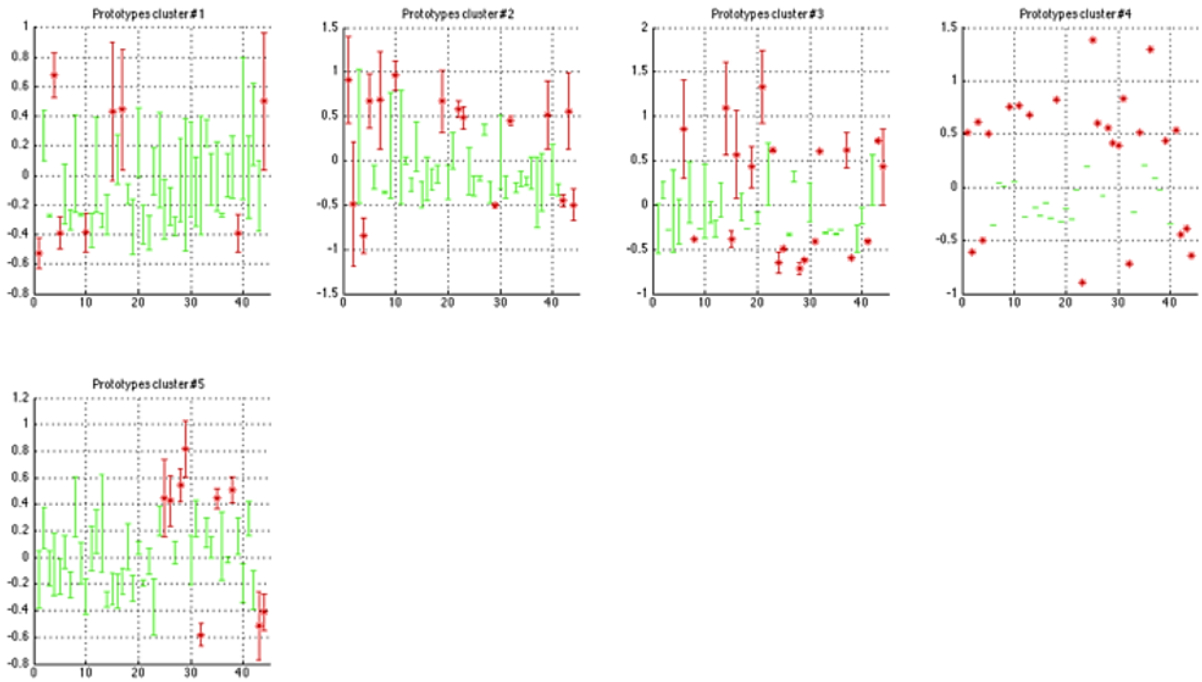
Variables distributed over the SOM surface are shown in Figure 7.1.5.2.

Figure 7.1.5.2. Conflicts variables distributed over the SOM surface (components).



The prototypes averaged within each cluster are shown in Figure 7.1.5.3.

Figure 7.1.5.3 – Conflict cluster prototypes.



The significant variables per cluster are shown in Table 7.1.5.4 and Table 7.1.5.5 (in brackets, the level of significance).

Table 7.1.5.4. Conflicts positive variables.

Cluster #1	**Cluster #2**	**Cluster #3**	**Cluster #4**	**Cluster #5**
CU/NEHP, (0.68026)	EDEP/PO, (0.9601)	HM/TOU, (1.3297)	DY/HY, (1.3841)	C/INF, (0.81717)
FR/URG, (0.50047)	CU/PH, (0.91264)	EDEP/WWM, (1.0902)	SC/LR, (1.2924)	C/LP, (0.54488)
PNB/NP, (0.44582)	EDEV/IZAE, (0.6833)	EDEP/TIUD, (0.85797)	C/VA, (0.83287)	ST/MANA, (0.50909)
PNB/PNS, (0.4331)	EDEP/PHI, (0.6718)	FR/DUR, (0.72247)	PNB/CHL, (0.82745)	DY/HY, (0.4483)
	HM/PIN, (0.66793)	ST/EN, (0.62046)	EDEP/NC, (0.76519)	SC/HY, (0.44328)
	HM/COM, (0.58069)	DY/CH, (0.61364)	EDEP/EPG, (0.75815)	DY/HUDE, (0.42552)
	FR/DUR, (0.56032)	SC/MIMA, (0.60471)	EDEP/PA, (0.67714)	
	ST/FOR, (0.51622)	PNB/PBH, (0.56979)	CU/EGP, (0.61364)	
	DY/CH, (0.48625)	FR/URG, (0.43128)	DY/HUDE, (0.5991)	
	SC/MIMA, (0.44819)	HM/PIN, (0.43044)	C/LP, (0.55478)	
			ST/MANI, (0.53329)	
			CU/PH, (0.51816)	
			SC/IMIMI, (0.51612)	

			EDEP/PHI, (0.50038)	
			ST/FOR, (0.43357)	
			C/INF, (0.41928)	
			C/AUT, (0.39757)	

Table 7.1.5.5. Conflicts negative variables.

Cluster #1	Cluster #2	Cluster #3	Cluster #4	Cluster #5
EDEP/PO, (-0.38585)	FR/CRI,(-0.4466)	PNB/PNS,(-0.38588)	FR/DUR,(-0.38611)	FR/URG,(-0.41118)
EDEP/PHI,(-0.38817)	CU/UGD,(-0.48619)	EDEP/WBR,(-0.38803)	FR/CRI,(-0.44228)	FR/DUR,(-0.51436)
ST/FOR,(-0.39281)	FR/URG,(-0.49699)	ST/MANI,(-0.40557)	CU/NEHP,(-0.49524)	SC/MIMA,(-0.58032)
CU/PH,(-0.52801)	C/INF,(-0.50354)	C/VA,(-0.40613)	CU/UGD,(-0.60498)	
	CU/NEHP,(-0.84818)	DY/HY,(-0.49292)	FR/URG,(-0.63894)	
		ST/MANA,(-0.60162)	SC/MIMA,(-0.7217)	
		C/INF,(-0.62055)	DY/CH,(-0.89385)	
		DY/AN,(-0.65036)		
		C/LP,(-0.7183)		

In order to establish a connection between SOM analysis of conflicts and territories, the results were integrated. The connection among European territories, which are described by S/E/E variables distributed over raster GIS, and the conflict classes was performed by isolating the specific area of the conflict and by counting the number of pixels belonging to each cluster. The determination of the conflict area was performed by means of a graphic mask that defined the position and the extension of the conflict on the map.

Table 7.1.5.6 in *Annex to Chapter 7* shows different conflict cases, their relations with the SOM's SOM clusterisation and the associated conflict cluster. The percentage amount of pixels contained in the conflict area for any cluster is in red. The association of Asian territories and Funchal macro-areas with conflict classes is straightforward (see Table 7.1.5.3).

2. SECOA Modelling:
A Feed-Forward Multilayer Perceptron approach

2.1. Structure of SECOA data for modelling

As already mentioned in chapter 7.1, social-economic-environmental (S/E/E) data that describe the territories where the conflicts are present had some dissimilarities. Whereas European countries involved in SECOA case studies and Israel data are geo-referenced, although with some missing data, the Asian countries (India and Viet Nam) data describe only the situation in urban macro-areas. Because of this significant difference, two different modelling approaches have been conducted, as follows:

- European and Israeli quantitative data modeling:

  Four groups of variables, amounting to 52 variables describe the territories of European countries.

  1. 19 geo-referenced variables describing the S/E/E aspect of every territory in the period 2001-2009 are: average income, net migrants, motorization rate, number households, number of residents, number of tourists, unemployment rate, employees agriculture, employees industrial, employees services, use agriculture, use airports, use industrial, use industrial-commercial, use natural, use open space, use residential mixed, flood hazard, and sea level rise.
  2. The same 19 variables were collected for the period 1991-2000. In order to account for the variability aspect of the phenomena between the two periods, the difference between the second period and the first period variables had to be calculated. Variables that indicate differences were labelled by adding a D after the variable name.
  3. 10 Integrated Coastal Zone Management binary variables (EIA: Environmental Impact Assessment, PH: Planning Hierarchy, SL: Setback Lines, MSP: Marine Spatial Planning, RC: Regulatory Commission) for each of the two periods were created.
  4. 4 Political System of Governance binary variables (Centralized, Decentralized, Unitary, Federal) were also added.

Out of four groups describing the territories of European countries, as mentioned above, the first group had many missing data, an extrapolation approach was performed by using a Self-

Organizing Map (SOM) applied as a pre-processing and homogenizing tool. A presence-absence table of European + Israeli variables is shown in Table 7.2.1.1.

Table 7.2.1.1. Presence of variables in European and Israeli urban areas.

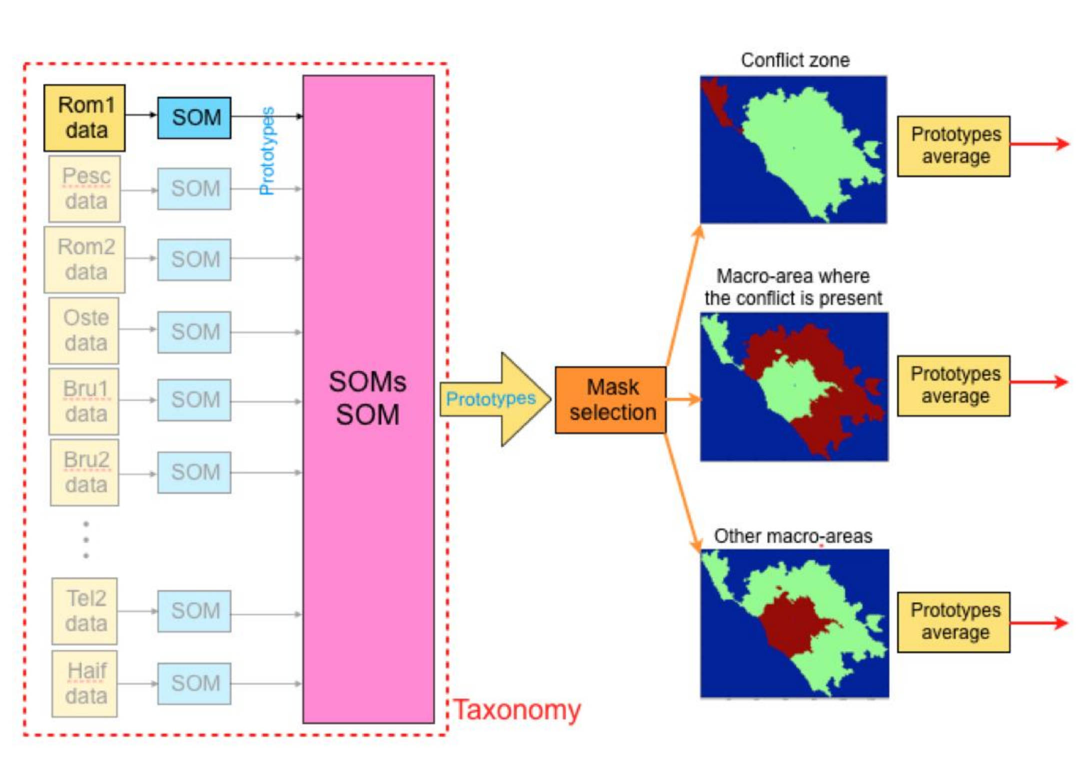
	Rom IT	**Chie IT**	**Lisb PT**	**EAlg PT**	**Func PT**	**Port UK**	**Tham UK**	**Brug BE**	**Oste BE**	**Got SE**	**Mal SE**	**TelA IL**	**Haif IL**
average income	1	1	1	1	1	1	1	1	1	1	1	1	1
net migrants	1	1	1	1	1	1	1	1	1	1	1	0	0
motorization rate	1	1	1	1	1	1	1	1	1	1	1	0	0
number households	1	1	1	1	1	1	1	1	1	1	1	1	1
number of residents	1	1	1	1	1	1	1	1	1	1	1	1	1
number of tourists	1	1	1	1	1	0	0	1	1	1	1	0	0
unemployment rate	1	1	1	1	1	1	1	1	1	1	1	1	1
employees agriculture	1	1	1	1	1	1	1	1	1	1	1	1	1
employees industrial	1	1	1	1	1	1	1	1	1	1	1	1	1
employees services	1	1	1	1	1	1	1	1	1	1	1	1	1
use agriculture	1	1	1	1	1	1	1	1	1	1	1	1	1
use airports	1	1	1	1	0	1	1	0	0	1	1	0	0
use industrial	1	0	1	1	1	1	1	0	0	1	1	0	0
use industrial commercial	1	1	1	1	1	1	1	1	1	1	1	0	0
use natural	1	1	1	1	1	1	1	1	1	1	1	0	0
use open-space	1	0	1	1	1	1	1	1	1	1	1	1	1
use residential mixed	1	1	1	1	1	1	1	1	1	1	1	1	1
flood hazard	1	1	1	1	0	1	0	1	1	0	0	1	1
sea level rise	1	1	0	0	0	0	0	1	1	1	1	1	1

The SOM procedure was divided in two steps:

1. Each territory described by 52-variables, each one related to any single GIS pixel, was analysed by a SOM machine, and prototypes related to any pixel were extracted.
2. A general SOMs SOM analysis was performed by feeding the network with the previously extracted prototypes, weighted according to their relative abundance.

The SOM procedures hereby described are identical to the ones used in the Taxonomy described in chapter 7.1. As it will be shown in the next section, the SECOA model will be built by dividing each territory into three parts, depending on the conflict under analysis: the conflict area; the macro-area of the urban territory, which contains the conflict case; and the remaining macro-areas. Every variable will take on a single value for each one of these sub-territories in order to describe the S/E/E+ICZM+Gov situation in that part of the area. In order to accomplish this task, a procedure was built to associate any GIS pixel with its corresponding prototype from SOMs SOM. Finally, the values attributed to any sub-territory were determined by the average of prototypes related to any pixel contained in the specific sub-territory area. The final diagram of the entire procedure is described in Figure 7.2.1.1 (Civitavecchia case study).

Figure 7.2.1.1. Diagram of data organization in European + Israeli countries case studies.



- All SECOA countries qualitative data modelling

The main difference among SECOA territories is in geographic scale description. While European and Israeli data represent the territories at 1x1 Km pixel scale (GIS), Indian and Vietnam data describe the S/E/E situation at a macro-area level. Even if the model should be applied at a macro-area level, the Asian data described the territorial aspects of S/E/E by using different variables lacking S/E/E sections. Indian and Vietnam variables included:

1. India: Employees in agriculture, employees in industrial, employees in services, net migration, average tourists, rainfall, slum population, temporary dwellings, vehicles, population, immigrants, international migrants, national immigrants, bus arrivals, household size, total area, agriculture area, built up area, natural area, open space area, road net length, sustainability, and sea level rise.
2. Vietnam: Population, international immigrants, national immigrants, immigrants from ring, immigrants from core, emigrants, vehicles, land area, economic area, housing area, dwellings, temporary dwellings, households, total area, agriculture area, industrial-commercial area, mixed area, residential area, rainfall, sustainability, tourists, and inundation.

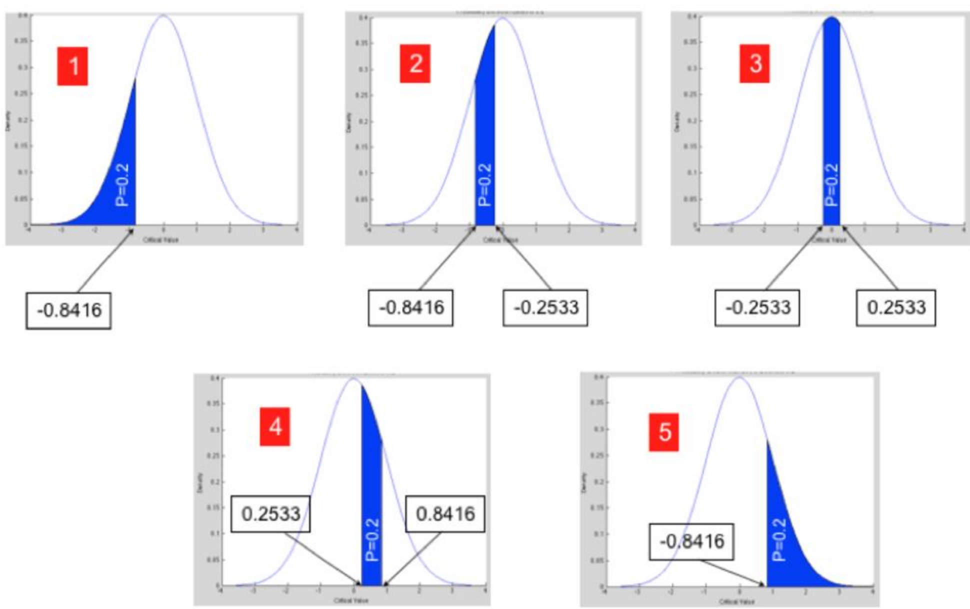
A large part of variables used in the description of European + Israeli data was missing, while several variables were specifically related to the Asian countries.
In order to make the variable sets homogeneous, we decided to describe every single variable in terms of qualitative parameters. Then, the two sets were transformed into qualitative sets, as follows:

1. *Europe and Israel*: after identifying the conflict area, the macro areas involved in conflict case and the remaining macro areas for every case study, the qualitative value defining each variable was determined using the following steps:
   a) the mean and the variance of all the prototype values over all territories referring to a specific variable were calculated;
   b) a normal Gaussian value Z was calculated by every averaged prototype associated with any single territory:

$$Z_{\overline{P}} = \frac{\overline{P} - \mu_P}{\sigma_P} \tag{7.2.1.1}$$

c) a five-level rank was then identified by evaluating the equiprobable interval that contains the variable *Z* based on the Gaussian normal distribution (Figure 7.2.1.2).

Figure 7.2.1.2. Transformation of averaged prototypes in 5-level ranking qualitative data.



2. *India and Vietnam*: as mentioned previously, Asian data were only partially described by the same variables as were European data, and the variables considered only the urban macro areas. Because of this last aspect, it was not possible to convert useful variables into qualitative values because of the lack of referring data distribution (as in Figure 7.2.1.2). To overcome these difficulties and to obtain a full set of comparable variables, some experts from India and Vietnam were recruited to fill the rank table for all S/E/E variables in the two periods.

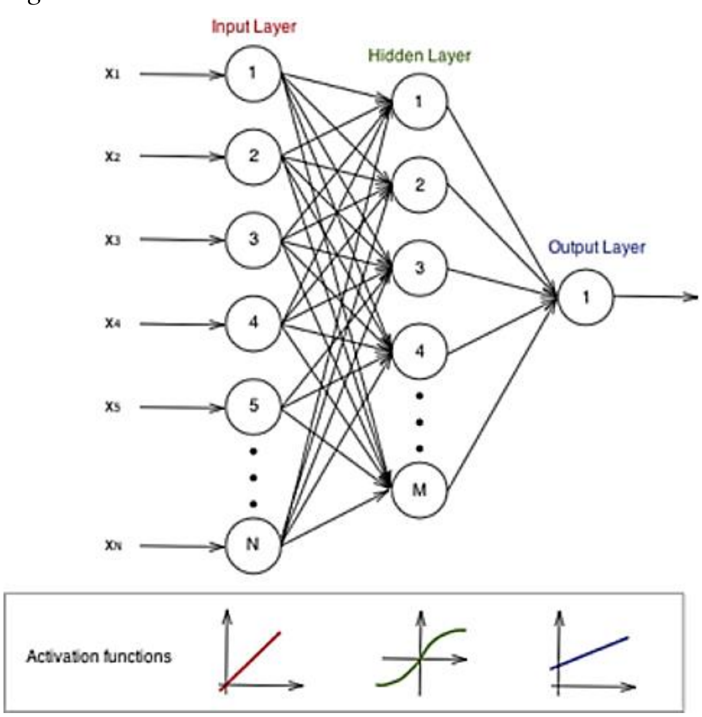
The S/E/E variables were the same as shown in quantitative section, except the "use airport" parameter that was considered useless in case of sub-territories analysis. Therefore, the final amount of qualitative variables was 50 (18 second period S/E/E, 18 variation S/E/E, 10 ICZM, 4 Governance).

The last obstacle in defining a complete set of qualitative variables was the lack of values presents in the European and Israeli data (Table 7.2.1.1). Therefore, even in the case of qualitative data, a SOM approach was applied in order to fulfil the gaps and obtain a full set of values for the modelling phase. Again, every sub-territory was associated with a prototype as a complete data pattern needed for the modelling phase.

2.2. Modelling SECOA data: A Feed-Forward neural network algorithm

As stated previously, the SECOA data-modelling phase was achieved by means of feed-forward neural networks. Pre-processed quantitative and qualitative data (see section 7.2.1) were used to train networks until the overall error was smaller than a given reference value. In order to reduce the network complexity, it was decided to implement a single neural network for each one of the SECOA conflict parameters. Therefore, a single model (quantitative or qualitative) comprised 44 neural networks trained on the associated S/E/E+ICZM+Gov pre-processed data.
After some preliminary attempts, the best neural network structure was determined as a network composed of a single hidden layer containing 30 symmetric sigmoid neurons and an output linear neuron (Figure 7.2.2.1),

Figure 7.2.2.1. A SECOA Neural Network Structure.



where N is the number of input variables (52 in the quantitative case, 50 in the qualitative case) and M is the number of hidden neurons (30).

The final neural equation is described by the following relation:

$$\boldsymbol{F}_{NN}^{(k)} = \sum_{h=1}^{M} \boldsymbol{w}_h^{(k)} \boldsymbol{f}\left[\sum_{j=1}^{N} \boldsymbol{w}_{hj}^{(k)} \boldsymbol{x}_j + \boldsymbol{b}_h^{(k)}\right] + \boldsymbol{b}^{(k)} \tag{7.2.2.1}$$

where k represents the specific network associated with the *k*-th conflict parameter, w_h and w_{hj} are respectively the synaptic weights connecting the hidden layer to the output layer and the weights connecting the input layer to the hidden layer. f(x) represents the symmetric sigmoid function:

$$\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{tanh}(\boldsymbol{\beta x}) \tag{7.2.2.2}$$

It can be noted that the neural network equation (eq. 7.2.2.1) is continuous and differentiable. The Jacobian matrix can be evaluated as following:

$$\frac{\partial \boldsymbol{F}_{NN}^{(k)}}{\partial \boldsymbol{x}_j} = \sum_{h=1}^{M} \boldsymbol{w}_h^{(k)} \boldsymbol{w}_{hj}^{(k)} \boldsymbol{f}'\left[\sum_{l=1}^{N} \boldsymbol{w}_{hl}^{(k)} \boldsymbol{x}_l + \boldsymbol{b}_h^{(k)}\right] \tag{7.2.2.3}$$

where $f'(x)$ is the derivative of the sigmoid function:

$$\boldsymbol{f}'(\boldsymbol{x}) = \frac{\boldsymbol{df}}{\boldsymbol{dx}} = \boldsymbol{\beta}\left[\mathbf{1} - \boldsymbol{tanh}^2(\boldsymbol{\beta x})\right] = \boldsymbol{\beta}\left[\mathbf{1} - \boldsymbol{f}^2(\boldsymbol{\beta x})\right] \tag{7.2.2.4}$$

The Jacobian matrix, fully analytically calculable, is a powerful and useful tool for evaluating the model trends, which will be discussed in the next section.

The learning phase was based on the Levemberg-Marquardt training algorithm, which is a numerical method for minimizing an error function, and it can be considered as a middle-point between the Gauss-Newton Algorithm and the Gradient Descent. The Levemberg-Marquardt algorithm can be represented by the following meta-language script (for the generic *k*-th conflict parameter):

1. For p=1 to P (number of patterns: all sub-territories data)
2. Evaluating the network output y (equation 1)
3. Evaluating and accumulating the squared weighted error between y and the target t
4. If the sum of partial squared errors is less than a reference value, go to 7.
5. Considering the sum of partial squared errors, the neural network weights and biases modification are applied according to Levemberg-Marquardt algorithm
6. Go to 1.
7. Exit

In the SECOA modelling, the evaluation of the *p*-th error is given by the difference between the target (*k*-th conflict parameter) and the current network output multiplied by a modifier factor w_p in order to take into account the relevance of the sub-territory under analysis. This choice was adopted to take into account different relevance of each sub-territory in the description of the conflict. In fact, the generic urban area case was divided into three parts: the conflict area, the macro-areas containing the conflict area and the remaining macro-areas. In order to give the network the ability to consider different influences of such sub-territories, the effect of the error associated with the conflict area pattern influenced the full network parameter modification (100%) while in case of the remaining two sub-territories, such influence was reduced to 20% and 5%, respectively. This approach was translated in terms of algorithmic language by adopting the modifier factor w_p with values 1, 0.2 and 0.05, respectively. In this case, the influence of territories far from the conflict area are considered in the general network training phase, but their influence cannot perturb excessively the suitable learning task of the conflict area situation.

2.3. Experimental results

A class of feed-forward neural networks was applied to evaluate the model of the relations connecting the S/E/E+ICZM+Gov sub-territories variables to the conflict parameters. The final model has to be intended as a class of N-dimensional hyper-surface (N=52 for quantitative data and N=50 for qualitative data), each of them interpolating the SECOA patterns for every single conflict parameter. Therefore, the geometrical characteristics of the model represent the global behaviour of the relations territory-conflict. In this context, a well-described model provides indications about territorial situations and conflict typologies not included in the data set used to model the system (generalization).

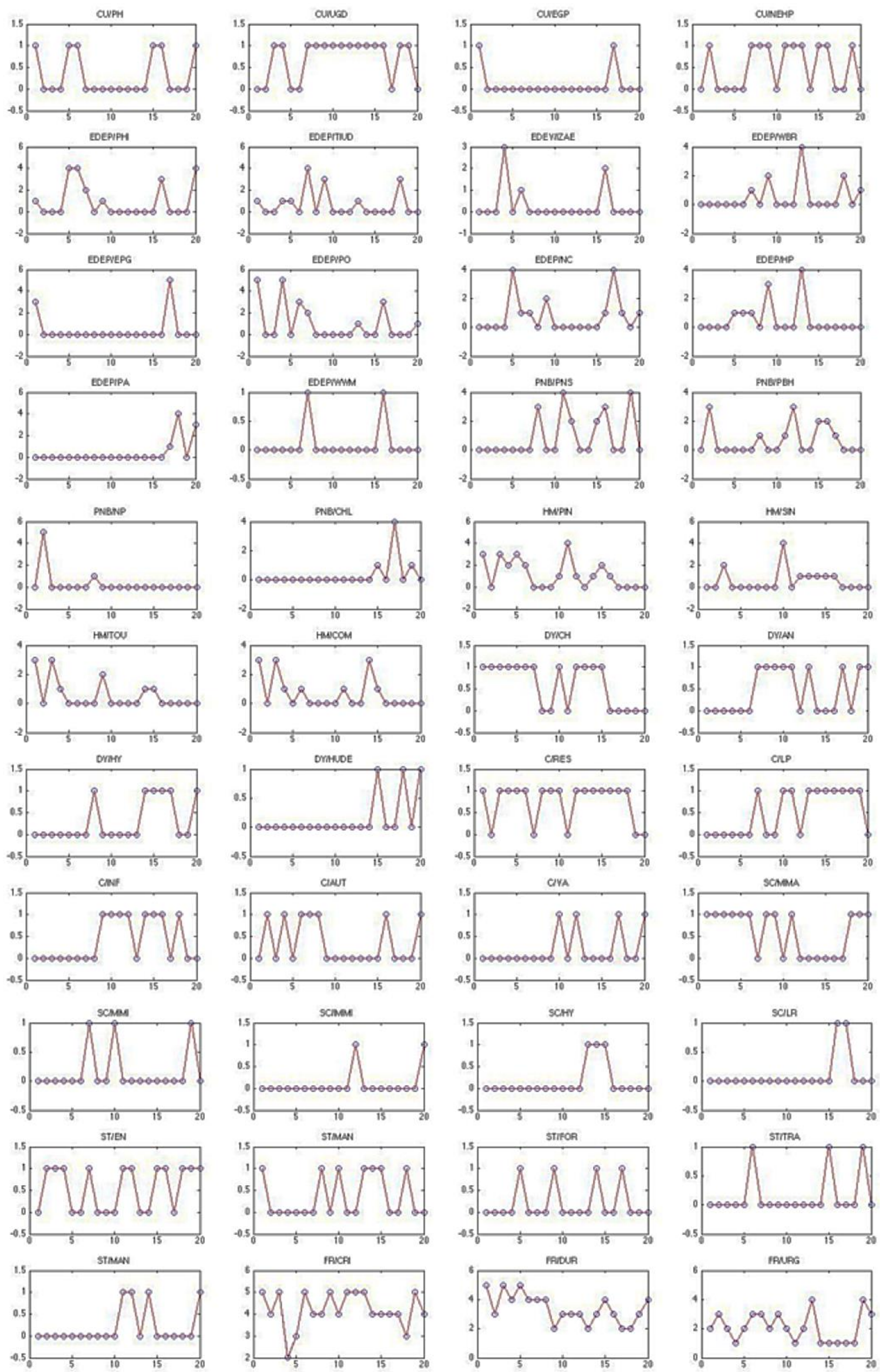
As mentioned in the previous section, the feed-forward neural networks were applied to both geo-referenced European and Israeli quantitative data and to all SECOA qualitative data. The results are described as follows:

1. European and Israeli quantitative data modeling

   Overall, 44 feed-forward neural networks with 52 input units, 30 hidden sigmoid units and 1 linear output unit were implemented. Each neural network was trained on a single conflict parameter. The training phase stopped when the mean squared weighted error dropped below 10^{-6} or if the learning gradient decreased under 10^{-6}. This last constraint was adopted to avoid the training when the network reached a local minimum with error greater than 10^{-6}. Every training phase was repeated 10 times and the best network in terms of learning error was chosen as referring model.

The 44 network outputs (blue points) and the associated target (red points) are represented in Figure 7.2.3.1.

Figure 7.2.3.1. Quantitative SECOA data - 44 network outputs (blue points) and the associated target (red points).



