Analysis of land change with parameterised multi-level class sets

Exploring the semantic dimension

Louisa J.M. Jansen

Thesis Committee

Thesis supervisor:
Prof. Dr. Ir. A. Veldkamp
Professor of Land Dynamics
Wageningen University, The Netherlands

Other members:
Prof. Dr. Ir. A.K. Bregt, Wageningen University, The Netherlands
Prof. Dr. A.J. Comber, Leicester University, United Kingdom
Prof. Dr. M. Herold, Wageningen University, The Netherlands
Prof. Dr. Ir. A. Stein, University of Twente, The Netherlands

Analysis of land change with parameterised multi-level class sets

Exploring the semantic dimension

Louisa J.M. Jansen

Thesis
Submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof. Dr. M.J. Kropff,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Wednesday 3 November 2010
at 4 p.m. in the Aula

Louisa J.M. Jansen
Analysis of land change with parameterised multi-level class sets: exploring the semantic dimension
Pages: 229
Ph.D. Thesis, Wageningen University, Wageningen, The Netherlands
With summaries in English and Dutch

ISBN: 978-90-8585-828-7

Preface

Chances, passion, talent and hard work.

Malcolm Gladwell, Outliers (2008)

No one is given the map to their dreams,
All we can do is to trace it.
See where we go to, know where we've been,
Build up the courage to face it.

Sandy Denny, One Way Donkey Ride (1977)

In 1989 the Royal Tropical Institute (KIT), The Netherlands, gave me the opportunity to go to the *Département de recherche sur les systèmes de production rural* (DRSPR) - *volet Fonsébougou* project in Sikasso, Mali. This period and project has remained a 'soft spot' in my memories and experiences. It provided my entrance into the world of international co-operation, life abroad and a first real work environment. Moreover, I started to work on a subject that in the second half of the 1990s would gain momentum, land-change analysis[1], with a tool present at the DRSPR *Opération haute volet* project in Bamako that nowadays is considered common, a Geographic Information System. I was also fortunate in collaborating with the Information Technology (IT) expert of the latter project and thus became profoundly aware of the immediate advantages of the interaction between subject-matter and IT experts.

Early in my professional working life some persons suggested I should use the results of my work to do a PhD and I kept their suggestion in mind. When I was employed at FAO none of my colleagues wanted to provide any money for some publications I had prepared. Looking back I should perhaps be grateful for their refusal because then the option to publish in scientific journals remained, as that would only take my time. From then onwards the flow of publications has grown steadily. Mostly because I really enjoyed the writing up and taking things a step further than was possible within the limited timeframe of projects. Furthermore, because writing means sharing, confronting my work with that of others, receiving their feedback and learning from them.

The option to use a selection of publications with a joint theme for a PhD emerged again. There has always been only one possible place that I would consider for doing a PhD: Wageningen University. Not only because it was the place in which I have

[1] Jansen, L., Diarra, S., 1992. Mali-Sud, étude diachronique des surfaces agricoles. Quantification des superficies agricoles et de la dégradation pour quatre terroirs villageois de 1952 à 1987. Version révisée. Institut d'Economie Rurale/Institut Royal des Tropiques. KIT Publications, Amsterdam. 57pp.

studied with great pleasure and the place that gave me the opportunity to touch a great variety of subjects in many different places around the world, but because wherever I was in the world ‘Wageningen’ would stand for something that is internationally recognized. Wageningen is worldwide on the ‘map’ and it is, even after having left many years ago, still on my ‘map’.

Having travelled and worked in several continents and spoken my languages, the subject of land change in combination with semantics, the study of meaning, is not so far-fetched as it might seem. Prolific experiences in describing and categorising phenomena, or using systems developed by others, proved a good foundation for reflecting on how different systems could ‘communicate’ better than they actually do. A ‘common language’ can act as a means of knowledge communication. In early civilisations such written accounts were needed to support development. Dante Alighieri wrote in 1320 that a work can also be *polysemantic*, which is of many senses. But what if these many senses are not clear to the present day reader? Rome, *cosi bella*, has proven to be an excellent place to contemplate communication and to bring it into practice. Communication is essential if we want to make progress in understanding each other and the world at large.

I have drawn some more lines on the ‘map’ to my dreams with the completion of this thesis that was written in-between consulting assignments and many spare hours in the evenings and weekends. It is the result of the big and small chances that arose in my life, the profound passion for my work, a bit of talent and certainly lots of hard work. Every time I thought to have gained knowledge I realised to be at the very beginning, so perseverance in the undertaking was certainly a decisive factor.

Some say that knowledge is something that you never have
Some say that knowledge is something sat in your lap
I must admit, just when I think I ‘m king, I just begin
Just when I think I’m king, I must admit I just begin.

Kate Bush, Sat In Your Lap (1981)

This thesis is spanning a range of years and there are too many names to thank everyone individually. But, I want all of you to know that I greatly appreciate all your direct or indirect contributions.

Acknowledgements

As the majority of publications used for this thesis reflect the results of the projects I worked in, or were prepared for conferences I was invited to, I would like to sincerely thank the following persons and institutions (in chronological order):

- The FAO Land and Water Development Division's Regular Programme on Land-Use Characterisation and Classification that reintroduced me into the world of land use in 1995.
- The Institute for Terrestrial Ecology and World Conservation Monitoring Centre for the discussions on land-use and land-cover categorisation approaches, 30-31 May 1996, Cambridge, UK. In particular Mr Barry Wyatt is thanked.
- The U.S. Federal Geographic Data Committee – Vegetation Subcommittee and Earth Cover Working Group for their invitation to present the parameterised land-cover categorisation approach, 15-16 October 1996, in Washington, DC. In particular Mr Gyde Lund, former Liaison Officer of the US Department of Agriculture - Forest Service, is thanked for his support.
- The Canadian Centre for Remote Sensing for their invitation to present the parameterised land-cover categorisation approach, 18 October 1996, in Ottawa, Canada. In particular Messrs Denny Kalensky and Stefan Palko are thanked for their support.
- Mr Christophe Duhamel, previously CESD Communautaire and LANDSIS, for the invitations to participate in the European Commission (EC) concerted action LANES *Development of a harmonised framework for multi-purpose land-cover/land-use information systems* in 1997 and the EC seminar on *Land-cover and land-use information systems for European Union policy needs*, held 21-23 January 1998, Luxembourg, Luxembourg. Both events underlined the increasing awareness on difficulties arising when constructing and comparing land-cover and land-use class sets and introduced the parameterised approach in the EC context.
- The Istituto Agronomico per l'Oltremare of the Italian Ministry of Foreign Affairs, Florence, Italy, which executed the *Projet de formation en gestion des ressources naturelles et sécurité alimentaire* in Niger and Senegal that allowed me to bring into practise my ideas on land-change analysis that were to become more elaborated and refined in successive projects.
- Mr Geoff Groom of the National Environmental Research Institute, Kalø, Denmark, who coordinated the *Nordic Landscape Monitoring* project of the Nordic Council of Ministers in which we first experimented with the feasibility and practice of harmonisation of the semantic contents of land-cover/use class sets and the quantification of harmonisation results.
- The Directorate General of Forests and Pastures of the Ministry of Agriculture and Food, Albania, and Agrotec S.p.A., Italy, which executed the *Albanian National Forest Inventory* (Loan/credit 2846 ALB) project in association with the Department of Forest Science and Environment of the University La Tuscia, Viterbo, Italy.

- The Ministry of Agriculture and Food, Tirana, Albania, the EU PHARE Programme Management Unit, Albania, and AGROTEC S.p.A., Italy, which executed the *Land-Use Policy II* (EU PHARE AL98-0502) project in consortium with Kadaster International, The Netherlands.
- The Ministry of Agriculture, Food and Forestry, Romania, the Romanian Space Agency and the Romanian Centre for Remote Sensing Applications in Agriculture for making available the data from the FAO *Land-Cover/Land-Use Inventory by Remote Sensing for the Agricultural Reform* (TCP/ROM/2801) project.
- Prof. Ryutaro Tateishi for his invitation to the 12^{th} Center for Environmental Remote Sensing (CEReS) international symposium entitled *Maximizing the use of satellite data for understanding the Earth environment*, held 13-14 December 2005, Chiba, Japan, that allowed me to collect all my thoughts on land-use (change) harmonisation.
- The Land-Dynamics Group of the Department of Environmental Sciences of Wageningen University, The Netherlands, for making available the elaborated data for Romania used in the EURURALIS project. In particular Mr Tom Veldkamp is thanked for his initial long-distance support and his coaching at the end to complete this thesis. Mr. Kasper Kok is thanked for the discussions on scale dimensions.

Table of contents

List of figures

List of tables

1 General introduction

The land is the simplest form of architecture.
Frank Lloyd Wright (1869-1959)

Un paysage quelconque est un état de l'âme.
Henri-Fréderic Amiel (1821-1881)

1.1 Relevance and rationale

The extent and intensity of land-cover change and land-use change, in short land change, increased in the 20th century (Lambin and Geist 2006). Land-based changes support over six billion people with food, fibre, water and other benefits, and supports the highest global average per capita consumption ever known (Turner *et al.* 2008). Land change has implications for sustainable development, livelihood systems and biodiversity as well as contributing to changes in the biogeochemical cycles of the Earth. Thus, land change is central to global environmental change (Meyer and Turner 1994; Turner *et al*. 1995; Walker *et al*. 1997; Lambin *et al*. 2000). Land-cover types and changes in them are sources and sinks for most of the material and energy flows that sustain the biosphere and geosphere, including trace gas emissions and the hydrological cycle. The value and use of land, in addition to the quality of other resources (e.g., water or minerals), are critical to the discussion of viable and sustainable development. Trajectories of land change involve both positive and negative human-environment interactions. Understanding change dynamics does not only help to identify vulnerable places, but also vulnerable (groups of) land users that on their own are incapable to respond in the face of environmental processes and problems.

The examination of pathways of land-use change is crucial for designing appropriate land-use policy interventions aimed at achieving sustainable management of ecosystems and rural development. For understanding the causes and effects of these land-use changes, it is critical to study the interaction between the temporal dynamics and the spatial pattern of land use. Interactions arise from feedbacks in the human-environment system, heterogeneity in the biophysical and human environment and the influence of land-use history (Verburg *et al.* 2004a; Lambin and Geist 2006). Land-use patterns form architecture in that most lands are managed and thus their use is designed, *de facto* or *de jure* (Turner 2010). But as with most building architecture, the land architecture in one place does not render similar results if duplicated in other places, at other times or at other scales.

The recognition that land use and land cover are closely related has called for a coupled human-environment, or social-biophysical, system analysis resulting in a joint project of the International Geosphere-Biosphere Programme (IGBP) and International Human Dimensions Programme on Global Environmental Change (IHDP): the Land-Use and Land-Cover Change (LUCC) project, followed up in 2006 by the Global Land Project (GLP 2005). Two crosscutting activities of the integrated research foci of the LUCC international research community were defined that are important in the context of this thesis (Turner *et al.* 1995; Lambin *et al.* 1999):

1. *Categorisation and data* examines data availability and data quality, and devises a categorisation structure suitable for the various research needs. Over time knowledge advances, technology develops and policy objectives change resulting in the fact that every survey, rather than being part of a sequence, creates a new baseline data set. Whereas before survey maps were an illustration that accompanied a descriptive memoir, nowadays maps are understood as data sets. This poses a problem as to the attached class labels that are often rather cryptic and unrelated to any categorisation system where the user may learn the concepts and criteria behind the class labels (Comber *et al.* 2005a; Wadsworth *et al*. 2006). Differences in the naming of classes, changes in class definition and adding or removing classes in data sets covering the same area in different periods will create difficulties in the interpretation of actual changes over time from changes in category definition. Also the exploration of category relationships within a categorisation system can reveal how well classes are separated or if there is a risk of confusion between classes, a situation that may be problematic from a data accuracy perspective (Ahlqvist 2005a and 2005b). Sokal (1974) defines classification, or categorisation, as "*the ordering or arrangement of objects into groups or sets on the basis of relationships. These relationships can be based upon observable or inferred properties*". Thus, classification provides a systematic categorisation framework and it is at the same time a simplification as it represents only part of the complexity of reality. Categorisation acts as a means to create order and consistency for knowledge communication. However, it is important to emphasize that categorisation is still a dynamic, ordered structure covered with ambiguity and vagueness (Ahlqvist 2008a). Land-cover and land-use categorisations can be designed that comprise a range of classes valuable for the understanding of processes and patterns of change. Categorisation as such is scale-neutral, i.e. the classes at all levels of the categorisation should be applicable at any scale or level of detail. Application of categorisation will result in a data set comprising only a limited number of classes occurring in a specific area, related to the method and means of observation, organised in a (geo) database and described in the metadata. The spatial and thematic quality of these data should be analysed with statistical means and properly documented. An international consensus on characterisation of land uses is inexistent and it is, therefore, difficult to make existing local or regional (spatial) land-use data consistent. The

availability of tools like remote sensing facilitates observation and collection of land-cover data over large areas in a systematic manner. In practice, various degrees of land-cover aspects are incorporated in land-use class sets and vice versa. Understanding of the land-cover/land-use relationships may improve these data sets and subsequently analysis results. A categorisation is also a means for data standardisation (for new data sets) and data harmonisation (correspondence between existing data sets), as well as being an instrument in contributing to the harmonisation of land-use and land-cover *change*, as we need to understand change processes in order to make informed decisions (McConnell and Moran 2001). Data harmonisation is related to spatial data integration and data interoperability but rarely provides a quantified measure for correspondence between harmonised data sets and the development of such a measure requires further examination.

2. *Scalar dynamics* recognize that land-use and land-cover change observed at any spatio-temporal scale involves complex synergy with changes observed at other analytical scales (Walsh *et al*. 1999; Lambin *et al.* 2000; Hoshino 2001; Veldkamp and Lambin 2001; Veldkamp *et al*. 2001b; Verburg *et al.* 2002; Evans and Kelley 2004; Overmars and Verburg 2006). The scale at which an analysis is conducted may affect the type of explanation given to the observed phenomenon as at each scale different processes have a dominant influence on land use. Furthermore, the scales over which processes operate do not necessarily correspond to the spatial extent of the observation of such processes (Pereira 2002). So the different scales at which land-change processes operate, and the different scales at which they are analysed, pose major impediments to developing a comprehensive understanding. Scales have extent and resolution: extent refers to the magnitude of a dimension used in measuring (e.g., area covered on a map), whereas resolution refers to the precision used in this measurement (e.g., grain size) (Kok and Veldkamp 2001; Verburg *et al.* 2004b). Extent and resolution are mostly linked in observations resulting in studies at large spatial extent that invariably have a relatively coarse resolution owing to our methods for observation, data analysis capacity and costs. While features that can be observed in case studies with a small spatial extent are generally not observable in studies for larger regions. Aggregation of detailed scale processes does not straightforwardly lead to a proper representation of the higher-level process. Changes are often non-linear and thresholds play an important role (Verburg *et al.* 2004b). Different change processes also have different temporal dynamics. Each change process has its own temporal resolution and their interconnection may change over time. The history of land-use change is composed of periods whose within-period change rate is quite stationary but the cross-period change rate is considerably different (Liu and Anderson 2004; Bakker and Van Doorn 2009).

The effects of spatial resolution are relatively well studied; the effect of changes in spatial extent is a dimension of scalar effects that is much less studied. The influence of spatial scale was reported by Verburg and Chen (2000) working at province and national levels in China. Kok and Veldkamp (2001) report that the change in spatial resolution for a case study of five countries in Central America does not greatly influence the analysis, whereas the effect of changing spatial extent is substantial. The key underlying assumptions at the national and supra-national levels are different. With mainly national policies in place, the country level is the largest extent that can be analysed.

Wu and Li (2006) distinguish three *dimensions of scale*: (1) space, (2), time, and (3) organisational levels, or organisational hierarchy, as constructed by the observer. The latter is synonymous with the variation in semantic contents of class sets. Though spatio-temporal dynamics are studied, the variation in the semantic contents of data expressed as differences in categorisation has received very little attention (Feng and Flewelling 2004). This aspect can be regarded as the joint result of categorisation and scalar dynamics. Categorisation produces data sets comprising classes that have different semantic contents (e.g., class labels). So the classes present in data sets and used in change dynamics analysis can also affect the type of explanation given to observed phenomena. Most researchers would probably admit that this might indeed have an impact except that such an impact has never been analysed in a systematic way. Moreover, unaware of what possible influence the variation in semantic contents of class sets may have in change modelling, one may well ask what possible consequences this may have for analysis of preferred pathways and future trajectories.

The semantic contents of land-use/cover (change) data are recently receiving more attention. Measuring semantic similarity of categories, either before or after data collection or between existing data sets, is an emerging area of research (Ahlqvist 2005a). There are various initiatives dealing with the changing context of access to spatial data (e.g., Spatial Data Infrastructures such as the European Union INSPIRE Directive) and the broad recognition that spatial data integration is an essential step in land-change modelling and initiatives (e.g., planning and decision making) that aim to respond to land change (Comber *et al.* 2005a). Increasingly data users become interested in understanding the wider meaning of data, i.e. the concepts adopted and categorisations used. Thus data integration, data interoperability and data harmonisation are linked and underline the importance of categorisation. Semantic interoperability goes beyond attempts to homogenise differences through standards (Harvey *et al*. 1999). Current metadata standards convey nothing about the semantic contents of class sets (Comber *et al*. 2005b; Schuurman and Leszczynski 2006). Especially remotely sensed data derived land-cover products report the technical aspects (e.g., scale, spatial resolution, accuracy) but the meaning of semantic contents is ignored (Comber *et al* 2005c).

The parameterised categorisation approach developed is in particular relevant to Europe. European integration and globalisation processes are accelerating with the extension of the European Union (EU) with 10 countries from Central and Southern Europe in 2004 and two countries from Eastern Europe in 2007[2]. This results in a larger internal market and the challenge to bridge socio-economic differences between Member States. But also global developments like trade liberalisation, population growth and world food demand influence EU policies and the environment (Eickhout *et al.* 2007). All these processes have an impact on the European landscapes: spatial development and planning policies have to keep pace with and attempt to provide some control over these developments (e.g., Common Agricultural Policy, Bird and Habitat Directives, Water Framework Directive). Rural areas comprise about 80% of the European Union and agriculture is the spatially dominant land use. Agriculture plays a key role in the quality of the wider environment (Rounsevell *et al.* 2003; Klijn *et al*. 2005; Van Meijl *et al.* 2006; Verburg *et al*. 2006; Westhoek *et al.* 2006). Surprisingly, perhaps, the greatest density of cropland is in Eastern Europe (Turner *et al.* 2008). Changes like global warming and rising sea levels are likely to result in a long-term impact resulting in a decline of suitable agricultural land (Bullard 2000).

In particular in the Central and Eastern European Countries (CEEC) where Governments implemented a comprehensive package of social and economic reform policies, spatial developments have been rapid in the period of transition from the centrally planned towards a market-oriented economy. Such developments are related to the land-reform choices Governments made in these countries that caused a considerable expansion of individual semi-subsistence holdings (e.g., as a result of decollectivisation) (Swinnen 1999; Kuemmerle *et al*. 2008). These land reforms did not only deal with transfer of property rights and ownership, they dealt with the structures of agrarian economy. Often those who received land were unprepared for their new status as landowners and unfamiliar with becoming independent farmers (Bullard 2000). In the CEEC, greater social equity has been achieved but land fragmentation emerged as a consequence of land reform with detrimental implications for sustainable economic growth and social development in rural areas. Fragmentation obstructs spatial planning in terms of land administration, land-use planning[3] and management. In rural areas the relation to land has profound implications for agricultural productivity, environmental sustainability and the social and economic status of the rural households. In Western Europe land abandonment occurs as a result of EU policies, in the CEEC it is often due to lack of resources to farm profitably. Matching land use and land tenure with the aim to reach a

---

[2] From 1 January 2007 onwards the EU comprises 27 Member States. Before 1 May 2004 the EU consisted of 15 countries (Austria, Belgium, Denmark, Finland, France, Germany, Greece, Italy, Ireland, Luxembourg, Netherlands, Portugal, Spain, Sweden and United Kingdom,). On 1 May 2004 10 countries joined (Cyprus, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Malta, Poland, Slovakia and Slovenia). On 1 January 2007 Bulgaria and Romania joined.

[3] Land-use planning is defined as "*the systematic assessment of land and water potential, alternative patterns of land use and other physical, social and economic conditions, for the purpose of selecting and adopting land-use options that are most beneficial to land users without degrading the resources or the environment, together with the selection of measures most likely to encourage such land use*" (FAO, 1993).

better socio-economic structure therefore becomes crucial (Larsson 2002). Nowadays urban centres rather than rural areas have become the dynamic engines of development and rural areas depend largely on the rural-urban linkages (Lambin *et al.* 2001; Riddell and Rembold 2002). For the illustration of the various aspects of the methodological approach mainly examples from Europe (Nordic countries and the CEEC) have been selected with emphasis on the developments in rural areas.

1.2 Objectives

The main objective of this thesis is an improved understanding of how class sets using a parameterised approach can contribute to the improved understanding of the spatio-temporal and semantic dimensions of land-change dynamics. Basically these are the three *dimensions of scale* mentioned by Wu and Li (2006). In particular, the semantic dimension of class sets is highlighted in categorisation, data harmonisation and standardisation, change detection, in addition to how the variation in semantic contents of class sets might influence model dynamics.

The main research questions are:

1. Can with the use of a parameterised approach the categorisation of land cover and land use be improved resulting in comprehensive data sets and time series that contribute to and that are functional in the understanding of land-change dynamics?
2. Is harmonisation of land cover and land use feasible and facilitated with a parameterised class set as bridging or reference system; can harmonisation of change be achieved?
3. In particular modifications are infrequently captured in land-change studies. Can parameterised class sets contribute to the analysis of the spatio-temporal and semantic dimensions of change dynamics and change processes such as conversions and modifications?
4. How does variation in semantic contents of class sets influence land-use modelling dynamics and what are the consequences for the analysis of preferred pathways of change and future trajectories (e.g., from a policy or decision-making point of view)?

1.3 Research methodology

In land-change research broad categories can be distinguished (Rindfuss *et al.* 2003; Turner *et al.* 2008):

- Observation and monitoring of land change (including remote sensing, categorisation systems, quantification of changes in the past);
- Understanding of these changes as a coupled human-environment system through identification of drivers of change and the factors that determine the land-use pattern to describe causal processes (e.g., socio-economic and biophysical driving forces); and
- Spatially explicit modelling of land change with computer models that enable combining categories 1 and 2 in a dynamic and integrated manner.

In addition, Turner *et al.* (2008) distinguish assessment of system outcomes, such as vulnerability, resilience, or sustainability.

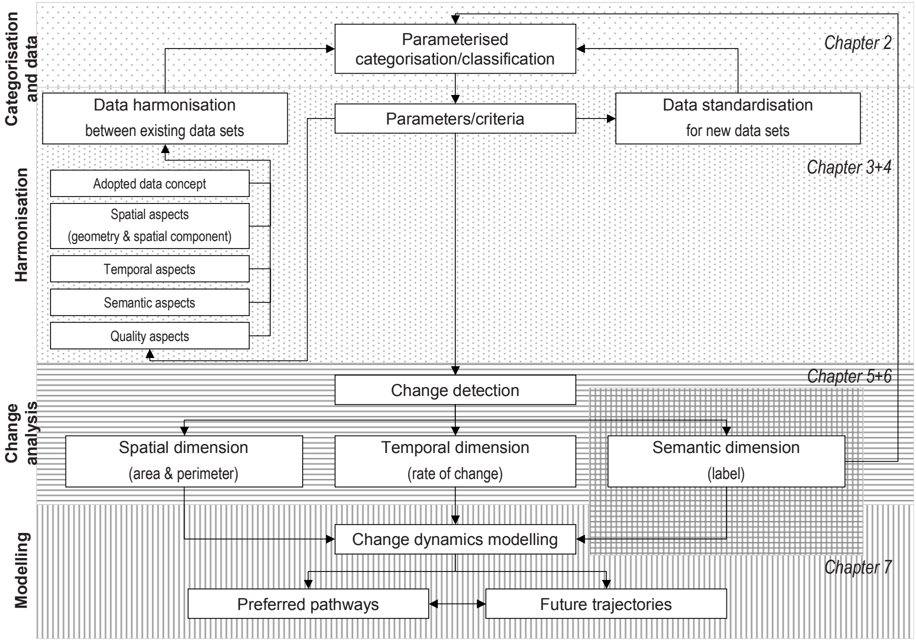
The first three categories are discussed below following the sequence of: (1) categorisation and data (paragraph 1.3.1), (2) harmonisation (paragraph 1.3.2), (3) land-change analysis (paragraph 1.3.3) and (4) modelling dynamics (paragraph 0). The identification of drivers of change is limited and presented only in connection with either change detection or modelling. An assessment of system outcomes is not included in this thesis.

1.3.1 Categorisation and data

Many land-cover and land-use description systems exist throughout the world but none is internationally accepted. Many systems are limited in the features they describe, they are inconsistent in the criteria used at various levels and they commonly mix inherent criteria with non-inherent criteria. The term 'land use' and 'land cover' have different meanings across disciplines and, as a result, imply a set of mostly unidentified parameters. Consequently, a common terminology is lacking. These different perspectives on land use and land cover are, however, all valid as the multi-dimensionality of geographical categories is a property that reflects human cognition of geographical phenomena. Different perceptions by different disciplines show that it is important to know the concepts behind class sets. These notions were first expressed in the 1950s by Guttenberg (1959 and 2002) and Shapiro (1959).

In the 1990s, different groups worked on the development of universally applicable land-cover and land-use categorisations that would contribute to standardisation of the criteria used for description and consequently categorisation, in addition to harmonisation between existing data sets (UNEP/FAO 1994; Wyatt *et al.* 1994 and 1998; LANES 1998) plus harmonisation of change (McConnell and Moran 2001). Emphasis was on the selection of a parameterised approach, i.e. using a set of explicit independent criteria resulting in a flexible data set that can be used as a uniform basis for description, in addition to the use of (part of) these parameters for land-change detection and monitoring (Figure 1-1). Thus not class labels are vital but the applied explicit set of parameters.

Figure 1-1. Categorisation and data, harmonisation, change analysis and modelling taking a parameterised approach



An *overarching concept* for a universally applicable system for land-cover categorisation and spatially (geographically) explicit data collection based upon a *structural-physiognomic approach* was developed by FAO after analysis of existing systems (Danserau 1961; Fosberg 1961; Eiten 1968; UNESCO 1973; Mueller-Dombois and Ellenberg 1974; Kuechler and Zonneveld 1988; UNEP/FAO 1994; CEC 1999). Explicit definition of the overarching concept, or *Weltanschauung*, assists users in understanding the concepts of the categorisation system and the meaning of classes. All land covers can be described though the level of detail for vegetated and cultivated areas is more elaborated than that for bare areas and water bodies.

A conceptual approach for land-use categorisation and spatially explicit data collection was developed based upon analysis of existing class sets and systems in various sectors (Guttenberg 1959 and 1965; Urban Renewal Administration 1965; Anderson *et al.* 1976; IGU 1976; Kostrowicki 1977, 1983a, 1983b, 1992a and 1992b; UN-ECE 1989; UN 1989; CEC 1993; UN 1998; Duhamel 1998; Wyatt *et al.* 1998; APA 1999). The overarching concept combines two key criteria: *function*, grouping all land used for a similar purpose, with *activity*, grouping all land undergoing a certain process resulting in a homogeneous type of products (UN 1989). The 'function' approach relates both to the intended (or primary) and unintended (or secondary) land-uses.

1.3.2 Harmonisation

Apart from developing new systematic categorisation frameworks for land cover and land use respectively, a link had to be built to existing class sets in order to analyse time series for the detection of changes in the environment. Spatial data integration concerns many aspects of data harmonisation. The latter being defined as "*the intercomparison of data collected or organised using different classifications dealing with the same subject matter*" (McConnell and Moran 2001). Data harmonisation will allow countries and institutions to continue to use existing data systems and class sets but when definitions are imprecise, ambiguous or absent problems may arise. Moreover, if many class sets are involved the number of pair-wise class combinations becomes excessive because comparison of n data sets requires $n(n-1)/2$ comparisons to be made.

Data harmonisation is based upon (Figure 1-1):

- The *data concept adopted* that includes the worldview embodied in the class set and its wider meaning;
- The *spatial aspects* that comprises the geometry (e.g., a description of the form of the entities through geographic primitives or through a structured geometry (e.g., topology)) and a spatial component (e.g., zero, one, two of three dimensions for point, line, two-dimensional or three-dimensional areas respectively);
- The *temporal aspects* (e.g., for geological data a different timeframe is applicable compared to land cover or land use that change more frequently);
- The *semantic aspects*, i.e. the label attached to the feature denoting the categorisation framework; and
- The *quality aspects* that concern all the above-mentioned aspects in order to produce quality data. Metadata may provide some measure of positional and thematic accuracy, or uncertainty, but many other measures of quality should be specified as they may limit interoperability but that currently are excluded from metadata requirements.

Often existing data need to be harmonised and/or regrouped before detection and analysis of changes can be performed but there is neither a standardised methodology for data harmonisation nor a standard means for quantifying the quality of harmonisation results though various techniques have been examined and research is progressing (Rodríguez *et al.* 1999; Wyatt and Gerard 2001; Kavouras and Kokla 2002; Ahlqvist 2005b; Comber *et al.* 2004a and 2005a; Wadsworth *et al*. 2006).

Data standardisation is defined as "*the use of a single standard basis for classification of a specific subject*" (McConnell and Moran 2001). It allows direct comparison of class sets but disregards the financial and intellectual investments made in established methods and data sets. Therefore, data standardisation seems more appropriate when dealing with new data sets. However, it assumes that the advances in knowledge, technological developments and changing policy objectives would not have an impact on the existing systematic categorisation framework. Lessons from the past impart the message that this may be an unrealistic expectation, so data standardisation seems to be feasible only in part.

Land cover and land use have three major semantic differences that affect their interoperability (Brown and Duh 2004) and these refer to the above-mentioned aspects of data harmonisation (paragraph 1.3.1):

- The *category definitions* of land cover and land use are different (e.g., 'undeveloped forest' is a clear-cut area that continues to be used for forestry (Lund 1999));
- Land cover and land use have different *geometric expressions* consequently a categorisation cross-walk approach to semantic interoperability, which defines interrelations between categorisation schemes without redefining spatial objects and like that proposed and implemented for alternative vegetation/land-cover classifications by the IGBP (Loveland *et al*. 2000), may be an inadequate solution for translation between land use and land cover (as Cihlar and Jansen (2001) pointed out: the spatial objects might need to change in addition to the class definitions); and
- Land cover and land use have different *spatial rules* to assign attributes to land-use/cover features because land-use class definitions tend to integrate information about activities taking place within a spatial unit (e.g., cadastral parcel or zone), while land-cover class definitions assess the static and *in situ* conditions. Thus, the entities of a land-cover data set (e.g., polygons) usually show more spatial variation than those of a spatially explicit land-use data set (assuming both data sets are compiled based on sources of the same level of detail).

1.3.3 Land-change analysis

The above-described categorisation concepts (paragraph 1.3.1) are geared towards identification of the two main types of changes indicating the type of process taking place and enabling their detailed description and in-depth analysis: *conversions* where evident changes occur that cannot be (easily) reversed and *modifications* where changes can be reversed. These two processes are driven by the interaction in space and time between biophysical and human dimensions (Turner *et al*. 1995).

Change dynamics manifesting themselves at three different dimensions of scale can be analysed in terms of (Figure 1-1):

- Changes in geometry (area and perimeter), i.e. spatial dimensions x and y;
- Changes in the rate of change, i.e. temporal dimension t; and
- Changes in class label, i.e. semantic dimension s that in a parameterised approach may range from a change in the composition of characteristics measured to a change in any of the measured characteristics. The semantic dimension relates to the constructed organisational hierarchy.

Scale is a central issue in land-change dynamics. Not only the three dimensions of scale above-mentioned are crucial. According to Wu and Li (2006) change dynamics manifest themselves also at different *kinds of scales*, i.e. the observational scale (e.g., polygon versus cadastral land parcel) and policy scale (e.g., national and district level versus commune level), and the *components of scale* in the land-cover and land-use change analyses are also different as there are:

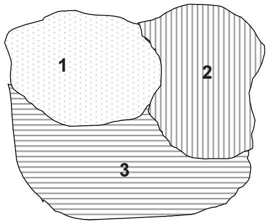
1. Differences in *cartographic scale* (e.g., land cover at 1:100,000 versus land use at 1:2,500);
2. Differences in *grain* (e.g., minimum mapping unit versus cadastral land parcel unit);
3. Differences in *extent* (e.g., country-wide area coverage versus the pilot area coverage); and
4. Differences in *coverage*, i.e. sampling intensity (e.g., random stratified sampling of land-cover polygons versus sampling of every cadastral land parcel in a commune).

The parameterised categorisation approach is combined with an object-oriented database approach for land-change analysis. In most change studies the state of land cover or land use at a certain point in time is compared to the state at a later moment thereby focussing on a representation of temporal data in which snapshots (e.g., remote sensing interpretations) are created for each moment in time. The land-cover state at t_1 is overlain with t_2, t_2 with t_3, t_3 with t_{3+k}, etc. The result of such overlays (in raster format) is a sequential representation of the dispersion of a class in other classes and such a representation usually does not contain the immediate link between the spatial and temporal dimensions of the changes that occurred between t_1 and t_{1+k}. However, an object-oriented approach to databases also documents the relationships between identified states and the processes that led up to these relationships.

The object-oriented approach in information systems is defined as the "*collection of co-operative objects, treating individual objects as instances of a class within a hierarchy of classes*" (Booch 1994). Figure 1-2 shows an example of such an approach using the relation 'Object.GetParent_ID (byValue)' to establish parent-child relationships between the polygons.

Figure 1-2. An example of an object-oriented approach to databases

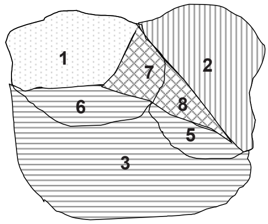
1996



ID	P_ID	Obj	Area	Year
1		A	200	1996
2		B	250	1996
3		C	400	1996

Object.GetParent_ID(byValue)

2003



ID	P_ID	Obj	Area	Year
1		A	200	1996
2		B	250	1996
3		C	400	1996
5	**2**	**C**	**75**	**2003**
6	**1**	**C**	**50**	**2003**
8	**2**	**D**	**75**	**2003**
7	**1**	**D**	**50**	**2003**
2		B	120	2003
3		C	400	2003
1		A	100	2003

1.3.4 Modelling dynamics

Land-use/cover change modelling is an important technique for the projection of future pathways of change and for conducting experiments that enhance our understanding of key processes and for describing the latter in quantitative terms (Turner *et al*. 1995; Lambin *et al.* 1999 and 2000; Veldkamp and Lambin 2001; Lambin and Geist 2006). Much of the integration of knowledge on land-use change takes place through (spatial) models that aim at explaining the causes, locations, consequences and trajectories of land-use change (Verburg and Veldkamp 2005). Models, like categorisations, simplify the complexity of the reality. Modelling efforts comprise general equilibrium models and optimisation models (Fischer and Sun 2001; Rounsevell *et al.* 2003; see Van Tongeren *et al.* (2001) for a review of global models), agent-based or multi-agent models (Bousquet and Le Page 2004; Evans and Kelley 2004) and spatially explicit macro-models using cellular automata (White and Engelen 1997; Wu and Webster 1998; Torrens 2002), artificial neural networks (Zhou and Civco 1996) or regression techniques (Ives *et al*. 1998; Pontius *et al*. 2001; Verburg *et al*. 2002). The two main sets of drivers (social and biophysical) each seem to play a prevalent role at a specific scale: social drivers are associated with finer-scale spatial patterns and biophysical drivers with coarser-scale spatial patterns. As a result two complementary approaches are advocated: from pattern-to-process (top-down) and from process-to-pattern (bottom-up) (Walsh *et al.* 1999; Laney 2004; Castella *et al.* 2007; Castella and Verburg 2007; Overmars *et al.* 2007).

The spatial organisation in most models is a geographical grid with cells of a certain size (e.g., in the EURURALIS study the grid cell size is 1x1km (Verburg *et al*. 2006 and 2008)). These grid cells offer convenient uniform units of analysis that can be aggregated to a series of higher artificial aggregation levels, considered proxies for different scales (De Koning *et al*. 1999). The grid-based data format facilitates 'harmonisation' of data from different sources and different formats (e.g., spatial polygon-based versus census-based data). This type of data rearrangement is a form of data harmonisation as it is related to the spatial aspects discussed in paragraph 1.3.2. This entails that modelling would be inevitably linked to harmonisation.

To make a realistic evaluation of the variation in semantic contents of class sets -expressed as differences in categorisation- on modelling dynamics, the drivers of change in the study areas have been identified. The empirically determined relations between land use and its driving factors are subsequently used for the simulation of future land-use changes. The CLUE (Conversion of Land Use and its Effects at Small regional extent) model is used in this systematic evaluation. The model has been thoroughly tested and validated in several parts of the world (Pontius *et al*. 2008). Furthermore, the spatial and temporal dimensions of land-use change have been explored with this model and therefore the examination of the variation in semantic contents of class sets is complementary to the earlier research. CLUE combines an empirical analysis with dynamic, multi-scale simulations to be able to handle different scenario conditions that may deviate from the historic trend (Veldkamp *et al.* 2001b; Verburg *et al.* 2002). The latter is important in the Central and Eastern European Countries, as the historical trend in the period of transition has been disrupted.

1.4 Outline

This thesis is a collection of research papers describing a methodology developed over a number of years in different institutional settings and project environments. Progress in methodology development and its applications was examined in operational projects and reported subsequently not just in technical project reports but also in peer-reviewed papers published in international scientific journals or books. The examples used to illustrate the various aspects of the developed methodology are actually not 'case studies' but 'real life cases' though the former wording will be used.

In Part I (Chapter 2), the concepts of parameterised systems to be used in land-change detection are described. The Land-Cover Classification System (LCCS) categorisation methodology has been tested, modified and validated in several international projects in order to evaluate its applicability in different environmental settings, its use at different data collection scales and with different means of data collection, its usefulness for data harmonisation and in land-cover change analysis. LUCC has endorsed the methodology (McConnell and Moran 2001). LCCS has been applied by the European Commission's Global Land Cover 2000 (GLC2000) project (Mayaux *et al.* 2004 and 2006), the FAO Africover project (10 countries) (Kalensky 1998) and FAO projects in the CEEC and the Commonwealth of Independent States (CIS) (e.g., Azerbaijan, Bulgaria (Travaglia *et al.* 2001), Romania and Moldova), in projects financed by the Italian Ministry of Foreign Affairs (e.g., Niger (Jansen *et al.* 2003c), The Gambia, Mozambique (Jansen *et al.* 2008a) and Senegal (Jansen and Ndiaye 2006)), by the Nordic Council of Ministers' Nordic Landscape Monitoring project (e.g., Estonia, Denmark, Norway and Sweden) (Groom 2004), and by a World Bank financed project in Albania (Jansen *et al.* 2006a). The concepts of a parameterised approach to the categorisation of land use are described that were tested by FAO in a pilot study in Kenya (Jansen and Di Gregorio 2003) and in a more advanced version by an EU PHARE Programme project in Albania (Jansen 2003 and 2006). The results of some of these projects are used in successive chapters to illustrate the developed methodologies.

In Part II (Chapters 3 and 4), the various issues concerning harmonisation of class sets are described for land cover and land use. For land cover, the various aspects of harmonisation are discussed with emphasis on methods to assess semantic similarity between classes (Chapter 3). An example is provided from the Nordic Landscape Monitoring project in which five class sets from the Nordic countries were 'translated' (in GIS the term 'reclassified' is often used though this is also the case when regrouping classes and this is not meant here) with the use of LCCS as reference system in which the original classes find their more or less corresponding class. The computation of semantic similarity between classes in a special module of LCCS is critically reviewed. For land use, the description for a parameterised land-use categorisation concept has been applied in a EU PHARE Programme project in Albania (Chapter 4). This case study is used to illustrate how a parameterised land-use class set can contribute to harmonisation, thereby contributing to harmonisation of land-use *change*. Both case studies further illustrate the necessity to develop measures to assess the quality of harmonisation results, i.e. a measure to express the accuracy of correspondence.

In Part III (Chapters 5 and 6), the use of parameterised land-cover and land-use class sets -comprising the categorisation concepts above-described- for change analysis is described. In the context of decision making the micro to meso-level dimension (e.g., the individual land user, community) tends to be the most appropriate, whereas biophysical driving forces are meso to macro-level processes expressing themselves in different ways across spatio-temporal scales (McConnell and Moran 2001). The results of two projects in Albania are used: (1) a country-wide analysis of land-cover change based upon remote sensing that provides an insight in the land-cover conversions and modifications at national and district levels in the period of transition (World Bank financed project) (Chapter 5); and (2) a cadastral-parcel based analysis of land-use changes at commune level in the period of transition using an object-oriented database approach (EU PHARE Programme project, see also Chapter 4) (Chapter 6). The latter includes an analysis of preferred pathways of change using cellular automata and the data mining technique. Spatial land-use data collection can also make use of multi-purpose cadastres that should be seen as an integral part of land management systems. The smallest spatially explicit land unit, the land parcel, is a legal unit subject to land-use policies and it can be adopted as the basic unit for data modelling and change analysis. The use of cadastral information for urban planning is an established practice, whereas it is much less common in rural land-use planning. The two case studies illustrate clearly that at aggregated data levels the local variability of spatially explicit land-cover/use changes may be obscured, whereas patterns can be shown that at more detailed data levels may remain invisible, and vice versa (Veldkamp *et al.* 2001b). Thus, there is an obvious need to make complementary multi-scale analyses in order to detect the change dynamics at different levels.

In Part IV (Chapter 7) the evaluation of the variation in semantic contents of class sets in model dynamics is discussed. Parameterised LCCS land-cover data from FAO projects in Romania and the elaborated EURURALIS land-cover/use data set were used in combination with the potential driver data from the EURURALIS project. Three levels of variation in semantic contents in the LCCS data set were used to analyse what drivers would best explain a certain land-cover type. The similarities and differences of those three levels of variation in semantic contents are discussed. The elaborated EURURALIS land-cover/use data set and resulting drivers of change is then compared to the results of the LCCS land-cover data sets at various levels of organisational hierarchy. A link with the validation of remotely sensed derived data is made. The use of different hierarchical organisations in the land-cover data investigated has implications for land-change analysis in Romania. In this chapter a land-use data set is not investigated.

Chapter 8 concludes this thesis with some remarks on the implications of the results for land-change analysis and indicates directions for future research.

PART I - CATEGORISATION AND DATA

To classify is human.

G.C. Bowker and S.L. Star

Classification is easy:
it is something you just do.

F.C. Bawden

Abstract

Systematic description of the environment for detection of land changes and the human-related causes and responses is essential in land-cover and land-use change studies. The combined use of land-cover and land-use data allows detection of where certain changes occur, what type of change, as well as how the land is changing. Existing systems for categorisation, or classification, of land cover or land use are limited in the storage of the number of classes and are often internally inconsistent. Therefore, the Land-Cover Classification System, a comprehensive parameterised categorisation system based upon systematic description of classes with the use of a set of independent quantifiable diagnostic criteria according to an overarching concept was developed. With this approach land-cover change detection becomes possible at the level of conversion of a class, whereas modification within a certain class type becomes immediately identifiable by a difference in parameter, or through the use of additional parameters as is shown in a series of examples illustrating the application of the approach to primarily vegetated areas. The development of a similar categorisation approach for land use is still a need. The proposed approach here combines function, grouping all land used for a similar economic purpose, with activity, grouping all land undergoing a certain process resulting in a homogeneous type of products. These concepts have been tested in several applications that have shown that such a categorisation system can be used as a bridging system that will ensure compatibility with, and bridge, existing systems. Furthermore, by providing (part of) the diagnostic criteria the system contributes to providing a uniform basis for land-change detection and these criteria contribute, in turn, to standardisation. Land-cover boundaries do not necessarily coincide with land uses and the land-cover/land-use relationship needs more study to understand its complexity.

Based on: Jansen, L.J.M., Di Gregorio, A., 2002. Parametric land-cover and land-use classifications as tools for environmental change detection. *Agriculture, Ecosystems & Environment* 91(1-3): 89-100.

2 Parameterised class sets as tools for land-change analysis

2.1 Introduction

The understanding of the interactions between land cover and land use in their spatial and temporal appearances is fundamental to comprehension of land-use and land-cover change. Land cover is an expression of human activities and, as such, changes with changes in land use and management. Hence, land cover may form a reference base for applications including forest and rangeland monitoring, production of statistics for planning and investment, biodiversity, climate change and desertification control (Di Gregorio and Jansen 1998). Detection of changes in the environment and the human-related causes and responses may be used to predict changes and project future trajectories. Land-use and land-cover change is a multi-disciplinary subject where bio-physicists and socio-economists meet one another. It is important that an integrated approach is taken with multiple partners involved to come to a widely accepted reference base for land-use and land-cover categorisation.

Without categorisation, phenomena would remain merely a bewildering multiplicity and the precise and unambiguous communication of ideas and concepts concerning these phenomena would be impossible. Categorisation of relevant phenomena is essential if generalisations are to be made concerning these phenomena. The prime interest is in general truths, i.e. truths related to classes or kinds rather than to their individual members. A truth discovered about such a member is always implicitly applied to the entire group to which the member in question belongs. Without categorisation such generalisations would also be impossible. And, finally, the evolution of a body of reliable knowledge concerning any set of phenomena through the process of accumulation would be extremely difficult without categorisation (Shapiro 1959).

Categorisation, or classification, is defined as "*the ordering or arrangement of objects into groups or sets on the basis of relationships. These relationships can be based upon observable or inferred properties*" (Sokal 1974). Another, and even earlier, definition by Shapiro (1959) reads "*the sorting of a set of phenomena composed of generally-alike units into classes or kinds, each class or kind consisting of members having definable characteristics in common*" is also interesting but does not underline the importance of relationships. It is important to note that categorisation is an abstraction in the sense that it depicts a representation of the reality (Di Gregorio and Jansen 2000).

With the combination of land cover and land use, change detection will provide the location of occurring changes and type of change, as well as the manner in which the land is changing. Change detection is of course related to the time of observation, the time span, the means and methods of observation. The scope of this chapter is the concept of change detection and the contribution parameterised categorisation systems can make by providing (part of) the diagnostic criteria. With the standardisation of the diagnostic criteria, such a categorisation system would provide a uniform basis for change detection.

2.2 Land-cover change detection

Land-cover change detection has to recognize that changes take two forms:

- *Conversion* from one land-cover category to another (e.g., from forest to grassland);
- *Modification* within one category (e.g., from rainfed cultivated area to irrigated cultivated area).

These two forms of change have implications for the methodology used to describe and classify land cover. Conversion implies an evident change, whereas modifications are much less apparent. The latter requires a greater level of detail to be accommodated.

The broader and fewer the categories used to describe land cover, the fewer the instances of conversion from one to another. If land-cover classes are as broad as 'Forest and woodland' and 'Permanent meadows and pastures' as in the FAO Production Yearbook (FAO 1990-1995), then forest fragmentation and changes in species cover composition due to overgrazing (e.g., bush encroachment) will not register as conversion. Conversion is reasonably well documented in change studies as in the FAO Forest Resources Assessment (FAO 1996). If conversion totals alone are used to measure change, it may occur that apparently no land-cover change appeared at all (Meyer and Turner 1992).

Broad categories cannot be used to measure any type of modification. If stands of single species (monocultures) replace stands of multiple species, the category 'Arable land' in the FAO Production Yearbook will not allow registration of such a change. Contrary to conversion, modification is not as well studied and at the global scale, often ignored. The ecological consequences, however, are as important in the case of conversion as in the case of modification. The subtle changes of modification do not always result in degraded ecosystems (Turner *et al.* 1995); it is a result of evolutional aspects. The increase in planting densities on cropped land is an example of modification that does not lead per se to degradation. Intensification of cultivation may lead to a longer period in which crops cover the land, however this will be highly localised and difficult to observe.

Categorisations are not just limited to an analysis of the current situation but categorisation results can be instrumental in understanding and changing existing circumstances. The interest in global and regional land-cover and land-use change (e.g., IGBP-IHDP Land-Use and Land-Cover Change Project and its successor the Global Land Project) requires a set of tools that will enable meaningful comparisons. Categorisation systems can be tools for change detection when they offer the capability to describe classes through a set of well-defined independent diagnostic criteria (parameters), which allow building up these classes, rather than being based upon the traditional system using descriptive class names without explicitly mentioning the criteria used. The parameters should be independent from the land in order to analyse changes related to either the bio-physical features, i.e. land cover, or to the human use of the land, i.e. land use, or both.

2.3 Existing categorisations

Traditional categorisations dealing with land cover and/or land use (Danserau 1961; Fosberg 1961; Trochain 1961; Eiten 1968; UNESCO 1973; Mueller-Dombois and Ellenberg 1974; Anderson *et al.* 1976; Kuechler and Zonneveld 1988; UN-ECE 1989; UNEP/FAO 1994; CEC 1995; Duhamel 1995; Thompson 1996) are limited in their capacity of storage of classes and often do not contain the whole variety of occurring land covers or land uses. Some describe (semi-) natural vegetation in great detail while accommodating cultivated areas in a single class or vice versa. More important, they are based upon the approach of class names and class descriptions that do not consistently use a set of criteria to make class distinctions (Jansen and Di Gregorio 1998a). Furthermore, the criteria used are often not inherent characteristics but describe the environmental setting of the land cover and land use, respectively. The distinction between land cover and land use is not always appreciated or adhered to in the above-mentioned categorisations.

The above is demonstrated using the CORINE Land Cover (CEC 1995) and UN-ECE Standard International Classification of Land Use (1989) as examples.

2.3.1 Inconsistent application of land-cover or land-use criteria

The application of land-use criteria in a land-cover categorisation, or land-cover criteria in a land-use categorisation, results in a system that fails to make a clear distinction between land cover and land use. In CORINE the classes ‘1.2.1. Industrial or commercial zones’, ‘1.4.2. Sports and leisure facilities’ and ‘2.3.1. Pastures’ each contain a set of different land-cover types for a specific land use. In the UN-ECE, a land-use categorisation, the distinction between the classes ‘2.1 Land under coniferous forest’ and ‘2.2 Land under non-coniferous forest’ is based upon land cover.

2.3.2 Inconsistent use of criteria at same level of categorisation

The inconsistent use of criteria at the same level within one major class. In CORINE, ‘2.1. Arable land’ is subdivided at the third level into ‘2.1.1 Non-irrigated arable land’ and ‘2.1.2. Permanently irrigated land’ using the practice of irrigation as criterion, whereas ‘2.1.3. Rice fields’ is distinguished from the crop species. It is not explained why the criterion to distinguish the latter has the same weight as irrigation. In the UN-ECE, the ‘5. Dry open land with special vegetation cover’ class is further subdivided into ‘5.1. Heathland’, ‘5.2 Dry tundra’ and ‘Mountainous grassland’ using criteria related to the vegetation type, vegetation type plus absence of water, and vegetation type plus landform. It is not explained why two of the three classes are distinguished using additional criteria to vegetation type.

2.3.3 Use of different criteria between classes

The use of different criteria between related major classes. In CORINE, ‘3.1 Forests’ are distinguished at the third level based upon the criteria of leaf type, whereas the related class ‘3.2 Scrub and/or herbaceous associations’ only uses the same criterion for distinction of ‘3.2.3 Sclerophyllous vegetation’. The other third level classes use their state (‘3.2.1 Natural grasslands’ and ‘3.2.4 Transitional woodland-scrub’) or vegetation types (‘3.2.2 Moors and heathland’) as criterion. One may wonder why these criteria are introduced at this level. The same occurs in UN-ECE where ‘1. Agricultural land’ is further subdivided using criteria related to cropping and the period during which the crops cover the land, whereas ‘2. Forest and other wooded land’ are subdivided at the third level according to major function. A user may question why the function criterion was not applied to agriculture.

2.3.4 Use of non-inherent characteristics

The use of characteristics that are not inherent to land cover: ‘5.2 Marine waters’ is subdivided at the third level into ‘5.2.1 Coastal lagoons’ and ‘5.2.2 Estuaries’ using geomorphologic criteria. The same occurs in the UN-ECE where ‘7.2 Tidal waters’ are subdivided into ‘7.2.1. Coastal lagoons’ and ‘7.2.2. Estuaries’. In both categorisations it is not explained why exogenous criteria are introduced for further division of water.

Although the underlying reasons for making the subdivisions based upon different criteria, described in the previous paragraphs, may be valid, they show that criteria do not always have the same weight in making distinctions. Such decisions are usually not well documented in the accompanying reports of the categorisations. It will be difficult for any user to trace back the origins of these unsystematic descriptions and to re-interpret the class descriptions and criteria used in the absence of sufficient documentation. This hampers harmonisation of categorisation results, as these interpretations are likely to differ between persons within one country and between countries. The actual categorisations make an insufficient contribution to data standardisation and harmonisation. Efforts to increase standardisation and harmonisation do not necessarily lead to loss of pragmatic decisions on the choice of criteria, as the focus should be on the logical and functional consistent application of a set of inherent criteria that are clearly separated from non-inherent criteria.

2.4 Criteria for change detection and categorisation systems

Change detection should be established on a sound base. Preferably, this base should use a common reference system established upon objective measurable and replicable criteria. At present, data collection and compilation is often accomplished for one single purpose, thereby limiting the use of the products to those that have a similar aim. Data collection is, however, time consuming and expensive.

The existing categorisation systems are limited in the number of classes thereby restricting the possibilities for change detection. Elements used at a high level of aggregation are few, thus, only this set of diagnostic elements can be considered for monitoring. A parameterised approach can be a tool for change detection because it describes classes through a set of independent quantifiable diagnostic criteria, rather than being based upon descriptive class names. The individual parameters provide, consequently, the elements for monitoring change. Those parameters should be selected on the basis of objective measurement and only the number of parameters in the system would limit the elements to be monitored. At the same time they would be standardised parameters contributing to harmonisation of criteria used for change detection.

The compilation of data sets has been greatly facilitated by Geographical Information Systems (GIS) owing to the minimum aggregation and maximum flexibility. However, the ease with which data sets are compiled did not contribute to being useful to a wider public and to a more diverse range of applications. To date, most databases still take a sectoral approach. In order to facilitate exchange of available data and to permit a comprehensive assessment of land-cover and land-use change on a uniform basis, data should be converted to a common basis that allows correlation and comparison. The different applications may diverge from this common reference base and add more specific purpose-related criteria.

A parameterised approach makes the criteria for categorisation more explicit than any traditional system, as well as the consistent application of selected criteria. The set of diagnostic criteria should be limited to those identifying a certain object and distinguishing it from other objects, but additional criteria may be used to add more detail to the description of the object or even describe the environmental setting of the class. The latter type of attribute should be clearly distinguishable from those that describe inherent characteristics. The use of explicit criteria will enhance comparison of change statistics by providing the set of parameters that may be analysed and monitored. These parameters can be measured by field observation, census and/or statistical methods.

Criteria that are linked to the means or scale of observation should never be included in a categorisation system as it is by definition independent of tools (e.g., spectral reflectance characteristics are related to the satellite instrument and the pixel size will determine what features can be detected).

A further requirement of a categorisation system is the applicability at various scales, from global, regional, national to local, i.e. being accommodated by having different levels of aggregation in the system. At each level the set of criteria applied will be different and criteria used at higher levels should not be repeated at lower levels. Original data should always be maintained to allow full desegregation, return from boundaries to gradients and, if necessary, reclassification of the original data for other purposes.

By increasing the number of classes in a system to be able to accommodate any land cover or land use occurring anywhere in the world, the problem of determination of clear class boundary definitions arises, as they will be based on very slight differences. The wrong, or different, designation of the same feature to different classes will affect the standardisation process, a principal objective of categorisation. The attempt to harmonise categorisation results will fail if the diagnostic criteria are not determined in a clear, meaningful and unambiguous manner (e.g., the meaning of a parameter may change with the type of environment). Therefore, strict and unambiguous class boundaries are a prerequisite. Furthermore, classes should be as neutral as possible in the description of a land-cover feature in order to answer the needs of a wide variety of end-users.

2.5 Land-Cover Classification System

2.5.1 Conceptual approach

The set of diagnostic criteria for the parameterised categorisation approach followed in the Land-Cover Classification System (LCCS) is based upon examination of criteria commonly used in existing categorisations that identify and describe land cover in an impartial, measurable and quantitative manner (FAO 1997; Di Gregorio and Jansen 1998 and 2000; Jansen and Di Gregorio 1998a). However, the definition of categorisation provided in FAO (2005) "*classification is an abstract representation of the situation in the field using well-defined diagnostic criteria: the classifiers*" confuses categorisation with an abstract representation of a categorisation example given in Kuechler and Zonneveld (1988) and completely overlooks the fact that categorisation is the basic cognitive *process* of arranging objects into classes or categories as well as the *act* of distributing objects into classes or categories of the same type.

The developed approach to categorisation aims at a logical and functional hierarchical arrangement of the parameters, thereby accommodating different levels of information, starting with broad-level classes that allow further systematic subdivision into more detailed subclasses. At each level the defined classes are mutually exclusive. Criteria used at one level of the categorisation are not to be repeated at other levels. The increase of detail in the description of a class is linked to the increase in the number of parameters used. In other words, the more parameters are added, the more detailed the class. The class boundary is then defined either by the different number of parameters, or by the presence of one or more different types of parameters. Emphasis is not given to the derived class name, the traditional method, but to the set of parameters used to define this land-cover class.

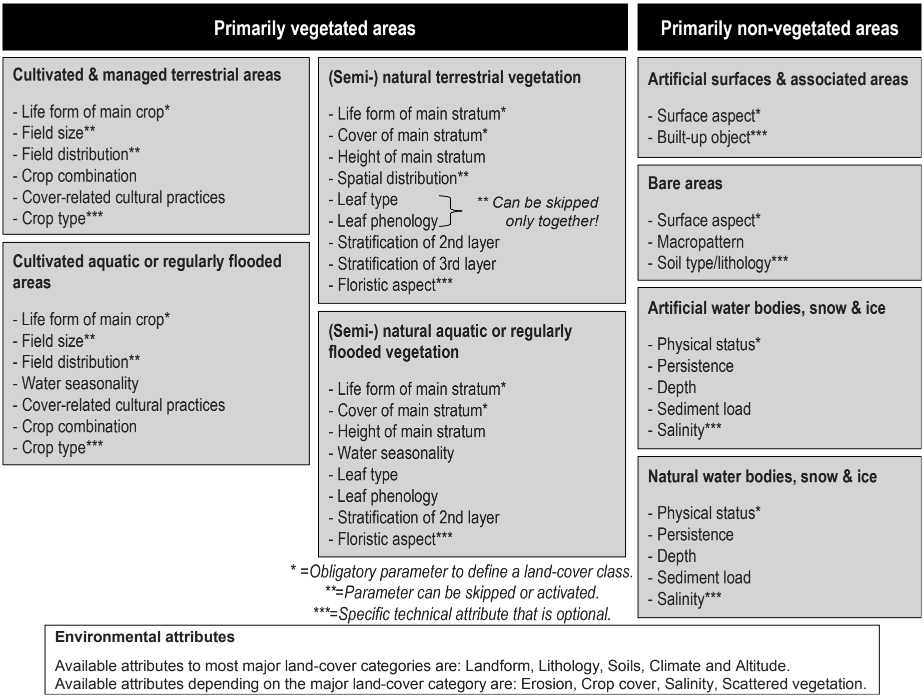
Many current categorisation systems are not suitable for mapping and subsequently monitoring purposes. In the developed parameterised approach, the use of diagnostic criteria and their hierarchical arrangement to form a land-cover class, are a function of geographical accuracy. The arrangement of parameters will assure at the highest levels of the categorisation, i.e. the most aggregated levels, a high degree of geographical accuracy.

Land cover should describe the whole observable biophysical environment and is, thus, dealing with a heterogeneous set of classes. Evidently, a forest is defined with a set of parameters different from those to describe snow-covered areas. Therefore, the definition of classes by parameters is not using the same set of parameters for description of every class because it would be impractical. In the new approach, the parameters are tailored to each of the eight major land-cover features identified (Figure 2-1).

According to the general concept of an *a priori* categorisation, it is fundamental to the system that all combinations of the parameters are accommodated in the system independent of scale and tools used to identify objects (e.g., human eye, statistics, aerial photographs or satellite remote sensing). By tailoring the set of parameters to the land-cover feature, appropriate combinations of sets of predefined parameters can be made without the likelihood of impractical combinations of parameters. Two distinct land-cover features having the same set of parameters may differ in the hierarchical arrangement of these parameters in order to ensure a high geographical accuracy.

Having all pre-defined classes included in the system is the intrinsic rigidity of this type of categorisation. However, it is the most effective way to produce standardisation of categorisation results between user-communities. The disadvantage is that in order to be able to describe any land cover occurring anywhere in the world in a consistent way, a huge number of pre-defined classes are needed and that users should describe a specific land-cover feature in a similar way. This led to the development of the application software that assists users in determination of parameters in a stepwise selection procedure that aggregates parameters to derive the land-cover class. Two examples of this procedure are shown in Figure 2-1.

Figure 2-1. The major land-cover categories of LCCS (version 2.0) grouped under the primarily vegetated and primarily non-vegetated area distinction



Correlation with other existing categorisations becomes a matter of translating the existing classes back into the parameters of the system. Comparison of individual classes, as well as the used parameters forming this class, becomes feasible. However, to be able to translate existing classes, documentation is needed on the criteria used. Individual class names are insufficient for any meaningful translation and differences in understanding of the concepts behind categorisation systems may differ between experts (Comber *et al*. 2005d).

LCCS may replace previous methods or act as a bridging system that allows translation of a previous categorisation into LCCS terminology (e.g., such as the FAO/UNESCO Soil Legend (FAO 1988) or Soil Taxonomy (US Soil Conservation Service 1975) in soil science).

Table 2-1. Formation of LCCS classes by use of a set of parameter options with increasing level of detail of the class

Parameters used	Boolean formula[4]	Standard class name
Natural And Semi-Natural Terrestrial Vegetation		
A. Life Form and Crown Cover[5]	A3A10	Closed forest
B. Height	A3A10B2	High closed forest
C. Spatial Distribution	A3A10B2C1	Continuous closed forest
D. Leaf Type	A3A10B2C1D1	Broadleaved closed forest
E. Leaf Phenology (leaf longevity)	A3A10B2C1D1E2	Broadleaved deciduous forest
F. Stratification[6]:		
F. 2nd layer, Life Form, Crown Cover & G. Height	A3A10B2C1D1E2F2F5F7 G2	Multi-layered broadleaved deciduous forest
F. 3rd layer, Life Form, Crown Cover & G. Height	A3A10B2C1D1E2F2F5F7 G2F2F5F10G2	Multi-layered broadleaved deciduous forest with emergents
Cultivated And Managed Terrestrial Areas		
A. Life Form	A4	Graminoid crop(s)
B. Spatial Aspects:		
Field Size	A4B1	Large-to-medium sized field(s) of graminoid crop(s)
Field Distribution	A4B1B5	Continuous large-to-medium sized field(s) of graminoid crop(s)
C. Crop Combination	A4B1B5C1	Monoculture of large-to-medium sized field(s) of graminoid crop(s)
D. Cover-related Cultural Practices:		
Water Supply	A4B1B5C1D1	Rainfed graminoid crop(s)
Cultivation Time Factor	A4B1B5C1D1D8	Rainfed graminoid crop(s) with fallow system

2.5.2 Implementation of LCCS in various projects

The LCCS categorisation methodology has been tested, modified and validated in several international projects in order to analyse its applicability in different environmental settings, its use at different data collection scales and with different means of data collection, its usefulness for data harmonisation and in land-cover change analysis. LUCC has endorsed the methodology (McConnell and Moran 2001). It has been applied by the European Commission's Global Land Cover 2000 (GLC2000) project (Mayaux *et al.* 2004 and 2006), the FAO Africover project (10 countries) (Kalensky 1998) and FAO projects in the CEEC and the Commonwealth of Independent States (CIS) (e.g., Azerbaijan, Bulgaria (Travaglia *et al.* 2001), Romania and Moldova), in projects financed by the Italian Ministry of Foreign Affairs (e.g., Niger (Jansen *et al.* 2003c), The Gambia, Mozambique (Jansen *et al.* 2008a) and Senegal (Jansen and Ndiaye

[4] String of parameter codes selected; each code comprises a letter referring to the parameter and a figure referring to the parameter option selected.
[5] In LCCS this parameter is called 'Cover' but in this text the term 'Crown cover' has been preferred.
[6] If an additional layer is present, the 'Life form', 'Crown cover' and 'Height' need to be determined concurrently.