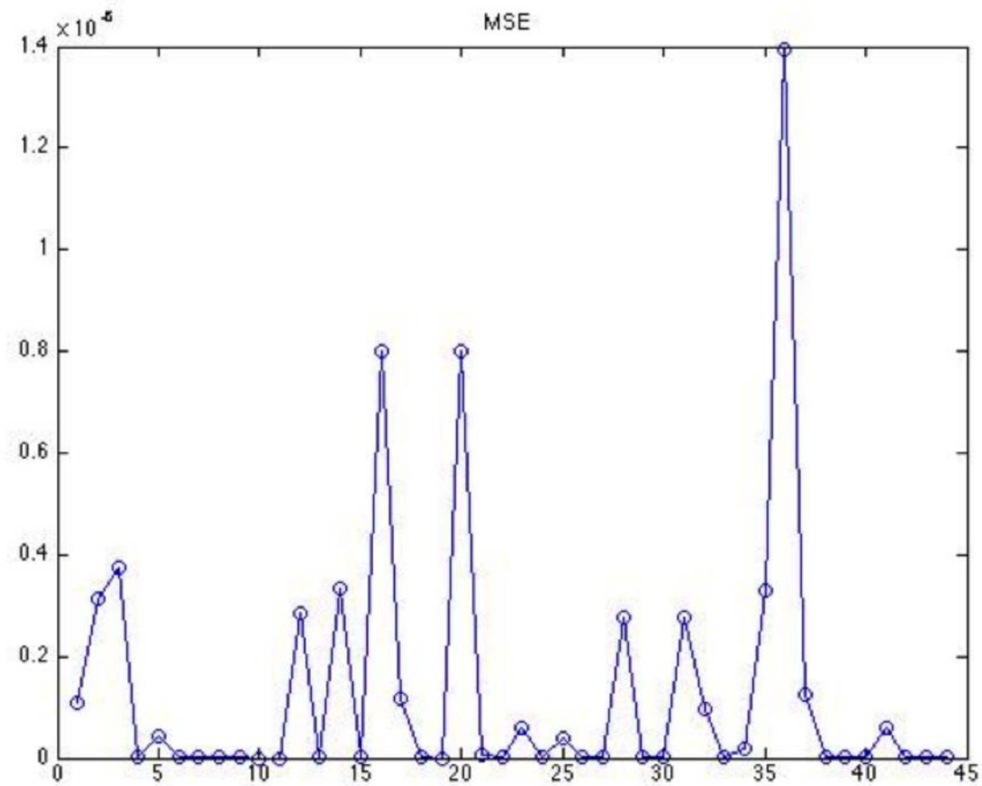
The 20 territories on the horizontal axis are listed in Table 7.2.3.1.

Table 7.2.3.1. European and Israeli territories.

1	2	3	4
Rome	Chieti-Pescara	Rome	Ostend
Civitavecchia	The "Costa Teatina" National Park	Ostia water-use and management	Ostend Airport
5	6	7	8
Zeebrugge	Zeebrugge	Lisbon	Algarve
Schipdonk canal	Zeebrugge harbour	Trafaria and Costa da Caprarica	Barrier Islands [Ria Formosa Natural Park]
9	10	11	12
Funchal	Thames Gateway	Thames Gateway	Portsmouth
Funchal Bay [Madeira Island]	Barking Riverside	Lower Thames Crossing	Langstone Harbour / Farlington Marshes
13	14	15	16
Portsmouth	Malmö	Malmö	Gothenburg
Tipner Regeneration	Managing Urban Sprawl	Falsterbo-Peninsula: Vellinge municipality	Torsviken
17	18	19	20
Gothenburg	Tel Aviv	Tel Aviv	Haifa
Kungsbacka	Palmachim beach	Netanya sandstone cliffs	Haifa Port

The mean squared error in evaluating the 44 conflict parameters is represented in Figure 7.2.3.2.

Figure 7.2.3.2. Mean squared error for the 44 conflict parameters in quantitative model (Europe + Israel).

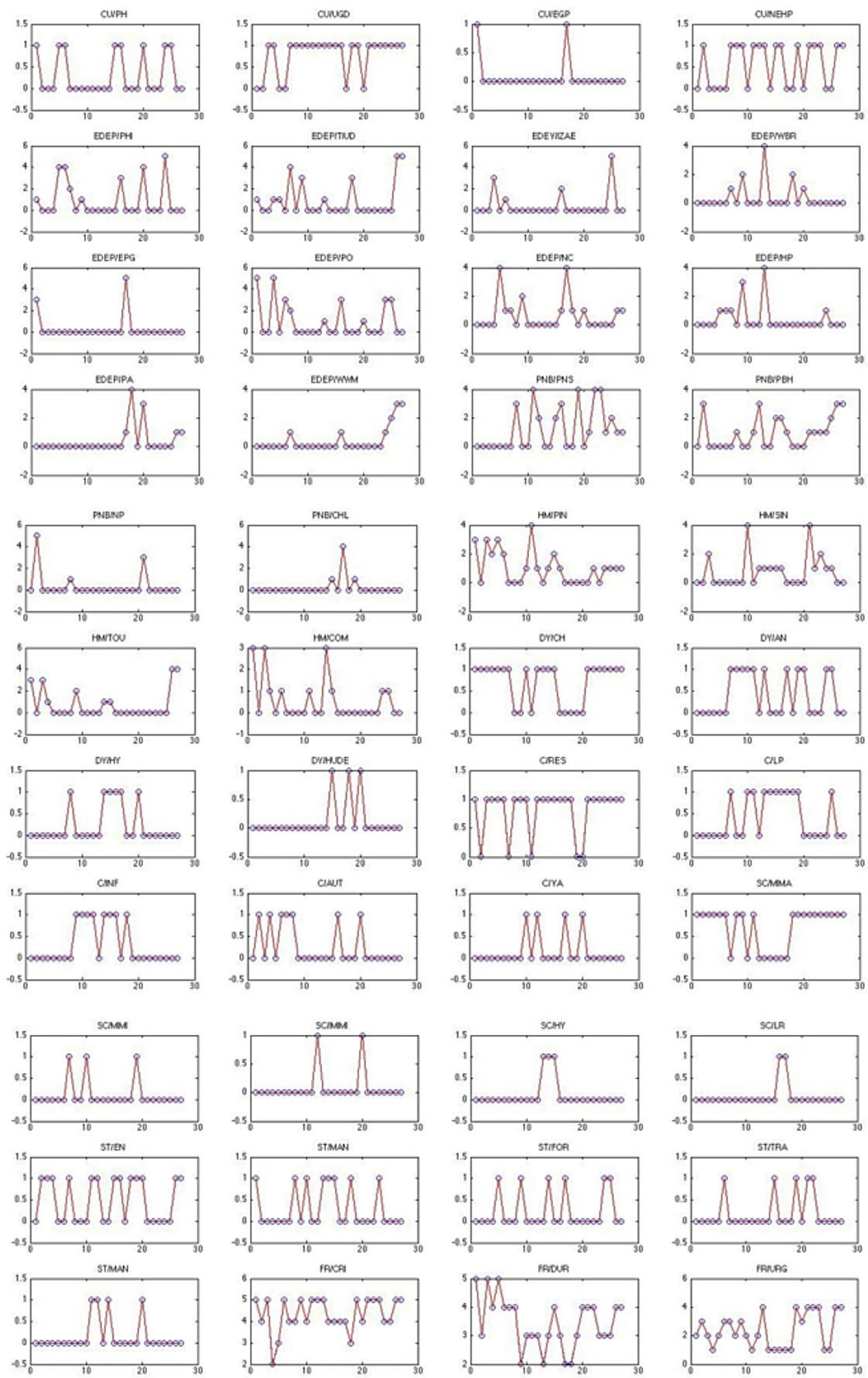


2. All SECOA countries qualitative data modelling

   Overall, 44 feed-forward neural networks with 50 input units, 30 hidden sigmoid units and 1 linear output unit were implemented. Each neural network was trained on a single conflict parameter. The training phase stopped when the mean squared weighted error dropped below 10^{-6} or if the learning gradient decreased under 10^{-6}. This last constraint was adopted to avoid the training when the network reached a local minimum with error greater than 10^{-6}. Every training phase was repeated 10 times and the best network in terms of learning error was chosen as referring model.

   The 44 network outputs (blue points) and the associated target (red points) are represented in Figure 7.2.3.3.

Figure 7.2.3.3. Qualitative SECOA data - 44 network outputs (blue points) and the associated target (red points).



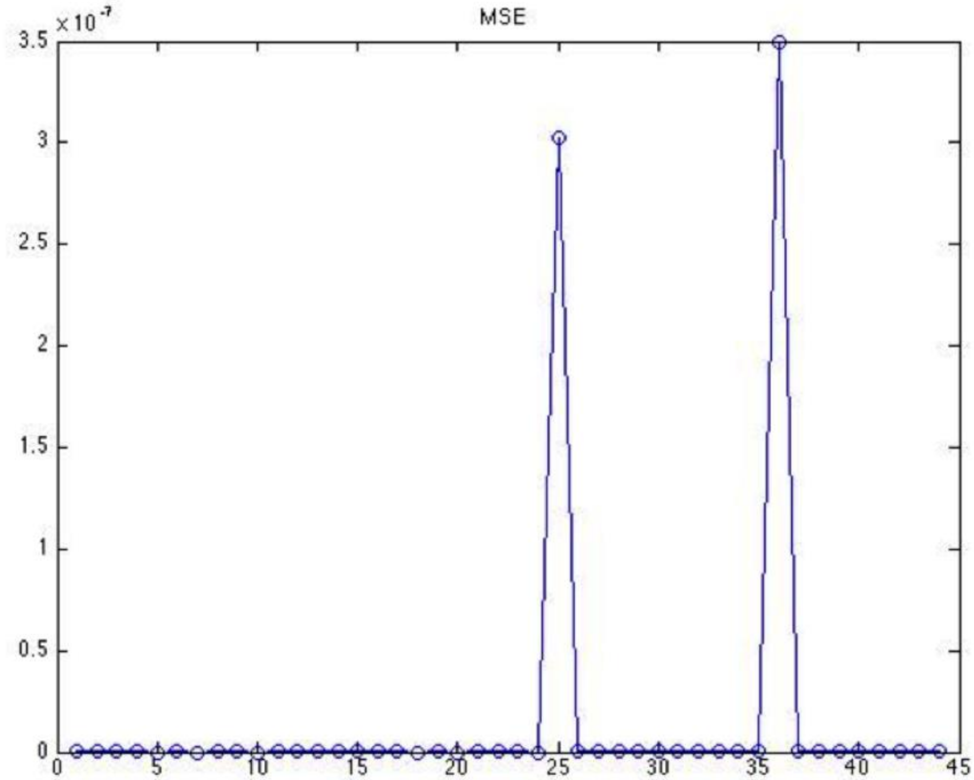
The 27 territories on the horizontal axis are listed in Table 7.2.3.2.

Table 7.2.3.2. SECOA territories.

1	2	3	4	5	6	7
Rome	Chieti-Pescara	Rome	Ostend	Zeebrugge	Zeebrugge	Lisbon
Civitavecchia	The "Costa Teatina" National Park	Ostia water-use and management	Ostend Airport	Schipdonk canal	Zeebrugge harbour	Trafaria and Costa da Caprarica
8	9	10	11	12	13	14
Algarve	Funchal	Thames Gateway	Thames Gateway	Portsmouth	Portsmouth	Malmö
Barrier Islands [Ria Formosa Natural Park]	Funchal Bay [Madeira Island]	Barking Riverside	Lower Thames Crossing	Langstone Harbour / Farlington Marshes	Tipner Regeneration	Managing Urban Sprawl
15	16	17	18	19	20	21
Malmö	Gothenburg	Gothenburg	Tel Aviv	Tel Aviv	Haifa	Mumbai
Falsterbo-Peninsula: Vellinge municipality	Torsviken	Kungsbacka	Palmachim beach	Netanya sandstone cliffs	Haifa Port	Sanjay Gandhi National Park
22	23	24	25	26	27	
Mumbai	Chennai	Haiphong	Haiphong	Haiphong	Nha Trang	
Mangrove forest	Pallikaranai Marshland	Haiphong port	Industrial Zone	Cat Ba	Nha Trang	

The mean squared error in evaluating the 44 conflict parameters is represented in Figure 7.2.3.4.

Figure 7.2.3.4. Mean squared error for the 44 conflict parameters in qualitative model (all SECOA countries).



2.4. Test analysis

The SECOA neural network model was also analysed in terms of property generalization. To evaluate the ability of the model to predict a conflict case when an untrained territorial pattern is presented at the network input, four new classes of neural networks were trained both for quantitative and qualitative data. In each of these new networks, a particular urban territory and the associated conflict case was excluded in the training phase and after a complete training, the network was asked to evaluate the conflict parameters of the untrained data. The comparison between the predicted values and the known ones resulted in an index of the neural model ability to describe the global aspects of SECOA conflict system.

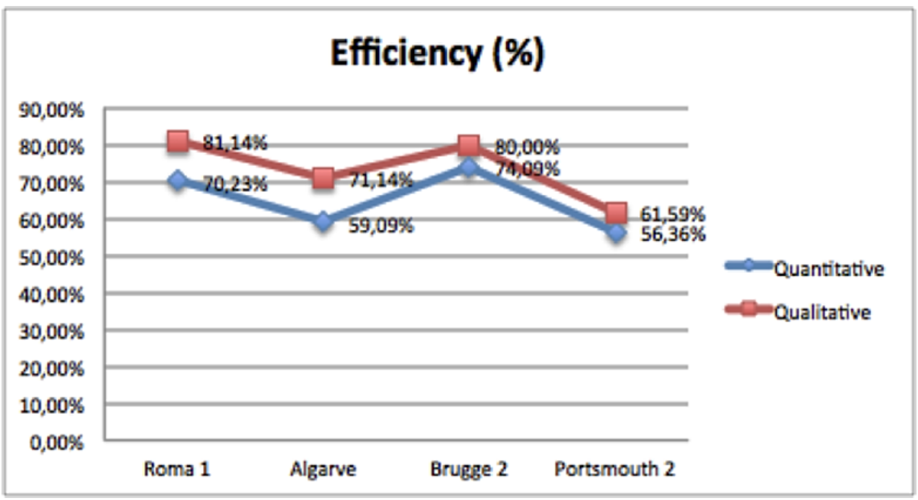
The trainings were performed using the same experimental conditions as in the previous section: 50 (or 52) by 30 by 1 neurons network, maximum mean error of 10^{-6}, 10 repetitions of the same conflict parameter and 19 (or 26) conflict case patterns to learn. The Civitavecchia case (Rome1), Ria Formosa case (EAlg), Zeebrugge Harbor case (Bru 2) and Tipner case (Por 2) were excluded by the training phase both for quantitative and qualitative case. After the training was performed, a parameter called efficiency was calculated using the formula:

$$\text{Efficiency} = 1 - \frac{1}{44}\sum_{k=1}^{44}\left(\frac{t^{(k)} - y^{(k)}}{\max\{t^{(k)}\}}\right)^2 \tag{7.2.4.1}$$

where the sum runs over the 44 conflict parameters, $t^{(k)}$ is the k-th conflict parameter target, $y^{(k)}$ is the k-th neural network outcome, and $\max\{t^{(k)}\}$ is a normalization factor, which is 1 for binary and 5 for multilevel conflict parameters.

The comparison between efficiencies of quantitative and qualitative cases is presented in Figure 7.2.4.1.

Figure 7.2.4.1. Comparison between quantitative and qualitative test analysis.



The same analysis was performed on every single conflict class (Table 7.1.5.2) and is shown in Figure 7.2.4.2a (Rome 1), 7.2.4.2b (Algarve), 7.2.4.2c (Zeebrugge 2) and 7.2.4.2d (Portsmouth 2).

Figure 7.2.4.2a. Efficiency of network model when tested on Rom 1 conflict case (Civitavecchia).

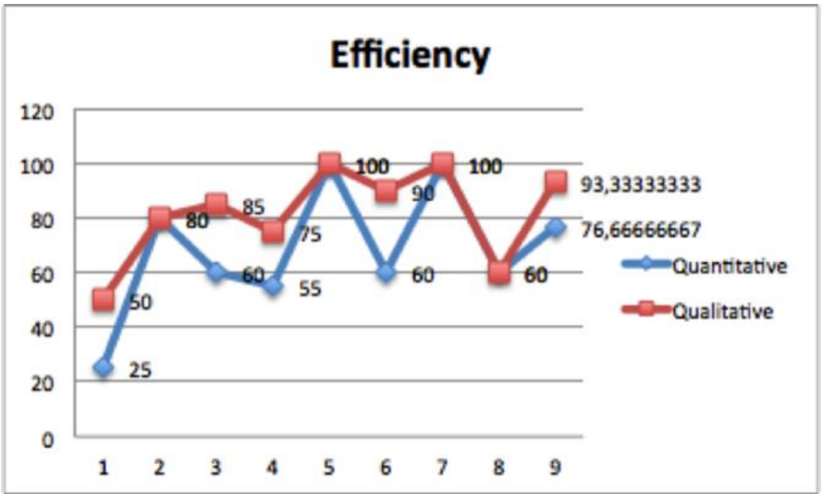


Figure 7.2.4.2b. Efficiency of network model when tested on Algarve conflict case (Ria Formosa).

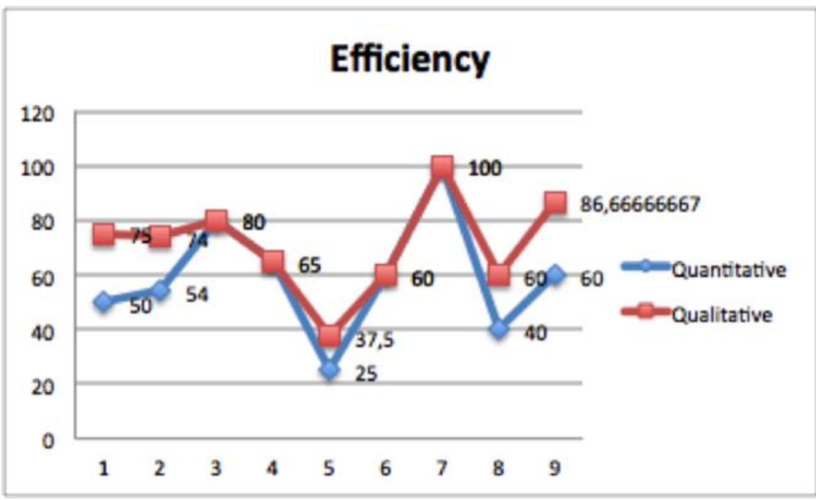


Figure 7.2.4.2c. Efficiency of network model when tested on Zeebrugge 2 conflict case (Zeerbrugge Harbor).

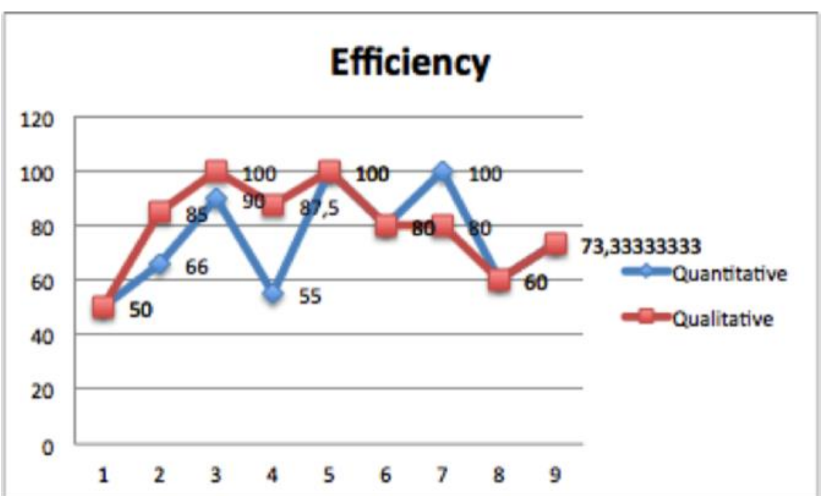
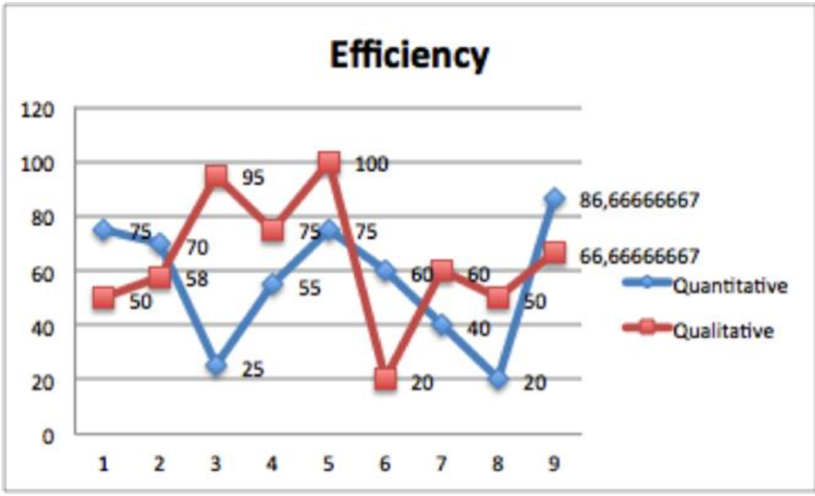


Figure 7.2.4.2d. Efficiency of network model when tested on Portsmouth 2 conflict case (Tipner).



The horizontal axis is divided into 9 categories listed in Table 7.2.3.3.

Table 7.2.3.3. Conflict classes used in efficiency analysis.

1	2	3	4	5
Category of uses	Economic Development Vs. Environmental protection	Preservation of natural sites and biodiversity	Human mobility and contrast for use of resources	Dynamics/manifestation over time
6	7	8	9	
Underlying cause/substance	Scale	Stage	Final Ranking	

Figures 7.2.4.1 and 7.2.4.2a to 7.2.4.2d highlight the best performance of the qualitative approach to analysing SECOA data with respect to the quantitative case. The qualitative data approach comprises a larger number of representative cases. The modelling system can define a better geometrical interpolating hyper-surface of which high-dimensional form is able to represent the entire SECOA system in a more accurate and precise way.

3. Analysis of SECOA modelling: Trend Analysis

One of the most relevant uses that can be done with a descriptive model is the prediction task. In SEOCA environment, a real temporal prediction cannot be addressed, since data were gathered in only two temporal periods and the instant of collection was remarkably dissimilar across cases. Usually a temporal prediction can be done when a large time course is available in data description, offering the possibility to apply several techniques as embedding space reconstruction, definition of minimal dimensionality and so on. In SECOA modelling, the only available prediction technique is the Trend Analysis. As discussed in section 7.2.3 on "Modelling SECOA data", once the class of neural networks converges in terms of training, every component of model equation is identified, including all weights and biases (Eq. 7.2.3.1). This means that every point in the SECOA space can be related to a particular conflict aspect. Since the number of training points is rather limited, an excessive lack of local information can affect the evaluation of new territorial conflict aspects that are far from the trained ones in terms of distance in the SECOA S/E/E variables space. On the other hand, the trained points and the high-dimensional geometry that the neural network was able to rebuild around such points are well defined in terms of mutual relations. Accordingly, we decided to emphasise the role of the slopes around each SECOA data point that may be related to the local conflict trend. This operation is easily performed by Eq. 7.2.3.3 that allows calculating the Jacobian matrix (or local partial derivatives). Therefore, it was possible to build a table for every conflict case. The trend (or tendency) of each conflict parameter related to all SECOA variables, and it was described with positive, negative or negligible slopes. A table like that can be read by answering the question: in the SECOA context, how does a particular conflict parameter change if a specific variable increases? In this framework, a complete overview of the SECOA trends can be offered to an expert policy-maker in order to suggest the suitable technique to approach the conflict solving. In our analysis, the Trends Table was built in the case of Qualitative Data since the modelling phase appeared to be more consistent with the testing phase.

Trends were calculated for all 44 conflict parameters in relation to the 52 S/E/E input variables, and for every case study.

Table 7.3.1. Initial part of the trends for Rome Civitavecchia case study.

	CU/PH	CU/UGD	CU/EGP	CU/NEHP	EDEP/PHI	EDEP/TIUD	EDEV/IZAE	EDEP/WBR	EDEP/EPG	EDEP/PO
	Ports and harbours related uses	Urban Growth and Development in terms of specific urban functions including tourism	Energy generation / production	Natural environment and habitat including national parks and protected areas	Port / harbour restructuring, expansion & infrastructure	Tourism lead infrastructure & urban development	Industrial zones and airports expansion	Waterfront & brownfields regeneration	Energy / power generation	Pollution [air, soil, water] associated
average income 2	0,03	-0,03	-0,10	0,30	-0,48	0,12	0,01	-0,52	-0,23	-0,05
net migrants 2	0,06	0,11	0,14	-0,25	0,39	0,42	0,20	-0,28	0,13	1,75
motorization rate 2	0,10	-0,13	0,35	-0,16	0,53	0,61	-0,31	0,37	1,04	1,34
number households 2	-0,03	-0,10	0,04	-0,25	-0,17	-0,50	-0,01	-0,11	-0,15	0,33
number of residents 2	-0,09	0,20	-0,04	0,11	0,47	-0,13	0,22	-0,20	-1,10	0,16
number of tourists 2	-0,38	0,25	-0,35	0,38	-0,65	0,36	-0,09	0,17	-1,72	-1,63
unemployment rate 2	0,11	-0,08	0,11	0,00	0,05	0,85	-0,10	0,10	0,50	0,34
employees agriculture 2	-0,10	0,03	-0,04	0,03	-0,76	-0,32	0,25	0,41	0,15	-1,16
employees industrial 2	0,02	-0,10	-0,09	0,13	0,09	-0,05	0,58	-0,28	-1,00	0,05
employees services 2	-0,06	0,12	-0,04	0,29	-0,08	-0,47	0,06	0,22	-0,03	-0,64
use agriculture 2	-0,01	-0,01	0,02	-0,09	0,22	-0,11	-0,03	0,08	0,10	-0,56
use industrial 2	0,00	-0,06	-0,10	0,07	0,44	-0,24	0,21	0,16	0,29	-0,32
use industrial commercial 2	0,07	-0,03	0,13	0,04	-0,13	0,30	-0,33	-0,07	0,21	-0,30
use natural 2	0,04	-0,04	0,05	-0,11	-0,31	0,03	-0,15	0,01	0,10	-0,14
use openspace 2	0,16	-0,04	0,03	-0,02	0,20	-0,20	-0,53	-0,10	0,58	0,15
use residential mixed 2	0,07	-0,13	0,03	-0,15	0,15	-0,29	0,27	0,10	0,06	0,00
floodhazard 2	-0,23	-0,22	0,15	-0,09	0,48	-0,17	0,22	-0,32	-0,01	-0,23
slr 2	0,02	-0,13	0,04	-0,02	-0,38	0,28	-0,36	0,16	0,49	-0,49
average income D	0,01	-0,27	0,42	-0,27	0,30	-0,31	-0,38	-0,22	0,97	0,04

In Table 7.3.1, an example of Trends is shown for the Rome - Civitavecchia case. Slopes are represented numerically and by a colour coding:

- high positive slope – dark green
- moderate positive slope – light green
- zero slope – white
- moderate negative slope – light red
- high negative slope – dark red

Trends represent a crucial tool in defining the possible evolution of a conflict parameter when some S/E/E variables change because of a policy applied to the specific area. Although the neural model is capable of providing a global description of the SECOA system, the sampled case studies involved in SECOA do not provide a uniform coverage of the multidimensional S/E/E variables space. The evaluation of any point describing a potential S/E/E territorial situation that is far from the sampled points may be affected by an amount of uncertainty due to a lack of sufficient information. Conversely, the evaluation of the trends gives sufficiently reliable information about conflict changes resulting from slight modifications in the S/E/E variables. In particular, a percentage of conflict parameter change can be described as in equation 7.3.1:

$$\frac{\Delta \boldsymbol{p_k}}{\boldsymbol{p_k}} \approx \left[\frac{\boldsymbol{x_j}}{\boldsymbol{p_k}}\right] \frac{\partial \boldsymbol{F_{NN}^{(k)}}}{\partial \boldsymbol{x_j}} \frac{\Delta \boldsymbol{x_j}}{\boldsymbol{x_j}} \quad (7.3.1)$$

where p_k is the *k*-th conflict parameter, x_j is the *j*-th S/E/E variable, and $\frac{\partial F_{NN}^{(k)}}{\partial x_j}$ is the Jacobian of the *k*-th neural model calculated over the *j*-th variable.

4. Conclusion

Artificial neural networks allowed extracting useful and significant information by territorial S/E/E and case study conflicts of SECOA system. The main steps defined in the SECOA analysis process can be summarized in the following points:

- *Taxonomy*. The special behaviour offered by Self-Organized Maps (SOM), in particular the Kohonen map, allowed to solve the lack of information data of some case study by filling the *holes* with a substitution inferred by the whole SECOA scenario. This particular aspect allows dealing with a problem usually considered hard for a classical statistical system. Furthermore, the SOM ability to reduce the data dimensionality was a useful tool to distribute the several case studies on a two-dimensional surface allowing the visual definition of conceptual distance among the different territories and conflicts. Cluster analysis for the taxonomy of territories and conflicts also taken advantage of SOM distribution property in order to define a suitable clustering among the high-dimensional SECOA data.
- *Modelling*. The relation between territories and conflict parameters has been addressed through the use of multi-layered feed-forward neural networks. The final model was able to correctly describe the connection between any specific territory of SECOA database and the conflict aspects in that particular area. Moreover, one of the abilities of a neural model is to give a general description of the whole space of variables by defining a nonlinear hyper-surface passing through all the data points used in the training phase. The regions that are not represented by some of the SECOA data are anyhow described by the nonlinear neural function, giving a hypothesis coherent with the whole data set of the general law underlying the SECOA manifestations.
- *Trends analysis*. The evaluation of the linearized tangent space around any SECOA case study allowed the definition of an evolutionary rule for the forecasting of SECOA conflicts. The trends extracted by the nonlinear neural function give a description of the modification of any conflict parameter due to the variation of every single variable of the SECOA database. In this framework, it is possible to evaluate the future evolution of the conflicts basing on the variables describing territories and the other supporting variables as ICZM, collected on two past periods of time. The application of different evolution of S/E/E variables may lead to different scenarios depending on the policies governing the variable evolutions.

Annex to Chapter 7.

Figure 7.1.2.1. SOM clusters: Rome.

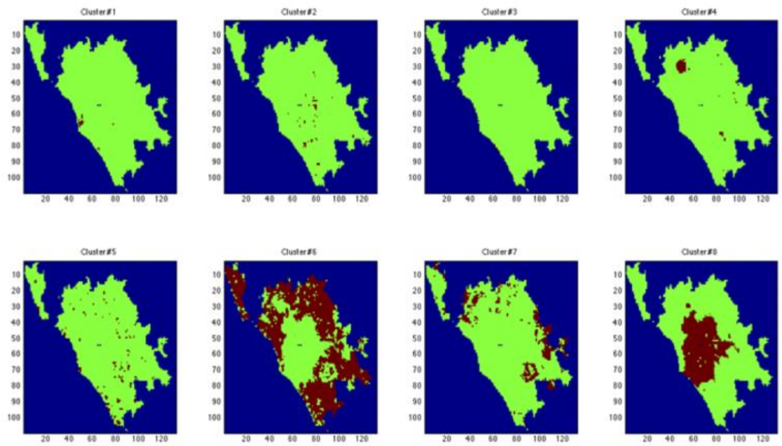


Figure 7.1.2.2. SOM clusters: Chieti-Pescara.