2006), by the Nordic Council of Ministers' Nordic Landscape Monitoring project (e.g., Estonia, Denmark, Norway and Sweden) (Groom 2004), and by a World Bank financed project in Albania (Jansen *et al.* 2006a). The results of some of these projects will be used for illustration purposes in subsequent chapters.

2.5.3 Example: primarily vegetated areas

For global application, the two major diagnostic elements for description of land cover are physiognomy and structure (Kuechler and Zonneveld 1988). These two elements are easily observable and can be quantitatively defined. These basic elements can be consolidated by additional species information in the case of primarily vegetated areas (e.g., 'Floristic aspect' and 'Crop type') or additional information on primarily non-vegetated areas (e.g., soil description, water quality).

In the present categorisation, primarily vegetated areas are classified using a pure *physiognomic-structural* method. The aspects considered are (1) physiognomy (i.e., the overall appearance of the vegetation), and (2) vertical and horizontal arrangement of the plants. This concept has been adopted with the confidence that only a pure physiognomic-structural representation of vegetation is able to incorporate, without any confusion of terms, floristic aspects of vegetation together with information on structure, which may be combined with environmental attributes (e.g., landform, climate, etc.). The proposed categorisation allows the user to add freely these attributes, which are not inherent characteristics of land cover, at any level of the created physiognomic-structural land-cover class.

The approach selected for categorisation of primarily vegetated areas in a land-cover categorisation system poses a challenge with regard to categorisation of other than (semi-) natural vegetated areas: the cultivated and urban vegetated areas. These managed vegetated areas are also characterised by plant communities containing growth forms and taxa, having a structure and a floristic composition. Therefore, the adopted physiognomic-structural approach is equally applicable to this type of area. Using the same approach to describe and classify this type of area at a certain level of detail has the advantage that all primarily vegetated areas can be compared.

A forestry application, for instance, will compile data that may also be useful for agricultural purposes. Information on vegetation structure and pattern may be used in both applications, whereas timber volume may only be useful in the first application. The latter type of data can be added by defining a so-called 'User-defined attribute' in LCCS.

2.5.4 Application for land-change detection

The advantages of the parameterised approach are that change detection becomes possible at the level of conversion of a class and that modification within a certain class type becomes immediately identifiable by a difference in parameter or through the use of additional parameters. Table 2-2 shows the conversion of a forest into a coffee (*Coffea* spp.) plantation (1) and a shrubland converted into a built-up area (2). Table 2-3 shows examples of modifications within the major land-cover type but with a change in domain (e.g., the change of a single parameter (1) leads to a less rigid change in domain than several changed parameters (2)), whereas Table 2-4 shows a land-cover modification within the domain (e.g., the change of a single parameter or the use of additional parameters).

Table 2-2. Detection of land-cover conversion showing defined classes with their set of parameters and the parameter options selected[7]

Parameter	Parameter option	Parameter	Parameter option
Land-cover conversion:			
1. from 'Multi-layered forest' (left) to 'Continuous large-sized field(s) of shrub crop(s) - coffee' (right):			
Life form of main layer:	Trees	Life form of main crop:	Shrubs
Crown cover:	Closed	Field size:	Large
Height:	> 30m	Field distribution:	Continuous
Macro pattern:	Continuous	Crop combination:	-
Leaf type:	-	Cover-related cultural practices:	-
Leaf phenology (leaf longevity):	-		
2nd layer Life form:	Trees		
2nd layer Crown cover:	Closed		
2nd layer Height:	High		
Floristic aspect:	-	Crop Type:	Coffee *(Coffea* spp.)
2. from 'Open shrubs (Shrubland)' (left) to 'Built-up area(s)' (right):			
Life form of main layer:	Shrubs	Surface aspect:	Built-Up area(s)
Crown cover:	Open		
Height:	5 - 0.3m		
Macro pattern:	-		
Leaf type:	-		
Leaf phenology (leaf longevity):	-		
2nd layer Life form:	-		
2nd layer Crown cover:	-		
2nd layer Height:	-		
Floristic aspect:	-		

[7] The '-' indicates that the parameter was not used.

The Land-Cover Classification System will register modifications within the land-cover type, that is from one domain to another (e.g., from 'Forest' to 'Woodland', from 'Shrubland' to 'Sparse vegetation' or from 'Tree crops' to 'Herbaceous crops') or within the domain (e.g., from 'Multi-layered forest' to 'Single-layered forest', from 'Small-sized fields of graminoid crops' to 'Large-sized fields of graminoid crops'). The more parameters used at the beginning of the monitoring process, the greater the detail of the defined class and the greater the possibility for detection of changes in any of the used parameters. The latter, however, is dependent on the method of measuring change.

The scale of the survey becomes an important issue concerning the number of used parameters. Both scale and the means of surveying (e.g., interpretation of satellite imagery, field plot sampling or statistical methods) determine which criteria can be used, thus where the limits are placed.

Table 2-3. Detection of land-cover modification within the major land-cover type showing defined classes with their set of parameters and the parameter options selected[8]

Parameter	Parameter option	Parameter	Parameter option
Land-cover modification within the major land-cover type:			
1. from 'Continuous closed forest' (left) to 'Continuous open forest (Woodland)' (right):			
Life form of main layer:	Trees	Life form of main layer:	Trees
Crown cover:	Closed	Crown cover:	Open
Height:	> 30m	Height:	> 30m
Macro pattern:	Continuous	Macro pattern:	Continuous
Leaf type:	-	Leaf type:	-
Leaf phenology (leaf longevity):	-	Leaf phenology (leaf longevity):	-
2nd layer Life form:	-	2nd layer Life form:	-
2nd layer Crown cover:	-	2nd layer Crown cover:	-
2nd layer Height:	-	2nd layer Height:	-
Floristic aspect:	-	Floristic aspect:	-
2. from 'Fragmented open high forest (Woodland)' (left) to 'Sparse trees and sparse shrubs' (right):			
Life form of main layer:	Trees	Life form of main layer:	Trees
Crown cover:	Open	Crown cover:	Sparse
Height:	High	Height:	> 30m
Macro pattern:	Fragmented	Macro pattern:	Parklike patches
Leaf type:	-	Leaf type:	-
Leaf phenology (leaf longevity):	-	Leaf phenology (leaf longevity):	-
2nd layer Life form:	-	2nd layer Life form:	Shrubs
2nd layer Crown cover:	-	2nd layer Crown cover:	Sparse
2nd layer Height:	-	2nd layer Height:	5 - 0.3m
Floristic aspect:	-	Floristic aspect:	-

[8] The '-' indicates that the parameter was not used.

Table 2-4. Detection of land-cover modification within the land-cover domain showing defined classes with their set of parameters and the parameter options selected[9]

Parameter	Parameter option	Parameter	Parameter option
Land-cover modification within the land-cover domain:			
from 'Small-sized field(s) of herbaceous crop(s)' (left) to 'Large-sized field(s) of irrigated herbaceous crop(s)' (right)			
Life form of main crop:	Herbaceous	Life form of main crop:	Herbaceous
Field size:	Small	Field size:	Large
Field distribution:	-	Field distribution:	Continuous
Crop combination:	-	Crop combination:	-
Cover-related cultural practices:	-	Cover-related cultural practices:	Irrigated
Crop type:	-	Crop type:	-

2.6 Land-use categorisation

Definition of sustainable and environmentally sound land-use systems requires information on current land uses as a starting point for any modification of current practices, in accordance with the wishes of the stakeholders and land managers' concerns and priorities. A major constraint to sound planning is the lack of an internationally agreed land-use categorisation that would enable the description of what the different types of land uses are in the world at large and in the individual countries in particular, where they are located, and how they are changing. Thus, at present comparisons are difficult to make, if at all possible.

Land cover and land use are not the same, though they are linked. The development of a parameterised approach for land use can benefit from the lessons learnt from the development of land-cover categorisation systems. Land use is related to the arrangements, activities and inputs people undertake in a certain land-cover type to produce, change or maintain it. Therefore, the parameters describing land use are linked to aspects, which may sometimes, but not always, be inferred from the resulting land cover (e.g., a field pattern indicating farming). However, interviews with land users and ground truthing are essential in any land-use data collection and cost contributing factors. Land use has many aspects that go beyond land cover (e.g., socio-economics, cultural and legal aspects) and therefore trying to systematically link land cover to land use embodies the risk of not capturing several aspects of land use. The complexity of the land-cover/land-use relationship (Cihlar and Jansen 2001) is such that the land-cover/land-use relationship varies not only between areas of interest but also within such areas.

The need for a structure to aggregate land-use data is imminent once data is collected and groupings of observations are needed. Therefore, a system is required

[9] The '-' indicates that the parameter was not used.

that accommodates arrangement of land-use data at various levels and which provides a reference base for identification and comparison. To date, neither a common underlying principle for land-use description, nor agreement on the inherent characteristics, exist.

Two widely applied approaches for characterisation of land uses can be identified (Jansen and Di Gregorio 1998b):

1. The *function* approach to describe land uses in an economic context; this type of approach answers the aim or 'what for' of land uses. This approach is commonly used for sectoral descriptions of land use (e.g., agriculture, forestry, fisheries, etc.). Land uses can be grouped together that do not possess the same set of observable characteristics but serve the same purpose, the so-called polythetic view (Sokal 1974). An example of such land uses is 'agriculture' that may come in many forms, dealing with plants or animals, related to extraction, production or service characteristics. These 'agricultural' land uses share a large proportion of characteristics but do not necessarily agree in any one characteristic (e.g., bee-keeping versus annual rainfed maize (*Zea mays* L.) cropping).
2. The *activity* approach to describe what actually takes place on the land in physical or observable terms. Activity is defined as "*the combinations of actions that result in a certain type of product*" (UN 1989) and refers to a process. A variety of activities may serve a single function (e.g., both farm housing and farming activities serve agriculture) and this approach is independent from the function approach.

Function groups all land used for the same economic purpose independent of the type of activities taking place, whereas *activity* groups all land undergoing a certain process resulting in a homogeneous type of products but that may serve different functions. A preliminary outline of this concept has been tested in a selected area in Kenya for identification of land-use parameters and derivation of land uses taking the land cover as basis (Jansen and Di Gregorio 2003). A result of this study and one in Lebanon (Jansen and Di Gregorio 2004) is that land-cover boundaries do not necessarily coincide with land-use boundaries (e.g., recreation). Similar land-cover types may contain different land uses and vice versa. The relation between land cover and land use is complex and needs careful examination in each situation (Cihlar and Jansen 2001). A more advanced version of the land-use categorisation concept was applied in an EU PHARE Programme project in Albania (Jansen 2003 and 2006). These project results are used in successive chapters to illustrate the developed methodology (Chapters 4 and 6).

The level of data collection increases notably from the function to the activity approach. The proposed use of the function approach at first level is also a pragmatic choice as most major functional groupings can be detected with limited investment of resources, whereas the activity approach would require substantial investments in data acquisition. A consolidated effort to catalyse the further development of a comprehensive land-use categorisation system is still necessary.

2.7 Conclusions

Land-change detection is a multi-disciplinary subject and it is important that an integrated approach is taken with several partners involved to agree upon a widely accepted reference base for land-use and land-cover categorisation. The parameterised approach of the Land-Cover Classification System has proven to be pragmatic and serves a variety of users in their needs. The independent diagnostic criteria, the parameters, standardise the description of classes in a systematic way. In turn, these criteria can be verified individually during the field sampling (or in future with Google Earth assuming that very high resolution images will become available for every place on Earth) and they can be analytically used in change studies. The parameterised land-cover categorisation developed contributes to standardisation of the systematic description of land-cover classes and the diagnostic criteria provide a uniform basis for detection of land-cover changes.

The parameterised approach allows detection of changes related to land-cover conversion, as well as the more difficult detectable land-cover modifications, based upon the diagnostic criteria used in the Land-Cover Classification System. Comparison of classes is possible at two levels: (1) the individual classes and (2) the used parameters. However, monitoring land-cover changes alone is not sufficient; it needs to be linked to land use in order to improve our understanding of why certain changes occurred, as well as analyse land-cover/land-use trends. A conceptual approach for land-use categorisation is proposed that combines function, which groups all land used for the same economic purpose independent of the type of activities taking place, with activity, which groups all land undergoing a certain process resulting in a homogeneous type of products but that may serve different functions. This approach was successfully tested in a case study and in a more advanced form in a project (see also Chapter 4 and paragraphs 6.1 to 6.4) but more work is required on definition of the diagnostic criteria in order to develop a reference base, as well as understanding the complexity of land-cover/land-use relations.

The proposed concept for future database development, using standardised categorisations as a reference base, will facilitate comparison and correlation. However, it does not solve the problems of time series analysis using existing data sets made with different categorisations. For these harmonisation will be a key issue as will be discussed next in Part II in Chapters 3 and 4.

PART II - HARMONISATION

We would never have learned anything if we had never thought:
"This object resembles this other, and I expect it to manifest the same properties".

Bertrand de Jouvenel

Everything is related to everything else,
but near things are more related than distant things.

Waldo Tobler's (1970, p. 236) first law of geography

Abstract

Land cover and land use are two key elements that describe the environment in natural and human-activity related terms. Land cover and land use are closely linked, however they are not identical. Therefore chapters 3 and 4 consider land cover and land use each in its own right in the context of harmonisation of class sets. In Chapter 3 harmonisation of land-cover class sets is discussed with particular emphasis on quantification of harmonisation results at the semantic contents level. In Chapter 4 harmonisation of land use is discussed with the emphasis on the harmonisation of land-use change that is a prerequisite in the analysis of environmental processes and problems.

Chapter 3. Harmonisation of land-cover class sets and quantification of harmonisation results

Harmonisation of land-cover data relates to spatial data integration and needs therefore to consider the data concepts adopted and the spatial, temporal, semantic and quality aspects of the data. Differences in semantic concepts are often considered the key obstacle to data integration and interoperability. If the problem of harmonisation is limited to the variation in the semantic contents of data expressed as differences in categorisation, then various approaches have been developed to address the methodological issues and for computing semantic similarity. Five Nordic class sets were selected for establishing correspondences between their semantic class contents using the parameterised Land-Cover Classification System (LCCS) as a reference system. Subsequently, semantic similarities between pair-wise classes were calculated using a module of LCCS. This part of the chapter first examines the aspects of land-cover harmonisation and the LCCS methodologies for categorisation and semantic similarity. It then discusses the functioning of LCCS as a reference system in which the more or less corresponding class of the original Nordic classes was determined and the semantic similarity indices computed. Suggestions are provided for improvements in the LCCS methodology, both in establishing correspondences and for

Based on: Jansen, L.J.M., Groom, G.B., Carrai, G., 2008. Land-cover harmonisation and semantic similarity: some methodological issues. *Journal of Land-Use Science* 3(2-3): 131-160.

Jansen, L.J.M., 2006. Harmonisation of land-use class sets to facilitate compatibility and comparability of data across space and time. *Journal of Land-Use Science* 1(2-4): 127-156.

computing semantic similarity. Recommendations are given for the way forward in land-cover harmonisation and for measures to express the quality of harmonisation of the semantic contents of class sets at class set level and individual class level.

Chapter 4. Harmonisation of land use and land-use change class sets

Harmonisation of land-use class sets should consider both space and time, as the objective should include harmonisation of land-use change to analyse environmental processes and problems. Analysis of major class sets reveals that a parameterised approach with two parameters may suffice: 'function' that describes land use in an economic context and 'activity' that is defined as the combination of actions resulting in a certain type of product.

An example of land-use harmonisation in Albania illustrates how creation of a reference system based upon the principles of categorisation, using the synergy between categorisation and information technology concepts and based upon the 'function' and 'activity' parameters can facilitate harmonisation between land-use class sets in parallel to achieving land-use change harmonisation.

As not only data quality is of paramount importance, further research is necessary to define quantitative measures to determine harmonisation results both at class level and between class sets.

3 Harmonisation of land-cover class sets and quantification of harmonisation results

3.1 Introduction

Land cover defined as "*the observed (bio)physical cover on the Earth's surface"* (Di Gregorio and Jansen 2000) is widely perceived as an important component of environmental and ecological systems and is considered central to understanding global environmental change (Meyer and Turner 1994; Turner *et al*. 1995; Walker *et al*. 1997; Lambin *et al*. 2000). Spatial variability is a fundamental quality of land-cover data with important implications for environmental and ecological modelling and analysis (Ahlqvist and Shortridge 2006). Environmental models become increasingly more sophisticated and with that the importance of accurate, meaningful and current data on land use and land cover to support these models increases (DeFries and Belward 2000). Many geographic entities have undertaken surveys of land cover; often these have been made on the basis of a particular *categorisation system* (often termed classification system) or *class set* (usually termed classification, nomenclature, legend). A common problem is however, that as over time knowledge advances, technology develops and policy objectives change, each survey with a class set designed for its purpose, rather than being part of a sequence creates a new baseline data set. Whereas in the past survey maps were illustrations that accompanied a descriptive memoir, nowadays, in the era of geo-informatics, maps are understood as primary data sets (Fisher 2003). This poses a further problem regarding the associated class labels that are often rather cryptic and unrelated to any categorisation system where the user may learn the concepts and criteria behind the class labels (Comber *et al.* 2005; Wadsworth *et al*. 2006). Differences in the naming of classes, changes in class definition and addition or removal of classes in data sets covering the same area in different periods create difficulties in the separation of actual changes over time from apparent changes in category definitions.

In practise, results from different surveys do need to be harmonised over time and space (e.g., in relation to trans-boundary issues), and reference to existing information is often required to verify new results (e.g., regarding urban sprawl and landscape changes). Data harmonisation, being defined as "*the intercomparison of data collected or organised using different classifications dealing with the same subject matter*" (McConnell and Moran 2001), thus becomes a prerequisite for many data analyses. Harmonisation will allow countries and institutions to continue to use established methods and data sets made with certain financial and intellectual investments (UNEP/FAO 1994; Wyatt and Gerard 2001).

Development of the general-purpose Land Cover Classification System (LCCS) (Di Gregorio and Jansen 2000) has led to the common belief that once such a categorisation system becomes widely adopted for new surveys the problem of data harmonisation would be overcome because new data sets would be collected using a single standard system allowing direct comparison of new class sets, whilst existing class sets could be 'translated' into the adopted system making possible direct class comparison with new class sets. However, this stance that is geared towards data standardisation, defined as "*the use of a single standard basis for classification of a specific subject*" (McConnell and Moran 2001), assumes falsely that the continuous advances in either knowledge, technological developments and/or changing policy objectives will not have any impact on a categorisation framework or its application. With each data collection effort lessons are learnt that leave their imprint on successive efforts (e.g. CORINE[10] land cover 1990 versus 2000 (Büttner *et al*. 2004)). Data standardisation may thus be an unrealistic expectation and only partly feasible with the need for data harmonisation always present.

With the emphasis shifting from *static* land-cover mapping towards more *dynamic* environmental monitoring and modelling (Lambin *et al*. 2000; McConnell and Moran 2001; Dolman *et al*. 2003), it is necessary to examine how far research has progressed in data harmonisation methodologies. In this paper, first the various aspects of land-cover harmonisation are examined with particular emphasis on semantic contents of classes and semantic similarity (paragraphs 3.2-3.4). This is followed by examination of the methodology of a particular tool, the Land Cover Classification System (paragraph 3.5), and the experiences with this tool by the Nordic Landscape Monitoring (NordLaM) project of the Nordic Council of Ministers. LCCS was used as a reference system for establishing correspondences between semantic contents of classes from five Nordic class sets and for computing semantic similarity between those classes (paragraph 3.6). Whilst the harmonisation results of the Nordic class sets will be of interest to essentially a Nordic audience, the methodological issues being addressed are relevant to the wider context of spatial data integration, interoperability, land-cover harmonisation and standardisation. Suggestions are provided to improve the methodologies implemented in LCCS (paragraph 3.7) and recommendations are provided for measures that would quantify correspondence results as well as discussing some general research questions that are still open (paragraph 3.8). Taking stock of land-cover harmonisation and semantic similarity methodologies is especially required in the context of Spatial Data Infrastructure (SDI) initiatives that change data access (e.g., the European Commission's 'Infrastructure for spatial information in Europe' (INSPIRE) Directive[11]) and in the context of the UN that promotes the use of modern

[10] CORINE stands for 'Co-ordination of information on the European environment'.
[11] Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an 'Infrastructure for Spatial Information' in the European Community (INSPIRE) was published in the official Journal on the 25 April 2007 and entered into force on 15 May 2007 (http://inspire.jrc.ec.europa.eu/).

information technologies in developing countries (e.g., UNCED Agenda 21 and the World Summit on Sustainable Development).

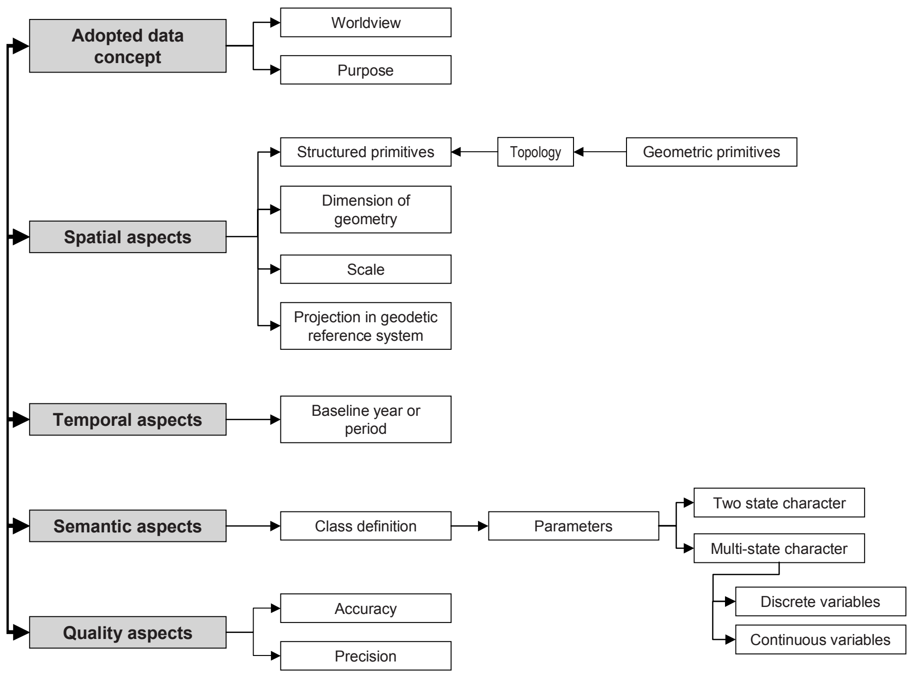
3.2 The aspects of harmonisation

Land-cover harmonisation touches the issue of spatial data integration, when it concerns spatially explicit data. The recognition that spatial data integration is an essential step in land-change modelling and initiatives (e.g., planning and decision making) that aim to respond to land change is broadening (Comber *et al*. 2005). Increasingly, the need is recognized for a deeper understanding of the wider meaning of data stored in geo-databases (Ahlqvist 2004). The latter is particularly important in the context of data interoperability, i.e. the exchange of meaningful information between multiple information sources (Vckovski 1999).

In the definition of McConnell and Moran (2001) given above two distinct levels of harmonisation should be identified:

1. The intercomparison of *classes* belonging to different categorisation systems;
2. The intercomparison of the *data* collected with the use of these categorisation systems.

Figure 3-1. The five interrelated aspects of land-cover data harmonisation



The first level deals with how classes are defined and named, whereas the second level deals with how data were collected and represented (e.g., methods, scale, time, coordinates, etc.). Land-cover harmonisation, therefore, has without exception to consider the following interrelated aspects (Figure 3-1):

- The adopted data concepts (level 1 and 2);
- The spatial aspects (level 2);
- The temporal aspects (level 2);
- The semantic aspects (level 1); and
- The quality aspects (level 1 and 2).

These different aspects of land-cover harmonisation are discussed in more detail below.

3.2.1 Adopted data concepts

Comparison of data sets should include comparison of the *Weltanschauung*, or worldview, embodied in the data (Comber *et al.* 2004a). Differences in the way that land cover may be conceptualised are not addressed by the stated objectives of SDIs nor by current metadata or data quality reporting paradigms. Only very few of the available categorisation systems explicitly mention the concept for description of classes. For example, the LCCS states that it is based upon a *structural-physiognomic approach* (Jansen and Di Gregorio 2002). Statements that any categorisation system will allow a 'neutral' description of land cover ignore the fact that human beings always look at the object land cover in a specific way (e.g., an economist will look at it differently compared to an agronomist or an ecologist). Furthermore, data are collected for an intended purpose and this leads to a particular or prevalent view. Related to the view with which data are collected is the meaning of the data. The latter may be obvious to the data producer, but it is rarely as clear to the data users unless they were part of the data collection process. Access to data through SDI initiatives implies that countless potential users may be reached. However, in almost all cases to date the metadata do not provoke users to consider the wider meaning of the data (Comber *et al.* 2004a; Schuurman and Leszczynski 2006). Land-cover data collected in the context of a forest inventory will focus on description of different parameters than land-cover data collected for surveillance and monitoring of habitats, although these data may have some parameters in common. In addition, the purpose for which data are collected may relate to a design with higher thematic and spatial accuracies for certain classes than for others. The lower accuracies may be insufficient for some data uses but the metadata do not provide such information. Worse still, some data sets have been collected without proper validation (e.g., the FAO Africover data, some country data sets for CORINE 1990). Any data collection without proper validation remains an *untested hypothesis* (Strahler *et al.* 2006).

3.2.2 Spatial aspects

The International Organization for Standards Technical Committee 211 (ISO TC/211) (www.iso.org) and the Open Geodata Interoperability Specification

Consortium (www.opengis.org) are two sources that have developed numerous standards concerning the spatial aspects of data. The spatial or geometric aspects of the data set comprise the description of the form of the entities through geographic primitives or through a structured geometry (e.g., topology). In general the spatial aspect considers cases of different representations of the same object. For instance, a road network can be represented by polygons of the road surface, as a network of road axes and nodes, or as an information level where the road is represented by the sequence of borders such as walls and façades of buildings (e.g., cadastre). Each way of representing a road network follows a different set of conceptual and technical practices.

The spatial component describes the dimensions of the geometry (i.e. two dimensions for areas, one dimension for lines and no dimension for points) in relation to the scale and projection in a geodetic reference system. Differences in scale can be overcome by geometric generalisation but this may imply loss of information; generalisation means also reorganisation of the semantic attributes (see paragraph 3.2.4). For instance, elaborating the example used above one could imagine a matching of road maps of two neighbouring countries. In country A, the roads are depicted at scale 1:5,000 by polygons projected in a local coordinate system, whereas in country B the roads are depicted at scale 1:100,000 by lines projected in UTM WGS84. Data harmonisation in such a case should consider three spatial aspects: (1) how to depict the roads (e.g., as polygons or lines), (2) the scale to be adopted and (3) the geodetic reference system to be used. Harmonisation of data sets that are represented by either polygons or grid cells (raster), such as many thematic data sets, do not represent significant problems of geometric harmonisation since usually these can be restructured using topological procedures.

3.2.3 Temporal aspects

It is necessary to consider the temporal aspects of data sets because certain themes undergo more changes with time than others, and data harmonisation between class sets covering the same subject matter but from different periods may not be meaningful. For example, the first CORINE land-cover data set, CLC1990, spans the period 1986-1998 and the second data set, CLC2000, the year 2000±1 year (Büttner *et al*. 2004). Harmonisation of the temporal aspects should provide a baseline period or year. For example, one may well question whether harmonisation of land-cover data from country A from the year 1995 with those of country B from the year 2007 is at all meaningful since in the period represented changes are also likely to have occurred in country A and, as mentioned previously, knowledge and technology have advanced and policy objectives will also have changed.

3.2.4 Semantic aspects

Since different applications have different worldviews and semantics, interoperability is primarily understood as a semantic modelling problem (Bishr *et al*. 1999). The variation in the semantic content of data expressed as differences in categorisation has received limited attention until recently (Feng and Flewelling 2004). Comber *et al.* (2004) report that differences in semantic concepts are often *the* major barrier to data integration. Achieving semantic interoperability in order to use existing data sets at a satisfactory level has therefore become a key issue.

Describing land cover is to account for its character, and here different concepts may co-exist in a single class or single categorisation system (EC 2001):

1. Two-state character, i.e. present/absent, 1/0, positive/negative, etc. (e.g., in LCCS the dichotomous phase uses the two state character for primarily vegetated/primarily non-vegetated, terrestrial/aquatic and artificial/(semi) natural land covers); and
2. Multi-state character that can be subdivided into quantitative and qualitative, or so-called terminological, variables. Two different types of quantitative variables can be distinguished: discrete (e.g., in LCCS in the modular-hierarchical phase vegetation can be described using growth form, leaf type and leaf longevity) and continuous variables (e.g. continuous fields that allow a more precise description of vegetation gradients and mixtures).

Harmonisation between classes that represent a mixture of these characters will be difficult, as will harmonisation between classes that have a two-state character but represent a mixture of quantitative and qualitative variables.

Class descriptions contribute to the definition of boundary conditions that should be applied unequivocally and consistently when establishing correspondence between class sets in order to avoid errors in data interpretation. The level of certainty with which such class correspondence is established is highest when the same parameters have been applied; a difference in the applied parameters, and thus in boundary conditions, results in a lower certainty level.

It may be necessary to 'translate' a class set into a third system, a so-called *reference system* that functions like a bridge between two class sets: each class in the original class sets will find its more or less corresponding class in the reference system. The use of a reference system may be a sensible choice when many class sets are involved as the number of pair-wise class combinations becomes excessive with comparison of n class sets requiring $n(n\text{-}1)/2$ comparisons to be made. As Wyatt and Gerard (2001) point out, the use of a reference system requires a single 'translation' from each original class set into the reference system and obviates the need for pair-wise class comparisons between every class set of interest.

3.2.5 Quality aspects

The quality aspects concern all the above-mentioned aspects in order to produce quality land-cover harmonisation results. Harmonisation requires the analysis of data quality because correspondence between data sets having very different levels of quality may not be meaningful (Jansen 2006). Often the metadata of a land-cover data set provide information concerning the positional and thematic accuracy. However, there are many other measures of quality and uncertainty that should be specified as they may limit interoperability but that currently are excluded from metadata requirements. One such measure is discussed in the next paragraph.

3.3 Methodologies for assessing semantic similarity

One step towards achieving semantic interoperability is to measure the degree of semantic similarity between categorisations. Various practical solutions for overcoming semantic differences have been proposed:

- Use of a standard set of parameters to overcome semantic divergences in categorisation systems (Wyatt *et al*. 1994; Wyatt and Gerard 2001; Jansen and Di Gregorio 2002);
- Bayesian probabilities based on a variety of metrics of geometric and semantic similarity to identify areas of change (Jones *et al.* 1999);
- Use of a similarity function to determine semantic neighbourhoods and distinguishing features (Rodríguez *et al*. 1999; Rodríguez and Egenhofer 2003);
- Uncertain conceptual spaces to represent uncertainty between spatially coincident but semantically divergent data (Ahlqvist *et al.* 2000; Ahlqvist 2004 and 2005b) or between spatially and semantically divergent data using the parameterised LCCS as a reference system to mediate between two class sets (Ahlqvist 2005c);
- The mathematical theory of concept lattices to link semantics from different data ontologies[12] and reveal interrelationships between categories (Kavouras and Kokla 2002);
- Use of similarity indices to describe the extent to which descriptions of classes match (Jansen *et al*. 2003c; Feng and Flewelling 2004);
- Semantic statistical approaches using expert knowledge to reconcile the uncertainty between different ontologies using expert descriptions of semantic relations (Comber *et al.* 2004a and 2005; Wadsworth *et al.* 2006);
- Use of a fuzzy logic framework and expert knowledge to reconcile inconsistent land-cover data (Fritz and See 2005);

[12] An ontology is an explicit specification of a conceptualisation (Gruber, 1993). In both computer and information sciences, an ontology is a data model that represents a set of concepts within a domain and the relationships between those concepts. It is used to reason about the objects within that domain. In the context of this paper one can consider ontology to be synonymous with categorisation system.

- Semantic variograms based on semantic similarity metrics to measure spatial variability of categorical data (Ahlqvist and Shortridge 2006); and
- Application of weighting by the semantic distance calculated from the four most discriminant LCCS parameters to the confusion matrix in a validation scheme (Mayaux *et al*. 2006).

Mainly the computer and information sciences have developed ways of computing semantic similarity that provide a quantitative measure to the user as to which categories are more similar and which categories are more dissimilar (semantically distant). Measuring semantic similarity of categories, either before or after data collection or between existing data sets, is an emerging area of research. The exploration of category relationships within a categorisation system can reveal how well classes are separated or if there is a risk of confusion between classes, a situation that may be problematic from a data accuracy perspective (Ahlqvist 2005b and 2005d). Semantic similarity addresses the issue of accuracy also in another way: if complete correspondence between classes from different class and data sets is not always possible then how accurate is class correspondence and how accurate are data harmonisation results? To date, research has not led to any widely accepted methodologies for land-cover harmonisation, or to an accepted means for quantifying the quality of harmonisation results.

The points discussed above make it clear that land-cover harmonisation is a multifaceted issue that concerns both geo-informatics, and statistical and subject matter specialists. A solution offered by any of these without involvement of the others will probably fall short in addressing the complexity of the problem. Efforts that are limited to a crosswalk 'translation' effort between categorisation systems and/or class sets ignore such complexity. For example, overviews in which FAO shows that country land-cover *maps* are 'translated' into LCCS (e.g., as shown by Herold *et al*. 2006b) offer the wrong impression of data harmonisation as such efforts have been limited to correspondence of original classes (legends) with LCCS rather than having examined the full meaning of the data.

3.4 Semantic differences that affect the interoperability of land-cover and land-use data

One specific issue that affects the interoperability of land-cover data sets is that land-cover class sets often contain land-use elements. Thus harmonisation of land cover may in different cases imply either a need to make harmonisation of land-use categories or the decision to leave the land-use elements out of the established correspondences, in which case part of the data richness is lost. Though land cover and land use are related, they are not the same (Jansen and Di Gregorio 2002). Nowadays it is advocated to separate the two but in the practise of much survey work this is frequently not the case for various justifiable reasons often related to the intended purpose for which data are collected.

According to Brown and Duh (2004) land cover and land use have three major semantic differences that affect their interoperability:

- The *category definitions* of land cover and land use are different. Land cover describes what you see on the surface of the Earth while land use may relate to an intended purpose that is not necessarily directly observable. For example, 'undeveloped forest' is a clear-cut area that continues to be used for forestry (Lund 1999);
- Land cover and land use have different *geometric expressions*; consequently a classification cross-walk approach to semantic interoperability that defines interrelations between categorisation systems or class sets without redefining spatial objects, as has been applied for alternative vegetation/land-cover class sets by the IGBP (Loveland *et al*. 2000), may be an inadequate solution for translation between land use and land cover (i.e. the spatial objects might need to change in addition to the class definitions); and
- Land cover and land use have different *spatial rules* to assign attributes to features because land-use class definitions tend to integrate information about activities taking place within a spatial unit (e.g., cadastral parcel or zone), while land-cover class definitions assess the static and *in situ* conditions. Thus, the entities of a land-cover data set (e.g., polygons) usually show more spatial variation than those of a spatially explicit land-use data set (assuming both data sets are compiled based on sources of the same level of detail).

Cihlar and Jansen (2001) pointed out that the complex relationships between land cover and land use should be considered from a spatial and thematically consistency viewpoint: in one-to-one and one-to-many land-cover/land-use relationships the relationship is thematically and spatially unique, whereas in many-to-one land-cover/land-use relationships the relationship is either not thematically unique but spatially consistent throughout the domain of interest, or the relationship is not thematically unique and not spatially consistent throughout the domain of interest. In addition, these relationships may change over time in the domain of interest, as well as vary between different domains of interest.

3.5 Methodological issues: the case of LCCS

The objective of the parameterised approach of the LCCS, developed by FAO and UNEP, was to have a consistent and pragmatic methodology for land-cover description in several countries representing different types of environments (Di Gregorio and Jansen 2000). Subsequently the methodology and its software application have been endorsed by the Land-Use and Land-Cover Change

(LUCC) core project of the International Geosphere Biosphere Programme (IGBP) and International Human Dimensions Programme on Environmental Change (IHDP) (McConnell and Moran 2001). More recently, FAO's attention has shifted from land-cover *mapping* to land-cover *harmonisation*. LCCS is, as a basis of a harmonisation strategy, recommended by the Global Observations of Forest Cover – Global Observation of Land Dynamics (GOFC-GOLD) and the Global Terrestrial Observing System (GTOS) (Herold *et al.* 2006b). Since the LCCS categorisation methodology was never critically reviewed in the scientific literature now seems a timely moment to do so, taking into account also the more recent methodological developments discussed above. The focus in this paper is on (a) methodological issues in the categorisation that might have repercussions on class comparisons when used as a reference system and (b) on the semantic similarity methodology.

3.5.1 Categorisation issues in the LCCS methodology

The main documents available that describe the LCCS (Di Gregorio and Jansen 2000, updated in FAO 2005) lack a formal definition of the categorisation rules. This represents a problem because, as the software source is not open, there is no possibility to (easily) understand the behaviour of the software application. The underlying logic can only be derived experimentally by using the software intensively and by defining classes step-by-step with the software to know if they are correct. Furthermore, identical Boolean class codes are used in LCCS for dissimilar parameters (e.g., in each main land-cover category the first parameter is coded 'A' followed by a number, thus 'A1' occurs eight times), though numerical class codes are unique. This means that researchers cannot refer to a comprehensive model that would allow them to make comparisons with other categorisation systems in order to evaluate LCCS. It also means that it is not possible to propose modifications as the formal definition of classes is missing and thus it is impossible to adequately describe LCCS (Di Costanzo and Ongaro 2004). All this has far-reaching consequences for the use of LCCS as a reference system with existing class sets because it implicitly requires adoption of a parameterised approach of which, for the user, the underlying rationale is mainly a 'black box'.

A parameterised approach is used to define classes organised hierarchically in a tree-like structure. The hierarchical order of parameters is justified in terms of the ease with which the orders are observed, but it would be more correct to speak of a hierarchical tree-like structure with more inclusive and abstract concepts at the top and more detailed concepts further down the hierarchy, a structure that can be modelled with set inclusion ("*is a*") relations (Feng and Flewelling 2004; Ahlqvist 2004 and 2005c). The "*is a*" relation captures superordinate-subordinate relations between two categories.

To shed some light on the 'black box' one could formulate, using simple terms, categorisation in the LCCS Classification Module (LCCS-CM) of category A as follows:

$$A = \sum p_1 + p_2 + ... + p_n$$

Equation 1

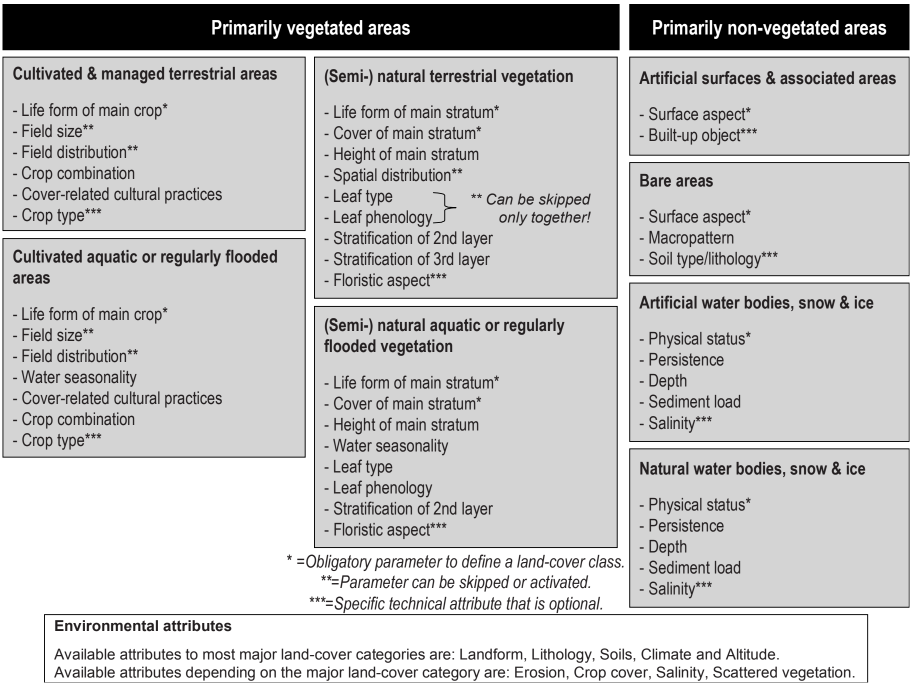
In this equation p stands for parameter. Thus, classes are defined by summing up parameters. However when *modifiers* are used that further refine an already used parameter the class definition is for example of the type:

$$A = \sum p_1 + (p_2 + mp_2) + ... + p_n$$

Equation 2

where parameter$_2$ is accompanied by its modifier (mp_2). It is important to understand that in such a case two codes are found that relate to the same defining element.

Figure 3-2. The major land-cover categories of LCCS (version 2.0) grouped under the primarily vegetated and primarily non-vegetated area distinction



The first land-cover parameter, or first and second for the (semi-) natural vegetation categories, is an obligatory element to define a class (Figure 3-2). But since codes are not exclusively assigned to a specific parameter, one needs to know to which of the set of eight main land-cover categories the summation of parameters, with or without modifiers, belongs. Thus one needs in addition the establishment of the land-cover category in this set in order to understand the meaning of the codes:

$$\text{Land-cover category} = \{1, 2, ..., 8\}$$

Equation 3

Figure 3-2 shows for each of the eight main land-cover categories the set of land-cover parameters and the specific technical attribute. Parameters are ordered hierarchically and when defining a class there is a top-down approach to be followed for the land-cover parameters with only a few of them that can be skipped, or activated when in principle the parameter is inactive (e.g., ‘Spatial distribution’ and ‘Field distribution’ in the second version of the LCCS-CM are inactive unless the user activates these parameters), in order to continue definition of a class. The position of a parameter in the hierarchical order can be considered as a *salience weight* (Ahlqvist 2004). For example, ‘Crop combination’ occurs in the fourth position in cultivated terrestrial areas and in the sixth position in aquatic or regularly flooded cultivated areas. In the latter type of environment the parameter is considered to have less weight in the class definition. The optional specific technical attributes and any environmental attributes can be added after having defined a land-cover class with at least the obligatory land-cover parameter(s). Codes for all attributes are unique and the order of appearance is linked to their coding and not to any weighting (e.g., landform with ‘L’ codes appears always before altitude with ‘P’ codes).

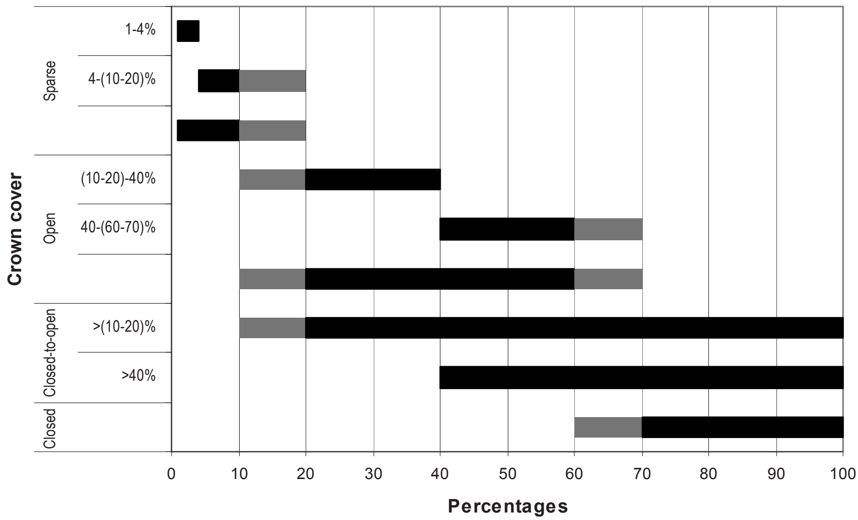
At first sight LCCS-CM may appear to be a so-called ‘crisp’ categorisation system with mutually exclusive parameter options but this is not always true. Figure 3-3 provides an example for the parameter ‘Crown cover’[13] used in the primarily vegetated area land-cover categories. The grey areas in the figure indicate threshold values for definition that are formed by a range and at these percentages the crown cover can be either sparse or open (10-20%), or, open or closed (60-70%). Two parameter options added in the second version of LCCS-CM introduce inconsistencies in the adopted concept: (1) ‘Closed-to-open’ defined as “*between 100 and 15%*” (FAO 2005) does not correspond to the options closed plus open as the range 10-15% is missing; and (2) ‘Closed-to-open’ defined as “*between 100 and 40%*” uses the value 40% that only exists as a threshold value for modifier options. Both parameter options ignore the range 60-70% as a threshold value. For the ‘Crown cover’ definition and other similar parameter definitions a fuzzy representation would be more suitable, as suggested by Ahlqvist (2005c).

The basic principle adopted in LCCS “*that a given land-cover class is defined by the combination of a set of independent diagnostic attributes*” (FAO 2005,

[13] In LCCS this parameter is called ‘Cover’ but in this text the term ‘Crown cover’ has been preferred.

p.12) may be true in most cases but is clearly not in all, as can be demonstrated by the use of the parameters 'Life form' (growth form) and 'Height' since they are interlinked in the definitions (see also paragraph 3.5.2). These interlinkages determine also which options are valid for the parameter 'Stratification' (vegetation layering) further down the hierarchy.

Figure 3-3. Overlapping definitions of crown cover parameters and modifiers in LCCS[14]



Thus the LCCS-CM methodology is not so clear-cut as appears to be the case when reading the available documentation. There are quite a number of exceptions to the categorisation rules (e.g., to give some examples from the LCCS glossary: herbaceous bamboos can be considered 'Woody' under (semi-) natural vegetation, or a succulent plant such as pineapple can be considered 'Shrub' under cultivated and managed areas) and restrictions (e.g., vegetation layering cannot describe more than three layers whereas in tropical areas more vegetation layers may occur) that limit the potential multitude of classes. These exceptions and limitations may be the result of the adopted concepts and common sense, but they make LCCS less easily understandable and less easy in its application despite its software.

3.5.2 Semantic similarity issues in the LCCS methodology

When using LCCS as a reference system the step-by-step definition of a class, or category, described above is also the first step in the 'translation' procedure. The 'translation', i.e. to find for each class in the original class set its more or less

[14] The definitions of 'Closed-to-open' have been added in the second version of LCCS using the threshold value of 15% where in other definitions the range 10-20% is given. For consistency's sake use has been made of the latter.

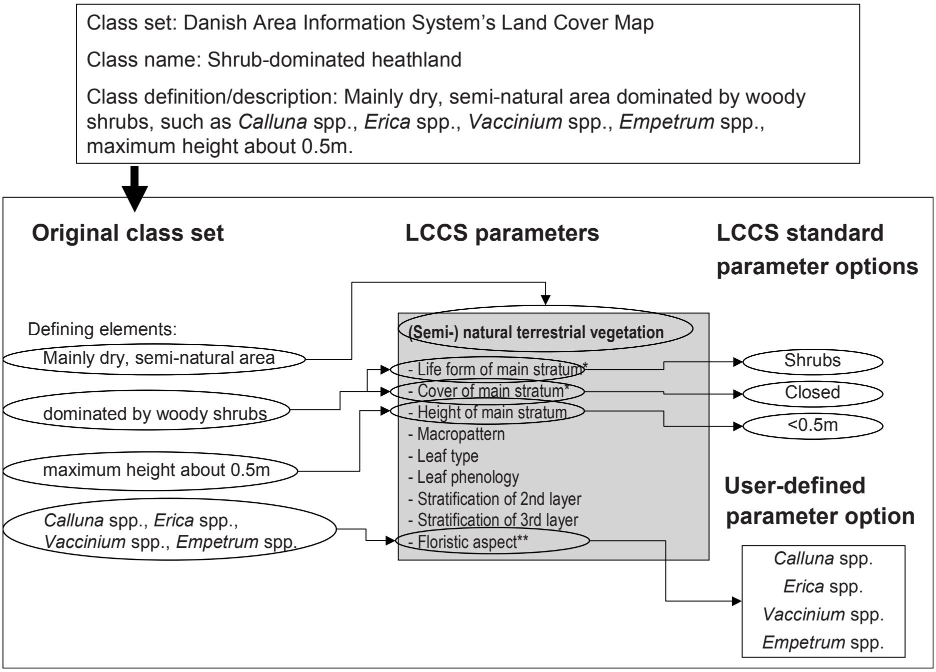
corresponding class in LCCS-CM, begins actually by *defining* rather than finding the corresponding class in the LCCS-CM.

Figure 3-4 illustrates the parameterisation procedure using a class of the Danish Area Information System as an example. This step introduces uncertainty as defining the corresponding class may differ from person to person. After definition of a class in LCCS-CM, followed by storage in the LCCS Legend Module that is not further discussed here, import of all classes into the LCCS Translator Module (LCCS-TM) can take place. The name *Translator* Module is misleading, as one does not actually 'translate' in the LCCS-TM but one compares and the term 'translator' does not refer to an automated thesaurus or text mining procedure for words used in the category names and definitions. In this module computation of the semantic similarity between the corresponding classes in LCCS terminology can be executed, this is called 'similarity assessment'. Similarity provides a quantified measure stating how much of a definition is included in another definition and this occurs between parameter options or a group of parameter options in the case of a class (e.g., how much of the definition of parameter option 'Graminoids' is found in 'Herbaceous', or how much of the definition of a deciduous forest is found in an evergreen forest). Also in this case the LCCS software application is a 'black box' because the methodology is not explained in the manual (both versions).

For the calculation of semantic similarity between classes within LCCS-TM, there are various issues influencing the computation. Each parameter used in the definition of the class has the same weight. Weighting is not implemented since the relative importance of the individual parameters in LCCS-CM is linked to their hierarchical order as explained before. The parameters at the top levels of the classification system are those that define broad classes (e.g., the parameters 'Life form' and 'Crown cover' are used to define 'Closed trees'); subsequent parameters further refine the defined class (e.g., 'Broadleaved deciduous closed trees' or 'Multi-layered closed trees'). The order of the parameters in this way supports class comparisons because comparison will first relate to the broadly defined land-cover type to which the class belongs and then relate to differences within the land-cover type.

In the similarity computation the values attached to the parameter options are of a two-state character, i.e. either 1 (similar) or 0 (dissimilar). In the various methodologies listed in paragraph 3.3, the values often comprise the full range from 1 to 0 in order to express partial similarity (semantic distance). For example, in cases where properties are imposed on ordinal, interval or ratio scales, similarity can also be expressed as an exponential decay function of semantic distance (e.g., 'Very open (10-20)-40%' is less distant from 'Sparse' than from 'Closed') but such functions are not included in the LCCS-TM. Exceptions to the two-state character of the values in the LCCS-TM however exist. For example, the parameter 'Life form' has the option 'Woody' that is further subdivided into 'Trees' and 'Shrubs', the option 'Herbaceous' further subdivided into ‘Graminoids’ and ‘Forbs’, and the option ‘Lichens/Mosses’ further subdivided into ‘Lichens’ and ‘Mosses’. One has to ask “to what degree are ‘Woody’ and ‘Trees’ similar”? In such cases the LCCS-TM uses the arbitrary value of 0.5 (half similar or dissimilar) for semantic similarity between either ‘Trees’ and ‘Woody’ or ‘Shrubs’ and ‘Woody’, etc., and vice versa (Di Gregorio and Jansen 2000).

Figure 3-4. The stepwise procedure of defining the corresponding class of a Nordic class in LCCS



Wyatt *et al*. (1994) distinguished three types of matches (Figure 3-5): (1) source and target classes match exactly; (2) cases where the source class is a subset of the target class; and (3) cases where the target class is a subset of the source class. If one compares ‘Trees’ with ‘Woody’ the situation would resemble case 2, whereas comparison of ‘Woody’ with ‘Trees’ resembles case 3. The type of match is different and so too could be the value used for computing semantic similarity in LCCS-TM. The above-mentioned cases are further complicated as the parameter ‘Height’ also plays a role in their definitions (see also paragraph 3.5.1). Closer examination of the parameters ‘Life form’ and ‘Height’ reveals overlaps between the lower height limit for trees and the upper limit for shrubs (Figure 3-6). Similar overlaps exist also between the minimum height for shrubs and the maximum height for herbaceous life forms. This type of partial overlap is not considered in the similarity computation.

Figure 3-5. Type of matches between classes or between parameters

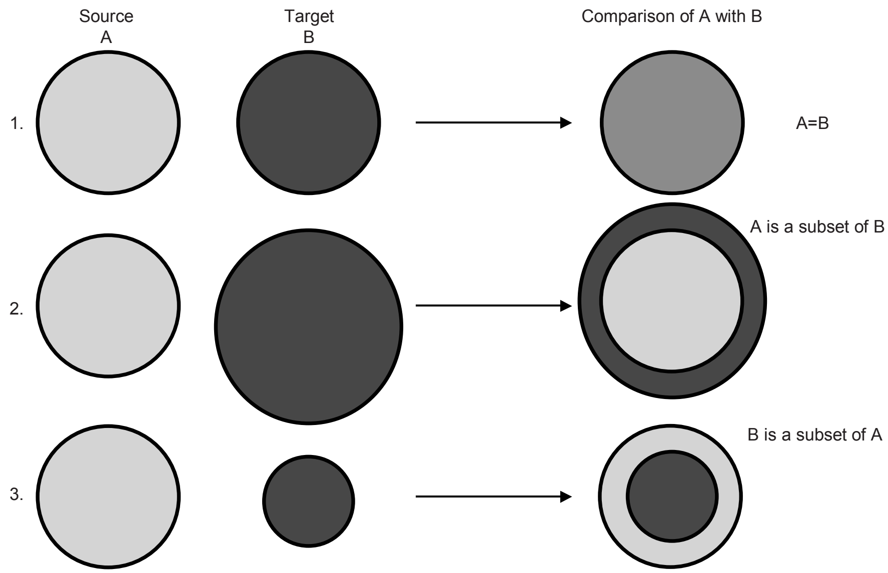
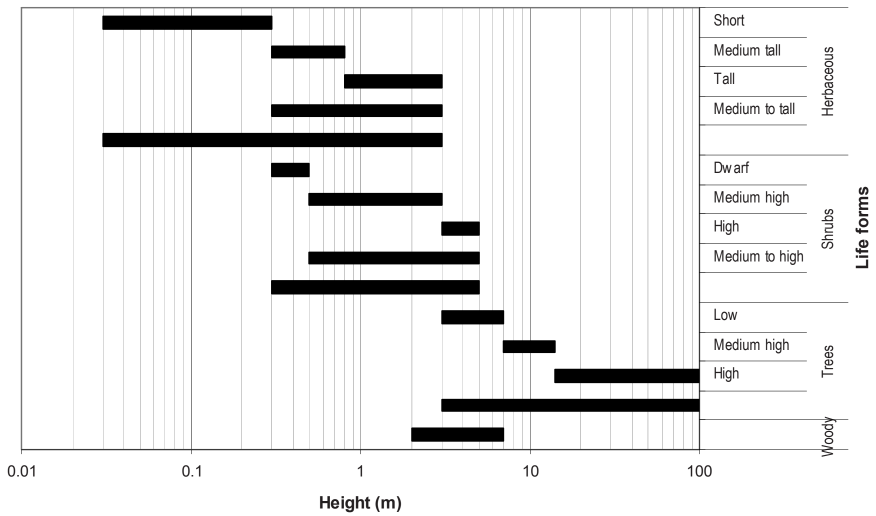


Figure 3-6. Overlaps in the LCCS parameter and modifier threshold values for the parameter 'Height' grouped according to the 'Life form' to which they apply [15]



[15] The definition of 'Dwarf shrubs' in LCCS is given as smaller than 0.5m, which can mean either 0-0.5m or 0.3-0.5m. Given that the lower threshold value for 'Shrubs' is set at 0.3m, use of 0.3-0.5m has been made for consistency's sake.

During the computation in LCCS-TM, the software application analyses similarity in two steps: first it will look for the source class parameter in the target class; second if the source class parameter is present in the target class, comparison will take place between source and target class. For example, the first step will analyse if the parameter 'Life form' from the source class is present in the target class, if this is so then the options will be compared and this could result in 'Trees' being compared to 'Graminoids' that are obviously dissimilar and the value 0 will be assigned to the parameter. This process based upon *commonalities* between two classes can be represented in mathematical form by:

$$S(\mathrm{a},\mathrm{b}) = \frac{|\mathrm{A} \cap \mathrm{B}|}{|\mathrm{A}|}$$

Equation 4

In this equation, *S* is the semantic similarity of the two categories, or classes, a and b being compared. A and B refer to the set of parameters belonging to category a and category b, respectively. $|A \cap B|$ refers to the number of parameters that belong to both category a and category b. The result of this equation is a similarity index on an interval scale, ranging from zero (dissimilar) to one (similar) that when multiplied with 100 gives the similarity in percentage. The semantic similarity is thus calculated based on the parameters that two classes share and the total number of parameters in the source class. If two categories a and b are compared then the perspective of the situation with category a as source is different from one with category b as source. If semantic similarity is the result of the *commonalities* and *differences* between two classes then the mathematical expression used by Feng and Flewelling (2004) is suitable:

$$S(\mathrm{a},\mathrm{b}) = \frac{|\mathrm{A} \cap \mathrm{B}|}{\left\| \mathrm{A} \cap \mathrm{B}| + \alpha(\mathrm{a},\mathrm{b})|\mathrm{A/B}| + (1-\alpha(\mathrm{a},\mathrm{b}))|\mathrm{B/A} \right\|} \text{, for } 0 < \alpha < 1$$

Equation 5

In this equation the same symbols are used as in Equation 4. In addition, use is made of: (1) A/B (or B/A) referring to parameters that belong to A (or B) but do not belong to B (or A); and (2) α is used to capture the possible asymmetries in semantic similarity between the two categories, where α(a,b) refers to the weight assigned to differences of parameters between a and b, while 1- α(a,b) refers to the weight assigned to differences of parameters between b and a. Also the result of this equation is a similarity index on an interval scale, ranging from zero to one. The equation requires input of parameters of the two categories being compared and the weights assigned to the two categories. The issue of assigning weights will be continued in paragraph 3.6.2.

Though the information richness of classes that include consideration of parameters other than those of land cover or environment can mostly be preserved in LCCS using user-defined parameters, the latter are excluded when computing semantic similarity and thus for such classes the index does not represent all the defining elements.

In general, the semantic similarity is highest in a group of classes describing the same land-cover category as they share the same set of parameters. Indices between classes belonging to different land-cover categories are in general small with the exception of cultivated areas and (semi-) natural vegetation because these land-cover categories contain a number of identical parameters to describe plants and their vertical and horizontal arrangements. As a consequence, a similarity can be found between for instance graminoid crops with a herbaceous type of (semi-) natural vegetation. Other land-cover categories, such as bare areas and built-up areas are dissimilar to any other land cover. Artificial water bodies, snow and ice can be compared to natural water bodies, snow and ice (see Figure 3-2).

3.6 Experiences of the NordLaM project with LCCS

The Nordic Landscape Monitoring (NordLaM) project of the Nordic Council of Ministers decided in 2002 to examine LCCS in the context of land-cover harmonisation at the semantic level using five different Nordic class sets from Denmark, Estonia, Norway (two class sets) and Sweden that are used in landscape monitoring (Groom 2004). LCCS was selected to act as mediator, or bridge, between the different ontologies of the class sets. These class sets with a total of 152 classes are from countries with similar types of landscapes but represent different approaches to land cover. The previous findings (Jansen 2004a) are critically re-assessed from the methodology viewpoint in this paper to underline the importance of (variation in) the semantic content of classes in harmonisation efforts.

The five Nordic class sets include both specific-purpose and general-purpose class sets. What follows is a short description of each class set to give a general idea to the reader:

- The ‘Area Information System’s Land-Cover Map’ (AIS-LCM) of Denmark comprises a general-purpose description of various land-cover types (11 classes) used in land-related research and administrative applications (Groom and Stjernholm 2001).
- The ‘Land-Cover Classification Scheme’ (EELC) of Estonia follows CORINE Land Cover until the third level with a fourth level comprising detailed vegetation descriptions for wetland land-cover types such as mires, (transitional) bogs and fens (Meiner 1999). Only a subset of the first 21 classes that refer to the coastal zone and wetlands was analysed.

- The ‘Monitoring Agricultural Landscapes’ (3Q) of Norway is a class set (57 classes) developed to monitor agricultural land-use patterns, biodiversity and cultural heritage; it contains a mixture of land-cover and land-use characteristics recording also events like trees blown over by strong wind, damage by hailstorms or area burnt by fire (Fjellstad *et al.* 2001).
- The ‘Digital Field Basis Map’ (DMK) of Norway covers 55% of the country excluding the area above the forest limit. The system focuses on land as a resource for agriculture and forestry (e.g., productivity, degree of cultivation, ploughing depth) and it is thus more geared to land use than land cover (8 classes)(www.skogoglandskap.no/filearchive/Dokument_02_03_nynorsk.pdf).
- The ‘Land-Cover Data’ (SMD) of Sweden is a general-purpose class set (55 classes) based upon CORINE Land Cover until the third level with a country specific fourth level including mires, the age and/or height of forest stands and land-use parameters for description of urban areas (Ahlcrona *et al.* 2001).

Detailed definitions of the CORINE land-cover classes to the third level, as are included in the Estonian and Swedish class sets, are provided by Bossard *et al.* (2000).

3.6.1 Categorisation issues using LCCS as a reference system

Harmonisation of the semantic contents of classes from different class sets using a reference system can be achieved on the condition that the reference system is flexible, can accommodate different classes and allows for acceptable compromises where the correspondence between original class and reference class is less than 100%. In order to be able to define corresponding classes in the LCCS-CM:

- The main LCCS parameters should coincide with the main parameters used in the original class sets;
- The hierarchical order of LCCS parameters should not impede defining the corresponding class;
- The LCCS threshold values in the definition of parameters and parameter options should coincide with those used in the original class sets;
- Information richness of the original classes should be maintained in the corresponding class;
- There should be fully developed concepts for whichever land-cover types are present; and
- In the original class sets definitions should be present and they should be unambiguous in order to establish correspondence.