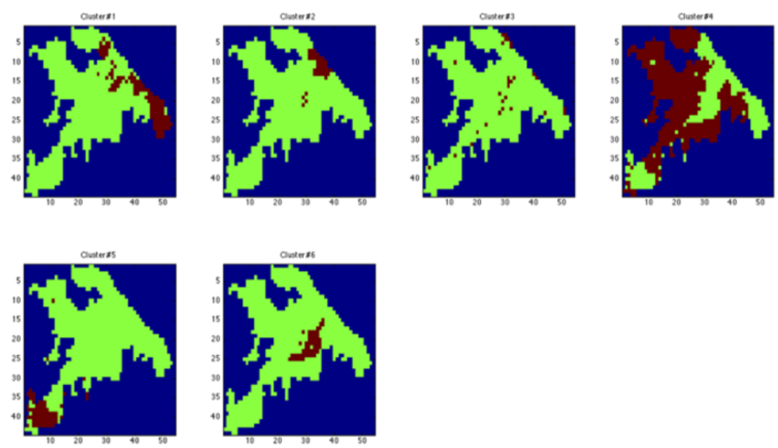


Figure 7.1.2.3. SOM clusters: Thames gateway.

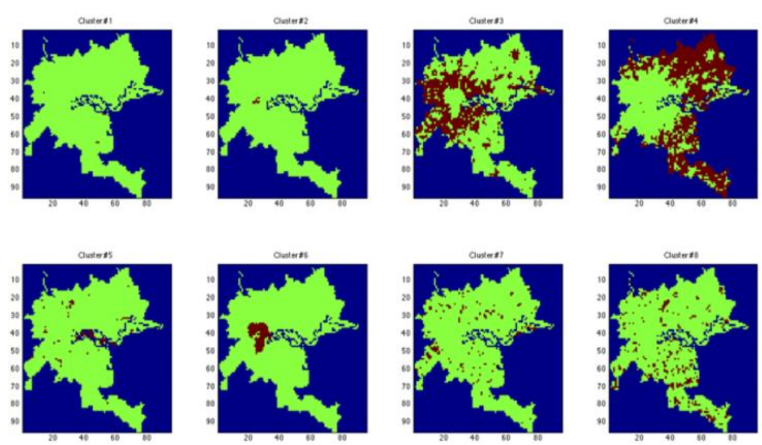


Figure 7.1.2.4. SOM clusters: Portsmouth.

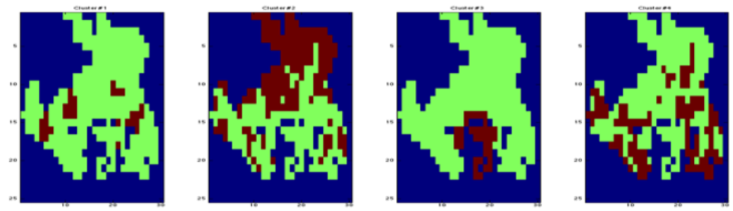


Figure 7.1.2.5. SOM clusters: Gothenburg.

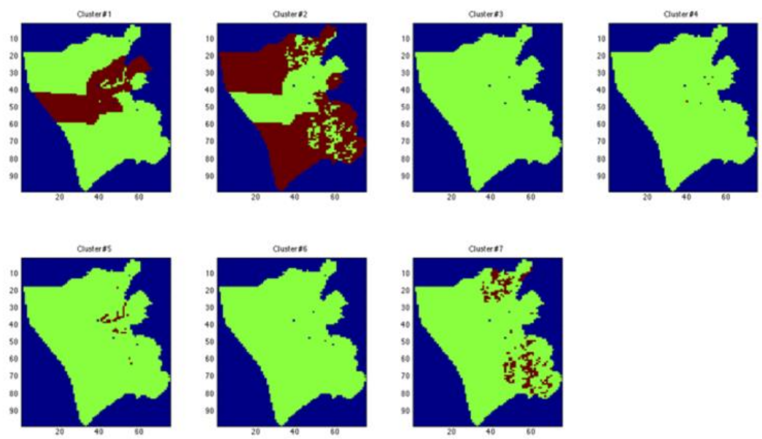


Figure 7.1.2.6. SOM clusters: Malmö.

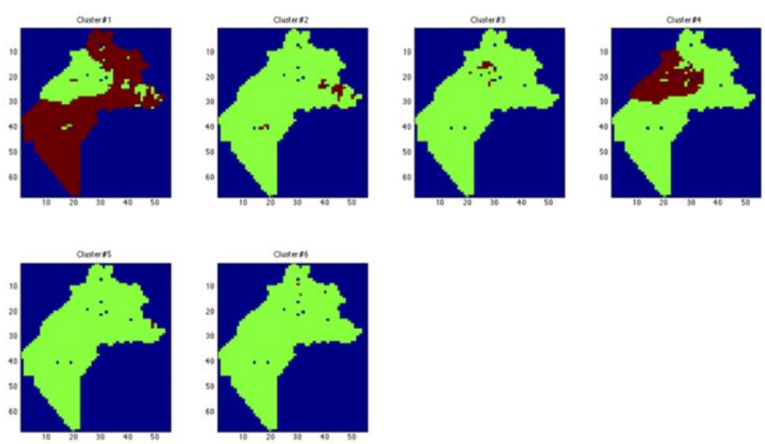


Figure 7.1.2.7. SOM clusters: Zeebrugge.

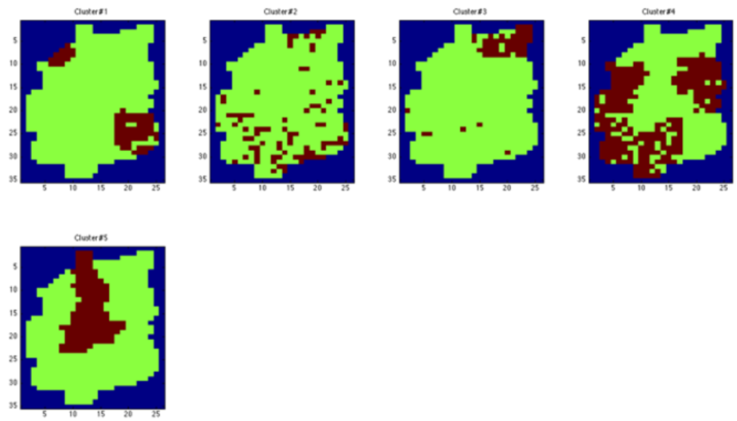


Figure 7.1.2.8. SOM clusters: Ostend.

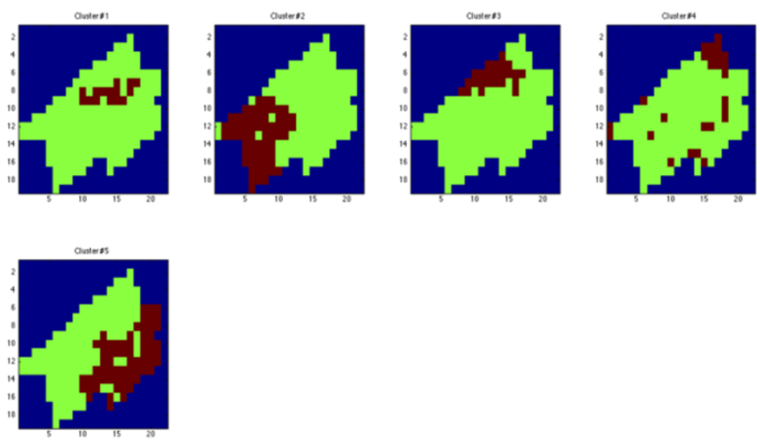


Figure 7.1.2.9. SOM clusters: Lisbon.

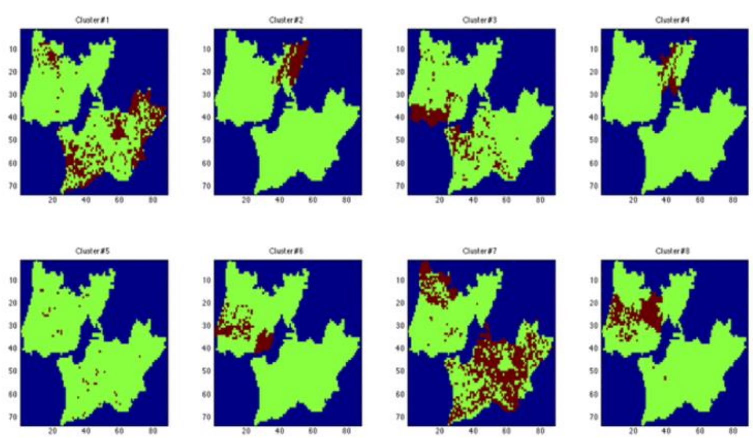


Figure 7.1.2.10. SOM clusters: Algarve.

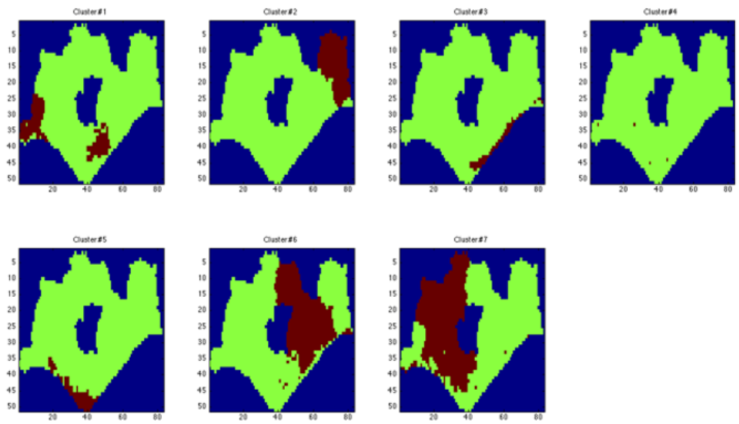
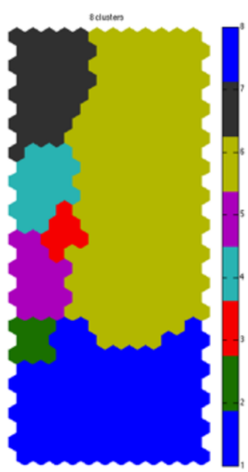
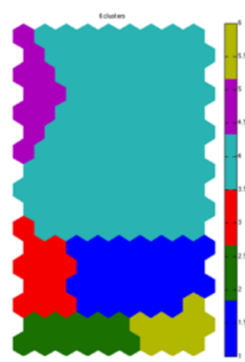


Figure 7.1.2.11 (a to j). Territories SOM clusters.

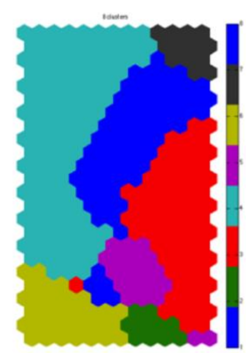
a. Rome.



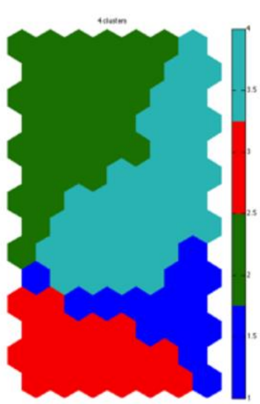
b. Chieti-Pescara.



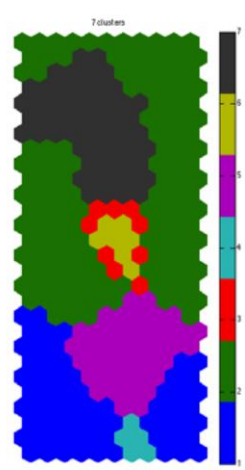
c.. Thames Gateway.



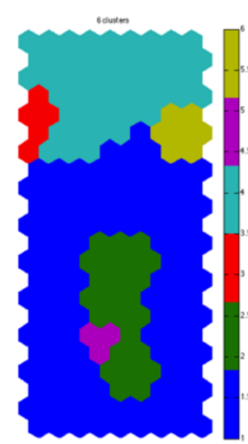
d. Portsmouth.



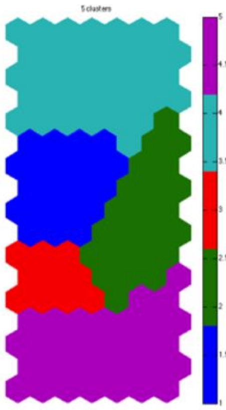
e. Gothenburg.



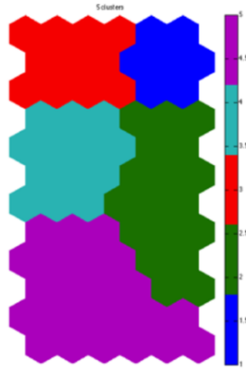
f. Malmö.



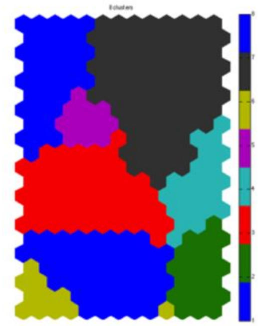
g. Zeebrugge.



h. Ostend.



i. Lisbon.



j. Algarve.

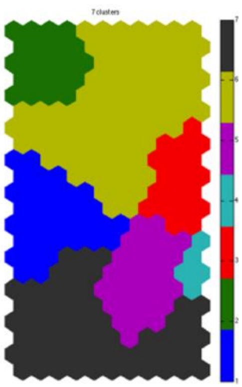


Figure 7.1.2.12. Rome: Prototypes.

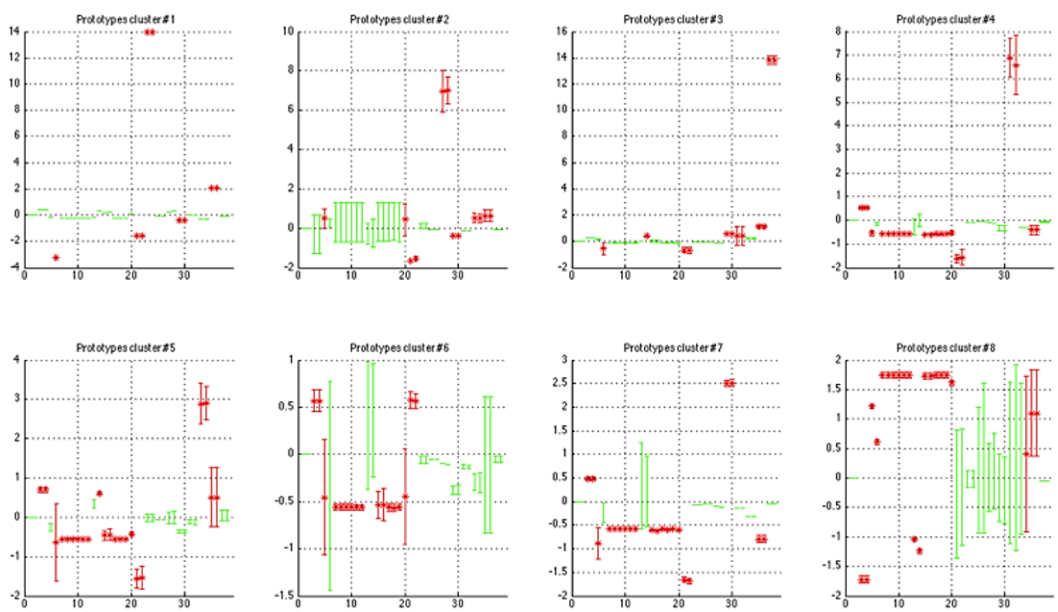


Table 7.1.2.5. Rome: Positive variables.

Cluster #1	Cluster #2	Cluster #3	Cluster #4	Cluster #5	Cluster #6	Cluster #7	Cluster #8
use airports 1, (13.9818)	use industrial commercial 2, (7.0241)	slr1, (13.8201)	use openspace 1, (6.8887)	use residential mixed 2, (2.9013)	use agriculture 1, (0.5727)	use natural 2, (2.5131)	number of tourists 1, (1.7408)
use airports 2, (13.9818)	use industrial commercial 1, (6.9702)	slr2, (13.8201)	use openspace 2, (6.5874)	use residential mixed 1, (2.8853)	net migrants 1, (0.56838)	use natural 1, (2.5049)	number of tourists 2, (1.7408)
floodhazard 1, (2.0866)	floodhazard 1, (0.63732)	floodhazard1, (1.094)	net migrants 1, (0.54658)	net migrants 1, (0.69718)	net migrants 2, (0.56838)	net migrants 1, (0.48856)	employees services 1, (1.7408)
floodhazard 2, (2.0866)	floodhazard 2, (0.63732)	floodhazard2, (1.094)	net migrants 2, (0.54658)	net migrants 2, (0.69718)	use agriculture 2, (0.56565)	net migrants 2, (0.48856)	number households 1, (1.7407)
	use residential mixed 1, (0.52136)	use natural 2, (0.54895)		unemployment rate 2, (0.61017)			number of residents 1, (1.7407)
	motorization rate 1, (0.50534)	use natural 1, (0.54544)		floodhazard1, (0.50722)			number of residents 2, (1.7407)
	use residential mixed 2, (0.50187)	use openspace 1, (0.42124)		floodhazard2, (0.50722)			number households 2, (1.7407)
	employees services 2, (0.44249)	use openspace 2, (0.41424)					employees industrial 1, (1.7397)
		unemployment rate 2, (0.39947)					employees industrial 2, (1.7392)
							employees agriculture 1, (1.7282)
							employees agriculture 2, (1.722)
							employees services 2, (1.6143)
							motorization rate 1, (1.2128)
							floodhazard1, (1.095)
							floodhazard2, (1.095)
							motorization rate 2, (0.61292)
							use residential mixed 2, (0.40685)

Table 7.1.2.6. Rome: Negative variables.

Cluster #1	Cluster #2	Cluster #3	Cluster #4	Cluster #5	Cluster #6	Cluster #7	Cluster #8
use natural 1, (-0.38598)	use natural 2, (-0.39298)	motorization rate 2, (-0.55876)	floodhazard1, (-0.40317)	employees services 2, (-0.43344)	employees services 2, (-0.44746)	number of tourists 1, (-0.57022)	unemployment rate 1, (-1.0446)
use natural 2, (-0.3888)	use natural 1, (-0.39425)	use agriculture 1, (-0.66237)	floodhazard2, (-0.40317)	employees agriculture 2, (-0.44153)	motorization rate 1, (-0.45541)	number of tourists 2, (-0.57211)	unemployment rate 2, (-1.242)
use agriculture 1, (-1.57)	use agriculture 2, (-1.5561)	use agriculture 2, (-0.67488)	employees services 2, (-0.51537)	employees agriculture 1, (-0.46628)	employees agriculture 2, (-0.53477)	employees services 1, (-0.57572)	net migrants 1, (-1.7192)
use agriculture 2, (-1.5746)	use agriculture 1, (-1.6792)		motorization rate 1, (-0.54568)	employees industrial 2, (-0.54701)	employees agriculture 1, (-0.53886)	number households1, (-0.57694)	net migrants 2, (-1.7192)
motorization rate 2, (-3.2255)			number of tourists 1, (-0.5686)	number households2, (-0.54815)	number households2, (-0.56032)	number of residents 1, (-0.57811)	
			number of tourists 2, (-0.56911)	number of residents 2, (-0.54964)	number of residents 2, (-0.56037)	number of residents 2, (-0.5797)	
			employees services 1, (-0.57422)	employees industrial 1, (-0.55051)	number of residents 1, (-0.56048)	number households2, (-0.58004)	
			number households1, (-0.57474)	number of residents 1, (-0.55252)	employees services 1, (-0.56057)	employees industrial 1, (-0.58817)	
			number of residents 1, (-0.57568)	number households1, (-0.55385)	number households1, (-0.56072)	employees industrial 2, (-0.59105)	
			number of residents 2, (-0.57646)	employees services 1, (-0.55461)	number of tourists 2, (-0.56179)	employees agriculture 1, (-0.60965)	
			number households2, (-0.57661)	number of tourists 2, (-0.55702)	number of tourists 1, (-0.56223)	employees services 2, (-0.61224)	
			employees industrial 1, (-0.58283)	number of tourists 1, (-0.56175)	employees industrial 1, (-0.56363)	employees agriculture 2, (-0.61674)	
			employees industrial 2, (-0.58456)	motorization rate 2, (-0.64017)	employees industrial 2, (-0.56404)	floodhazard1, (-0.79526)	
			employees agriculture 1, (-0.60072)	use agriculture 2, (-1.5271)		floodhazard2, (-0.79526)	
			employees agriculture 2, (-0.60444)	use agriculture 1, (-1.5663)		motorization rate 1, (-0.88069)	
			use agriculture 2, (-1.5634)			use agriculture 1, (-1.6563)	
			use agriculture 1, (-1.607)			use agriculture 2, (-1.6843)	

Figure 7.1.2.13. Chieti-Pescara: Prototypes.

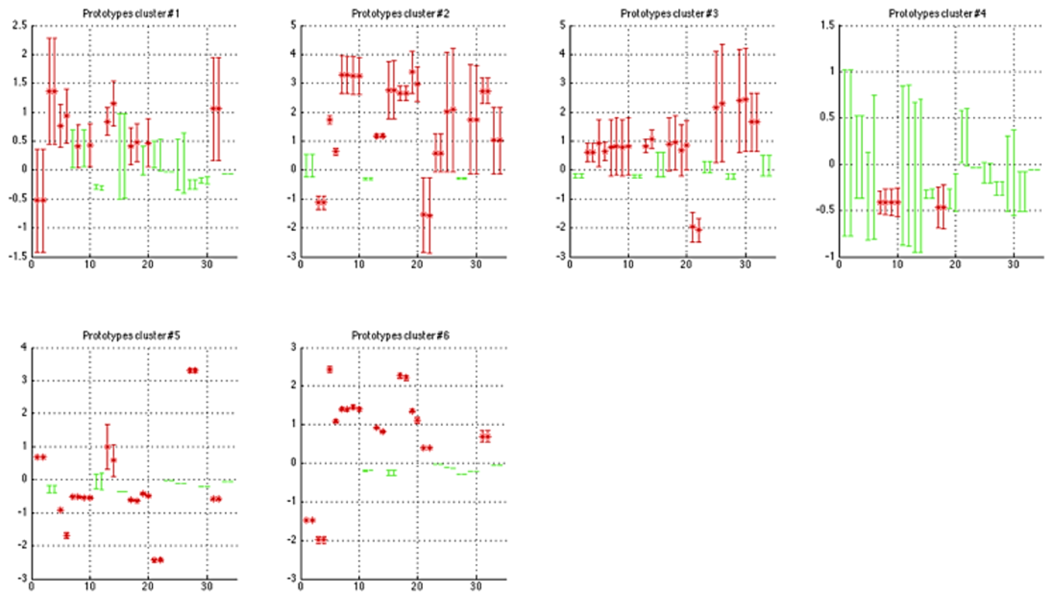


Table 7.1.2.7. Chieti-Pescara: Positive variables.

Cluster #1	Cluster #2	Cluster #3	Cluster #4	Cluster #5	Cluster #6
net migrants 1, (1.3707)	employees services 1, (3.3829)	use residential mixed 2, (2.4386)		use natural 1, (3.3044)	motorization rate 1, (2.4358)
net migrants 2, (1.3707)	number households1, (3.2973)	use residential mixed 1, (2.4)		use natural 2, (3.3044)	employees industrial 1, (2.2625)
unemployment rate 2, (1.1511)	number households2, (3.2931)	use industrial commercial 2, (2.3133)		unemployment rate 1, (0.98457)	employees industrial 2, (2.2101)
floodhazard 1, (1.0594)	number of residents 1, (3.2731)	use industrial commercial 1, (2.1786)		average income 2, (0.67307)	number of residents 1, (1.4478)
floodhazard 2, (1.0594)	number of residents 2, (3.2638)	floodhazard 1, (1.6661)		average income 1, (0.67307)	number households1, (1.409)
motorization rate 2, (0.9332)	employees services 2, (2.9764)	floodhazard 2, (1.6661)		unemployment rate 2, (0.57834)	number of residents 2, (1.405)
unemployment rate 1, (0.83797)	employees agriculture 2, (2.768)	unemployment rate 2, (1.07)			number households2, (1.3944)
motorization rate 1, (0.76067)	employees agriculture 1, (2.7616)	employees industrial 2, (0.95227)			employees services 1, (1.3584)
employees industrial 2, (0.47943)	floodhazard1, (2.7445)	motorization rate 1, (0.93276)			employees services 2, (1.114)

employees services 2, (0.46734)	floodhazard2, (2.7445)	employees industrial 1, (0.89432)			motorization rate 2, (1.0939)
number of residents 2, (0.42891)	employees industrial 1, (2.6565)	employees services 2, (0.85684)			unemployment rate 1, (0.91445)
number households2, (0.41536)	employees industrial 2, (2.6499)	unemployment rate 1, (0.83498)			unemployment rate 2, (0.82563)
employees industrial 1, (0.41042)	use industrial commercial 2, (2.075)	number of residents 2, (0.82987)			floodhazard1, (0.69944)
	use industrial commercial 1, (2.0092)	number households2, (0.81781)			floodhazard2, (0.69944)
	use residential mixed 1, (1.7575)	number of residents 1, (0.78015)			use agriculture 2, (0.39401)
	motorization rate 1, (1.75)	number households1, (0.77286)			use agriculture 1, (0.39027)
	use residential mixed 2, (1.7297)	employees services 1, (0.67027)			
	unemployment rate 2, (1.1878)	motorization rate 2, (0.6512)			
	unemployment rate 1, (1.1701)	net migrants 1, (0.61125)			
	slr1, (1.0198)	net migrants 2, (0.61125)			
	slr2, (1.0198)				
	motorization rate 2, (0.63039)				
	use airports 1, (0.58811)				
	use airports 2, (0.58811)				

Table 7.1.2.8. Chieti-Pescara: Negative variables.

Cluster #1	Cluster #2	Cluster #3	Cluster #4	Cluster #5	Cluster #6
average income 1, (-0.53017)	net migrants 1, (-1.1175)	use agriculture 1, (-1.9732)	number households2, (-0.40608)	employees services 1, (-0.4279)	average income 1, (-1.4804)
average income 2, (-0.53017)	net migrants 2, (-1.1175)	use agriculture 2, (-2.0763)	number households1, (-0.40882)	employees services 2, (-0.48812)	average income 2, (-1.4804)
	use agriculture 1, (-1.544)		number of residents 2, (-0.41101)	number households1, (-0.50716)	net migrants 1, (-1.9842)
	use agriculture 2, (-1.58)		number of residents 1, (-0.41215)	number households2, (-0.52111)	net migrants 2, (-1.9842)
			employees industrial 2, (-0.46118)	number of residents 1, (-0.53902)	
			employees industrial 1, (-0.46309)	number of residents 2, (-0.55464)	
				floodhazard1, (-0.58634)	
				floodhazard2, (-0.58634)	
				employees industrial 1, (-0.60397)	
				employees industrial 2, (-0.62955)	
				motorization rate 1, (-0.91991)	
				motorization rate 2, (-1.6972)	
				use agriculture 2, (-2.4166)	
				use agriculture 1, (-2.4401)	

Figure 7.1.2.14. Thames Gateway: Prototypesl.

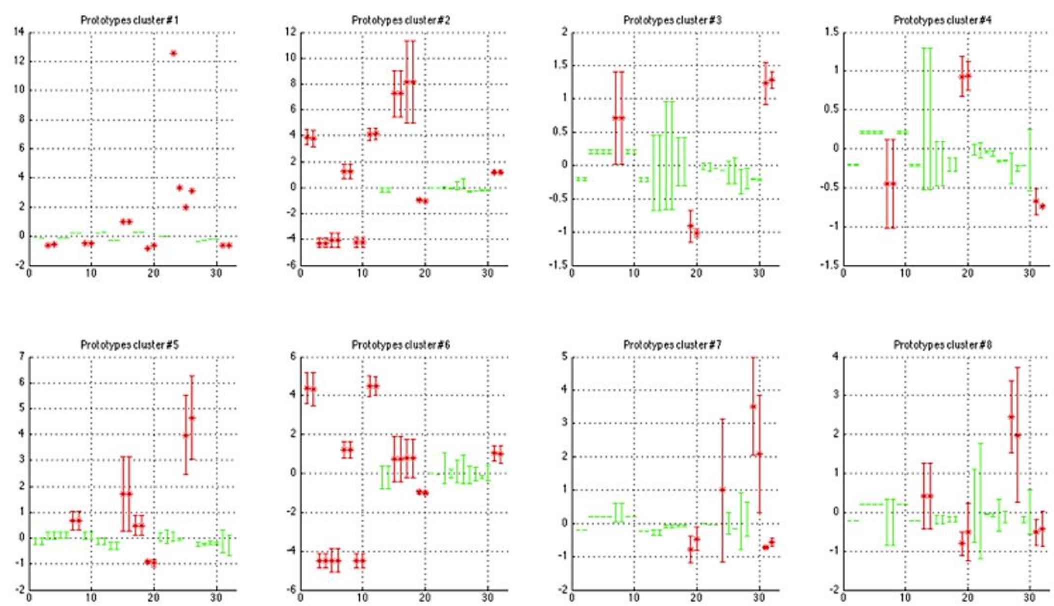


Table 7.1.2.9. Thames Gateway: Positive variables.