The main parameters in two LCCS-CM land-cover categories of major interest for landscape level monitoring were analysed. The relevant classes in the Nordic class sets show that specific-purpose class sets use almost the full range of parameters to describe the cultivated area classes (Table 3-1), whereas in the (semi-) natural vegetated area classes the parameter use is more dispersed (

Table 3-2). Certain parameters or attributes have not been used at all or with only very limited use by these five class sets (e.g. ‘Field size’, ‘Crop type’ and ‘Spatial distribution’). This may be caused by the difficulty to apply such a parameter, or its being not clearly defined or explained, or its being not considered to be of (any) importance. The latter would justify moving such a parameter further down the LCCS-CM parameter hierarchy or ensure that such a parameter can be left out, as is indeed the case already for ‘Field size’ and ‘Spatial distribution’, while ‘Crop type’ is optional.

Use of the parameters ‘Leaf type’ and ‘Leaf phenology’ (leaf longevity) is more complicated because one cannot skip ‘Leaf type’ to define ‘Leaf phenology’ only. As a result, if ‘Leaf type’ could not be defined, consequently one could not add ‘Leaf phenology’. More flexibility in LCCS-CM would be required in this case. The almost complete absence of the use of the parameter ‘Stratification’ is noteworthy. One explanation of this is that the applications for which the Nordic class sets were created are not interested in the layering of the groups of life forms. Or, it may be that little layering is present in the described vegetation types possibly associated with climate. Or, possibly the use of the ‘Stratification’ parameter in LCCS-CM is not seen as straightforward and therefore was passed-over by the translators? Part of the answer may also be that the Nordic class sets are used with remote sensing data applications in which any layering underneath the highest crown cover cannot be identified on the satellite image or aerial photograph. The actual reasons for these patterns in parameter use are not evident.

Table 3-1. Overview of the use of parameters and two specific attributes in the major land-cover category cultivated terrestrial areas by the different class sets[16]

		Class set			
		AIS-LCM (Denmark)	3Q (Norway)	DMK (Norway)	SMD (Sweden)
Relevant number of classes:		2	14	4	3
Parameter:	Life form (obligatory)	2	14	4	3
	Field size	0	0	0	0
	Field distribution	2	14	0	0
	Crop combination	0	13	0	0
	Water supply	2	14	0	0
	Cultivation time factor	2	14	0	0
Attributes:	Crop cover	0	9	0	0
	Crop type	0	3	0	0

[16] The EELC class set is not represented as the first 21 classes do not contain relevant classes for this category.

Table 3-2. Overview of the use of parameters and one specific attribute in the major land-cover categories (semi-) natural vegetation (terrestrial (A12) and aquatic or regularly flooded (A24)) by the different class sets

	Class set									
	AIS-LCM (Denmark)		EELC (Estonia)		3Q (Norway)		DMK (Norway)		SMD (Sweden)	
Relevant number of classes	A12	A24	A12	A24	A12	A24	A12	A24	A12	A24
per category:	6	1	0	15	8	5	3	1	16	6
Parameter:										
Life form (obligatory)	6	1	0	15	8	5	3	1	16	6
Crown cover (obligatory)	6	1	0	15	8	5	3	1	16	6
Height	6	1	0	14	8	5	3	0	11	4
Spatial distribution (A12)	2	NA	0	NA	6	NA	0	NA	0	NA
Water seasonality (A24)	NA	1	NA	14	NA	2	NA	0	NA	3
Leaf type	2	0	0	0	4	1	3	0	9	0
Leaf phenology	2	0	0	0	4	1	3	0	9	0
Stratification 2nd layer	0	0	0	0	3	0	0	0	0	0
Stratification 3rd layer (A12)	0	NA	0	NA	1	NA	0	NA	0	NA
Attribute:										
Floristic aspect	0	0	0	15	0	0	0	0	0	0

Threshold values are related to the purpose of a class set and thus problems were encountered in establishing correspondence as the purposes of the class sets and LCCS-CM differ. The lack of coincidence in threshold values at high levels of the LCCS-CM has a much bigger impact on establishing correspondence than differences at lower hierarchical levels. For instance, the first parameter in the LCCS-CM is to distinguish between primarily vegetated and primarily non-vegetated areas with a threshold value of 4% vegetation cover for at least two months a year. The Norwegian 'Monitoring of Agricultural Landscapes' uses 25% crown cover to make the same distinction. If at least 4% of an area is vegetated LCCS considers the rest of the area to be empty, i.e. there are no other structures or occupied surfaces. If this were not so one would have to speak of a mixed class in which the vegetated area is subordinate to other land-cover classes (e.g., bare surfaces). Thus, the definition used by 'Monitoring Agricultural Landscapes' will either encourage creation in LCCS-CM of mixed classes or else disregard extremely sparse vegetation. In the case of creating mixed classes correspondence is established as a one-to-many relationship. A second example is the threshold value used for 'Crown cover', which is the second parameter in the (semi-) natural vegetated areas categories. In the Swedish 'Land-Cover Data' 30% is used for distinction between tree-dominated classes (Table 3-3). Thus, this class set comprises a parameter with a definition that came close to the 'Crown cover' modifier option of 40% in the first version of LCCS-CM. It is important to note that in such a case the use of a parameter option with a modifier is required. However, in the second version of LCCS

'Closed-to-open' with the option "*more than 40%*" has been added. LCCS contains more options for indication of 'Crown cover' than most existing categorisation systems but even so differences of 5 to 10% in threshold values occur. Table 3-3 further shows different thresholds between tree heights used in LCCS-CM and the 'Swedish Land-Cover Data': certain tree heights would lead to a different class (e.g., 6m). Such differences cannot be ignored.

In the Norwegian class set, the 'Digital Field Basis Map', the thresholds for 'Mixed forest' are unusual, with 20% crown cover as the lower limit and 50% crown cover as the higher limit for needleleaved trees; thus a forest area with 40% broadleaved trees is classified as needleleaved forest. This is probably due to the larger economic value of needleleaved trees and therefore this class is of prime interest to the forester. A cultivated area where trees are also present is described as 'forest' in this class set, whereas this would be a mixed class in LCCS-CM. In this class set an area is called 'forest' also when it is for the time being without trees as it will be used again for forestry purposes, thus substantiating that *category definitions* of land cover and land use are different. However, this 'forest' situation is analogous to the description of cultivated areas without crops in LCCS-CM, whilst (semi-) natural areas are described by their static and *in situ* land cover. Thus the spatial rules for land-cover description in LCCS-CM are distinct for different land-cover categories.

Table 3-3. Different threshold values of various parameters used in LCCS and SMD

Parameter	Parameter and modifier options	LCCS	SMD
Crown cover	Closed trees	>(60-70) %	
	Closed-to-open trees	>40%	>30%
		>(10-20) %	
	Open trees	(10-20)-(60-70) %	
		40-(60-70) %	
		(10-20)-40 %	
	Sparse trees	1-(10-20) %	
		4-(10-20) %	
		1-4%	
Tree height classes	High trees >15 m	>14-30m	>15m
	Medium trees	7-14m	>5m
	Low or young trees	3-7m	2-5m

Establishing correspondence between class definitions may lead to the case in which the original Nordic class found correspondence in several LCCS classes due to differences in threshold values, semantic ambiguity or occurrence of two different objects in a class. Thus the result of the correspondence is a one-to-many relationship. This was the case when a range was included in the definition, especially for the parameter 'Crown cover' being closed-to-open. This occurred in several vegetation types of the Danish 'Area Information System's Land-Cover Map' and the Swedish 'Land-Cover Data' class sets and in the 'Forest' class of the Norwegian 'Digital Field Basis Map' where the

vegetation can be either closed or open. In the first version of LCCS-CM one is being forced to create a mixed class creating a one-to-many relationship. In the second version of LCCS-CM the option 'Closed-to-open' has been included so one-to-one relationships can be established.
In other class names two clearly distinct types of objects co-occur, such as 'Fruit trees and berry plantations' in the Swedish 'Land-Cover Data' class set, and in various classes of the 'Area Information System's Land-Cover Map' class set. In such cases various options are possible; taking the 'Land-Cover Data' class as an example:

- A mixed class of fruit tree plantations with berry plantations is created because due to the mapping scale and/or the arrangement of fields these two types of fields cannot be spatially distinguished (one-to-many correspondence relationship).
- A single class is created containing the dominant crop fruit trees with the berries as a second crop because they occur in the same field. In this case it is a single class containing a multiple crop (N.B. it may also be possible that a single class exist in which the berries are dominant over the trees) (one-to-one correspondence relationship).
- A mixed class is created combining the two above-mentioned options (one-to-many correspondence relationship).

The best practise in such cases depends on how the two components are arranged spatially but this information was unavailable for the above-mentioned class. It may also happen that all options occur in practise but that this is not reflected in the original class set and thus poses problems when establishing correspondence. Here the problem is how to establish correspondence, as several options are available. Furthermore, as two types of correspondence relationships may be established one may well ask if these would influence the semantic similarity indices.

Information richness in the original classes should be maintained when defining the corresponding classes in the LCCS-TM. Therefore, the occurrence of land-use terminology in some classes related to the monitoring of land change (e.g., development of land-use patterns) calls for compromises in establishing correspondence. The difficulty is that semantic differences affect the interoperability between land cover and land use (Brown and Duh 2004). LCCS-CM is not dealing with land-*use* though some management related parameters are accommodated for cultivated area and built-up area classes. In the Norwegian class sets grasslands occur that are managed and thus belong to the cultivated areas category of LCCS-CM and some land-cover related cultural practises could be described. Grasslands that are abandoned and invaded by natural vegetation belong to the (semi-) natural vegetation category and thus land-cover related cultural practises could not be described. The other class sets also contain classes for which it was difficult to establish any correspondence such as 'Construction sites', 'Clear felled areas' and 'Burned areas' from the Swedish 'Land-Cover Data' class set that are relating to a future cover or an event that has removed and/or affected the cover. Here again the static and *in* *situ* description of land cover is requested when actually this is unknown for the classes concerned. For making correspondence, for ‘Construction sites’ a mix was chosen between built-up areas, unconsolidated and consolidated materials, whereas the classes concerning burned and clear felled areas were translated as a closed woody vegetation type with an added LCCS user-defined parameter explaining that it refers to a clear felled or burned area. In the draft document of the translation of the CORINE land-cover class set in LCCS (Herold *et al.* 2006a) these classes have been translated in a different manner. ‘Burned areas’ are translated as ‘(Semi-) natural terrestrial vegetation’ with a user-defined attribute and ‘Construction sites’ as an arbitrary mixed class of ‘Built-up areas’ with ‘Bare areas’. ‘Clear felled areas’ is a fourth level class that is specific for the Swedish class set. All these solutions are very subjective for two reasons: (1) such a solution depends heavily on who is establishing the correspondence; and (2) whether this solution or another one was adopted entails introducing uncertainty in the correspondence. One cannot expect to find a perfect match between an *actual* and a *potential* land cover. These type of phenomena and also damage due to hail storms or wind (e.g., in ‘Monitoring Agricultural Landscapes’), cannot be accommodated by any other means than adding user-defined parameters to preserve the information richness of a class, but user-defined parameters are not standardised.

Fully developed concepts are a prerequisite in a reference system. However, the occurrence of lichen-dominated areas with trees cannot be accommodated in LCCS-CM. The lichens concept that is adopted is extremely limited. It is impossible to link this ‘Life form’ with any ‘Stratification’. This is a significant drawback for the correct establishment of correspondence of vegetation types that include lichens. LCCS-CM cannot claim to be universally applicable as vegetation types in which lichens with trees occur, as are widespread in Nordic countries, cannot be appropriately described.

In the original class sets definitions should be present and unambiguous in order to establish correspondence. A problem occurs when a definition is not given, as occurred for the class ‘Sparsely vegetated areas’ in the Swedish ‘Land-Cover Data’ class set. These areas were translated as a mixed class containing unconsolidated materials and herbaceous open vegetation. The definition of ‘sparsely’ as used in the ‘Land-Cover Data’ class set is lacking and depending on it it could be argued that ‘sparse herbaceous vegetation’ should have been selected for the class correspondence. The corresponding class gives the impression that there are two elements present: (1) bare areas/bare soils with (2) open vegetation. But the concept of sparse vegetation in LCCS is not the same as a mixed class of bare soil with vegetation. Whichever solution is adopted, it means introducing uncertainty.

As illustrated with the above examples several problems were identified in establishing correspondences related to both the original class sets and the

LCCS-CM resulting in several cases in which a questionable solution for ‘translation’ was adopted and where (further) uncertainty was introduced.

3.6.2 Semantic similarity issues using LCCS as a reference system

Pair-wise calculation of semantic similarity between the corresponding classes of the original Nordic classes was performed in order to quantify the similarity, and inversely dissimilarity, of their semantic class contents. As an example the comparison of corresponding classes of the Danish ‘Area Information System’s Land-Cover Map’ is shown (Table 3-4). The darker the grey shading the more similar the classes. The matrix shows clearly that comparison of class A (source) with class B (target) results in a different semantic similarity as the comparison of class B (source) with class A (target); thus the indices in the matrix on both sides of the diagonal are asymmetrical. The matrix is based upon correspondence with the first version of LCCS. For a number of classes one-to-many correspondence relationships were established as the original Danish classes comprised either a range (e.g., open-to-closed in the forest classes) or two different objects (e.g., shrubs and grass in ‘Shrub and grass heath land’) or made no distinction where in LCCS a distinction is made (e.g., the kind of water bodies). Similarity indices showing high similarity further from the diagonal axis are mostly linked to the aquatic (semi-) natural vegetation type of class ‘Marshland’. This class can be regrouped with the terrestrial vegetation classes under the category of (Semi-) natural vegetation independent of the environment (e.g., aquatic or terrestrial) in which the vegetation type occurs (grey line in Table 3-4). The other, lower (but non-zero) off-diagonal similarity indices are caused by the occurrence of similar ‘Life forms’ between classes. Table 3-4 illustrates that the similarity within a category is in general higher than between categories. One should note that in case of mixed classes being present, only one element of a mixed class could be selected as source class in the computation. When selecting the source class in LCCS one can set which element of the mixed class will be the source class (e.g., see division of mixed classes 6, 7, 8, 10 and 11 where the figures 1 and 2 refer to the first and second element of the mixed class). This source class can be compared only to the first element of a mixed target class (comparison of 6.1 with mixed classes 7, 8, 10 and 11 refers to comparison of 6.1 with 7.1, 8.1, 10.1 and 11.1 respectively). Thus, in a case of mixed classes the semantic similarity calculated addresses only part of the classes present in both source and target class. The question raised earlier when there were various options for correspondence of the class ‘Fruit trees and berry plantations’ in the ‘Land-Cover Data’ class set can now be answered: in LCCS-TM the option selected for correspondence has repercussions for the semantic similarity.

Of more interest than the full correspondence matrices is, in the context of this paper, to better understand the computation of semantic similarities within LCCS-TM. A series of examples will illustrate how the choices made in the implemented algorithm influence the computed indices. Table 3-5, Table 3-6 and Table 3-7 show how the similarity is calculated and illustrate at the same time that the source class greatly determines the type and number of parameters in the computation. One should note that two types of null values occur: (1) null value to indicate that although the parameter or modifier is shared the options are dissimilar; and (2) the parameter or modifier of the source class is not present in the target class. The treatments of these different null values in the similarity calculation are identical. Examples 1 and 3 in Table 3-5 show the influence of the number of parameters of the source class in the computation. In both examples only one parameter is common. A parameter that is present in the target class but not in the source class is not considered. Example 2 and 4 in Table 3-5 are selected to show the influence of the decision in LCCS that whether 'Graminoids' is compared to 'Herbaceous' or 'Herbaceous' compared to 'Graminoids' the value is always 0.5. Example 2 shows the case that the source parameter is a subset of the target parameter, whereas Example 4 shows the case that the target parameter is a subset of the source parameter. Furthermore, comparing Example 1 with 2 and Example 3 with 4, one can see the influence of the arbitrary value of 0.5 in the resulting semantic similarity indices.

Table 3-6 illustrates how *modifiers* that define a parameter option in more detail influence the computation since when used they have the same weight in the computation as a parameter option. In fact one could argue whether it is correct to count a single element twice when a parameter option is present with its modifier. Table 3-7 shows another incongruity in that the weights of all parameters in the computation in LCCS are equal: if the first and most important parameter 'Life form' is dissimilar and only the 'Crown cover' similar a very high semantic similarity is calculated between two distinct vegetation types. Here the issue of assigning weights to parameters is important (see paragraph 3.5.2). When calculating semantic similarity according to Feng and Flewelling (2004) (Equation 5) weights can be assigned in two different manners: the α or the weights that are assigned to each pair of $A \cap B$, A/B, and B/A in Equation 5. For estimating α of categories within a single categorisation system, Rodriguez *et al*. (1999) suggested that the number of links from both categories to the immediate category that includes both categories can be used. But when using different categorisation systems this is impossible and a value of 0.5 can be assigned to α (thus 1- α is also 0.5). Weights assigned to each pair of $A \cap B$, A/B, and B/A can be related to the *depth* and *density* of the categories in the hierarchical categorisation system. Depth of two categories can affect semantic similarity measures because categories at lower hierarchical levels are more refined than those at higher levels. This implies that two categories at lower hierarchical levels are more similar in semantics than those at higher levels. The density of the categorisation system can also affect semantic similarity measures. This is because categories in a denser portion of the categorisation system (e.g., in LCCS in the primarily vegetated area land-cover categories) are closer in meaning than those in a less dense portion. If the denser portion of the categorisation system has many more categories than the less dense portion, the

Table 3-4. Semantic similarity of corresponding classes of the Danish AIS-LCM class set showing one-to-many correspondence relationships used as source class but absent from the target class (based upon LCCS version 1 with similarity in percentages)

	Target class										
Source class	1	2	3	4	5	6.1	7.1	8.1	9	10.1	11.1
1. Seasonally vegetation covered ground	100	100	0	0	25	0	0	0	25	0	0
2. Grazed or mown grass	100	100	0	0	25	0	0	0	25	0	0
3. Scrub or woodland	0	0	100	40	0	0	60	60	0	0	0
4. Shrub dominated heath land	0	0	40	100	0	60	20	20	0	0	0
5. Grass dominated heath land	25	25	0	0	100	25	0	0	75	0	0
6.1 Shrub and grass heath land (shrub component)	0	0	0	75	25	100	0	0	25	0	0
6.2 Shrub and grass heath land (grass component)	25	25	0	0	100	25	0	0	75	0	0
7.1 Deciduous forest (closed crown cover)	0	0	60	20	0	0	100	60	0	0	0
7.2 Deciduous forest (open crown cover)	0	0	40	0	20	20	80	40	20	0	0
8.1 Evergreen forest (closed crown cover)	0	0	60	20	0	0	60	100	0	0	0
8.2 Evergreen forest (open crown cover)	0	0	40	0	20	20	40	80	20	0	0
9. Marshland	25	25	0	0	75	25	0	0	100	0	0
10.1 Unvegetated ground (in built-up areas)	0	0	0	0	0	0	0	0	0	100	0
10.2 Unvegetated ground (in non built-up areas)	0	0	0	0	0	0	0	0	0	0	0
11.1 Open water (natural water bodies)	0	0	0	0	0	0	0	0	0	0	100
11.2 Open water (artificial water bodies)	0	0	0	0	0	0	0	0	0	0	100

Table 3-5. Semantic similarity between cultivated area classes of the 3Q and (semi-) natural vegetated area classes of the SMD class sets

Example	LCCS parameters	Source class	LCCS parameters	Target class	Similar parameter options	Semantic similarity
1		*3Q (Norway) 'A1ko'* [17]		*SMD (Sweden) 'Grass heath'*		25%
	A. Life form:	Graminoid crops	A. Life form:	Graminoids	1	
			A. Crown cover:	Closed-to-open		
	B. Spatial distribution:	Continuous			0	
	C. Crop combination:	(not used)				
	D. Water supply:	Rainfed cultivation			0	
	D. Cultivation time factor:	Permanent cropped			0	
2		*3Q (Norway) 'A1ko'*		*SMD (Sweden) 'Meadow'*		13%
	A. Life form:	Graminoid crops	A. Life form:	Herbaceous	0.5	
			A. Crown cover:	Closed-to-open		
	B. Spatial distribution:	Continuous			0	
	C. Crop combination:	(not used)				
	D. Water supply:	Rainfed cultivation			0	
	D. Cultivation time factor:	Permanent cropped			0	
3		*SMD (Sweden) 'Grass heath'*		*3Q (Norway) 'A3sb'* [18]		50%
	A. Life form:	Graminoids	A. Life form:	Graminoid crops	1	
	A. Crown cover:	Closed-to-open			0	
			B. Spatial distribution:	Continuous		
			C. Crop combination:	Single crop		
			D. Water supply:	Rainfed cultivation		
			D. Cultivation time factor:	Permanent cropped		
4		*SMD (Sweden) 'Meadow'*		*3Q (Norway) 'A3sb'*		25%
	A. Life form:	Herbaceous	A. Life form:	Graminoid crops	0.5	
	A. Crown cover:	Closed-to-open			0	
			B. Spatial distribution:	Continuous		
			C. Crop combination:	Single crop		
			D. Water supply:	Rainfed cultivation		
			D. Cultivation time factor:	Permanent cropped		

[17] A1ko=Fully cultivated or surface cultivated land covered by plants of which the seeds are used by people or animals.
[18] A3sb=Pasture; correspondence resulted in creation of a mixed class including natural and artificial grass covers (one-to-many relationship) of which only the first class is shown in this table.

semantic similarity measures between categories in these two portions of the categorisation system may be skewed compared with measures that are made on two categories from the same portion of the taxonomy. To account for these factors, it has been suggested that the number of links coming out of a category can be used as an estimate for density and the number of levels down in the categorisation system can be used as an estimate for depth. A weight can then be calculated based on the combination of these two estimates (Feng and Flewelling 2004). Table 3-5, Table 3-6 and Table 3-7 illustrate bias present in the algorithm implemented in LCCS. It is important that the current algorithm be changed to one that takes better into account the importance of parameters used in the definition of classes, the position of the class in the categorisation hierarchy, the type of match and especially one that includes a semantic distance function for partial overlap.

Table 3-6. Semantic similarity of (semi-) natural terrestrial vegetation and (semi-) natural aquatic or regularly flooded vegetation classes from the 3Q class set with (semi-) natural terrestrial vegetation classes of the SMD class set

LCCS parameters and modifiers	Source class:	Target class:	Similar paramete r options	Semantic similarity
Example 1.	*3Q (Norway) 'F2ri'* [19]	*SMD (Sweden) 'Thickets'*		25%
A. Life form:	Shrubs	Shrubs	1	
A. Crown cover:	Closed	Closed-to-open	0	
B. Height:	5-0.3m		0	
B. Height: (*modifier*)	<0.5m		0	
Example 2.	*3Q (Norway) 'M2bu'* [20]	*SMD (Sweden) 'Moors and heath land'*		25%
A. Life form:	Shrubs	Shrubs	1	
A. Crown cover:	Closed	Closed-to-open	0	
B. Height:	5-0.3m	5-0.3m	1	
B. Height: (*modifier*)	3-0.5m	<0.5m	0	
C. Water seasonality:	Waterlogged soil		0	
D. Leaf type:	Broadleaved		0	
E. Leaf phenology:	Deciduous		0	
F. Stratification:	Single layer		0	

[19] F2ri=Heath vegetation dominated by heather and brushwood.
[20] M2bu=Beach swamps that are at least 50% covered by bushes that are over 1m high.

Table 3-7. Semantic similarity between two (semi-) natural terrestrial vegetation classes from the Swedish SMD class set

LCCS parameters	Source class	Target class	Similar parameter options	Semantic similarity
	'Meadow'	*'Mixed forest not on mires or bare rock'* [21]		50%
A. Life form:	Herbaceous	Trees	0	
A. Crown cover:	Closed-to-open	Closed-to-open	1	
B. Height:		>30-3m		
C. Spatial distribution:		(not used)		
D. Leaf type:		Broadleaved		
E. Leaf phenology:		Deciduous		

3.7 Discussion

Differences in semantic concepts are often the key obstacle to data integration. One should realise that inconsistencies that hamper establishment of correspondence occur both in the Nordic class sets and in the LCCS-CM and that with the use of LCCS-CM as reference system a further level of uncertainty is introduced compared to direct comparison of the Nordic class sets. Correspondence can be either complete, partial or approximate at best, and in all cases it would be extremely useful being able to quantify the level of correspondence as this would give an idea not only of how much information was lost but also of uncertainty. It is unrealistic to expect that no information losses will occur but it is important that such losses are within acceptable, preferably quantified limits.

Establishment of correspondence has also implications for the class set structure because the corresponding class is not necessarily of a similar hierarchical level than the original class. More complicated is the situation in which one-to-many relationships are needed to establish correspondence. In such cases the corresponding classes may be of different hierarchical levels. Such changes in class set structure lead inevitably to changes in the data structure and hence to the scale hierarchy that were not examined in this paper.

In the assessment of LCCS as a reference system the impression prevailed in the NordLaM project that instead of defining the correspondence between a Nordic class and LCCS, one was establishing how much of LCCS was in the original class. A prerequisite for a reference system would be an approach in which classes can be accommodated that may call for compromises in the adopted

[21] The class 'Mixed forest not on mires or bare rock' contains a mixture of broadleaved with coniferous trees in which "*the share of coniferous or broadleaved species does not exceed 25% in the canopy closure*" (Bossard *et al*., 2000), whereas in LCCS 'mixed' is defined as "*each of the two components occupies at least 25 percent of the area"*. Consequently a mixed class needed to be created for correspondence of which the first class is shown in the table.

concepts and structure of the reference system in order to conserve information richness, but this type of flexibility is lacking in LCCS. The only way at present to store information richness that does not correspond or coincide to LCCS parameters is the application of user-defined parameters. However, these are not considered in the semantic similarity computation. In such cases it would be more meaningful to compute semantic similarity directly between the original Nordic classes.

In order to support the use of the LCCS categorisation methodology as a reference system there should be clear convincing advantages to counterbalance its semblance to a 'black box', since a formal definition of the categorisation rules is lacking. In the NordLaM project the use as a reference categorisation system alongside the existing system in the country allowed the user to fall back upon the well-known existing system and because both systems are used at the same time the learning curve of understanding the LCCS categorisation methodology may be less steep. When introduced as a new categorisation system, the user has no fallback option and thus has either to come to terms with the steep learning curve of a 'black box' or by depending on FAO for support. In the latter situation it would be important to reflect upon the implications of such dependence. For example, what should a user do who wants to apply LCCS in environmental monitoring and modelling using software packages without making a direct link to the LCCS software? Or, a user that basically needs to integrate spatial data? As shown by Ahlqvist (2005c) and Mayaux *et al.* (2006), one can take from the full LCCS methodology those elements that are useful in a specific application.

Though it seems that quite a number of class sets meanwhile have been translated (according to Herold *et al*. 2006b), the discussion of encountered problems and adopted solutions in these crosswalk 'translations' have not been made available to the scientific public apart from the example of 26 classes provided in McConnell and Moran (2001) and the draft document of Herold *et al*. (2006a) that was made available to the authors. The encountered problems and solutions, however, are a basis for further discussion in order to reach consensus. Feedback from the user and scientific communities will be indispensable in order to assess and enhance the current methodology.

The current semantic similarity algorithm in LCCS (version 2.0) is too simplistic to deal with the complexity of semantic similarity. It seems that many recent developments in the field of semantic similarity metrics have been overlooked in LCCS. The parameterised concept definitions of LCCS could be used to bridge between concepts in different categorisation systems and class sets. However, as Ahlqvist (2004) rightly points out LCCS-CM uses standard set theoretic representations without recognizing a semantic *space* underlying the concept representation, thus limiting the possibilities to measure in the LCCS-TM semantic similarity based on concept distance. Examination of the semantic similarity metrics in literature makes it evident that a thorough revision of the implemented LCCS-TM methodology is necessary and until that has happened its use is *not* recommended. Moreover, implementation of various methodologies for semantic similarity should be considered as each methodology has its own merits.

The NordLaM project selected LCCS as a reference system in order to establish semantic correspondence between Nordic class sets used in landscape monitoring with quantification of semantic similarity as the ultimate goal. However, in establishing correspondence with LCCS-CM uncertainties were introduced that could not be quantified, whereas the semantic similarity indices resulted in startling findings. Introduction of unknown quantities of uncertainty hamper the proper distinction between real changes from changes in categorisation. As a result there were no apparent convincing methodological advantages in using LCCS as a reference system, other than that the use of a reference system reduces the number of pair-wise class comparisons to be made.

Currently there is an urgent need to make the formal definition of the full methodology implemented in LCCS available to the user and scientific communities. The suggestion at the expert consultation in Artimino, Italy, in 2002 to set up a technical panel (FAO 2002 p.16; Jansen 2004b) that would receive feedback from the user and scientific communities and that would propose in a participatory way improvements of LCCS has, as far as we know, not been realised but such a panel could act as a forum to channel methodological improvements of the system. Such processes are important as FAO's intention is that the LCCS categorisation methodology should become an ISO standard (pers. comm. FAO). The critical examination in this paper, however, shows that there is not only room for improvement of LCCS but there is a real need, as there are various methodological issues raised in this paper that seem significant. If the LCCS is recommended as a basis for a land-cover harmonisation strategy one should be aware that the implemented methodology has a series of problems and shortcomings.

3.8 Recommendations and open research questions

The way forward to land-cover harmonisation is probably adoption of a parameterised approach such as implemented in LCCS. The advantages of such an approach are that the parameters with which classes are defined become explicit and class comparisons can be made in a systematic manner. However, correspondence needs to be accompanied by a mathematical theory that addresses uncertainty. As the NordLaM project experiences show there are quite a number of methodological issues for which in each individual harmonisation attempt so-called 'best practises' are developed but a wider consensus of such practises is lacking. Therefore each 'translated' land-cover class or data set risks to be a result that cannot be replicated by others in exactly the same way, no matter how many official organisations endorse such a 'translation'. This is a scientifically and practically unsound situation. More research is needed to improve existing and further develop methodologies.

It seems unrealistic to expect that land-cover standardisation will lead to worldwide adoption of a single categorisation system. Each class set has its own worldview and this is also true for LCCS as a categorisation system. Land-cover standardisation would lead to adoption of a single worldview, whereas land-cover harmonisation allows different worldviews to co-exist. The latter seems not only a much more flexible approach but also one that makes the world richer.

Semantic uncertainty is an inseparable companion of almost any information and that is certainly the case for harmonisation efforts in which different types of uncertainty accompany each other. At present, there is no accepted way to derive an overall score of the semantic similarity between two class sets and no measure to establish the success of correspondence at class level. Such indices would be particularly important when translating existing class sets into a reference system's terminology as they could indicate if the correspondence is close to the original class set and how well the fit between original class and corresponding class is. Such a quality assessment of the correspondence per class as well as per class set is suggested as analogous to the thematic accuracy assessment and has been suggested by Ahlqvist *et al.* (2000) and by Jansen (2006) for land use. Such quality statements are important if correspondence results are to be linked to semantic similarity indices as discussed here, and in the case that they need to be linked to land-cover *change* dynamics and the boundary conditions verified in the data validation effort are involved. If one wants to monitor gradual changes at the landscape level, then one has the necessity to be able to distinguish between real changes and changes in categorisation definitions. Quantitative semantic similarity metrics may help to better assess such differences, whereas at present there is often no explicit recognition of semantic differences in cases where two different class sets or categorisation systems are involved.

Methodologies for semantic similarity metrics should be evaluated using a single class set so as to assess the merits of the different methodologies. The cited examples from the literature describe each their own methodology applied in a particular area and it is difficult to compare the advantages and disadvantages of these methodologies. Furthermore, it would be interesting to see how semantic similarity indices of the different methodologies vary when applied to the same class set.

The possibility that metadata accompanying land-cover *data* sets should be extended to comprise more information on data and class accuracy, including the various levels of class (set) correspondence, should be further assessed and discussed. Especially in computer and information sciences a number of useful suggestions have been made but none seems to have become part of a metadata standard. This is, however, important in the context of SDI as it would inform data users much better as to what one can and cannot do with data.

The distinct aspects of spatial data integration that this paper has discussed briefly, i.e. adopted concept, spatial, temporal and quality aspects, should be considered *in parallel* to the semantic aspects. Harmonisation of land-cover data that deals solely with the semantic contents of the classes is a misrepresentation of the complexity of land-cover harmonisation. If the two different class and data sets to be harmonised are seen as two different 'objects', the harmonisation is the establishment of relationships between the two objects. The relationship between any two objects encompasses the assumptions that each makes about the other, including what operations can be performed and what behaviour results (Booch 1994).

4 Harmonisation of land use and land-use change class sets

4.1 Introduction

Land-use change knowledge has become increasingly important in order to analyse environmental processes and problems, such as uncontrolled urban development, deteriorating environmental quality, loss of prime agricultural lands, expansion of agriculture into areas that comprise either fragile ecosystems (e.g., wetlands and steep lands) or a high value with respect to biodiversity (e.g., humid tropical forests) or areas with a high incidence of diseases (e.g., malaria, river blindness). These processes and problems must be understood if living conditions and standards are to be improved or maintained at current levels (Anderson *et al.* 1976; Dumanski and Pieri 2000). Land-use change, as one of the main driving forces of (global) environmental change, is central to sustainable development (Meyer and Turner 1994; Walker *et al.* 1997; Walker 1998; Lambin *et al.* 2000). It is, therefore, essential to have detailed and in-depth knowledge of not only land-use processes and problems but also of land uses. Such information is required at multiple scales to support local, regional, state and cross-border co-operation.

Nowadays emphasis is shifting from static land-use data collection, and representation as maps, towards more dynamic environmental modelling in order to understand the past, monitor the present situation and to predict future trajectories (Lambin *et al.* 2000; McConnell and Moran 2001; Dolman *et al.* 2003). This suggests it is important to re-examine existing land-use data sets and attempt to harmonise them in order to make comparisons within and between countries and to compile time series with which to analyse the change dynamics and detect trends. Instead of a universally applicable land-use categorisation there is a need to develop tools aimed at facilitating the linkage between existing class sets. Data harmonisation will be required as it is unrealistic to work only with new standardised class sets, with major financial and intellectual investments having been made in existing class sets and survey programmes that use established methods of categorisation (Wyatt *et al.* 1998; Wyatt and Gerard 2001).

The Land-Use and Land-Cover Change (LUCC) programme element of the International Geosphere-Biosphere Programme (IGBP) and International Human Dimensions Programme on Global Environmental Change (IHDP) mentioned in their science/research plan that categorisation and data are a cross-cutting integrating activity for which data availability and data quality need to be analysed and a categorisation structure suitable for various requirements need to be devised (Turner *et al.* 1995). In addition, McConnell and Moran (2001) highlight two key issues:

- Both space and time considerations are essential for making land-use data compatible and hence comparable.
- Harmonisation of land-use categorisations includes harmonisation of land-use *change,* as we need to understand land-use change processes for decision making as explained above.

Furthermore, in a LUCC Workshop, held 8-10 February 2006 in Rome, a parameterised approach to harmonisation was advocated, as existing class sets are too label-oriented. Any discussion on harmonisation of land-use class sets should address not only existing or proposed (parameterised) categorisations but also data quality, space and time dimensions and land-use change.

In this chapter the harmonisation of land-use class sets, or correspondence between land-use class sets, will emphasize the *semantic* aspect of class sets consisting of the class definitions as these imply the parameters used in the formation of classes. Class descriptions contribute to the definition of boundary conditions that should be applied unequivocally and consistently when establishing correspondence between classes belonging to different class sets in order to avoid errors in data interpretation. The level of confidence with which such class correspondence is established is highest when the same parameters have been applied; differences in the applied parameters, and thus in boundary conditions, produce lower confidence levels. Complete correspondence is not always obtainable when harmonising data, thus there is a need to establish rules in order to reach the highest level of confidence possible.

4.2 Definition of the domain of interest

4.2.1 Land use

An international agreement on the definition and categorisation of land use is to this day inexistent, although many attempts were made previously (Guttenberg 1965; IGU 1976; Kostrowicki 1977 and 1992a; UNEP/FAO 1994; Baulies and Szejwach 1998; Duhamel 1998; McConnell and Moran 2001; Jansen and Di Gregorio 2002). Consequently, a common terminology is lacking. The term ‘land use’ has different meanings across disciplines and, as a result, implies a set of mostly unidentified parameters. These different perspectives on land use are, however, all valid. In the context of the present part of the chapter, land use is defined as “*the type of human activity taking place at or near the surface*” (Cihlar and Jansen 2001).

Land use is determined by natural, socio-economic, institutional, cultural and legal factors. In general, possible land uses are limited by biophysical constraints. These include climate, topography, soils and the geological substrate, presence or availability of water and the type of vegetation. Agricultural practices differ from one region to another and different types of land uses are practised on the same type of land in different areas, depending on the history, local traditions and way of life, apart from the biophysical constraints (Cihlar and Jansen 2001). The location of an area with respect to other types of land uses, such as residential and industrial areas, is also an important factor (e.g., the location of a commune close to main urban centres and its proximity to, for example, an airport) (Jansen 2003). Economic incentives as part of policy (e.g., the EU Common Agricultural Policy) can affect land-use patterns.

4.2.2 Categorisation

Categorisation, or classification, is defined as "*the ordering or arrangement of objects into groups or sets on the basis of relationships. These relationships can be based upon observable or inferred properties*" (Sokal 1974). Thus, categorisation denotes a process. The term 'classification' embodies two meanings (Duhamel 1998): (1) establishment of groupings of all objects in a given field (according to Sokal's definition); and (2) using the established groupings in order to decide the membership status of other objects (e.g., in remote sensing the imagery is used for the identification process of objects). The term 'classification system' includes not only the definition of the domain investigated and the categorisation process of the objects, but also a considered set of principles, or methodology, to assign individual land uses to land-use classes and their arrangement according to a set of adopted rules. Furthermore, it includes information for evaluating the reliability of assignment of objects to the various classes. Thus, the quality of the data should be documented.

Describing land use is to account for its character and different types of characteristics exist that can co-exist in a single class set (EC 2001):

- Two-state character (e.g., present/absent, 1/0, positive/negative, etc.); and
- Multi-state character: subdivided into:
  - Quantitative, that is discrete or continuous variables; and
  - Qualitative or so-called terminological variables.

Classifying all the objects in the domain of interest requires some basic principles, which have been described in detail elsewhere (e.g., EUROSTAT 1991; UNEP/FAO 1994; FAO 1997; LANES 1998; Duhamel 1998; Jansen and Di Gregorio 1998a; Di Gregorio and Jansen 2000; EC 2001). The key principles are:

- Completeness and absence of overlap of classes;
- Existence of definitions and explanatory notes;
- Existence of an index of objects;
- Spatial and temporal consistency; and
- Independence from scale and data collection tools.

Since many existing categorisations and map legends do adhere only in part to these principles, as will be demonstrated later, the use of the term ‘class sets’ has been preferred in the present part of the chapter.

4.2.3 Data standardisation and data harmonisation

Data standardisation is defined as “*the use of a single standard basis for classification of a specific subject*”, whereas data harmonisation is defined as “*the intercomparison of data collected or organised using different classifications dealing with the same subject matter*” (McConnell and Moran 2001). The understanding between data standardisation and data harmonisation is fundamental:

- Data standardisation will allow direct comparison of class sets but would disregard the financial and intellectual investments made in established methods and data sets; and
- Data harmonisation will allow countries and institutions to continue to use existing data systems and categorisations but when definitions are imprecise, ambiguous or absent problems may arise. Moreover, if many class sets are involved the number of pair-wise class combinations becomes excessive because comparison of n data sets requires $n(n-1)/2$ comparisons to be made.

The problem of excessive pair-wise class combinations can be resolved by developing a common reference system. Correspondence between classes may then be inferred from the explicit record of how each class relates to the reference system. The advantage is that translation of class sets into the reference system would be required just once. In addition, such a reference system would be well suited to form the basis for a generally accepted categorisation that could be promoted as future standard. At the same time a reference system could form the sound basis for a data model for use in geo-databases needed to manage information on land (Wyatt *et al.* 1993; McConnell and Moran 2001; Jansen *et al*. 2008b).

4.3 Basic units of measurement

Land use lacks a common unit of analysis, the so-called basic unit of measurement. The definition of this unit differs according to the purpose of data collection and/or analysis. Sometimes a statistical sample area is used, sometimes a mapping unit at a particular scale (e.g., minimum mapping unit in the case of thematic mapping), sometimes the cadastral parcel is used and sometimes a pixel or a grid cell is used in modelling and monitoring efforts. These four basic units of measurement are discussed in more detail below.

4.3.1 Cadastral land parcel unit

In many countries, the smallest land unit that one can define coincides with the cadastral land parcel unit, which is the lowest-level unit of the cadastre and thus has a legal status. In the cadastral system not only the spatial extent of these land parcels is recorded and their ownership but often also the occurring land-use or land-cover related information (e.g. arable land with building). In order to have a flexible approach in which different units of measurement can be aggregated, the cadastral land parcel can be selected as the basic unit of measurement for land use. These cadastral land parcel units can be regrouped according to ownership, by cadastral zone and by the various levels of administrative units (e.g., village, commune or district level). Furthermore, the land parcel units may be regrouped according to similar type of uses and socio-economic properties in order to identify land-use systems (e.g., if the different cadastral parcels are grouped at the level of ownership and/or leasing, the level of a socio-economic unit can be reached in which also the availability and use of technology can be incorporated). Thus, there is flexibility in the use and regrouping of the data that will serve different levels of decision making in land-use planning and policy. Another advantage is that land-use change analyses will be possible at a level that corresponds with decisions made by the individual landowner or landholder (e.g., in agent-based modelling (Bousquet and Le Page 2004)).

Land registration and the cadastre need to be seen as part of the process of natural resources planning and management. They deal with two of the world's major resources, i.e. land and information. Land information is necessary in many Government activities. The registers may be used for land taxation, the rights over public utilities over private land or along public roads for facilities such as electricity and water may need to be protected, infrastructures need to be maintained and/or improved, restrictions may be necessary where misuses occur, etc. The cadastre should therefore be seen as an integral part of the land management system (Dale 1995).

The use of the land is closely related to land rights, which may be associated with certain limitations or constraints. In addition, the period over which certain land rights are held is important. An owner that has land rights for a long period may be more inclined to make investments than one who has land rights for a very restricted period. Access to land and ownership may thus impede or restrict the use of the land. Land rights constitute a condition under which land use develops. Land rights may restrict the choice of the various options of land use and it is, therefore, an important determinant of what type of actual uses may be found in a particular place and time. The type of land rights and who is holding these land rights (e.g., individual, family or private company) are recorded in the cadastral system.

4.3.2 Land-cover polygon

Land use describes the use of the object 'land' and thus needs to be tied to a methodology in which the object is defined. This has led to the common practice to combine land use with land cover in the same class set, thereby attaching use to what you see because of what people do on the surface of the Earth and that can be observed by Earth-observing systems. However, land use has many aspects that go beyond land cover (e.g., socio-economics, cultural and legal aspects). Therefore, too much emphasis on land cover embodies the risk of not capturing several aspects of land use.

The advantages of using land-cover polygons as basic unit of measurement are that cover can be observed and that tools such as remote sensing and geographic information systems can help in a first stratification of the land-cover-related uses. Consequently, a spatial relationship is established between land use and land cover. A problem arises where land-use delineations do not concur with land-cover polygons. Several uses may take place within one land cover (e.g., in a building), as well as one land use may be applied to various land-cover types (e.g., certain types of free grazing). In the cases where the boundaries concur, one can aggregate either the land uses or the land covers. However, a land use may be confined to part of a land cover or parts of several land-cover polygons. In such cases, a further analysis and delineation would be required. In practice, most of the land has been designated a certain function that applies to the whole unit under consideration. The cases that a land cover with a specific function does not concur with the land use are rare (e.g., certain types of recreation or tourism) (Jansen and Di Gregorio 2003 and 2004a). A methodology for recording land use based upon available land-cover polygons is described by Cihlar and Jansen (2001). One should note that the land-cover/land-use relationship may change with time, thus establishment of the relationship alone is not enough.

4.3.3 Statistical sample unit

Statistics are often based upon a selection of areas that are representative for a much larger area, the so-called statistical sample unit. In Table 4-1 for instance, the TER-UTI class set uses an area of $9m^2$ distributed in a systematic manner over the country territory to do annual systematic observations. This methodology has also been applied in Bulgaria besides France. This provided, among other projects, the experience integrated into the Land-Use/Cover Area Frame Statistical Survey (LUCAS) launched by EUROSTAT and the Directorate General Agriculture (discussed in more detail in paragraph 4.7.3). LUCAS is making observations using a systematic grid: on a regular grid of 18 by 18km, each grid element contains 12x30 rectangular primary sampling units covering 90ha. In addition, there are 10 secondary sampling units per primary sampling unit. The secondary sampling unit area is considered as being equal to $7m^2$ (a circle with a diameter of 3m). These sampling units are revisited on a regular basis in order to describe them anew and analyse any changes. In 2005, this methodology has been revised in a regular grid of 1 by 1km covering the entire EU providing the base sample. From this base sample, the LUCAS master sample is extracted corresponding to a regular grid of 2 by 2km (e.g., 1 million points) where each point is photo-interpreted in order to stratify the sample in seven generic strata. From the stratified master sample, a sub-sample will be extracted for categorisation by field visit according to the full LUCAS class set (pers. comm. C. Duhamel, LANDSIS g.e.i.e).

4.3.4 The pixel and grid cell

The remote sensing community working with satellite images sometimes uses the pixel as basic unit of measurement. The basic character of digital satellite data is a two-dimensional array of discrete pixels. The value of each pixel corresponds to the average radiance measured electronically over the ground area corresponding to each pixel (Lillesand and Kiefer 2000).

Spatially explicit modelling of human land-use decisions and subsequent land-cover changes is often based upon a cellular model to which a number of spatial modelling techniques are applied (Parker *et al.* 2002). For instance in Albania, of which the harmonisation aspects are discussed in paragraph 4.7, cellular automata have been applied as described by Jansen *et al.* (2007). The basic unit of measurement in such cases is a grid cell as part of a two-dimensional array of discrete raster cells. The grid cell is also widely used in monitoring efforts (e.g., monitoring and forecasting of crop yields in early warning and food security applications).

4.4 Previous attempts at land-use harmonisation and standardisation

An important effort for establishment of an international recognized statistical system was made by the United Nations Statistical Division with the publication of the International Standard Classification of all Economic Activities (ISIC) in 1948 with four major revisions in 1958, 1968, 1989 and 2008 (UN 2008). A fourth revision is currently taking place and a revised draft is available on-line (www.unstats.un.org) showing the dynamic character of this categorisation using activity as main parameter.

The International Geographic Union established the Commission on World Land-Use Survey in 1949 (IGU 1976). A class set (legend) for a world map at a scale of 1:1,000,000 was developed combining land-cover characteristics with function. This scale was quickly abandoned in favour of national land-use surveys at much larger scales in Great Britain, Italy, Japan, Malaysia, Poland and Sri Lanka. Furthermore, the IGU established the Commission on

Agricultural Typology that tried in the period 1964-1976 to produce a system dedicated to agriculture. The work of this Commission was discontinued after 1976 though some of its members continued and completed the Types of Agriculture Map of Europe in 1983 (Kostrowicki 1977, 1984 and 1992b). Contacts with FAO were made in the early 1970s when the interest in a world agricultural classification increased due to the growing food crisis. In 1983, Kostrowicki proposed a land-use categorisation system, including non-agricultural land uses, which was a prime mover behind a proposal to UNESCO in 1987 for a world land-use map (Kostrowicki 1983a, 1983b and 1992a). However, nothing came of it.

The American Society of Planning Officials identified different dimensions of land use at an early stage (Guttenberg 1959, 1965 and 2002). The choice of the individual ownership parcel as basic unit of measurement, laid the foundation of a conceptually interesting and methodologically innovative categorisation system named 'Multiple Land-Use Classification System' in which land use is defined as a relationship between variables (Guttenberg 1959 and 2002). In a way this methodology is conceptually closer to geo-database systems than to just a 'classification'. Guttenberg (1965) also identified different 'modes' for categorisation: referential, appraisive and prescriptive (Figure 4-1). However, most of the existing categorisations remain in the referential mode, as it is the most neutral one, and frequently deal with observable characteristics, such as land cover and actual activity, and derived characteristics, such as function and legal aspects. The appraisive mode casts land use in the light of social interests and values that differ according to local prevailing customs.

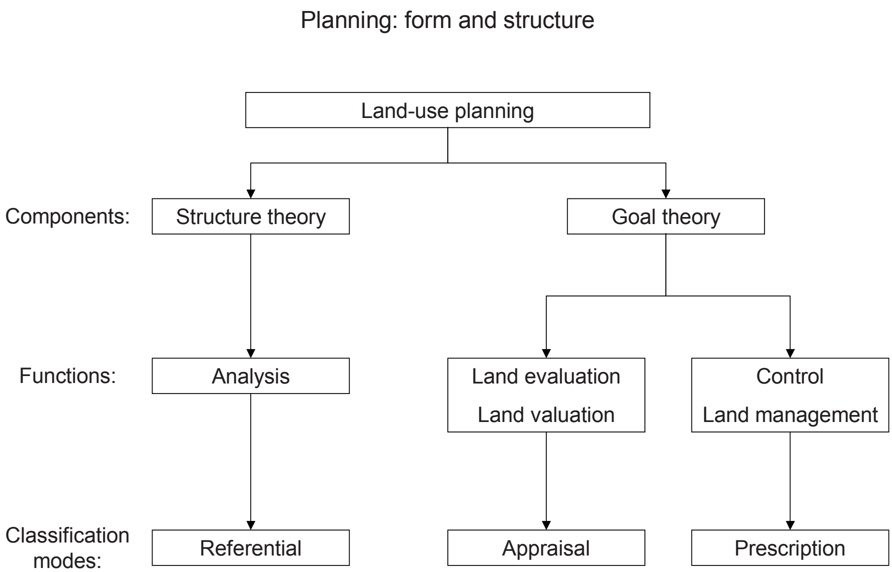
In the period 1969-1971, a study was made by the Commission on Geographic Applications of Remote Sensing of the Association of American Geographers. The results were published in 1971 by Anderson and further elaborated in 1976 (Anderson *et al.* 1976). This remote-sensing driven categorisation was based upon the World Land-Use Survey system (Paludan 1976) and evolved in the period of the first LANDSAT launch. The system represents the traditional subdivision in land-use terminology for built-up and agricultural lands, and land-cover terminology for natural vegetation, water, snow and ice.

The Economic Commission for Europe of the United Nations proposed a Standard International Classification of Land Use that would allow comparison of national land-use statistics (UN-ECE 1989). However, this is a mixture of land-cover and land-use terminology and the classes are not exhaustive.

The interest in reviewing and updating the U.S. Standard Land-Use Coding Manual (Urban Renewal Administration 1965) led to the initiative of the Land-Based Classification Standards (LBCS) project, co-ordinated by the Research Department of the American Planning Association in corporation with several U.S. departments and agencies (APA 1999). This effort is based upon recognition of various categories in which land use is traditionally classified:

activity, function, structure-type, site development character and ownership. These categories have each there own set of characteristics and categorisation takes place across these multiple categories. The effort addresses many of the problems that previous systems had but remains at the level of a system divided into several descriptive classes. The choice of categories may be disputed.

Figure 4-1. Analysis of land-use planning (adapted from Guttenberg 1965)



The Land Utilization Type (LUT), as developed by FAO in the Framework for Land Evaluation (FAO 1976) and in the Guidelines for Land-Use Planning (FAO 1984), has been widely used as a generalised description of agricultural land use in terms of inputs, two levels only, and outputs for which suitability could only be defined imprecisely. This concept was based upon a shortened list of the land-use variables identified by the IGU, the difference being the application of a qualitative land-use description in the Framework. The concept was too imprecise to be applied at farm level or for production planning, it contained only one (plot) level and reflected more a potential than an actual land-use class, while being qualitative in nature. One should note, though, that this concept was adapted to the requirements of a land evaluation system and as such, it has been used in numerous regional or district crop suitability, capability and pre-feasibility studies (pers. comm. F.O. Nachtergaele, FAO). The matching of potential agricultural land uses with the land through a series of decisions and ratings yielded into a quite complicated expert system, thus the methodology became the reverse of being transparent. In the late 1980s, at FAO attempts were undertaken to improve the LUT concept. The matching of precisely defined qualities and characteristics of the land unit with broadly, usually qualitative LUTs resulted in the limited use of the quantitative land resource data. A series of FAO commissioned studies was initiated as well as collaboration with UNEP (Remmelzwaal 1989; Adamec 1992; Muecher *et al.* 1993; UNEP/FAO 1994; ITC/FAO/WAU 1996; Wyatt *et al.* 1998). Adamec (1992) was the first to define the agricultural land-use type as "*a series (or sequence) of operations (or activities) carried out (or undertaken) to produce (or harvest) products or benefits for consumption or sale*" but he recognized at the same time the difficulty to apply individual operations or their sequence and dates of execution as parameters plus the inputs already employed. Nonetheless, this definition was adopted by the ITC/FAO/WAU effort resulting in the Land-Use Database (1996). In this database, the primarily agricultural land-use class is independent from scale, the basic unit being the plot. The database permits user-defined hierarchical structures, comparison, and a number of standardised parameters are included. However, the database allows users to add or change parameters and definitions along with the order of parameters to fit a specific aim. If the objective of categorisation is a contribution to data harmonisation and data standardisation, another approach should be selected. The study by Wyatt *et al.* (1998) was an effort at outlining the parameters to be used for globally applicable definition of land uses. The idea of analysis of existing systems in order to extract the set of parameters to be used for building a reference system would have been valid if existing categorisations were used. However, the analysis was based upon a number of legends, hence indicating gaps in the completeness of land-use classes and parameters used. Duhamel (1998) clearly identified that the above-mentioned studies and some selected national class sets suffer from the lack of systematic analysis of what defines land use, in addition to the insufficient adherence to the fundamental principles of categorisation mentioned earlier (Table 4-1).

The current view of the way forward is to promote a *parameterised approach* to categorisation. The explicit use of quantitative parameters will facilitate harmonisation between class sets if the same set of parameters is used. In many existing class sets one will find (Jansen and Di Gregorio 2002):

- Inconsistent application of land-cover or land-use parameters, i.e. land-cover parameters are being used to distinguish land uses and vice versa;
- Inconsistent use of parameters at same level of categorisation, i.e. in one category a certain parameter is used and in a related category a completely different one is used;
- Use of different parameters between classes, i.e. for subdivision of a class into three subclasses more than one parameter is used; and
- Use of non-inherent characteristics, i.e. using characteristics that are not related to the subject but describe, for instance, its environment (e.g., climate, physiography, altitude from a DEM, etc.).

Table 4-1. Overview of adherence of selected systems to the general principles of categorisation (based upon Duhamel 1998).

	Land-use class sets							
Principles	Anderson (1976)	UN-ECE (1989)	UK Land use (1972)	TER-UTI (1994)	Remmelzwaal (1989)	Adamec (1992)	Young (1993)	Muecher *et al.* (1993)
Completeness	Only applicable to USA (as intended)	More applicable to northern European countries than Mediterranean	Fulfilled	Fulfilled	Not fulfilled, too much geared towards agriculture	Fulfilled	Not fulfilled, some land uses are missing	Fulfilled
Absence of overlap	Absent	Absent	Fulfilled	Fulfilled	Potential confusion between mixed classes	Absent	Potential confusion between certain classes	Fulfilled
Observation unit	Not addressed	Not addressed	Spatial unit, hereditary and zone	A circle of $9m^2$	Not addressed	Not addressed	Management unit; population census unit	Unit of biophysical management (e.g., plot)
Tool independent	Remote sensing dependent	Not addressed	Independent	Independent	Not addressed	Not addressed	Not addressed	Independent
Definitions and explanatory notes	No systematic reporting, definitions read like comments	Definitions exist; unsystematic explanatory notes	Non-existent	No definitions; explanations are given	No systematic definitions; no explanatory notes	Non-existent	Non-existent	Non-existent
Interpretation rules	LC and LU are mixed	General usage of mixed classes; Dominance is not defined	All uses are recorded, no weighing in mixed classes	Hierarchy of uses requested but no rules exist to define	Multiple uses are not discussed	Non-existent	Multiple uses are recognized: primary use refers to the value added to the holding	Multiples uses exist; multiple sequences of operations are not dealt with
Inclusion of new objects	Not mentioned, there is often an Other category	No systematic approach followed	Possible	Possible	Not addressed	Not addressed	Not addressed	Not addressed
Index of objects	Non-existent	Non-existent	Existing	Existing	Non-existent	Non-existent	Non-existent	Non-existent
Correspondence with other systems	Non-existent	Tentative correspondence with socio-economic systems	National Standard Industrial Classification (SIC)	Table of correspondence with earlier versions	Non-existent	Non-existent	Non-existent	Non-existent

Although the underlying reasons for making subdivisions based upon different parameters may be valid, they show that parameters do not always have the same weight in making distinctions. Such decisions are usually not well documented in the accompanying reports of the class sets. This hampers harmonisation of class sets, as re-interpretations of not well-documented decisions are likely to differ between persons within one country and between countries. The actual class sets make an insufficient contribution to data harmonisation and standardisation. Efforts to increase harmonisation and standardisation do not necessarily lead to less pragmatic decisions on the choice of parameters. The focus should be on the logically and functionally consistent application of a set of inherent land-use parameters that are clearly separated from non-inherent parameters (Jansen and Di Gregorio 2002; Jansen 2003).

However, if an international agreement on the definition of land use is or cannot be reached and a common terminology found, data harmonisation will remain an impossible task, let alone attempting data standardisation. It is therefore important to underline commonalities in the existing approaches and identify a set of commonly used parameters in existing and widely applied class sets.

4.5 Major parameters for harmonisation of class sets

A set of necessary parameters to describe land use could form a basis for facilitating the linkage between existing class sets. These parameters, once identified and defined uniformly, will allow -through combinations- the definition and grouping of land uses for a variety of class sets. Some ranking may be proposed to limit the number of parameters.

An analysis of several existing class sets shows that statistical data are often collected on the basis of economic purpose and/or activities (UN-ECE 1989; UN 1989; UN 1998; UN 2008), natural resources related disciplines tend to amalgamate land-cover characteristics with activity or function (Anderson 1976; IGU 1976; CEC 1995; FAO 1998), while legal aspects are described by land rights or patents and other related legal conditions (FAO 1998; UN 1998). Table 4-2 provides an overview of the most commonly used major parameters applied by various international systems. ‘Function’ refers to economic purpose, ‘activity’ refers to a process resulting in a similar type of products, ‘biophysical’ refers to the material and immaterial environment (e.g., vegetation, land cover, geology, etc.) and ‘legal’ refers to the context of existing laws and regulations.

Table 4-2 shows that the major land-use parameters utilised by sectoral class sets are limited. Though the meaning of land use varies widely among sectors, the set of major parameters is apparently not so broad. Just two parameters suffice to describe any land use: ‘function’ and ‘activity’. The function approach describes land uses in an economic context. This type of approach answers the aim of land uses and is commonly used in sectoral land-use descriptions (e.g., agriculture, forestry, fisheries, etc.). The approach is able to group land uses together that do not possess the same set of observable characteristics but serve the same purpose, the so-called polythetic view (Sokal 1974). An example of such land uses is 'agriculture' that may come in many forms, dealing with plants or animals, related to extraction, production or service characteristics. These 'agricultural' land uses share a large proportion of characteristics but do not necessarily agree in any one characteristic (e.g., bee-keeping versus annual rainfed maize cropping agree in the removal of biomass but differ in many other characteristics). The activity approach describes what actually takes place on the land in physical or observable terms. Activity is defined as "*the combinations of actions that result in a certain type of product*" (UN 1989) and refers to a process. The term 'activity' does not mean that one needs to witness the activity as observer at the moment that it is being carried out, but one may observe the results and infer the activity. It is important to note that the function approach is independent of the activity approach: a variety of activities may serve a single function (e.g., both farm housing and farming activities serve agriculture). Thus, 'function' and 'activity' could form the key parameters for an underlying overarching concept to describe whatever type of land use.

Table 4-2. Analysis of land-use characteristics used by several main class sets [22]

Main sector	Land-use characteristics			
	Function	Activity	Biophysical	Legal
Agriculture	x	x	x	
Fisheries	x	x	x	
Forestry	x	x	x	x
Economics	x	x		
Sociology	x	x		
Statistics	x	x	x	
Industry	x	x		x
Housing	x	x	x	x
Services		x		x

At the lower levels where the 'activity' approach is used, parameters could be based upon the three main elements that characterise 'activity': (1) input of resources, (2) production process and (3) output product(s). The concept of input-output could also be termed import-export. This concept is able to address various issues among disciplines such as cycles, fluxes, emissions and intensities needed in assessments of interactions between land-water, land-atmosphere, etc.

Widely known and used systems for economic activities are: (1) the 3rd revision of the ISIC of the United Nations Statistical Commission (UN 1989) (Table

[22] Based upon: World Land-Use Survey (IGU, 1976), Anderson (Anderson *et al*., 1976), ISIC 3rd revision (UN, 1989), Standard International Classification of Land Use (UN-ECE, 1989), NACE 1st revision (CEC, 1993), Central Product Classification (UN, 1998), FAOSTAT (FAO, 1998), Land-Based Classification Standard (APA, 1999). For 'forestry', use was also made of http://home/att.net/~gklund/DEFpaper.htm.

4-3), which ensures harmonisation with other main economic categorisations, such as the Central Product Classification (UN 1998) (the CPC was developed for the purpose to measure outputs, i.e. products and services. Each category is accompanied by a reference to the ISIC class where the output is mainly produced (industrial origin parameter), categorisation of products is based on the physical characteristics of the goods or the nature of services rendered); and (2) the Nomenclature des Activités de la Communauté Européenne (NACE) of the Commission of the European Communities, which first two levels are compatible with ISIC (CEC 1993).

Table 4-3. The main categories of ISIC 3rd revision and the draft for the 4th revision

International Standard Classification of All Economic Activities			
Code	3rd revision (UN 1998)	Code	4th revision (UN 2008)
A	Agriculture, hunting and forestry A-01 Agriculture, hunting and related service activities A-02 Forestry, logging and related service activities	A	Agriculture, forestry and fishing A-01 Crop and animal production, hunting and related service activities A-02 Forestry and logging A-03 Fishing and aquaculture
B	Fisheries	B	Mining and Quarrying
C	Mining and quarrying	C	Manufacturing
D	Manufacturing	D	Electricity, gas, steam and air conditioning supply
E	Electricity, gas and water supply	E	Water supply; sewerage, waste management and remediation activities
F	Construction	F	Construction
G	Wholesale and retail trade	G	Wholesale and retail trade; repair of motor vehicles and motorcycles
H	Hotels and restaurants	H	Transportation and storage
I	Transport, storage and communication	I	Accommodation and food service activities
J	Financial intermediation	J	Information and communication
K	Real estate, renting and business activities	K	Financial and insurance activities
L	Public administration and defence	L	Real estate activities
M	Education	M	Professional, scientific and technical activities
N	Health and social work	N	Administrative and support service activities
O	Other community, social and personal service activities	O	Public administration and defence; compulsory social security
P	Private households with employed persons	P	Education
Q	Extra-territorial organizations and bodies	Q	Human health and social work activities
		R	Arts, entertainment and recreation
		S	Other service activities
		T	Activities of households as employers; undifferentiated goods- and services-producing activities of households for own use
		U	Activities of extraterritorial organizations and bodies

The usefulness of the ‘function’ and ‘activity’ parameters is apparent. ‘Function’ groups all land used for the same economic purpose independent of the type of activities taking place, whereas ‘activity’ groups all land undergoing a certain process resulting in a certain type of product that may serve different functions. The result of a combined approach will be a flexible data set where re-grouping of parameters can take place for a wide variety of queries.