Cluster #1	Cluster #2	Cluster #3	Cluster #4	Cluster #5	Cluster #6	Cluster #7	Cluster #8
use industrial 1, (12.5408)	employees services 1, (8.1632)	use residential mixed 2, (1.2864)	use agriculture 2, (0.93776)	use industrial commercial 2, (4.6472)	unemployment rate 2, (4.5054)	use openspace 1, (3.5129)	use natural 1, (2.4488)
use industrial 2, (3.2985)	employees services 2, (8.1632)	use residential mixed 1, (1.2299)	use agriculture 1, (0.92925)	use industrial commercial 1, (3.9772)	unemployment rate 1, (4.4929)	use openspace 2, (2.0777)	use natural 2, (1.9894)
use industrial commercial 2, (3.1312)	employees industrial 1, (7.2492)	number households1, (0.714)		employees industrial 1, (1.705)	average income 1, (4.382)	use industrial 2, (0.99786)	employees agriculture 1, (0.42034)
use industrial commercial 1, (1.9861)	employees industrial 2, (7.2492)	number households2, (0.714)		employees industrial 2, (1.705)	average income 2, (4.3365)		employees agriculture 2, (0.42034)
employees industrial 1, (1.0035)	unemployment rate 2, (4.1506)			number households1, (0.68421)	number households1, (1.2274)		
employees industrial 2, (1.0035)	unemployment rate 1, (4.1081)			number households2, (0.68421)	number households2, (1.2274)		
	average income 1, (3.8858)			employees services 1, (0.49684)	use residential mixed 1, (1.052)		
	average income 2, (3.8125)			employees services 2, (0.49684)	use residential mixed 2, (0.98104)		
	number households1, (1.252)				employees services 1, (0.78366)		
	number households2, (1.252)				employees services 2, (0.78366)		
	use residential mixed 1, (1.2101)				employees industrial 1, (0.74231)		
	use residential mixed 2, (1.188)				employees industrial 2, (0.74231)		

Table 7.1.2.10. Thames Gateway: Negative variables.

Cluster #1	Cluster #2	Cluster #3	Cluster #4	Cluster #5	Cluster #6	Cluster #7	Cluster #8
number of residents 1, (-0.46358)	use agriculture 1, (-0.9509)	use agriculture 1, (-0.90933)	number households1, (-0.44887)	use agriculture 1, (-0.92807)	use agriculture 1, (-0.95695)	use agriculture 2, (-0.46223)	use residential mixed 2, (-0.43212)
number of residents 2, (-0.46908)	use agriculture 2, (-1.0255)	use agriculture 2, (-1.0101)	number households2, (-0.44887)	use agriculture 2, (-0.95853)	use agriculture 2, (-0.99341)	use residential mixed 2, (-0.57329)	use agriculture 2, (-0.51031)
net migrants 2, (-0.57684)	motorization rate 1, (-4.0508)		use residential mixed 1, (-0.67637)		motorization rate 1, (-4.47)	use residential mixed 1, (-0.71299)	use residential mixed 1, (-0.51533)
use agriculture 2, (-0.59224)	motorization rate 2, (-4.0536)		use residential mixed 2, (-0.73549)		motorization rate 2, (-4.4713)	use agriculture 1, (-0.78002)	use agriculture 1, (-0.80616)
net migrants 1, (-0.59736)	number of residents 1, (-4.228)				net migrants 1, (-4.4836)		
use residential mixed 1, (-0.6014)	number of residents 2, (-4.2299)				net migrants 2, (-4.4892)		
use residential mixed 2, (-0.64919)	net migrants 2, (-4.261)				number of residents 2, (-4.5083)		
use agriculture 1, (-0.86833)	net migrants 1, (-4.2652)				number of residents 1, (-4.5089)		

Figure 7.1.2.15. Portsmouth: Prototypes.

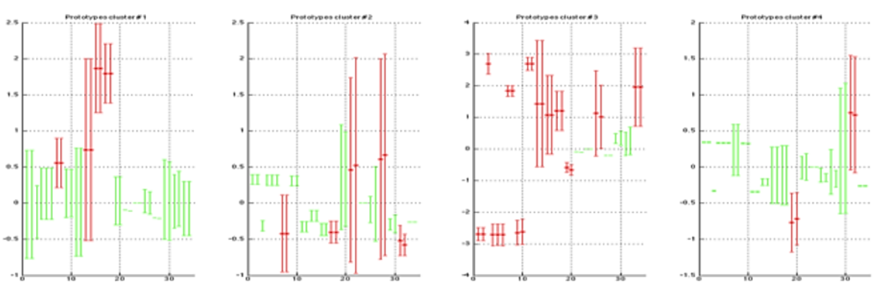


Table 7.1.2.11. Portsmouth: Positive variables.

Cluster #1	Cluster #2	Cluster #3	Cluster #4
employees industrial 1, (1.8667)	use natural 2, (0.66917)	net migrants 1, (2.7017)	use residential mixed 1, (0.75309)
employees industrial 2, (1.8667)	use natural 1, (0.60879)	unemployment rate 1, (2.6965)	use residential mixed 2, (0.72316)
employees services 1, (1.7946)	use airports 2, (0.52417)	unemployment rate 2, (2.6965)	
employees services 2, (1.7946)	use airports 1, (0.45906)	floodhazard1, (1.9551)	
employees agriculture 1, (0.73932)		floodhazard2, (1.9551)	
employees agriculture 2, (0.73932)		number households1, (1.8395)	
number households1, (0.55543)		number households2, (1.8395)	
number households2, (0.55543)		employees agriculture 1, (1.4364)	
		employees agriculture 2, (1.4364)	
		employees services 1, (1.2066)	
		employees services 2, (1.2066)	
		use industrial commercial 1, (1.1279)	
		employees industrial 1, (1.0806)	
		employees industrial 2, (1.0806)	
		use industrial commercial 2, (1.0191)	

Table 7.1.2.12. Portsmouth: Negative variables.

Cluster #1	Cluster #2	Cluster #3	Cluster #4
	employees services 1, (-0.40078)	use agriculture 1, (-0.58418)	use agriculture 2, (-0.71469)
	employees services 2, (-0.40078)	use agriculture 2, (-0.65717)	use agriculture 1, (-0.76754)
	number households1, (-0.42126)	number of residents 2, (-2.6173)	
	number households2, (-0.42126)	number of residents 1, (-2.6449)	
	use residential mixed 1, (-0.51823)	average income 2, (-2.6965)	
	use residential mixed 2, (-0.57826)	average income 1, (-2.6965)	
		net migrants 2, (-2.7127)	
		motorization rate 2, (-2.713)	
		motorization rate 1, (-2.7154)	

Figure 7.1.2.16. Gothenburg: Prototypes.

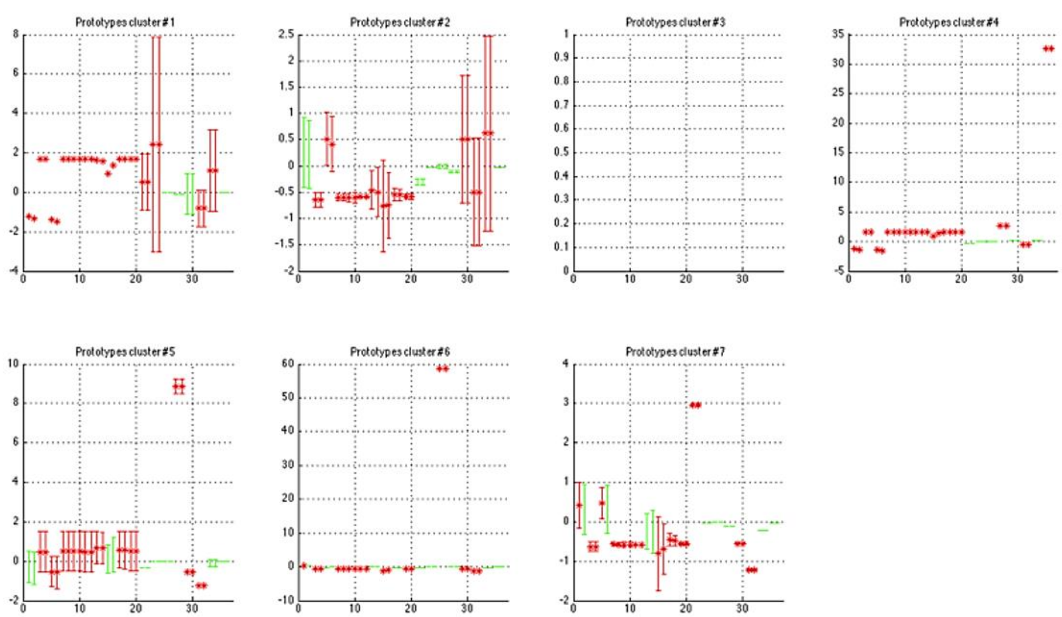


Table 7.1.2.13. Gothenburg: Positive variables.

Cluster #1	Cluster #2	Cluster #3	Cluster #4	Cluster #5	Cluster #6	Cluster #7
use airports 1, (2.4215)	use residential mixed 1, (0.61813)		slr1, (32.6059)	use industrial commercial 1, (8.8489)	use industrial 1, (58.7065)	use agriculture 1, (2.968)
use airports 2, (2.4215)	use residential mixed 2, (0.61813)		slr2, (32.6059)	use industrial commercial 2, (8.8489)	use industrial 2, (58.7065)	use agriculture 2, (2.968)
number of tourists 2, (1.6846)	motorization rate 1, (0.50916)		use industrial commercial 1, (2.6845)	unemployment rate 1, (0.68424)	average income 1, (0.4347)	motorization rate 1, (0.47643)
number of tourists 1, (1.6846)	use natural 1, (0.50664)		use industrial commercial 2, (2.6845)	unemployment rate 2, (0.67481)		average income 1, (0.42219)
employees services 1, (1.6827)	use natural 2, (0.50664)		number of tourists 2, (1.6849)	employees ndustrial 1, (0.58704)		
employees services 2, (1.6822)	motorization rate 2, (0.41583)		number of tourists 1, (1.6849)	employees industrial 2, (0.5695)		
number households1, (1.6807)			employees services 1, (1.683)	employees services 2, (0.51624)		
employees industrial 2, (1.6802)			employees services 2, (1.6825)	number households1, (0.50949)		
number households2, (1.6794)			number households1, (1.681)	employees services 1, (0.50891)		
employees industrial 1, (1.6785)			employees industrial 2, (1.6805)	number of residents 1, (0.50523)		
number of residents 1, (1.6769)			number households2, (1.6797)	number households2, (0.50447)		
number of residents 2, (1.6757)			employees industrial 1, (1.6788)	number of residents 2, (0.50255)		
net migrants 1, (1.6666)			number of residents 1, (1.6773)	number of tourists 2, (0.48286)		
net migrants 2, (1.6666)			number of residents 2, (1.6761)	number of tourists 1, (0.48286)		
unemployment rate 1, (1.6362)			net migrants 1, (1.6669)	net migrants 1, (0.47927)		

Unemployment rate 2, (1.5749)			net migrants 2, (1.6669)	net migrants 2, (0.47927)		
employees agriculture 2, (1.3477)			unemployment rate 1, (1.6364)			
use residential mixed 1, (1.1017)			unemployment rate 2, (1.5751)			
use residential mixed 2, (1.1017)			employees agriculture 2, (1.348)			
employees agriculture 1, (0.93913)			employees agriculture 1, (0.93942)			
use agriculture 1, (0.53876)						
use agriculture 2, (0.53876)						

Table 7.1.2.14. Gothenburg: Negative variables.

Cluster #1	Cluster #2	Cluster #3	Cluster #4	Cluster #5	Cluster #6	Cluster #7
use openspace 1, (-0.82006)	unemployment rate 1, (-0.45914)		use openspace 1, (-0.56702)	motorization rate 1, (-0.53192)	employees services 2, (-0.4961)	employees industrial 1, (-0.45308)
use openspace 2, (-0.82006)	use openspace 1, (-0.49692)		use openspace 2, (-0.56702)	use natural 1, (-0.54939)	employees services 1, (-0.52073)	employees industrial 2, (-0.47978)
average income 1, (-1.2223)	use openspace 2, (-0.49692)		average income 1, (-1.2225)	use natural 2, (-0.54939)	number households 1, (-0.52325)	use natural 1, (-0.55143)
average income 2, (-1.321)	unemployment rate 2, (-0.4989)		average income 2, (-1.3211)	motorization rate 2, (-0.55335)	number of residents 1, (-0.54312)	use natural 2, (-0.55143)
motorization rate 1, (-1.4002)	employees industrial 1, (-0.5377)		motorization rate 1, (-1.4005)	use openspace 1, (-1.2283)	number households 2, (-0.54356)	employees services 2, (-0.55505)
motorization rate 2, (-1.4969)	employees industrial 2, (-0.55111)		motorization rate 2, (-1.4971)	use openspace 2, (-1.2283)	number of residents 2, (-0.55333)	employees services 1, (-0.56302)
	employees services 2, (-0.58552)				use natural 1, (-0.55791)	number households 1, (-0.56488)
	employees services 1, (-0.58656)				use natural 2, (-0.55791)	number households 2, (-0.57381)
	number of tourists 2, (-0.59132)				number of tourists 2, (-0.59334)	number of residents 1, (-0.57401)
	number of tourists 1, (-0.59132)				number of tourists 1, (-0.59334)	number of residents 2, (-0.57927)
	number households 1, (-0.59319)				net migrants 1, (-0.64049)	number of tourists 2, (-0.59334)
	number households 2, (-0.59903)				net migrants 2, (-0.64049)	number of tourists 1, (-0.59334)
	number of residents 1, (-0.60323)				employees griculture 2, (-0.77699)	net migrants 1, (-0.62603)
	number of residents 2, (-0.60735)				employees agriculture 1, (-1.1454)	net migrants 2, (-0.62603)
	net migrants 1, (-0.63692)				use openspace 1, (-1.1514)	employees agriculture 2, (-0.69273)
	net migrants 2, (-0.63692)				use openspace 2, (-1.1514)	employees griculture 1, (-0.80347)
	employees agriculture 2, (-0.74472)					use openspace 1, (-1.2148)
	employees agriculture 1, (-0.76636)					use openspace 2, (-1.2148)

Figure 7.1.2.17. Malmö: Prototypes.

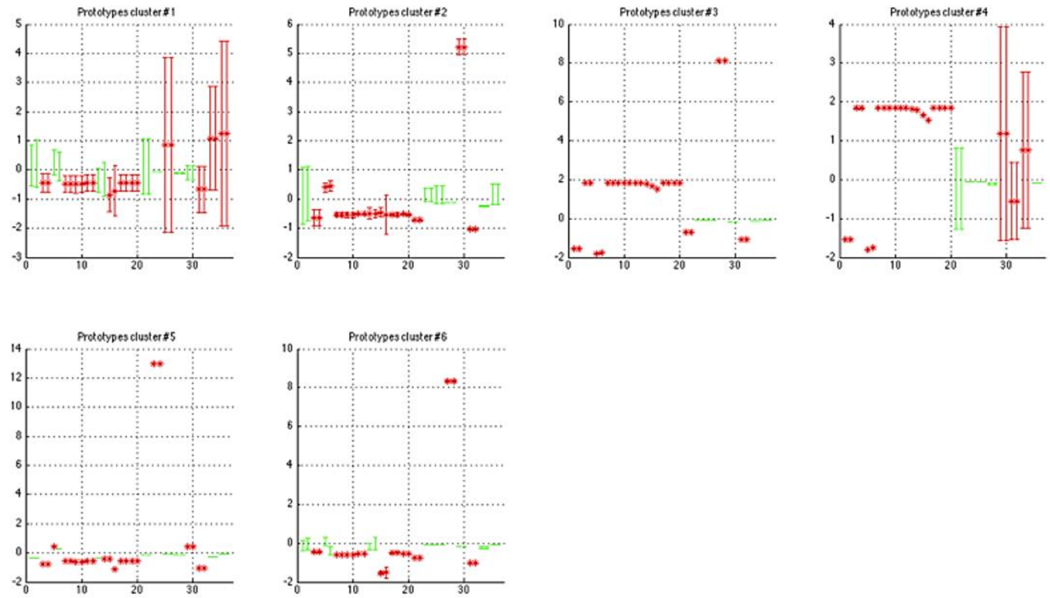


Table 7.1.2.15. Malmö: Positive variables.

Cluster #1	Cluster #2	Cluster #3	Cluster #4	Cluster #5	Cluster #6
slr1, (1.2524)	use natural 1, (5.2206)	use industrial commercial 1, (8.1282)	number of tourists 2, (1.8517)	use airports 1, (12.9806)	use industrial commercial 1, (8.3317)
slr2, (1.2524)	use natural 2, (5.2206)	use industrial commercial 2, (8.1282)	number of tourists 1, (1.8517)	use airports 2, (12.9806)	use industrial commercial 2, (8.3317)
use residential mixed 1, (1.0799)	motorization rate 2, (0.45175)	number of tourists 2, (1.8527)	employees services 1, (1.8516)	motorization rate 1, (0.43097)	
use residential mixed 2, (1.0799)	motorization rate 1, (0.40885)	number of tourists 1, (1.8527)	employees services 2, (1.8516)	use natural 1, (0.41285)	
use industrial 1, (0.85448)		employees services 1, (1.8526)	employees industrial 1, (1.8514)	use natural 2, (0.41285)	
use industrial 2, (0.85448)		employees services 2, (1.8526)	employees industrial 2, (1.8511)		
		employees industrial 1, (1.8524)	number households1, (1.8505)		
		employees industrial 2, (1.8521)	number households2, (1.8503)		
		number households1, (1.8514)	number of residents 2, (1.8496)		
		number households2, (1.8512)	number of residents 1, (1.8491)		
		number of residents 2, (1.8504)	net migrants 1, (1.8348)		

		number of residents 1, (1.8499)	net migrants 2, (1.8348)		
		net migrants 1, (1.8358)	unemployment rate 1, (1.8148)		
		net migrants 2, (1.8358)	unemployment rate 2, (1.7866)		
		unemployment rate 1, (1.8164)	employees agriculture 1, (1.6661)		
		unemployment rate 2, (1.7885)	employees agriculture 2, (1.5316)		
		employees agriculture 1, (1.666)	use natural 1, (1.1888)		
		employees agriculture 2, (1.531)	use natural 2, (1.1888)		
			use residential mixed 1, (0.75823)		
			use residential mixed 2, (0.75823)		

Table 7.1.2.16. Malmö: Negative variables.

Cluster #1	Cluster #2	Cluster #3	Cluster #4	Cluster #5	Cluster #6
employees industrial 2, (-0.43733)	employees agriculture 1, (-0.45213)	use agriculture 1, (-0.72297)	use openspace 1, (-0.54906)	unemployment rate 2, (-0.43079)	net migrants 1, (-0.45172)
number of tourists 1, (-0.45115)	unemployment rate 1, (-0.48621)	use agriculture 2, (-0.72297)	use openspace 2, (-0.54906)	employees agriculture 1, (-0.44804)	net migrants 2, (-0.45172)
number of tourists 2, (-0.45115)	unemployment rate 2, (-0.50263)	use openspace 1, (-1.0537)	average income 2, (-1.5416)	number of tourists 1, (-0.53902)	employees industrial 2, (-0.47008)
employees industrial 1, (-0.45355)	employees services 1, (-0.51707)	use openspace 2, (-1.0537)	average income 1, (-1.5462)	number of tourists 2, (-0.53902)	employees industrial 1, (-0.51488)
net migrants 1, (-0.45575)	number of tourists 1, (-0.51943)	average income 2, (-1.5435)	motorization rate 2, (-1.7555)	employees services 1, (-0.54205)	number of tourists 1, (-0.52563)
net migrants 2, (-0.45575)	number of tourists 2, (-0.51943)	average income 1, (-1.5481)	motorization rate 1, (-1.8077)	employees industrial 1, (-0.54661)	number of tourists 2, (-0.52563)
employees services 1, (-0.45748)	employees services 2, (-0.52224)	motorization rate 2, (-1.7575)		employees services 2, (-0.55038)	employees services 2, (-0.54002)
employees services 2, (-0.45804)	employees industrial 1, (-0.52562)	motorization rate 1, (-1.8093)		employees industrial 2, (-0.57011)	employees services 1, (-0.54296)
number households1, (-0.47641)	employees agriculture 2, (-0.53497)			number households1, (-0.58661)	number households1, (-0.57716)

number households2, (-0.47824)	number households1, (-0.54042)			number households2, (-0.59024)	number households2, (-0.58274)
number of residents 2, (-0.48447)	employees industrial 2, (-0.54122)			number of residents 2, (-0.60231)	number of residents 2, (-0.59664)
number of residents 1, (-0.48959)	number households2, (-0.54131)			number of residents 1, (-0.60683)	number of residents 1, (-0.60683)
use openspace 1, (-0.67177)	number of residents 2, (-0.54645)			net migrants 1, (-0.73924)	use agriculture 1, (-0.74597)
use openspace 2, (-0.67177)	number of residents 1, (-0.54858)			net migrants 2, (-0.73924)	use agriculture 2, (-0.74597)
employees agriculture 2, (-0.71091)	net migrants 1, (-0.63389)			use openspace 1, (-1.0536)	use openspace 1, (-1.0313)
employees agriculture 1, (-0.85138)	net migrants 2, (-0.63389)			use openspace 2, (-1.0536)	use openspace 2, (-1.0313)
	use agriculture 1, (-0.72478)			employees agriculture 2, (-1.0886)	employees agriculture 2, (-1.5177)
	use agriculture 2, (-0.72478)				employees agriculture 1, (-1.5334)
	use openspace 1, (-1.0445)				
	use openspace 2, (-1.0445)				

Figure 7.1.2.18. Zeebrugge: Prototypes.

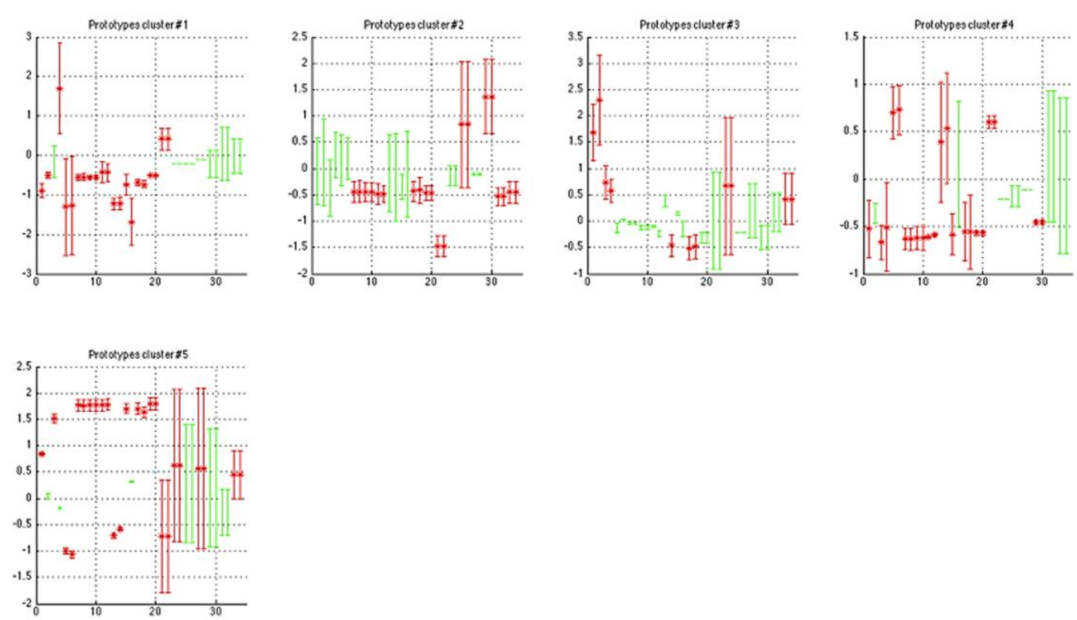


Table 7.1.2.17. Zeebrugge: Positive variables.

Cluster #1	Cluster #2	Cluster #3	Cluster #4	Cluster #5
net migrants 2, (1.6983)	use residential mixed 1, (1.3672)	average income 2, (2.3017)	motorization rate 2, (0.73177)	employees services 1, (1.7917)
use agriculture 1, (0.41009)	use residential mixed 2, (1.3672)	average income 1, (1.6854)	motorization rate 1, (0.69726)	employees services 2, (1.7915)
use agriculture 2, (0.41009)	use natural 1, (0.83637)	net migrants 1, (0.73088)	use agriculture 1, (0.59896)	number of tourists 2, (1.7812)
	use natural 2, (0.83637)	use industrial commercial 1, (0.66961)	use agriculture 2, (0.59896)	number of residents 1, (1.7719)
		use industrial commercial 2, (0.66961)	unemployment rate 2, (0.53809)	number of tourists 1, (1.7705)
		net migrants 2, (0.5804)	unemployment rate 1, (0.38792)	number of residents 2, (1.7701)
		slr1, (0.4209)		number households1, (1.7673)
		slr2, (0.4209)		number households2, (1.7655)
				employees agriculture 1, (1.7027)
				employees industrial 1, (1.7025)
				employees industrial 2, (1.6341)
				net migrants 1, (1.5229)
				average income 1, (0.84174)
				use industrial commercial 1, (0.62094)
				use industrial commercial 2, (0.62094)
				use openspace 1, (0.57371)
				use openspace 2, (0.57371)
				slr1, (0.4497)
				slr2, (0.4497)

Table 7.1.2.18. Zeebrugge: Negative variables.

Cluster #1	Cluster #2	Cluster #3	Cluster #4	Cluster #5
number of tourists 1, (-0.41764)	employees industrial 2, (-0.4102)	unemployment rate 2, (-0.45986)	use residential mixed 1, (-0.45843)	unemployment rate 2, (-0.57828)
number of tourists 2, (-0.42883)	employees industrial 1, (-0.43241)	employees industrial 2, (-0.48008)	use residential mixed 2, (-0.45843)	unemployment rate 1, (-0.70905)
average income 2, (-0.49715)	number of residents 2, (-0.43954)	employees industrial 1, (-0.5141)	net migrants 2, (-0.50524)	use agriculture 1, (-0.72255)
employees services 1, (-0.50932)	number of residents 1, (-0.44097)		average income 1, (-0.5252)	use agriculture 2, (-0.72255)
employees services 2, (-0.50932)	number households2, (-0.44128)		employees industrial 2, (-0.55746)	motorization rate 1, (-0.99522)
number households2, (-0.54665)	number households1, (-0.44234)		employees industrial 1, (-0.5576)	motorization rate 2, (-1.0652)
number households1, (-0.55311)	slr1, (-0.45045)		employees services 1, (-0.56784)	
number of residents 2, (-0.556)	slr2, (-0.45045)		employees services 2, (-0.56889)	
number of residents 1, (-0.5649)	employees services 2, (-0.4673)		employees agriculture 1, (-0.58235)	
employees industrial 1, (-0.68667)	employees services 1, (-0.46842)		number of tourists 2, (-0.5918)	
employees industrial 2, (-0.73146)	number of tourists 1, (-0.4809)		number of tourists 1, (-0.61243)	
employees agriculture 1, (-0.74532)	number of tourists 2, (-0.48829)		number of residents 1, (-0.61938)	
average income 1, (-0.88575)	floodhazard1, (-0.52851)		number of residents 2, (-0.62399)	
unemployment rate 2, (-1.2081)	floodhazard2, (-0.52851)		number households1, (-0.62963)	
unemployment rate 1, (-1.2236)	use agriculture 1, (-1.4675)		number households2, (-0.63299)	
motorization rate 2, (-1.2577)	use agriculture 2, (-1.4675)		net migrants 1, (-0.66725)	
motorization rate 1, (-1.3079)				
employees agriculture 2, (-1.6789)				

Figure 7.1.2.19. Ostend: Prototypes.

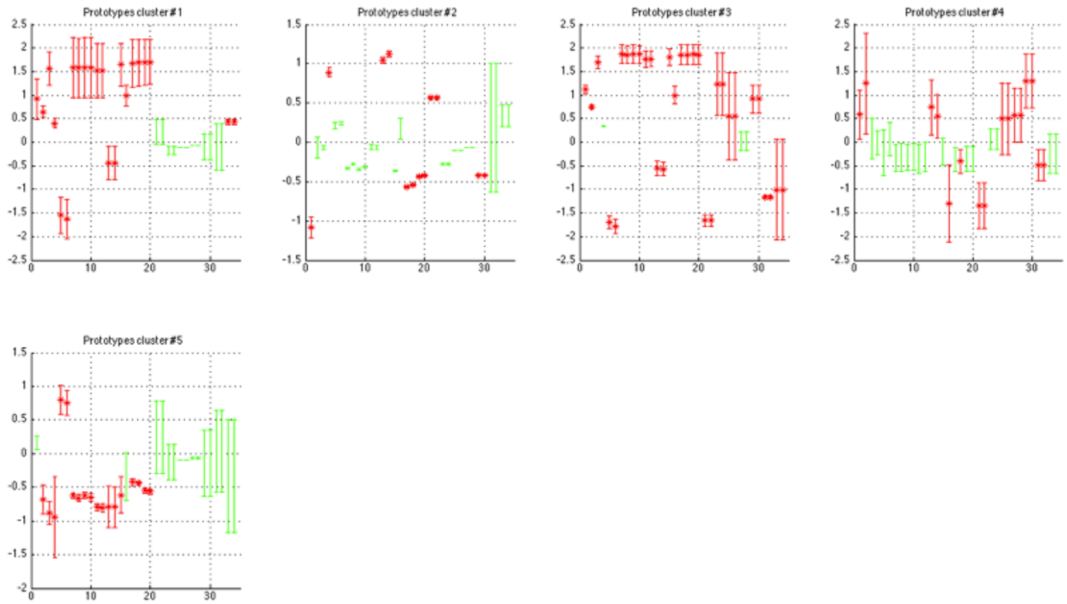


Table 7.1.2.19. Ostend: Positive variables.

Cluster #1	Cluster #2	Cluster #3	Cluster #4	Cluster #5
employees services 2, (1.7009)	unemployment rate 2, (1.1223)	number of residents 1, (1.8714)	use residential mixed 1, (1.2982)	motorizationrate 1, (0.80025)
employees services 1, (1.7004)	unemployment rate 1, (1.0432)	number households1, (1.8687)	use residential mixed 2, (1.2982)	motorizationrate 2, (0.75183)
employees industrial 2, (1.6873)	net migrants 2, (0.88808)	number of residents 2, (1.8639)	average income 2, (1.2499)	
employees industrial 1, (1.6813)	use agriculture 1, (0.56251)	employees services 1, (1.8615)	unemploymentrate 1, (0.74394)	
employees agriculture 1, (1.6454)	use agriculture 2, (0.56251)	employees services 2, (1.8608)	average income 1, (0.58605)	
number of residents 1, (1.5839)		number households2, (1.8572)	use openspace 1, (0.57126)	
number households1, (1.5836)		employees industrial 2, (1.8559)	use openspace 2, (0.57126)	
number of residents 2, (1.5776)		employees industrial 1, (1.8536)	unemployment rate 2, (0.54823)	
number households2, (1.5752)		employees agriculture 1, (1.8013)	use natural 1, (0.49614)	
net migrants 1, (1.5603)		number of tourists 1, (1.7711)	use natural 2, (0.49614)	

number of tourists 1, (1.5178)		number of tourists 2, (1.7707)		
number of tourists 2, (1.5152)		net migrants 1, (1.6905)		
employees agriculture 2, (0.99193)		use industrial commercial 1, (1.228)		
average income 1, (0.91638)		use industrial commercial 2, (1.228)		
average income 2, (0.64306)		average income 1, (1.1185)		
slr1, (0.43394)		employees agriculture 2, (1.001)		
slr2, (0.43394)		use residential mixed 1, (0.91411)		
net migrants 2, (0.40431)		use residential mixed 2, (0.91411)		
		average income 2, (0.74797)		
		use natural 1, (0.54858)		
		use natural 2, (0.54858)		

Table 7.1.2.20. Ostend: Negative variables.

Cluster #1	**Cluster #2**	**Cluster #3**	**Cluster #4**	**Cluster #5**
unemployment rate 1, (-0.43801)	use residential mixed 1, (-0.42185)	unemployment rate 1, (-0.54713)	employees industrial 2, (-0.40397)	employees industrial 1, (-0.42602)
unemployment rate 2, (-0.44083)	use residential mixed 2, (-0.42185)	unemployment rate 2, (-0.56571)	floodhazard1, (-0.48187)	employees industrial 2, (-0.44167)
motorization rate 1, (-1.5458)	employees services 2, (-0.42344)	slr1, (-1.0098)	floodhazard2, (-0.48187)	employees services 1, (-0.54648)
motorization rate 2, (-1.636)	employees services 1, (-0.43858)	slr2, (-1.0098)	employees agriculture 2, (-1.2966)	employees services 2, (-0.56096)
	employees industrial 2, (-0.54268)	floodhazard1, (-1.1598)	use agriculture 1, (-1.3409)	employees agriculture 1, (-0.61647)
	employees industrial 1, (-0.56995)	floodhazard2, (-1.1598)	use agriculture 2, (-1.3409)	number of residents 1, (-0.61679)
	average income 1, (-1.0837)	use agriculture 1, (-1.6584)		number households1, (-0.62562)

		use agriculture 2, (-1.6584)		number of residents 2, (-0.6464)
		motorization rate 1, (-1.6952)		number households2, (-0.66204)
		motorization rate 2, (-1.784)		average income 2, (-0.68472)
				unemployment rate 1, (-0.79061)
				number of tourists 1, (-0.79338)
				unemployment rate 2, (-0.79543)
				number of tourists 2, (-0.80427)
				net migrants 1, (-0.88309)
				net migrants 2, (-0.94058)

Figure 7.1.2.20. Lisbon: Prototypes.

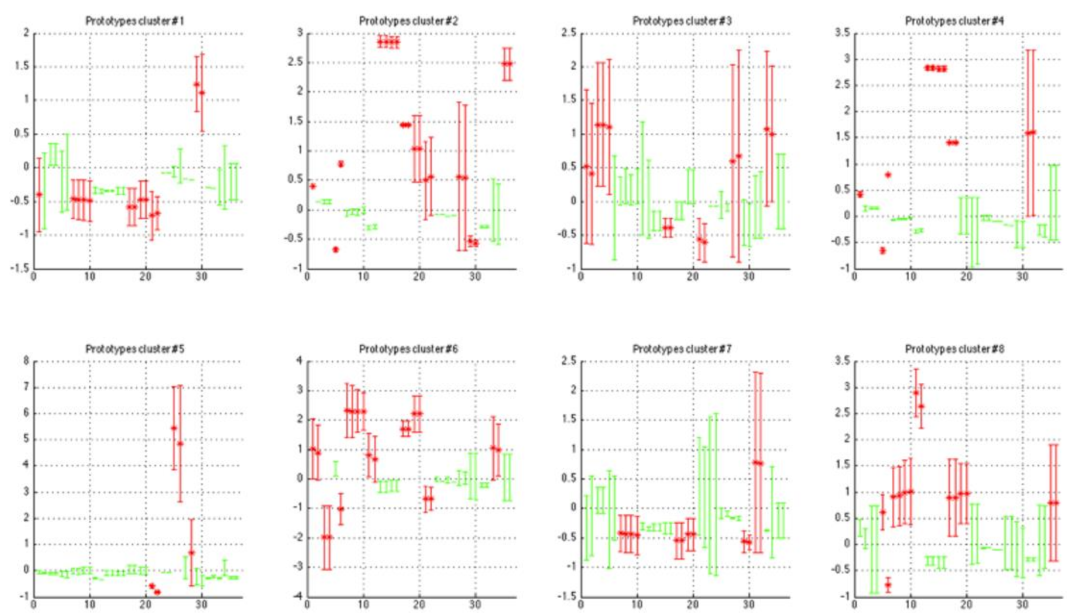


Table 7.1.2.21. Lisbon: Positive variables.

Cluster #1	Cluster #2	Cluster #3	Cluster #4	Cluster #5	Cluster #6	Cluster #7	Cluster #8
use natural 1, (1.2432)	unemployment rate 1, (2.8573)	net migrants 1, (1.1407)	unemployment rate 1, (2.84)	use industrial 1, (5.4397)	number households1, (2.3224)	use openspace 1, (0.77884)	number of tourists 1, (2.8984)
use natural 2, (1.117)	unemployment rate 2, (2.8573)	net migrants 2, (1.1407)	unemployment rate 2, (2.84)	use industrial 2, (4.8452)	number of residents 1, (2.3074)	use openspace 2, (0.76909)	number of tourists 2, (2.6397)
	employees agriculture 1, (2.8447)	motorization rate 1, (1.1084)	employees agriculture 1, (2.8256)	use industrial commercial 2, (0.67582)	number households2, (2.2972)		number of residents 2, (1.0046)
	employees agriculture 2, (2.8447)	use residential mixed 1, (1.0814)	employees agriculture 2, (2.8256)		number of residents 2, (2.2907)		number of residents 1, (1.0011)
	floodhazard1, (2.4728)	use residential mixed 2, (1.0018)	use openspace 2, (1.6024)		employees services 1, (2.2073)		employees services 1, (0.96713)
	floodhazard2, (2.4728)	use industrial commercial 2, (0.6758)	use openspace 1, (1.58)		employees services 2, (2.2073)		employees services 2, (0.96713)
	employees industrial 1, (1.4464)	use industrial commercial 1, (0.60095)	employees industrial 1, (1.4094)		employees industrial 1, (1.7075)		number households2, (0.9253)
	employees industrial 2, (1.4464)	average income 1, (0.51454)	employees industrial 2, (1.4094)		employees industrial 2, (1.7075)		number households1, (0.90446)
	employees services 1, (1.0327)	average income 2, (0.40719)	motorization rate 2, (0.7883)		use residential mixed 1, (1.0465)		employees industrial 1, (0.88728)
	employees services 2, (1.0327)		average income 1, (0.42623)		average income 1, (1.0261)		employees industrial 2, (0.88728)
	motorization rate 2, (0.7786)				use residential mixed 2, (0.98066)		floodhazard1, (0.79387)
	use agriculture 2, (0.56569)				average income 2, (0.89034)		floodhazard2, (0.79387)
	use industrial commercial 1, (0.56543)				number of tourists 1, (0.8051)		motorization rate 1, (0.61912)
	use industrial commercial 2, (0.5464)				number of tourists 2, (0.6694)		
	use agriculture 1, (0.50632)						
	average income 1, (0.40333)						

Table 7.1.2.22. Lisbon: Negative variables.

Cluster #1	Cluster #2	Cluster #3	Cluster #4	Cluster #5	Cluster #6	Cluster #7	Cluster #8
average income 1, (-0.40303)	use natural 1, (-0.53696)	employees agriculture 1, (-0.39051)	motorization rate 1, (-0.65323)	use agriculture 1, (-0.59282)	use agriculture 2, (-0.68182)	number households1, (-0.42096)	motorization rate 2, (-0.77442)
number households1, (-0.46513)	use natural 2, (-0.56159)	employees agriculture 2, (-0.39051)		use agriculture 2, (-0.82299)	use agriculture 1, (-0.68348)	employees services 1, (-0.43864)	
employees services 1, (-0.47002)	motorization rate 1, (-0.67257)	use agriculture 1, (-0.56121)			motorization rate 2, (-1.0263)	employees services 2, (-0.43864)	
employees services 2, (-0.47002)		use agriculture 2, (-0.61061)			net migrants 1, (-1.9744)	number households2, (-0.44034)	
number households2, (-0.47739)					net migrants 2, (-1.9744)	number of residents 1, (-0.44117)	
number of residents 1, (-0.48058)						number of residents 2, (-0.45886)	
number of residents 2, (-0.49388)						employees industrial 1, (-0.5422)	
employees industrial 1, (-0.58701)						employees industrial 2, (-0.5422)	
employees industrial 2, (-0.58701)						use natural 1, (-0.56229)	
use agriculture 2, (-0.67842)						use natural 2, (-0.5708)	
use agriculture 1, (-0.71029)							

Figure 7.1.2.21. Algarve: Prototypes.

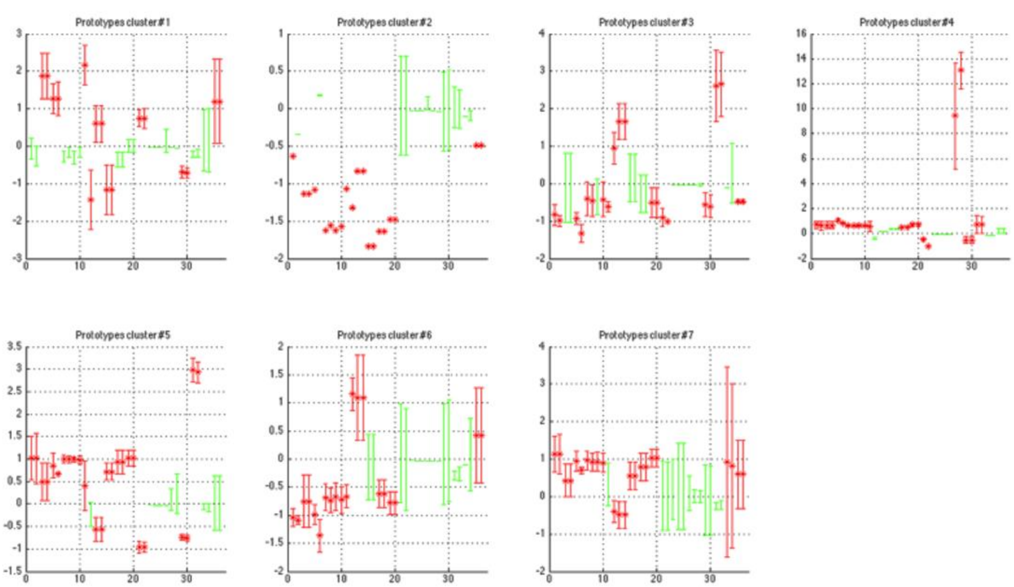


Table 7.1.2.23. Algarve: Positive variables

Cluster #1	Cluster #2	Cluster #3	Cluster #4	Cluster #5	Cluster #6	Cluster #7
number of tourists 1, (2.1769)		use openspace 2, (2.6525)	use industrial commercial 2, (13.0841)	use openspace 1, (2.9817)	number of tourists 2, (1.1543)	average income 1, (1.1367)
net migrants 1, (1.8843)		use openspace 1, (2.6054)	use industrial commercial 1, (9.4296)	use openspace 2, (2.9362)	unemployment rate 1, (1.098)	average income 2, (1.1353)
net migrants 2, (1.8843)		unemployment rate 1, (1.6652)	motorization rate 1, (1.0636)	average income 1, (1.0299)	unemployment rate 2, (1.098)	employees services 1, (1.0208)
motorization rate 2, (1.2771)		unemployment rate 2, (1.6652)	motorization rate 2, (0.82821)	employees services 1, (1.0145)	floodhazard1, (0.42146)	employees services 2, (1.0208)
motorization rate 1, (1.2595)		number of tourists 2, (0.94033)	use openspace 1, (0.75324)	employees services 2, (1.0145)	floodhazard2, (0.42146)	number households1, (0.96148)
floodhazard1, (1.1988)			use openspace 2, (0.72414)	average income 2, (1.0116)		motorization rate 1, (0.95725)
floodhazard2, (1.1988)			average income 1, (0.70505)	number households1, (1.0012)		number households2, (0.92889)
use agriculture 1, (0.75292)			employees services 1, (0.70069)	number households2, (0.9917)		number of residents 1, (0.92703)
use agriculture 2, (0.73971)			employees services 2, (0.70069)	number of residents 1, (0.99027)		use residential mixed 1, (0.91953)
unemployment rate 1, (0.59753)			net migrants 1, (0.68736)	number of residents 2, (0.98118)		number of residents 2, (0.9049)
unemployment rate 2, (0.59753)			net migrants 2, (0.68736)	employees industrial 1, (0.92738)		use residential mixed 2, (0.81199)
			number households1, (0.64957)	employees industrial 2, (0.92738)		employees industrial 1, (0.78085)
			number of residents 1, (0.63253)	motorization rate 1, (0.85558)		employees industrial 2, (0.78085)
			average income 2, (0.62295)	employees agriculture 1, (0.71739)		motorization rate 2, (0.70763)
			number households2, (0.62176)	employees agriculture 2, (0.71739)		floodhazard1, (0.59093)
			number of residents 2, (0.61862)	motorization rate 2, (0.67656)		floodhazard2, (0.59093)
			number of tourists 1, (0.58062)	net migrants 1, (0.49008)		employees agriculture 1, (0.54758)
			employees industrial 1, (0.52928)	net migrants 2, (0.49008)		employees agriculture 2, (0.54758)
			employees industrial 2, (0.52928)	number of tourists 1, (0.40897)		net migrants 1, (0.42734)
						net migrants 2, (0.42734)

Table 7.1.2.24. Algarve: Negative variables.

Cluster #1	Cluster #2	Cluster #3	Cluster #4	Cluster #5	Cluster #6	Cluster #7
use natural 1, (-0.68112)	floodhazard1, (-0.49324)	number households1, (-0.39357)	use agriculture 1, (-0.49109)	unemployment rate 1, (-0.55584)	employees industrial 1, (-0.61751)	number of tourists 2, (-0.40394)
use natural 2, (-0.70376)	floodhazard2, (-0.49324)	number of residents 2, (-0.416)	use natural 1, (-0.49821)	unemployment rate 2, (-0.55584)	employees industrial 2, (-0.61751)	unemployment rate 1, (-0.48481)
employees agriculture 1, (-1.1639)	average income 1, (-0.62949)	number households2, (-0.45256)	use natural 2, (-0.50735)	use natural 1, (-0.7441)	number of tourists 1, (-0.66364)	unemployment rate 2, (-0.48481)
employees agriculture 2, (-1.1639)	unemployment rate 1, (-0.83625)	floodhazard1, (-0.4746)	use agriculture 2, (-1.0085)	use natural 2, (-0.75327)	number of residents 1, (-0.6698)	
number of tourists 2, (-1.4336)	unemployment rate 2, (-0.83625)	floodhazard2, (-0.4746)		use agriculture 1, (-0.95464)	number households1, (-0.67961)	
	number of tourists 1, (-1.0719)	employees services 1, (-0.51454)		use agriculture 2, (-0.95637)	number of residents 2, (-0.72602)	
	motorization rate 1, (-1.0796)	employees services 2, (-0.51454)			number households2, (-0.73371)	
	net migrants 1, (-1.1286)	use natural 1, (-0.55064)			net migrants 1, (-0.74954)	
	net migrants 2, (-1.1286)	number of tourists 1, (-0.60351)			net migrants 2, (-0.74954)	
	number of tourists 2, (-1.3246)	use natural 2, (-0.60589)			employees services 1, (-0.78375)	
	employees services 1, (-1.482)	average income 1, (-0.82587)			employees services 2, (-0.78375)	
	employees services 2, (-1.482)	use agriculture 1, (-0.91312)			motorization rate 1, (-0.98923)	
	number households2, (-1.5566)	motorization rate 1, (-0.93164)			average income 1, (-1.0404)	
	number of residents 2, (-1.5664)	average income 2, (-0.98861)			average income 2, (-1.0838)	
	number households1, (-1.6201)	use agriculture 2, (-1.0051)			motorization rate 2, (-1.3652)	
	number of residents 1, (-1.6277)	motorization rate 2, (-1.3254)				
	employees industrial 1, (-1.6362)					
	employees industrial 2, (-1.6362)					
	employees agriculture 1, (-1.8361)					
	employees agriculture 2, (-1.8361)					

Figure 7.1.3.3. SOM's SOM Cluster 1.

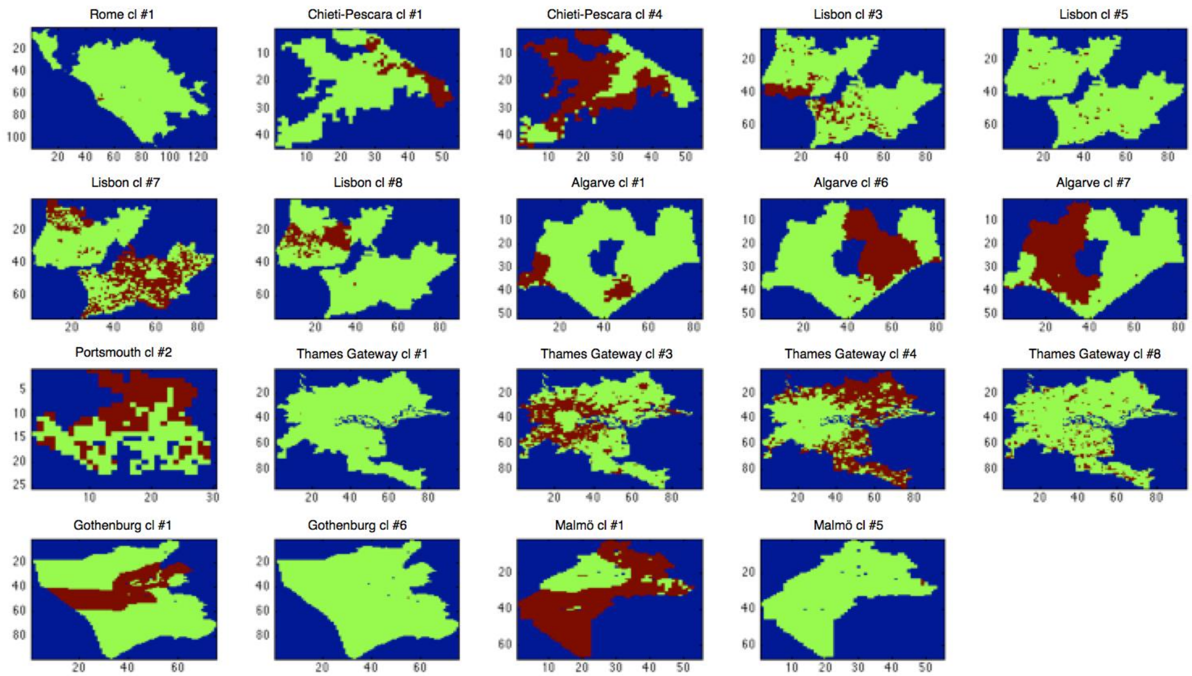


Figure 7.1.3.4. SOM's SOM Cluster 2.

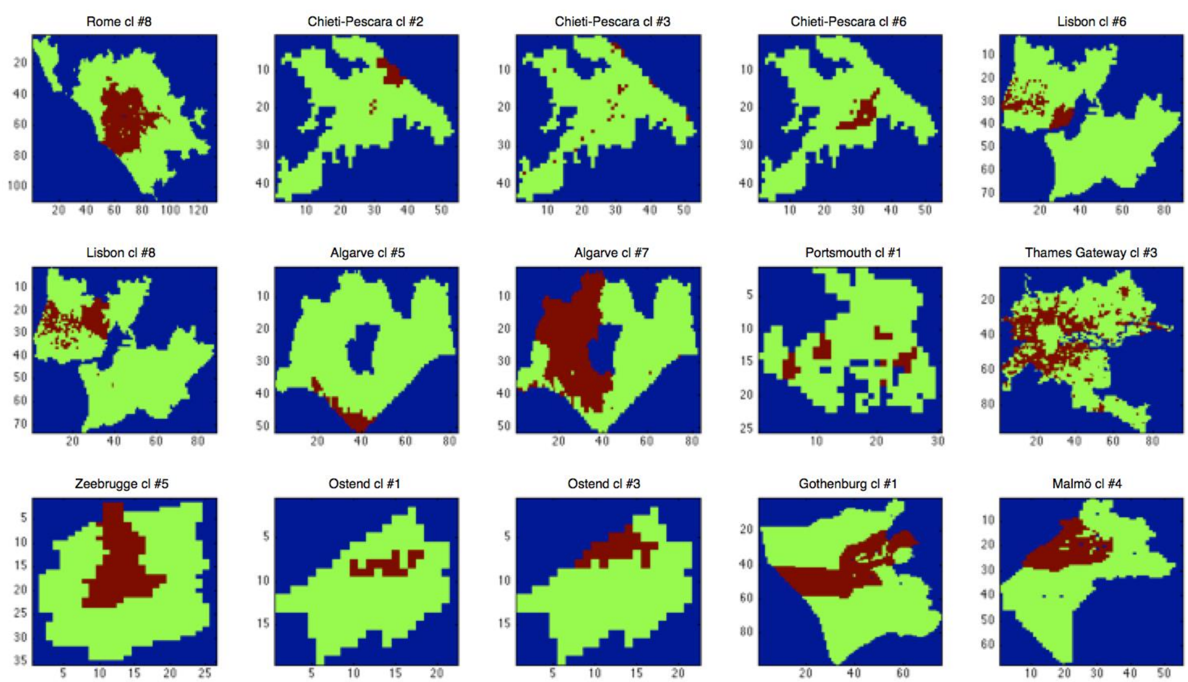


Figure 7.1.3.5. SOM's SOM Cluster 3.

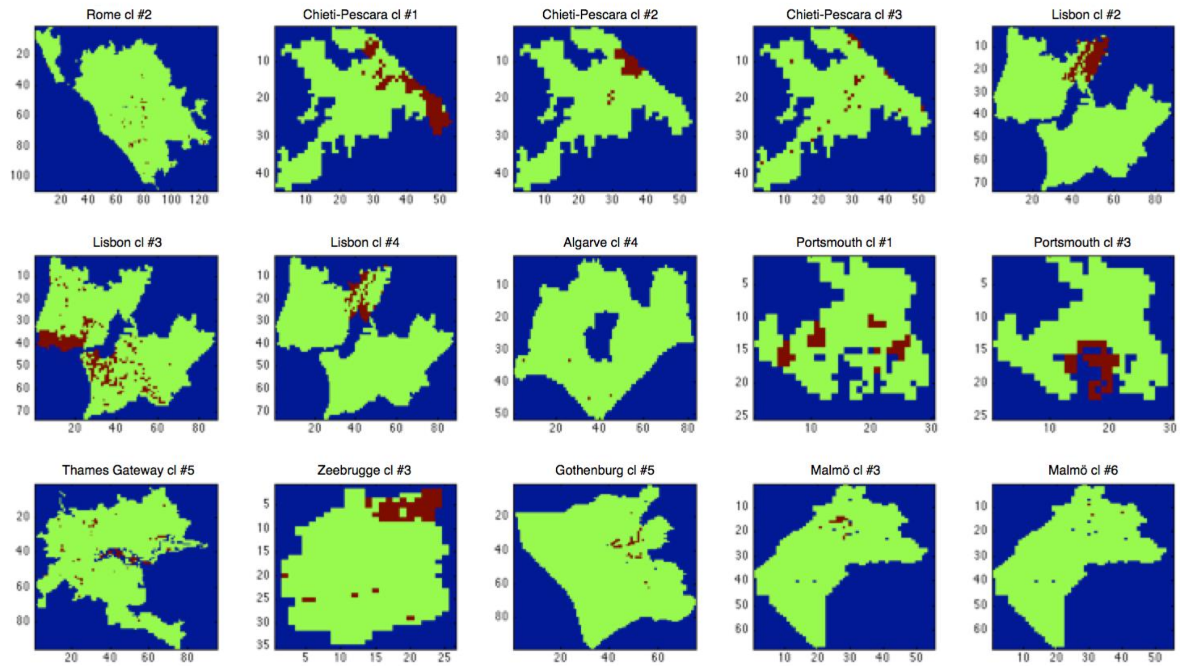


Figure 7.1.3.6. SOM's SOM Cluster 4.

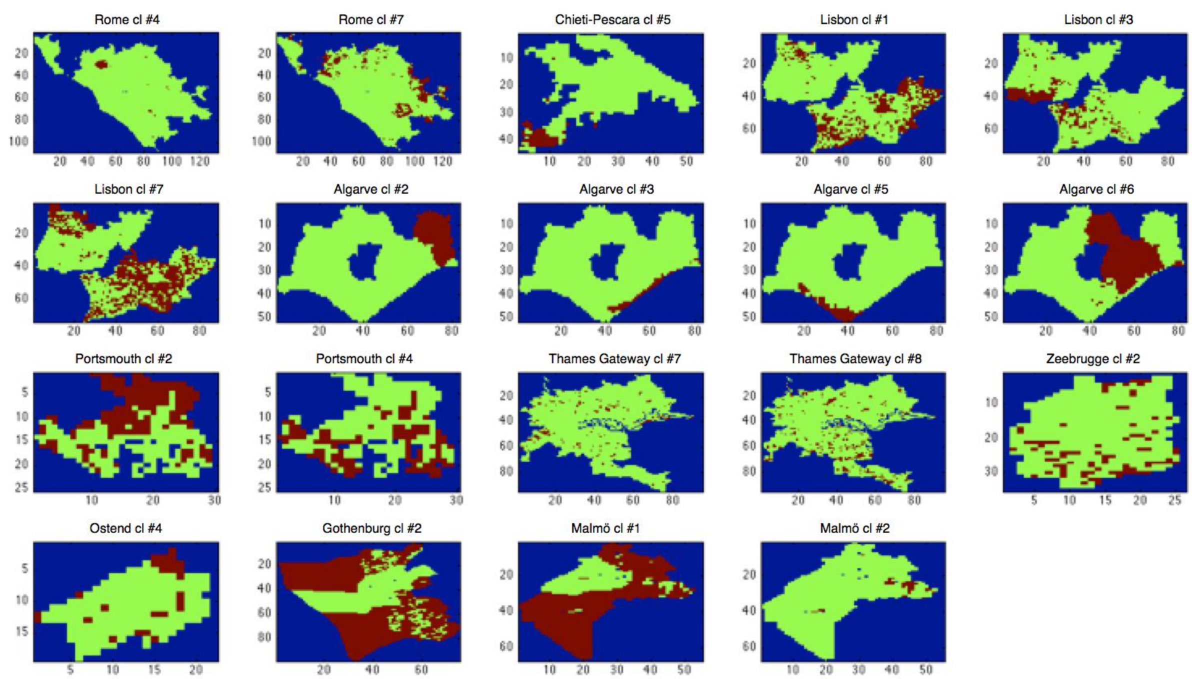


Figure 7.1.3.7. SOM's SOM Cluster 5.

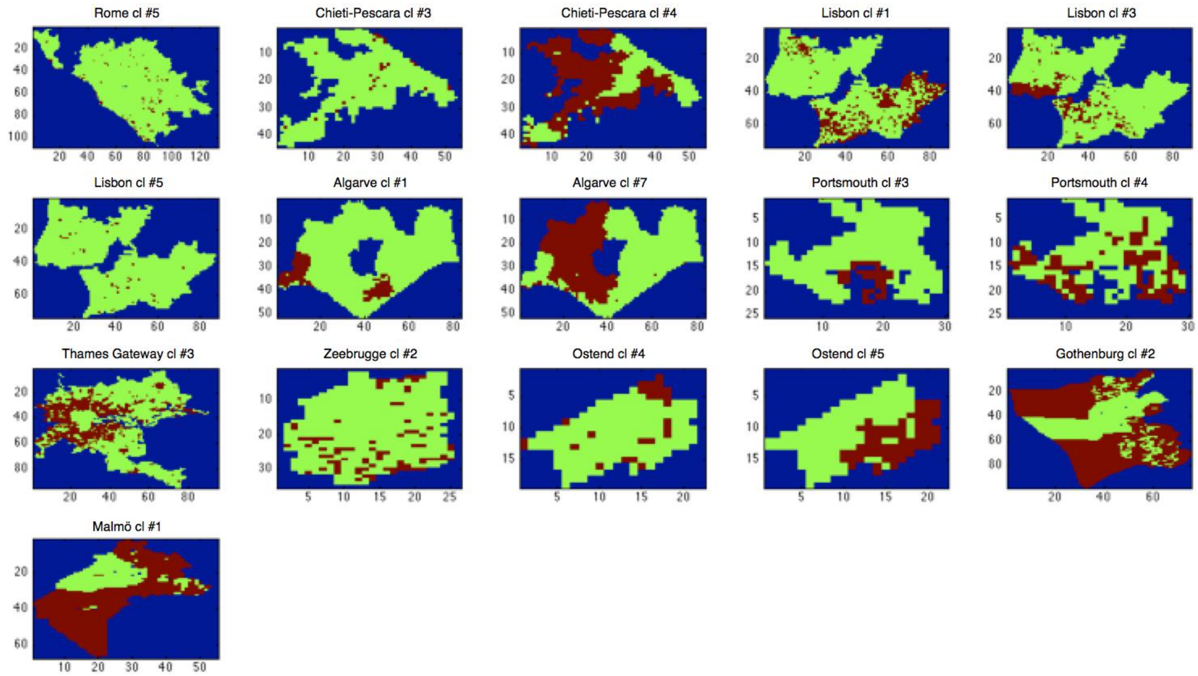


Figure 7.1.3.8. SOM's SOM Cluster 6.