The level of data collection increases notably from the ‘function’ to the ‘activity’ concept. The use of the ‘function’ parameter as first level parameter is proposed as a pragmatic choice as most major functional groupings can be detected with limited investment of human and financial resources, whereas the ‘activity’ parameter would require substantial investments in data acquisition.

4.6 Data quality

Harmonisation of class sets requires the analysis of data quality because correspondence between two class sets having two very different qualities may not be meaningful. In the metadata of each class set, parameters should be described related to the *positional and thematic accuracy*. The positional accuracy when using remote sensing can be divided into:

- Geo-referencing, i.e. the technical solutions for projecting the imagery onto the selected projection and spheroid aiming at providing for each pixel on the image its position on the ground by the means of a tern of coordinates.
- Location control, i.e. the correspondence between the coordinates of any arbitrary chosen point on the image and its position on the ground by the confrontation with better accuracy source data.
- Registration, i.e. the precision of the drawing/digitising system adopted defined as the difference between the same lines when interpretation is repeated of the same feature.

A statistically valid design for estimating accuracy parameters has three parts: (1) the response design specifies which data are to be collected at each sample location; (2) the sampling design specifies the locations at which the response data are to be acquired; and (3) the analysis lays out the formulas and tests to be applied to the observations (Strahler *et al.* 2006).

One of the most common means of expressing thematic accuracy in remote sensing is the preparation of a classification *confusion matrix*, sometimes called *error matrix* or *contingency table*. The confusion matrix compares on a class-by-class basis, the relationship between known reference data, i.e. the ground truth, and the corresponding results of classification either in the form of pixels, cluster of pixels, polygons or groups of polygons (Lillesand and Kiefer 2000).

Semantic harmonisation of class sets should consider the data quality aspect in a comprehensive manner and would need to address also the following two aspects that are still at the level of research (Jansen *et al*. 2008b):

- A quantitative measure should be provided of the harmonisation result of a class. In existing examples, the impression is often given that class correspondence is 100%, whereas more often than not the result will be much lower.
- A quantitative measure should be provided for the overall correspondence between two class sets similar to the overall accuracy calculated from the confusion matrix.

4.7 Example: land-use harmonisation in Albania

4.7.1 Use of a reference system and a data model

The land-use data harmonisation process is illustrated with an example form the EU PHARE Land-Use Policy II (LUP II) project in Albania based upon the cadastral parcel as basic unit of measurement (Jansen 2003; Jansen *et al.* 2007). The LUP II results are compared to the World Bank Albanian National Forest Inventory (ANFI) project based upon the land-cover polygon as basic unit of measurement and a class set defined with the Land-Cover Classification System (Jansen *et al.* 2006a). The Albanian Government needed an analysis of land-use change dynamics to better understand the past, monitor the current situation and to predict future trajectories in order to plan land uses and develop and implement appropriate policies. In the example, data quality aspects have not been quantified, as the basis for the harmonisation effort is the cadastre, where in the past land use has been systematically recorded, implying high data accuracy.

A standard hierarchical methodology for description of land use has been developed for Albania, as there was no such methodology available or an international land-use reference system. The developed Land-Use Information System of Albania (LUISA) adopts the ‘function’ and ‘activity’ parameters for systematic description and has been developed in complete synergy by the subject-matter specialist and information technology specialist.

Harmonisation between class sets can be achieved on the condition that the data structure of existing data sets is integrated in the newly developed class set. Here, problems may arise and if so they should be overcome. It may mean having to compromise and accommodate certain classes in a specific position in the class set that is neither the most suitable when considering the concepts adopted nor enhancing the class set’s internal consistency. Adoption of a hierarchical system will allow the applicability at various scales, from national, regional, to local. In addition, the class set structure is linked to a data structure, so one should not only be familiar with the subject matter of land use and the principles of categorisation, but also with information technology concepts (e.g., relational databases or object-oriented approaches). In the above discussion, it is assumed that a common set of attributes distinguishes the classes to be compared and that class differences are primarily due to differences in boundary conditions. In the case of land use, this is a reasonable assumption (Wyatt *et al.* 1998).

In the context of the LUP II project, four data sets covering the period 1991-2003 (e.g., under socialist Government, before and after privatisation (i.e. transfer of ownership)) are important:

1. Statistical data from the Institute of Statistics (INSTAT) comprising seven classes;
2. Cadastral data from the Immovable Property Registration System (IPRS Kartela) comprising 41 classes (spatially explicit data);
3. Commune data comprising 14 classes (spatially explicit data); and
4. LUISA data comprising 48 classes where the most detailed levels of the hierarchy were used for land-use data collection (spatially explicit data).

Correspondence between classes of the available class sets has been inferred from the explicit record of how each class relates to LUISA using the available definitions. Three class sets would lead to three comparisons to be made for each class, whereas four class sets would request six comparisons per class. It was therefore more efficient to use LUISA as reference system. During its development, LUISA has been systematically and thoroughly tested. For the purpose of the LUP II project the land-use categories have been limited to four that each are linked to a set of laws in the country. Each of these categories branches out into different levels, each level having its own set of classes and use of parameters, definitions and guidelines as described in detail in Jansen (2003) (Figure 4-2).

The classes presented in Figure 4-2 remain at a general level because more detail can be provided in combination with other data in the developed geo-database. For example, the LUISA Agricultural land uses can be linked to agronomic and land tenure data that are kept in separate data layers to have more flexibility in the geo-database. Combination of a class such as ‘Arable land’ with crop specific data will give more information about the crop type, fertilizers and pesticides used (inputs used), yield (output), etc. In addition, combination of a class such as ‘Actually not cultivated land’ with land tenure information such as ‘Not registered’ may indicate causes or constraints why agricultural land is not being used. This is of particular importance in the context of land-use planning and policy (Jansen 2003).

A link that is often ignored at an early stage of categorisation comprises the structure of data resulting from categorisation in a geo-database. The *data model* developed for the LUP II project distinguishes *spatial features* (e.g., land use and soils) from *linear features* (e.g., roads and channels) (Carrai 2003; Jansen

2003). The latter two classes are however also related to land use because roads form the transport network, whereas channels form the drainage and irrigation network. This division has important implications for the way in which roads and channels appear in LUISA. In the data model, linear features have been split into several segments; for each road or channel segment data is collected that deal with their actual state and maintenance. The advantage of having such segment information is that the user of the data can identify, for example, if anywhere on a road used for transporting agricultural products to the nearby market there is a segment that is in such a bad state that a vehicle cannot pass. If the road would be a single feature in the database, such an analysis would be impossible. Another example can be given using channels. In many class sets, one will find the class 'irrigated agriculture' where the parameter irrigation is applicable to the whole polygon. In practice, irrigation channels may function only in part due to their maintenance state but such a polygon would still carry the parameter irrigated. A much more flexible approach is to separate irrigation channels from the agricultural fields and to split the channels in segments. Such a distinction permits the user to identify those fields that are actually irrigated from those that cannot be irrigated due to segment information on the state of the irrigation channels. One will thus not find every single possibility of a class in LUISA because of the data model adopted. It is sufficient to record roads and channels as land *cover* because the segment information can be combined with these features at a later stage of the data integration process in order to define land use.

Once correspondence with LUISA was established for each class of the class sets, land-use change could be analysed using just LUISA. Using different class sets with several classes results in numerous land-use changes making a meaningful analysis difficult. LUISA does not only act as a reference system for harmonisation of land-use class sets, it also acts as a reference system for harmonisation of land-use change. The LUISA class structure, i.e. the data structure, is tailored in an efficient and logical manner in order to identify land-use change processes. In principle, *land-use modifications* occur within a land-use category and the degree of modification depends on the level of the class (e.g., at Level IV modification is small, at Level III medium and at Level II high) and *land-use conversion* occurs between land-use categories. The exceptions are the Non-agricultural land-use classes, where modifications occur within one group (e.g., within Urban uses, within Transport, within Utilities, etc.) and conversions between groups (e.g., from Unproductive to Urban uses, or from Water bodies & waterways to Extraction & mining). In the Agricultural, Forests and Pasture & Meadows land-use categories conversions occur between categories, whereas modifications occur within a single category within and between groups (e.g., within the Agricultural Land-uses modifications exist within Permanent Crop Cultivation or between Temporary Crop Cultivation and Permanent Crop Cultivation, etc.). For the interpretation of land-use change a piece of software was written, the Land-Use Change Analyses (LUCA), that groups the changes

Figure 4-2. Overview of the LUISA class set with the four main categories of land use (Jansen 2003; Jansen *et al.* 2007)

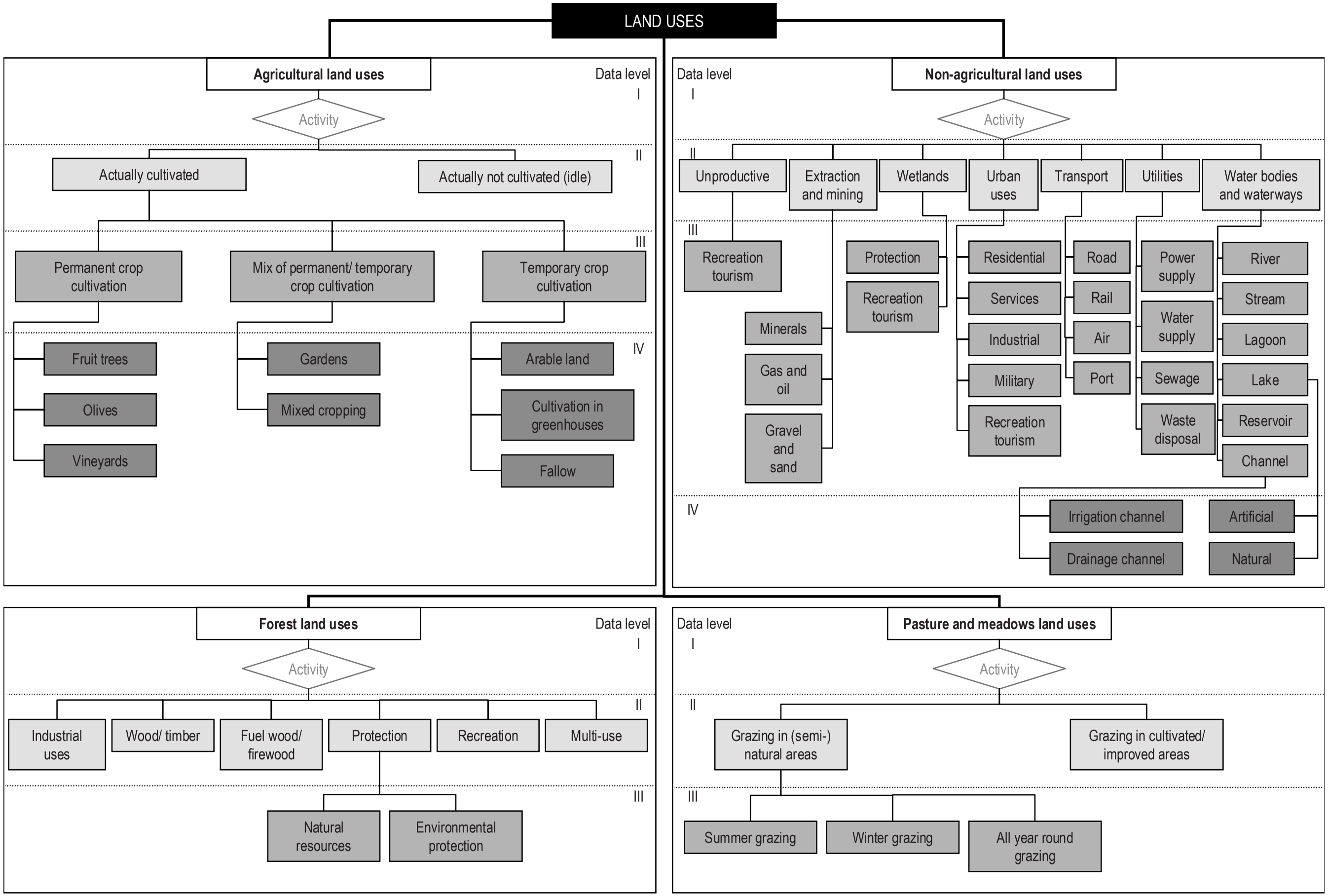
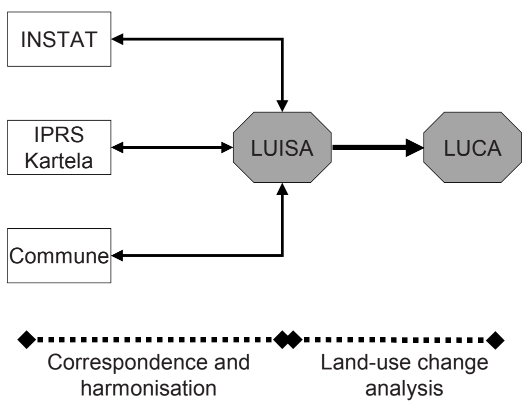


Table 4-4. Grouping of the land-use changes according to LUCA (Jansen 2003; Jansen *et al.* 2007)

Type of land-use change			Code
No change		Correspondence	1
Modifications	Low level	Low level modification in Agriculture	201
		Low level modification in Forests	202
		Low level modification in Pastures	203
		Low level modification in Non-Agriculture	204
	Medium level	Medium level modification in Agriculture	301
		Medium level modification in Forests	302
		Medium level modification in Pastures	303
		Medium level modification in Non-Agriculture	304
	High level	High level modification in Agriculture	401
		High level modification in Forests	402
		High level modification in Pastures	403
		High level modification in Non-Agriculture	404
Conversions		Agriculture-to-Forest	5
		Agriculture-to-Pasture	6
		Agriculture-to-Non Agriculture	7
		Forest-to-Pasture	8
		Forest-to- Agriculture	9
		Forest-to-Non Agriculture	10
		Pasture-to-Agriculture	11
		Pasture-to-Forest	12
		Pasture-to-Non Agriculture	13
		Non Agriculture-to-Agriculture	14
		Non Agriculture-to-Forest	15
		Non Agriculture-to-Pasture	16
Unknown		No correspondence (land-use change is unlikely to occur)	99

Figure 4-3. Harmonisation of class sets in Albania using a reference system (LUISA) and harmonising land-use change (LUCA)



according to the land-use change processes modification and conversion as shown in Table 4-4 (Jansen 2003; Jansen *et al.* 2007).

The harmonisation process, between the different class sets and for harmonisation of land-use change using LUISA as reference system, is shown in Figure 4-3.

4.7.2 Results of correspondence between the class sets

Correspondence between the classes of the four systems is important when using existing data sets coming from different sources at different levels of detail and trying to integrate and harmonise them in a geo-database. A table of correspondence has been prepared (Table 4-5) that shows that correspondence is often of the type one-to-many or many-to-many, especially when classes used at national level (e.g., INSTAT) are correlated with classes used at more detailed levels (e.g., IPRS Kartela, Commune and LUISA). However, if one looks at the more detailed level of the cadastral parcel unit of the IPRS Kartela and LUISA class sets, the many-to-many relationships occur less frequently and one gets a better idea about the correlation of single classes (Table 4-6 and Table 4-7).

Some classes of IPRS Kartela do not correspond with a class of LUISA because they either occur below ground (e.g., 130, 332) and have not been included, or do not address a land use (e.g., 344). Other classes are of a more generic nature than the detail of the classes used in LUISA resulting in a one-to-many relationship (e.g., 108, 110 and 118) or vice versa (e.g., LUISA classes 91, 92 and 93 with various IPRS Kartela classes). Other classes are more closely related to a land *cover* than a land *use* (e.g., 118, 119, 135 and 336) and the relation with land use is not always apparent.

LUISA classes 95, 113, 114, 122, 124 and 133 do not correspond with any IPRS Kartela class. More detail has been introduced in the description of Agricultural and Non-agricultural land uses. Suitable agricultural lands are limited in Albania and it is regarded as important to know why they might not be utilised for production of agricultural goods and/or services in the current agricultural year or for longer periods. The LUISA classes distinguished in the Forests and Pastures and Meadows categories have been introduced to better distinguish their range of uses instead of focussing mainly on their different land-cover type.

Table 4-5. Correspondence between land-use classes from different class sets at a generic level (Jansen 2003)

Legal categories	Land-use classes	Class sets INSTAT [23]	IPRS Kartela [24]	Commune [25]	LUISA [26]
Agriculture	Used agricultural area	b	-	1	1, 2, 3, 4, 5, 6, 7
	Area with arable land crops	c, d	101, 102	1a	6, 7
	Area with permanent crops	f	116, 125, 128, 131, 148	1b, 1c, 1d	1, 2, 3, 4, 5
	Non-utilised agricultural area	e	-	-	8, 9
Pastures and meadows	Grassland and pastures	g	108, 110, 153	2, 3a	51, 52, 53, 54, 55
Forests	Forests	h	118	3	31, 32, 33, 34, 35, 36, 37
Non-agricultural lands	Water bodies	-	107, 109, 111, 120, 138, 153	4a	131, 132, 133, 134, 135, 136, 137, 138
	Wetlands	-	336	-	81, 82
	Built-up areas	-	100, 103, 106, 114, 121, 129, 130, 136, 144, 152, 213, 261, 332, 337, 338, 339, 340, 341, 342	4b, 4c, 4d, 4e	91, 92, 93, 94, 95, 111, 112, 113, 114, 121, 122, 123, 124
	Barren	-	119, 135	4f	61
	Mining/extraction	-	117, 343	-	71, 72, 73

[23] a=Total Area (not represented in the table), b=Used agricultural area (UAA), c=Cultivated area with arable land crops, d=Main crops (the first ones), e=Non-utilised agricultural area, f=Area with permanent crops, g=Grassland and pasture, h=Forests.

[24] For explanation of the codes see Table 4-6. Classes 130, 332 and 344 not included.

[25] 1=Agriculture, 1a=Arable, 1b=Vineyards, 1c=Fruit trees, 1d=Olives, 2=Pastures, 3=Forest, 3a=Brush land, 4=Non-agricultural, 4a=Water body, 4b=Built-up, 4c=Cemetery, 4d=Roads, 4e=Railway, 4f=Barren.

[26] For explanation of the codes, see Table 4-7.

Table 4-6. Correspondence between land-use classes at the level of the cadastral parcel unit (Jansen 2003)

IPRS Kartela land-use classes [27] Code	Class names	LUISA Class codes [28]
100	Apartment	91
101	Arable	6, 7, 8, 9
102	Arable + garden	6, 7, 8, 9
103	Water treatment facility	123
106	Building non-residential	92, 93, 94
107	Channel	137, 138
108	Pasture	51, 52, 53, 54, 55
109	Lake	134, 135
110	Meadows	51, 52, 53, 54, 55
111	River	131
114	Block of flats	91
116	Fruit trees	1
117	Oil well	72
118	Forest	31, 32, 33, 34, 35, 36, 37
119	Barren	61
120	Reservoir	136
121	Road	111
125	Garden (of private building)	4
128	Olives	2
129	Cemetery	92
130	Tunnel, underground	-
131	Vineyards	3
135	Rocky	61
136	Public area	92
138	Stream	132
144	Transformer building (step-up or step-down)	121
148	Fruit trees + garden	5
152	Railroad	112
153	Barrier (natural or artificial)	51, 52, 53, 54, 137, 138
213	Building for residential purpose	91
261	Sport field	92
332	Underground	-
336	Marsh	81, 82
337	Sidewalk	111
338	Unit (consisting of small shop or bar)	92
339	Garage	91
340	Studio	91
341	Power plant	121
342	Area associated to power plant	121
343	Mine area	71, 73
344	Transport equipment	-

[27] The IPRS Kartela classes do not have a hierarchical data structure, their structure is flat.

[28] For the explanation of the codes see Table 4-7.

Table 4-7. Correspondence between the land-use classes of LUISA and IPRS Kartela[29] (Jansen 2003)

Category	LUISA Code	Description	IPRS Kartela Code
Agricultural Land uses	1	Fruit trees	116
	2	Olives	128
	3	Vineyards	131
	4	Gardens	125
	5	Mixed cropping	148
	6	Arable lands	101, 102
	7	Cultivation in greenhouse	101, 102
	8	Fallow lands	101, 102
	9	Actually not cultivated (idle and abandoned) lands	101, 102
Forests	31	Industrial forests uses	118
	32	Forests for wood/timber production	118
	33	Forests for fuel wood/firewood	118
	34	Protection of natural resources	118
	35	Forests for environmental protection	118
	36	Forests for recreation	118
	37	Multi-use forests	118
Pastures and Meadows	51	Grazing in (semi-) natural areas	108, 110, 153
	52	Summer grazing in (semi-) natural areas	108, 110, 153
	53	Winter grazing in (semi-) natural areas	108, 110, 153
	54	All-year-round grazing in (semi-) natural areas	108, 110, 153
	55	Grazing in cultivated/improved areas	108, 110
Non-agricultural land uses	61	Recreation/tourism in unproductive areas	119, 135
	71	Mineral extraction and mining	343
	72	Gas and oil extraction	117
	73	Gravel and sand extraction/mining	343
	81	Protection of wetlands	336
	82	Recreation/tourism in wetlands	336
	91	Residential area	100, 114, 213, 339, 340
	92	Services	106, 129, 136, 261, 338
	93	Industrial area	106
	94	Military area	106
	95	Recreation/tourism in urban areas	-
	111	Road	121, 337
	112	Railroad	152
	113	Airport	-
	114	Port	-
	121	Power supply	144, 341, 342
	122	Water supply	-
	123	Sewage	103
	124	Waste disposal	-
	131	River	111
	132	Stream	138
	133	Lagoon	-
	134	Natural lake	109
	135	Artificial lake	109
	136	Water reservoir	120
	137	Irrigation channel	107, 153
	138	Drainage channel	107, 153

[29] See Figure 4-2 for the LUISA hierarchical structure; see Table 4-6 for the IPRS Kartela codes.

4.7.3 Comparison with the ANFI remotely sensed land-cover/use data set

The World Bank financed Albanian National Forest Inventory (ANFI) project provided an analysis of spatially explicit land-cover/use change dynamics in the period 1991-2001 using the Land-Cover Classification System for codification of classes, satellite remote sensing and field survey for data collection and elements of the object-oriented geo-database approach to handle changes as an evolution of land-cover/use objects, i.e. polygons, over time to facilitate change dynamics analysis (Jansen *et al.* 2006a). Land-cover/use changes are the results of many interacting processes and each of these operates over a range of scales in space and time (Verburg *et al.* 2003). The detailed LUISA land-use data can be compared to the coarser ANFI data (scale 1:2,500 and 1:100,000 respectively), as far as space and time considerations both data sets represent more or less the same period (1991-2003 and 1991-2001 respectively), but the analysis of each data set gives a somewhat different view on the change dynamics at detailed versus aggregated data levels. At aggregated data levels the local variability of spatially explicit land-cover/use changes may be obscured, whereas patterns can be shown that at more detailed data levels may remain invisible and vice versa (Veldkamp *et al.* 2001b).

The LUISA data set permits analysis of changes at the level of the individual cadastral parcel unit, thereby highlighting changes at the level of the landowner and/or land user. The ANFI data set provides a national overview of the major change processes, such as deforestation, urbanisation and increased pasture, but cannot provide conclusive evidence on especially the use of agricultural land (Jansen *et al.* 2006a). The LUISA data set provides an insight into the non-use of low productivity areas in hilly terrain and the extensive forms of agriculture practised on prime agricultural land because of the lack of fertilizer use and the breakdown of irrigation systems (Jansen *et al.* 2007). These two spatially explicit data sets are therefore complementary when analysing change dynamics.

It is important to note that the use of remote sensing for land cover is a common approach. Interpretation of satellite images can provide a quick overview of the type and location of different land-cover types. Often land-use elements are inferred from land cover (e.g., detection of a field pattern results in the class ‘agriculture’). However, the above example clearly demonstrates that land use requires a different approach because it contains many aspects that go beyond land cover. Even with the use of the most detailed satellite images, such aspects will not be captured.

4.7.4 Correspondence with an international class set

Land-use harmonisation should also make reference to an internationally established class set used to describe national level data. Such a reference was not immediately related to the work of the LUP II project, but the value of the project outputs will be enhanced if correspondence to especially EU wide operational systems is assured. This will facilitate accession of Albania into the EU and continuity in data collection routines.

Land use is of high importance in the definition and evaluation of common sectoral policies in the EU, e.g. on environment, agriculture, transport and the integration of those policies in a comprehensive assessment and planning of the territory. EUROSTAT, the Statistical Office of the European Communities, has the mission to provide the EU with high quality statistical information services. To support policy formulation, EUROSTAT launched in co-operation with the Directorate General for Agriculture in 2000 the Land-Use/Cover Area Frame Statistical Survey (LUCAS) project that has been applied in the period 2001-2005 and will be applied in a revised form in 2006 in 23 EU Member States.

Overall objectives of this survey are (EUROSTAT 2001):

1. Collection of harmonised data (i.e. unbiased estimates) at EU level of the main land-use and land-cover areas and changes.
2. Inclusion not only of the usual agricultural domain but also the aspects linked with environment, multi-functionality, landscape and sustainable development.
3. A common sampling base (e.g., sampling frame, class set and data management) that interested Member States can use to obtain representative data at national, but also regional, level by increase of the sampling rate while respecting the general LUCAS approach.
4. Evaluation of the strengths and weaknesses of a point area frame survey as one of the pillars of the future Agriculture Statistical System (area frame means that the observation units are territorial subdivisions instead of agricultural holdings as used in the Farm Structure Survey).

Table 4-8. LUCAS version 1.0 (EUROSTAT 2001)

Land-Use/Cover Area Frame Statistical Survey Code	Land-use category name
U11	Agriculture
U12	Forestry
U13	Fishing
U14	Mining – Quarrying
U21	Energy Production
U22	Industry – Manufacturing
U31	Transport, Communication, Storage, Protective Works
U32	Water, Waste Treatment
U33	Construction
U34	Commerce, Finance, Business
U35	Community services
U36	Recreation, Leisure, Sport
U37	Residential
U40	Unused

Table 4-9. Correspondence between the land-use classes of LUISA and LUCAS 1.0 (Jansen 2003)

Category	LUISA [30] Code	Land-use category code	LUCAS 1.0 Code
Agricultural land uses	1	Fruit trees	U11
	2	Olives	U11
	3	Vineyards	U11
	4	Gardens	U11
	5	Mixed cropping	U11
	6	Arable lands	U11
	7	Cultivation in greenhouse	U11
	8	Fallow lands	U11
	9	Actually not cultivated (idle and abandoned) lands	U11
Forests	31	Industrial forests uses	U12
	32	Forests for wood/timber production	U12
	33	Forests for fuel wood/firewood	U12
	34	Protection of natural resources	U12
	35	Forests for environmental protection	U12
	36	Forests for recreation	U12
	37	Multi-use forests	U12
Pastures and Meadows	51	Grazing in (semi-) natural areas	U11
	52	Summer grazing in (semi-) natural areas	U11
	53	Winter grazing in (semi-) natural areas	U11
	54	All-year-round grazing in (semi-) natural areas	U11
	55	Grazing in cultivated/improved areas	U11
Non-agricultural land uses	61	Recreation/tourism in unproductive areas	U36
	71	Mineral extraction and mining	U14
	72	Gas and oil extraction	U14
	73	Gravel and sand extraction/mining	U14
	81	Protection of wetlands	?
	82	Recreation/tourism in wetlands	U36
	91	Residential area	U37
	92	Services	U34, U35, U36
	93	Industrial area	U22
	94	Military area	U35?
	95	Recreation/tourism in urban areas	U36
	111	Road	U31
	112	Railroad	U31
	113	Airport	U31
	114	Port	U31
	121	Power supply	U21
	122	Water supply	U32
	123	Sewage	U32
	124	Waste disposal	U32
	131	River	U13, U32
	132	Stream	U13, U32
	133	Lagoon	U13, U32
	134	Natural lake	U13, U32
	135	Artificial lake	U13, U32
	136	Water reservoir	U13, U32
	137	Irrigation channel	U32
	138	Drainage channel	U32

[30] See Figure 4-2 for the hierarchical data structure of LUISA.

The main LUCAS land-use categories (version 1.0) are shown in Table 4-8. Correspondence with LUISA is shown, although based upon a different basic unit of measurement, in Table 4-9 at the individual class level. The LUISA class ‘Protection of wetlands’ does not find a corresponding class in LUCAS, whereas the correspondence of the LUISA class ‘Military area’ with LUCAS U35 may raise some questions. There are few one-to-many correspondences that are mainly concerning land uses related to water and services (Jansen 2003).

4.8 Discussion and conclusions

Land use has been defined and interpreted in many different ways depending on the sector. The multi-disciplinary nature of the subject has hampered the development of a standardised methodology for categorisation as well as harmonisation of land uses worldwide. Existing class sets have been reviewed in order to distil the key elements but there is a genuine lack of consistency in applied methodology and adherence to the principles of categorisation, and a variety of basic units of measurement are used. Evaluation of the main parameters used in existing class sets leads to the conclusion that the combination of just two parameters may suffice: ‘function’ together with ‘activity’. ‘Function’ is centred on the purpose of land uses, whereas ‘activity’ groups all land undergoing a certain process resulting in a certain type of product.

The example in Albania shows how the use of a reference system, based upon the ‘function’ and ‘activity’ parameters and using the cadastral parcel as basic unit of measurement, may facilitate harmonisation of class sets in parallel with achieving harmonisation of land-use change. This reference system can form the basis for future standardisation of land-use class sets in Albania. In addition, the use of synergies between categorisation and information technology concepts (e.g., data model and resulting geo-database structure) should be enhanced.

Comparison of the cadastral-parcel-based class set of Albania with a polygon-based class set at coarser resolution shows that different levels of detail are needed when analysing land-use change. Remote sensing is a useful tool for gaining a quick overview of land-cover related land uses but the potential for a detailed and in-depth knowledge of land use is limited as other aspects, such as socio-economics, institutional, cultural and legal factors, are not captured by remotely sensed based land cover. Therefore, remote sensing can make a valuable contribution but its limits should be clear and complementary approaches should be used. Understanding land-use change dynamics does not only help to identify vulnerable places, but also vulnerable (groups of) land users (or landowners when working with cadastral data) that on their own are incapable to respond in the face of environmental processes and problems.

The way forward for harmonisation of land-use class sets is to promote and fully develop a parameterised approach to categorisation. Commonalities in existing approaches should be emphasized and a set of commonly used parameters should be identified. Lessons can be learnt from harmonisation attempts at local, regional and national levels that are equally valid for a globally applicable land-use categorisation. The successor of the IGBP/IHDP LUCC, the Global Land Project (GLP 2005) could provide the necessary platform. Furthermore, a quantitative measure should be defined to express the harmonisation result of a class and between class sets as it is not only important to document data quality but also harmonisation quality.

PART III - LAND-CHANGE ANALYSIS

We have at any rate one advantage over Time and Space.
We think *them* whereas it is extremely doubtful whether *they* think us!

John Cowper Powys (1872-1963)

Abstract

As scale is an important issue in change analysis, three different scales are considered in chapters 5 and 6: the national and district levels for land-cover change (Chapter 5) and the commune level for land-use change (Chapter 6). The land-cover and land-use change analyses are complementary in the understanding of land-change processes and patterns.

Chapter 5. Change analysis with parameterised land-cover class sets at national and district levels

In the turmoil of a rapidly changing economy the Albanian Government needs accurate and timely information for management of their natural resources and formulation of land-use policies. The transformation of the forestry sector has required major changes in the legal, regulatory and management framework. The World Bank financed Albanian National Forest Inventory project provides an analysis of spatially explicit land-cover/use change dynamics in the period 1991-2001 using the Land-Cover Classification System for codification of classes, satellite remote sensing and field survey for data collection and elements of the object-oriented geo-database approach to handle changes as an evolution of land-cover/use objects, i.e. polygons, over time to facilitate change dynamics analysis.
Analysis results at national level show the trend of natural resources depletion in the form of modifications and conversions that lead to a gradual shift from land-cover/use types with a tree cover to less dense tree covers or even a complete removal of trees. Policy failure (e.g., corruption, lack of law enforcement) is seen as the underlying cause. Another major trend is urbanisation of areas near large urban centres that change urban-rural linkages. Furthermore, after privatisation agricultural areas increased in the hills where environmental effects may be detrimental, while prime agricultural land in the plains is lost to urbanisation.

Based on: Jansen, L.J.M., Carrai, G., Morandini, L., Cerutti, P.O., Spisni, A., 2006. Analysis of the spatio-temporal and semantic aspects of land-cover/use change dynamics 1991-2001 in Albania at national and district levels. *Environmental Monitoring & Assessment* 119: 107-136.

Based on: Jansen, L.J.M., Carrai, G., Petri, M., 2007. Land-use change dynamics at cadastral parcel level in Albania: an object-oriented geo-database approach to analyse spatial developments in a period of transition (1991-2003). In: Koomen, E., Bakema, A., Stillwell, J., Scholten, H. (Eds.). Modelling land-use change – progress and applications. GeoJournal Library Vol. 90. Springer Publishers, Berlin Heidelberg. Pp. 25-44.

At district level, the local variability of spatially explicit land-cover/use changes shows different types of natural resources depletion. The distribution of changes indicates a regional prevalence, thus a decentralised approach to the natural resources management could be advocated.

Chapter 6. Change analysis with parameterised land-use data sets at commune level

A case study in Albania is presented based on the EU PHARE Land-Use Policy II project results where GIS-oriented instruments and innovative methodologies were implemented to support decision-making for land-use policy and planning. The developed Land-Use Information System for Albania allows the logical and functional hierarchical arrangement of land uses and data harmonisation with other land-use description systems. It is linked to the object-oriented Land-Use Change Analyses methodology that groups changes into conversions and modifications. The preferred change patterns indicate that land users take rational decisions when changing land use, even in the absence of any regulating plan, as is the case in post-communist Albania.

5 Change analysis with parameterised land-cover class sets at national and district levels

5.1 Introduction

In the early 1990s, Albania entered a period of transition from a central-based planned economy to a market economy. Early efforts to introduce democracy and build a market economy were severely undermined by the socio-economic crisis and generalised unrest that followed the financial collapse of 1997. The lack of a democratic culture, the absence of dialogue between different political tendencies and a limited understanding of the concept of national interest amongst political leaders have often prevented the development and implementation of sound policies to address the many issues that Albania faces (CEC 2002).

Despite disruptions in production caused by energy shortages (e.g., even in the capital Tirana there is no 24 hours regular power supply), real Gross Domestic Product (GDP) growth in 2001 reached the target of 7.3%. However, GDP per capita remains one of the lowest in Europe (around 1,400 €) and the overall impact of economic growth remains limited on the poorest layers of the population (CEC 2002). Around a quarter of the population is considered to be living below the poverty line (World Fact Book 2005). Construction and services mainly contributed to the GDP increase, with growth rates of 17% and 12% respectively. Industrial production grew around 6% but this sector is weak and its contribution to the overall GDP growth limited. Industries are often obsolete, non-viable and incapable of competing with European industry. Efforts of Government to improve the poor national road and railway networks, a long-standing barrier to sustained economic growth, are very slow. Agriculture, which still accounts for slightly more than 50% of Albania's GDP, grew by 3.5%. The privatisation process led to the break up of the former 550 collective farms, which catered for the state processing and marketing agencies, into 467,000 smallholder farms that operate very often at little more than subsistence level. Although for their type quite productive, they are not price competitive and about 75% of farm production is home consumed. These growth figures are not fully reliable, since official figures provide inadequate coverage of private sector activity (CEC 2002).

In Albania, the post-collectivisation ownership status was identical for collective and state farmlands due to the nationalisation of all land after the Second World War. Therefore the Government could apply the same land reform procedures to collective and state farmlands: (physical) distribution of collective (76% of Total Agricultural Lands (TAL)) and state farmlands (24% of TAL) (Swinnen 1999). The pre-collectivisation landownership distribution was highly unequal (3% of the population owned the land) resulting in historical justice and social equity being conflicting objectives. Being the poorest and most rural economy of the Central and Eastern European Countries, Albania has 50% of its working population employed in agriculture and agriculture has a prominent economic role. Government decided to redistribute the land to the rural households on an equal per capita basis with partial financial compensation for former owners. This choice is consistent with equity considerations in choosing a land reform procedure. The distribution of even a small piece of land to farm workers has had an important effect on their income and food security situation. Land distribution was also a preferable choice from the efficiency point of view: low technology agriculture, labour-intensive farming structures and imperfect (or missing) capital markets (Swinnen 2000).

The change to a market-oriented economy had also an impact on the natural resources and their management, not only due to privatisation, but also because of the strong land fragmentation as a result of the land distribution and increased urbanisation. For the first time in 50 years people were free to move around. The rural population, particularly in mountainous areas, sharply decreased because of urban drift or migration abroad. The increasing pastoral economy and husbandry caused landscape degradation and natural resources depletion in many regions of the country. Uncontrolled timber harvesting, overgrazing and overexploitation of wood (in a country with a permanent energy shortage) and other forest products have changed environmental assets. The depletion of forest resources, particularly in accessible areas, has become alarming. Scarce possibilities of control and a lenient policy caused severe, sometimes even irremediable, damages to the natural resources of Albania.

The agrarian reform in its first phase led to a fast increase in the construction of (illegal) buildings and new roads. In a subsequent phase many new small farms were abandoned followed by rapid urbanisation as more and more people left the rural areas to become resident in urban centres. These urban centres, however, were not prepared to receive the massive influx of people. In the turmoil of such a changing economy and the spatial and temporal dynamics of land cover/use that are continuously evolving, it is important for the Albanian Government to have accurate and timely information for natural resources management, land-use planning and policy development, as a prerequisite for monitoring and modelling land use and land change and as a basis for land-use statistics. Land-cover/use change, as one of the main driving forces of (global) environmental change, is central to sustainable development (Meyer and Turner 1994; Walker and Steffen 1997; Walker 1998; Lambin *et al.* 2000). In spite of the many achievements in institutional and policy reform, reliable estimates are missing and great uncertainty exists on the actual, real economic potentialities of the natural resources. The quality and quantity of resources at various points in time, the rates by which they have changed, the overall distribution of the land-cover/use types, etc., are not precisely known. Therefore, there are many uncertainties about the strategy to be adopted by Government in order to plan a sustainable use of natural resources while preserving biological richness and diversity.

The objective of the study presented in this part of the chapter was an inventory of the land-cover/use types in Albania, their location, extent and distribution and an understanding of the change dynamics in the period 1991-2001 at both national and district levels in order to provide to Government spatially explicit data and information for a sustainable management of natural resources. While natural resources management policies are formulated at the national level by different ministries, they are mostly executed at district or even commune level. The responsibility for forests and pastures is with the Directorate General of Forests and Pastures (DGFP) of the Ministry of Agriculture and Food (MoAF). The DGFP has District Forest Service Directories (DFSD) in the 36 districts. The Albanian Forestry Project, of which the Albanian National Forest Inventory (ANFI) project that executed the study is a sub-component, is carried out under agreement between the World Bank and MoAF and will transfer the responsibility of about 40% of the forest area directly to the communes. This transfer process is based on the Law 'On Forests and Forest Service Police' (No. 7623 dated 13 October 1992) and Regulation 'On the Transfer of Forests and Pastures in Use to Communes (No. 308, dated January 1996). According to these legal acts 'the communal forests and pastures would be given to users who are permanent inhabitants of the Commune. The agreement –signed contract between the Commune and the users- gives the latter the full rights to all benefits from communal forests and pastures transferred to the Commune' (SDC/FAO/World Bank/Ministero degli Affari Esteri 2003).

5.2 Materials and methods

5.2.1 Remote sensing materials used

Digital LANDSAT 7 Enhanced Thematic Mapper (ETM+) imagery has been used to produce the baseline interpretation of 2001 using on-screen digitising and visual interpretation. For the 1991 visual interpretation use has been made of LANDSAT 5 Thematic Mapper (TM) images (Table 5-1). For interpretation purposes the multiple view approach was selected combining multi-stage sensing (i.e. high-resolution satellite data is analysed in combination with low altitude data such as topographic maps, forest type maps and field survey data), multi-spectral sensing (i.e. data are acquired simultaneously in several spectral bands) and multi-temporal sensing (i.e. data about the terrain is collected at different dates). The 2001 images have been geo-referenced using the topographic maps of the Albanian Military Geographic Institute at scale 1:100,000 (image-to-map approach) and the 1991 images have been geo-referenced according to the geo-referenced 2001 October set (image-to-image approach).

Table 5-1. LANDSAT 7 ETM+ and 5 TM frames used for interpretation

LANDSAT Path-Row	Acquisition Date LANDSAT 7 ETM+		LANDSAT 5 TM	
185-032	25 October 2000	19 June 2000	23 September 1991 9 September 1992	3 June 1991
186-031	3 October 2001	9 May 2000	30 September 1991	9 May 1991
186-032	3 October 2001	28 May 2001	14 September 1991	9 May 1991

A Brovey fusion procedure was applied to the False Colour Composite (FCC) 453 multi-spectral imagery at 30m resolution with panchromatic band at 15m resolution (Table 5-2). The result is imagery that is characterised by the pixel resolution of 15m of the panchromatic band and the spectral resolution of the multi-spectral bands of the FCC. The procedure enhances the visual quality of the imagery and consequently facilitates detection of different vegetation types. In addition, band 3 has been filtered with an edge-sharpening filter, kernel 3x3, to reduce fuzzy noise. The same procedure has been applied to the FCC 432 (Jansen *et al*. 2003a).

Table 5-2. The Brovey fusion formula

R=4, G=5, B=3, I=panchromatic band		
Red layer	Green layer	Blue layer
[R/(R+G+B)]*I	[G/(R+G+B)]*I	[B/(R+G+B)]*I

In the interpretation process various levels of complexity exist, from simple direct recognition of objects in the scene to inference of site conditions. The interpreters use the process of convergence of evidence to successfully increase the accuracy and detail of the interpretations. During the interpretation process special attention was paid to: (1) the spatial coherence of polygons, i.e. are the boundaries in the appropriate place and have the same logical and functional thinking been applied in a consistent manner in the area of interpretation; and (2) the thematic coherence, i.e. is the label given to the polygons correctly describing their contents and are other areas with similar features described in the same manner. A continuous crosschecking of the 1991 and 2001 interpretations was necessary in order to guarantee spatial and thematic coherence within the interpretations and between them. The 2001 interpretation has been validated using 431 field observations and 111 additional observations. The overall thematic accuracy of the 2001 interpretation at the level of LCCS domains and land-cover groups, discussed in the next paragraph, is 85% (Jansen *et al*. 2003a).

5.2.2 Land-cover/use categorisation applied

The 1991 and 2001 land-cover/use interpretations apply the Land-Cover Classification System (LCCS), endorsed by the Land-Use and Land-Cover Change (LUCC) program element of the International Biosphere-Geosphere

Programme (IGBP) and International Human Dimensions Programme on Global Environmental Change (IHDP), for definition of classes to ensure harmonisation of the data with existing data sets at international level while at the same time standardising the method used for description of land-cover/use features (Di Gregorio and Jansen 2000; LANES 1998; McConnell and Moran 2001; FAO 2002; Jansen and Di Gregorio 2002). The defined classes used in both interpretations are shown in Table 5-3 with their groupings at different levels of aggregation (Jansen *et al*. 2003a and 2003b).

Land-cover/use change has to recognize that changes come in two types: (1) conversion from one land-cover category to another (e.g., from cultivated to built-up area); and (2) Modification within one category (e.g., from forest to woodland, from thicket to shrubland, etc.). These two types of change have implications for the methodology used to describe and classify land cover/use (Jansen and Di Gregorio 2002). Conversion implies an evident change, whereas modifications are much less apparent. The latter requires a greater level of detail to be accommodated. With a system based upon class names the latter type of change cannot be captured unless the system contains an ample set of classes.

In the past the emphasis of change studies has been on conversions, whereas more recently there has been increased recognition of the processes of modification (Lambin *et al.* 2003).

The logical ordering of classes in the LCCS Legend Module facilitates the analysis phase because classes are grouped according to major land-cover category, followed by occurring land-cover domains. A matrix with these groupings of classes filled with change dynamics statistics facilitates the interpretation of identified changes. Three different areas can be identified in the matrix (Table 5-4): (1) the areas where no land-cover/use change occurred; (2) the areas where modifications within or between domains occurred; and (3) various types of conversion (e.g., reforestation or deforestation). The same matrix can be used for the interpretation of the likely causes of land-cover/use change such as deforestation, forest fragmentation, afforestation, reforestation, etc.

The analysis of the spatial extent of the different types of land-cover/use changes will permit a further inside in the prevailing land-cover/use change trends. The spatial and temporal land-cover changes should be linked to socio-economic developments in order to understand land-use changes. Land-use characterises the human use of the land-cover type. For example, forests can be used for selective logging, for recreation, or not at all.

Table 5-3. The land-cover/use classes with overall accuracy of 85% at land-cover group level

LCCS Category	LCCS LC domain	LC Group	User description	User label
Dichotomous Phase	Aquatic Vegetation (AV)		Aquatic vegetation	0AQ
Cultivated areas and managed lands (A11)	Tree and Shrub Crops (TC)		Broadleaved arboriculture mainly for wood/timber production. Plantation with *Populus* spp. and *Juglans* spp.	1AB
			Coniferous arboriculture mainly for wood/timber production. Plantation with *Pinus* spp.	1AC
			Fruit trees or shrub crops; when the crop is young and still unproductive another crop may be intercropped on the same field.	1FR
	Herbaceous Crops (HC)		Cultivated areas with herbaceous crops on level land; when not in actual use the fallow field area may be used for pasture.	1CU
			Cultivated areas with herbaceous crops on sloping land; when not in actual use the fallow field area may be used for pasture.	1CS
	Managed Lands (ML)		Vegetated Urban Area(s)/Park(s) (green areas inside built-up areas).	1UP
(Semi-) natural vegetation (A12 and A24)	Forests (FO)	Broadleaved (FOB)	Broadleaved evergreen forest	2BE
			Broadleaved deciduous forest. *Quercus* spp. and/or *Ostrya* spp. are dominant, usually coppice	2BD
			Creek and riverine broadleaved deciduous forest	2CR
			Broadleaved deciduous forest with dominant *Fagus silvatica*. 'Mixed' means a mixture of broadleaved species.	2FS
		Coniferous (FOC)	Coniferous forest (Med.) on level land	2CA
			Coniferous forest (Sub-Med.) on sloping land	2CB
			Coniferous forest (Alpine) on steep land	2CC
			Coniferous forest (Med.) on level land / Beaches (The code for this class is: 2CA/6BC).	COB
		Mixed (FOM)	*Fagus silvatica* pure and mixed with coniferous forest (<30%)	2FC
	Forests or Woodlands		Broadleaved deciduous forest. *Quercus* spp. and/or *Ostrya* spp. dominant, usually coppice / Cultivated areas with herbaceous crops on sloping land (The codes for this mixed class are: 2BD/1CS or 2BD/1FR or 2DO/1CS or 2DO/1FR; rare to have 1CU inside).	CXB
	Woodlands (WL)	Broadleaved (WLB)	Broadleaved evergreen woodland	2EO
			Broadleaved deciduous open forest. *Quercus* spp. are dominant (coppice)	2DO
			Broadleaved deciduous open forest with *Fagus silvatica* dominating. 'Mixed' means a mixture of broadleaved species.	2FB
		Coniferous (WLC)	Coniferous open forest (Med.) on level land	2OA
			Coniferous open forest (Sub-Med.) on sloping land	2OB
			Coniferous open forest (Alpine) on steep land	2OC
		Mixed (WLM)	*Fagus silvatica* pure and mixed with coniferous open forest (<30%)	2FO
	Thickets and Shrublands (TS)		Mediterranean macchia	2MM
			Mediterranean macchia/Broadleaved deciduous forest. *Quercus* spp. and/or *Ostrya* spp. are dominant, usually coppice. (The codes for this mixed class are: 2MM/2BD or 2MM/2DO or 2MG/2BD or 2MG/2DO).	MXB
			Maquis and garigue (incl. low Med. macchia)	2MG
			Maquis and garigue (incl. low Med. macchia) / Cultivated areas with herbaceous crops on sloping land (The codes for this mixed class are: 2MG/1CS or 2MG/1FR or 2MM/1CS or 2MM/1FR; rare to find 1CU inside).	CXM
	Grasslands/Pastures		Sparse trees and shrubs with open to closed grass cover and rock outcrops; these areas are used as pastures.	2SR

LCCS Category	LCCS LC domain	LC Group	User description	User label
	(GL)		Sparse trees and shrubs with very open grass cover and rock outcrops; these areas are used as pastures.	2GR
Artificial Surfaces (B15)	Built Up Areas (BU)		Built-up areas. Urban and industrial areas (including road network)	5UR
Bare areas (B16)	Bare Areas (BA)		Beaches	6BC
			Bare rocks and/or soils (includes bare areas due to human activities such as mining or environmental degradation).	6RS
Water bodies (B27 and B28)	Water Bodies (WB)		Artificial perennial water bodies	7AA
			Natural perennial water bodies	8AN

Table 5-4. Change matrix

Class codes[31]	1a	1b	1c	2d	2e
1a	(1)	(2)	(2)	(3)	(3)
1b	(2)	(1)	(2)	(3)	(3)
1c	(2)	(2)	(1)	(3)	(3)
2d	(3)	(3)	(3)	(1)	(2)
2e	(3)	(3)	(3)	(2)	(1)

5.2.3 Applied change mapping procedure in the geo-database

In most change studies the state of land cover at a certain point in time is compared to the state at a later moment. This approach is basically focussing on a representation of temporal data in which snapshots (e.g., satellite image interpretations) are created for each moment in time. In what one could call a Geographic Information System (GIS) Overlay Approach this means that the land-cover state at t_1 is overlain with t_2, t_2 with t_3, t_3 with $t_{3+k,}$ etc. The result of such overlays (in raster format) is a sequential representation of the dispersion of a class in other classes. However, such representations usually do not contain the immediate link between the spatial and temporal dimensions of the changes that occurred between t_1 and t_{1+k}.

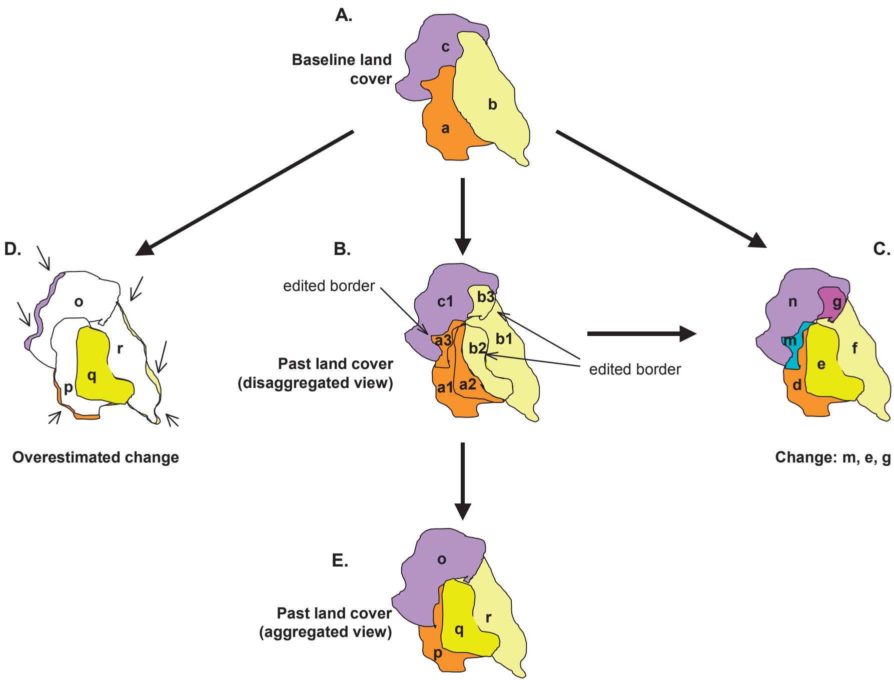
The approach to spatial and temporal analyses is more integrated when developing a different approach to databases that do not only store the state of land cover/use at different moments in time but also documents the relationships between such states (Figure 5-1). Thus, the database may not only contain relationships but also the processes that led up to these relationships. Versions are described in relation to a key state that can be located at a certain time, and through version identifiers one can store just changes instead of a complete version. The events and states are modelled as object classes with different roles; ‘events’ are used to describe what happened, is happening or will happen during the lifespan of an object, whereas ‘states’ certify what has changed, is changing or will change. Events can be modelled independently from these states; the object is depicted by different instances to different classes of events and states. Furthermore, if all this information is contained in a database one can reduce the amount of data to be stored and easily track the history of a polygon. This Object-Oriented Analysis (OOA) approach is closely linked to databases and data modelling.

Polygons are defined by a set of boundaries. Land-cover/use boundaries are linear objects demarcating different land-cover/use faces that experience a succession of changes in their positions during their lifespan. The history of each Land-Cover/use Boundary (LCB) is unique and shows the geographical, i.e. in the sense of relative position referenced to a baseline map, significance of

[31] The number indicates the land-cover category and the letter indicates the land-cover class.

a LCB over the development of environments, landscapes, anthropological activities, socio-economic aspects, etc. Each LCB has its own space-time path that represents its lifespan (Figure 5-2). A LCB begins to exist when it is for the first time mapped (creation) and has an existence period (existence) along which alterations can occur due to the evolution of the environment; it may happen that a LCB ceases to exist (demise) when it shares common characteristics on both of its sides (i.e. left area has the same semantic attribute of right area). The detection of change in LCB involves the description of the evolution of LCB at an earlier time (i.e. 1991) that accounts for the LCB being the way it is at a later point in time (i.e. 2001). The development of the boundary follows a longitudinal configuration (or a sequential one) without any branching; in other words, a boundary can only be alive and unique but it cannot become something else.

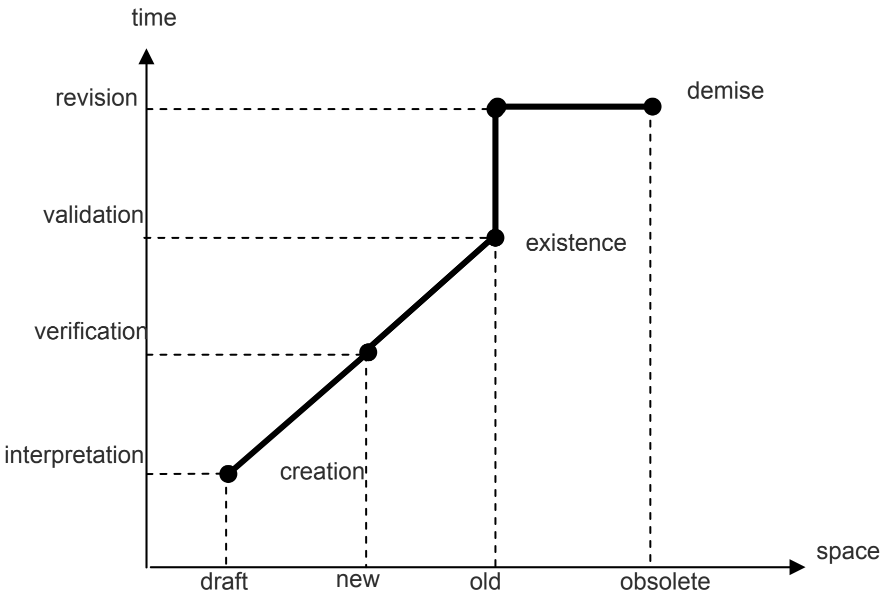
Figure 5-1. Physical implications of land-cover/use change mapping



A combination of the GIS Overlay approach and the OOA approach has been adopted at polygon level for the land-cover/use change analysis in order to be able to handle changes as an evolution of land-cover/use objects over time. The 1991 polygons are described by what has changed in their state, i.e. the spatial extent of the polygon formed by a set of land-cover/use boundaries and/or the polygon label (land-cover/use class) vis-à-vis their state in 2001. This allows quick identification of 'hotspots' of change. The GIS approach has been applied for a quick general overview, whereas the OOA approach has been used to understand the relationships between changes.

Past land cover/use (i.e. 1991) has been interpreted starting from the present validated land-cover/use layer (i.e. 2001). In practice, the interpretation of 2001 has been overlain on the 1991 satellite images. Where change was identified polygon boundaries and/or labels were updated in order to match the 1991 land cover/use and the change in state of the polygons recorded. This approach allowed minimising errors induced by creating a new layer. In fact, a certain amount of difference in physically drawing a new layer is to be accounted for (Figure 5-1). This would lead to an overestimation of change in terms of especially spatial extent.

Figure 5-2. A possible space-time path for a Land-Cover Boundary



5.3 Different aspects of land-cover/use changes

Land-cover/use changes are the results of many interacting processes and each of these operates over a range of scales in space and time (Verburg *et al.* 2003). Methods for detection and measuring of spatially explicit land-cover/use changes from remote sensing depend on comparisons between data sets acquired at known intervals of time, i.e. 1991 and 2001 in the case of this study. The accuracy of these data sets influences the analysis that can be made. Various sources of potential error exist such as spatial and temporal effects and the extent to which a given land-cover/use class may be recognized unambiguously from its radiometric properties. The key issue is that not only the right combination of data acquisition and data interpretation techniques must be selected, but also the right mixture of remote sensing with conventional techniques must be identified. Remote sensing is a tool and like any other tool its capacity to detect change is limited. This limit is related both to the accuracy with which the land cover/use will be identified on the image and consequently mapped at a certain point in time t_k and also to the rate and extent of change on the ground (Wyatt 2000).

Land-cover/use change analysis manifesting themselves at three different dimensions of scale can be analysed in terms of:

- Changes in geometry (area and perimeter), i.e. spatial aspects *x* and *y*;
- Rate of change, i.e. temporal aspect *t*; and
- Changes in class label, i.e. semantic aspect *s* that in a parameterised approach may range from a change in the composition of characteristics measured to a change in any of the measured characteristics. The semantic aspect relates to the constructed organisational hierarchy.

Spatial aspects influence the capacity to detect change in two ways. First, one should consider the spatial resolution of the images in relation to the scale of the changes to be observed: LANDSAT 5TM and 7ETM+ both at 30m resolution. Furthermore, the high degree of fragmentation of the landscape in Albania is important in the choice of imagery to detect changes. The occurrence of mixed pixels on different images of the same area may suggest change when there is no apparent change on the ground. Second, geo-referencing of images will cause errors however slight. This type of error is independent of the manner in which geo-referencing is undertaken, i.e. image-to-map (2001 images) or image-to-image (1991 images). Consequently, a proportion of apparent differences between images is due to mis-registration. The interpretations have an actual positional accuracy of not more than 34.5 meters on the ground as average with a standard deviation of 18.6 meters (Jansen *et al*. 2003a).

Temporal aspects should be considered when one tries to reconcile time and frequency of remote sensing data acquisition with the rate of change in the features of interest, i.e. the natural resources. Another problem to be considered is cloud incidence. Therefore the set of images used for a vast area usually covers different periods of time and it is difficult to establish a precise baseline against which to measure changes. Table 5-1 shows the used images for 1991 from September, whereas the 2001 images are all from October except one image that is from the year before.

The semantic aspects in a parameterised approach may range from a change in the composition of characteristics measured or a change in a single characteristic measured. For example, vegetation changes through the seasons or in the case of cultivated areas there exists an alternation between crops and the land lying bare but such changes are not considered to be of the type of a land-cover/use change. However, due to various reasons the canopy cover of a vegetation type can become less dense, so the characteristic of canopy cover changed. At the same time it is possible that due to the more open canopy cover the species composition changed.

A parameterised categorisation such as the LCCS facilitates land-cover/use change studies because the criteria used to define classes function at the same time as the parameters to be observed over time (Jansen and Di Gregorio 2002 and 2004b; Jansen 2004a). It therefore assists in determining the *s* aspect of change. Most existing classifications and legends in Europe (UN-ECE 1989; CEC 1995 and 1999) are based upon class names thereby not facilitating the use of these systems for monitoring purposes. It is more difficult to interpret a change in class names than comparing two sets of parameters.

5.4 Change dynamics in the period 1991-2001

5.4.1 Analysis of changes at national level

At aggregated data levels the local variability of spatially explicit land-cover/use changes may be obscured, whereas patterns can be showed that at more detailed data levels may remain invisible and vice versa (Veldkamp *et al.* 2001b). The change dynamics show that at national level an area of almost 330,000 hectares of the territory, i.e. 11.5%, is subject to land-cover/use change. The results at aggregated land-cover group or LCCS domain levels show immediately that the most significant changes occur in the vegetation type classes (Table 5-3). The area subject to land-cover/use change dynamics in the vegetation groups and domains comprises 91.2%, whereas 70.0% of this change is redistributed over these classes as modifications. This means that 21.2% is related to change that is unrelated to vegetation, thus changes that can be attributed to land-cover/use conversion. The domains comprising the Cultivated Areas have an area of 7.9% subject to land-cover/use change dynamics and 19.6% of the total land-cover/use change has become agricultural area. The net gain is significant with 11.7%. The two LCCS major land-cover categories (Semi-) Natural Vegetation (A12) and Cultivated Areas and Managed Lands (A11) (Table 5-3) explain 99.1% of total change with a redistribution of the areas subject to change of 89.6% to the same categories.

Figure 5-3 indicates clearly that two types of changes are spatio-temporally dominant: (1) from Broadleaved Forests into Broadleaved Woodlands (50,352 ha); and (2) from Broadleaved Forests into Herbaceous Crops (56,977 ha). These changes are followed by changes of a more limited extent like Broadleaved Woodlands into Grasslands (20,660 ha), Broadleaved Forests into Grasslands (14,545 ha) and Herbaceous Crops into Built-up Areas (14,121Ha).

Broadleaved Forests are the land-cover group with the largest spatio-temporal aspect of change dynamics (139,829 ha).

A closer look at the land-cover domains is taken by calculating the type of change, i.e. modification within its domain or across related domains, or conversion between non-related domains. The interest of the study is in particular in the forest and woodland (open forest) vegetation types, as well as in those types that can be used for pasture. A better insight is gained by studying what type of change prevails in these LCCS domains in order to discover possible trends (Table 5-5):

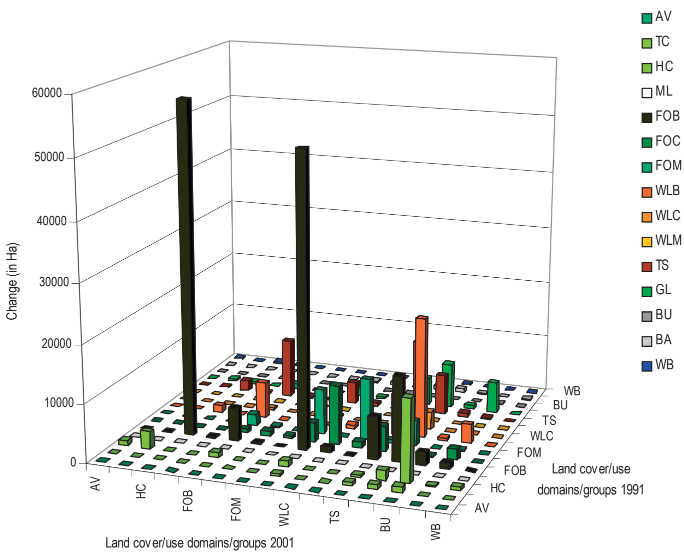
- In the Forest (FO) domain modification into Woodland (WL) is prevalent as 25.6% of change can be explained by it. The second most important change as mentioned above, consisting of 17.3%, is conversion into Herbaceous Crops (HC).
- In the Woodland domain the modification to Grassland (GL) consist of 7.2%.
- In the Thicket and Shrubland (TS) domain the most important change is modification within the domain with 3.9%, followed by modification into Forest with 3.2%.
- In the Grassland domain conversion to Bare Areas (BA) with 1.6% and modification within the domain with 1.4% are the most significant among the changes.

From the above it seems that there is a gradual shift from Forests to Woodlands, Woodlands to Grasslands and Grasslands to Bare Areas. The highest percentages of change are found in the vegetation types were trees were, or still are but to a (much) lesser degree, dominant. In each of these vegetation types especially -but not only!- the tree layer has become less dense with time. Analysis of the semantic aspects of the LCCS parameter options reveals that in all cases the parameter canopy cover of the life form trees has changed either from closed to open, closed to sparse, or from open to sparse. Since the change study is based upon remote sensing no statement can be made about the height and state of the vegetation. A logical explanation would be that these natural resources have been depleted as a result of deforestation (e.g., illegal cutting). Forests also show that a considerable part of them have been converted into agricultural fields, a change with a more permanent character. The Thickets and Shrublands, though, show a different development and sometimes even a return to a tree dominated vegetation type. Human or animal pressure on the environment may sustain certain vegetation types; if this pressure falls away, the vegetation might regenerate.

Analysis of the Broadleaved Forests and Woodlands (Open Forests) at the class level improves the understanding of their change dynamics, especially since Broadleaved Forests change into Broadleaved Woodlands and the latter in turn change into Grasslands. The following classes are involved (note that percentages given concern total change and are not relative to the country territory):

1. 'Broadleaved deciduous forest (*Quercus* spp. and/or *Ostrya* spp. are dominant) usually coppice' mixed with 'Cultivated areas with herbaceous crops on sloping land' (CXB) with 21.0% (17.1% converted into 'Cultivated areas with herbaceous crops on sloping land' (1CS));
2. 'Broadleaved deciduous forest (*Quercus* spp. and/or *Ostrya* spp. are dominant) usually coppice' (2BD) with 15.0% (10.0% going to 'Broadleaved deciduous open forest (*Quercus* spp. are dominant) usually coppice' (2DO));
3. 'Broadleaved deciduous open forest (*Quercus* spp. are dominant) usually coppice' (2DO) with 11.5% (with 3.8 and 2.0% going to 'Sparse trees and shrubs with open to closed grass cover and rock outcrops (pastures)' (2SR) and 'Sparse trees and shrubs with very open grass cover and rock outcrops (pastures)' (2GR) respectively);
4. '*Fagus silvatica* pure and mixed with coniferous forest' (2FC) with 7.8% (with 3.2% going to '*Fagus silvatica* pure and mixed with coniferous open forest' (2FO)); and
5. 'Broadleaved deciduous forest with dominant *Fagus silvatica*' (2FS) with 5.9% (with 3.6% going to 'Broadleaved deciduous open forest with *Fagus silvatica* dominating' (2FB)).

Figure 5-3. Land-cover/use change by LCCS domain or land-cover group[32]



[32] AV=Aquatic vegetation, TC=Tree and shrub crops, HC=Herbaceous crops, ML=Managed lands, FOB=Broadleaved forests, FOC=Coniferous forests, FOM=Mixed forests, WLB=Broadleaved woodlands, WLC=Coniferous woodlands, WLM=Mixed woodlands, TS=Thickets and shrublands, GL=Grasslands, BU=Built-up areas, BA=Bare areas, WB=Water bodies.

The only other significant change dynamic at class level is the conversion from 'Cultivated areas with herbaceous crops on level land' (1CU) to 'Built-up areas' (5UR), comprising 3.3% of total change, where cultivated fields have been replaced by constructions. This is an important change as in case 1 above the cultivated areas on slopes are increasing, whereas the cultivated field areas on level land are decreasing. This means that especially in the sloping and hilly areas of Albania particular land-cover/use types have changed in favour of cultivated areas. At the same time this may mean that this change occurred where less favourable environmental conditions exist (e.g., shallow soils, steep(er) slopes, difficult access, etc.) and where environmental effects may be detrimental (e.g., land degradation and soil erosion).

Table 5-5. Modification and conversion of the vegetated areas subject to change at LCCS domain level[33]

LCCS Domains	Modifications	Area Ha	%	Conversions*	Area Ha	%
Forests	within FO	10351	3.1	from FO to HC	56977	17.3
	from FO to WL	84390	25.6	from FO to BU	2518	0.8
	from FO to TS	12026	3.7	from FO to BA	3384	1.0
	from FO to GL	25896	7.9			
Woodlands	within WL	983	0.3	from WL to HC	1161	0.4
	from WL to FO	6621	2.0	from WL to BU	367	0.1
	from WL to TS	8417	2.6	from WL to BA	3457	1.0
	from WL to GL	23717	7.2			
Thickets & Shrublands	within TS	12683	3.9	from TS to HC	1802	0.5
	from TS to FO	10456	3.2	from TS to BU	663	0.2
	from TS to WL	4281	1.3	from TS to BA	334	0.1
	from TS to GL	6937	2.1			
Grasslands	within GL	7760	2.4	from GL to HC	397	0.1
	from GL to FO	817	0.2	from GL to BU	742	0.2
	from GL to WL	2223	0.7	from GL to BA	5346	1.6
	from GL to TS	4702	1.4			

But if certain classes lost area to change, other classes can be attributed large parts of the areas subject to change. Such classes are:

- 'Cultivated areas with herbaceous crops on sloping land' (1CS) with 18.3% (17.1% coming from 'Broadleaved deciduous forest (*Quercus* spp. and/or *Ostrya* spp. are dominant) usually coppice' mixed with 'Cultivated areas with herbaceous crops on sloping land' (CXB));

[33] See Figure 5-3 for the codes used.

- ‘Broadleaved deciduous open forest (*Quercus* spp. are dominant) usually coppice’ (2DO) with 14.1% (10% coming from ‘Broadleaved deciduous forest (*Quercus* spp. and/or *Ostrya* spp. are dominant) usually coppice’ (2BD));
- ‘Sparse trees and shrubs with open to closed grass cover and rock outcrops (pastures)’ (2SR) with 12.4% (3.8% from ‘Broadleaved deciduous open forest (*Quercus* spp. are dominant) usually coppice’ (2DO));
- ‘Maquis and garigue (incl. low Med. macchia)’ (2MG) with 8.5%;
- ‘Sparse trees and shrubs with very open grass cover and rock outcrops (pastures)’ (2GR) with 8.2%;
- ‘Broadleaved deciduous forest (*Quercus* spp. and/or *Ostrya* spp. are dominant) usually coppice’ (2BD) with 6.4% (2.8% from ‘Mediterranean macchia/Broadleaved deciduous forest (*Quercus* spp. and/or *Ostrya* spp. are dominant) usually coppice’ (MXB) and 1.9% from ‘Broadleaved deciduous open forest (*Quercus* spp. are dominant) usually coppice’ (2DO));
- ‘Broadleaved deciduous open forest with *Fagus silvatica* dominating’ (2FB) with 6.4% (3.6% from ‘Broadleaved deciduous forest with dominant *Fagus silvatica*’ (2FS)); and
- ‘Built-up areas’ (5UR) with 6.0% (3.3% from ‘Cultivated areas with herbaceous crops on level land’ (1CU)).

A number of figures illustrate where the main land-cover/use changes are found in Albania. TheFigure 5-4, Figure 5-5 and Figure 5-6 show the spatial distribution of the three main changes at class level with the areas where the change in class occurs in dark colour and with all polygons subject to change shown in light grey (they do not concern the total area of change but polygons where the boundaries and/or label changed). Figure 5-4 shows the change ‘Broadleaved deciduous forest (*Quercus* spp. and/or *Ostrya* spp. are dominant) usually coppice’ mixed with ‘Cultivated areas with herbaceous crops on sloping land’ (CXB) converted into ‘Cultivated areas with herbaceous crops on sloping land’ (1CS) that occurs mainly in part of the south of the territory. Figure 5-5 shows the change ‘Broadleaved deciduous forest (*Quercus* spp. and/or *Ostrya* spp. are dominant) usually coppice’ (2BD) into ‘Broadleaved deciduous open forest (*Quercus* spp. are dominant) usually coppice’ (2DO) that occurs mainly in the north and south, whereas Figure 5-6 shows two related change types ‘Broadleaved deciduous open forest (*Quercus* spp. are dominant) usually coppice’ (2DO) into ‘Sparse trees and shrubs with open to closed grass cover and rock outcrops (pastures)’ (2SR) and ‘Sparse trees and shrubs with very open grass cover and rock outcrops (pastures)’ (2GR) respectively that occur mainly in the south. From these figures as well as further analysis of where certain change types are found, it seems that changes are regional in their occurrence. The class ‘Broadleaved deciduous open forest’ is in the south subject to two types of changes: in part it occurs where before the forest canopy cover was closed and in part it is transformed into Grasslands.

Figure 5-4. Spatial distribution of the change 'Broadleaved deciduous (open) forest, usually coppice' mixed with 'Cultivated areas with herbaceous crops on sloping land' into 'Cultivated areas with herbaceous crops on sloping land' (dark colour)

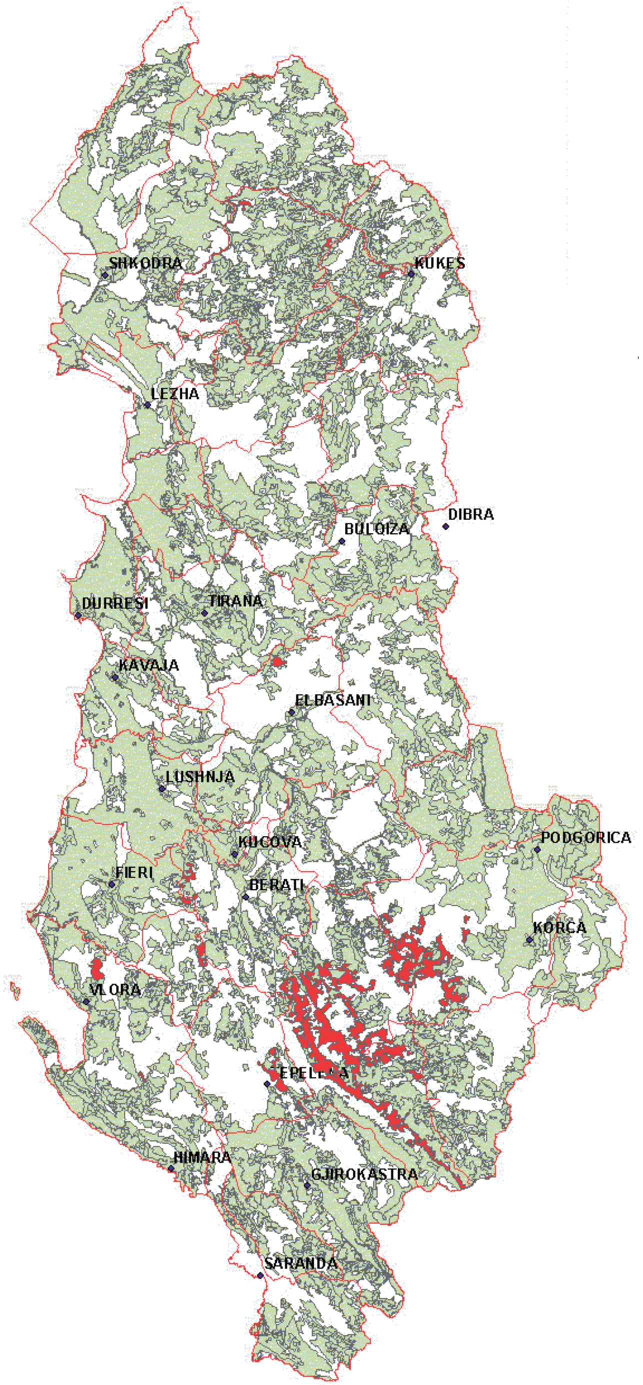
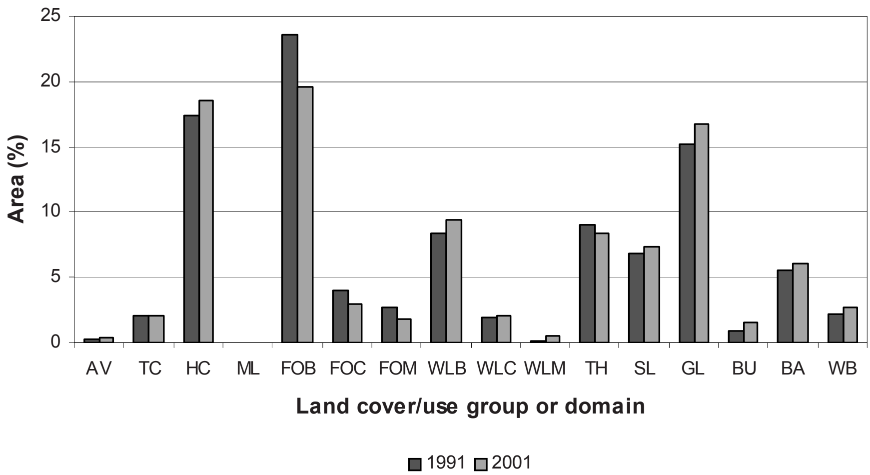


Figure 5-5. Spatial distribution of the change ‘Broadleaved deciduous forest’ to ‘Broadleaved deciduous open forest’ (dark colour)

Figure 5-6. Spatial distribution of the changes ‘Broadleaved deciduous open forest’ into ‘Sparse trees and shrubs with open to closed grass cover and rock outcrops’ and into ‘Sparse trees and shrubs with very open grass cover and rock outcrops’ (dark colour)

It is not sufficient to look at the area subject to change dynamics alone. The changes should also be analysed in relation to the presence of a land-cover/use type in the country territory. Figure 5-7 shows the land-cover groups or domains in 1991 and 2001. From this figure it becomes clear that Albania is dominated by Broadleaved Forests, Herbaceous Crops and Grasslands. Changes in any of these land-cover/use types may have replications to the full area covered by these land-cover/use types. In fact from the analysis of the area subject to change it has become clear that all these land-cover groups or domains are indeed subject to significant change dynamics.

Figure 5-7. Changes in spatial extent of the LCCS domains[34]



However, the change dynamics of the individual land-cover/use group or domains becomes clearer when one calculates their absolute increase or decrease over the period 1991-2001 irrespective of how much of the country territory they occupy as is illustrated in Figure 5-8A. The percentages of absolute increase or decrease are much more pronounced. Forests decrease most, followed by Thickets that decrease at a lower rate; Built-up Areas clearly increase followed by Woodlands and Grasslands, but the latter two at a much lesser magnitude. Figure 5-8A shows clearly that urbanisation is one of the main land-cover/use changes in the country though the area subject to this change is relatively small.

However, if cultivated land with high crop production capacities is lost, these area losses may have relatively important consequences for the total agricultural crop production. The increase in Water Bodies, due to the class Artificial Water Bodies, is neglected in this interpretation as the water levels in the reservoirs depend more on meteorological factors and/or water uses than on change. Many

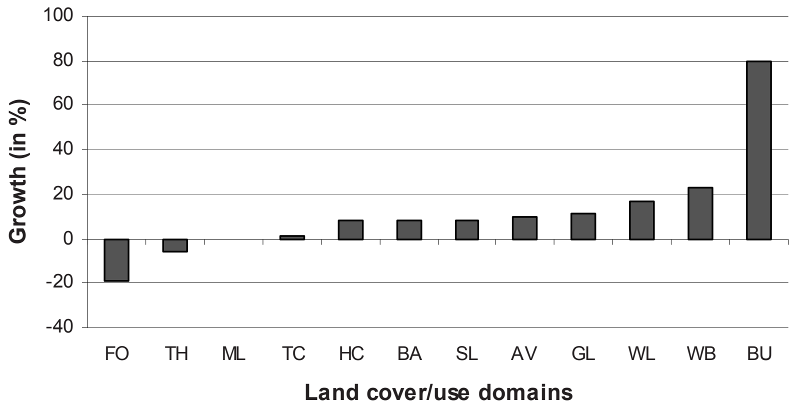
---

[34] See Figure 5-3 for explanation of the codes.

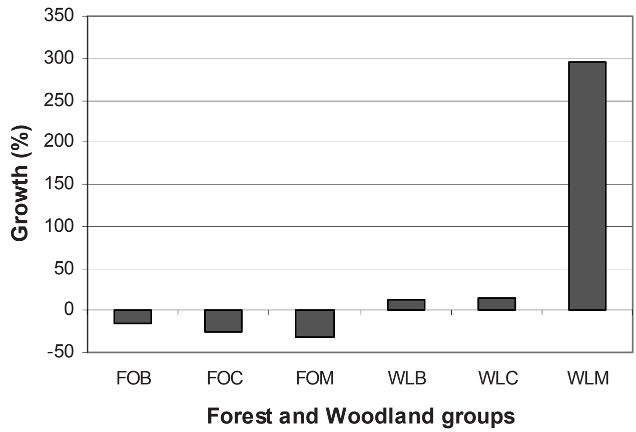
irrigation systems are malfunctioning or have even broken down after the change in economy, thus the amounts of water used have diminished. This change seems more related to the change in agricultural practises than to any real land-cover change, but likely it has greatly influenced the agricultural production of irrigated crops.

Figure 5-8. Absolute changes in spatial extent of the LCCS domains (A) and the Forest and Woodland groups in particular (B) [35]

(A)



(B)



Since the better understanding of the change dynamics concerns in particular the Forests and Woodlands, a closer analysis is made. The growth percentages of these land-cover/use domains, as shown in Figure 5-8A, may hide individual differences at group level. Figure 5-8B shows that the Forests groups are subject

[35] See Figure 5-3 for the codes used.

to a decrease in the range of 16 to 32% with the Mixed Forests occupying only a small part of the country territory but suffering relative greater losses than the Coniferous and Broadleaved Forests respectively. The Broadleaved and Coniferous Woodlands show an increase around 13% while the Mixed Woodlands show an enormous increase of 295%. However, the spatial extent of Broadleaved Woodlands is many times that of Mixed Woodlands (270,444 ha versus 15,797 ha; see also Figure 5-3). Figure 5-8A and B show that the change dynamics may have a greater effect on relative minor land covers/uses in the country territory and consequently may indicate priority areas for the sustainable management of those land-cover/use types that show an unwanted development and where policy interventions may be required (e.g., land cover/use with a great risk to disappear or growth rates with environmental implications).

5.4.2 Analysis of changes at district level

The land-cover/use change dynamics at district level have been analysed in order to examine local variability and an overview is provided of the most significant changes (Figure 5-9). The percentages provided are related to total change within the district and not to the district territory. As expected the most significant change dynamics at national level are also found at district level: changes related to the depletion of natural resources, in particular deforestation, such as the change from Broadleaved Forests into Broadleaved Woodlands (FOB to WLB), or into Grasslands (FOB to GL), or into Herbaceous Crops (FOB to HC), and the change from Broadleaved Woodlands into Grasslands (WLB to GL). But also the change from Coniferous Forests into Grasslands (FOC to GL) is important at district level.

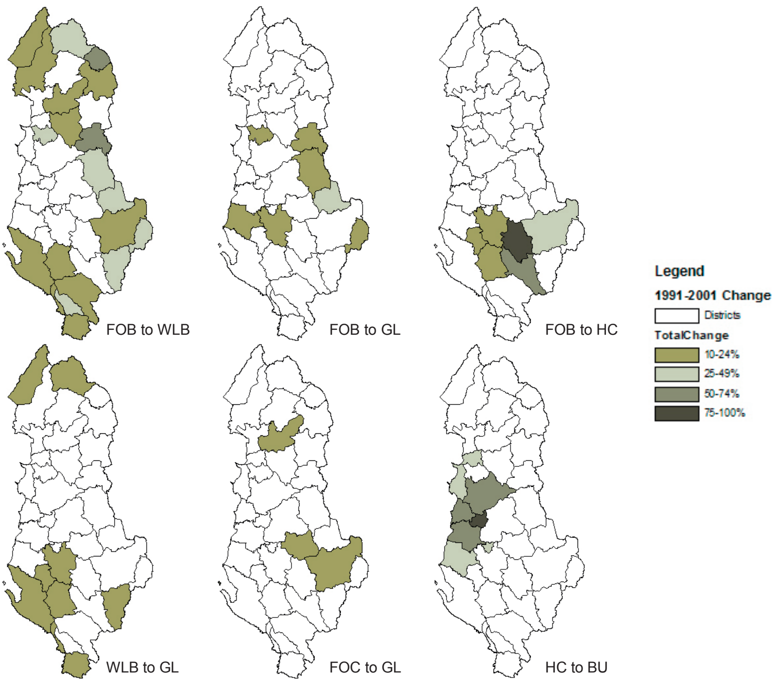
Furthermore, the change related to urbanisation is found, i.e. from Cultivated Areas into Built-up Areas (HC to BU), in those regions where large urban centres are found: the Durres-Kavaja-Tirana triangle, which is the economic centre of the country, but also around Lushnja, Fier and Kucova.

In several districts the most important change dynamic does not correspond to a change that is particularly pronounced at national level. Examples are the change from Broadleaved Forests into Shrublands in Shkoder (30%) and from Mixed Forests into Mixed Woodlands in Diber and Kukes (both 35%). These two changes once again illustrate deforestation. Furthermore, there is the change from Shrublands into Grasslands in Kukove (50%) that may be related to pasture and, finally, the change from Thickets into Broadleaved Forests in Elbasan and Gramsh (48 and 50% respectively) where the tree canopy cover increases substantially.

Modifications within the Broadleaved Forests are important in Has, Librazhd and Tropoje (21, 17 and 11% respectively), within the Grasslands in Durres and Devoll (28 and 31% respectively) and within Shrublands in Fier (25%).

These districts, all of them close together in a specific part of the country, could approach one another when designing and implementing policies and plans for natural resources management. Neighbouring districts that may have experienced, or are experiencing, the same type of change to a lesser degree could learn from such experiences. A decentralised approach seems justified by the data of the study, especially as the Albanian Government is transferring the responsibility for forest areas.

Figure 5-9. Overview of the most important change dynamics at district level (percentages refer to total change within the district)



5.5 Discussion and conclusions

The spatio-temporal and semantic aspects of land-cover/use dynamics in the period 1991-2001 have been analysed for the first time for the whole of Albania through an analysis of spatially explicit data collected through remotely sensed data interpretation and field validation. This analysis has confirmed the major trends of natural resources depletion, in particular deforestation, and urbanisation while at the same time showing that trends are location specific in the country.

The analysis of underlying causes of the observed changes in Albania is limited by the scarce or unavailable spatially explicit data of potential drivers. As a result of not being able to perform a quantitative driver analysis, the interpretation of underlying causes is mostly qualitative, or speculative or based upon other studies.

The transformation of the Albanian forestry sector from the centrally planned and state-implemented model to a market-oriented economy has required major changes in the legal, regulatory and management framework for the sector. Policy failure seems to be one of the essential underlying causes for the widespread natural resources depletion (e.g., corruption, weak or no law enforcement) though policy reforms are progressing and operational and management capacities are being strengthened with international assistance programs (REC for Central and Eastern Europe 2000; SDC/FAO/World Bank/Ministero degli Affari Esteri 2003). However, the conditions of the forests as measured by the forest inventory of the ANFI project give reason for concern. The changes in structure of the forests (e.g., from closed forests to more open forests and from high forests to coppice to shrubs) and the unbalanced age class distribution indicate over-utilisation of the forest resources thus jeopardising the sustainable management. Albania is one of few European countries where forest resources declined in recent decades, in particular during the transition period, as reported by UN-ECE (2001) and the Forestry Project (SDC/FAO/World Bank/Ministero degli Affari Esteri 2003). According to The State of Environment Report almost 30% of the forests and about 50% of the pastures were turned into cultivated areas between 1960 and 1980 showing that forest and pasture resources in the past have been sacrificed to economic development based on intensive agriculture (UNEP 1999). Government has put less and less fuel wood for sale on the market in the 1990s and as a result, given the lack of alternatives for energy supply and the widespread rural poverty, induced the rural population to cut illegally (UNEP 1999; UN-ECE 2001).

Furthermore, a notable rise in pasture activities has taken place as the number of cattle increased considerably more than the increase in area and the increased grazing intensity has caused other forms of natural resources depletion, namely deterioration of the productive capabilities of pastoral areas and environmental degradation (e.g., as a result of low vegetation cover and the trampling of animals the increased manifestation of run-off and soil erosion) and consequently a demand for new pastoral areas. The special study on grazing impact on wooded lands carried out in the context of the ANFI project reports that pastures and meadows have a poor range condition and the stocking rates are four times the grazing capacity, having thus not only implications for the pastures but also for animal productivity (Papanastasis 2003). Important in this context is also the increased cropping of alfalfa for periods of up to five years on the same plot and the increased use of cultivated areas for grazing. This may alleviate in part the pressure on the pastures.

Though the ANFI project confirmed and quantified the poor state of forests and pastures in Albania, this pessimistic situation also provides opportunities. The young stage of forests can be taken as starting point for increasing carbon stocks in Albanian forests through sustainable forest use. Also degraded lands and abandoned cultivated areas form potential areas for afforestation and reforestation. Although Albania adopted the United Nations Framework Convention on Climate Change in 1995, it did ratify the Kyoto Protocol only on the 1st of April 2005. Ratification is a necessary step to enter into the emission trading schemes, including carbon credits, generated by afforestation and reforestation programmes. If these activities are conducted on the basis of synergies among environmental principles, i.e. biodiversity conservation, combating land degradation and carbon sequestration, these may contribute to develop win-win opportunities between environmental protection and conservation, sustainable development and economic growth.

Urban areas occupies only a small area in the country but changes in spatial extent of built-up areas *per se* do not appear to be central to this type of land-cover/use change. However, it is a misconception to think that a change can be ignored if the area involved is only small. The importance of urbanisation lies in the fact that it changes urban-rural linkages (Lambin *et al.* 2001). Consumption expectations in urban centres are higher and will have an impact on areas much bigger than the cities themselves and located at distance (e.g., fuel wood that will need to be brought from the forested areas).

With the presented data on land-cover/use dynamics no statements can be made as to the state of land-cover/use classes because the tool of remote sensing is not sufficient. For the (Semi-) Natural Vegetation classes one should consider that factors such as plant species composition (pastures), plant height or wood volumes (forestry) very likely have changed over time. Nothing can be said in the executed change analysis about the state of vegetation types (e.g., degenerated or not, decrease in tree height or not, deterioration in species composition or not, etc.), the state of cultivated fields (e.g., in active use, fallow or abandoned), or the state of urbanisation (e.g., increase in the number of floors of buildings, decrease in the number of habitants per house, etc.) because these features cannot be derived from satellite remote sensing. However, the forest inventory and grazing impact studies carried out in the context of the ANFI project have provided information on the current state of forests and pastures that can be used for monitoring purposes.

Privatisation of agricultural land has changed agricultural production considerably. Changes in the intensity of use, (mal-) functioning of irrigation systems and land fragmentation should be considered when analysing changes at the level of land use. However, these factors cannot be measured with remote sensing. Combination of the present study results with socio-economic data can provide more conclusive evidence.

Land-cover/use changes do not always occur in a progressive or gradual manner, but they may show periods of rapid and abrupt change followed either by a quick recovery of ecosystems or a non-equilibrium trajectory (Lambin *et al.*

2003). In the present study only two years are available: 1991 describing the land-cover/use situation under the centralised Government and 2001 in a market-oriented economy. The mid 1990s are not represented but stand for the moment in which the land was distributed to rural households and registration as private property took place. It would have been interesting to see how the change dynamics evolved before and after registration as a study at detailed level indicates (Jansen *et al.* 2007).

Considering the above-described limitations of remote sensing for analysis of land-cover/use dynamics, one could state that the present results are more likely an underestimation of change than an overestimation. If more land-use aspects and information would be integrated into the study, the area subject to land change would be likely to be more extensive.

The establishment of permanent forest inventory plots by the ANFI project together with the remote sensing based national inventory of land-cover/use types provides DGFP and DSFD with the technical capability to continue state-of-the-art forest and pasture resources assessments and monitoring programs. The applied inventorying system allows replication of measurements and observations both in the field and through remote sensing. The results show that it is not only important to monitor the extent of natural resources areas but also the quality of these resources. This monitoring should be executed at regular intervals, which hitherto has not been the case. The monitoring system should have a national and a district component as the first is the level at which policies are formulated and the latter is the level at which management takes place and laws should be enforced. Collaboration with the National Environmental Monitoring Program of the Ministry of Environment should be strengthened in order to apply international monitoring methodologies and to enhance the use of limited monitoring equipment. The monitoring and information flow, however, should be focussed on the production of elements for decision making in natural resources management. People's participation in this democratic dialogue should be promoted by increasing the influence of civil society in the decision-making processes.

6 Change analysis with parameterised land-use class sets at commune level

6.1 Introduction

In Albania, the Government has distributed land to rural households instead of restitution of most of the fertile land to a small number of families that would have restored the highly unequal, pre-reform land distribution (Swinnen 1999 and 2000). The transition from 550 large agricultural co-operatives to 467,000 smallholder farms was associated with the fragmentation of land into 1.5 million parcels that often have limited or no access to infrastructure and mechanisation. Most of the agricultural land lies in sloping areas with soils having high erosion risk potentials. Most of the farms are subsistence ones and about 75% of farm production is for home consumption. The lack of information, inadequate extension services, almost no access to bank credit, lack of marketing channels and difficult access to transport are the major constraints for the Albanian farmer. Since around half of the Albanian population is employed in the agricultural sector, a national development priority is a sound land-use policy, allocating land to uses that prevent degradation and yield high long-term returns. The land users should ensure the long-term quality of land for human use, minimise social conflicts and protect ecosystems. All user categories should have enough land with an infrastructure balanced against environmental threats, at reasonable cost and having a well-defined tenure.

The EU PHARE Land-Use Policy (LUP) II project provided GIS-oriented instruments and innovative methodologies to support decision making for land-use policy and planning to the Ministry of Agriculture and Food in Albania. These methodologies and tools have been applied in three representative pilot communes in the northwest, centre and southeast of the country. This chapter illustrates the concepts adopted and results obtained for the analysis of land-use change dynamics over the period 1991-2003. Land-use change is one of the main driving forces of (global) environmental change and therefore central to sustainable development (Meyer and Turner 1994; Walker *et al.* 1997; Walker 1998). Thus, analysis of past land uses and understanding processes and preferred pathways of change will support informed decision making for improved, sustainable and environmentally sound land uses in future.

6.2 Methodology

This paragraph gives a short description of the information system and its basic unit that were used in this study and briefly introduces the methods that were used in the analysis of the land-use changes. The methodology is described more extensively in two LUP II project documents (Carrai 2003; Jansen 2003).

6.2.1 The cadastral land parcel as a basic unit

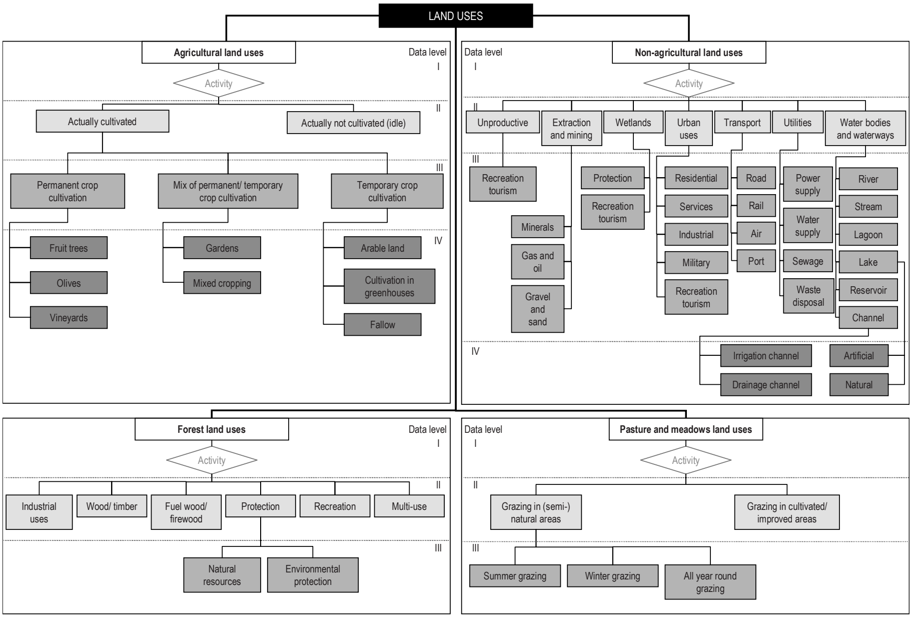
For each piece of land, individuals choose a type of use from which they expect to derive the most benefits in the context of their knowledge, the individual's household, the community, the bio-physical environment and the political structure to which the individual may be subject. These choices vary in space and time resulting in a spatial pattern of land uses. The analysis at the level of the spatially explicit legal parcel unit of the multi-purpose cadastre may show the variability at the level of each cadastral zone while the aggregated level of the commune may show patterns that remain invisible at the detailed scale, and vice versa (Veldkamp *et al.* 2001b). The aggregated level of the commune is important in the land-use policy and planning process while the cadastral parcel unit is a level that corresponds with the decisions made by the individual landowner or land user. It should be clear though, that such decisions may be related to the size of the group that the individual belongs to (Verburg *et al.* 2003). Individuals interact to form groups and organise collective action (e.g. farmer associations).

In general, land registration and the cadastre should be seen as part of the process of natural resources planning and management. The multi-purpose cadastre should therefore be seen as an integral part of the land management system. It is therefore important to establish linkages with a wider range of land-related data, especially those relating to the environment. In this manner, managing land and land information come together (Dale 1995; Larsson 2002).

6.2.2 The Land-Use Information System for Albania

There is significant diversity of opinion about what constitutes a land use (UNEP/FAO 1994). In the context of the project land use is defined as "*the type of human activity taking place at or near the surface*" (Cihlar and Jansen 2001). The developed Land-Use Information System for Albania (LUISA) has adopted, as guiding principles, two criteria that are commonly applied in international systems (Anderson *et al.* 1976; IGU 1976; UN-ECE 1989; UN 1989 and 1998; CEC 1993, 1995 and 1999; FAO 1998; APA 1999): (1) *function* that refers to the economic purpose of the land use and can group many different land-use types in a single category; and (2) *activity* that refers to a process resulting in a similar type of product and is used at the lower levels of the hierarchy (Jansen and Di Gregorio 1998b and 2002). The adopted concept builds upon and exceeds experiences gained in two case studies (Jansen and Di Gregorio 2003 and 2004a). Furthermore, LUISA arranges in a logical and functional manner land uses at different levels of detail and allows data harmonisation with other land-use description systems in use in the country (e.g. statistical office, cadastre and communes).

Figure 6-1. Overview of the LUISA class set with the four main categories of land use (Jansen 2003)



Categories present in the current version of LUISA represent the key categories of the Albanian law on the land: 'agricultural', 'forests', 'pastures and meadows' and 'non-agricultural' land uses (Figure 6-1, see also Figure 4-2). The set of classes in this legend is only a proportion of what one may actually find in Albania. The cadastre in Albania contains information on 1.5 million parcel units with an average size of less than 1 ha. Because of the scale of observation selected, i.e. the cadastral parcel unit, and in order to create in a timely manner a pragmatic land-use database of manageable size (i.e. all records created will need to be maintained and updated at regular intervals), the decision was made that only one land-use class is attached to each parcel unit. At aggregated cadastral parcel levels, mixed classes can be introduced but they do not exist at the most detailed level of LUISA.

The LUISA data, together with other data sets, have been structured according to the European Environmental Agency's Infrastructure for Spatial Information in Europe Directive (the INSPIRE Directive entered into force from 15 May 2007 onwards).

6.2.3 The Land-Use Change Analyses methodology

LUISA contains many classes and thus will result in numerous possible land-use changes that do not facilitate a meaningful interpretation if not grouped in a functional and systematic manner. The developed object-oriented Land-Use Change Analyses (LUCA) methodology arranges the potential land-use changes in three main groups per land-use category in order to underline the change processes: (1) land-use *conversion*, i.e. where a certain land use has been changed into a land use that is very different and the change cannot easily be reversed; (2) land-use *modification*, i.e. changes that are related to one another and where the situation can be reversed; and (3) *no change*, i.e. areas that have remained under the same land use. The parent-child relationships created facilitate the analysis of the spatio-temporal dimensions, i.e. area and perimeter over time (Booch 1994).

In principle, land-use modifications occur within a land-use category and land-use conversion occurs between land-use categories. The exception is the 'non-agricultural' land-use category that contains a larger variety of classes than the other categories; in this category modifications occur within one group (e.g. within 'urban uses') and conversions between groups (e.g. from 'unproductive' to 'urban uses'). Unlikely changes such as a 'residential area' having changed into 'arable land' have been excluded from the change analysis.

6.2.4 Knowledge Discovery in Databases

The Knowledge Discovery in Databases (KDD) process is an iterative procedure of selection, exploration and modelling of large amounts of data that was used to detect *a priori* unknown relationships in the data. The KDD process comprises many elements of which the two most important in the context of this chapter are (Bonchi and Pecori 2003):

1. Data-mining: the most important phase in which, through the use of specific algorithms, previously unknown patterns are extracted from the data that are channelled into a data model;
2. Pattern evaluation: an interpretation and evaluation of the identified patterns and data model in order to create new knowledge.

Some preliminary statistics on correlations between parameters were performed using the On-Line Analytical Process (OLAP) cube for multi-dimensional analysis in order to better understand which parameters to use in the KDD process. OLAP was performed with the following variables: (1) land-use change class, (2) land-use change period, (3) slope class and (4) land suitability.

The variables used as inputs into the decision tree that belongs to the data-mining phase of KDD have been used with the assumption that one of the variables, i.e. land use in 2003, is dependent on the other variables. The use of the variables to construct the decision tree is such that one starts at the initial node with all the available data; then at each step groups are created on the basis of an explanatory variable and in the successive step, each group created will be further subdivided by another explanatory variable and so on until the terminal node. Once a variable has been used, it cannot be used in successive steps (Lombardo *et al.* 2002). From the initial node to the terminal node, a series of decision rules can be extracted of the type IF-THEN. Each decision rule is characterised by a weight and a confidence level that measure the frequency and strength of the decision rule respectively. Decision rules that are valid for many cells have a major weight, whereas those that repeat themselves in the same manner have more significance. The method requires several runs in order to create groups that maximise the internal homogeneity and the external heterogeneity. To create the groups at each level of the procedure, a function is used as an efficiency index known as the 'function segmentation criteria' (Han and Kamber 2000).

6.2.5 Pilot area selection

The choice of pilot communes illustrates the diversity in landforms and (agro-) ecological conditions plus the variety in socioeconomic settings. The choice of Preza Commune was also governed by the fact that it already served as a pilot area in the LUP I project. The availability of suitable digital data sets was a prime criterion for selection.

6.3 Results

6.3.1 The temporal changes in the communes

Each of the three land-use data sets available represents a critical moment in time: (1) the 1991 data represent the land uses under the former centralised Government; (2) the 1996 data represent the time when distribution and registration of the land to the family households took place; and (3) the 2003 data represent the actual land uses in the market-oriented economy.

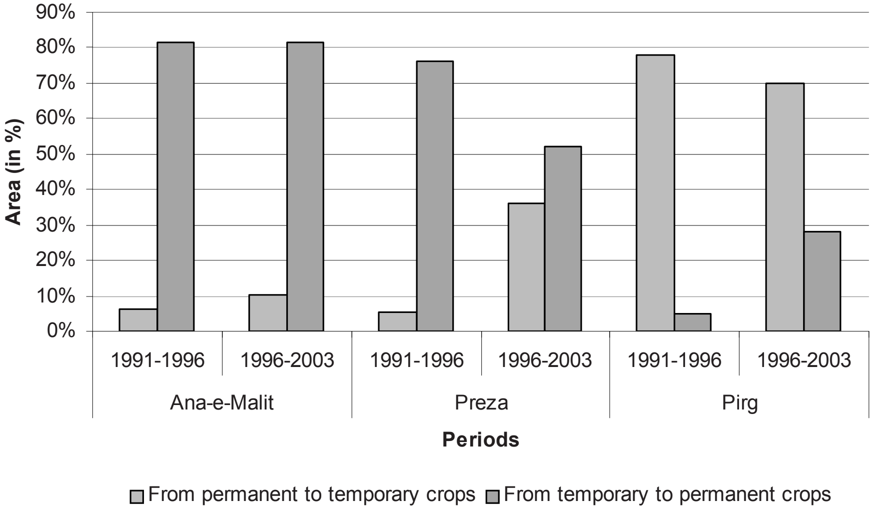
Table 6-1 shows the different types of land-use changes aggregated for the three communes, i.e. Preza, Ana-e-Malit and Pirg, in 1991-1996 and 1996-2003. The communes comprise 2552, 3357 and 2150 ha and are situated in the centre, northwest and southeast of the country respectively. In all three communes, the intensity of changes in 1991-1996, before the land distribution, is higher than in 1996-2003. The majority of parcels were not subject to any change in either period. In Ana-e-Malit and Pirg the area not subject to change increases in the second period, but in Preza it decreases. The main change in land use in both periods involves a land-use modification and in all three communes it is the 'medium-level-modification-in-agriculture', which means that classes in the 'agricultural' land-use category changed at level III, i.e. from permanent into temporary crop cultivation or vice versa. However, the extent of this modification is diminishing in 1996-2003 in Ana-e-Malit and Pirg, whereas Preza shows a clear increase. Land-use conversions are much less important in terms of their extent but their impact may be bigger than that of land-use modifications. The most common conversion is 'agriculture-to-nonagriculture', except in Preza in 1991-1996 where 'pasture-to-agriculture' conversion is dominant. The second most common conversion is ‘agriculture-to-pasture’ in Preza and Ana-e-Malit in both periods and in Pirg in 1996-2003. In Pirg, ‘nonagriculture-to-agriculture’ conversion is important in 1991-1996. It seems that in 1996-2003 in particular, agricultural lands were converted, whereas overall changes were affecting fewer parcels. In this period, land was privatized and apparently many new owners did not want or did not have the means to continue agricultural activities.

Table 6-1. Predominant types of land-use changes (claiming over 1% of the total area) in Preza, Ana-e-Malit and Pirg in 1991-1996 and 1996-2003

Type of land-use change	Preza		Ana-e-Malit		Pirg	
	1991-96	1996-03	1991-96	1996-03	1991-96	1996-03
No change	86.5	80.2	71.7	90.2	81.3	91.9
Medium level modification in Agriculture	4.9	7.6	9.8	1.9	8.2	3.9
High level modification in Non-Agriculture		1.8			1.5	
Agriculture-to-Forest			1.3			
Agriculture-to-Pasture	1.6	1.1	5.6			1.8
Agriculture-to-Nonagricultural		1.1	2.5	2.1	1.4	
Forest-to-Pasture		1.1	2.9			
Forest-to- Agriculture		3.2				
Pasture-to-Agriculture	1.2		1.5			
Nonagricultural-to-Agriculture					2.5	

Concerning the most important change, ‘medium-level-modification-in-agriculture’, more insight is gained when analysing what type of land-use classes result in this type of change. Selection of this change type in the three communes and grouping the class combinations of this change shows that in Preza and Ana-e-Malit in 1991-1996 the trend is to go from temporary to permanent crops, whereas in Pirg the trend in the same period is from permanent to temporary crops (Figure 6-2). In 1996-2003, the trend in Ana-e-Malit remains more or less the same. In Preza, however, the majority of changes still involve the change from temporary to permanent crops though the rate of change is at a lower level than in the previous period, while the change from permanent to temporary crops increases. In 1996-2003, the main trend in Pirg remains the change from permanent to temporary cropping but at a lower level than in the previous period and the change to permanent crops increases. In Pirg, many terraces with fruit trees, the main crop production system, were destroyed in the 1990s; in Preza and Ana-e-Malit projects are underway to plant useful trees (e.g. fruit trees, olives).

Figure 6-2. Detailed analysis of the LUCA change type 301 'medium-level-modification-in-agriculture'



The identified change dynamics have some important repercussions: the permanent cultivation land-use types are usually found on man-made terraces or in landscapes with slopes where the trees stabilise and protect the environment. Further analysis combining the land-use change data with a digital terrain model shows that one of the adverse affects of the change from permanent to temporary crops is increased erosion in hilly areas. Furthermore, there seems to be a shift in agricultural land uses because the area lost in one place and gained in another affects different parts of the commune territory. From the three-dimensional analysis of where such changes are found, it becomes clear that parts of the flat or almost flat areas favourable for agriculture are lost, whereas areas where less or even unfavourable terrain conditions (e.g. steep slopes) exist are gained. This consumption of prime agricultural land, in plains and river valleys of peri-urban areas, blurs the distinction between cities and countryside (Lambin *et al.* 2003).

6.3.2 The spatial distribution of changes in the communes

As physical and social characteristics of communities vary in space and time, so do land-use choices, resulting in a spatial pattern of land-use types (Cihlar and Jansen 2001). If one shows the land-use changes not in the format of statistics but as maps, one can easily identify in each commune areas that were more prone to land-use changes than others.

Figure 6-3. Distribution of land-use changes in the commune of (A) Preza, (B) Ana-e-Malit and (C) Pirg, in 1991-2003 (communes are not shown at same scale)

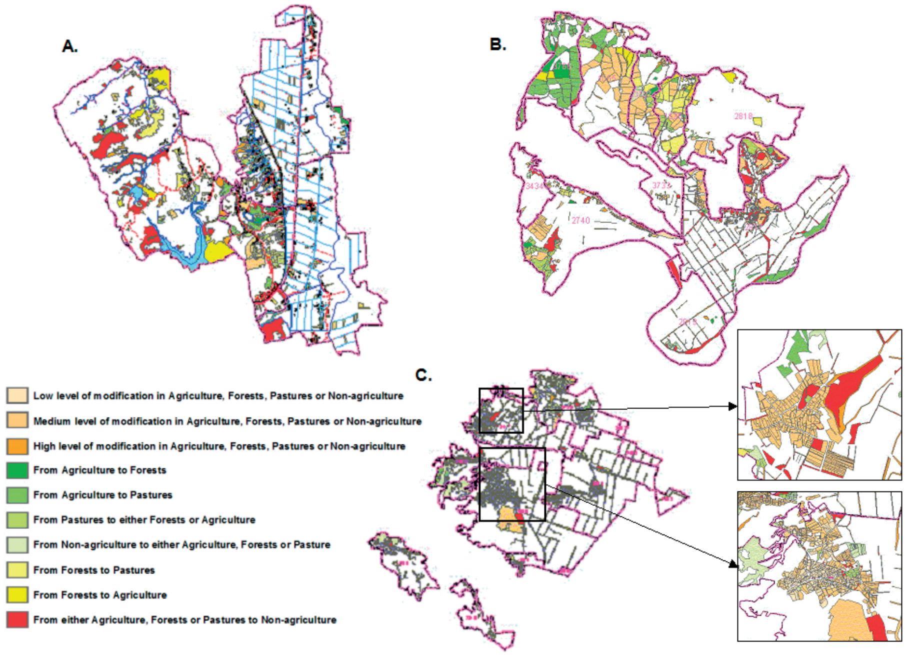


Figure 6-3 shows the distribution of changes over the territory of the communes ranked according to the environmental impact of the change and the fact that Albanian law protects agricultural land, forests and pastures from other uses. The changes with the strongest adverse environmental impact, occurring in protected lands are indicated at the bottom of the figure in the darkest colours. The changes in Preza seem to be divided clearly over the territory: most conversions are found in the western part that consists mainly of hills, whereas most modifications occur in the eastern part that consists of foothills and a plain (indicated by the channel system). In Ana-e-Malit, modifications occur mainly on the foothills and close to the main village of the commune where also the frequency of conversions is highest. In the flatter areas, indicated by the channel system, few changes occur. In Pirg, modifications occur in areas where the land parcels have been divided into many very small parcels close to the villages as shown in the two detailed windows. Also conversions occur in these areas but of a type that is considered to have a positive environmental impact. Large parcels are more often subject to conversions considered to have a negative impact than small land parcels. Also in this commune, the flat areas with channel systems are not subject to many changes.

The areas where land-use conversions occurred that cannot be easily reversed are mainly in the sloping and hilly parts of the communes. In the plains, land-use modifications were dominant, whereas the residential areas grew at the cost of neighbouring land uses.

6.3.3 Preferred pathways of change in Preza Commune

The change dynamics can be related to the landscape position of the cadastral parcel within the terrain and the land suitability for irrigated agriculture, as the communes are predominantly agricultural ones, as well as a set of variables related to what is found in or close to the land parcel. The area of Preza Commune that changed in 1991-1996 and/or 1996-2003 was examined more closely.

A preliminary statistical analysis using OLAP showed that:

- In 1991-1996, more stability concerning land uses exists with around 39% of the total area being classified as no land-use change or 'medium-level-modification-in-agriculture' changes homogeneously distributed within the area involving the various slope and land suitability classes.
- In the same period, transformations are uniformly distributed between the different land-use classes and slope categories. Moreover, there are no major conversions of land use but only some medium-level-modifications.
- In 1996-2003, contrary to the changes in the previous period, a portion of steep sloping lands has been abandoned (20%); this is probably related to abandonment of terraced areas.
- Moreover, in the same period, privatisation of agricultural lands led to encroachment of fields at the costs of forests. Conversion from forests into pastures and meadows is around 10%.
- It is interesting to note that there is a strong relation between slope class and land-use class, i.e. steep lands are always related to land uses like forestry and pastures and meadows.

The data for Preza Commune was used as input into the KDD process in order to identify which variables in the extracted decision rules are important and lead to specific pathways of change. The rules with major weights were chosen first, followed by those with high significance. The territory of Preza Commune was divided in cells of 50 by 50 metres to which a series of attributes are linked from the available data sets. The analysis aims at explaining which factors in or near the cells are important in a specific type of change in either period.

The analysis concerns in particular the 'medium-level-modification-in-agriculture' land-use change and focuses on areas that are either not cultivated or fallow, regrouped under uncultivated, as (temporary) abandonment of cultivated and especially terraced areas is a problem. For the two periods, a set of decision rules was extracted that describe the pathways of change. The complete set of rules for 1996-2003 is almost twice the number of the previous period (719 versus 366 rules), though there are less changes in that period. Two types of rules are extracted, i.e. *transformation* and *inertial rules*, with the description of their conditions (e.g. IF LU_1 and [conditions 1, 2, ...] THEN LU_2). Transformation rules describe a land-use change ($LU_1 \neq LU_2$), whereas inertial rules describe a land use not subject to change ($LU_1 = LU_2$). The extracted rules show that in 1996-2003, the vicinity of the examined cell does not influence the land-use change dynamics in particular. In 1991-1996, one finds the opposite, i.e. the vicinity of the cell is very important for change dynamics. One should also note that in 1991-1996, the extracted rules are essentially inertial rules and transformation rules are few and related to only a few cells, whereas in 1996-2003, there are more transformation rules than inertial rules. Furthermore, the transformation rules for 1991-1996 contain one principal condition that leads to a certain land-use change. In 1996-2003, a principal condition accompanied by more than one set of sub-conditions leads to the same land-use change. The preferred pathways of change are much more complex in the second period. Table 6-2 to Table 6-5 show those rules related to permanent cropping, temporary cropping and uncultivated areas. A change that becomes more evident is that remote areas with either permanent or temporary cropping, often on steeper terrain, and with a lack of infrastructure tend to become uncultivated. So, in these areas the agricultural intensity has decreased dramatically.

Application of the set of decision rules for 1996-2003 to the original data of 1996 resulted in a predicted land use for 2003 with a correlation coefficient of 0.75 with the observed 2003 data. The difference between the average square root of classification (0.15) and average absolute error (0.04) is low, which means the absence of classification outliers. In addition, the accuracy of prediction for each land-use class is above 0.70 with the exceptions of services and industrial areas because the first class is barely present in 1996 and the second absent at that date. Low values for these two classes, however, do not imply that the extracted decision rules involving these classes are erroneous, but they do indicate that these rules are not easily tested and evaluated.

6.3.4 Factors in the decision-making process that drive land management

The land-use change dynamics discussed previously are related to changes in land management that, in turn, are driven by changes in decision-making processes. This decision-making is influenced by factors at different levels with direct or indirect causes (Lombardo *et al.* 2002). A number of such factors, relevant for our case, are discussed below. This inventory is based on the findings of the LUP II project inventories and workshops.

The change in economic system in Albania has forced changes at all levels of organisation. Many land users have a sceptical approach to any form of collective action and at receiving advice from Government related services. Farmers, for example, are reluctant to organise themselves on a voluntary basis in farmer associations and they hardly use the free agricultural extension services. The general lack of information hampers informed and strategic decision making by the rural households. Economic factors and policies, such as taxes, subsidies, credit access, technology, production and transportation costs, define a range of variables that have a direct impact on the decision making by land users. Market access is largely conditioned by Government investments in transportation infrastructure and is identified as one of the major problems and constraints in the communes (Table 2-6). The lack of market access in certain areas has greatly influenced the agricultural production, identified as another major problem and constraint. With mainly semi-subsistence farming and no external demand (or the impossibility to respond to any external demand), the agricultural intensity has decreased dramatically. In the pilot areas, results from the socioeconomic study report that the production of most crops has declined drastically (e.g. wheat by 50%; tobacco, sunflower, sugar beet and soya by 25-33%), whereas the area of forage crops (e.g. alfalfa) increased by 17% and so did livestock production. The only crops experiencing an increase in area and production are vegetables, though mainly used for self-sufficiency purposes. Another result of the land distribution was the changed access to non-land assets such as agricultural equipment. If farmers have no or little access to machinery and labour needs to be executed manually, agricultural production will suffer. Thus, the tendency of rural households active in farming is to move towards a mixture of livestock and forage production. Crop types that are in competition with imports from EU countries in the internal market especially lose out in this competition and, as a result of their low quality and the lack of facilities, cannot be exported to an external market (e.g. CIS countries). It should therefore not come as a surprise that because of the many difficulties, 47% of the rural households in the pilot communes decided to be active in agriculture only part-time. The low agricultural productivity levels can be seen as an indicator of the non-ability of the land users to adapt to changed circumstances as described by Lambin *et al.* (2000).

Erosion and land degradation, flooding and sedimentation (especially in the floodplain of Ana-e-Malit) and pollution and solid waste problems mentioned in Table 6-6 can be seen as other indicators of the fact that, in the pilot communes, the ability to adapt to changed circumstances is very limited. Another factor influencing the decision making of the land users is land tenure. The farm sizes in the pilot communes are very small: 78% of households have a farm smaller than 1 ha distributed over 3 to 5 land parcels. Correcting land fragmentation is therefore considered important in Albania, as in many other parts of Central Europe (Van Dijk 2003a). Graefen (2002) confirms that land fragmentation is putting an additional burden on farm management. But the question is if land consolidation is meaningful considering the average farm size of a rural household, i.e. if four parcels of less than 1 ha farm are re-allocated one can still not make a decent living. In such cases, off-farm income can supplement the revenues from the farm, thus overcoming the farm size restriction. Small farms may make sense in some labour-abundant agricultural economies in the short run; in the longer run, the transition to a modern state means that farm size must be sufficiently large (Rozelle and Swinnen 2004).

Table 6-2. Transformation rules for land-use groups in 1991-1996

1991	Principal conditions	1996	Cells
Permanent cropping	High level of uncultivated in vicinity	Uncultivated areas	167
Temporary cropping	High level of olive trees in vicinity	Permanent cropping	28
	Medium level of uncultivated in vicinity AND Original parcel medium-to-low in size AND No water in vicinity	Uncultivated areas	15
	Low number of buildings in 500m AND High level of residence in vicinity	Urban areas	68

Table 6-3. Inertial rules for 1991-1996

Land use	Principal conditions	Secondary conditions, if applicable	Cells
Permanent Crops	Near to urban secondary road AND No pastures/meadows in vicinity AND No residence in vicinity AND No water in vicinity AND No uncultivated in vicinity AND No transport in vicinity		173
	Medium-to-high level of olive trees in vicinity		54
	Low level of uncultivated in vicinity AND Medium level of olive trees in vicinity AND No forests in vicinity		40
Temporary Crops	High level of crops in vicinity	AND No olive trees in vicinity AND No uncultivated in vicinity AND No water in vicinity AND No fruit trees in vicinity AND No pastures/meadows in vicinity	2245
		AND Low number of buildings in 500m	223
		AND Low level of fruit trees in vicinity	137
		AND No fruit trees in vicinity	125
	Medium-to-high level of crops in vicinity AND No buildings within 500m AND Poor road condition	AND No water in vicinity	283
		AND No distance from the edge of cell AND No olive trees in vicinity AND No uncultivated in vicinity	103
		No residence in vicinity AND Small distance to the edge of cell	93
	Medium-to-high level of crops in vicinity	AND No buildings within 500m AND Low erosion risk AND No olive trees in vicinity AND No forests in vicinity AND No fruit trees in vicinity AND Original parcel low in size AND Small distance to artificial watering canal	150
		AND No residence in vicinity AND Fair road condition	96
	Medium level of crops in vicinity AND No buildings within 500m AND No olive trees in vicinity AND No uncultivated in vicinity AND No water in vicinity		264

Table 6-4. Transformation rules for land-use groups in 1996-2003

1996	Principal conditions	Secondary conditions, if applicable	2003	Cells
Permanent cropping	Medium-to-high level of fruit trees in vicinity	AND Original parcel medium in size	Uncultivated	48
		AND No additional condition		23
Temporary cropping	100%>slope>75% AND Near to urban main road AND No pastures/meadows in vicinity		Permanent cropping	21
	100%>slope>75% AND Low number of buildings in 500m AND No transport in vicinity AND Medium-to-low erosion risk	AND Original parcel low in size AND Small distance to nearest road	Uncultivated	45
		AND Original parcel high in size AND No pastures/meadows in vicinity		36
	100%>slope>75% AND Gravel loose road in vicinity AND Medium-to-low erosion risk	AND Low number of buildings in 500m AND No transport in vicinity AND Original parcel medium in size AND No residence in vicinity		71
		AND Very low drainage value AND Medium-to-small distance to natural watering canal		19
	Gravel loose road in vicinity AND No Residence in vicinity	AND Low number of buildings in 500m AND No transport in vicinity AND Original parcel low in size AND No pastures/meadows in vicinity AND Low level of olive trees in vicinity		48
		AND 75%>slope >50% AND Original parcel medium in size		26
	Slope < 25% AND Low number of buildings in 500m AND No forests in vicinity AND Near to urban secondary road AND No transport in vicinity AND No residence in vicinity AND Low drainage value AND No water in vicinity			53
	Medium-to-high distance from artificial watering canal AND Medium distance from nearest road			21
Uncultivated	Paved road in vicinity AND Original parcel medium in size	AND No crops in vicinity	Permanent cropping	49
		AND No pastures/meadows in vicinity AND No forests in vicinity		28

Table 6-5. Inertial rules for 1996-2003

Land use	Principal conditions	Secondary conditions, if applicable	Cells
Permanent cropping	No Fruit trees in vicinity AND Paved road in vicinity AND Medium-to-high distance from artificial watering canal AND Medium-to-low number of buildings in 500m		87
	No Fruit trees in vicinity AND No transport in vicinity AND No Water in vicinity AND No forests in vicinity AND No crops in vicinity AND Medium distance from artificial watering canal		54
Temporary cropping	Slope < 25% AND Original parcel low in size AND No buildings inside	AND Small distance to nearest road AND No water in vicinity AND Low number of buildings in 500m AND No residence in vicinity AND Unpaved road in vicinity	133
		AND Paved road in vicinity AND Medium distance from artificial watering canal	59
Uncultivated	Unpaved road in vicinity AND Low number of buildings in 500m AND Medium distance from nearest road		64

Table 6-6. Main problems and constraints as identified in the pilot areas (Agrotec S.p.A. 2004)

Constraints and problems	Preza	Ana-e-Malit	Pirg
Agricultural production	xxx	xx	xxx
Marketing	xx	xxx	xxx
Land tenure (security and size)	xx	xx	xx
Settlement and peri-urban development	xxx	x	xx
Erosion and land degradation	xx	x	xxx
Flooding and sedimentation	x	xxx	x
Pollution and solid waste	xxx	xx	xx

xxx - very serious; xx - serious problem; x – moderate

6.4 Conclusions

For the first time in Albania, the temporal and spatial magnitude of change dynamics at cadastral level was studied in three pilot areas.

Modification is the predominant land-use change type and concerns agricultural lands where temporary crops are replaced by permanent crops or vice versa. In the understanding of the change processes of modification, the decision-making processes of the land users play a key role. Development of future trajectories that include intensification of agriculture should consequently include the decision-making processes of these farmers though policies usually address more aggregated levels (e.g. district or national levels). A study carried out at national and district levels may obscure the existing local variability of spatially explicit land-use changes, whereas it may show patterns that, at more detailed data levels, remain invisible (Jansen *et al.* 2006a). Understanding land-use change dynamics is foremost concerned with the quantities of change, i.e. the amount of area changed and the amounts of inputs used and/or production per unit area gained or lost as a function of management level.

In 1991-1996, the observed changes were still influenced by a central planning policy, most likely due to the persisting influence of former officials, technicians and experts still considered to be a reference in land use. With the collapse of central Government, the absence of any planning authority and without any improvement in the land market, land uses were mainly preserved where environmental conditions were more favourable, and degradation occurred where environmental conditions were less favourable. With the beginning of a land market and corresponding lack of regulation and legislation in 1996-2003, land-use changes were more dynamic. The greater number of pathways for 1996-2003 seems to confirm that the new landowners of the cadastral parcels each went their own way without any level of Governmental land-use planning involved.

The analysis of preferred pathways of change in Preza Commune indicates that the land users take rational decisions when they change land use because of, for example, low suitability or unsuitable soils for a particular use and they seem to abandon steep lands where erosion phenomena manifest themselves. The socioeconomic evolution confirms that before 1991 agricultural output is mainly increased by bringing more (terraced) land into production followed by the intensification of production through fertilizer use and/or irrigation. After 1996, the costs of maintenance of these terraced areas and, more important, the division of this area not according to contour lines but perpendicular to the terracing led to the prevalent use of these areas for pasture. Furthermore, the areas most suitable to agriculture, well served with infrastructure and close to urban centres, have in general maintained their production characteristics. In the case of urbanisation, green areas around buildings have been maintained for production of fruit and vegetables for self-sufficiency purposes of the family household. These developments are especially surprising in the absence of any regulating plan. Thus the farmers' perception of such areas has guided their decisions on crop allocation.

Trajectories of land-use change involve both positive and negative human-environment interactions. The interactions between the driving factors and impacts of land change are often referred to as feedback mechanisms that operate over different spatial and temporal scales (Veldkamp *et al*. 2001a). The extracted rules, i.e. the pathways of change, for Preza Commune could be particularly critical when both types of rules indicate negative developments at national level such as the trend confirming that individuals tend to exploit better environmental conditions for their own benefit while a planning policy should distribute resources and exploitations over the area in a well-balanced manner. Indirectly, these results should stimulate the Albanian Government to develop a land-use policy and strongly invest in land-use planning to prevent the permanent deterioration of the environment with non-reversible transformations. Also non-spatial policies like subsidies could play a role.