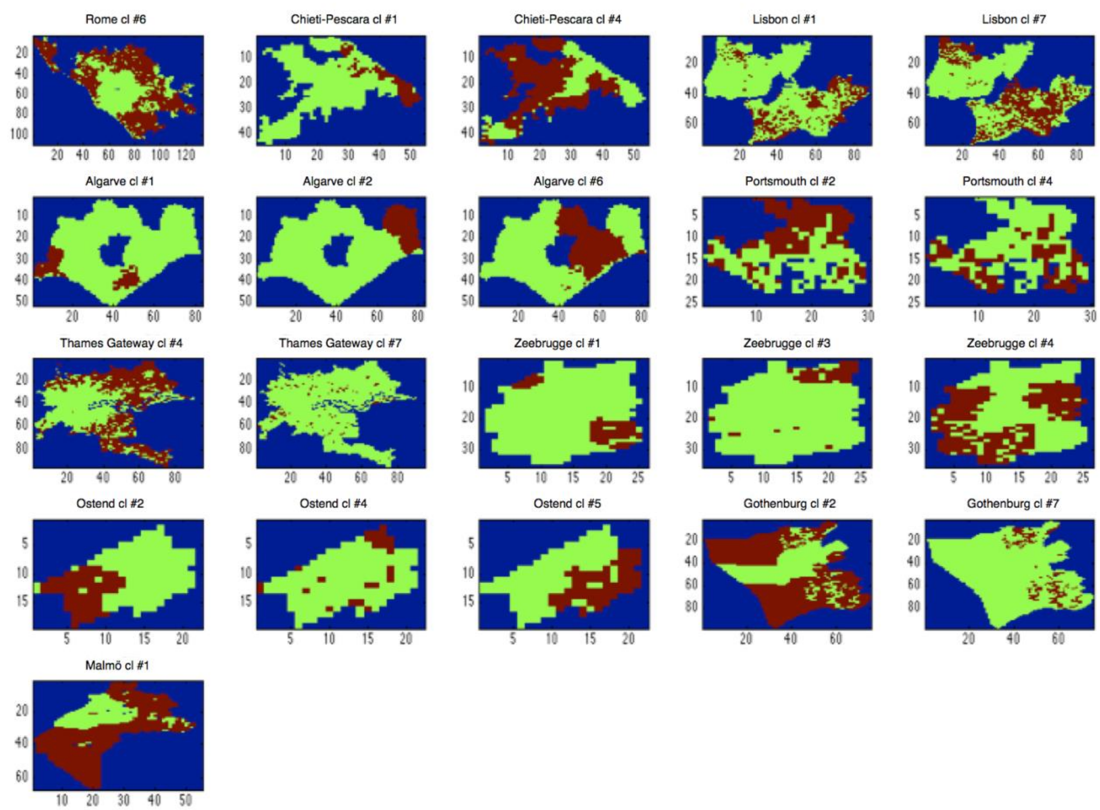


Figure 7.1.3.9. SOM's SOM Cluster 7.

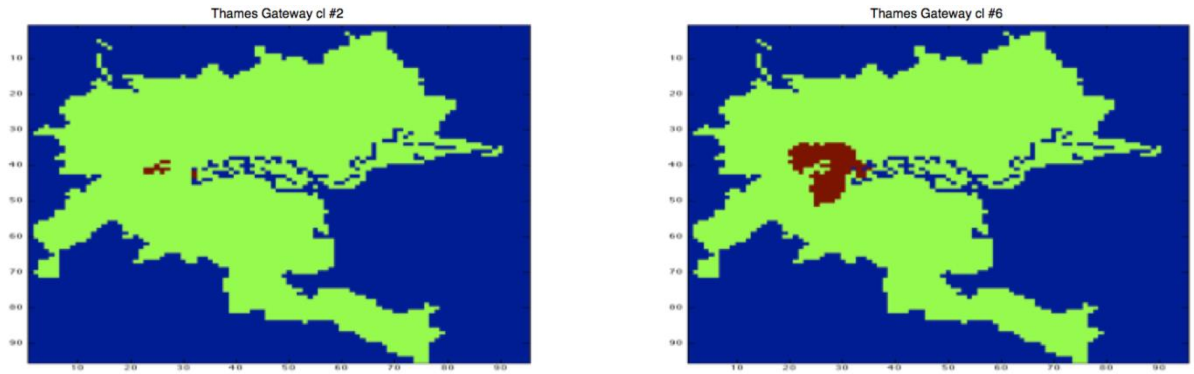


Figure 7.1.3.10. SOM's SOM Cluster 8.

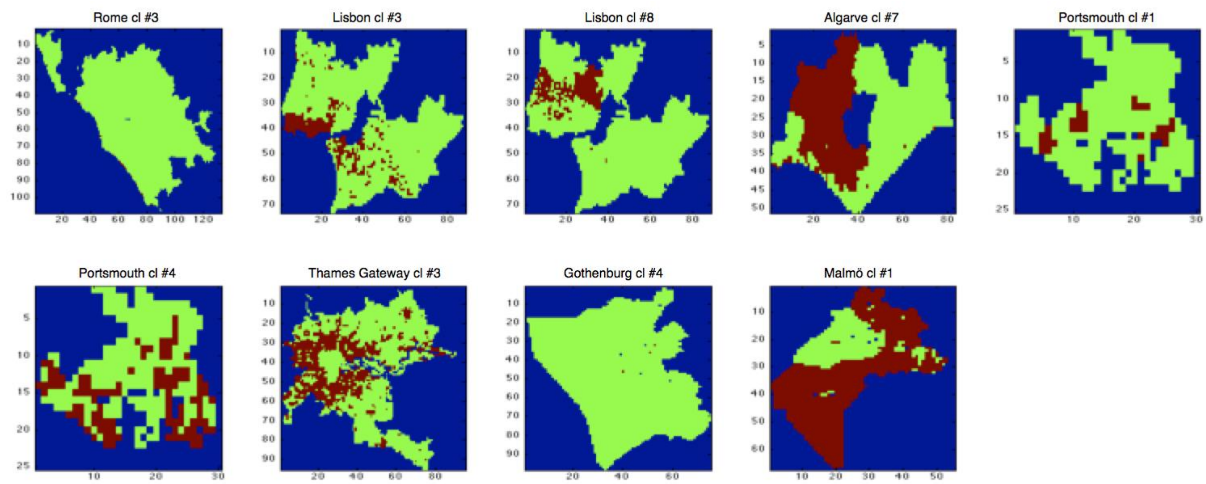


Figure 7.1.3.11. SOM's SOM prototypes.

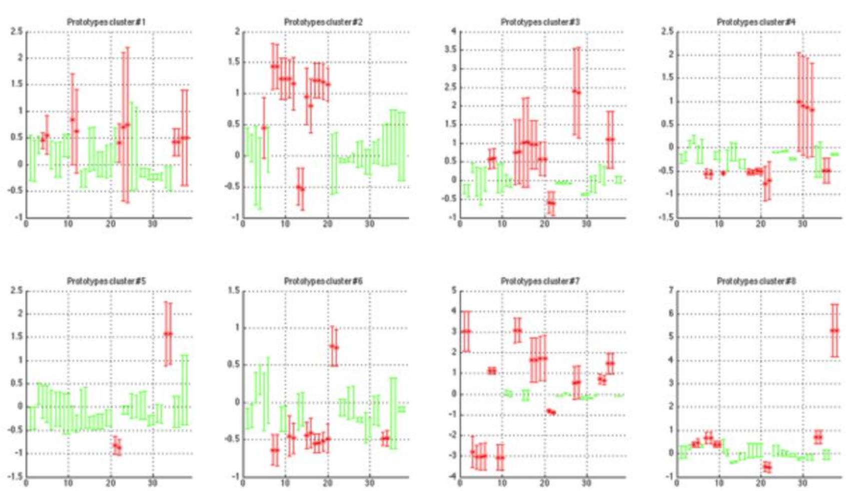


Table 7.1.3.2. SOM's SOM: positive variables.

Cluster #1	Cluster #2	Cluster #3	Cluster #4	Cluster #5	Cluster #6	Cluster #7	Cluster #8
number of tourists 1, (0.84952)	number households 1, (1.4405)	use industrial commercial 1, (2.3932)	use natural 1, (0.99549)	use residential mixed 1, (1.5715)	use agriculture 1, (0.75646)	unemployment rate 1, (3.0962)	slr 1, (5.3021)
use airports 2, (0.74798)	number households 2, (1.4392)	use industrial commercial 2, (2.3563)	use natural 2, (0.9092)	use residential mixed 2, (1.5694)	use agriculture 2, (0.73468)	unemployment rate 2, (3.0815)	slr 2, (5.3021)
use airports 1, (0.7105)	number of tourists 1, (1.2411)	floodhazard 1, (1.0946)	use openspace 1, (0.86253)			average income 1, (3.0529)	use residential mixed 2, (0.72331)
number of tourists 2, (0.62585)	number of residents 1, (1.2393)	floodhazard 2, (1.0946)	use openspace 2, (0.81695)			average income 2, (3.0299)	use residential mixed 1, (0.72172)
motorization rate 1, (0.55685)	number of residents 2, (1.2387)	employees agriculture 2, (1.0295)				employees services 2, (1.7394)	number households 2, (0.6829)
slr 1, (0.49993)	employees industrial 1, (1.2175)	employees agriculture 1, (1.0072)				employees services 1, (1.736)	number households 1, (0.67863)
slr 2, (0.49993)	employees industrial 2, (1.2127)	employees industrial 2, (0.96825)				employees industrial 2, (1.6398)	motorization rate 1, (0.46741)
net migrants 2, (0.45181)	employees services 1, (1.1909)	employees industrial 1, (0.96142)				employees industrial 1, (1.6394)	number of residents 2, (0.39751)
floodhazard 1, (0.42489)	number of tourists 2, (1.1613)	unemployment rate 2, (0.77371)				floodhazard 1, (1.487)	number of residents 1, (0.39499)
floodhazard 2, (0.42489)	employees services 2, (1.149)	unemployment rate 1, (0.76056)				floodhazard 2, (1.487)	net migrants 2, (0.39174)
use agriculture 2, (0.40552)	employees agriculture 1, (0.95071)	number households 2, (0.59249)				number households 1, (1.1337)	
	employees agriculture 2, (0.79982)	employees services 2, (0.58556)				number households 2, (1.1313)	
	motorization rate 1, (0.4463)	employees services 1, (0.58395)				use residential mixed 1, (0.73816)	
		number households 1, (0.58199)				use residential mixed 2, (0.67337)	
						use industrial commercial 2, (0.56893)	
						use industrial commercial 1, (0.52986)	

Table 7.1.3.3. SOM's SOM: negative variables.

Cluster #1	Cluster #2	Cluster #3	Cluster #4	Cluster #5	Cluster #6	Cluster #7	Cluster #8
	unemployment rate 1, (-0.5022)	use agriculture 1, (-0.58478)	floodhazard 1, (-0.48526)	use agriculture 1, (-0.82334)	employees agriculture 2, (-0.40825)	use agriculture 1, (-0.8279)	use agriculture 1, (-0.55296)
	unemployment rate 2, (-0.53964)	use agriculture 2, (-0.61768)	floodhazard 2, (-0.48526)	use agriculture 2, (-0.86852)	employees agriculture 1, (-0.44807)	use agriculture 2, (-0.87746)	use agriculture 2, (-0.59755)
			employees services 1, (-0.49464)		number of tourists 1, (-0.45027)	net migrants 1, (-2.7964)	
			employees services 2, (-0.50856)		number of tourists 2, (-0.47827)	motorization rate 2, (-3.0087)	
			employees industrial 1, (-0.51338)		use residential mixed 2, (-0.48153)	motorization rate 1, (-3.0432)	
			employees industrial 2, (-0.51513)		employees services 2, (-0.48317)	net migrants 2, (-3.0468)	
			number of tourists 1, (-0.54607)		use residential mixed 1, (-0.48913)	number of residents 2, (-3.056)	
			number households 1, (-0.5489)		employees services 1, (-0.52612)	number of residents 1, (-3.0567)	
			number households 2, (-0.56281)		employees industrial 2, (-0.54762)		
			use agriculture 2, (-0.70177)		employees industrial 1, (-0.54896)		
			use agriculture 1, (-0.77145)		number households 2 (-0.63723)		
					number households 1, (-0.63733)		

Figure 7.1.4.1. Cluster distribution on SOMs for Asian territories + Funchal.

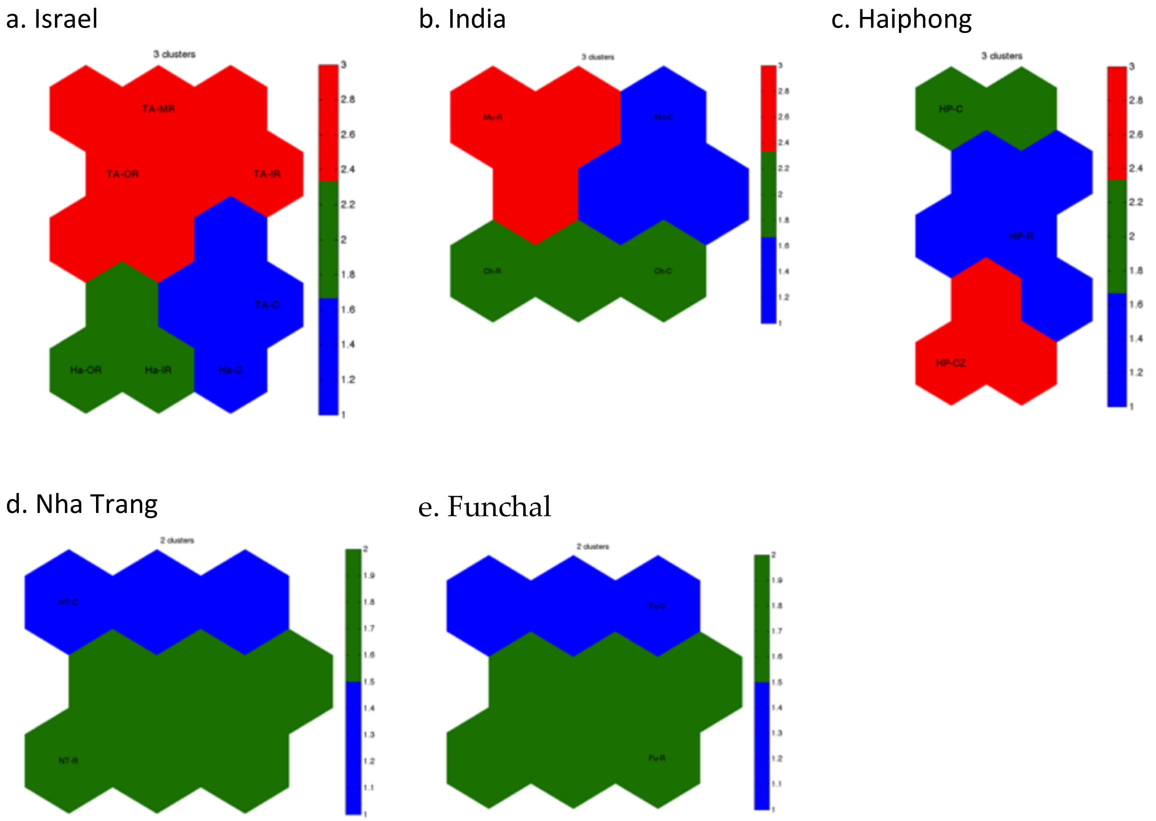


Figure 7.1.4.2. Israel: Prototypes.

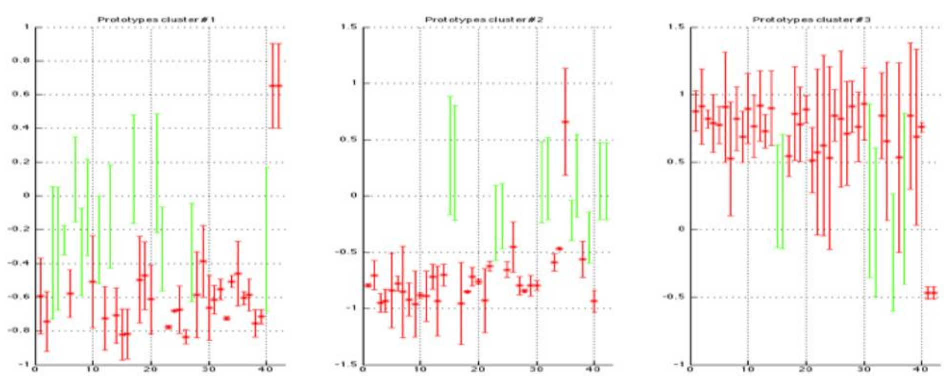


Table 7.1.4.3. Israel: Positive variables.

Cluster #1 Tel Aviv Core, Haifa Core	**Cluster #2** Haifa Inner Ring, Haifa Outer Ring	**Cluster #3** Tel Aviv Inner Ring, Tel Aviv Middle Ring, Tel Aviv Outer Ring
SLR 1, (0.65278)	Natural Area 1, (0.65811)	Employees Public Services 2, (0.93237)
SLR 2, (0.65278)		Residents Middle Social 2, (0.91556)
		Population 2, (0.91269)
		Employees Private Services 2, (0.91171)
		Changed Address 0-17 2, (0.90642)
		Residents Lower Social 2, (0.89875)
		Residents High Social 2, (0.89582)
		Commuting Out 2, (0.89017)
		Population 1, (0.87768)
		Commuting In 2, (0.85959)
		Employees Industrial 1, (0.84608)
		Mixed Industrial-Commercial Area 1, (0.84527)
		Residential Area 1, (0.84213)
		Changed Address >=18 2, (0.82121)
		In-migrants 2, (0.82003)
		Employees Industrial 2, (0.81904)
		Out-migrants 2, (0.78729)
		Commuting Out 1, (0.77792)
		Changed Address 0-17 1, (0.77431)
		Residents Middle Social 1, (0.76512)
		Employees Public Services 1, (0.762)
		Extr. Rainfall 2, (0.75931)
		Residents Lower Social 1, (0.72939)
		Employees Private Services 1, (0.71149)
		Residents High Social 1, (0.68768)
		Road Net Length 1, (0.68571)
		Government Area 1, (0.65616)
		Employees Agriculture 1, (0.62246)
		Peak Passengers Out 1, (0.57377)
		Commuting In 1, (0.54392)
		Urban Open Space Area 1, (0.53326)
		Employees Agriculture 2, (0.53093)
		Changed Address >=18 1, (0.52427)
		Peak Passengers In 1, (0.51311)

Table 7.1.4.4. Israel: Negative variables.

Cluster #1 Tel Aviv Core, Haifa Core	**Cluster #2** Haifa Inner Ring, Haifa Outer Ring	**Cluster #3** Tel Aviv Inner Ring, Tel Aviv Middle Ring, Tel Aviv Outer Ring
Employees Public Services 1, (-0.38835)	Employees Industrial 2, (-0.45561)	SLR 1, (-0.46883)
Natural Area 1, (-0.46116)	Government Area 1, (-0.4671)	SLR 2, (-0.46883)
Commuting Out 1, (-0.47309)	Residential Area 1, (-0.56481)	
Commuting In 2, (-0.49635)	Mixed Industrial-Commercial Area 1, (-0.59221)	
Government Area 1, (-0.50648)	Peak Passengers Out 1, (-0.62541)	
Residents High Social 2, (-0.50875)	Employees Industrial 1, (-0.65484)	
Agriculture Area 1, (-0.55425)	Residents Lower Social 2, (-0.70304)	
Changed Address 0-17 2, (-0.5784)	Population 2, (-0.70457)	
Employees Private Services 2, (-0.5847)	Commuting Out 1, (-0.71947)	
Open Space Area 1, (-0.58606)	Residents Middle Social 2, (-0.71964)	
Population 1, (-0.59278)	Commuting Out 2, (-0.76067)	
Urban Open Space Area 1, (-0.60515)	Changed Address 0-17 2, (-0.7792)	
Commuting Out 2, (-0.61226)	Employees Public Services 2, (-0.79547)	
Total Area 1, (-0.61653)	Population 1, (-0.79582)	
Employees Public Services 2, (-0.66304)	Employees Public Services 1, (-0.79695)	
Employees Industrial 1, (-0.67455)	Employees Private Services 1, (-0.79718)	
Employees Agriculture 2, (-0.68009)	Changed Address 0-17 1, (-0.83981)	
Residents Lower Social 2, (-0.70845)	Employees Private Services 2, (-0.84258)	
Road Net Length 1, (-0.71568)	Commuting In 2, (-0.85275)	
Residents Middle Social 2, (-0.72493)	Changed Address >=18 1, (-0.85341)	
Mixed Industrial-Commercial Area 1, (-0.72607)	Residents High Social 2, (-0.88549)	
Population 2, (-0.74326)	Residents Middle Social 1, (-0.8905)	
Residential Area 1, (-0.75514)	Changed Address >=18 2, (-0.92021)	
Employees Agriculture 1, (-0.77742)	Peak Passengers In 1, (-0.93006)	
Household size 2, (-0.81609)	Out-migrants 2, (-0.93165)	
Household size 1, (-0.82233)	Residents Lower Social 1, (-0.93248)	
Employees Industrial 2, (-0.83462)	Extr. Rainfall 2, (-0.93552)	
	In-migrants 2, (-0.9521)	
	Commuting In 1, (-0.95626)	
	Residents High Social 1, (-0.96153)	

Figure 7.1.4.3. India: Prototypes.

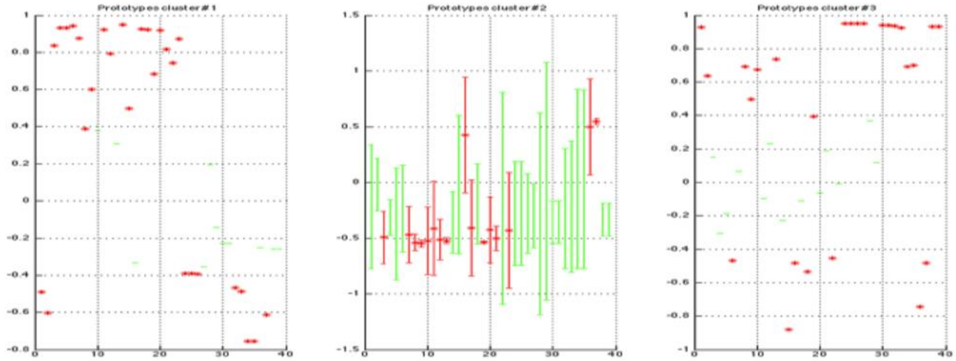


Table 7.1.4.5. India: Positive variables.

Cluster #1 Mumbai Core	**Cluster #2** Chennai Core, Chennai Ring	**Cluster #3** Mumbai Ring
Temporary Dwellings 2, (0.94754)	Sustainability Index 2, (0.54681)	Agriculture Area 1, (0.95148)
Employees in Services 2, (0.94122)	Sustainability Index 1, (0.49821)	Total Area 1, (0.95076)
Employees in Industrial 2, (0.93341)	Vehicles 2, (0.4256)	Total Area 2, (0.95076)
Employees in Services 1, (0.93232)		Agriculture Area 2, (0.94983)
Population 1, (0.92448)		Natural Area 1, (0.93859)
Population 2, (0.92243)		Natural Area 2, (0.9385)
Slum Population 1, (0.92106)		Open Space Area 1, (0.93527)
International In migrants 2, (0.91902)		SLR 1, (0.93195)
Net Migration 1, (0.87505)		SLR 2, (0.93195)
Household Size 2, (0.87251)		Employees in Agriculture 1, (0.92864)
Employees in Industrial 1, (0.83662)		Open Space Area 2, (0.92432)
National In migrants 2, (0.815)		Temporary Dwellings 1, (0.73706)
Slum Population 2, (0.79335)		Road Net Length 2, (0.6995)
Household Size 1, (0.74146)		Road Net Length 1, (0.69411)
In migrants 2, (0.6839)		Net Migration 2, (0.69169)
Rainfall 1, (0.5997)		Rainfall 2, (0.67576)
Vehicles 1, (0.49699)		Employees in Agriculture 2, (0.63848)
Net Migration 2, (0.38718)		Rainfall 1, (0.49584)
		In migrants 2, (0.39341)

Table 7.1.4.6. India: Negative variables.

Cluster #1 Mumbai Core	**Cluster #2** Chennai Core, Chennai Ring	**Cluster #3** Mumbai Ring
Total Area 1, (-0.39132)	Population 1, (-0.40669)	Household Size 1, (-0.45311)
Total Area 2, (-0.39132)	Slum Population 1, (-0.41256)	Employees in Services 2, (-0.46955)
Agriculture Area 1, (-0.39257)	International In migrants 2, (-0.42771)	Sustainability Index 2, (-0.48112)
Open Space Area 1, (-0.46677)	Household Size 2, (-0.43253)	Vehicles 2, (-0.48224)
Open Space Area 2, (-0.48518)	Net Migration 1, (-0.4706)	Population 2, (-0.53535)
Employees in Agriculture 1, (-0.49001)	Employees in Industrial 1, (-0.49412)	Sustainability Index 1, (-0.74475)
Employees in Agriculture 2, (-0.60286)	National In migrants 2, (-0.50225)	Vehicles 1, (-0.87972)
Sustainability Index 2, (-0.6125)	Slum Population 2, (-0.51352)	
Road Net Length 2, (-0.75384)	Temporary Dwellings 1, (-0.52216)	
Road Net Length 1, (-0.75676)	Rainfall 2, (-0.52698)	
	In migrants 2, (-0.53866)	
	Net Migration 2, (-0.53943)	
	Rainfall 1, (-0.54777)	

Figure 7.1.4.4. Haiphong: Prototypes.

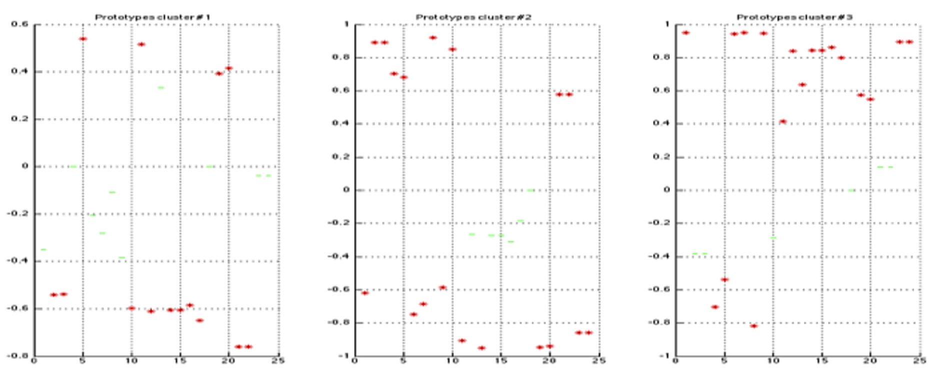


Table 7.1.4.7. Haiphong: Positive variables.

Cluster #1 Haiphong Ring	**Cluster #2** Haiphong Core	**Cluster #3** Haiphong Coastal Zone
In Migrants from Core 2, (0.53853)	Dwellings 1, (0.91943)	Housing Area 2, (0.95151)
Total Area 2, (0.51653)	National In Migrants 1, (0.89275)	Population 1, (0.95028)
Sustainability 2, (0.41575)	International In Migrants 1, (0.89207)	Temporary Dwellings 1, (0.94654)
Sustainability 1, (0.39287)	Households 1, (0.85211)	Land Area 2, (0.94347)
	In Migrants from Ring 2, (0.70361)	Inondation 1, (0.89327)
	In Migrants from Core 2, (0.68167)	Inondation 2, (0.89327)
	Tourists 1, (0.57993)	Residential Area 1, (0.86222)
	Tourists 2, (0.57993)	Industrial-Commercial Area 1, (0.84413)
		Industrial-Commercial Area 2, (0.84413)
		Agriculture Area 1, (0.84043)
		Residential Area 2, (0.7997)
		Agriculture Area 2, (0.63677)
		Sustainability 1, (0.57358)
		Sustainability 2, (0.54745)
		Total Area 2, (0.41654)

Table 7.1.4.8. Haiphong: Negative variables.

Cluster #1 Haiphong Ring	**Cluster #2** Haiphong Core	**Cluster #3** Haiphong Coastal Zone
National In Migrants 1, (-0.5388)	Temporary Dwellings 1, (-0.58482)	In Migrants from Core 2, (-0.53853)
International In Migrants 1, (-0.53998)	Population 1, (-0.61887)	In Migrants from Ring 2, (-0.70361)
Residential Area 1, (-0.5844)	Housing Area 2, (-0.68602)	Dwellings 1, (-0.81731)
Households 1, (-0.59704)	Land Area 2, (-0.74832)	
Industrial-Commercial Area 1, (-0.60636)	Inundation 1, (-0.85817)	
Industrial-Commercial Area 2, (-0.60636)	Inundation 2, (-0.85817)	
Agriculture Area 1, (-0.61051)	Total Area 2, (-0.90475)	
Residential Area 2, (-0.65005)	Sustainability 2, (-0.9404)	
Tourists 1, (-0.76048)	Sustainability 1, (-0.94491)	
Tourists 2, (-0.76048)	Agriculture Area 2, (-0.95145)	

Figure 7.1.4.5. Nha Trang: Prototypes.

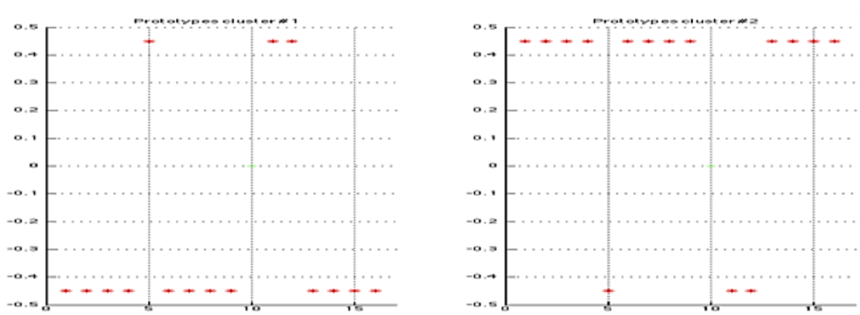


Table 7.1.4.9. Nha Trang: Positive variables.

Cluster #1 Nha Trang Core	**Cluster #2** Nha Trang Ring
Agriculture Area 1, (0.44912)	Households 2, (0.44912)
Agriculture Area 2, (0.44912)	Population 1, (0.44912)
Land Area 1, (0.44912)	Households 1, (0.44912)
	Tourists 1, (0.44912)
	Population 2, (0.44912)
	In Migrants 2, (0.44912)
	Out Migrants 2, (0.44912)
	Land Area 2, (0.44912)
	Total Area 2, (0.44912)
	Tourists 2, (0.44912)
	Inundation 1, (0.44912)
	Inundation 2, (0.44912)

Table 7.1.4.10. Nha Trang: Negative variables.

Cluster #1	**Cluster #2**
Households 2, (-0.44912)	Land Area 1, (-0.44912)
Population 2, (-0.44912)	Agriculture Area 1, (-0.44912)
In Migrants 2, (-0.44912)	Agriculture Area 2, (-0.44912)
Out Migrants 2, (-0.44912)	
Land Area 2, (-0.44912)	
Total Area 2, (-0.44912)	
Tourists 1, (-0.44912)	
Inundation 1, (-0.44912)	
Inundation 2, (-0.44912)	
Population 1, (-0.44912)	
Households 1, (-0.44912)	
Tourists 2, (-0.44912)	

Figure 7.1.4.6. Funchal: Prototypes.