Land-use change analyses assist the Government in defining those areas where certain land-use processes and patterns are undesired or cause negative environmental impacts that need to be mitigated. It will assist in prioritising areas for the definition of land-use planning interventions in the three pilot communes and development of sustainable future land-use trajectories. Spatial analysis can thus be instrumental in land-use planning and informed decision-making. In addition, an analysis of change may not only help to identify vulnerable places but also vulnerable (groups of) people that on their own are incapable of responding in the face of environmental change.

PART IV - MODELLING DYNAMICS

Time present and time past
Are both perhaps present in time future,
And time future contained in time past.

T.S. Eliot (1888-1965)

Abstract

To define and explain the interaction of human-environment systems, understanding the scale of interaction and the scale of different environmental and social processes is of paramount importance. There are three dimensions of scale: space, time and the organisational hierarchy as constructed by the observer. The latter dimension of scale has received little attention. The variation in semantic contents of data expressed as differences in categorisation is synonymous with organisational hierarchy. In this chapter the relationship between semantic contents of data with modelling dynamics is explored using two land-cover data sets for Romania, one based upon the Land-Cover Classification System (LCCS) and the other as used in the EURURALIS study. Three levels of semantic contents of the LCCS data and the single semantic level present in the EURURALIS data are used to establish empirical relations between the land-cover class and its driving factors. The methodology of the CLUE model is used as the spatial and temporal dimensions of land change have been explored with this model and the examination of the variation in semantic contents of data is complementary to the earlier research. The results show that variation in semantic contents of data within one data set and between two data sets lead to different sets of spatial determinants. There is no pattern recognizable when establishing the organisational hierarchy. Future policy and decision making depend to a great extent on which organisation hierarchy is present in the data used to formulate a policy or to make an informed decision. This would mean that if the same results would be found in other data sets using different models not only multi-scale but also multi-semantic analysis are needed in order to make meaningful predictions of spatially explicit land change.

Based on: Jansen, L.J.M., Veldkamp, A., 2010. Evaluation of the variation of semantic contents of class sets on modelling dynamics. *International Journal of Geographical Information Science* under review.

7 Evaluation of the variation in semantic contents of class sets on modelling dynamics

7.1 Introduction

An improved understanding and projections of the dynamics of land-use and land-cover change as inputs to and consequences of environmental change, and as elements of sustainable development, was the objective of the International Geosphere Biosphere Programme (IGBP) and International Human Dimensions Programme on Global Environmental Change (IHBP) Land-Use and Land-Cover Change (LUCC) project (Turner *et al*. 1995). This project was followed up in 2006 by the Global Land Project (GLP 2005). Two crosscutting issues were defined:

1. *Data and categorisation* examined data availability, data quality and categorisation. Differences in the naming of classes, changes in class definition and adding or removing classes in data sets covering the same area in different periods will create difficulties in the interpretation of actual changes over time from changes in category definition.
2. *Scalar dynamics* recognize that land-change patterns observed at any spatio-temporal scale are caused by complex synergy with changes observed at other analytical scales (Veldkamp and Fresco 1996; Walsh *et al*. 1999; Lambin *et al.* 2000; Hoshino 2001; Veldkamp and Lambin 2001; Veldkamp *et al*. 2001b; Verburg *et al.* 2002; Evans and Kelley 2004; Overmars and Verburg 2006). The scale at which an analysis is conducted may affect the type of explanation given to the observed phenomenon as at each scale different processes have a dominant influence on land use or land cover.

Scale is defined as "*both the limit of* resolution *where a phenomena is discernable and the* extent *that the phenomena is characterised over space and time*" (White and Running 1994). Scale, in the sense of the dimensions of space and time, has been examined in various studies (Verburg and Chen 2000; Kok and Veldkamp 2001; Liu and Anderson 2004; Verburg *et al.* 2004b; Bakker and Van Doorn 2009). Features observed in case studies with a small spatial extent are generally not observable in studies for larger regions. Aggregation of detailed scale processes does not straightforwardly lead to a proper representation of the higher-level process. Changes are often non-linear and thresholds play an important role. Different change processes also have different temporal dynamics. The history of land change is composed of periods whose within-period change rate is quite stationary but the cross-period change rate is considerably different.

Wu and Li (2006) and Jansen *et al.* (2006a) distinguish *three* dimensions of scale: (1) space, (2) time, and (3) organisational hierarchy as constructed by the observer. From these three dimensions, the third has received very little attention. In fact so little that this dimension is not even included in the definition of scale cited above. The organisational hierarchy is synonymous with the variation in the semantic contents of data expressed as differences in categorisation (Feng and Flewelling 2004). The variation in semantic contents of data can be regarded as the joined result of the two crosscutting issues above defined. Categorisation produces data sets comprising classes that have different semantic contents (e.g., class labels). So the classes present in data sets and used in land-dynamics analysis can also affect the type of explanation given to observed phenomena. Many researchers would probably admit that this might indeed have an impact except that such an impact has never been analysed in a systematic way (Parker *et al*. 2002). Moreover, unaware of what possible influences the variation in semantic contents of class sets may have in land-dynamics analysis, this might strongly affect the analysis of ‘preferred’ pathways and future trajectories.

The semantic contents of land-use/cover (change) data are recently getting more attention. Measuring semantic similarity of categories, either before or after data collection or between existing data sets, is an emerging area of research (Ahlqvist 2005a; Jansen *et al.* 2008b). There are various initiatives dealing with the changing context of access to spatial data (e.g., Spatial Data Infrastructures such as the European Union INSPIRE Directive) and the broad recognition that spatial data integration is an essential step in land-change modelling and initiatives (e.g., planning and decision making) that aim to respond to land change (Comber *et al.* 2005a). Increasingly data users become interested in understanding the wider meaning of data, i.e. the concepts adopted and categorisations used.

As with the existence of numerous categorisations, the diversity of modelling approaches seems to indicate that modelling is not of a one-size-fits-all nature for the understanding of spatially explicit land-change dynamics. In fact, nowadays the use of multiple models is advocated because the complexity of land dynamics cannot be addressed by a single model (Castella *et al.* 2007; Overmars *et al.* 2007). LUCC has, from its onset, advocated the development and application of (spatial) models as pivotal tool to understand land dynamics (Lambin *et al.* 1999). Models are a method to identify and explore possible futures (Kok *et al*. 2007). So contrary to change analysis using remote sensing that is essentially looking back in time, models allow looking into the (near) future. But is there a relationship between modelling dynamics and semantics? Until now, no one looked at models as a tool to examine the influence of the variation of semantic contents of class sets, expressed as differences in categorisation. If the explaining factors of land-cover patterns will change with changing resolution and extent, what will they do with variation in semantic contents? This chapter is going to examine if variation in semantic contents of land cover influences the underlying explaining factors, and thus modelling dynamics.

7.2 Study area: land change in Romania in the transition period

In the context of the case study and in order to be able to place the occurrence of certain land-cover types, a broad understanding of land change in Romania in the transition period is necessary.

7.2.1 Land-reform choices

In Romania, land reform is central to the democratisation process and the consolidation of civil society in general (Van Meurs 1999). In order to understand the land-change dynamics in Romania it is necessary to understand the land-reform choices the Government made that resulted in changes in the land-use systems and consequently land cover. The transition from a centrally planned to a market-oriented economy involved *privatisation* of agricultural lands meaning the shifting of ownership of land from collectives and state to private persons (Van Dijk 2003a). The objective is creation of competition in agricultural production, bringing about an increase in efficiency and production. In Romania the principal land-reform procedures were (Swinnen 1999):

- Restitution and distribution (physical) of collective farmlands (58% of Total Agricultural Lands (TAL)); and
- Restitution in part of the state farmlands (28% of TAL).

Collectivised land remained legally privately owned throughout the socialist Government period. With the establishment of collectives only part of the rights to land were transferred from the owners in the collective: the right to use and the right to alienate. Thus, the actual ownership titles in principle remained with the members. Separate land parcels were merged in huge tracks of land that hid the legal patchwork underneath. As a consequence, Government could not but use a land restitution process. Otherwise they would have taken away the legal ownership rights from the ‘former’, but also ‘formal’ or ‘legal’, owners. With restitution the effective property rights transferred to those who possessed the (legal) ownership rights (Swinnen 1999; Van Dijk 2003a). Under socialist Government, farm workers were allowed to have small plots where they cultivated especially fruits and vegetables that played an important role in their fragile food security situation. Thus private farming remained existent under socialist Government and occupied 15.6% of TAL in Romania (van Dijk 2003b).

The land reform started in 1991. The *Land Law* 18/1991 defined the conditions of the end of cooperatives and the redistribution of collectivised land. It made provision for a restitution of the property to the former owners or their heirs (imposing a maximum limit of 10 and a minimum limit of 0.5 hectares), and for a distribution of land to the members of cooperatives who did not own any land before the collectivisation process (0.5 ha). Few months after adoption of the Land Law, law 36/1991 enabled the beneficiaries of land restitution or distribution to form production associations named 'agricultural societies'. This law had the purpose to maintain some economies of scale in agricultural production and to limit the dispersion of the capital of cooperatives (Amblard and Colin 2009).

Whereas 'agricultural societies' were created on the basis of the former cooperatives, 'commercial societies' generally appeared from the privatisation of state farms. These were first converted into 'commercial societies', whose capital was entirely held by the state in shares (law 15/1990). According to law 18/1991, former owners of the land farmed by state farms became shareholders of the 'commercial societies', their number of shares depending on their land acreage. This status gave the landowners the right to receive dividends of the results of the society. Their property rights were thus recognized, but unlike collectivised land, the physical restitution of land was not yet considered. It was only 10 years later that land was returned (up to a ceiling of 50 ha per owner) with the law 1/2000 (Amblard and Colin 2009). Previously, the *Leasing Law* 16/1994 and *Law on Legal Circulation of Land* 54/1998 legalised land sales and land rentals.

The implementation of the Land Law led to the restitution of land to mainly an older and urban-based population: 57% of them are more than 60 years old and only 8.2% are less than 39 years old; 43.1% of them are urban people while 39.1% are employed or retired in rural areas. Only 17.8% of these landowners work in agriculture (Dumitru 2002).

Transfer of ownership of land resulted in the formation of very small land parcels and unfavourable shapes (Riddell and Rembold 2002). This happened in all countries in Eastern Europe and in Table 7-1 the fragmentation indicators for Romania and Albania (see also Chapters 4, 5 and 6), two countries with large rural populations and large agricultural areas with respect to total land area (62.3 and 39.6% according to FAOSTAT (2000; 2001)), are provided.

Table 7-1. Physical land-fragmentation indicators in Romania and Albania (Sabates-Wheeler 2002)

Fragmentation indicator (unit)	Romania (2000)	Albania (2001)
Total size agricultural area (Ha) (FAOSTAT 2000; 2001)	14,857,000	1,139,000
Total number of farms (million)	4.70	0.48
Total number of parcels (million)	40	1.8
Average farm size (Ha)	2.30	1.25
Average plot size (Ha)	0.43	0.25
Average number of parcels (No.)	4-5	3-4
Private farms with more than 1 hectare of land (%)	40	42

But the present landscape features are different from when collectivisation took place because new roads, irrigation and drainage channels were built, and buildings with industrial and residential functions were constructed. The original parcels have often been consolidated during collectivisation making it difficult to locate the exact amounts of land in the same places as they were before collectivisation. Therefore reform laws specify that former owners be restituted land in historical boundaries, if possible, or they receive property rights to a plot of land of comparable size and quality.

Ideally land consolidation should have taken place simultaneously with the land reforms, as it would have reduced the changes that have and will continue to take place in order to accomplish a land parcelling structure adapted to current farming techniques. Those who received land were often unprepared for their new status as landowners and unfamiliar with becoming independent farmers (Bullard 2000). The privatisation of arable land was not linked to the privatisation of the machinery and equipment needed to work the land effectively and profitably (Van Meurs 1999). Furthermore, the social structure of the 1990s is very different from that of the collectivisation period. The economy is no longer rural led but urban centres have become the dynamic engines of development. Rural areas depend largely on the rural-urban nexus (Lambin *et al*. 2001; Riddell and Rembold 2002).

7.2.2 Economic situation of rural areas

After the collapse of socialism in 1989, a large outflow of the population from rural areas could be observed. This flow declined over subsequent years (Rusu *et al*. 2002). This migration to urban centres led to a predominance of elder people in the rural population with a large proportion of pensioners. The relation to land in rural areas has profound implications for agricultural productivity, environmental sustainability, and the economic and social status of rural households. Sustainable development is becoming increasingly important to ensure that land currently used for agriculture will be available in future and resources will be available for future generations.

State farms were allotted the best farmlands and received more Government support to invest in infrastructure and technology. Thus, land distribution meant higher costs of disruption for state farms than the more labour-intensive, low technology collective farms. The state farms were more capital intensive than collective farms and their workers' incomes better. Though differences between collective farm members and state farm members became with time smaller in most CEEC, in Romania these income differences were still relatively large in 1991.

Swinnen (1999 and 2000) and Van Dijk (2003a) note that countries such as Romania with low productivity on collective farms have a significantly higher degree of decollectivisation than those where collective farm productivity was higher. Where collective farm productivity was too low to provide for the basic food security of its members, these members left. Furthermore, there is a positive correlation between the 1993 share of agriculture in the economy and the decollectivisation index. Romania, having 22% of the labour force employed in agriculture, shows a higher degree of decollectivisation compared to CEEC where agricultural employment is less than 10% of the labour force.

Romania has a large rural population and very low incomes in collective farming. The means of production are very poor and irrigation networks are little used by private owners because of their high costs of operation (Rusu *et al*. 2002). With the maximum limit of ten hectares for restitution and distribution of the remaining share of its collective farmland to collective farm workers, the Romanian Government combined equity and efficiency considerations (Swinnen 1999). In combination with the distribution of the rest of the land to collective farm workers, land reforms created a fairly equitable land and welfare distribution. However, the transfer of ownership of land was not accompanied by the transfer of ownership of both upstream and downstream input suppliers and output procurers of agricultural machinery. These facilities, crucial to efficient production, remained the property of the state, which meant that the new owners faced huge constraints in their ability to farm profitably (Rusu *et al*. 2002).

In the 1990s, ‘transition’ to a market economy caused a considerable expansion of individual semi-subsistence holdings. In the socialist era the household plots or micro farms provided additional income, mainly in kind, to wages or pensions. The extension of individual semi-subsistence holdings, in number and in area, came logically from the initial conditions of the post-socialist transition (Pouliquen 2001): (1) fall in employment and in wages, particularly in rural areas; (2) restitution and distribution of plots of land; and (3) reduced subsidies and decline of the state agri-food distribution chains. The maintenance of individual semi-subsistence holdings depends primarily on a family transfer of non-farm incomes from other sources. The loss-making agricultural activity is, however, completely rational from the micro-economic point of view, as these losses are covered in large part by welfare transfers of budgetary origin (e.g., pensions) to the households concerned.

The lack of financing on holdings is due to the very low profitability of agriculture on average, and to the narrow limits of possible budgetary support, combined with difficult and costly access to loans. These are the major direct causes of (Pouliquen 2001): (1) the extensification of techniques and productive orientations of agriculture, following the end of the high subsidisation of the socialist era; and (2) the very limited character of their re-intensification. As a result, the average capital per agricultural employed person remains lower than average EU level.

7.3 Methodology and materials

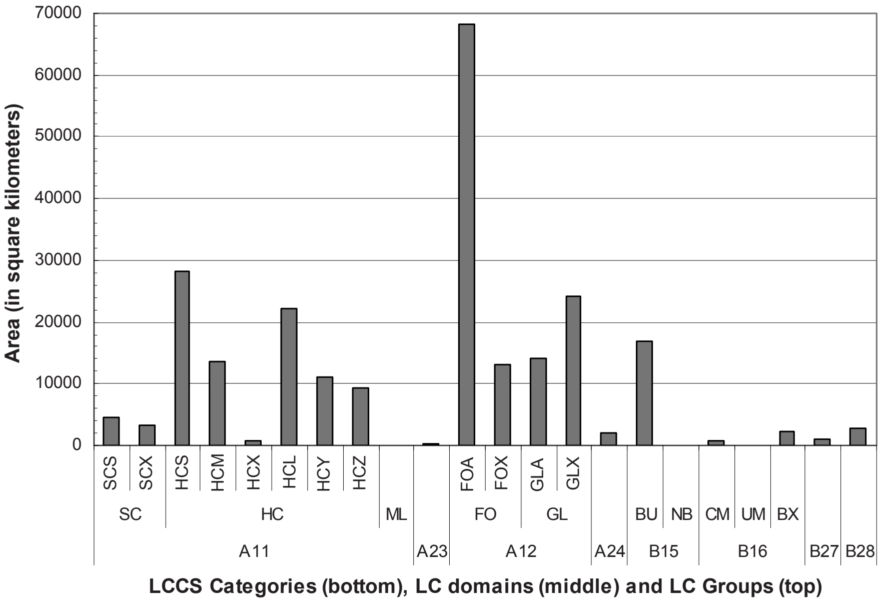
To make a realistic evaluation of the variation in semantic contents of class sets -expressed as differences in categorisation- on modelling dynamics, the drivers of change in the study area, i.e. Romania, have been identified. The drivers are grouped into biogeographical, geomorphological, demographic, accessibility factors and soil variables. The approach selected was more an empirical data-driven approach than a knowledge-based approach. The CLUE (Conversion of Land Use and its Effects) framework is used in this systematic evaluation (Veldkamp and Fresco 1996; Verburg *et al*. 2002). This model has been thoroughly tested and validated in several parts of the world (Pontius *et al*. 2008). The CLUE methodology consist of two parts (Verburg *et al*. 2003): (1) empirically determined relations between land cover and its driving factors explicitly taking scale dependencies into account (e.g., through regression analysis); and (2) dynamic modelling to simulate future land-cover changes. The multi-scale analysis of the driving factors of land change is based on the analysis of spatial patterns of actual land cover. Characteristic for the CLUE methodology is also that no *a priori* levels of analysis are imposed. Instead the analysis is repeated at a selection of artificial resolutions, imposed by the gridded data structure (Verburg *et al*. 2003). For the evaluation of variation in semantic contents of class sets only the first part of the CLUE methodology was used.

The spatial and temporal dimensions of land-cover change have been explored with different versions of the CLUE model and therefore the examination of the variation in semantic contents of class sets is complementary to the earlier research as it adds the semantic dimension. CLUE combines an empirical analysis with dynamic, multi-scale simulations to be able to handle different scenario conditions that may deviate from the historic trend (Veldkamp *et al*. 2001b; Kok and Winograd 2002; Verburg *et al.* 2002). The latter is important in Romania as one of the CEEC, where the historical trend in the period of transition was disrupted.

7.3.1 Dependent variable: land cover

The comprehensive parameterised land-cover data set from the Food and Agriculture Organization of the United Nations (FAO) *Land-Cover/Land-Use Inventory by Remote Sensing for the Agricultural Reform* (TCP/ROM/2801-3001) project for the year 2003 was made available by the Ministry of Agriculture, Food and Forestry of Romania (beneficiary), the Romanian Space Agency (ROSA) (implementing agency) and the Centre for Remote Sensing Applications in Agriculture (CRUTA) (technical co-ordination). This digital data set was produced for a nominal scale of 1:50,000 using LANDSAT Thematic Mapper and Enhanced Thematic Mapper satellite images of 2003 together with ancillary data (e.g., crop calendar, digital terrain model, field data). For definition of classes the Land Cover Classification System (LCCS) was used. The original data set shown in Table 7-2 comprises 68 classes that have been regrouped into eight (categories, first column of the table), 14 (domains, second column of the table, or, when not distinguished category) and 22 classes (groups, third column, or when not distinguished either domain or category). The distribution of the regroupings at the three distinguished levels is shown in Figure 7-1.

Figure 7-1. Distribution of LCCS land-cover types in Romania in 2003[36]



As the interest in land change in Romania is linked to land reform, i.e. *privatisation* meaning the shifting of ownership of land from collectives and state to private persons, particular importance is given to the grouping of field sizes in the cultivated areas. The size of the fields may still be linked to ownership. So there is more detail left in the regrouping of these classes than in other land-cover types. The distinction between permanent and arable crops, often made in statistics, is also maintained. Orchards and vineyards (3.32% of total area) are kept separate from arable crops (35.74% of total area). Forested areas have not been classified in great detail because of the variety of species and complex geomorphology. One could add that for a reliable interpretation of different forest classes multi-temporal imagery would be needed that was not available to the FAO project.

[36] See Table 7-2 for explanation of used codes.

Table 7-2. The LCCS land-cover class set for Romania for the year 2003[37]

LCCS Category	LC Domain	LC Group	User Name	LCCS Code
Cultivated areas & managed lands (A11)	Shrub Crops (SC)	Shrub crops (SCS)	Orchards	10494-1891-W8
			Vineyards	10566-1891-S0610
		Mixed units (SCX)	Vineyards/Orchards	10582-1891-S0610 / 10161-W8
			Orchard/Grasslands	10510-1891-W8 / 20203
			Vineyard/Grasslands	10582-1891-S0610 / 20203
	Herbaceous Crops (HC)	Small-sized fields (HCS)	Small-sized herbaceous fields (< 2 ha)	10110
			Gardens with vegetables (< 2 ha)	11135
		Medium-sized fields (HCM)	Medium-sized herbaceous fields (2-5 ha)	10100-11971
		Medium to large-sized irrigated fields (HCX)	Permanently irrigated herbaceous fields (2-50 ha)	10675
		Large-sized fields (HCL)	Large-sized herbaceous fields (5-50 ha)	10100-11341
			Very large-sized herbaceous fields (>50ha)	10100-11341(1)[Z1]
		Mixed unit: Small to medium-sized fields (HCY)	Small-sized herbaceous fields/Medium-sized herbaceous fields (<2-5 ha)	10110 / 10104-11971
			Medium-sized herbaceous fields/Small-sized herbaceous fields (<2-5 ha)	10100-11971 / 10112
			Fragmented small-sized herbaceous fields (<2 ha)/Grassland	10112 / 20204-L4L8
			Gardens with vegetables (<2 ha)/Grasslands (cultivated garden with grass strips)	11135 / 20202
		Mixed unit: Medium to large-sized fields (HCZ)	Medium-sized herbaceous fields/Large-sized herbaceous fields (2-50 ha)	10100-11971 / 10104-11341
			Large-sized herbaceous fields/Medium-sized herbaceous fields (2-50 ha)	10100-11341 / 10104-11971
			Irrigation network on not-cultivated agricultural fields	10236(1)[Z12] / 5002-7(1)[Z13]
	Managed Lands (ML)		City Parks (forest plots in the city)	11177
Cultivated aquatic areas (A23)			Rice fields	3025-S0308
			Cultivated reed fields in Danube Delta (used for heating)	3003-S1201(1)[Z6]

[37] LCCS Category and LCCS Code are standard provided by the LCCS software application; User Names are taken from the FAO project.

LCCS Category	LC Domain	LC Group	User Name	LCCS Code
(Semi-) natural vegetation (A12)	Forest and woodlands (FO)	Forests (FOA)	Forests	20007
		Mixed units: Forests dominated (FOX)	Trees/Shrubs	20016 / 20023
			Trees/Shrubs/Grassland	20008 / 20022 / 20038
			Riverine vegetation (includes trees, shrub & grass)	20008(3)[Z3] / 20022 / 20038
			Woodland/Grassland	20014 / 20046
			Sub alpine (grl/jnepeni) (*Pinus Mugo*)	20120-P3Zt1 / 20039-P3 / 6002-1-P3
			Shrubland/Grassland	20023 / 20203
	Grasslands (GL)	Grasslands (GLA)	Grassland	20194
			Mountain grassland (>1500 m altitude)	20203-P3
		Mixed units: Grasslands dominated (GLX)	Grassland/Trees/Shrubs	20194 / 20016 / 20023
			Grassland/Shrubland	20027 / 20023
			Grassland/Cattle stables/Building for shepherds	20040 / 5003-A44Zp7 / 5003-A44Zp3P2
			Grassland/Shrubs/Trees	20194 / 20023 / 20016
			Grassland/Trees	20203 / 20016
			Cottages (grass/gardens/houses)	20204-14301 / 10112 / 5003-15-A44Zp4
(Semi-) natural aquatic vegetation (A24)			Wetlands	40048-L13 / 40054-L13
			Peat marshes or peat bogs (at high altitudes)	40219-4732-P3
			Open swamp vegetation in Danube Delta	40413-287-P5Zt2(2)[Z15]
		Mixed unit	(Semi-) natural aquatic vegetation in Danube Delta	40413-1-P5Zt2(2)[Z15]/40168-287-P5

LCCS Category	LC Domain	LC Group	User Name	LCCS Code
Artificial surfaces (B15)	Built-up areas (BU)		Built up areas (all, which is not IBA, ABA or CRA)	5003
			Greenhouses	5003-8--A44Zp1
			Airports	5003-8--A21
			Cemeteries	5003-A23
			Maritime ports	5003-8--A32
			Industrial built-up areas	5003-8(2)[Z10]
			Agricultural built-up areas	5003-8(2)[Z11]
			Medium density urban areas	5003-14
		Mixed unit: Built-up areas	Low density residential areas/Gardens (villages)	5003-15 / 11155
	Non built-up areas (NB)		Dump sites	5004-1
			Tailing pond (near industrial areas)	5004-1-A44Zp2
			Extraction sites	5004-2
Bare areas (B16)	Consolidated materials (CM)		Bare rocks (top of mountains, road excavation)	6002-1-U1
			Bare areas	6005
			Bare soil (mainly along rivers)	6005-U3
			Terra Rossa (from the erosion of limestone)	6005-M299Zm1
			River bed	6005(3)[Z4]
	Unconsolidated materials (UM)		Beach and sandbars (on the sea coast)	6006-P5(3)[Z9]
	Mixed units: Bare areas (BX)		Sands/Bare soil (Grinduri) (in the Danube Delta)	6004-P5U1(1)[Z14](1)[Z16] / 6005-P5
			Bare areas/Grassland	6005 / 20203
			Bare soil/Waterlogged vegetation/Grassland (Grinduri 2) (in the Danube Delta)	6005-P5(1)[Z16] / 40064-P5R2 / 20059-P5
Artificial water bodies (B27)			Channels (only in the Danube Delta)	7001-1
			Reservoirs	7002-5
			Fishponds (also those in Danube Delta)	7002-5-U1(2)[Z5]
			Impoundments (shallow water bodies 10-20 ha)	7013-5-U1
Natural water bodies (B28)			Oxbows (old branches of the river)	8001-5-U1(1)[Z8]
			Rivers (only the big rivers, others - as lines)	8002-1
			Lakes	8002-5

Table 7-3. The EURURALIS land-cover/use data set (Hellmann and Verburg 2006)

Code	EURURALIS land-cover class
0	Built-up area
1	Arable land (non-irrigated)
2	Pasture
3	(Semi-) natural vegetation (including natural grasslands, scrublands, regenerating forest below 2m, and small forest patches within agricultural landscapes)
4	Inland wetlands
5	Glaciers, snow, sands and sparsely vegetated areas
6	Irrigated arable land
7	Recently abandoned farmland (includes very extensive farmland not reported in agricultural statistics, herbaceous vegetation, grasses and scrubs below 30cm)
8	Permanent crops
9	Forest

The elaborated land-cover/use data from the EURURALIS project has been described in detail by Verburg *et al.* (2006; page 47) and Hellman and Verburg (2006). It contains 10 classes (Table 7-3). Important to know is that an effort was made to preserve heterogeneous agricultural areas in this data set as these represent landscapes with high spatial variability. This non-parameterised land-cover data set is used for reasons of comparison: it comprises a set of classes with different semantics.

7.3.2 Independent variables

The other elaborated thematic data for Romania come from the EURURALIS project and were made available by the Land-Dynamics Group of the Department of Environmental Sciences of Wageningen University and Research, The Netherlands. The variables in this data set are described in detail by Verburg *et al*. (2006 and 2008).

Table 7-4 lists the broad selection of independent variables of possible determinants that are used in the statistical analyses. It provides a short description in addition to the origin of the data. The explaining factors are divided into five functional groups to facilitate interpretation:

- *Biogeographical factors* include factors that describe the elevation and slope, environmental regions and factors based on temperature that influence the growing season as well as water deficit in the growing season.
- *Demographic factors* include population potential and the proportion rural-urban population.
- *Geomorphology* is based upon average height differences in the terrain.
- *Soil variables* include variables that influence the productivity of crops; and
- *Accessibility factors* expressed as travel time to cities with a certain number of inhabitants, travel time to major roads and major airports, and ports with a certain number of tons of freight per year.

Table 7-4. Information on the variables included in the regression analysis

Group	Variable	Description	Source
Bio-geographical	ENVMAP05 ENVMAP06 ENVMAP08	Environmental region 'Alpine South' Environmental region 'Continental' Environmental region 'Pannonian'	Metzger *et al*. (2005)
	DEM SLOPE	Height (in m) Slope (based on DEM) (in degrees)	USGS GTOPO30 (edcdaac.usgs.gov/gtopo30)
	MEAN_TEMP	Mean temperature 1961-1990 (in ^{0}C)	New *et al*. (1999)
	T_MIN0 T_PLUS15	Count of months with average temperature < 0 0C Count of months with average temperature >15 ^{0}C	ELPEN database (www.macauley.ac.uk/elpen)
	DDW_SHORTAGE RAIN_WC_5M RAIN_WC_YR	Water deficit growing season Accumulated rainfall of March, April, May, June and July Accumulated rainfall per year	Hijmans *et al*. (2005) (www.worldclim.org)
Demographic	LANDSC SZ_LANDSC_RUR	Population density (in persons/km^2) If LANDSC >100, 100, LANDSC	ORNL LandScan 2004™ (www.ornl.gov/sci/gist/landscan/)
	POPPOT_SUM POPPOT_LOG POPPOT_1MI	Gaussian population potential Gaussian population potential ; logarithmic Gaussian population potential; maximum value set at 1,000,000	Dobson *et al*. 2000
Geo-morphology	GEOMORF01 GEOMORF02 GEOMORF03 GEOMORF04 GEOMORF05	Average height difference of 0-20m (flat) Average height difference of 20-80m (rolling) Average height difference of 80-200m (hilly) Average height difference of 200-400m (mountainous) Average height difference of > 400m (very mountainous)	Computed from the 1000m DEM from SRTM 3 arc-second resolution data from NASA
Soil	CLAYCONT IL PEAT SALINITY SOILDEPTH STONINESS SWAP WR	Soil clay content (percentage) Presence of an impermeable layer within the soil profile Presence of peat in the soil profile Saline soils Soil depth Stoniness Soil water available to plants Soils with water restriction	Soil Geographical Database of the European Soils Bureau (CEC 1985; King *et al*. 1994)
Accessibility	ACCESS_1 ACCESS_2 ACCESS_3 ACCESS_4 ACCESS_5 ACCESS_6 ACCESS_7	Travel time to cities with more than 100,000 inhabitants Travel time to cities with more than 500,000 inhabitants Travel time to ports with more than 15,000 kTon/year of freight Travel time to cities with more than 650,000 inhabitants Airline distance to nearest main road (mainly highways) Travel time to major airports Travel time to major airports and major ports	Accessibility analysis based on GISCO database infrastructure

7.3.3 Data format

The projections of the land-cover data set and the EURURALIS land-cover/use and driver data are provided in Table 7-5. For the analyses the projection of the regrouped LCCS land-cover data set was transformed into the one of the EURURALIS data. Subsequently, the data was rasterised at the high resolution of 1x1km. Each cell in the raster is assigned only one value. Thus, at this pixel size 1x1 km the class that is present in the data is the one with the maximum combined area in the pixel, other classes were lost. This process resulted in the loss of the Managed Lands (ML) class. Thus in the three data sets representing different levels of aggregation the number of classes is 8 (Category level), 13 (Domain level) and 21 (Group level).

Table 7-5. Projections of the data sets

LCCS land-cover data for the year 2003		EURURALIS data set (Verburg *et al.* 2006)	
Projection	STEREO 70	Projection	Albers Equal Area Conic
Datum	Pulkovo 1942	Datum	WGS72
Spheroid	Krassovski	Spheroid	
False easting	500000,00 meters	First standard parallel	32 30 00
False northing	500000,00 meters	Second standard parallel	54 30 00
Longitude of origin	25.00 (25° E)	Central meridian	22 39 00
Scale factor	0.999750	Latitude of origin	51 24 00
Latitude of origin	46.00 (46° N)	False easting	0.0
Linear unit	meters	False northing	0.0
		Unit of measure	meters

7.3.4 Multi-collinearity

Independency between variables is a prerequisite of the statistical method employed. To avoid the effects of multi-collinearity of all pairs of variables with a correlation over 0.80, one is omitted from the analysis. The use of the stepwise regression procedure solves remaining multi-collinearity problems. The omitted variables concern: DDW_SHORTAGE, MEAN_TEMP, RAIN_WC_YR and T_MIN0. Pairs that show very high correlation values are RAIN_WC_5M and RAIN_WC_YR (+0.971), DEM and MEAN_TEMP (-0.944), T_PLUS15 and T_MIN0 (-0.862), DDW_SHORTAGE and RAIN_WC_5M (+0.824).

7.3.5 Spatial autocorrelation

The regression coefficient and significance of the contribution of individual variables are sensitive for the presence of autocorrelation. Overmars (2000) analysed the influence of spatial autocorrelation using a multi-resolution dataset. Results indicated the presence of spatial autocorrelation at the most detailed resolution of 9.5x9.5 km and a rapid decrease at coarser resolutions. Most regression equations varied very little when a spatial autoregressive model was used. The three most important determinants were identical in most cases. Therefore, the possible effects of spatial autocorrelation have been disregarded and the interpretation of equations has been limited to the three most important variables in terms of standardised betas similar to the approach taken by Kok and Veldkamp (2001).

7.3.6 Statistical analysis

The CLUE model is a simulation model to spatially allocate land-use changes (see for CLUE-CR Veldkamp and Fresco (1996) and for CLUE-S Verburg *et al*. (2002)). The quantities of change of the demand-driven land uses are to be determined outside the model. The location preference for the different land-use types is based on the spatial variation of the location factors that are assumed to be important determinants of the land-cover pattern (Table 7-4).The relations are estimated by logistic regression analysis using the land-cover data as dependent variable. Logistic regression is a frequently used methodology in land-use and land-cover change research (Lesschen *et al*. 2005). The same methodology was used by Verburg *et al*. (2004c) to analyse the factors determining land-use patterns in the Netherlands. The occurrence of most land-cover types can be explained by the location factors as indicated by the area under the ROC curve (fit of model). A random model would have a value of 0.50 and a perfect model a value of 1.0 (Swets 1988). The estimated probabilities based on the regression model are used as a proxy for the location preference for the considered land cover.

The relationships between land cover and the selected variables are quantified in a two-step procedure using logistic regression. First, significantly contributing variables are selected with a stepwise regression procedure in SPSS 15.0 using the 0.05 significance criterion. Second, this set of variables is used to construct regression equations. This procedure is repeated for every land-cover class of the three regroupings and the EURURALIS land-cover/use data set. The adjusted coefficient of determination (R^2) serves as a measure for the amount of variation explained. The standardised regression coefficients (standardised betas or β_{st}) are used to indicate the relative importance of individual variables in a given equation.

7.4 Results

7.4.1 Coefficients of determination

The coefficients of determination (R^2) for the statistical models are provided in Table 7-6. A statistically significant model is always established with the highest coefficient of determination for 'Medium to large-sized irrigated fields' (HCX with 0.987) and the lowest coefficients of determination for 'Forest-dominated mixed units' (FOX with 0.651), followed by grasslands at two semantic levels (GLA and GL). As stated before the focus of the dataset was not on forest classes because multi-temporal images to detect different forest types were not available. From forest to grasslands there is a gradient of different vegetation types (e.g., woodlands, thickets, shrublands). As long as trees are dominant the assignment to the corresponding land-cover class (FOA) poses no major problems but below a certain threshold of the presence of trees there is a problem whether assigning the class to the 'Forests-dominated mixed unit' class or to the Grasslands land-cover classes. This results in lower accuracy values for the FOX and GLA land-cover classes. It appears that this, subsequently, results in lower coefficients of determination in the statistical model.

The confusion between more open tree-dominated vegetation types with thickets and shrublands is also reported in detail by Jansen *et al.* (2003a) and Jansen *et al.* (2006b). The latter contains land-cover data at two different spatial scales with the same land-cover class set and at both scales the confusion occurs. Furthermore, great differences also occur between spatial and census data for classes such as grasslands, shrubs and woodlands (Pelorosso *et al*. 2009).

The coefficients of determination are highest for the aggregated classes at category level compared to the domain level with the exception of:

- A11 where HC is better explained at the domain level, whereas SC is less well explained;
- B15 where BU is slightly better explained at the domain level.
- B16 where BX is better explained at the domain level, whereas CM is less well explained.

A12 is much better explained at the aggregated land-cover category level. The classes dominated by water, i.e. A23, A24, B27 and B28, exist only at land-cover category level.

The coefficient of determination at domain level compared to the group level explains:

- SC less well than the group level classes SCS and SCX, though the latter is only slightly better explained.
- HC better for all group level classes with the exception of HCX, the only irrigated class, which is better explained at group level.

- FO is better explained at the domain level.
- GL is better explained than the group level class GLA, but less well than GLX.

The lower explanatory power at a more detailed semantic level demonstrates the importance of underlying driving forces that are more difficult to quantify with very coarse location characteristics. This is clear for A12 at category level compared to the domain and group levels. It is less clear for A11 that shows a clear difference for the SC and HC classes at domain level. The HCX class is best explained at group level, this may plead for the distinction of irrigated cultivated areas from non-irrigated areas as underlying drivers clearly differ.

It is clear that there is no overall preferred semantic level that furnishes a statistical model with the highest explanatory power for all land-cover classes at that level. It depends very much on the type of land-cover class distinguished. At each of the three distinguished semantic levels some land-cover classes are better explained than at the other two levels. This may be surprising because at the level of spatial scale the coefficients of determination at national level explained substantially more than at regional level, and they performed better at coarse resolution compared to fine resolution (Kok and Veldkamp 2001).

7.4.2 Variable importance

In Table 7-6, the three most important variables in terms of standardised betas are listed. These three variables account at least for 75% of the total explaining power.

If one looks in more detail at the A11 and A12 categories then it seems logical that at category level 'Cultivated areas' (A11) have a negative relationship with mountainous geomorphology, a positive relationship with the environmental region 'Continental' and a negative relationship with the presence of an impermeable layer within the soil profile. The relationships of the '(Semi-) natural vegetation' category (A12) seem complementary: a positive relationship with environmental region 'Pannonian', a positive relationship with the presence of an impermeable layer within the soil profile and a negative relationship with saline soils. Most plants do not grow (well) on saline soils.

At the domain level, a negative relationship exists with saline soils and flat geomorphology for 'Shrub crops' (SC) and 'Forests' (FO), two types of permanent vegetated areas. For 'Herbaceous crops' (HC) flat to hilly geomorphology is the main explaining variable, whereas for 'Grasslands' the relationship with flat geomorphology is negative, as well as with very mountainous geomorphology. A negative relationship with saline soils exists also for this class. At group level, both subclasses of SC, SCS and SCX, have the same three most explaining variables but in a slightly different order (number 2 and 3 reverse).

The subclasses of HC show a variety in the three most explaining factors and this is interesting as the subdivision of these classes is mainly on field size and irrigation. It is shown that with different field sizes the three most explaining variables vary. Areas with small-sized, medium-sized and large-sized fields (HCS, HCM and HCL) show that:

- Small-sized fields have a positive relationship with saline soils, a positive relationship with environmental region ‘Continental’ and a negative relationship with the count of months with an average temperature higher than 15 degrees. From these relationships one may infer that small-sized fields are found under less favourable conditions for cultivation.
- Medium-sized fields have a positive relationship with the count of months with an average temperature higher than 15 degrees, a negative relationship with both flat geomorphology and slope. So this type of fields occurs probably in a landscape with (slightly) rolling geomorphology where conditions for cultivation are more favourable.
- Large-sized fields have a positive relationship with the presence of peat in the soil, the only time that this variable occurs as most explaining factor, a negative relationship with environmental region ‘Alpine South’ and a positive relationship with rolling geomorphology. So these are the most favourable conditions for cultivation that occur in Romania.

The two classes with a mixture of field sizes, small to medium (HCY) and medium to large (HCZ), show also intermediate results because the first class shows two identical relationships (i.e. positive relationships with saline soils and with environmental region ‘Continental’), whereas the second shows only one identical relationship (a positive relationship with rolling geomorphology). The other two most explaining variables for HCZ show an inversed relationship compared to the pure medium-sized and large-sized classes: positive relationships with environmental region ‘Alpine South’ and flat geomorphology. So the mixture of medium to large-sized fields is found in a different landscape.

The irrigated class HCX with medium to large-sized fields is the only one that shows a positive relationship with the presence of an impermeable layer within the soil profile. This is not surprising if one considers that then water remains longer in the soil. Furthermore, it is the only HC subclass showing a relationship with the demographic variable potential population: irrigated cultivated areas are found in areas where people live that need to attend to the irrigation system. The negative relationship with height is what one would expect under irrigated conditions.

Table 7-6. The three most explaining factors in terms of standardised betas in order of importance per land-cover class at the distinguished semantic levels

Levels with land-cover classes			The most explaining factors in order of importance at each semantic level									R^2		
Cat.	Dom.	Grp	Category level			Domain level			Group level					
A11			GEOMORF04-	ENVMAP06+	IL-							0.878		
	SC					SALINITY-	ENVMAP05-	GEOMORF01-					0.749	
		SCS							ENVMAP05-	GEOMORF03+	GEOMORF01-			0.786
		SCX							ENVMAP05-	GEOMORF01-	GEOMORF03+			0.750
	HC					GEOMORF02+	GEOMORF01+	GEOMORF03+					0.884	
		HCS							SALINITY+	ENVMAP06+	T_PLUS15-			0.761
		HCM							T_PLUS15+	GEOMORF01-	SLOPE-			0.822
		HCX							IL+	POPPOT_LOG+	DEM-			0.987
		HCL							PEAT+	ENVMAP05-	GEOMORF02+			0.868
		HCY							SALINITY+	GEOMORF04-	ENVMAP06+			0.775
		HCZ							ENVMAP05+	GEOMORF01+	GEOMORF02+			0.850
A12			SALINITY-	ENVMAP08+	IL+							0.907		
	FO					SALINITY-	GEOMORF01-	ENVMAP08+					0.895	
		FOA							GEOMORF05-	SALINITY-	GEOMORF01-			0.892
		FOX							ENVMAP05-	GEOMORF05+	GEOMORF01-			0.651
	GL					GEOMORF05-	SALINITY-	GEOMORF01-					0.676	
		GLA							ENVMAP05-	IL+	ENVMAP06-			0.669
		GLX							GEOMORF01-	GEOMORF05-	SALINITY-			0.731
A23			DEM-	RAIN_WC_5M+								0.971		
A24			ENVMAP05+	SALINITY-	GEOMORF01+							0.968		
B15			T_PLUS15-	GEOMORF01-	POPPOT_LOG+							0.863		
	BU					T_PLUS15-	GEOMORF01-	SZ_LANDSC_RUR+					0.860	
	NB					No significant model								
B16			SALINITY-	T_PLUS15+	POPPOT_LOG-							0.735		
	CM					GEOMORF05+	POPPOT_LOG+						0.660	
	UM					No significant model								
	BX					SALINITY+	POPPOT_LOG-						0.756	
B27			ENVMAP08-	GEOMORF01-	SLOPE-							0.897		
B28			ENVMAP05+	ENVMAP08-	GEOMORF01-							0.967		

The relationships for the group-level classes of ‘Forests’ (FOA and FOX) and ‘Grasslands’ (GLA and GLX) demonstrate that:

- FOA has a negative relationship with very mountainous geomorphology, whereas FOX has a positive relationship. One might infer that mixed forests stands occur under more adverse forestry conditions.
- GLA has negative relationships with the two environmental regions ‘Continental’ and ‘Pannonian’, whereas GLX has negative relationships with flat and very mountainous geomorphology. The variable environmental region is more important for pure grasslands, whereas the variable geomorphology determines grasslands mixed with tree and/or shrubs.

For the non-vegetated categories ‘Artificial areas’ (B15) and ‘Bare areas’ (B16) it is as expected that demographic variables appear in each model: a positive relationship with ‘Built-up areas’ where the (rural) population lives and a negative one for mixed bare areas where people do not live.

The group of accessibility variables does not occur in the statistical models. The other four groups contribute in different manners to the total explaining power. These variables have been regrouped in Table 7-7 according to the count and percentage of the three most explaining factors. At the level of percentages one can compare between the three semantic levels. It is clear that biogeographical factors and geomorphology are the most important explaining variables in the statistical models.

Biogeographical factors are most important at the category level for explaining the occurrence of land cover with 47.8%. Geomorphology and soil are of equal importance at this level and demographic variables are least important. But at the domain level geomorphology variables are explaining land cover with 47.4%. The second most important variable group is soil, followed by biogeographical and demographic variables with equal importance. At the group level geomorphology is still the most important variable group explaining land cover somewhat less than at the domain level (41.7%), closely followed by biogeographical variables with 36.1%. Geomorphology and soil variables follow *ex aequo* and demographic variables are least important.

The effect of increasing semantic detail in the classes on the importance of the variable groups shows that there is a shift from biogeographical variables at category level to geomorphology at domain and group levels. Furthermore, at each semantic level it is always GEOMORF01 that occurs most, but never as first most explaining variable (with the exception of land-cover class GLX). SALINITY occurs less often than GEOMORF01, but it occurs more often as first or second most explaining variable. ENVMAP05 occurs especially at the group level as first or second most explaining variable.

Table 7-7. Variable importance in the regression models in view of the three most explaining variables

Variable group	LC Category		LC Domain		LC Group	
	Number	Percentage	Number	Percentage	Number	Percentage
Biogeographical:	11	47.8	3	15.8	13	36.1
ENVMAP05	2	8.7	1	5.3	6	16.7
ENVMAP06	1	4.3	-	-	3	8.3
ENVMAP08	3	13.0	1	5.3	-	-
DEM	1	4.3	-	-	1	2.8
SLOPE	1	4.3	-	-	1	2.8
T_PLUS15	2	8.7	1	5.3	2	5.6
RAIN_WC_5M	1	4.3	-	-	-	-
Demographic:	2	8.7	3	15.8	1	2.8
SZ_LANDSC_RUR	-	-	1	5.3	-	-
POPPOT_SUM	2	8.7	2	10.5	1	2.8
Geomorphology:	5	21.7	9	47.4	15	41.7
GEOMORF01	4	17.4	5	26.3	7	19.4
GEOMORF02	-	-	1	5.3	2	5.6
GEOMORF03	-	-	1	5.3	2	5.6
GEOMORF04	1	4.3	-	-	1	2.8
GEOMORF05	-	-	2	10.5	3	8.3
Soil:	5	21.7	4	21.1	7	19.4
IL	2	8.7	-	-	2	5.6
PEAT	-	-	-	-	1	2.8
SALINITY	3	13.0	4	21.1	4	11.1
TOTAL	23	100.0	19	100.0	36	100.0

7.4.3 Comparison with EURURALIS land-cover/use classes

The EURURALIS data set contains just one semantic level in which among the three most explaining factors in terms of standardised betas biogeographical variables occur nine times, geomorphology variables 10 times and soil factors four times (Table 7-8). Demographic factors do not occur as one of the three most explaining factors, as well as the variables SLOPE and PEAT. The available variables were identical to the ones used in the regression models with the LCCS data set. Also for the EURURALIS data set biogeographical and geomorphology factors are the most important explaining factors.

In Table 7-8 the order of the classes is following more or less that of the LCCS data set, so vegetated area classes are grouped followed by the non-vegetated area classes.

Table 7-8. The three most explaining factors in terms of standardised betas in order of importance per EURURALIS land-cover/use class

Code	EURURALIS land-cover/use class	The most explaining factors in order of importance			R^2
8	Permanent crops	No significant model			-
1	Arable lands (non-irrigated)	GEOMORF04-	GEOMORF03-	ENVMAP06+	0.894
6	Irrigated arable lands	GEOMORF01-	ENVMAP05-	GEOMORF02-	0.762
9	Forest	GEOMORF05+	SALINITY+	IL-	0.865
2	Pasture	ENVMAP08-	GEOMORF03+	T_PLUS15-	0.691
3	(Semi-) natural vegetation	SALINITY+	GEOMORF01-	GEOMORF04+	0.685
7	Recently abandoned farmland	SALINITY-	GEOMORF05-	GEOMORF01-	0.780
0	Built-up area	ENVMAP08-	ENVMAP06-	T_PLUS15-	0.909
4	Inland wetlands	DEM-	RAIN_WC_5M+	-	0.957
5	Glaciers, snow, sands and sparsely vegetated areas	No significant model			-

For most land-cover/use classes a statistically significant model is established with the exception of two classes, 'Permanent crops' and 'Glaciers, snow, sands and sparsely vegetated areas', with the highest coefficient of determination for 'Inland wetlands' and the lowest coefficient of determination for '(Semi-) natural vegetation'. Not only are the values different from the LCCS class set, the highest and lowest values also occur for other class types, though one can argue that the class FOX is part of '(Semi-) natural vegetation' but the latter exists in the LCCS data set at category level, i.e. A12, and then has a high coefficient of determination.

How do the classes in the EURURALIS data set compare to the LCCS data set? The class 'Inland wetlands' seems to be fully compatible with A23 and shows exactly the same two most explaining factors although the coefficient of determination is lower (0.957 versus 0.971). The class 'Irrigated arable lands' should be compatible with HCX as these represent irrigated crops in both data sets but not only the three most explaining factors are very different, so is the coefficient of determination (0.762 versus 0.987). The class HCX stands out among the 'Herbaceous crops' (HC) as a class with a different set of most explaining variables, this is not true for 'Irrigated arable lands' that has the same type of explaining variables as non-irrigated classes in the EURURALIS data set. For the class 'Arable lands (non-irrigated)' the type of explaining variables is the same as for the different classes in the LCCS data set but the combination of the three most explaining factors is different. The class 'Forest' contains as most explaining factors very mountainous geomorphology and saline soils like the class FOA but the relationship here is positive. The class 'Pasture' shows a set of different explaining factors compared to the class GL, but resembling in having both low coefficients of determination. The relationship between grassland *cover* and pasture *use* shows high variability (Bakker and Veldkamp 2008). The class '(Semi-) natural vegetation' has a positive relationship with saline soils; this is striking because in the LCCS data set this relationship, when occurring in any of the A12 classes, is always negative. The class 'Recently abandoned farmland' has the same set of three most explaining variables as the class GLX but in reversed order and with a higher coefficient of determination (0.780 versus 0.731). The class 'Built-up area' is compatible with BU but only the occurrence of the variable T_PLUS15 is the same. The coefficient of determination is higher.

Thus with different semantics, but an identical set of variables, the three most explaining factors in terms of standardised betas are different. Not only factors differ, sometimes the relationship with a factor changes from positive into negative or *vice versa*. The latter seems illogical unless one considers that the semantic contents of classes are much more different than their names seem to suggest. The characteristics of the larger component (category level) are not simple combinations of attributes of smaller components (domain and group levels).

7.5 Consequences of organisational hierarchy on modelling dynamics

Different organisational hierarchies lead to different semantic contents of the grid of 1x1 km used in the regression analyses. Moreover, this variation in semantic contents leads to different sets of spatial determinants. Future policy and decision making depend to a great extent on which organisational hierarchy is present in the data set used to formulate a policy or to make an informed decision. Policy makers try to develop 'good' land governance but often the implemented regulation leads to adverse effects. One of the reasons for this scaling and governance problem is the scale mismatch when the units of regulation do not match the functional units where the process operates (Veldkamp 2009). The scales over which processes operate do not necessarily correspond to the spatial extent of the observation of such processes (Pereira 2002).

As the focus is on the rural areas in Romania, where cultivated (35.74% of total area) and (semi-) natural vegetated areas (34.08%) form the two most important spatially explicit categories, a more in-depth examination of these two categories is made. In the rural areas the consequences of land reform and the transition to a market economy are quite different compared to urban areas (Van Meurs 1999):

- In urban areas, privatisation and the liberalisation of the economy mean business opportunities and new patterns of ownership.
- In rural areas privatisation generally meant the restitution of property to former owners or their heirs. Cultivated land, the primary means of production, is a fixed commodity. Thus, future patterns of ownership seem to be predetermined.

Romania has the smallest farms among the new European Union Member States: an average farm size of just over two hectares, with a share of cultivated land in farms below five hectares of 58% (Vidican 2008). An improvement of the agricultural production would make a considerable contribution to the welfare of the rural population as agricultural products comprise one of the few genuine export options in a competitive world market (Van Meurs 1999).

From the LCCS data set it should be clear that the category and domain levels lump together all types of farms or farmers producing crops, whereas at group level it becomes clear that there are different explaining factors for: (1) shrub and herbaceous crops, and (2) within the classes of herbaceous crops depending on field size or irrigation. These different field sizes occur in different parts of the landscape.

The diversity in the relationships between the spatial determinants of herbaceous crops with field sizes or irrigation seem to suggest that further investigation into these field sizes might reveal that they are related to different farm(er) types. Bakker and Van Doorn (2009) have shown that the relationships between landscape factors and land changes are different for each farmer type. They also showed that the relationship between a landscape characteristic and the probability[38] of land change is significantly different for the different farmer types. So farmer types are a key issue in understanding land change as they reflect the heterogeneity of human behaviour and decisions. Particularly in the context of Romania where limited landowner experience, restricted economic and technical resources and the socio-economic context in which these human behaviour and decisions occur are all factors contributing to this diversity. A relevant approach to analyse such heterogeneity in farmers' decision making is the use of typology (Valbuena *et al.* 2008) and to distribute the defined farmer types spatially over the country.

Distribution of farmer types may make (more) explicit land fragmentation from the point of land use and not landownership. This is important in Romania where selling or leasing of land is possible according to the law but where the choice, in terms of land re-allocation, for the existing farmer association often prevails before participation in land transactions (Vidican 2009). Small farms may make sense in some labour-abundant agricultural economies in the short run; in the longer run, the transition to a modern state means that farm size must be sufficiently large (Rozelle and Swinnen 2004). The farmer associations provide security of tenure and capital access, allowing landowners to draw on the benefits of economies of scale. Because economic diversification is lacking in the rural areas, few of the landowners are willing, or able, to leave agriculture. As Vidican (2009) points out: leasing-out is a viable alternative for younger

[38] The classical definition of probability is identified with the works of Pierre-Simon Laplace. As stated in his book *Théorie analytique des probabilité* (1912), the probability of an event is the ratio of the number of cases favourable to it, to the number of all cases possible when nothing leads us to expect that any one of these cases should occur more than any other, which renders them, for us, equally possible.

landowners who can engage in non-farming activities, as well as for (older) landowners that are resource constrained (e.g., physical, financial and human capital) for working the land themselves.

From the LCCS data set it is clear that the three most important spatial determinants for shrub crops at group level are identical, only their order slightly differs. So in this case the organisational hierarchy at domain level may suffice. Land with permanent crops forms an important role in the landscape and in the perception of farmers. In a way they form a kind of capital that is not abundant (3.32% of total area) and generally in areas with more difficult access. According to Fraser and Stringer (2009) such stocks are usually better taken care of than easily accessible, abundant stocks.

For the (semi-) natural areas in the LCCS data set the spatial determinants at group level show a diversity greater than the determinants at domain level (two are identical). The subdivision in the two forest and two grassland types seems, thus, sensible. The abundance of forests in the rural areas may be perceived by the rural population as a source of capital in times of need. After the state-owned enterprises were broken up, people returned to the rural areas. As a consequence the demands for fuel wood, wood for construction purposes and cooking increased. Short-lived security may be brought by tree logging (pers. comm. Alexandru Badea, CRUTA) with detrimental environmental effects (e.g., land degradation and erosion) similar to the behaviour observed in Albania (Jansen *et al.* 2006a; Chapter 5). Thus, both the abundance of forests and their perception by the rural population are key issues in future policy and decision-making. The two grassland types (16.03% of total area) are used as either pastures or meadows. In the transition period the forests were also heavily grazed in the absence of designated pastures (Fraser and Stringer 2009). Hence policy and decision making concerning grasslands does also affect the forested areas.

For modelling dynamics the LCCS class set renders a model for every cultivated area and (semi-) natural vegetated area land-cover type with coefficients of determination that are higher than for the EURURALIS class set. The EURURALIS class set has less clear distinctions in terms of semantics although the main land-cover types are included. This class set does not, however, make any distinction related to field size. In the case of Romania this is an important omission.

Inclusion of farmers' perception in the change dynamics modelling is important as their decisions are not always linked to what spatially or landscape-wise would be sensible. In agriculture the 'portfolio theory', usually applied in financial investment, suggests that under uncertain conditions land users will spread their risks through diversification. This has not been the case in Romania. High levels of natural capital (e.g., fertile soils, forests and natural grazing areas) combined with limited amounts of off-farm employment (non-agrarian livelihoods are few in the rural areas) and a high degree of socio-economic uncertainty give ample cause for concern since the existing land-use systems of the semi-subsistence holdings are vulnerable.

7.6 Discussion

The overall results of the regression analysis are satisfactory (the domain level having the lowest values with an average R^2>0.783). Coefficients of determination are (very) high, although some statistical models yield a relatively low coefficient. Remarkable is that these are classes known in remote sensing based data-collection efforts with validation as classes having lower levels of confidence (Scepan 1999; Loveland *et al*. 2000; Jansen *et al.* 2003a and 2006b). Confusion between assigning certain pixels to certain classes implies uncertainty in the data that seems to propagate in the statistical models.

The LCCS and EURURALIS data sets are based on interpretation of different satellite images (LANDSAT TM versus SPOT) by different persons and although there may be agreement on the type of classes identified, this does not necessarily mean that these classes have been identified at the same location (Hansen and Reed 2000). Moreover, these thematic data sets are only a model, or simplification, and hence a flawed representation of reality (Foody 2001). As data are abstracted from their 'raw' form (e.g., a remote sensing image comprising pixels) to the higher representations used by GIS (spatial land-cover object), it passes through a number of different conceptual models via a series of transformations. Each model and each transformation process contributes to the overall uncertainty present within the data. There is, as a result, a continuum of abstraction (Gahegan and Ehlers 2000). Categorisation is only one type of abstraction in the whole process. But the interest in data accuracy should not only comprise the spatial and temporal viewpoint, it should also focus on the semantic perspective.

The approach used emphasizes land-intensive activities related to land-cover patterns and underlines an implicit assumption that these are the only activities of economic or other importance. This presumption is neither generally valid nor justified because less land-intensive activities (e.g., services) or linear activities (e.g., transportation) may be important determinants of spatial development (Briassoulis 2008). Furthermore, the analysis is based upon 1x1km pixels. This is a unit of measurement that does not correspond to environmental (e.g., landscape, watershed) or human decision units (e.g., household).

The country level analysis assures the same level of administration at the three semantic levels. This is the level at which most policies apply and consequently the most appropriate level for analysis and modelling. One should note, however, that levels of administration are not based on scale but on definition (e.g., formal definitions of national, regional and local levels).

The spatial determinants included in the statistical models for the LCCS class set have the expected sign. This strongly suggests that for most land-cover types the set of variables used includes the most important spatial determinants of land cover in Romania. Biogeographical variables are most important at category level, whereas geomorphology is most important at domain and group levels. Soil and demographic variables are less important, whereas accessibility factors did not occur in any of the statistical models but they might become more important at other organisational hierarchies. Thus, biogeographical factors, geomorphology, soil variables and demographic factors should be included in any analysis of the distribution of land-cover types at country level independent of the semantic contents of the data. One should note, though, that these biophysical drivers of change are factors that land-use policy and planning *cannot* influence.

The three most explaining variables are not always found at the semantic organisational level one would expect. A variable such as environmental region (ENVMAP) would be expected at the most aggregated level, where it occurs, but it is more markedly present at the most detailed semantic level examined. Geomorphology would probably be expected at the most aggregated semantic level, where it is present, but it becomes more pronounced with each semantic level. Soil variables, like impermeable layer and saline soils, would be mainly expected at the most detailed semantic level examined but they are clearly present at the least detailed semantic level. This may be related to the pattern of soil variability in Romania. So one does not always find the explaining variable at the semantic level expected. Similar results have been found for explaining factors and scale (pers. comm. Kasper Kok, WUR Land-Dynamics Group).

The objectives of the LCCS and EURURALIS data sets were different: in the first data set the agricultural and (semi-) natural areas have been subdivided according to type of plants (permanent versus temporary) and in the cultivated areas field size and irrigation were used as criteria to create homogeneous types of classes. In the second data set more or less the same parameters (permanent versus temporary plants, irrigation) were used but at the same time an effort was made to preserve heterogeneous agricultural areas in the classes. This means that the semantics are very different and consequently the regression models. That the classes are more heterogeneous in the EURURALIS class set leads to overall lower coefficients of determination for the statistical models. Homogeneous type of classes may shed light on which driving variables best explain a certain land-cover, whereas heterogeneous classes may explain what are the driving variables in a heterogeneous landscape. Which approach to take depends on the objective of the research.

7.7 Conclusions

If one wishes to define and explain the interaction of human-environment systems, understanding the scale of interaction and the scale of different environmental and social processes is of paramount importance (Engel-Di Mauro 2009). Changing *resolution* did not greatly influence the analysis of explaining factors of land cover in Central America (Kok and Veldkamp 2001), though it did in Ecuador (De Koning *et al*. 1998) and Java (Verburg *et al*. 1999). However, the effect of changing *extent* from national to supra-national level was substantial (Kok and Veldkamp 2001). Complementary to these analyses of the spatial dimension of scale and the temporal dimension (Liu and Anderson 2004; Bakker and Van Doorn 2009), is the *variation in semantic contents*, the third often forgotten dimension of scale.

The statistical models at three semantic contents levels demonstrate that the spatial determinants vary with different organisational hierarchy. Semantic contents play an important role in land-change analysis. They play different roles at different levels of organisation, as well as that different class sets lead to different statistical models in regression analysis using an identical set of variables. These models in turn will lead to different outcomes in modelling exercises. There is no pattern recognizable when establishing the organisational hierarchy. The coefficients of determination do not explain more at one semantic level than another; the explaining power varies per group of classes and the classes within the group. Thus, when establishing the organisational hierarchy in a class set it is important to consider interactions between the land-cover types. These interactions are known to operate over different spatial and temporal scales. Based upon the results here presented, the semantic aspect of scale should be added leading to multi-scale and multi-semantic analyses.

Another important aspect is uncertainty propagation. Certain classes in the class set are known to be established with more confusion when using remote sensing than other classes. The uncertainty between especially more open tree-dominated vegetation types with thickets and shrublands propagates in the establishment of statistical significant models with consequences for the outcomes in (spatially explicit) models.

The conclusion can be drawn from the case study in Romania that the semantic dimension does play an underestimated role in land-change dynamics next to the spatial and temporal dimensions. If the same conclusion would be found when using different case studies in other areas with other models then this conclusion could be extended to have a more general relevance. If such were the case it would mean that before a realistic simulation can be made, a thorough analysis of the effects of variation in semantic contents on the predictions of spatially explicit land changes needs to be included.

8 Synthesis

The distinction between past, present and future is only an illusion, however persistent.

Albert Einstein (1879-1955)

8.1 Methodological issues

The focus of this dissertation has been on methodology: the parameterised approach to categorisation to create multi-level class sets for two subjects, land cover and land use, and the use of such class sets in harmonisation efforts, land-change analysis and modelling. The terms ‘land use’ and ‘land cover’ are often used interchangeably but there is a relationship: land cover can be a cause, constraint or consequence of land use (Cihlar and Jansen 2001). Land cover and land use almost never match one to one and consequently data analysis almost always results in scattered non-linear relationships (Bakker and Veldkamp 2008).

Two crosscutting activities of land-change research, i.e. *categorisation and data* and *scalar dynamics*, were important in the context of this thesis. The first examines data availability, data quality and categorisation structure, whereas the second recognizes that land-change patterns observed at any spatio-temporal scale are caused by complex synergy with changes observed at other analytical scales. The variation in the semantic contents of data, expressed as differences in categorisation, can be regarded as the joint result of *categorisation and data* and *scalar dynamics*.

8.1.1 Categorisation and data

To classify is human as Bowker and Star (1999) stated. Few categorisations take formal shape or any formal algorithm, even fewer categorisations are standardised. Yet, we all use (in)formal categorisations on a daily basis, intentionally or inadvertently. The knowledge about which categorisation will be useful under certain conditions and at a given moment is embodied in our responsibilities and routines in a certain context. At the level of policy, categorisation of areas, uses and covers plays an equally important role. The categorisation of an area as either nature reserve or industrial will have a clear impact on future economic decisions. Thus the relation between categorisation and decision-making may be invisible but is evidently powerful. Nowadays in the information era, scientists work on the design, description and choice of categorisation systems embodying choices that create people's identities. But few people realise how much impact a categorisation may have. In the context of land cover, in Europe the CORINE Land Cover contributed in creating a European identity; the LCCS may contribute in creating a UN identity.

Categorisations embody a worldview and each category and class in it values this specific viewpoint. This is in itself not critical as long as it is recognized that another viewpoint may be silenced. From the analysis of semantic information and used definitions one can deduce something about this view and the intent of the data producers, but much more transparency is needed. And not only the latter because what are truly needed are more insights in the design of categorisation systems and research examining their impact. The effort of attaching objects to categories and the ways in which those categories are ordered into systems is often disregarded. In the land-cover domain, for instance, several class definitions in CORINE Land Cover (CEC 1999; Bossard *et al.* 2000) or in LCCS (FAO 2005) are described by taking a bird's eye view, or *map* view, rather than a *geographic entity* view probably because these systems are used in remote sensing.

Categorisation facilitates the communication of knowledge concerning specific phenomena (e.g., land use and land cover) between individuals. Ideally categorisations are able to travel across the borders of (scientific) communities, of which the individuals are part, and maintain some sort of constant identity. Categorisations can be tailored to meet the needs of any one community, though having, at the same time, common identities across settings. In order to represent multiple constituencies, categorisations incorporate ambiguity, i.e. leaving terms open for multiple meanings across different worlds. Categorisations are thus inherently vague, ambiguous and constant. Communication is interesting in that it, generally speaking, must reside in more than one context. Dante Alighieri wrote in 1320 that his work *Divina Commedia* is 'polysemantic', that is of many senses; the first sense in his *Divina Commedia* is that which comes from the letter, the second is that which is signified by the letter. These multiple interpretations are primary, not accidental nor incidental.

The tangible results of categorisation are *classes* and *categories* that serve as the vehicles for communication of meaning (Ahlqvist 2008b). *Parameters* used in the categorisation are usually not tangible, simply because they remain more often than not unmentioned. The members in each class or category have definable characteristics in common and with the use of categorisation one can discover general truths related to the distinguished classes or categories rather than to their individual members (Shapiro 1959). Categorisation is, at the same time, a simplification because it represents only part of the complexity of reality (like models represent simplifications of the real world). Different perspectives, or so-called 'scapes', to categorisation can be taken that are all equally valid and valuable (Veldkamp 2009; paragraph 0) One needs to recognise, therefore, that no categorisation reflects accurately the social or the natural world (Bowker and

Star 1999). Categorisations arise out of social communication needs but they serve specific purposes: not only do they reflect the ideas of a certain community or institution, but they can also be the end-result of negotiating and reconciling individual, group and institutional differences (Ahlqvist 2008b).

Definitions are the main, and usually the only, descriptions of categories and classes, since other elements that could contribute to the semantic definition of categories (e.g., the parameters or criteria used) are often absent (Chapter 2). Rich narratives are needed that further specify and clarify what is included in a parameter, a class or category because anyone using a parameter, class or category will have to interpret their semantic definition and may therefore introduce bias. A parameter, class or category needs to be understood in a similar manner by the data producer (the generator of data sets), the distributor (the subsequent distribution of data sets) and user (in the end the user of data sets).

Definitions expressed in natural language associated by sub-type/super-type relationships, i.e. hierarchical relationships, are called *terminological ontologies* (Sowa 2000). Almost all land-use and land-cover categorisations to date are terminological ontologies (e.g., CORINE Land Cover and LCCS). Ontology is an explicit specification of a conceptualisation to represent shared knowledge (Gruber 1993; Ahlqvist 2008b). Semantic information can be determined from the definitions of the ontology and the representation of categories can be enriched with semantic properties (e.g., purpose, time, location, etc.) and relations (e.g., "is-a", "is-a-part-of", "associated-with", etc.) in order to reveal similarities and heterogeneities (Kavouras *et al*. 2005). Recognition of semantic heterogeneity is the basis for creating sound data linkages between multiple data sets that are needed for land-change analysis, monitoring and modelling for land-use planning, policy and informed decision-making (Chapters 3 and 4).

Especially in change analysis, monitoring and modelling, semantics often form a problem due to the limited description of how exactly class labels should be understood (Comber *et al.* 2004a) and expert opinions by definition differ (Comber *et al*. 2005a). Moreover, data sets from the same area but from different times often need to be integrated in a geo-database while at the same time each is based upon a (slightly) different categorisation (Comber *et al*. 2004b). Similarity in terms does not necessarily imply equivalent category terms.

Land cover and land use are socially mediated constructs (Comber *et al.* 2007; Ahlqvist 2008b), as described above, without agreed fundamental units. In fact various types of units of measurement are possible (see paragraph 4.3). Categorisations are used for communication of knowledge by being dynamic though ordered structures immersed, at the same time, with vagueness and ambiguity (Chapter 2). Operationally, though, categorisation often makes a straightforward unproblematic leap from concept to class, eliminating any traces of concept ambiguity by stating mutually exclusive and crisp classes (Ahlqvist 2008b). The latter is certainly true for CORINE Land Cover and LCCS.

As Cihlar and Jansen (2001), Comber *et al*. (2005b and 2007) and Ahlqvist (2008b) point out: manifold ways to conceptualise and communicate knowledge exist according to the disciplines of (groups of) experts, professions, etc., so that there are necessarily many-to-many relationships between classes and thus inherent ambiguity in any categorisation. Categorisations contribute to communication of knowledge and in making joint progress in that knowledge by facilitating communication. However, they can only make such contributions by being *dynamic* in nature. By keeping the voices of parameters and their constituents present, as is the case in parameterised categorisations, the maximum flexibility of the system is retained. This includes the key ability to be able to change with changing knowledge, technological developments and changing policy objectives.

Collection of *data* leads to the creation of categories. Contrary to old hierarchical class and data sets (or databases), where relations had to be decided once for all the time of original creation, many class and data sets today incorporate object-oriented views whereby different parameters can be selected and combined on the fly for different purposes (Bowker and Star 1999). Parametric Object-Oriented Data Models (POODM) should take the place of old-fashioned categorisation systems like CORINE Land Cover and LCCS because they allow an unprecedented flexibility and capability in the design and use of very complex information systems and land change requires such an information system. Such a POODM should use the Unified Modelling Language (UML) as standard, something both CORINE Land Cover and LCCS miss and thereby neglect fulfilling the ISO 19100 standard. These parameterised multi-level class and data sets put more emphasis on the parameters to be used than on the structure in which these are organised. This approach is dynamic, easily adaptable under changing circumstances.

With the choice of categories, data quality becomes an issue. Certain classes in the class set are known to be established with more confusion than other classes (Chapter 7). Especially the confusion between more open tree-dominated vegetation types with thickets and shrublands when using remote sensing (Jansen *et al.* 2003a; Jansen *et al.* 2006b) or between spatial and census data for grasslands, shrubs and woodlands classes (Pelorosso *et al*. 2009). Such uncertainty in the categories propagates in the establishment of statistical significant models with consequences for the outcomes in (spatially explicit) models (see also paragraph 8.1.4).

With the progress in computer and information sciences there seems to be a real need to improve existing categorisation concepts (e.g., abandonment of mutually exclusive and crisp classes in the standard set theory in favour of fuzzy set theory as will be explained in paragraph 8.1.3) and the operational use of such categorisations, and to accept that categorisations make a contribution to communication of knowledge by being *dynamic* in nature. This is particularly true for land-change analysis, monitoring or modelling efforts that are inherently dynamic. These aspects will be further elaborated in paragraphs 8.1.2 to 8.1.4. But whatever categorisation or data model -the link between reality and the database- will be adhered to in future, it should link to standard software applications and not require the use of non-standard software.

Based on the fact that categorisations are dynamic in nature one could argue that the definitions of Shapiro (1959), Sokal (1974) (discussed in paragraph 2.1), and FAO (2005) (discussed in paragraph 2.5.1) or even the definition applied in the documents of LCCS[39] submitted to ISO/TC 211 should be abandoned in favour of a modified version of the definition of Bowker and Star (1999): *categorisation is a spatial, temporal, or spatio-temporal, and organisational hierarchy based segmentation of the world*. This definition emphasizes that the dimensions of time and space are imperative in determining a categorisation, as well as the organisational level. In the case of a non-hierarchical system one could speak of zero organisational hierarchy, analogue to zero tillage when no tillage occurs.

8.1.2 Harmonisation

Categorisation and standards are closely related, but not identical. Standards are a way of categorising the world with a set of agreed-upon principles, spanning more than one community of practice, persisting over time, making something function over distance and heterogeneous measurements, and often enforced by a legal body. Categorisations may or may not become standardized; if they do not, they are *ad hoc*, limited to an individual or a local community, and/or of limited duration (Bowker and Star 1999).

A categorisation is also a means for *data standardisation* (for new data sets) and *data harmonisation* (correspondence between existing data sets). Standardisation takes categorisation one-step further in that it fixes a categorisation (Ahlqvist 2008b). The development of accepted standards in science ensures repeatable experiments, exchange of findings, etc., and thus a standard can act as a ‘common language’. However, standardisation assumes that the advances in knowledge, technological developments and changing policy objectives will not have an impact on the existing systematic categorisation framework (Chapter 3). A major drawback is that firm establishment of a categorisation system runs the risk of becoming stale and out of phase with contemporary thinking. Therefore standards cannot represent the

[39] LCCS has been submitted to ISO in two parts (with ISO numbers 19144-1 and 19144-2 respectively): (1) "Classification Systems - Part 1, Classification system structure" is a generic standard for classification systems in general; and (2) "Classification Systems - Part 2, UN FAO - Land Cover Classification System (LCCS) - Conceptual Basis and Registration of Classifiers" is a specific standard for LCCS (http://www.glcn.org/act_7_en.jsp).

depth of knowledge held within a community (Ahlqvist 2008b; Comber *et al*. 2005b and 2007). The above means that standards also need to change over time, i.e. they need to be *dynamic* in nature, so they are not as 'standard' as their name seems to suggest. The problem is that many standards have significant inertia; it is difficult and expensive to change (e.g., the major revisions made in ISIC in 1958, 1968, 1989 and 2008).

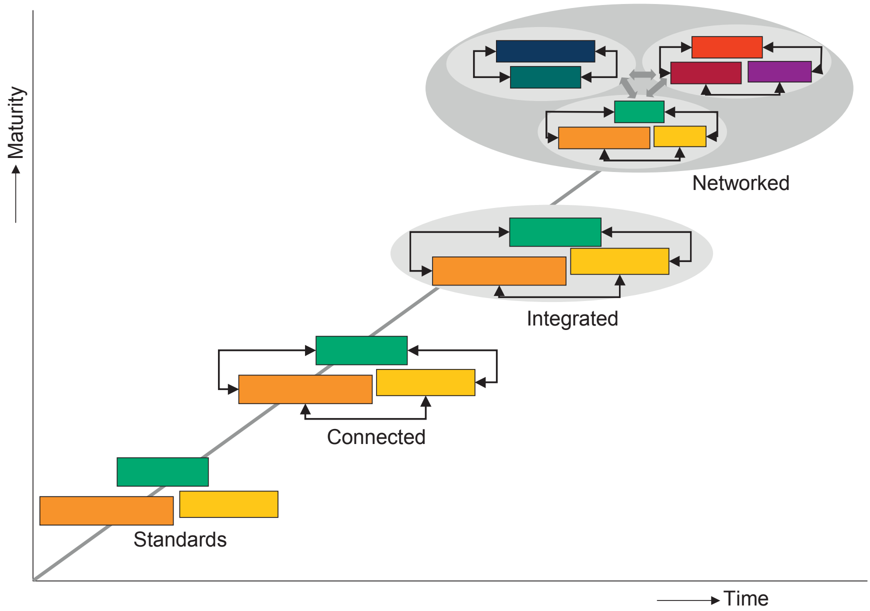
The International Organization for Standards Technical Committee 211 (ISO TC/211) (www.iso.org) and the Open Geodata Interoperability Specification Consortium (www.opengis.org) are two sources that have developed numerous standards concerning data interoperability between the data producer (the generator of data sets), the distributor (the subsequent distribution of data sets) and user (in the end the user of data sets). Those standards concerned mainly the development of exchange formats, data projection, spatial reference systems and measurement units. Thus, they were mainly concerned with specifying the *syntactic* and *schematic* aspects of interoperability (Ahlqvist 2008b). The efforts to create operational Spatial Data Infrastructures (SDI) have to deal also with the *semantic* dimensions of data. These dimensions have become apparent to the communities producing and using satellite-based land-cover and/or land-use data (Comber *et al* 2005b; Jansen *et al.* 2008), the data modelling community (Bishr 1998) and efforts based on ontology (Gruber 1993). As Ahlqvist (2008b) shows, ontology development and integration can be seen as a formal parallel to the social categorisation process.

Developed standards have to go through an adoption process. This can be compared to the five-step process of the diffusion of an innovation. This process is a type of decision-making. It occurs over a period of time through different forms of communication among members of a similar social system. The five steps distinguished are: (1) knowledge, (2) persuasion, (3) decision (adoption or rejection), (4) implementation, and (5) confirmation (Rogers 1983). Evidence of the 'knowledge' and 'decision' steps exists, but the evidence for the other stages is much less clear (Van den Ban and Hawkins 1985). Once the innovation is spread in the social system it will go from one decision-making unit to the next over time (e.g., individual, household, collective)(Röling 1988). In the case of categorisations dealing with land change, the implementation of a standard requires considerable additional learning and decision making on how to use this standard most effectively. In this and many other cases, one is not dealing with the adoption of one innovation, but with a whole package of innovations (Van den Ban and Hawkins 1985). Often innovations have to be adapted to the specific situation in which they will be used (e.g., CORINE Land Cover 1990 was adapted to be applied in more countries in 2000).

The expression "union in diversity" is a slogan of the European Union that can be used to demonstrate why especially harmonisation and not standardisation is needed. Unity in diversity is necessary and can be synthesised in two words: *complementarity* and *interdependence* (Banini 2006). Local systems are inclined to specialise their functions, and therefore unavoidably, they become more and more dependent on each other. Local systems, moreover, are nodes of global networks, forming part of the same global system interconnected by material and immaterial networks (Dematteis 2002). Application of a singled out worldview, that would be the case if a single categorisation standard were adopted worldwide, would hinder the scientific community from looking beyond. Many visions of interpreting reality exist, all of which are equally valid and necessary. A world in which several worldviews co-exist seems not only a much more flexible approach suited to the intended purpose of data collection, but also one that makes the world richer. Harmonisation of categorisation systems, which can be considered a way of data modelling when considering harmonisation of data from different sources and different formats, seems thus a more realistic avenue for future applications as the emphasis is shifting from *static* mapping towards more *dynamic* monitoring and modelling (Chapters 3 and 4). Investing in harmonisation would thus enhance communication between systems.

The concept of levels of maturity is used to show the importance of harmonisation (Figure 8-1). These levels of maturity are based upon the ‘Stages of Growth Model’ (Nolan 1979) and the application of this model to land administration (Van Oosterom *et al*. 2010). Every step in the model provides higher value and efficiency. None of the previous levels can be omitted because the subsequent level builds on the previous one. The ultimate step towards the highest level involves an important mind shift: it will place the categorisation in the context of current relevant global themes such as global environmental change, climate change, public safety, poverty reduction and food security. The first level comprises multiple standards (level 1) since through these different institutions can make a connection to exchange land-cover or land-use (change) information (level 2). Examples of the first level comprise CORINE Land Cover in the EU Member States and LCCS for the UN agencies. The INSPIRE Directive is an example of how the national level can connect with other national levels in the EU. After institutions or countries are connected they start acting as a unity (level 3). The ultimate level means a shift because no longer the organisation or country is in the central position but current global themes. This means that experts from the social and biophysical disciplines collaborating in global environmental change should understand each other despite their different disciplines and terminology. For this collaboration to be fruitful semantic harmonisation will be required.

Figure 8-1. Categorisation levels of maturity



The standardisation efforts by FAO and partners to formalise LCCS as an ISO standard may be related to their need of a body, i.e. the ISO, to enforce their system because without a mechanism of enforcement (top-down approach), or a grassroots movement (bottom-up approach), their system will fail (unlike CORINE Land Cover that is enforced by the European Environment Agency (EEA)[40]). FAO's standardisation efforts show clearly that they are tempted to believe that sharing their worldview and categorisation aims can be pursued inside progressively wider organisational demarcations until the whole globe is included. They consider that the smaller the scale, the more universal the categorisation worldview and aims are. But as Banini (2006) shows diversity, interconnection and complementarity are inside local categorisation systems. This means that local systems can act mainly as nodes of local networks, while other systems, also interacting locally, participate in the values and interests of various systems. The local categorisation system should be aware of: (1) its own specifics, as well as those of other systems; and (2) the fact that all these specifics are complementary and interconnected, in a mosaic of diversity that constitutes a global level. Thus, categorisation systems should not only be viewed from a *multi-scalar* perspective but also a *trans-scalar* one: categorisations should look beyond the local worldview and through other dimensions. The latter is a multi-perspective that is also called the different 'scapes' perspective (Veldkamp 2009).

[40] The EEA has currently 32 member countries (the 27 EU Member States together with Iceland, Liechtenstein, Norway, Switzerland and Turkey) and six co-operating countries (Albania, Bosnia and Herzegovina, Croatia, the former Yugoslav Republic of Macedonia, Montenegro and Serbia).

From the above it should be clear that the formalisation of any land-use or land-cover categorisation does not represent scientific progress. Especially not, if any such proposed system would already lag behind scientific developments. Nowadays Parametric Object-Oriented Data Models (POODM) allow an unprecedented flexibility and capability in the design of very complex information systems. There is also no natural law that the best standard wins; they do so for a variety of reasons hardly ever linked to merit. Furthermore, formalisation of any system would withdraw the attention from harmonisation that would allow the continued existence of many categorisations with their complementarity and interdependence. The existence of different categorisations enriches the understanding of our environment by taking different perspectives (Jansen 2009). These different perspectives -that include different choices of scale- reflect different reasons for analysing, and can provide equally valid, but non-equivalent descriptions of, the same system. It will be necessary, in many cases, to adopt more than a single perspective to reflect both the general complexity of the issue and the different perspectives of diverse stakeholders (Rothman 2002).

8.1.3 Land-change analysis

Land-change patterns observed at any spatio-temporal scale are caused by complex synergy with changes observed at other analytical scales. The manner in which land changes are recognized is based on a mathematical theory underlying a categorisation. However, this fact is usually neglected. Considering widely applied categorisations such as CORINE Land Cover and LCCS one may well wonder what mathematical theory underlies these systems (3.5.1).

If one looks at change analysis and monitoring from the semantic perspective then one can observe that it is often performed in a rather straightforward manner by constructing a change matrix for spatially explicit evaluation of changes. This approach is based on *standard set* theory in which the crisp class A has either changed in another crisp class or crisp class A remained unchanged. Changes of crisp class A into crisp class B or into crisp class C are treated in an identical manner though one change type may relate to a *conversion* and the other to a *modification*. A conversion means large semantic differences between classes (e.g., change from pasture into residential area), whereas modification means small semantic differences (e.g., change from low-density residential area into a high-density residential area). The land-cover change analysis in Albania using LCCS is an example of this approach (Chapter 5). A more detailed approach has been used in the EU PHARE Land-Use Policy II project in Albania (Chapters 4 and 6) where not only land-use changes were identified as either conversions between land-use categories or modifications within a land-use category but the categorisation hierarchy was used to distinguish for each type of modification three levels of intensity. It is surprising that the hierarchy of class sets, often carefully constructed, is hardly ever used in the analysis of either changes or correspondence between classes.

Another approach to change analysis uses *probabilistic reasoning* (Haenni 2005) instead of standard set theory though this approach considers the classes still as being crisp and unambiguous. A more sophisticated approach is to consider the notion of vagueness in the categorisation system using *fuzzy set* theory. The notion of category semantics and category similarity metrics (e.g., overlap and distance) is concerned with the vagueness inherent in category definitions and semantic relations between categories (Ahlqvist 2008b), thereby overcoming the traditional limitations on the exhaustiveness and mutually exclusivity of classes (Rocchini and Ricotta 2007).

The interpretation of a change matrix under the assumption of fuzzy categories will differ from the standard one where diagonal elements hold instances of 'no change' and off-diagonal elements hold instances of category gains and losses (Fisher *et al*. 2006; Pontius and Cheuk 2006). The diagonal can no longer be treated as holding instances of 'no change' and the use of category semantics and category similarity metrics should be considered. Ahlqvist (2007) shows that with such an approach not only a spatially explicit evaluation of changes can be given, but also a nuanced assessment on changes of heterogeneous class types.

It seems that many recent developments in the field of semantic similarity metrics and in the theory underlying a categorisation have been overlooked in systems such as CORINE Land Cover and LCCS. As Ahlqvist (2004) rightly point out LCCS uses standard set theoretic representations without recognizing a semantic *space* underlying the concept representation, thus limiting the possibilities to measure in LCCS semantic similarity based on concept distance. Also the interpretation of a change matrix under the assumption of a sophisticated mathematical theory such as fuzzy categories would provide a more nuanced interpretation of land change. The use of advanced mathematical theories in categorisations has become nearly compulsory to make scientific progress in the understanding of land change.

Three dimensions of scale can be distinguished (Wu and Li 2006): (1) space, (2) time, and (3) organisational levels, or organisational hierarchy, as constructed by the observer. Space and time are the most obvious and the most frequently used dimensions in land-change analysis. Organisational hierarchy as constructed by the observer is the less obvious dimension and the one that is most often taken for granted. Multi-level class sets are constructed with great care but the same care can usually not be observed in the analysis phase:

- When harmonisation is concerned class correspondence is often established with total neglect of the organisational hierarchy in the original class set (see paragraph 8.1.2).
- In assessing semantic similarity the organisational hierarchy could be used when *depth* and *density* of the classes in the hierarchical categorisation system would be calculated (paragraph 3.6.2). Feng and Flewelling (2004) suggested that the number of links coming out of a category could be used as

an estimate for density and the number of levels down in the categorisation as an estimate for depth.

- In land-change analysis the original number of classes is often grouped into a limited number of classes for reasons of thematic and/or positional accuracy. But when accuracy levels would allow, it would be worthwhile to analyse also if what is observed at the grouped level accounts for all group members (e.g., the example of the Forests and Woodlands in paragraph 5.4.1 clearly showed that change dynamics had a greater effect on relatively minor land covers/uses).

Several aspects of scale are pertinent. Not only the dimensions of scale mentioned above but also the different *kinds of scales* (e.g., observational and policy scales) and different *components of scale* (e.g., cartographic scale, grain, extent and coverage) (Wu and Li 2006).

Grouping of classes in the land-change analyses in order to identify change processes resulted in variation in the semantic contents of data expressed as differences in categorisation. It is important to make such groupings as it clearly shows that at aggregated data levels the local variability of spatially explicit land- changes may be obscured, whereas patterns can be shown that at more detailed data levels may remain invisible, and vice versa (Veldkamp *et al.* 2001b). Thus, there is an obvious need to make complementary multi-scale analyses in order to detect the change dynamics at different levels.

However, complementarity does not only exist in multi-scale analysis but also in the combination of land-use with land-cover change. In most studies land cover and land use are amalgamated but one ought to clearly distinguish one from the other (Jansen and Di Gregorio 2002) and lately this issue has received anew attention (GLP 2005; Bakker and Veldkamp 2008; Comber 2008; Comber *et al*. 2008). The analysis results in Albania clearly demonstrate that the two are complementary, i.e. the interpretation of land-cover change is strengthened by land-use change and vice versa (Chapters 5 and 6). Land cover is an aid in understanding patterns, whereas land use helps understanding processes. Patterns and processes operate at different scales and they should be linked in spatially explicit land-change science. The realisation that land cover and land use represent different dimensions, kinds and components of scale should be extended to a more integrated approach in land-change science.

Furthermore, choices of scale not only affect what can or will be analysed, but also what can or will be done (Rothman 2002). Thus, scale is linked to governance (Veldkamp 2009).

8.1.4 Modelling dynamics

For the human brain memories have two functions. Mostly one is tempted to immediately refer to the first function, i.e. to remember *the past.* Humans define concepts in order to process experiences, store them in memory and when necessary recall these experiences from memory. However, the second function, i.e. the ability to construct with elements from that past *the future* is often not considered (BBC Horizon documentary '*How does your memory work?*' broadcasted in the Netherlands on 14 September 2008). If one considers land-change analysis in a similar manner than one could argue that one needs to remember the land-change *past* in order to be able to imagine a land-change *future*. Thus the understanding of land-change dynamics at various times in the past ("what was") makes a contribution to the improvement of the understanding of current and future land changes ("what is" and "what will be"). The human brain is unique in being able to look both backwards and forwards. This ability could be used to improve projections and scenario building ("what if"). The scientific community could make progress in land-change analysis by looking more frequently at the land-change past to furnish those elements to project a land-change future. Preferred pathways from the past (Chapter 6) may thus make a contribution to future trajectories and to instruments such as scenario building and projections.

The variation in semantic contents of class sets has an influence in the analysis of change dynamics of the past and thus it has an influence on modelling dynamics representing the future. Thus, the organisational hierarchy of the class set has consequences for the analysis of preferred pathways and future trajectories of land cover and land use respectively. *A priori* no statement can be made as to what semantic level would be most appropriate. Also in the case of the semantic dimension it seems valid that at aggregated semantic levels the local variability of spatially explicit land changes may be obscured, whereas patterns can be shown that at more detailed semantic levels may remain invisible, and vice versa. This is analogue to what is true for data levels (Veldkamp *et al*. 2001b). There is an obvious need to make complementary multi-level analyses of the semantic dimension in order to detect the change dynamics and explaining factors at the different levels.

Though large progress in simulation techniques and data has been made, uncertainty levels remain high and the predictability of land change in most cases low (Pontius *et al*. 2008). This means that further efforts are required to improve our understanding and characterisation of land change (Verburg *et al*. 2009). Also in this case quantification of uncertainty would make a contribution.

8.2 Perspectives in land-change science

Only connect!

E.M Forster (1879-1970)

The most commonly used categorisation systems are hierarchically structured (e.g., plant taxonomy). To many ecologists it has been long apparent that ecological systems are structured as such (Egler 1942; Schultz 1967). Early on it was also acknowledged that "*it is not to be assumed that some one classification will one day be found, and all others will then be abandoned. Each classification serves a certain purpose, and will continue to exist by its own right*" (Egler 1942). Thus, there is not one categorisation that best characterizes land cover or land use. In addition, it seems not fruitful to go in search of the one hierarchy because there is no single, *a priori* parameter for developing such a hierarchy. Instead, a number of different hierarchies may be used to address different problems (see also Figure 8-1 where several standards are indicated). With standardisation one runs the risk of adopting a categorisation with a determined hierarchy that fits a predetermined purpose. Adopting such a categorisation for another purpose involves working with a system with a bias that might force our thinking into the framework (e.g., overarching concept) that was designed for, and is probably more appropriate for, another problem area. Currently, there is no *a priori* designation of hierarchy imposed by the social and biophysical sciences in such a way that no other manner of looking at either land use or land cover is feasible or useful. The hierarchy theory also includes that principles developed at one hierarchical level cannot be transposed to higher and lower levels. Clear distinction of type and category within the hierarchy will not lead to more scientific progress. It is the inherent and awkward ambiguities of land cover and land use that should be included in the more innovative approach of using fuzzy set theory as the mathematical theory underlying the categorisation. This means a move away from existing systems like CORINE Land Cover and LCCS.

As with the existence of numerous categorisations, the diversity of modelling approaches seems to indicate that modelling is not of a one-size-fits-all nature for the understanding of land-change dynamics. In fact, nowadays the use of multiple models is advocated because the complexity of land dynamics cannot be addressed by a single model (Castella *et al.* 2007; Overmars *et al.* 2007). Not only process and pattern need to be examined. The type of change model is also important for the type of changes one is interested in: models range from coarse (e.g., undeveloped to developed), to moderate (e.g., conversions), to fine (e.g., modifications) (Briassoulis 2008). The use of different systems also applies to categorisations as the semantic dimension plays a role. A specific semantic level will render different results in explaining factors than an other semantic level or even a totally different semantic organisation.

The role of cadastral data in land-change science has been very limited whereas it has proven to be a very useful basic unit of measurement, especially for land use. This may be partly due to the fact that not in every country it is easy to link cadastral data to land-change data (e.g., in The Netherlands the data resides in two different institutions), the temporal dimension of data may be different (e.g., in several Eastern European countries the cadastral data is not yet up to date, whereas farmers have to submit every year their requests for EU subsidies to the Paying Agency in their country) and the spatial dimension does not always coincide (e.g., the cadastral parcel unit is not identical to crop areas). However, more emphasis should be put on the consideration of the complementarity of such data types and the possible surplus value of the combination of such data sets for land-change science (e.g., direct linkage between landowner and/or land user with land (change) (Valbuena *et al*. 2010)). Parcel-based systems used in spatial econometric modelling may represent more realistically actors and their spatial relationships. Bakker and Van Doorn (2009) used the cadastral archives (data on property size, distance from the farmer's residence and his or her property) and landowners' data to make a realistic typology of farmer types. This typology could be made spatially explicit because of the cadastral data. The innovation of linking cadastre with land-change science would be the integrative aspects rather than advances in the disciplines being joined.

Land-change science should move beyond the analysis of land cover towards a focus on land use with the latter defined in terms of land function and activities. Land function is the capacity of land to provide goods and services related to the *intended* land use as well as the *unintended* land use (e.g., aesthetic beauty, cultural heritage, water retention and preservation of biodiversity). Especially the different land uses that are systematically linked through temporal (e.g., crop rotations) and/or spatial (e.g., agro-pastoralism) interactions in so-called *land-use systems* require more than land-cover observations alone (Verburg *et al*. 2009). The unintended land uses are not a by-product of rural land use. Nowadays spatial planning and rural policies are targeted at protecting and strengthening such functions. The (intensity of) interactions between people and their environment require innovative data collection methods focussing on land function that will become more important than just land cover. The progress of the land-change community to develop new data, methodologies and models from the mid 1990s to date shows that this community has the capacity to innovate.

The concept of multi-functional land use is again a topic of interest for policy and decision makers concerned with rural development (Willemen *et al*. 2010). Therefore research should abandon the idea of direct linkages between land use and land cover. Cihlar and Jansen (2001) have shown in a systematic manner that more often than not there are no one-to-one relationships. Bakker and Veldkamp (2008) pointed out that even if there is a one-to-one relationship the amount of commodity harvested could not be directly translated into land-cover extent. Dominant or *primary* land uses could be maintained for the traditional concept of close association of land use with land cover (see also Jansen and Di Gregorio 2004a) and usually correspond to the intended land use that directly affects and controls the land cover, whereas *secondary* land uses do not directly affect and control the land cover (see also Jansen and Di Gregorio 2003). Research should pay more attention to further investigation of the land-use/land-cover relationship and the consequences of certain types of relationships in analysis, monitoring and modelling of land change.

It is a mistake to assume that where relationships are found among aggregated data, these relationships will also be found among individuals or households, or vice versa. This is called the ‘ecological fallacy’ (Kok *et al*. 2010). This means that different scales comprise different information on relationships between driving factors and land use, or land cover. Furthermore, these relationships need to be determined at the scale they are used. Often *a priori* assumptions are made as to which driving factor will occur at what level. Both for Latin America (Kok and Veldkamp 2001) and Romania (Chapter 7) these assumptions have been challenged while studying in the first case the spatial and in the second case the semantic dimensions of land dynamics. Scale sensitivity remains an issue.

8.3 Conclusions

The focus has been on *methodology*: the parameterised approach to categorisation to create multi-level class sets for two subjects, land cover and land use, and the use of such class sets in harmonisation efforts, land-change analysis and modelling dynamics.

- The use of different perspectives in categorisation systems has shown to be of chief importance, in addition to the fact to accept that categorisations make a contribution to communication of knowledge by being dynamic in nature (Chapters 2 and 8). Therefore, a new definition of categorisation is proposed that includes the three dimensions of scale: *categorisation is a spatial, temporal, or spatio-temporal, and organisational hierarchy based segmentation of the world*.
- Categorisation is also a means for data standardisation and data harmonisation (Chapters 3, 4 and 8). Data standardisation assumes that the advances in knowledge, technological developments and changing policy objectives will not have an impact on the categorisation framework. If they would have an impact, the standard would show to be dynamic in nature. Therefore, more emphasis should be put on data harmonisation that embodies different perspectives that are complementary and interdependent. These different perspectives enrich the understanding of our environment.

- Harmonisation of land cover and land use is facilitated with a parameterised class set as bridging or reference system on the condition that the information in a class can be preserved in the corresponding class of the bridging system (Chapters 3 and 4). If this is the case also harmonisation of change can be achieved (Chapter 4). More research on developing a method to quantify the correspondence between classes and between data sets is necessary.
- The parameterised approach to categorisation can result in comprehensive data sets and time series if (1) the most recent data set is validated, (2) the same categorisation is used in the time series, and (3) the most recent data set is used as baseline to identify changes (Chapters 5 and 6).
- The parameterised class sets contribute to a better understanding of the spatio-temporal and semantic dimensions of land change (Chapters 5 and 6). Several aspects of scale are pertinent to these dimensions. The spatial dimension is determined by cartographic scale, grain, extent and coverage. The temporal dimension determines the rate of change. The semantic dimension distinguishes conversions, where semantic differences are big, and modifications, where semantic differences are small. The use of parameters to define a class is a tool to establish when conversion or modification occurs, as one can quickly see how much difference there is in the used parameters (Chapter 2). Furthermore, land use and land cover differ in the kinds and components of scale. As a result land-use and land-cover change analyses are complementary (Chapter 8).
- Land-change analysis could be further developed realising that a mathematical theory underlies the categorisation (Chapters 3 and 8). At the same time the hierarchy in the categorisation could be used not only to distinguish conversions and modifications but also the level of intensity of such a change (Chapter 4), and the depth and density of classes in the categorisation system could be used when assessing semantic similarity. The use of different levels in land-change analysis is necessary to discover whether local variability of spatially-explicit land changes is obscured at aggregated levels, whereas patterns could be shown that at more detailed levels would have remained invisible and vice versa (Chapters 5 and 6). Complementarity does not only exist in multi-scale analysis but also in the combination of land-use with land-cover change in order to link patterns and processes (Chapter 8).
- Land-change analysis is also complementary to modelling dynamics if one realises that knowledge of the land-change *past* can contribute to imagine a land-change *future* (Chapters 7 and 8). To be able to look both backwards and forwards can help to improve projections and scenario building.
- One should note that the gridding of data, than can be regarded as a form of harmonisation, does influence both land-change analysis and modelling dynamics (Chapters 3, 4, 5, 6, 7 and 8).

- The variation in semantic contents of class sets has an influence in the analysis of change dynamics of the past and thus on modelling dynamics representing the future (Chapters 7 and 8). Which semantic level is most appropriate cannot be stated *a priori*. Therefore, not only multi-scale but also multi-semantic analyses are necessary.
- Through the whole process of data collection, categorisation, harmonisation, change analysis and modelling dynamics uncertainty plays a key role (Chapters 3, 4, 5, 6, 7 and 8). Though progress in analysis techniques and data has been made, uncertainty levels remain high and predictability of land change in most cases low. Further efforts are needed to improve our understanding and characterisation of land change (Chapter 8).

Charles Darwin demonstrated in '*On the origin of species'* (1859) that nature is *dynamic* according to the temporal scale of geology; similarly our landscapes with their land-cover patterns and land-use processes evolve with our time and our history. It is important to realise that not only the temporal dimension is important, but that the spatial and semantic dimensions are equally important.

References

Adamec, J., 1992. Land-use classification study. First draft. Internal working paper. FAO, Rome.

Agrotec S.p.A., 2004. Final report. EU PHARE Land-Use Policy II project. AGROTEC S.p.A. Consortium, Rome. 125pp. (Carrai, G., Jansen, L.J.M., Remmelzwaal, A.)

Ahlcrona, E., Olsson, B., Rosengren, M., 2001. Swedish CORINE Land Cover. In: Groom, G., Reed, T., (Eds.). Strategic landscape monitoring for the Nordic countries. TemaNord 523. Nordic Council of Ministers, Copenhagen. Pp. 95-99.

Ahlqvist, O., 2004. A parameterized representation of uncertain conceptual spaces. *Transactions in GIS* 8(4): 493-514.

Ahlqvist, O., 2005a. Probing the relationship between classification error and class similarity. *Photogrammetric Engineering & Remote Sensing* 71(12): 1365-1373.

Ahlqvist, O., 2005b. Using semantic similarity metrics to uncover category and land-cover change. In: Rodríguez, M.A., Cruz, I.F., Egenhofer, M.J., Levashkin, S. (Eds.). Geospatial Semantics (GeoS) First International Congress, 29-30 November 2005, Mexico City, Mexico. Springer Publishers, Berlin Heidelberg. Pp. 107-119.

Ahlqvist, O., 2005c. Using uncertain conceptual spaces to translate between land-cover categories. *International Journal of Geographical Information Science* 19(7): 831-85.

Ahlqvist, O., 2005d. Probing the relationship between classification error and class similarity. *Photogrammetric Engineering & Remote Sensing* 71(12): 1365-1373.

Ahlqvist, O., 2008a. Extending post-classification change detection using semantic similarity metrics to overcome class heterogeneity: a study of 1992 and 2001 US National Land-Cover database changes. *Remote Sensing of Environment*, 112(3): 1226-1241.

Ahlqvist, O., 2008b. In search for classification that supports the dynamics of science: the FAO Land-Cover Classification System and proposed modifications. *Environment and Planning B* 35: 169-186.

Ahlqvist, O., Keukelaar, J., Oukbir, K., 2000. Rough classification and accuracy assessment. *International Journal of Geographic Information Science* 14(5): 475-496.

Ahlqvist, O., Shortridge, A., 2006. Characterising land-cover structure with semantic variograms. In: Riedl, A., Kainz, W., Elmes, G. (Eds.). Progress in spatial data handling: 12th international symposium on spatial data handling. Springer Publishers, Berlin Heidelberg. Pp. 401-415.

Amblard, L., Colin, J-P., 2009. Reverse tenancy in Romania: actor's rationales and equity outcomes. *Land-Use Policy* 26: 828-836.

Anderson, J.R., Hardy, E.E., Roach, J.T., Witmer, R.E., 1976. A land-use and land-cover classification system for use with remote sensor data. USGS Professional Paper No. 964. Washington, D.C.

APA [American Planning Association], 1999. Land-based classification standard. Second draft. Research Department. Chicago, US.

Bakker, M.M., Van Doorn, A.M., 2009. Farmer-specific relationships between land-use change and landscape factors: introducing agents in empirical land-use modelling. *Land-Use Policy* 26: 809-817.

Bakker, M.M., Veldkamp, A., 2008. Modelling land change: the issue of use and cover in wide-scale applications. *Journal of Land-Use Science* 3(4): 203-213.

Banini, T., 2006. Identity and surroundings: a critical reading in a transcalar perspective. In: Claval, P., Pagnini, M.P., Scaini, M. (Eds.). The cultural turn in geography. Proceedings of the conference, Gorizia Campus, 18-20 September 2003, Trieste, Italy. Edizione Università di Trieste. Pp. 59-68.

Baulies, X., Szejwach, G., 1998. LUCC Data requirements workshop – survey of needs, gaps and priorities on data for land-use/land-cover change research. Organised by IGBP/IHDP-LUCC and IGBP Data and Information Systems (DIS), 11-14 November 1997. LUCC Report Series No. 3, LUCC International Project Office, Barcelona.

Bishr, Y., 1998. Overcoming the semantic and other barriers to GIS interoperability. *International Journal of Geographical Information Science* 12: 299-314.

Bishr, Y.A., Pundt, H., Kuhn, W., Radwan, M., 1999. Probing the concept of information communities - a first step toward semantic interoperability. In: Goodchild, M., Egenhofer, M., Fegeas, R., Kottman, C., (Eds.). Interoperating geographic information systems. Kluwer Academic Publishers, Dordrecht. Pp. 55–69.

Bonchi, F. and Pecori, S., 2003. Data Mining ed estrazione della conoscenza da grandi database territoriali. In: Santini L.G., Zotta, D. (Eds.). Atti della Terza Conferenza Nazionale su Informatica e Pianificazione Territoriale - Input 2003, 5-7 June 2003, Pisa. Alinea, Florence (on CDROM).

Booch, G., 1994. Object-oriented Analysis and Design with Applications. Second edition. Benjamin/Cummings Publishing Company, Redwood City, MA.

Bossard, M., Feranec, J., Otahel, J., 2000. CORINE Land Cover Technical Guide – Addendum 2000. Technical Report 40. European Environment Agency, Copenhagen. 105pp.

Bowker, G.C., Star, S.L., 1999. Sorting things out: classification and its consequences. MIT Press, Cambridge MA.

Briassoulis, H., 2008. Land-use policy and planning, theorising, and modelling: lost in translation, found in complexity. *Environment and Planning B: Planning and Design* 35: 16-33.

Brown, D.G., Duh, J-D., 2004. Spatial simulation for translating from land use to land cover. *International Journal of Geographic Information Science* 18(1): 35-60.

Bullard, R.K., 2000. Rural development of land in the changing environment and culture of Central Europe. Land Markets and Land Consolidation in Central Europe Seminar, 22nd Urban and Regional Data Management Symposium (UDMS), Delft University of Technology, The Netherlands. Pp. 79-88 (www.landentwicklungmuenchen.de/aktuelle_aufsaetze_extern/aktuelle_aufsaetze_international.html).

Büttner, G., Feranec, J., Jaffrain, G., Mari, L., Maucha, G., Soukup, T., 2004. The European CORINE land cover 2000 project. XX Congress of the International Society for Photogrammetry and Remote Sensing, 12-23 July 2004, Istanbul, Turkey. 12pp.

Carrai, G., 2003. LUP geo-database model - version 1.0. Technical report No. 1. EU PHARE Land-Use Policy II project. AGROTEC S.p.A. Consortium, Rome. 55pp.

Castella, J-C., Kam, S.P., Quang, D.D., Verburg, P.H., Hoanh, C.T., 2007. Combining top-down and bottom-up modelling approaches of land-use/cover change to support public policies: application to sustainable management of natural resources in northern Vietnam. *Land-Use Policy* 24: 531-545.

Castella, J-C., Verburg, P.H., 2007. Combination of process-oriented and pattern-oriented models of land-use change in a mountain area of Vietnam. *Ecological Modelling* 202: 410-420.

CEC [Commission of the European Communities], 1985. Soil map of the European Communities. In: Commission of the European Communities, Luxembourg. 124pp.

CEC, 1993. Nomenclature des activités de la communauté européenne (NACE). Première révision. Règlement 761/93 du Conseil. Brussel.

CEC, 1995. CORINE – Guide technique. Commission of the European Communities, Brussels.

CEC, 1999. CORINE land cover. Brussels.

CEC, 2002. Albania - stabilisation and association report. Commission Staff Working Paper, 4 April 2002, Commission of the European Communities, Brussels.

Cihlar, J., Jansen, L.J.M., 2001. From land cover to land use: a methodology for efficient land-use mapping over large areas. *The Professional Geographer* 53(2): 275-289.

Comber, A.J., 2008. Land use or land cover? *Journal of Land-Use Science* 3(4): 199-201.

Comber, A.J., Fisher, P.F., Wadsworth, R.A., 2004a. Assessment of a semantic statistical approach to detecting land-cover change using inconsistent data sets. *Photogrammetric Engineering & Remote Sensing* 70 (8): 931–938.

Comber, A.J., Fisher, P.F., Wadsworth, R.A., 2004b. Integrating land-cover data with different ontologies: Identifying change from inconsistency. *International Journal of Geographical Information Science* 18(7): 691−708.

Comber, A., Fisher, P., Wadsworth, R., 2005a. Comparing statistical and semantic approaches for identifying change from land-cover data sets. *Journal of Environmental Management* 77: 47-55.

Comber, A., Fisher, P., Wadsworth, R., 2005b. What is land cover? *Environment and Planning B: Planning & Design* 23: 199-209.

Comber, A., Fisher, P., Wadsworth, R., 2005c. You know what land cover is but does anyone else?… an investigation into semantic and ontological confusion. *International Journal of Remote Sensing* 26(1): 223-228.

Comber, A., Fisher, P., Wadsworth, R., 2005d. Comparing the consistency of expert land-cover knowledge. *International Journal of* Applied *Earth Observation and Geoinformation* 7: 189-201.

Comber, A.J., Fisher, P.F., Wadsworth, R.A., 2007. Land cover: to standardise or not to standardise? Comment on ‘Evolving standards in land-cover characterisation’ by Herold *et al*. *Journal of Land-Use Science* 2(4): 283-287.

Comber, A.J., Wadsworth, R.A., Fisher, P.F., 2008. Using semantics to clarify the conceptual confusion between land cover and land use: the example of ‘forest’. *Journal of Land-Use Science* 3(2): 185-198.

Dale, P.F., 1995. Land reform, land registration and the cadastre: a comparative analysis between Bulgaria, Hungary and Rumania. (Available at www.oicrf.org)

Danserau, P., 1961. Essai de représentation cartographique des éléments structuraux de la végétation. In: Gaussen, H., 1961. Méthodes de la cartographie de la végétation. Centre National de la Recherche Scientifique. 97th International Colloquium, Toulouse, 1960. Pp. 233-255.

Darwin, C., 1859. On the origin of species - by means of natural selection or the preservation of favoured races in the struggle for life. The 150th anniversary landmark edition. Penguin Classics, London, UK. 516pp.

De Koning, G.H.J., Verburg, P.H., Veldkamp, A., Fresco, L.O., 1998. Land use in Ecuador: a statistical analysis at different aggregation levels. *Agriculture, Ecosystems & Environment* 70:231-247.

De Koning, G.H.J., Verburg, P.H., Veldkamp, A., Fresco, L.O., 1999. Multi-scale modelling of land-use change dynamics in Ecuador. *Agricultural Systems* 61(2): 77-93.

DeFries, R.S., Belward, A.S., 2000. Global and regional land-cover characterisation from satellite data: an introduction to the special issue. *International Journal of Remote Sensing* 21: 1083–1092.

Dematteis, G., 2002. Possibilità e limiti dello sviluppo locale. Becattini, G., Sforzi, F. (Eds.). Lezioni sullo sviluppo locale. Rosenberg & Sellier, Torino. Pp. 41-63.

Di Costanzo, M., Ongaro, L., 2004. The Land Cover Classification System (LCCS) as a formal language: a proposal. *Journal of Agriculture & Environment for International Development* 98(3-4): 117-164.

Di Gregorio, A., Jansen, L.J.M., 1998. A new concept for a land-cover classification system. *The Land* 2 (1): 55-65.

Di Gregorio, A., Jansen, L.J.M., 2000. Land-Cover Classification System (LCCS): classification concepts and user manual. FAO/UNEP/Cooperazione italiana, Rome. 177pp. (with CDROM)

Dumitru, M., 2002. Romania - country case studies on integrating land issues into the broader development agenda. Report commissioned by the World Bank, Bucharest.

Dobson, J.E., Bright, E.A., Coleman, P.R., Durfee, R.C., Worley, B.A., 2000. A global population database for estimating population at risk. *Photogramm. Eng. Remote Sensing* 66: 849-857.

Dolman, A.J., Verhagen, A., Rovers, C.A. (Eds.), 2003. Global environmental change and land use. Kluwer Academic Publishers, Dordrecht.

Duhamel, C., 1995. Programme télédétection et statistique. Cadre de travail statistique utilisation des sols. Draft. EUROSTAT/CESD-Communautaire, Luxembourg.

Duhamel, C., 1998. First approximation for a land-use classification. CESD-Communautaire, Luxembourg. FAO commissioned study (contains an analysis of previously FAO commissioned studies).

Dumanski, J., Pieri, C., 2000. Land quality indicators: research plan. *Agriculture, Ecosystems & Environment* 81:93-102.

EC [European Commission], 2001. Manual of concepts on land-cover and land-use information systems. EC-EUROSTAT Edition 2000 - Theme 5, Luxembourg. 106pp.

Eickhout, B., Van Meijl, H., Tabeau, A., Van Rheenen, T., 2007. Economic and ecological consequences of four European land-use scenarios. *Land-Use Policy* 24: 562-575.

Egler, F.E., 1942. Vegetation as an object of study. *Philosophy of Science* 9: 245-260.

Eiten, G., 1968. Vegetation forms. A classification of stands of vegetation based on structure, growth form of the components, and vegetative periodicity. *Boletim do Instituto de Botânica*, Vol. 4. San Paulo.

Engel-Di Mauro, S., 2009. Seeing the local in the global: political ecologies, world systems, and the question of scale. *Geoforum* 40: 116-125.

EUROSTAT, 1991. EUROSTAT spatial statistics. Task force internal report on land-use statistics, EUROSTAT, Luxembourg.

EUROSTAT, 2001. Land-use/cover area frame statistical survey (LUCAS). Version 1.0. EUROSTAT, Luxembourg.

Evans, T.P., Kelley, H., 2004. Multi-scale analysis of a household level agent-based model of land-cover change. *Journal of Environmental Management* 72: 57-72.

FAO [Food and Agriculture Organization of the United Nations], 1976. A framework for land evaluation. FAO Soils Bulletin No. 32. FAO, Rome.

FAO, 1984. Guidelines: land evaluation for rainfed agriculture. Soils Bulletin 52. FAO, Rome.

FAO, 1988. FAO-UNESCO Soil map of the world. Revised legend. FAO/UNESCO/ISRIC World Soil Resources Reports No. 60 (Reprinted 1990)

FAO, 1990-1995. FAO Production yearbook. Vol. 44-49. FAO Statistics Series No. 99, 104, 112, 117, 125 and 130. FAO, Rome.

FAO, 1996. Forest resources assessment 1990. Survey of tropical forest cover and study of change processes. FAO Forestry Paper No. 130. FAO, Rome.

FAO, 1993. Guidelines for land-use planning. FAO Development Series 1. FAO, Rome. 96pp.

FAO, 1997. Africover land-cover classification. FAO, Rome.

FAO, 1998. FAO Statistical Databases (FAOSTAT), FAO, Rome (www.fao.org).

FAO, 2002. Proceedings of the FAO/UNEP expert consultation on strategies for global land-cover mapping and monitoring. Artimino, Italy, 6-8 May 2002. FAO, Rome. 38pp.

FAO, 2005. Land Cover Classification System - classification concepts and user manual. Software version 2. Environment and Natural Resources Series 8, Rome. 190pp.

FAOSTAT, 2000; 2001. FAO Statistical Databases. FAO, Rome (www.fao.org).

Feng, C-C., Flewelling, D.M., 2004. Assessment of semantic similarity between land-use/land-cover classification systems. *Computers, Environment & Urban Systems* 28: 229-246.

Fisher, P.F., 2003. Multimedia reporting of the results of natural resource surveys. *Transactions in GIS* 7: 309-324.

Fisher, P., Arnot, C., Wadsworth, R., Wellens, J., 2006. Detecting change in vague interpretations of landscapes. *Ecological Informatics* 1(2): 163−178.

Fischer, G., Sun, L.X., 2001. Model-based analysis of future land-use development in China. *Agriculture, Ecosystems & Environment* 85: 163-176.

Fjellstad, W., Mathiesen, H., Stokland, J., 2001. Monitoring Norwegian Agricultural Landscapes – The 3Q Programme. In: Groom, G. , Reed, T., (Eds.). Strategic landscape monitoring for the Nordic countries. TemaNord 523. Nordic Council of Ministers, Copenhagen. Pp. 19-28.

Foody, G.M., 2001. GIS: the accuracy of spatial data revisited. *Progress in Physical Geography* 25(3): 389-398.

Fosberg, F.R., 1961. A classification of vegetation for general purposes. *Tropical Ecology* 2: 1-28.

Fraser, E.D.G., Springer, L.C., 2009. Explaining agricultural collapse: macro-forces, micro-crises and the emergence of land-use vulnerability in Southern Romania. *Global Environmental Change* 19: 45-53.

Fritz, S., See, L., 2005. Comparison of land-cover maps using fuzzy agreement. *International Journal of Geographical Information Science* 19(7): 787–807.

Gahegan, M., Ehlers, M., 2000. A framework for the modelling of uncertainty between remote sensing and geographic information systems. *ISPRS Journal of Photogrammetry & Remote Sensing* 55(3): 176-188.

Graefen, C., 2002. Land reform and land fragmentation and consequences for rural development in the CEE/CIS countries, Paper prepared for the International Symposium by FAO, GTZ, FIG, ARGE Landentwicklung und Technische Universität München, 25-28 February 2002, Munich.

Global Land Project [GLP], 2005. Science plan and implementation strategy. IGBP Report No. 53/IHDP Report No. 19. IGBP Secretariat, Stockholm. 64pp. (Ojima, D., Moran, E., McConnell, W., Stafford Smith, M., Laumann, G., Morais, J., Young, B. (Eds.))

Groom, G. (Ed.), 2004. Developments in strategic landscape monitoring for the Nordic countries. ANP2004:705. Nordic Council of Ministers, Copenhagen. 167pp.

Groom, G., Stjernholm, M., 2001. The Area Information system – a Danish national spatial environmental database. In: Groom, G. , Reed, T., (Eds.). Strategic landscape monitoring for the Nordic countries. TemaNord 523. Nordic Council of Ministers, Copenhagen. Pp. 81-88.

Gruber, T.R., 1993. Towards principles for the design of ontologies used for knowledge sharing. *International Journal Human-Computer Studies* 43: 907-928.

Guttenberg, A.Z., 1959. A multiple land-use classification system. *J. of the American Institute of Planners* 25: 143-50.

Guttenberg, A.Z., 1965. New directions in land-use classification. American Society of Planning Officials. Chicago.

Guttenberg, A., 2002. Multidimensional land-use classification and how it evolved: reflections on a methodological innovation in urban planning. *Journal of Planning History* 1(4): 311-324.

Haenni, R., 2005. Towards a unifying theory of logical and probabilistic reasoning. 4th International Symposium on Imprecise Probabilities and Their Applications (ECSQARU'05) , Pittsburgh, Pennsylvania.

Han, J., Kamber, M., 2000. Data Mining: Concepts and Techniques. Morgan Kaufmann Publishers, San Mateo, CA.

Hansen, M.C., Reed, B., 2000. A comparison of the IGBP-DIScover and University of Maryland 1km global land-cover products. *International Journal of Remote Sensing* 21(6 & 7): 1365-1373.

Harvey, F., Kuhn, W., Pundt, H., Bishr, Y., Riedemann, C., 1999. Semantic interoperability: a central issue for sharing geographic information. *The Annals of Regional Science* 33: 213-232.

Hellman, F., Verburg, P.H., 2006. Legend of European land-cover classes in Dyna-CLUE 2006. Wageningen University and Research, Wageningen. 27pp.

Herold, M., Hubald, R., Sarfatti, P., Di Gregorio, A., 2006a. Translating and evaluating the CORINE legend using the UN Land Cover Classification System (LCCS). Draft version 20 September 2006. 70pp.

Herold, M., Latham, J.S., Di Gregorio, A., Schmullius, C.C., 2006b. Evolving standards in land-cover characterisation. *Journal of Land-Use Science* 1(2-4): 157-168.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., 2005. Very high-resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25: 1965-1978.

Hoshino, S., 2001. Multilevel modelling on farmland distribution in Japan. *Land-Use Policy* 18: 75-90.

IGU [International Geographic Union], 1976. World land-use survey. Report of the Commission to the General Assembly of the International Geographic Union. *Geographica Helvetica* 1: 1-28.

ITC/FAO/WAU, 1996. The Land-Use Database. A knowledge-based software program for structured storage and retrieval of user-defined land-use data sets. Draft version, January 1996.

Ives, A.R., Turner, M.G., Pearson, S.M., 1998. Local explanations of landscape patterns: can analytical approaches approximate simulation models of spatial processes? *Ecosystems* 1: 35–51.

Jansen, L.J.M., 2003. A standard methodology for land-use description and harmonisation of data sets. Technical report 2. EU PHARE Land-Use Policy II project. AGROTEC S.p.A. Consortium, Rome. 81pp.

Jansen, L.J.M., 2004a. Thematic harmonisation and analyses of Nordic data sets into Land Cover Classification System (LCCS) terminology. In: Groom, G. (Ed.). Developments in strategic landscape monitoring for the Nordic countries. ANP 705. Nordic Council of Ministers, Copenhagen. Pp. 91-118.

Jansen, L.J.M., 2004b. Global land-cover harmonisation: report of the UNEP/FAO expert consultation on strategies for land-cover mapping and monitoring, 6-8 May 2002, Artimino, Italy. In: Groom, G. (Ed.). Developments in strategic landscape monitoring for the Nordic countries. ANP 705. Nordic Council of Ministers, Copenhagen. Pp. 75-89.

Jansen, L.J.M., 2006. Harmonisation of land-use class sets to facilitate compatibility and comparability of data across space and time. *Journal of Land-Use Science* 1(2-4): 127-156.

Jansen, L.J.M., 2009. Semantic characterisation of human settlement areas: critical issues to be considered. In: Gamba, P., Herold, M. (Eds.). Global mapping of human settlement: experiences, data sets and prospects. Taylor & Francis, UK. Pp. 241-257.

Jansen, L.J.M., Bagnoli, M., Chimenti, S., Dell'Agnello, A., Focacci, M., Monaco, S.A., Sorini, P., Stoppioni, S., Urbano, F., 2006b. Satellite image interpretation of land-cover types in Manica and Maputo at nominal scale of 1:250,000 and at national level at nominal scale 1:1,000,000. Technical report 1. Integrated Assessment of Mozambican Forests project. Agriconsulting S.p.A., Rome. 165pp.

Jansen, L.J.M., Bagnoli, M., Focacci, M., 2008a. Analysis of land-cover/use change dynamics in Manica Province in Mozambique in a period of transition (1990-2004). *Forest Ecology & Management* 254: 308-326.

Jansen, L.J.M., Carrai, G., Morandini, L., Cerutti, P.O., Spisni, A., 2003b. Analysis of the spatio-temporal and semantic aspects of land-cover/use dynamics. Technical report 2. World Bank Albanian National Forest Inventory project. Agrotec S.p.A., Rome. 40pp.

Jansen, L.J.M., Carrai, G., Morandini, L., Cerutti, P.O., Spisni, A., 2006a. Analysis of the spatio-temporal and semantic aspects of land-cover/use change dynamics 1991-2001 in Albania at national and district levels. *Environmental Monitoring & Assessment* 119: 107-136.

Jansen, L.J.M., Cerutti, P.O., Spisni, A., Carrai, G. and Morandini, L., 2003a. Interpretation of remote sensing imagery for land-cover/use mapping. World Bank Albanian National Forest Inventory project. Technical report 1. Agrotec S.p.A., Rome. 62pp.

Jansen, L.J.M., Carrai, G., Petri, M., 2007. Land-use change dynamics at cadastral parcel level in Albania: an object-oriented geo-database approach to analyse spatial developments in a period of transition (1991-2003). In: Koomen, E., Bakema, A., Stillwell, J., Scholten, H. (Eds.). Modelling land-use change – progress and applications. Springer Publishers, Berlin Heidelberg. Pp. 25-44.

Jansen, L.J.M., Di Gregorio, A., 1998a. The problems with current land-cover classifications: development of a new approach. Land-cover and land-use information systems for European policy needs, Luxembourg, 21-23 January 1998. EC EUROSTAT Studies and Research, Edition 1999, Theme 5: 93-102.

Jansen, L.J.M., Di Gregorio, A., 1998b. Land-use and land-cover characterisation and classification: the need for baseline data sets to assess future planning. Proceedings IGBP-DIS and IGBP/IHDP-LUCC data gathering and compilation workshop, 18-20 November 1998, Institute for Cartography of Catalonia, Barcelona, Spain.

Jansen, L.J.M., Di Gregorio, A., 2002. Parametric land-cover and land-use classifications as tools for environmental change detection. *Agriculture, Ecosystems & Environment* 91(1-3): 89-100.

Jansen, L.J.M., Di Gregorio, A., 2003. Land-use data collection using the 'Land-Cover Classification System' (LCCS): results from a case study in Kenya. *Land-Use Policy* 20(2): 131-148.

Jansen, L.J.M., Di Gregorio, A., 2004a. Obtaining land-use information from a remotely sensed land-cover map: results from a case study in Lebanon. *Int. J. of Applied Earth Observation and Geoinformation* 5: 141-157.

Jansen, L.J.M, Di Gregorio, A., 2004b. Land-Cover Classification System: basic concepts, main software functions and overview of the 'land system' approach. In: Groom, G. (Ed.). Developments in strategic landscape monitoring for the Nordic countries. Nordic Council of Ministers of Environment. ANP 2004/705: 64-73.

Jansen, L.J.M., Groom, G.B., Carrai, G., 2008b. Land-cover harmonisation and semantic similarity: some methodological issues. *Journal of Land-Use Science* 3(2-3): 131-160.

Jansen, L.J.M., Mahamadou, H., Sarfatti, P., 2003c. Land-cover change analyses using LCCS. *Journal of Agriculture & Environment for International Development* 97(1-2): 47-68.

Jansen, L.J.M., Ndiaye, D.S., 2006. Land-cover change dynamics 1978-1999 of (peri-) urban agriculture in the Dakar region. *Journal of Agriculture & Environment for International Development* 100(1-2): 29-52.

Jones, C.B., Ware, J.M., Miller, D.R., 1999. A probabilistic approach to environmental change detection with area-class map data. In: Agouris, P., Stefanidis, A. (Eds.). Integrated spatial databases, digital images and GIS. Lecture Notes in Computer Science Vol. 1737. Springer Publishers, Berlin Heidelberg. Pp. 122-136.