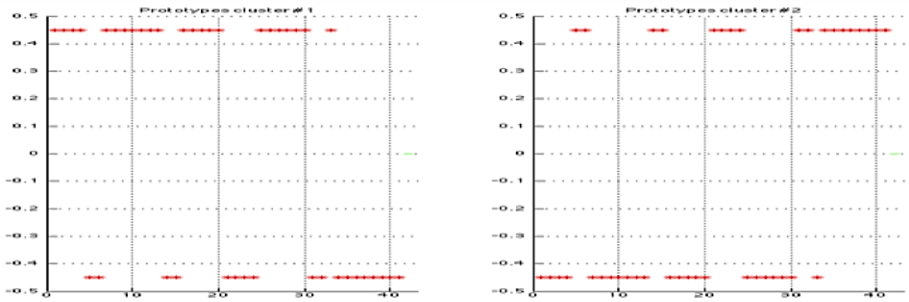


Table 7.1.4.11. Funchal: Positive variables.

Cluster #1 Funchal Core	**Cluster #2** Funchal Ring
Population 1, (0.44912)	Out migrants 1, (0.44912)
Population 2, (0.44912)	Lower Social Residents 2, (0.44912)
In migrants 1, (0.44912)	Second Homes 1, (0.44912)
High Social Residents 2, (0.44912)	Land Area 2, (0.44912)
Middle Social Residents 1, (0.44912)	Housing Area 2, (0.44912)
Lower Social Residents 1, (0.44912)	Agriculture Employees 2, (0.44912)
Second Homes 2, (0.44912)	Motorization rate 1, (0.44912)
Temporary Residents 1, (0.44912)	Households 2, (0.44912)
Temporary Residents 2, (0.44912)	Out migrants 2, (0.44912)
Tourists 2, (0.44912)	Undeveloped Land Area 2, (0.44912)
Vehicles 2, (0.44912)	Economic Area 2, (0.44912)
Temporary Dwellings 2, (0.44912)	Household Size 2, (0.44912)
In migrants 2, (0.44912)	Commuting Out 2, (0.44912)
International In migrants 1, (0.44912)	Industrial Employees 2, (0.44912)
International In migrants 2, (0.44912)	Households 1, (0.44912)
High Social Residents 1, (0.44912)	Household Size 1, (0.44912)
Middle Social Residents 2, (0.44912)	Rate of Unemployment 2, (0.44912)
Dwellings 1, (0.44912)	Motorization rate 2, (0.44912)
Dwellings 2, (0.44912)	
Apartments 1, (0.44912)	
Apartments 2, (0.44912)	
Temporary Dwellings 1, (0.44912)	
Commuting In 2, (0.44912)	

Table 7.1.4.12. Funchal: Negative variables.

Cluster #1	Cluster #2
Out migrants 2, (-0.44912)	In migrants 2, (-0.44912)
Economic Area 2, (-0.44912)	International In migrants 1, (-0.44912)
Household Size 2, (-0.44912)	International In migrants 2, (-0.44912)
Commuting Out 2, (-0.44912)	High Social Residents 1, (-0.44912)
Industrial Employees 2, (-0.44912)	Middle Social Residents 2, (-0.44912)
Households 1, (-0.44912)	Dwellings 1, (-0.44912)
Out migrants 1, (-0.44912)	Dwellings 2, (-0.44912)
Lower Social Residents 2, (-0.44912)	Apartments 1, (-0.44912)
Second Homes 1, (-0.44912)	Apartments 2, (-0.44912)
Land Area 2, (-0.44912)	Temporary Dwellings 1, (-0.44912)
Undeveloped Land Area 2, (-0.44912)	Commuting In 2, (-0.44912)
Housing Area 2, (-0.44912)	Population 1, (-0.44912)
Agriculture Employees 2, (-0.44912)	Population 2, (-0.44912)
Motorization rate 1, (-0.44912)	In migrants 1, (-0.44912)
Households 2, (-0.44912)	High Social Residents 2, (-0.44912)
Household Size 1, (-0.44912)	Middle Social Residents 1, (-0.44912)
Rate of Unemployment 2, (-0.44912)	Lower Social Residents 1, (-0.44912)
Motorization rate 2, (-0.44912)	Second Homes 2, (-0.44912)
	Temporary Residents 1, (-0.44912)
	Temporary Residents 2, (-0.44912)
	Tourists 2, (-0.44912)
	Vehicles 2, (-0.44912)
	Temporary Dwellings 2, (-0.44912)

Table 7.1.5.6. Conflicts and SOM's SOM clustering integration.

Country	Case	Conflict Name	Label		Conflict cluster		
ITA	Rome	Civitavecchia	Rom1		2		
Cluster Som's Som							
1	2	3	4	5	6	7	8
		Rome cl 2 (0,56%)	Rome cl 4 (0,28%)	Rome cl 5 (3.35%)	Rome cl 6 (84,64%)		Rome cl 3 (0,84%)
			Rome cl 7 (10,34%)				
		0,56%	10,62%	3,35%	84,64%		0,84%

Country	Case	Conflict Name	Label		Conflict cluster		
ITA	Rome	Ostia water-use and management	Rom2		3		
Cluster Som's Som							
1	2	3	4	5	6	7	8
	Rome cl 8 (97,37%)						Rome cl 3 (2,63%)
	97,37%						2,63%

Country	Case	Conflict Name	Label		Conflict cluster		
ITA	Chieti-Pescara	The Costa Teatina National Park	Pes		1		
Cluster Som's Som							
1	2	3	4	5	6	7	8
Pescara cl 1 (95,08%)	Pescara cl 3 (3,28%)	Pescara cl 1 (95,08%)		Pescara cl 3 (3,28%)	Pescara cl 1 (95,08%)		
Pescara cl 4 (1,64%)		Pescara cl 3 (3,28%)		Pescara cl 4 (1,64%)	Pescara cl 4 (1,64%)		
96,72%	3,28%	98,36%		4,92%	96,72%		

Country	Case	Conflict Name	Label		Conflict cluster		
PRT	Lisbon	Trafaria and Costa da Caparica	Lis		1		
Cluster Som's Som							
1	2	3	4	5	6	7	8
Lisbon cl 3 (26,67%)	0	Lisbon cl 3 (26,67%)	Lisbon cl 1 (40%)	Lisbon cl 1 (40%)	Lisbon cl 1 (40%)	0	Lisbon cl 3 (26,67%)
Lisbon cl 7 (33,33%)	0	0	Lisbon cl 3 (26,67%)	Lisbon cl 3 (26,67%)	Lisbon cl 7 (33,33%)	0	0
0	0	0	Lisbon cl 7 (33,33%)	0	0	0	0
60%		26,67%	100%	66,67%	73,33%		26,67%

Country	Case	Conflict Name	Label		Conflict cluster		
PRT	Algarve	Barrier Islands ria Formosa Natural Park	Alg		1		
Cluster Som's Som							
1	2	3	4	5	6	7	8
Algarve cl 1 (26,06%)	Algarve cl 5 (18,79%)	Algarve cl 4 (0,61%)	Algarve cl 3 (9,09%)	Algarve cl 1 (26,06%)	Algarve cl 1 (26,06%)		Algarve cl 7 (40,91%)
Algarve cl 6 (4,55%)	Algarve cl 7 (40,91%)		Algarve cl 5 (18,79%)	Algarve cl 7 (40,91%)	Algarve cl 6 (4,55%)		
Algarve cl 7 (40,91%)			Algarve cl 6 (4,55%)				
71,52%	59,70%	0,61%	32,43%	46,97%	30,61%		40,91%

Country	Case	Conflict Name	Label		Conflict cluster		
UK	Thames Gateway	Barking riverside	Tha1		5		
Cluster Som's Som							
1	2	3	4	5	6	7	8
		Thames cl 5 (100%)					
		100%					

Country	Case	Conflict Name	Label		Conflict cluster		
UK	Thames Gateway	Lower Thames crossing	Tha2		5		
Cluster Som's Som							
1	2	3	4	5	6	7	8
Thames cl 1 (1,59%)	Thames cl 3 (10,58%)	Thames cl 5 (7,94%)	Thames cl 7 (4,76%)	Thames cl 3 (10,58%)	Thames cl 4 (70,37%)		Thames cl 3 (10,58%)
Thames cl 3 (10,58%)			Thames cl 8 (4,76%)		Thames cl 7 (4,76%)		
Thames cl 4 (70,37%)							
Thames cl 8 (4,76%)							
87,30%	10,58%	7,94%	9,52%	10,58%	75,13%		10,58%

Country	Case	Conflict Name	Label		Conflict cluster		
UK	Portsmouth	Langstone Harbour/Farlington marshes	Por1		1		
Cluster Som's Som							
1	2	3	4	5	6	7	8
		Portsmouth cl 3 (100%)		Portsmouth cl 3 (100%)			
		100%		100%			

Country	Case	Conflict Name	Label		Conflict cluster		
UK	Portsmouth	Tipner regeneration	Por2		1		
Cluster Som's Som							
1	2	3	4	5	6	7	8
		Portsmouth cl 3 (100%)		Portsmouth cl 3 (100%)			
		100%		100%			

Country	Case	Conflict Name	Label		Conflict cluster		
BEL	Zeebrugge	Schipdonk canal	Bru1		2		
Cluster Som's Som							
1	2	3	4	5	6	7	8
			Zeebrugge cl 2 (20%)	Zeebrugge cl 2 (20%)	Zeebrugge cl 6 (80%)		
			20%	20%	80%		

Country	Case	Conflict Name	Label		Conflict Cluster		
BEL	Zeebrugge	Zeebrugge harbour	Bru2		2		
Cluster Som's Som							
1	2	3	4	5	6	7	8
	Zeebrugge cl 5 (100%)						
	100%						

Country	Case	Conflict Name	Label		Conflict Cluster		
BEL	Ostend	Ostend airport	Oste		2		
Cluster Som's Som							
1	2	3	4	5	6	7	8
	Ostend cl 3 (100%)						
	100%						

Country	Case	Conflict Name	Label		Conflict cluster		
SWE	Gothenburg	Torsviken	Got1		1		
Cluster Som's Som							
1	2	3	4	5	6	7	8
Gothenburg cl 1 (100%)	Gothenburg cl 1 (100%)						
100%	100%						

Country	Case	Conflict Name	Label		Conflict cluster		
SWE	Gothenburg	Kungsbacka	Got2		4		
Cluster Som's Som							
1	2	3	4	5	6	7	8
Gothenburg cl 1 (0,61%)	Gothenburg cl 1 (0,61%)	Gothenburg cl 5 (0,14%)	Gothenburg cl 2 (87,48%)	Gothenburg cl 2 (87,48%)	Gothenburg cl 2 (87,48%)		
					Gothenburg cl 7 (11,77%)		
0,61%	0,61%	0,14%	87,48%	87,48%	99,25%		

Country	Case	Conflict Name	Label		Conflict cluster		
SWE	Malmö	Managing urban sprawl	Mal1		1		
Cluster Som's Som							
1	2	3	4	5	6	7	8
Malmö cl 1 (74,43%)	Malmö cl 4 (21,25%)	Malmö cl 3 (1,17%)	Malmö cl 1 (74,43%)	Malmö cl 1 (74,43%)	Malmö cl 1 (74,43%)		Malmö cl 1 (74,43%)
Malmö cl 5 (0,21%)		Malmö cl 6 (0,21%)	Malmö cl 2 (2,74%)				
74,64%	21,25%	1,38%	77,17%	74,43%	74,43%		74,43%

Country	Case	Conflict Name	Label		Conflict cluster		
SWE	Malmö	Falsterbo	Mal2		5		
Cluster Som's Som							
1	2	3	4	5	6	7	8
Malmö cl 1 (99,42%)			Malmö cl 1 (99,42%)	Malmö cl 1 (99,42%)	Malmö cl 1 (99,42%)		Malmö cl 1 (99,42%)
			Malmö cl 2 (0,58%)				
99,42%			100,00%	99,42%	99,42%		99,42%

Alessandro LONDEI, Department of European, American and Intercultural Studies, Sapienza University of Rome, Italy.

email: alessandro.londei@gmail.com

Londei, Alessandro, "Application of Artificial Neural Networks to SECOA project". In: Armando Montanari (ed.), *Urban Coastal Area Conflicts Analysis Methodology: Human Mobility, Climate Change and Local Sustainable Development*. Cap. 7. *SECOA FP7 Research Project*, Vol. 5. Rome: Sapienza Università Editrice, 2013. *http://digilab-epub.uniroma1.it*. DOI: 10.13133/978-88-98533-01-5. Web. 30 September 2013.

ABSTRACT: This chapter discusses the application of artificial neural networks to the SECOA framework in order to accomplish two specific tasks, specifically, to determine the case studies taxonomy and the relationship of the S/E/E geo-referenced data with the conflict features defining SECOA modelling. Self-Organized Maps (SOM) revealed to be a suitable tool for extracting the SECOA taxonomy because of the non-linear dimensionality reduction that this kind of neural network offers and the spontaneous ability to reveal the similarity distribution of high-dimensional data. Moreover, as will be shown in this chapter, SOM is able to homogenise heterogeneous data that present lack information. Feed-forward neural networks are used to extract a descriptive model of the relations between territorial data and conflict features in order to define a well-structured artificial system able to correctly associate the training data and generalise the information passed to the network. Specifically, this last aspect is crucial for the development of a Spatial Decision Support System (SDSS). A SDSS may use the information provided by the neural network modelling to estimate the conflict evolution due to specific political actions by evaluating the linear tangent space of the neural system and the directions of greatest variation of conflict parameters.

KEYWORDS: Artificial Neural Networks, Self-Organizing Maps, Feed-Forward Neural Networks, Back-Propagation, Modelling, Taxonomy, Spatial Decision Support System.

Submitted: 6th June 2013
Accepted: 18th July 2013
Final Revision: 23th July 2013
Published: 30th September 2013

CHAPTER 8.

SECOA Territories and Processes: An Attempt of Conflicts' Interpretation

Barbara STANISCIA
Department of European, American and Intercultural Studies, Sapienza University of Rome, Italy

1. Introduction

The outputs generated by the Self-Organizing Maps (SOM) and Feed Forward Neural Networks (FFNN) related to the taxonomy and the model have been used, in the present study, to analyze and classify the SECOA territories and conflicts in a systematic and comparative way. An attempt has been made to interpret the SECOA realities. The present chapter provides an overview of the achieved results. A classification of the SECOA territories is provided, followed by a classification of the SECOA conflicts. Then, territories and conflicts are interpreted simultaneously and comparatively.

2. A Classification of Territories

2.1 The EU regions

Variables and data used for the SOM's SOM of the EU areas have been re-interpreted in order to achieve taxonomy of the SECOA territories. They are organized into four main categories: variables that explain characteristics related to individuals; variables related to enterprises; variables signifying positive environmental characteristics; and variables signifying negative environmental characteristics (Table 8.2.1).

Table 8.2.1. Interpretation of the SECOA variables and associated data.

<table>
<tr><th colspan="2">Groups of variables</th><th>Data</th></tr>
<tr><td colspan="3">Individuals' related variables</td></tr>
<tr><td>Characteristics of the settled population</td><td>Families' economic prosperity</td><td rowspan="3">Number of residents
Number of households
Average income
Number of tourists
Number of migrants
Unemployment rate
Land use: residential mix</td></tr>
<tr><td>Area attractiveness: for daily activities, temporary mobility, permanent mobility</td><td>Families' economic difficulties</td></tr>
<tr><td>Residential-oriented land use</td><td></td></tr>
</table>

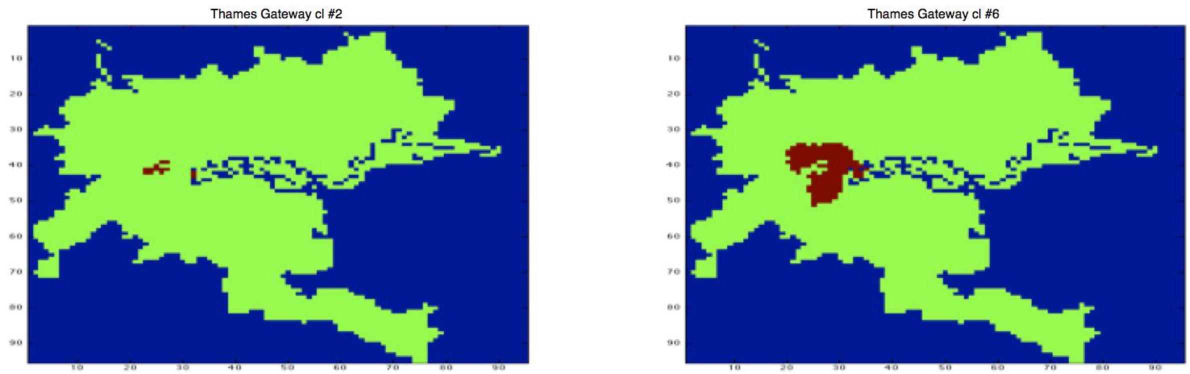
Businesses' related variables		
Enterprises' economic dynamism	Productive-oriented land use	Employees and local units of the enterprises Land use: industrial, commercial
Environment related variables		
Environment friendly land use	Agriculture-oriented land use	Land use: natural, open space, agriculture
Environmental problems related variables		
Natural hazards		
Pollution	Infrastructures with high impact	Flood hazard SLR Motorization rate

The first group includes variables that are able to describe individuals as well as households and their dwellings. The second group includes variables that are able to describe the economic activities, related to both enterprises and land use. The third group includes variables that are able to describe a positive environmental state. Agricultural land use has been interpreted as positive for the environment, since agricultural activity has been considered as a way to preserve the territory and the landscape. The fourth category includes variables that are able to describe a negative environmental state or environmental threats.

Following the outputs of the SOM's SOM, the territories of the metropolitan and urban regions located in the EU (except for Funchal) have been grouped in three main types: (i) central, urban areas; (ii) periurban and suburban areas; and (iii) rural and natural areas.

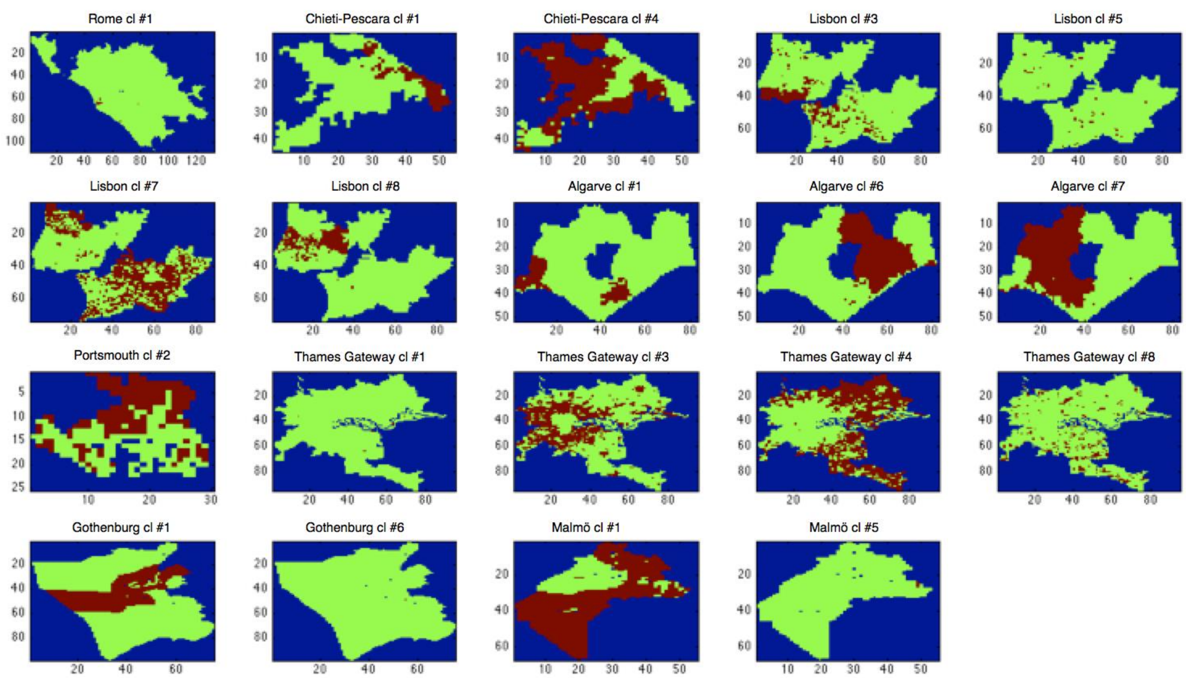
Central, urban areas could be further subdivided. There are, indeed, some areas that are characterized by high income, high unemployment, a high number of households and residential areas, a disproportionately low number of residents, high intensity of the services and industrial sector (no intensity of the agriculture sector), that are unattractive for the permanent component of human mobility, that present a low motorization rate and are particularly exposed to natural hazards; those areas show characteristics that can be found in central, gentrified urban areas (Figure 8.2.1.1);

Figure 8.2.1.1. EU areas: type i) – a).



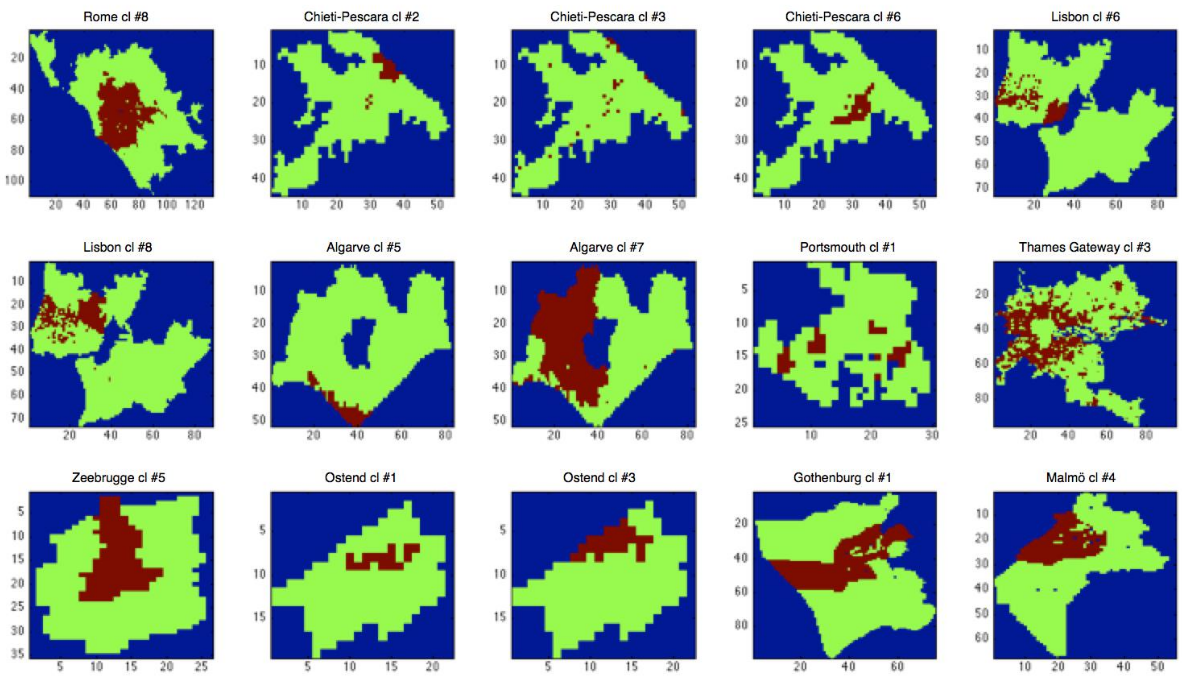
there are some attractive areas for both the permanent and temporary components of human mobility with important motorization rates and high impact of infrastructures and some traces of agriculture activity; those areas show characteristics that make them comparable to core areas (Figure 8.2.1.2);

Figure 8.2.1.2. EU areas: type i) – b).



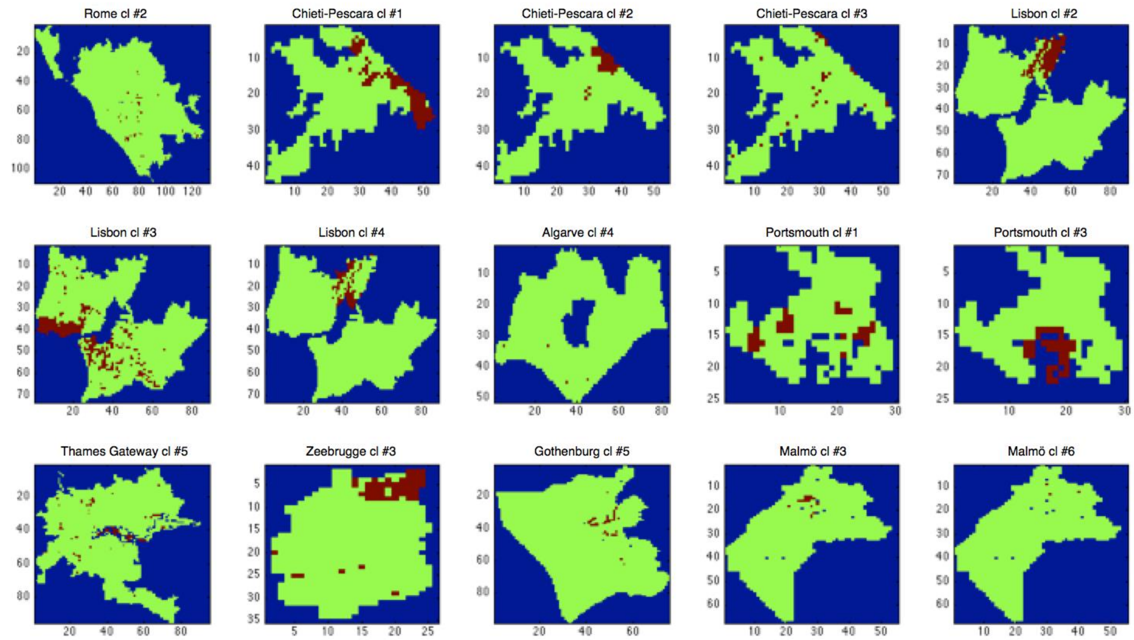
there are, finally, areas with an important presence of households and residents, attractive for tourists, with a high number of employees, that do not present unemployment problems, and have a high motorization rate; those areas show characteristics that can be considered similar to those of central, mature urban areas (Figure 8.2.1.3).

Figure 8.2.1.3 - EU areas: type i) – c).



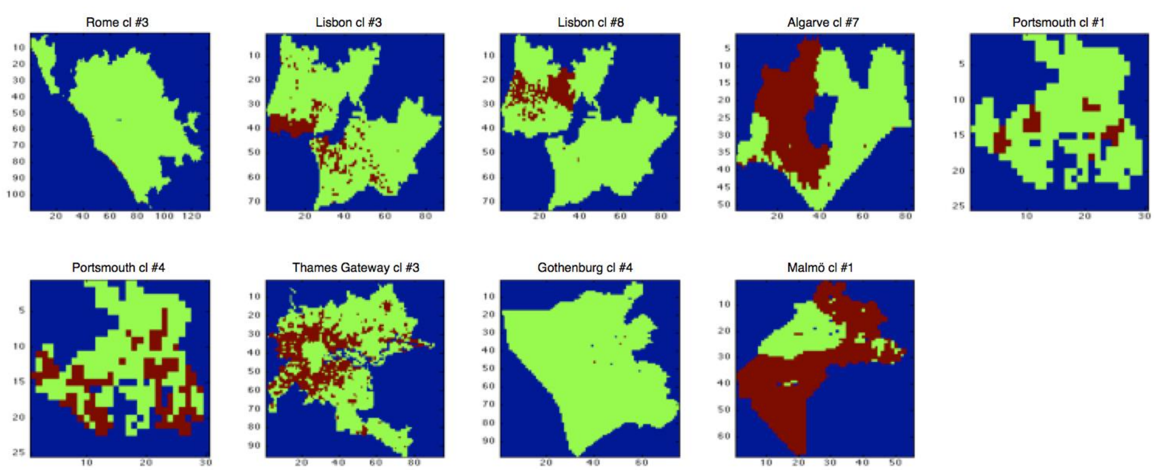
Among the periurban and suburban areas, different features can be highlighted. There are, indeed, areas with high intensity of commercial and industrial spaces (and no agriculture), dynamism of all economic sectors (but important unemployment rates), low presence of households, and subjected to natural hazards; those areas show characteristics that can be assimilated to those of peri-urban areas (Figure 8.2.1.4);

Figure 8.2.1.4. EU areas: type ii) – a).



there are other areas characterized by an important presence of residents and households, high motorization rate, that are attractive for the permanent component of human mobility, and subjected to the risks linked to the sea level rise; those areas present features that can be assimilated to those of suburban, residential, coastal areas (Figure 8.2.1.5);

Figure 8.2.1.5. EU areas: type ii) – b).



there are, finally, areas characterized by the importance of residential land use; they can be defined as residential, periurban and suburban (Figure 8.2.1.6).

Figure 8.2.1.6. EU areas: type ii) – c).

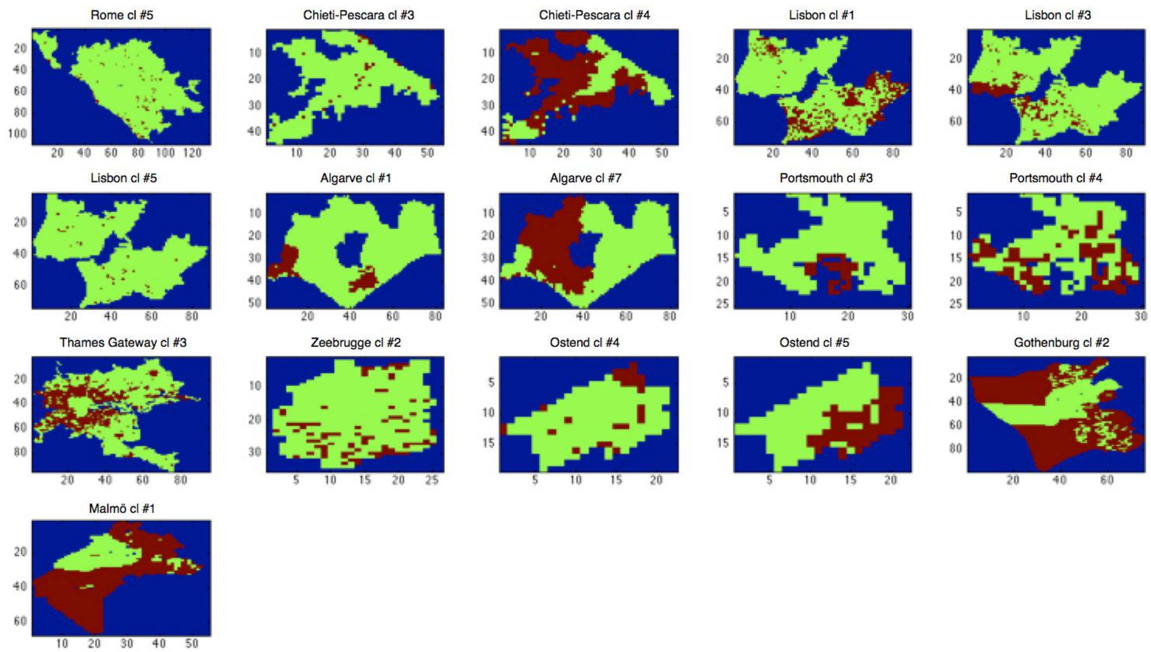
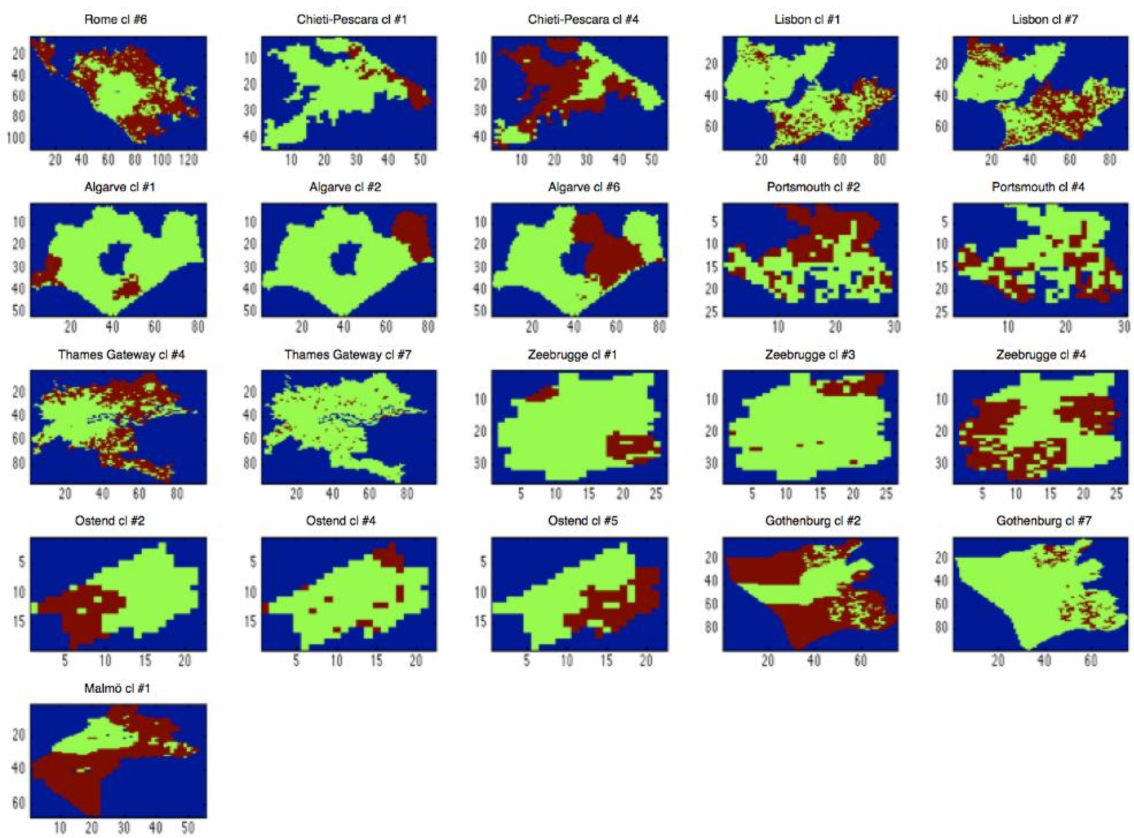
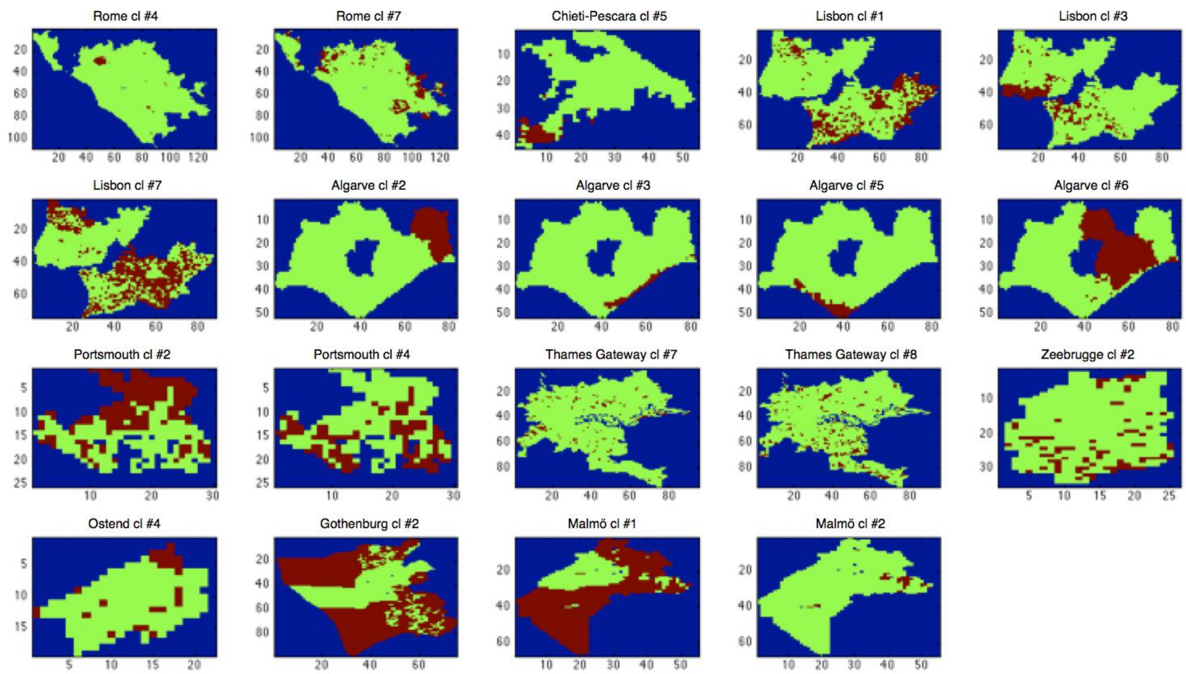


Figure 8.2.1.7. EU areas: type iii) – a).



Among the rural and natural regions, areas characterized by the dominance of agriculture over the other sectors of economic activities and low presence of households can be found (Figure 8.2.1.7) as well as areas characterized by an important presence of nature and open spaces (Figure 8.2.1.8);

Figure 8.2.1.8. EU areas: type iii) – b).



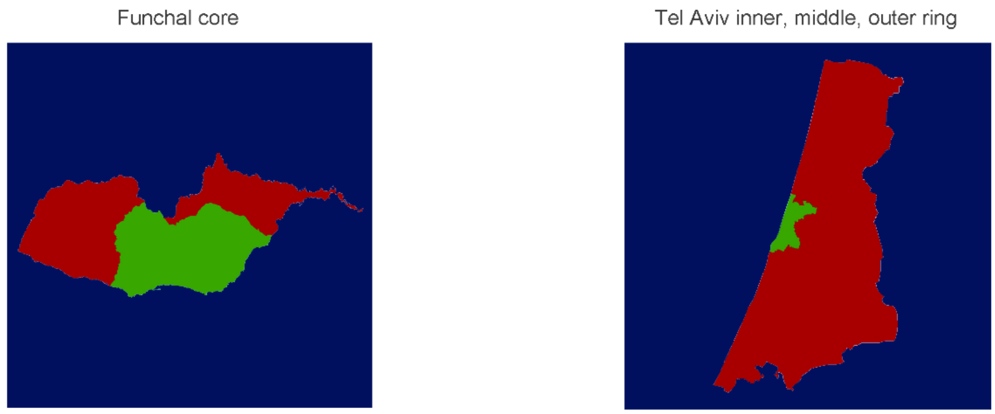
the first ones can be considered as peripheral, rural areas while the second ones can be considered as peripheral, natural areas.

8.2.2 The Asian regions and Funchal

Given the difference among the variables and data used, a different classification was elaborated for the Asian regions and Funchal. Interpreting the variables according to the scheme provided in Table 8.2.1 and the outcomes of the SOM's SOM, those areas could be grouped into five main types.

In the core of Funchal and in the inner, middle and outer rings of Tel Aviv, areas characterized by a high concentration of population and dwellings, attractive for both tourists and migrants, with the presence of economic activities concentrated in the tertiary sector are found (Figure 8.2.2.1);

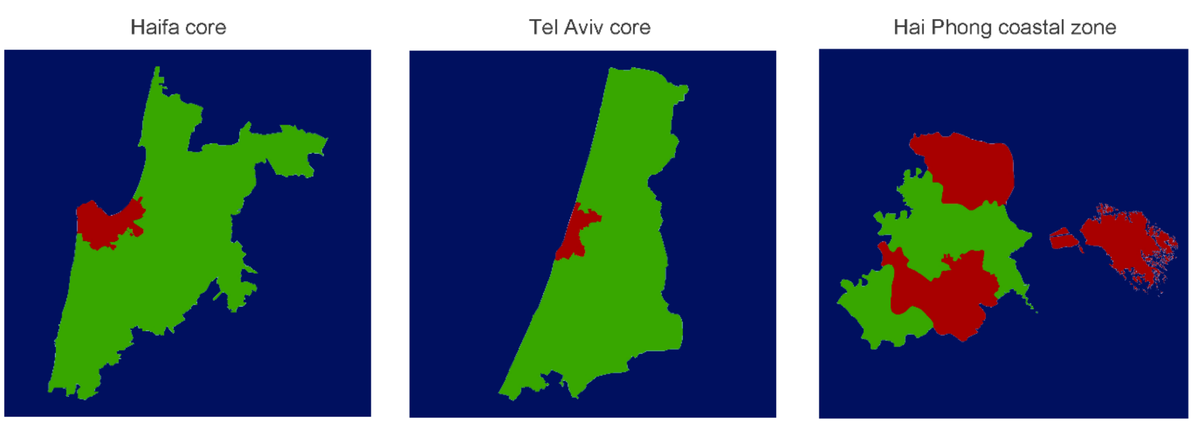
Figure 8.2.2.1. Asian regions and Funchal: type a).



they can be assimilated to the central, urban areas of the EU cases.

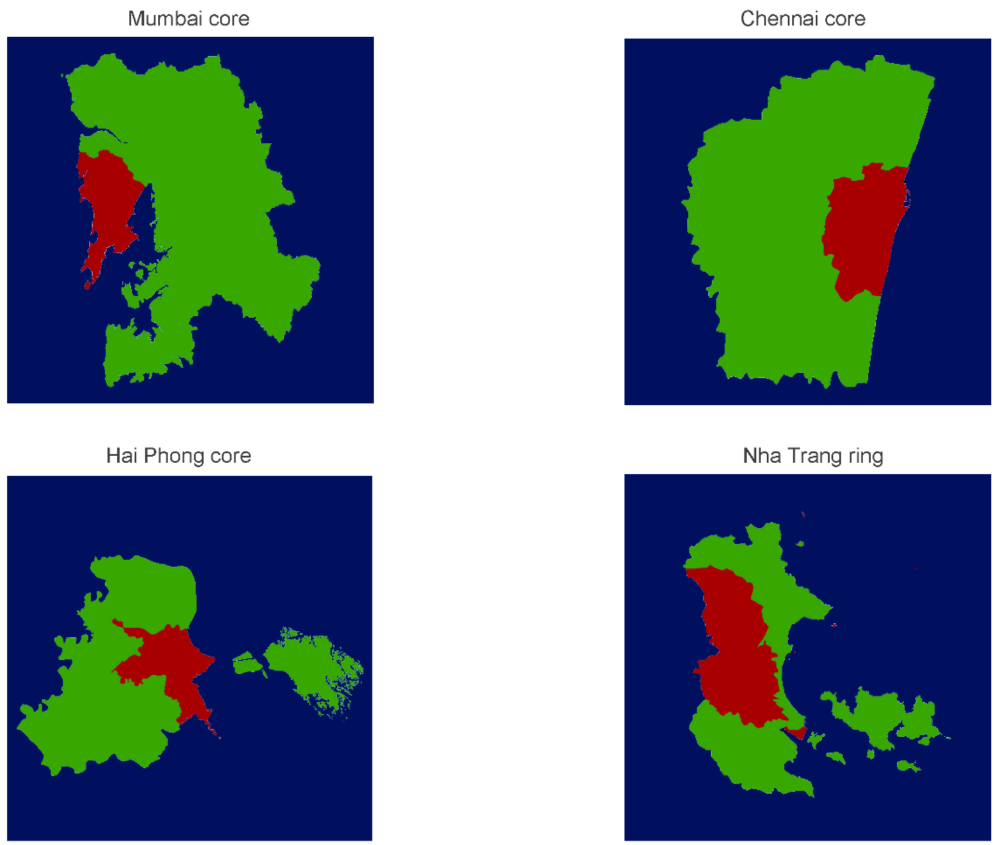
In the cores of Haifa and Tel Aviv and in the Hai Phong coastal zone there is the presence of areas with similar characteristics of the central, urban areas of the EU cases with high exposure to the damages of natural hazards (Figure 8.2.2.2).

Figure 8.2.2.2. Asian regions and Funchal: type b).



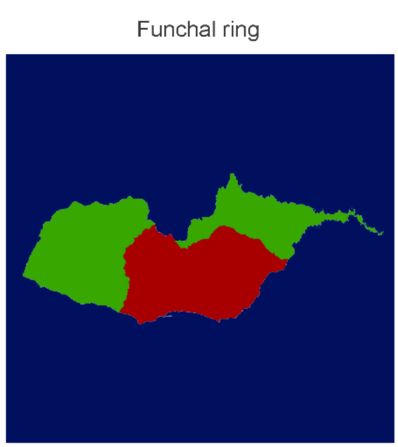
In the cores of Mumbai, Chennai and Hai Phong and in the ring of Nha Trang, there are areas with similar characteristics of the central, urban areas of the EU cases but with an important presence of informal and shanty dwellings (Figure 8.2.2.3).

Figure 8.2.2.3. Asian regions and Funchal: type c).



In the ring of Funchal, peri-urban and sub-urban areas characterized by a mixed use of residences and economic activities, high rates of motorization and commuting are recorded (Figure 8.2.2.4);

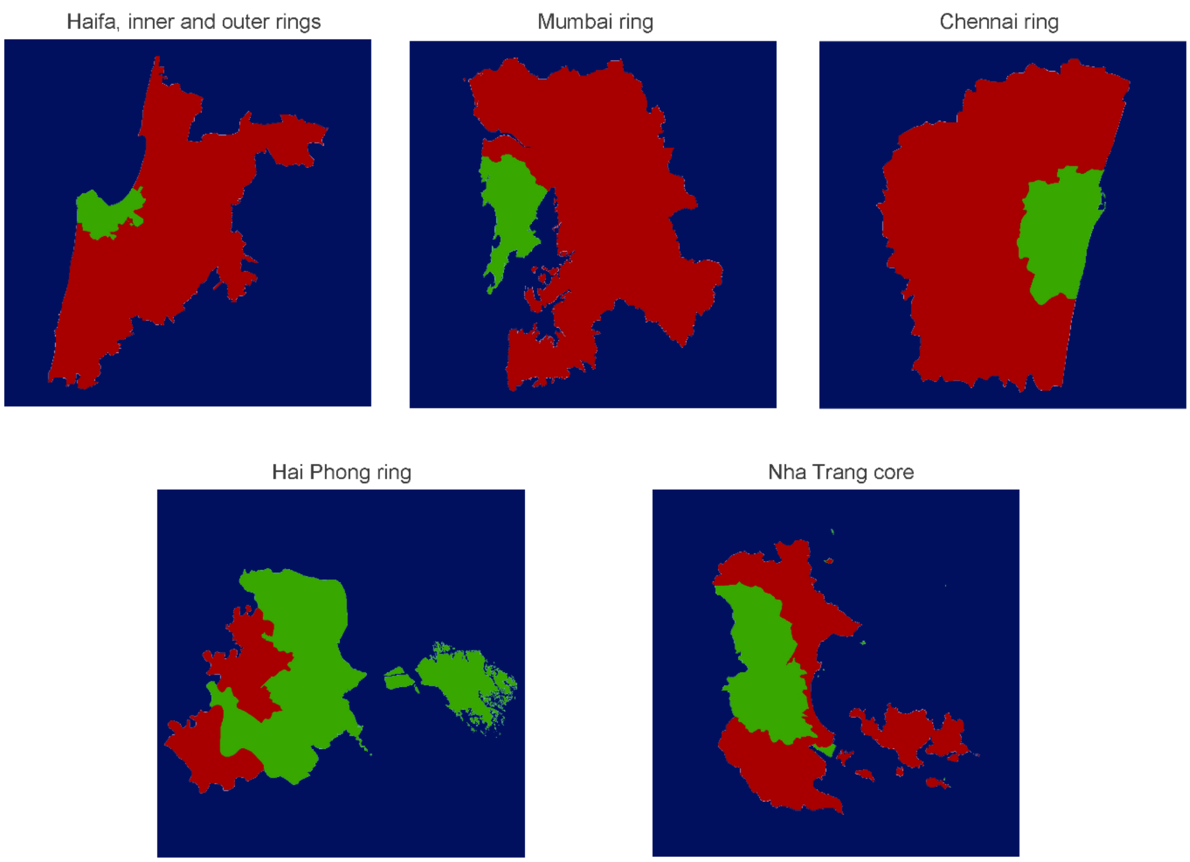
Figure 8.2.2.4. Asian regions and Funchal: type d).



they can be compared to the periurban and suburban areas of the EU cases.

Finally, rural and natural areas can be found in the inner and outer rings of Haifa, in the rings of Mumbai, Chennai and Hai Phong, in the core of Nha Trang (Figure 8.2.2.5).

Figure 8.2.2.5 - Asian regions and Funchal: type e)



3. A Classification of Conflicts

Conflicts have been classified using the variables and data described in chapter 4. Following the cluster analysis performed through the SOM, five groups of conflicts have been identified (Table 8.3.1).

Table 8.3.1 Classification of the SECOA conflicts.

Countries	Metropolitan/urban Region/areas	Conflicts	Types
BE	Ostend	Ostend airport	B
BE	Zeebrugge	Schipdonk canal	B
BE	Zeebrugge	Zeebrugge harbour	B
IL	Tel Aviv	Palmachim beach	D
IL	Tel Aviv	Netanya sandstone cliffs	A
IL	Haifa	Haifa port	E
IN	Mumbai	Sanjay Gandhi National Park	A
IN	Chennai	Pallikaranai marshland	A
IN	Mumbai	Mangrove forest	A
IT	Rome	Civitavecchia	B
IT	Chieti-Pescara	Costa Teatina National Park	A
IT	Rome	Ostia water-use and management	C
PT	Lisbon	Trafaria and Costa da Caparica	A
PT	Algarve	Barrier Islands (Ria Formosa Natural Park9	A
PT	Funchal	Funchal Bay (Madeira Island)	D
SE	Malmö	Managing urban sprawl	D
SE	Malmö	Falsterbo-Peninsula: Vellinge municipality	D
SE	Gothenburg	Torsviken	E
SE	Gothenburg	Kungsbacka	E
UK	Thames Gateway	Barking riverside	D
UK	Thames Gateway	Lower Thames crossing	A
UK	Portsmouth	Langstone Harbour/Farlington marshes	A
UK	Portsmouth	Tipner regeneration	A
VN	Haiphong	Haiphong port	B
VN	Haiphong	Industrial zone	B
VN	Haiphong	Cat Ba	C
VN	Nha Trang	Nha Trang	C

There are conflicts strongly marked by contrasts linked to the presence of natural parks, protected areas, and preservation of natural sites that urge a solution. There are, then, conflicts strongly marked by contrasts linked to the presence of impacting infrastructures (harbours, airports, industrial zones) and transportation systems supporting commuting and provoking pollution; some of them are chronic and long lasting, while others are in their formation stage. The scale of the conflict is between the micro and macro levels. In the third group we find conflicts linked to the competition between tourism development and nature protection and to resources management; the micro and macro levels are involved; they are chronic and long lasting and require an urgent response. The fourth category of conflicts grouped those mainly linked to legal or policy reasons or to infringements over access, that are in a stage of management, that are hushed, for which the solution is going to be deferred, or in which several territorial levels are involved. The fifth group, finally, includes conflicts characterized by high complexity, in which several issues, territorial and institutional levels, are involved; they are in a stage of formation or manifestation, they are neither chronic nor urgent and are not long lasting.

4. An Attempt to Interpret Territories and Conflicts

Taxonomy of the geographical spaces and conflicts studied by SECOA is reported in Table 8.4.1.

For each conflict, the typology according to the classification provided in section 8.3 is indicated; the typology of space in which that conflict takes place is identified according to the classification provided in section 8.2. Every conflict is associated with a type of space; each type of space has its own conflicts. The two colours used in the table provide an indication about the intensity of the relation between conflict and space: the stronger the colour is the more intense the relationship.

Table 8.4.1 Classification of the SECOA territories and conflicts.

Countries	**Conflicts**		**Spatial typologies**				
	Conflicts	**conflicts' typologies**	**central, urban areas**	**periurban and suburban areas**	**rural and natural areas**	**core areas very sensitive to naturalhazards (only in Asia)**	**core areas with an important presence of the informal city (only in India and Vietnam)**
IL	Netanya sandstone cliffs	A	X				
IN	Sanjay Gandhi National Park	A					X
IN	Mangrove forest	A				X	X
IN	Pallikaranai marshland	A			X		
IT	Costa Teatina National Park	A	X	X	X		
PT	Trafaria and Costa da Caparica	A			X		
PT	Barrier Islands (Ria Formosa Natural Park)	A	X				
UK	Langstone Harbour/Farlington marshes	A		X			
UK	Tipner regeneration	A		X			
UK	Lower Thames crossing	A	X				
BE	Schipdonk canal	B			X		
BE	Zeebrugge harbour	B	X				
BE	Ostend airport	B	X				
IT	Civitavecchia	B			X		
VN	Haiphong port	B				X	X
VN	Industrial zone	B			X	X	X
IT	Ostia water-use and management	C	X				
VN	Cat Ba	C				X	X
VN	Nha Trang	C			X		
IL	Palmachim beach	D	X				
PT	Funchal Bay (Madeira Island)	D	X				
SE	Managing urban sprawl	D	X	X	X		
SE	Falsterbo-Peninsula: Vellinge municipality	D	X	X	X		
UK	Barking riverside	D		X			
IL	Haifa port	E		X		X	
SE	Kungsbacka	E			X		
SE	Torsviken	E	X				

The territories and conflicts studied by SECOA were modelled through a bottom-up approach using FFNN. The resulting model is coherent and consistent with the taxonomy. It is an interpretative model that aims at describing the relationships and interactions existing between the spatial patterns of the SECOA metropolitan/urban areas/regions and the features of the (analyzed) conflicts taking place in those areas. It establishes the degree and the importance of each spatial variable in shaping the different aspects of each conflict and highlights the variables that are more likely to affect (positively or negatively) each conflicts' characteristic. The model indicates the weight of the influence of each variable on each aspect of each conflict and shows how the same variable has different effects on different aspects of the conflicts.

For each conflict, a sort of "fingerprint" has been obtained, given by the weight of the influence that each variable has on each parameter defining the conflict. The weight has to be considered in a relative sense. It can be positive or negative. In the first case, an increase in the value of the given variable provokes an increase of the value of the given parameter defining the conflict. In the second case, an increase of the value of the given variable determines a decrease in the value of the given parameter defining the conflict. It is important to note that the weight and influence of the variables on the parameters differ in each conflict; the weight and influence of the variables on the parameters are determined by the whole model; i.e. they are determined by all data and information considered in the model; those outcomes are specific to the SECOA model and cannot be generalized; the relationships between spatial variables and conflicts' parameters can be useful when, in defining appropriate policies, the policymakers have to decide which is the appropriate lever to use to bring a change in the conflict; the model outcomes can be used as a basis for building alternative future scenarios.

5. Conclusion

Developing taxonomy and a model from the variables and data available in SECOA is a great challenge. Extremely different territories as well as very heterogeneous availability and organization of data made the task difficult. The use of GIS and Artificial Neural Networks made the work possible. Taxonomy and model are tools to be used for other purposes such as the construction of scenarios. They explain and summarize complex situations but are unable to provide an interpretation of the diversities and specificities of the different realities. This interpretation should be made case by case and using specific disciplinary skills.

Taxonomy and model have been ways to synthesize and reduce the complexity of the results derived by different spatial analyses. Are ways to achieve general results - while still remaining within the given context – going over analytical details. Are ways to implement comparative research and to allow generalization. This inevitably leads to a loss of detailed information and a simplification of the analyzed realities. It is a way to create a logical bridge with further phases of research and of action, the ones involving the construction of alternative scenarios and dedicated policies.

Barbara STANISCIA, Department of European, American and Intercultural Studies, Sapienza University of Rome, Italy.

email: barbara.staniscia@uniroma1.it

Staniscia, Barbara, "SECOA Territories and Processes: an Attempt of Conflicts' Interpretation". In: Armando Montanari (ed.), *Urban Coastal Area Conflicts Analysis Methodology: Human Mobility, Climate Change and Local Sustainable Development*. Cap. 8. *SECOA FP7 Research Project*, Vol. 5. Rome: Sapienza Università Editrice, 2013. *http://digilab-epub.uniroma1.it*. DOI: 10.13133/978-88-98533-01-5. Web. 30 September 2013.

ABSTRACT: This chapter presents a systematic classification and comparison of the SECOA territories and conflicts. The variables and the associated data are interpreted and grouped into four main categories. The outputs generated by the Self-Organizing Maps and the Feed Forward Neural Networks, related to the taxonomy and the model, are interpreted from a territorial perspective. Five territorial typologies and five types of conflicts are identified. Different interpretative processes for EU and Asian regions are necessary due to differences in data availability. The 17 urban coastal regions and the 27 conflicts in SECOA are classified according to the identified typologies.

KEYWORDS: Classification, Territorial Typologies, Types of Conflicts, EU Regions, Asian Regions

---

Submitted:	6th June 2013
Accepted:	18th July 2013
Final Revision:	23th July 2013
Published:	30th September 2013

Annex I.

CHAPTER 9.

General Conclusions

Armando MONTANARI
Department of European, American and Intercultural Studies,
Sapienza University of Rome, Italy

1. Conclusions

This volume is the completed section of the process of analytical research and methodological comparisons undertaken by SECOA, a 48-month research project selected and funded by the EU under the FP7 program. Hence, while scientifically autonomous, the volume is a natural link between the different phases of analysis within SECOA, i.e. Work Packages (WPs) 1-5, and the interpretive and predictive values that are being drawn up by WPs 7 and 8. Within the overall scope of SECOA's research activity, this volume's task was to supply answers to questions that will undergo further study by research groups. These groups will subsequently have to create methods and tools to identify the most suitable policies to effectively manage environmental conflicts, use fragile and rare resources more efficiently, and develop administrative structures capable of dealing with the needs of a continuously evolving society (the wisdom stage). It was also deemed necessary to construct possible alternative scenarios in order to contribute to an enhanced vision of sustainable urban development in coastal areas (the understanding stage).

The findings of the research discussed in this volume are to be used to understand the relationships between the variables collected in the previous phases (WPs 1, 2, 3 and 4) of SECOA. The physical, environmental, economic and social contexts were interpreted using synthetic variables that, appropriately elaborated, made it possible to show the relationships and hierarchies of the values determining conflicts in urban coastal areas. In particular, this book looks at: (i) a list of the ways in which the sustainable development variables are related to the nature of the conflicts; (ii) an analysis of the way in which modifying the sustainable development variables could influence the nature of the conflicts; (iii) the measurement of the hierarchical levels of the influences mentioned in the previous point; (iv) the identification of the variables that principally influence the meaning and specificity of the conflicts. A better understanding of the context also highlighted the balance of power among the various types of stakeholders that make up civil society in the case studies selected by SECOA. The human component, habitual behaviours, cultural references and local administration methods vary considerably among the case studies; there are particularly marked differences between northern and southern European countries, and between European countries in general and the Asian countries that were part of the study. Nevertheless, for at least two values it was possible to find a pattern of uniform behaviour in the different case studies. By their very nature, global problems unfold in the same way and cause similar reactions, irrespective of local characteristics. There is undoubtedly a difference at the local level when it comes to the history and duration of conflicts. At this stage of the evolution of human civilization, the management of environmental conflicts is an indicator of the development of civic culture in each place. The term "contrast", a significant element of SECOA, was replaced in the operational phase of research with "conflict", which better describes a phenomenon that – while not involving the use of weapons – has all the hallmarks of sometimes protracted hostilities. The findings of the research project thus far include the realisation that it is precisely because of the way in which these conflicts – not contrasts – are perceived that it is near-impossible for them to ever be totally resolved. Environmental disputes and conflicts have become part and parcel of our post-modern society. So the term "solution", also a key initial concept of SECOA, had to be replaced by the more realistic term "mitigation". Environmental conflict has become such an ingrained part of the basic values of our society that it is no longer possible to root it out. For public administrations, one possibility – perhaps even a duty – is to try and mitigate existing conflicts, which will, however, continue to exist in other forms, and about other subjects.

Sustainable development is a system in which the balance of the elements that make it up – the environment, the economic system, the social and cultural structure – must be maintained to make it possible for the natural systems to survive. In the logical model of SECOA, the benchmark balance is applied to urban coastal areas that are particularly vulnerable to global processes such as climate change and human mobility because of their location and their fragility. At the local level, there is balance if the administration and the social and cultural traditions of a particular society manage to interact properly. The lack of balance is determined by an element of error. The error in this type of system creates conflicts of varying intensity and duration. Each conflict has specific aspects which, however, vary within certain variables, defined using typological criteria. The error layout for each conflict has been summarised into a colour layout alternating positive and negative values of different levels, defined "fingerprint". The fingerprint is the static layout of the error values. When the parameters are changed, the fingerprint takes on different values and meanings that can contribute to building predictive scenarios and form the basis for implementing the envisaged Spatial Decision Support System (SDSS).

List of Tables

Chapter 3.

Chapter 4.

Chapter 5.

Chapter 6.

Chapter 7.

Annex to Chapter 7.

Chapter 8.

List of Figures

Chapter 0.

Chapter 1.

Chapter 3.

Chapter 4.

Chapter 5.

Chapter 6.

Chapter 7.

Annex to Chapter 7.

Chapter 8.