Kalensky, Z.D., 1998. AFRICOVER land-cover database and map of Africa. *Canadian Journal of Remote Sensing* 24(3): 292-297.

Kavouras, M., Kokla, M., 2002. A method for the formalization and integration of geographical categorisations. *International Journal of Geographical Information Science* 16(5): 439-453.

Kavouras, M., Kokla, M., Tomai, E., 2005. Comparing categories among geographic ontologies. *Computers & Geosciences* 31: 145-154.

King, D., Daroussin, J., Taverbier, R., 1994. Development of a soil geographic database from the soil map of the European Communities. *CATENA* 21: 37-56.

Klijn, J.A., Vullings, L.A.E., Van den Berg, M., Van Meijl, H., Van Lammeren, R., Van Rheenen, T., Tabeau, A.A., Veldkamp, A., Verburg, P.H., Westhoek, H., Eickhout, B., 2005. The EURURALIS study: technical document. Alterra Report No. 1196, Alterra Wageningen, The Netherlands. 215pp. (See also www.eururalis.nl).

Kok, K., Veldkamp, A., 2001. Evaluating impact of spatial scales on land-use pattern analysis in Central America. *Agriculture, Ecosystems & Environment* 85: 205-221.

Kok, K., Veldkamp, A., Jansen, L.J.M., 2010. Spatial-scale sensitivity in modelling deforestation patterns: issues for modellers and implications for policy makers. *Ecology and Society* submitted.

Kok, K., Verburg, P.H., Veldkamp, A., 2007. Integrated assessment of the land system: the future of land use. *Land-Use Policy* 24: 517-520.

Kok K., Winograd M. 2002. Modelling land-use change for Central America, with special reference to the impact of hurricane Mitch. *Ecological Modelling* 149: 53-69.

Kostrowicki, J., 1977. Agricultural typology concept and method. *Agricultural Systems* 2: 33-45.

Kostrowicki, J., 1983a. Land-use survey, agricultural typology and land-use systems: introductory remarks. *Rural Systems* 1(1): 1-23.

Kostrowicki, J., 1983b. Land-use systems and their impact on environment. An attempt at a classification. *Adv. Space Res*. 2(8): 209-215.

Kostrowicki, J., 1984. Types of agriculture in Europe: a preliminary outline. *Geographia Polonica* 50: 131-149.

Kostrowicki, J., 1992a. A hierarchy of world types of agriculture. *In*: N. Mohammed (Ed.), New dimensions in agricultural geography. Vol. I - Historical dimensions of agriculture. Vedams Books international, New Delhi. Pp. 163-203.

Kostrowicki, J., 1992b. Types of agriculture map of Europe: concept, method, techniques. In: N. Mohammed (Ed.), 1992. New dimensions in agricultural geography. Vol. I - Historical dimensions of agriculture. Vedams Books international, New Delhi. Pp. 227-249.

Kuechler A.W., Zonneveld, I.S. (Eds.), 1988. Vegetation mapping. Handbook of vegetation science. Vol. 10. Kluwer Academic Publishers, Dordrecht.

Kuemmerle, T., Hostert, P., Radeloff, V.C., Van der Linden, S., Perzanowski, K., Kruhlov, I., 2008. Cross-border comparison of post-socialist farmland abandonment in the Carpathians. *Ecosystems* 11: 614-628.

Lambin, E.F., Baulies, X., Bockstael, N., Fischer, G., Krug, T., Leemans, R., Moran, E.F., Rindfuss, R.R., Skole, D., Turner II, B.L., Vogel, C., 1999. Land-use and land-cover change: implementation strategy, IGBP Report No. 48/IHDP Report No. 10, IGBP, Stockholm. 125pp.

Lambin, E.F., Geist, H.J. (Eds.), 2006. Land-use and land-cover change – local processes and global impacts. IGBP Series, Springer Publishers, Berlin Heidelberg. 222pp.

Lambin, E.F., Geist, H.J., Lepers, E., 2003. Dynamics of land-use and land-cover change in tropical regions. *Annual Reviews of Environmental Resources* 28: 205-241.

Lambin, E.F., Rounsevell, M.D.A., Geist, H.J., 2000. Are agricultural land-use models able to predict changes in land-use intensity? *Agriculture, Ecosystems & Environment* 82: 321–331.

Lambin, E.F., Turner, B.L., Geist, H.J., Agbola, S.B., Angelsen, A., Bruce, J.W., Coomes, O.T., Dirzo, R., Fischer, G., Folke, C., George, P.S., Homewood, K., Imbernon, J., Leemans, R., Li, X., Moran, E.F., Mortimore, M., Ramakrishnan, P.S., Richards, J.F., Skånes, H., Steffen, W., Stone, G.D., Svedin, U., Veldkamp, T.A., Vogel, C., Xu, J., 2001. The causes of land-use and land-cover change: moving beyond the myths. *Global Environmental Change* 11: 261-269.

LANES, 1998. Development of a harmonised framework for multi-purpose land-cover/land-use information systems derived from earth observation data. Final report. Concerted Action of the European Commission - research programme on the environment and climate. CESD-Communautaire, Luxembourg.

Laney, R.M., 2004. A process-led approach to modelling land change in agricultural landscapes: a case study from Madagascar. *Agriculture, Ecosystems & Environment* 101: 135-153.

Larsson, G, 2002. Matching land use and land tenure arrangements for developing programmes. *Landnutzung & Landentwicklung* 43: 71-77.

Lesschen, J.P., Verburg, P.H., Staal, S.J., 2005. Statistical methods for analysing the spatial dimension of changes in land use and farming systems. LUCC Report Series No. 7. International Livestock Research Institute and LUCC Focus 3 Office, Wageningen University, The Netherlands. Pp. 33-48.

Lillesand, T.M., Kiefer, R.W., 2000. Remote sensing and image interpretation. Fourth edition. John Wiley & Sons, New York.

Liu, X., Anderson, C., 2004. Assessing the impact of temporal dynamics on land-use change modelling. *Computers, Environment & Urban Systems* 28: 107-124.

Lombardo, S., Pecori, S., Santucci, A., 2002. The competition for land between urban and rural systems investigated by G.I.S. and data-mining techniques. Decision Making in Urban and Civil Engineering, Vol. 1 and 3, London.

Loveland, T. R., Reed, B. C., Brown, J. F., Ohlen, D. O., Zhu, Z., Yang, L., Merchant, J. W., 2000. Development of a global land-cover characteristics database and IGBP DISCover from 1 km AVHRR data. *International Journal of Remote Sensing* 21(6–7): 1303–1330.

Lund, H.G., 1999. A 'forest' by any other name… *Environmental Science & Policy* 2: 125-133.

Mayaux, P., Bartholomé, E., Fritz, S., Belward, A., 2004. A new land-cover map of Africa for the year 2000. *Journal of Biogeography* 31: 861-877.

Mayaux, P., Eva, H., Gallego, J., Strahler, A.H., Herold, M., Agrawal, S., Naumov, S., De Miranda, E.E., Di Bella, C.M., Ordoyne, C., Kopin, Y., Roy, P.S., 2006. Validation of the global land-cover 2000 map. *IEEE Transactions on Geoscience & Remote Sensing* 44 (7): 1728-1739.

McConnell, W.J., Moran, E.F. (Eds.), 2001. Meeting in the middle: the challenge of meso-level integration. An international workshop on the harmonisation of land-use and land-cover classification. 17-20 October 2000, Ispra, Italy. LUCC Report Series No. 5. Anthropological Center for Training and Research on Global Environmental Change – Indiana University and LUCC International Project Office, Louvain-la-Neuve.

Meiner, A. (Ed.), 1999. Land cover of Estonia. Implementation of CORINE Land Cover Project in Estonia. EEIC, Tallinn. 133pp.

Metzger, M.J., Bunce, R.G.H., Jongman, R.G.H., Muecher, C.A., Watkins, J.W., 2005. A climatic stratification of the environment of Europe. *Global Ecol. Biogeography* 14: 549-563.

Meyer, W.B., Turner, B.L., 1992. Human population growth and global land-use/cover change. *Annual Review of Ecology & Systematics* 23: 39-61.

Meyer, W.B., Turner, B.L. (Eds.), 1994. Changes in land use and land cover: a global perspective. Cambridge University Press, Cambridge, UK.

Muecher, C.A., Stomph, T.-J., Fresco, L.O., 1993. Proposal for a global land-use classification. Internal final report, February 1993. FAO/ITC/WAU, Wageningen.

Mueller-Dombois, D., Ellenberg, H., 1974. Aims and methods of vegetation ecology. John Wiley & Sons, New York and London. 547pp.

New, M., Hulme, M., Jones, P., 1999. Representing twentieth-century space-time climate variability. Part I: development of a 1961-1990 mean monthly terrestrial climatology. *J. Climate* 12: 829-856.

Nolan, R., 1979. Managing the crisis in data processing. *Harvard Business Review* 57: 115-126.

Overmars, K., 2000. Quantification of spatial autocorrelation and an application of a spatial autoregressive model in land-use analysis. Internal report. Wageningen University and Research Centre, Wageningen.

Overmars, K.P., Verburg, P.H., 2006. Multi-level modelling of land use from field to village level in the Philippines. *Agricultural Systems* 89: 435-456.

Overmars, K.P., Verburg, P.H., Veldkamp, A., 2007. Comparison of a deductive and an inductive approach to specify land suitability in a spatially explicit land-use model. *Land-Use Policy* 24: 584-599.

Paludan, C.T., 1976. Land-use surveys based on remote sensing from high altitudes. In: IGU, 1976. World land-use survey. Report of the Commission to the General Assembly of the International Geographic Union. *Geographica Helvetica* 1: 117-24.

Papanastasis, V., 2003, Special study on grazing impact on wooded lands, including fuel wood consumption assessment. Technical report. World Bank Albanian National Forest Inventory project. Agrotec S.p.A., Rome. 68pp.

Parker, D.C., Berger, T., Manson, S.M. (Eds.), 2002. Agent-based models of land-use and land-cover change. Report and review of an international workshop, October 4-7, 2001. LUCC Report Series No. 6. Anthropological Center for Training and Research on Global Environmental Change - Indiana University and LUCC International Project Office, Louvain-la-Neuve.

Pelorosso, R., Leone, A., Boccia, L., 2009. Land-cover and land-use change in the Italian Apennines: a comparison of assessment methods. *Applied Geography* 29(1): 35-48.

Pereira, G.M., 2002. A typology of spatial and temporal scale relations. *Geographical Analysis* 34(1): 21-33.

Pontius, R.G., Boersma, W., Castella, J-C., Clarke, K., De Nijs, T., Dietzel, C., Duan, Z., Fotsing, E., Goldstein, N., Kok, K., Koomen, E., Lippitt, C.D., McConnell, W., Sood, A.M., Pijanowski, B., Pithadia, S., Sweeney, S., Trung, T.N., Veldkamp, A., Verburg, P.H., 2008. Comparing the input, output and validation maps for several models of land change. *Annals of Regional Science* 42: 11-47.

Pontius, R. G., Cheuk, M. L., 2006. A generalized cross-tabulation matrix to compare soft-classified maps at multiple resolutions. *International Journal of Geographical Information Science* 20(1): 1−30.

Pontius, R.G., Cornell, J.D., Hall, C.A.S., 2001. Modelling the spatial pattern of land-use change with GEOMOD2: application and validation. *Agriculture, Ecosystems & Environment* 85: 191-2003.

Pouliquen, A., 2001. Competitiveness and farm incomes in the CEEC agri-food sectors. Implications before and after accession for EU markets and policies. Paper prepared for the European Commission.

REC for Central and Eastern Europe, 2000. Strategic environmental analysis of Albania, Bosnia & Herzegovina, Kosovo and Macedonia. Annex 1 – Country report Albania. Regional Environmental Center for Central and Eastern Europe.

Remmelzwaal, A., 1989. Classification of land and land use. First approach. Internal working paper. FAO, Rome.

Riddell, J., Rembold, F., 2002. Farm land rationalisation and land consolidation: strategies for improved land management in Eastern and central Europe. Keynote speech prepared for the International Symposium Land Fragmentation and Land Consolidation in CEEC: A Gate towards Sustainable Rural Development in the New Millennium, FAO/GTZ/FIG/ARGE Landentwicklung and Technische Universität München, 25-28 February 2002, Munich. (See http://www.landentwicklung-muenchen.de/cd_ceec_conference/4_keynote_speeches/ 4_rembold/rembold.html)

Rocchini, D., Ricotta, C., 2007. Are landscapes as crisp as we may think? *Ecological Modelling* 204(3-4): 535-539.

Rodríguez, M.A., Egenhofer, M.J., Rugg, R.D., 1999. Assessing semantic similarities among geospatial feature class definitions. Interoperating Geographic Information Systems (InterOp'99), Zurich, Switzerland. In: Vckofski, A., Brassel, K., Schek, H-J. (Eds.). Lecture Notes in Computer Science Vol. 1580, Springer Publishers Berlin Heidelberg. Pp. 189-202.

Rodríguez, M.A., Egenhofer, M.J., 2003. Determining semantic similarity among entity classes from different ontologies. *IEEE Transactions on Knowledge & Data Engineering* 15(2): 442-456.

Rogers, E.M., 1983. Diffusion of innovations. Free Press, New York.

Röling, N., 1988. Extension science – information systems in agricultural development. Cambridge University Press, Great Britain. 233pp.

Rothman, D.S., 2002. Editorial. *Integrated Assessment* 3(2-3): 97-99.

Rounsevell, M.D.A., Annetts, J.E., Audsley, E., Mayr. T., Reginster, I., 2003. Modelling the spatial distribution of agricultural land use at the regional scale. *Agriculture, Ecosystems & Environment* 95: 465-479.

Rozelle, S., Swinnen, J.F.M., 2004. Success and failure of reform: insights from the transition of agriculture. *Journal of Economic Literature* 42: 404-456.

Rusu, M., Florian, V., Popa, M., Marin, P., Pamfil, V., 2002. Land fragmentation and land consolidation in the agricultural sector. A case study from Romania. Paper prepared for the International Symposium Land Fragmentation and Land Consolidation in CEEC: A Gate towards Sustainable Rural Development in the New Millennium, FAO/GTZ/FIG/ARGE Landentwicklung and Technische Universität München, 25-28 February 2002, Munich.

Sabates-Wheeler, R., 2002. Consolidation initiatives after land reform: responses to multiple dimensions of land fragmentation in Eastern European agriculture. *Journal of International Development* 14: 1005-1018.

Scepan, J., 1999. Thematic validation of high-resolution global land-cover data sets. *Photogrammetric Engineering & Remote Sensing* 65(9): 1051-1060.

Schultz, A.M., 1967. The ecosystem as a conceptual tool in the management of natural resources. In: Cieriacy-Wantrup, S.V. (Ed.). Natural resources: quality and quantity. University of California Press, Berkeley. Pp.139-161.

Schuurman, N., Leszczynski, A., 2006. Ontology-based metadata. *Transaction in GIS* 10(5): 709-726.

SDC/FAO/World Bank/Ministero degli Affari Esteri, 2003. Forestry project Albania. Swiss Agency for Development and Co-operation/Food and Agricultural Organization of the United Nations/World Bank/Italian Ministry of Foreign Affairs.

Shapiro, I.D., 1959. Urban land-use classification. *Land Economics* 35(2): 149-155.

Sims, D., 1992. Description and classification of land-use types. FAO informal discussion paper.

Sokal, R., 1974. Classification: purposes, principles, progress, prospects. *Science* 185(4157): 1115-1123.

Sowa, J.F., 2000. Knowledge representation: logical, philosophical and computational foundations. Brooks Cole Publishing Co., Pacific Grove (CA). 594pp.

Strahler, A.H., Boschetti, L., Foody, G.M., Friedl, M.A., Hansen, M.C., Herold, M., Mayaux, P., Morisette, J.T., Stehman, S.V., Woodcock, C.E., 2006. Global land-cover validation: recommendations for evaluation and accuracy assessment of global land-cover maps. EC Joint Research Centre – Institute for Environment and Sustainability, Ispra, Italy. 48pp.

Swets, J.A., 1988. Measuring the accuracy of diagnostic systems. *Science* 240: 1285-1293.

Swinnen, J.F.M., 1999. The political economy of land reform choices in Central and Eastern Europe. *Economic of Transition* 7(3): 637-664.

Swinnen, J.F.M., 2000. Political and economic aspects of land reform and privatization in Central and Eastern Europe. Paper prepared for the UDMS Conference Land Markets and Land Consolidation in Central Europe, Delft.

Thompson, M., 1996. A standard land-cover classification for remote-sensing applications in South Africa. *South African Journal of Science* 92: 34-42.

Torrens, P.M., 2002. Cellular automata and multi-agent systems as planning support tools. In: Geertman, S., Stillwell J. (Eds.). Planning support systems in practice. Advances in Spatial Science. Springer Publishers, London. Pp. 205-222.

Travaglia, C., Milenova, L., Nedkov, R., Vassilev, V., Milenov, P., Radkov, R., Pironkova, Z., 2001. Preparation of land-cover database of Bulgaria through remote sensing and GIS. FAO Environment and Natural Resources Working Paper No. 6. FAO, Rome.

Trochain, J.L., 1961. Représentation cartographique des types de végétation intertropicaux africains. In: Gaussen, H. Méthodes de la cartographie de la végétation. Centre National de la Recherche Scientifique. 97th International Colloquium, Toulouse, 1960. Pp. 87-102.

Turner, B.L., 2010. Sustainability and forest transitions in the southern Yucatan: the land architecture approach. *Land-Use Policy* 27: 170-179.

Turner, B.L., Lambin, E.F., Reenberg, A., 2007. The emergence of land-change science for global environmental change and sustainability. *Proceedings of the National Academy of Sciences* 104(52): 20666-20671.

Turner, B.L., Skole, D., Sanderson, S., Fischer, G., Fresco, L.O., Leemans, R., 1995. Land-use and land-cover change science research plan. IGBP Global Change Report No.35/IHDP Report No. 7. IGBP/IHDP, Stockholm/Geneva.

UN [United Nations], 1989. International standard classification of all economic activities (ISIC). Third revision. United Nations Statistics Division, Statistical Classifications Section, New York.

UN, 1998. Central Product Classification (CPC). Version 1.0. United Nations Statistics Division, Statistical Classifications Section, New York.

UN, 2008. International standard classification of all economic activities (ISIC). Fourth revision (11 August 2008). United Nations Statistics Division, Statistical Classifications Section, New York.

UN-ECE [United Nations Economic Commission for Europe], 1989. Proposed ECE standard international classification of land use. Geneva.

UN-ECE, 2001. Environmental performance review of Albania. Report prepared for the eight session of the Committee on Environmental Policy, November 2002. Geneva.

UNEP [United Nations Environment Programme], 1999. State of the environment report Albania. United Nations Environmental Programme, Nairobi. (Available at http://enrin.grida.no/allsoe.cfm)

UNEP/FAO, 1994. Report on the UNEP/FAO expert meeting on harmonizing land-cover and land-use classifications. Geneva, 23-24 November 1993. GEMS Report Series No. 25. Nairobi.

UNESCO [United Nations Educational, Scientific and Cultural Organization], 1973. International classification and mapping of vegetation. Paris.

Urban Renewal Administration, 1965. Standard land-use coding manual: a standard system for identifying and coding land-use activities (SLUCM). Urban Renewal Administration, Housing and Home Financing Agency, and Bureau of Public Roads, Department of Commerce. Washington D.C.

U.S. Soil Conservation Service, 1975. Soil Taxonomy, by the Soil Survey Staff. USDA Agriculture Handbook No. 436. Washington, D.C.

Valbuena, D., Verburg, P.F., Bregt, A.K., 2008. A method to define a typology for agent-based analysis in regional land-use research. *Agriculture, Ecosystems & Environment* 128: 27-36.

Valbuena, D., Verburg, P.H., Veldkamp, A., Bregt, A.K., Ligtenberg, A., 2010. Effects of farmers' decisions on the landscape structure of a Dutch rural region: an agent-based modelling approach. *Landscape and Urban Planning* in press.

Van den Ban, A.W., Hawkins, H.S., 1985. Agricultural extension. Longman Scientic & Technical, UK. 328pp.

Van Dijk, T., 2003a. Scenarios of Central European land fragmentation, *Land-Use Policy* 20(2): 149-158.

Van Dijk, T., 2003b. Dealing with central European land fragmentation. A critical assessment on the use of Western European instruments. Eburon, Delft.

Van Meijl, J.C.M., Van Rheenen, T., Tabeau, A., Eickhout, B., 2005. The impact of different policy environments on land use in Europe. *Agriculture, Ecosystems & Environment* 114: 21-38.

Van Meurs, W., 1999. Land reform in Romania – a never-ending story. *Southeast Europe Review for Labour and Social Affairs* 2(2): 109-122.

Van Oosterom, P., Groothedde, A, Lemmen, C., Van der Molen, P., Uitermark, H., 2010. Landadministratie als hoeksteen van de geo-informatie infrastructuur – Deel 2 van 3. *GEO-INFO* 7(1): 31-35.

Van Tongeren, F., Van Meijl, H., Surry, Y., 2001. Global models applied to agricultural and trade policies: a review and assessment. *Agricultural Economics* 26: 149-172.

Vckovski, A., 1999. Interoperability and spatial information theory. In: Longley, P.A., Goodchild, M.F., Maguire, D.J., Rhind, D.W., (Eds.). Geographical information systems - principles and technical issues. Wiley, New York. Pp. 31-37.

Veldkamp, A., 2009. Investigating land dynamics: future research perspectives. *Journal of Land-Use Science* 4(1): 5-14.

Veldkamp, A., Fresco, L.O., 1996. CLUE-CR: an integrated multi-scale model to simulate land use change scenarios in Costa Rica. *Ecological Modelling* 91: 231-248.

Veldkamp, A, Kok, K., de Koning, G.H.J., Schoorl, J.M., Sonneveld, M.P.W., Verburg, P.H., 2001a. Multi-scale system approaches in agronomic research at the landscape level. *Soil and Tillage Research* 58: 129-140.

Veldkamp, A, Kok, K., de Koning, G.H.J., Verburg, P.H., Bergsma, A.R., 2001b. The need for multi-scale approaches in spatial specific land-use change modelling. *Environmental Modelling & Assessment* 6: 111-121.

Veldkamp, A., Lambin, E.F., 2001. Editorial: predicting land-use change. *Agriculture, Ecosystems & Environment* 85: 379–386.

Verburg, P.H., Chen, Y.Q., 2000. Multi-scale characterisation of land-use patterns in China. *Ecosystems* 3: 369-385.

Verburg, P.H., De Groot, W.T., Veldkamp, A., 2003. Methodology for multi-scale land-use change modelling - concepts and challenges. In: Dolman, A.J., Verhagen, A. and Rovers, C.A. (Eds.). Global environmental change and land use. Kluwer Academic Publishers, Dordrecht. Pp. 17-51.

Verburg, P.H., Veldkamp, A., Bouma, J., 1999. Land-use change under conditions of high population pressure. The case of Java. *Global Environmental Change* 9: 303-312.

Verburg, P.H., Eickhout, B., Van Meijl, H., 2008. A multi-scale, multi-model approach for analysing the future dynamics of European land use. *The Annals of Regional Science* 42 (1): 57-77.

Verburg, P.H., Ritsema van Eck, J.R., De Nijs, T.C.M, Dijst, M.J., Schot, P., 2004c. Determinants of land-use change patterns in the Netherlands. *Environment and Planning B* 31: 125-150.

Verburg, P.H., Schot, P.P., Dijst, M.J., Veldkamp, A., 2004b. Land-use change modelling: current practice and research priorities. *GeoJournal* 61: 309-324.

Verburg, P.H., Schulp, C.J.E., Witte, N., Veldkamp, A. 2006. Downscaling of land-use change scenarios to assess the dynamics of European landscapes. *Agriculture, Ecosystems & Environment* 114: 39-56.

Verburg, P.H., Soepboer, W., Veldkamp, A., Limpiada, R., Espaldon, V., Mastura, S.S.A., 2002. Modelling the spatial dynamics of regional land use: the CLUE-S model. *Environmental Management* 30: 391-405.

Verburg, P.H., Van de Steeg, J., Veldkamp, A., Willemen, L., 2008. From land-cover change to land-function dynamics: a major challenge to improve land characterisation. *Journal of Environmental Management* 90(3): 1327-1335.

Verburg, P.H., Veldkamp, A., 2005. Spatial modelling to explore land-use dynamics – Editorial. *International Journal of Geographical Information Science* 19(2): 99-102.

Verburg, P.H., Veldkamp, A., Willemen, L., Overmars, K.P., Castella, J-C., 2004a. Landscape-level analysis of the spatial and temporal complexity of land-use change. In: DeFries, R., Asner, G., Houghton, R. (Eds.). Ecosystems and land-use change. Geophysical Monograph Series 153, American Geophysical Union. Pp. 217-230.

Vidican, G., 2008. Land reform and land use in Romania. *Development & Transition* 11: 5-8. (Available at www.developmentandtransition.net)

Vidican, G., 2009. Assessing land re-allocation decisions during transition in Romania. *Land-Use Policy* 26: 1080-1089.

Wadsworth, R.A., Comber, A.J., Fisher, P.F., 2006. Expert knowledge and embedded knowledge: or why long rambling class descriptions are useful. In: Riedl, A., Kainz, W., Elmes, G. (Eds.). Progress in spatial data handling: 12th international symposium on spatial data handling. Springer Publishers, Berlin Heidelberg. Pp. 197-213.

Walker, B., 1998. GCTE and LUCC – a natural and timely partnership. *LUCC Newsletter* 3: 3-4.

Walker, B., Steffen, W., Canadell, J., Ingram, J. (Eds.), 1997. The terrestrial biosphere and global change: implications for natural and managed ecosystems. Synthesis volume. IGBP Book Series 4, Cambridge University Press, Cambridge.

Walsh, S.J., Evans, T.P., Welsh, W.F., Entwisle, B., Rindfuss, R.R., 1999. Scale-dependent relationships between population and environment in northeastern Thailand. *Photogrammetric Engineering & Remote Sensing* 65(1): 97–105.

Westhoek, H.J., Van den Berg, M., Bakkes, J.A., 2006. Scenario development to explore the future of Europe's rural areas. *Agriculture, Ecosystems & Environment* 114: 7-20.

White, R., Engelen, G., 1997. Cellular automata as the basis of integrated dynamic regional modelling. *Environmental Planning B* 24: 235–246.

White, J.D., Running, S.W., 1994. Testing scale-dependent assumptions in regional ecosystem simulations. *Journal of Vegetation Science* 5: 687-702.

Willemen, L., Hein, L., Van Mensvoort, M.E.F., Verburg, P.H., 2010. Space for people, plants, and livestock? Quantifying interactions between multiple landscape functions in a Dutch rural region. *Ecological Indicators* 10(1): 62-73.

World Fact Book, 2005. Available at http://www.odci.gov/cia/publications/factbook/geos/al.html.

Wu, J., Li, H., 2006. Concepts of scale and scaling. In: Wu, J., Jones, K.B., Li, H., Loucks, O.L. (Eds.). Scaling and uncertainty analysis in ecology: methods and applications. Springer Publishers, Berlin Heidelberg. Pp. 1-13.

Wu, F., Webster, C.J., 1998. Simulation of land development through the integration of cellular automata and multi-criteria evaluation. *Environmental Planning B* 25: 103–126.

Wyatt, B.K., 2000. Remote sensing of land-cover and land-cover change. In: Verstraete, M.M., Menenti, M., Peltoniemi, J. Observing land from space: science, customers and technology, Vol. 1-2. Kluwer Academic Publishers, Dordrecht.

Wyatt, B.K., Billington, C., De Bie, K., De Leeuw, J., Greatorex-Davies, N., Luxmoore, R., 1998. Guidelines for land-use and land-cover description and classification. ITE/WCMC/ITC/UNEP/FAO Internal unpublished working document. ITE, Monkshood, UK.

Wyatt, B.K., Gerard, F.F.G., 2001. What's in a name? Approaches to the intercomparison of land-use and land-cover classifications. In: Groom, G.B., Reed, T. (Eds.). Strategic landscape monitoring for the Nordic countries. Nordic Council of Ministers, TemaNord 2001:523. Pp. 113-121.

Wyatt, B.K., Greatorex-Davies, J.N., Hill, M.O., Parr, T.W., Bunce, R.G.H., Fuller, R.M., 1993. Comparison of land-cover definitions. Countryside 1990 series, Vol. 3. Report prepared for the Department of the Environment by the Institute for Terrestrial Ecology (ITE). Natural Environment Research Council, UK. 131pp. (with diskette).

Zhou, J., Civco, D., 1996. Using genetic learning neural networks for spatial decision making in GIS. *Photogrammetric Engineering & Remote Sensing* 62: 1287–1295.

Summary

Introduction and objectives

The extent and intensity of land-cover change and land-use change, in short land change, increased in the 20th century. This has implications for sustainable development, livelihood systems and biodiversity, as well as contributing to changes in the biogeochemical cycles of the Earth. Thus, land change is central to global environmental change. The recognition that land use and land cover are closely related has called for a coupled human-environment system analysis. For the integrated research on this system *data and categorisation* and understanding the *dynamics of scale* are important. Data and categorisation examines data availability and data quality, and devises a categorisation structure suitable for the various research needs. The dynamics of scale at which land-change processes operate, and the different scales at which they are analysed, pose major obstacles to developing a comprehensive understanding.

The overall objective of this thesis is an improved understanding of how class sets using a parameterised approach can contribute to the improved understanding of the spatio-temporal and semantic dimensions of land-change dynamics. Thus, the focus is on methodology. The immediate objectives formulated are:

1. Can a parameterised approach to the categorisation of land cover and land use result in comprehensive data sets and time series that contribute to and that are functional in the understanding of land-change dynamics?
2. Is harmonisation of land cover and land use feasible and facilitated with a parameterised class set as bridging or reference system; can harmonisation of change be achieved?
3. In particular modifications are infrequently captured in land-change studies. Can parameterised class sets contribute to the analysis of the spatio-temporal and semantic dimensions of change dynamics and change processes such as conversions and modifications?
4. How does variation in semantic contents of class sets influence modelling dynamics and what are the consequences for the analysis of preferred pathways of change and future trajectories (e.g., from a policy or decision-making point of view)?

Categorisation and data

Systematic description of the coupled human-environment system is needed in order to understand land-change dynamics. Land cover and land use are the two key elements that describe the environment in natural and human-activity related terms. An internationally accepted categorisation system for either land cover or land use does not exist. Existing categorisation systems fall short in their ability to store classes, they are often internally inconsistent and ambiguous, and mix land cover with land use or vice versa. There is an obvious need for the development of a comprehensive categorisation system based upon systematic description of classes. Such a system should use a set of independent quantifiable diagnostic criteria, the parameters, and be developed

according to an overarching concept. The FAO/UNEP Land-Cover Classification System (LCCS) intends to be such a system for land cover. It is based upon examination of criteria commonly used in existing categorisations that identify and describe land cover in an impartial, measurable and quantitative manner. LCCS is an *a priori*, hierarchically organised, parameterised categorisation system where a class is composed of measured or observed characteristics. These parameters have standard definitions. In addition, LCCS can also make a contribution to change detection at the level of conversion of a class, whereas modification within a certain class type becomes immediately identifiable by a difference in parameter, or through the use of additional parameters. Modifications can be reversed with time, thus they are temporal scale dependent. From the semantic viewpoint a conversion means large semantic differences between classes, whereas modification means small semantic differences.
An international agreement on the definition and categorisation of land use is to this day inexistent. Consequently, a common terminology is lacking. The term ‘land use’ has different meanings across disciplines and, as a result, implies a set of mostly unidentified parameters. These perspectives on land use are, however, all valid. Examination of major land-use parameters utilised by sectoral class sets shows that the two parameters occur in most existing systems: ‘function’ and ‘activity’. It is therefore proposed to combine the ‘function’ approach, describing land uses in an economic context, with the ‘activity’ approach, describing what actually takes place on the land in physical or observable terms. ‘Function’ groups all land used for the same economic purpose independent of the type of activities taking place, whereas ‘activity’ groups all land undergoing a certain process resulting in a certain type of product that may serve different functions.

Harmonisation

A common problem in land-change dynamics is that over time knowledge advances, technology develops and policy objectives change. This means that with each survey being executed with a class set designed for its purpose, a new baseline data set is created rather than a data sequence. Differences in the naming of classes, changes in class definition and addition or removal of classes in data sets covering the same area in different periods create difficulties in the separation of actual changes over time from apparent changes in category definitions. In practise, however, results from different surveys will need to be harmonised over time and space. But there is no commonly accepted methodology of how to achieve high quality harmonisation results. Development of the general-purpose LCCS has led to the belief that once such a categorisation system becomes widely adopted for new surveys the problem of data harmonisation would be overcome. But with each data collection effort lessons are learnt that leave their imprint on successive efforts (e.g., CORINE land cover 1990 versus 2000). Data standardisation may thus be an unrealistic expectation and only partly feasible with the need for data harmonisation always present.
The semantic aspect is just one of the aspects related to harmonisation and spatial data integration. It forms often the major barrier to data integration and interoperability. If the problem of harmonisation is limited to the semantic aspects, i.e. the variation in semantic contents of data expressed as differences in categorisation, then various approaches have been developed to address the methodological issues and for

computing semantic similarity. The examples provided by five Nordic class sets for which correspondence between classes was established, using the LCCS as reference system, shed light on a number of problems. The use of a reference system introduces an additional level of unknown uncertainty, although it limits the number of pair-wise class comparisons to be made. In LCCS standard set theoretic representations are used, so there is no semantic distance function for computing partial overlap. Correspondence was either complete, partial or approximate at best, and in all these cases it would have been useful to be able to quantify the level of correspondence as this would not only give an idea of how much information is maintained but also of uncertainty. Though LCCS is a hierarchical system, this hierarchy is neither used to establish the position of the class in the hierarchy, nor the importance of parameters used in the definition of classes or the type of match. In LCCS the proper distinction between real changes from changes in harmonisation is hampered.

This leads to the issue that harmonisation of class sets should consider both space and time, as the objective should include harmonisation of land *change* to analyse environmental processes and problems. The land-use change harmonisation process was illustrated with an example from Albania. The specifically created reference system based upon the 'function' and 'activity' parameters facilitated harmonisation of class sets at the semantic contents level in parallel with achieving harmonisation of land-use change. As data quality is of paramount importance for any harmonisation attempt, more research is needed to define quantitative measures to express the harmonisation results, i.e. harmonisation quality, both at class level and between class sets.

Land-change analysis

The use of parameterised categorisations facilitates land-change analysis because the parameters to define classes function at the same time as the parameters to be observed over time. It is more difficult to interpret a change in class name than comparing two sets of parameters. Inventory of land-change types, their location, extent and distribution and an understanding of the dynamics in a certain period at different organisational levels (e.g., national, district, commune) provides to decision makers spatially explicit data and information for sustainable management of natural resources.

Two case studies in Albania were used to analyse land change: one focuses on a countrywide analysis of land-cover change, the other on land-use changes in a number of pilot areas. To overcome some of the classical problems in land-change science, the data sets were created with the same land-cover or land-use class set with the most recent data set created first and statistically validated in the field before the data set of a previous year was created. In this manner changes in conceptualisation and application of categories were avoided.

The land-cover change analysis confirmed that at aggregated data levels the local variability of spatially explicit land changes was obscured, whereas patterns were shown that at more detailed data levels remained invisible and vice versa. At detailed data level various types of conversions and modifications could be shown related to natural resources depletion, in particular deforestation and urbanisation, while at the same time showing that trends are location specific.

The land-use changes in the communes concerned mainly modification of agricultural lands where temporary crops were replaced by permanent crops or vice versa. The intensity of the land-use change was determined using the hierarchy of the categorisation system. Analysis of the preferred pathways of change provided a better insight in the decision making of farmers. After privatisation of the agricultural land, land-use changes were more dynamic and the greater number of pathways (almost twice the number of the previous period) of factors leading to a certain change show that new landowners of the cadastral parcels each went their own way. The permanent deterioration of the environment in Albania should stimulate the Government to strongly invest in land-use planning to distribute resources and exploitations in a well-balanced manner and in non-spatial policies like for example subsidies.

Scale is a central issue in land-change dynamics. The two case studies in Albania clearly show differences in the kinds of scale (observational scale and policy scale) and components of scale (differences in cartographic scale, grain, extent and sampling intensity). All these differences between the land-cover and land-use change analyses add to their complementarity, thereby contributing to a better understanding of the linkages between land-cover patterns and land-use processes.

Modelling dynamics

Understanding the scale of interaction and the scale of different environmental and social processes is of paramount importance to the study of the interaction of human-environment systems. Three dimensions of scale are distinguished: space, time and organisational hierarchy as constructed by the observer. The latter is synonymous with the variation in semantic contents of data expressed as differences in categorisation. Classes present in data sets can also affect the type of explanation given to observed phenomena. In turn this might strongly affect the possible consequences for analysis of preferred pathways and future trajectories.

The relationship between semantic contents of data with modelling dynamics was explored using two land-cover data sets for Romania, one based on LCCS and the other as used in the EURURALIS study. The methodology of the CLUE model was used, as the spatial and temporal dimensions of land change have been explored with this model and the examination of the variation in semantic contents of data is complementary to the earlier research. The LCCS class set comprised three levels of semantic contents and the EURURALIS a single semantic level. Empirical relations between the land-cover class and its driving factors were established using the same set of driving factors. The results show that the variation in semantic contents of data within one data set and between two data sets lead to different sets of spatial determinants. There is no pattern recognizable when establishing the organisational hierarchy. Especially the distinction of field size seems important in Romania as these might reveal to be related to different farm(er) types. They are a key issue reflecting the heterogeneity of human behaviour and decisions. Farmers' perceptions and decisions are not always linked to what spatially or landscape-wise would make most sense. Future policy and decision making depend to a great extent on which organisational hierarchy is present in the data used to formulate a policy or to make an informed decision. This would mean that the semantic dimension does play an underestimated role in land-change dynamics, next to the spatial and temporal dimensions. If the same results would be found in other data sets using different models, not only multi-scale but also multi-semantic analysis will be needed in order to make meaningful predictions of spatially explicit land change.

Synthesis and conclusions

The focus has been on methodology: the parameterised approach to categorisation to create multi-level class sets for two subjects, land cover and land use, and the use of such class sets in harmonisation efforts, land-change analysis and modelling dynamics. The use of different perspectives in categorisation systems has shown to be of chief importance, in addition to the fact to accept that categorisations make a contribution to communication of knowledge by being dynamic in nature. Therefore, a new definition of categorisation is proposed that includes the three dimensions of scale: *categorisation is a spatial, temporal, or spatio-temporal, and organisational hierarchy based segmentation of the world*.

Categorisation is also a means for data standardisation and data harmonisation. Data standardisation assumes that the advances in knowledge, technological developments and changing policy objectives will not have an impact on the categorisation framework. Because if they would the standard would show to be dynamic in nature. Therefore, more emphasis should be put on data harmonisation that embodies different perspectives that complement each other and are interdependent. These different perspectives in the categorisation systems enrich the understanding of our environment. Land-change analysis could be further developed realising that a mathematical theory underlies the categorisation. At the same time the hierarchy in the categorisation could be used not only to distinguish conversions and modifications but also the level of intensity of such a change, and the depth and density of classes in the categorisation system could be used when assessing semantic similarity. The use of different levels in land-change analysis is necessary to discover whether local variability of spatially-explicit land changes is obscured at aggregated levels, whereas patterns could be shown that at more detailed levels would have remained invisible and vice versa. Complementarity does not only exist in multi-scale analysis but also in the combination of land-use with land-cover change in order to link patterns and processes. Land-change analysis is also complementary to modelling dynamics if one realises that knowledge of the land-change *past* can contribute to imagine a land-change *future*. To be able to look both backwards and forwards can help to improve projections and scenario building. One should note that the gridding of data, than can be regarded as a form of harmonisation, does influence both land-change analysis and modelling dynamics. The variation in semantic contents of class sets has an influence in the analysis of change dynamics of the past and thus on modelling dynamics representing the future. Which semantic level is most appropriate cannot be stated *a priori*. Therefore, not only multi-scale but also multi-semantic analyses are necessary.

Through the whole process of data collection, categorisation, harmonisation, change analysis and modelling dynamics uncertainty plays a key role. Though progress in analysis techniques and data has been made, uncertainty levels remain high and predictability of land change in most cases low. Further efforts are needed to improve our understanding and characterisation of land change.

Samenvatting

Inleiding en doelstellingen

In de 20e eeuw nam de omvang en intensiteit van verandering in landbedekking en landgebruik, kortom landdynamiek, toe. Dit heeft gevolgen voor duurzame ontwikkeling, levensonderhoudsystemen en de biodiversiteit, evenals bij te dragen aan veranderingen in de biogeochemische cycli van de aarde. Dus landdynamiekprocessen staan centraal in de wereldwijde veranderingen in het milieu. De erkenning dat landgebruik en landbedekking nauw verwant zijn, heeft opgeroepen tot een gekoppelde analyse van het mens-milieu systeem. Voor het geïntegreerde onderzoek van dit systeem zijn '*data en classificatie'* en inzicht in de '*schaal dynamiek'* belangrijk. Data en classificatie onderzoeken de beschikbaarheid van gegevens, de kwaliteit van data, en een classificatiestructuur die geschikt is voor de verschillende behoeften van het onderzoek. De schaal dynamiek waarop landdynamiekprocessen werken, en de verschillende schalen waarop zij worden geanalyseerd, vormen belangrijke obstakels voor de ontwikkeling van een alomvattend begrip.

De algemene doelstelling van dit proefschrift is een beter begrip hoe klassen, die gebruik maken van een geparametriseerde aanpak, kunnen bijdragen tot een beter vermogen om de spatio-temporele en semantische dimensies van landdynamiek te begrijpen. Het accent ligt hierbij op de methodologie. De specifieke doelstellingen zijn geformuleerd als volgt:

1. Kan een geparametriseerde aanpak van de classificaties van landbedekking en landgebruik leiden tot alomvattende datasets en tijdreeksen die bijdragen aan, en functioneel zijn in, het begrijpen van de landdynamiek?
2. Is harmonisatie van landbedekking en van landgebruik haalbaar en gefaciliteerd met een geparametriseerd stel klassen als overbrugging- of referentiesysteem; kan harmonisatie van veranderingen worden bereikt?
3. In het bijzonder modificaties worden zelden opgenomen in landdynamiek studies. Kan een geparametriseerd stel klassen bijdragen aan de analyse van de spatio-temporele en semantische dimensies van landdynamiekprocessen zoals conversies en modificaties?
4. Hoe beïnvloedt de variatie in de semantische inhoud van een stel klassen de modellering dynamiek en wat zijn de gevolgen voor de analyse van preferente wegen van verandering en toekomstige trajecten (bijvoorbeeld uit een beleid of besluitvorming oogpunt)?

Data en classificatie

Systematische beschrijving van het gekoppelde mens-milieu systeem is nodig om landdynamiek te begrijpen. Landbedekking en landgebruik zijn hierbij de twee belangrijkste elementen die het milieu in zowel natuurlijke als menselijke activiteit gerelateerde termen beschrijven. Een internationaal geaccepteerd classificatiesysteem voor landbedekking of landgebruik bestaat echter niet. Bestaande classificatiesystemen schieten te kort in hun vermogen om veel klassen op te slaan, ze zijn vaak intern

inconsistent en dubbelzinnig, en mengen landbedekking met landgebruik en omgekeerd. Er is een duidelijke noodzaak voor de ontwikkeling van een alomvattend classificatiesysteem op basis van een systematische beschrijving van de klassen. Een dergelijk systeem moet een verzameling van onafhankelijke, kwantificeerbare, diagnostische criteria gebruiken -de parameters- en worden ontwikkeld volgens een overkoepelend concept. Het FAO/UNEP 'Land-Cover Classification System' (LCCS) pretendeert een dergelijk systeem voor landbedekking te zijn. Het is gebaseerd op onderzoek van de criteria die gewoonlijk worden gebruikt in de bestaande classificaties die landbedekking identificeren en beschrijven op een onpartijdige, meetbare en kwantitatieve wijze. LCCS is een *a priori*, hiërarchisch georganiseerd, geparametriseerd classificatiesysteem waar een klasse bestaat uit gemeten of waargenomen kenmerken. Deze parameters hebben standaard definities. Daarnaast kan LCCS ook een bijdrage leveren aan de opsporing van veranderingen op het niveau van conversie van een klasse, terwijl modificatie binnen een bepaald type klasse direct herkenbaar wordt door een verschil in parameter, of door het gebruik van extra parameters. Modificaties kunnen ongedaan worden gemaakt met de tijd, dus zij zijn afhankelijk van de temporele schaal. Vanuit het semantische oogpunt betekent een conversie grote semantische verschillen tussen klassen, terwijl modificatie kleine semantische verschillen betekent.
Een internationaal akkoord betreffende de definitie en classificatie van landgebruik bestaat tot op heden niet. Als gevolg hiervan ontbreekt een gemeenschappelijke terminologie. De term 'landgebruik' heeft verschillende betekenissen in verschillende disciplines met als gevolg het gebruik van een reeks van veelal onbekende parameters. Deze verschillende perspectieven op landgebruik zijn echter alle geldig. Onderzoek naar de belangrijkste landgebruikparameters die in sectorale classificaties worden gebruikt laat zien dat er twee parameters voorkomen in bijna alle systemen: 'functie' en 'activiteit'. Het voorstel is dan ook om de 'functie'-benadering, waarin vormen van landgebruik in een economische context worden beschreven, te combineren met de 'activiteit'-aanpak, waarin wordt beschreven wat er werkelijk gebeurt op het land in fysieke of waarneembare termen. 'Functie' groepeert al het land dat wordt gebruikt voor hetzelfde economische doel onafhankelijk van de aard van de activiteiten die er plaatsvinden, terwijl 'activiteit' al het land groepeert dat een bepaald proces ondergaat resulterend in een bepaald type product dat verschillende functies kan dienen.

Harmonisatie

Een veelvoorkomend probleem in de landdynamiek is dat na verloop van tijd kennis vordert, technologie ontwikkelt en beleidsdoelstellingen veranderen. Dit betekent dat met elke kartering die wordt uitgevoerd, met een voor dat doel specifiek ontworpen classificatie, een nieuwe basis dataset wordt gemaakt in plaats van een continue gegevensreeks. Verschillen in de naamgeving van klassen, veranderingen in de definitie van de klasse, en de toevoeging of verwijdering van de klassen in de datasets over hetzelfde gebied in verschillende periodes leveren problemen op in de scheiding van de feitelijke veranderingen in de tijd van klaarblijkelijke veranderingen in de definities van categorieën. In de praktijk zullen de resultaten van verschillende onderzoeken echter moeten worden geharmoniseerd in tijd en ruimte. Er is echter geen

algemeen aanvaarde methodologie om een hoge kwaliteit van harmonisatieresultaten te bereiken.
De ontwikkeling van LCCS heeft geleid tot de overtuiging dat wanneer een dergelijk classificatiesysteem voor algemene doeleinden op grote schaal wordt geadopteerd voor nieuwe karteringen, het probleem van data harmonisatie zou worden overwonnen. Maar met elke kartering worden lessen geleerd die hun stempel drukken op opeenvolgende inspanningen (bijvoorbeeld CORINE Land Cover 1990 versus 2000). Data standaardisatie is dus een onrealistische verwachting en slechts ten dele haalbaar, maar de noodzaak van data harmonisatie blijft altijd aanwezig.
Het semantische aspect is slechts een van de aspecten die verband houden met harmonisatie en ruimtelijke data-integratie. Het vormt vaak de belangrijkste barrière voor data-integratie en interoperabiliteit. Als het probleem van de harmonisatie beperkt wordt tot het semantische aspect, namelijk de variatie in de semantische inhoud van de data uitgedrukt als verschillen in classificatie, dan zijn er verschillende benaderingen ontwikkeld om de methodologische problemen aan te pakken en voor de berekening van semantische gelijkenis tussen klassen. De voorbeelden van vijf Scandinavische datasets, waarvoor de overeenstemming tussen de klassen werd vastgesteld met behulp van LCCS als referentiesysteem, werpen licht op een aantal problemen. Het gebruik van een referentiesysteem introduceert een extra niveau van onbekende onzekerheid, hoewel het het aantal vast te stellen paarsgewijze overeenstemmingen tussen klassen beperkt. In LCCS worden zogenoemde ‘standaard set theoretische’ representaties gebruikt, er is dus geen semantische afstand functie voor het berekenen van overlap. Overeenstemming tussen klassen was ofwel volledig, ofwel gedeeltelijk of bij benadering, en in al deze gevallen zou het nuttig zijn geweest om te kunnen kwantificeren hoeveel overeenstemming er was, omdat dit niet alleen een idee geeft hoeveel informatie wordt gehandhaafd maar ook van de onzekerheid. Hoewel LCCS een hiërarchisch systeem is, wordt deze hiërarchie niet gebruikt om de positie van de klasse in de hiërarchie te bepalen, noch het belang van de parameters die worden gebruikt in de definitie van de klassen, noch de mate van overeenkomst. Dit belemmert het maken van het juiste onderscheid tussen echte veranderingen in het milieu en veranderingen in de data als gevolg van de harmonisatie.
Dit leidt tot het probleem dat de harmonisatie van datasets zowel ruimte en tijd moet beschouwen, omdat de doelstelling de harmonisatie van veranderingen moet omvatten om milieuprocessen en problemen te kunnen analyseren. Het harmonisatieproces van landgebruikverandering wordt geïllustreerd met een voorbeeld uit Albanië. Het speciaal gecreëerde referentiesysteem, gebaseerd op de 'functie' en 'activiteit' parameters, vergemakkelijkt de harmonisatie van klassen op het niveau van de semantische inhoud parallel met het bereiken van harmonisatie van landgebruikveranderingen. Omdat de data kwaliteit van groot belang is voor elke harmonisatiepoging is meer onderzoek nodig om een kwantitatieve maat voor harmonisatie resultaten, namelijk de harmonisatiekwaliteit, te definiëren zowel op klasniveau als tussen klassen.

Analyse van landdynamiek

Het gebruik van geparametriseerde classificaties vergemakkelijkt de analyse van landdynamiek omdat de parameters die klassen definiëren tegelijkertijd fungeren als de parameters die door de tijd kunnen worden geobserveerd. Het is moeilijker om een verandering in klasnaam te interpreteren dan het vergelijken van twee sets parameters. Inventarisatie van typen landdynamiek, hun locatie, omvang en spreiding, en een begrip van de dynamiek in een bepaalde periode op verschillende organisatorische niveaus (bijvoorbeeld nationaal, district, gemeente) biedt aan verantwoordelijken voor het nemen van besluiten ruimtelijk-expliciete data en informatie voor een duurzaam beheer van natuurlijke hulpbronnen.

Twee casestudies in Albanië werden gebruikt om landdynamiek te analyseren: de ene casestudie richt zich op een landelijke analyse van veranderingen in landbedekking, de andere op veranderingen in landgebruik in een aantal proefgebieden. Om enkele van de klassieke problemen in de wetenschap van landdynamiek te voorkomen, werden de datasets gecreëerd met dezelfde landbedekking- of landgebruikclassificatie beginnend met de meest recente data die in het veld statistisch werden gevalideerd alvorens de dataset van een vroeger jaar werd gemaakt. Op deze manier werden veranderingen in de conceptualisering en de toepassing van de categorieën vermeden.

De analyse van veranderingen in landbedekking bevestigde dat op geaggregeerd dataniveau de lokale variabiliteit van de expliciete ruimtelijk veranderingen werd verborgen, terwijl patronen werden aangetoond die op meer gedetailleerde dataniveaus onzichtbaar bleven en omgekeerd. Op het gedetailleerde dataniveau konden verschillende soorten conversies en modificaties worden aangetoond met betrekking tot uitputting van natuurlijke hulpbronnen, vooral ontbossing en verstedelijking, terwijl tegelijkertijd bleek dat deze trends locatie-specifiek zijn.

De landgebruikveranderingen in de gemeenten hebben voornamelijk betrekking op de modificatie van landbouwgronden waar tijdelijke gewassen worden vervangen door permanente gewassen of vice versa. De intensiteit van deze veranderingen in landgebruik werd bepaald met behulp van de hiërarchie van het classificatiesysteem. Analyse van de gewenste trajecten van verandering leverde een beter inzicht op in de besluitvorming van boeren. Na de privatisering van de landbouwgronden, hadden landgebruikveranderingen een hogere dynamiek en uit het grotere aantal trajecten van factoren (bijna twee keer het aantal uit de vorige periode) die leiden tot een bepaalde verandering bleek dat de nieuwe eigenaren van de kadastrale percelen elk hun eigen weg zijn gegaan. De permanente verslechtering van het milieu in Albanië zou de regering moeten stimuleren om stevig te investeren in de ruimtelijke ordening om middelen en exploitaties op een evenwichtige wijze te verdelen en in niet-ruimtelijk beleid, bijvoorbeeld het geven van subsidies.

Schaal is een centraal thema in de landdynamiek. De twee casestudies in Albanië maken verschillen in de aard van de schaal (observationele en beleid schaal) duidelijk en de schaalcomponenten (verschillen in cartografische schaal, basiseenheid van metingen, de omvang en bemonsteringintensiteit). Al deze verschillen tussen de analyses van veranderingen in landbedekking en in landgebruik voegen iets toe aan hun complementariteit, en aldus dragen ze bij aan een beter begrip van de verbanden tussen landbedekkingpatronen en landgebruikprocessen.

Modellering dynamiek

Inzicht in de omvang van de interactie en de omvang van de verschillende milieu- en sociale processen is van het allergrootste belang voor de studie van de interactie van mens-milieu systemen. Drie dimensies van schaal worden onderscheiden: ruimte, tijd en organisatorische hiërarchie zoals die door de waarnemer wordt geconstrueerd. Deze laatste is synoniem met de variatie in semantische inhoud van de data uitgedrukt als verschillen in classificatie. De klassen in datasets kunnen ook invloed hebben op de aard van de verklaring die wordt gegeven aan de waargenomen verschijnselen. Op zijn beurt kan dit sterk van invloed zijn op de mogelijke gevolgen voor de analyse van geprefereerde routes en toekomstige trajecten van verandering.

De relatie tussen de semantische inhoud van data met modeldynamiek werd onderzocht met behulp van twee landbedekking datasets voor Roemenië, de ene dataset op basis van LCCS en de andere zoals gebruikt in de EURURALIS studie. De methodologie van het CLUE model werd gebruikt omdat de ruimtelijke en temporele dimensies van veranderingen in landbedekking daarmee al zijn onderzocht. Het onderzoek van de variatie in de semantische inhoud van data is complementair aan het eerdere onderzoek. De LCCS data bestaan uit drie semantische niveaus en de EURURALIS data heeft een enkel semantisch niveau. Empirische relaties tussen de landbedekkingklasse en de drijvende factoren zijn vastgesteld met gebruikmaking van dezelfde set factoren. De resultaten tonen aan dat de variatie in de semantische inhoud van de data binnen een dataset of tussen twee datasets leidt tot verschillende sets van ruimtelijke factoren. Er is geen patroon te herkennen bij de vaststelling van de organisatorische hiërarchie. Vooral het onderscheid op de grootte van de landbouwpercelen is belangrijk in Roemenië, want die kan waarschijnlijk worden gerelateerd aan verschillende typen boeren en/of boerderijen. Zij zijn belangrijk omdat zij de heterogeniteit van het menselijk gedrag en menselijke beslissingen weerspiegelen. De percepties van boeren en hun besluiten zijn niet altijd gekoppeld aan wat ruimtelijk of landschappelijk het meest zinvol zou zijn. Toekomstig beleid en besluitvorming hangen voor een groot deel af van welke organisationele hiërarchie aanwezig is in de data die gebruikt wordt om een beleid te formuleren of om een geïnformeerd besluit te nemen. Dit zou betekenen dat de semantische dimensie een niet te onderschatten rol speelt in de landdynamiek naast de ruimtelijke en temporele dimensies. Als dezelfde resultaten zouden worden aangetroffen in andere data met gebruik van andere modellen dan zou niet alleen multi-schaal, maar ook multi-semantische analyse nodig zijn om zinvolle voorspellingen van landdynamiek te maken.

Synthese en conclusies

De aandacht was vooral gericht op de methodologie: de geparametriseerde aanpak van classificatie om stellen klassen te maken waarin meerdere niveaus worden onderscheiden voor twee onderwerpen, te weten landbedekking en landgebruik, en het gebruik van dergelijke klassen in harmonisatie-inspanningen, analyse van landdynamiek en modellering dynamiek. Het gebruik van verschillende perspectieven in de classificatiesystemen heeft aangetoond dat dit van grote betekenis is, naast het feit te aanvaarden dat classificaties een bijdrage leveren aan de communicatie van kennis door hun dynamische natuur. Daarom wordt een nieuwe definitie van classificatie

voorgesteld die ook de drie dimensies van schaal omvat: *classificatie is een ruimtelijke, temporele, of spatio-temporele, en op organisatorische hiërarchie gebaseerde segmentatie van de wereld*.
Classificatie is ook een middel voor data standaardisatie en data harmonisatie. Data standaardisatie gaat ervan uit dat de vooruitgang in kennis, de technologische ontwikkelingen en veranderende beleid doelstellingen geen impact zullen hebben op het classificatiekader. Want als ze wel een impact zouden hebben dan zou blijken dat de standaard een dynamisch karakter heeft. Meer nadruk zou moeten worden gelegd op data harmonisatie omdat het de verschillende perspectieven belichaamt die elkaar aanvullen en van elkaar afhankelijk zijn. Deze verschillende perspectieven in de classificatiesystemen verrijken het begrip van onze omgeving.
Analyse van landdynamiek kan verder worden ontwikkeld met het besef dat een wiskundige theorie aan de classificatie ten grondslag ligt. Tegelijkertijd kan de hiërarchie in de classificatie niet alleen worden gebruikt om conversies en modificaties te onderscheiden, maar ook het niveau van de intensiteit van een dergelijke modificatie, en de diepte en de dichtheid van de klassen in het classificatiesysteem kunnen worden gebruikt bij de beoordeling van semantische gelijkenis. Het gebruik van verschillende niveaus in de landdynamiek analyse is nodig om te ontdekken of de lokale variabiliteit van de veranderingen wordt verborgen op geaggregeerd niveau, terwijl patronen kunnen worden aangetoond die op meer gedetailleerde niveaus onzichtbaar zouden zijn gebleven en omgekeerd. Complementariteit bestaat niet alleen in de multi-schaal analyse, maar ook in de combinatie van landgebruik en landbedekking om patronen en processen te koppelen.
Analyse van landdynamiek is ook een aanvulling op de modellering dynamiek, indien men zich realiseert dat de kennis van landdynamiek uit het verleden kan bijdragen aan het verbeelden van toekomstige veranderingen. In staat zijn om zowel achteruit als vooruit te kijken kan helpen om prognoses en scenario's te verbeteren. Men moet er rekening mee houden dat het groeperen van data, dat eveneens kan worden beschouwd als een vorm van harmonisatie, invloed heeft op de analyse van landdynamiek en modellering dynamiek. De variatie in de semantische inhoud van klassen heeft invloed op de analyse van landdynamiek uit het verleden, en dus op de modellering dynamiek van de toekomst. Welk semantisch niveau het meest geschikt is kan niet *a priori* gesteld worden. Daarom zijn niet alleen multi-schaal, maar ook multi-semantische analyses nodig.
Door het hele proces van het verzamelen van data, classificatie, harmonisatie, analyse van landdynamiek, en modellering dynamiek speelt onzekerheid een belangrijke rol. Hoewel vooruitgang in de analysetechnieken en gegevens is geboekt, blijven de onzekerheidniveaus hoog en de voorspelbaarheid van veranderingen in de meeste gevallen laag. Verdere inspanningen zijn dan ook nodig om onze kennis en de classificatie van landdynamiek te verbeteren.

About the author

Louisa Johanna Maria Jansen was born in Breda on 21 December 1964. She completed secondary school (VWO) in 1984 at the St. Paulus Lyceum in Tilburg. In 1984 she started her study Biology at the Agricultural University in Wageningen, which she finished in 1991, with specialisations in vegetation science (landscape level), soil inventory & land evaluation, geology (geomorphology & geochemistry), hydrogeology and regional soil science. She spent one year abroad at the *Département de recherche sur les systèmes de production rural – Volet Fonsébougou* in Mali, and at the *Instituto Agronômico de Campinas – Seção de Pedologia* in Brazil. In addition, she spent long periods in the Limagne, France, and in Andalusia, Spain. She followed the course *Remote sensing applications for land resource surveys and rural development* at the International Institute for Aerospace Survey and Earth Sciences[41] (ITC) in Enschede, The Netherlands.

During her time as student, she worked at the School of Plant Biology of the University College of North Wales in Bangor, United Kingdom, executing plant ecological experiments and measurements and she worked for the Working Group on Geographical Information of the Agricultural University in collaboration with the Royal Tropical Institute, Amsterdam, on the use of GIS in rural development.

From 1991 to 1993, she returned to the ITC to work as researcher and lecturer for the Water Resources Surveys Division and Soils Division concerning herself with environmental monitoring and watershed management using satellite remote sensing and GIS.

In 1994, she joined the UNDP financed *Land-Use Planning for Rational Utilization of Land and Water Resources* project in Swaziland, executed by the Food and Agriculture Organization of the United Nations (FAO), and worked on the subjects soil conservation and ecology. In 1995, she joined the Land and Water Development Division at FAO headquarters in Rome, Italy, to work on soil resources and soil conservation. While working for the Regular Programme 'Land-Use Characterisation & Classification' her interests returned to land use and land cover. She developed a methodology to transform satellite remote sensing derived land-cover data into land-use information and tested it in two pilot projects (Kenya and Lebanon). In addition, she developed with others a parameterised land-cover categorisation[42] applied in various FAO projects (e.g., Azerbaijan, Bulgaria, Moldova, Romania and 10 Eastern African countries) and presented this methodology at international workshops and conferences.

From 2000 onwards, she worked as freelance senior advisor dividing her attention and eclectic interests mainly between Europe and Africa. In Europe, she worked in Malta

[41] Nowadays known as University of Twente – Faculty of Geo-information Science and Earth Observation.
[42] LCCS has been submitted to ISO in two parts (with ISO numbers 19144-1 and 19144-2 respectively (see http://www.glcn.org/act_7_en.jsp).

for the Priority Actions Programme-Regional Activity Centre of the UNEP Mediterranean Action Plan, in Denmark for the *Nordic Landscape Monitoring* project of the Nordic Council of Ministers co-ordinated by the National Environmental Research Institute, in Romania for the FAO *Land-Cover/Land-Use Inventory by Remote Sensing for the Agricultural Reform* project, in Albania for the World Bank *Albanian National Forest Inventory* project and the European Union PHARE *Land Use Policy II* project, in Moldova for the FAO *Building Capacity in Inventory of Land Cover/Use by Remote Sensing* project, and in Bulgaria for the European Commission *Assistance for the Paying Agency* project. Furthermore, as assessor she was involved in the evaluation of a restricted Call for Proposals from the European Commission for the 7th Framework Programme 'Environment and Sustainable Management of Natural Resources, Including Energy'.
In Africa, she worked for the *Projet de formation en gestion des ressources naturelles et sécurité alimentaire* of the Istituto Agronomico per l'Oltremare of the Italian Ministry of Foreign Affairs at the Centre régional AGRHYMET in Niger and the Centre du suivi écologique in Senegal, for the *Integrated Assessment of Mozambican Forests* project in Mozambique, and as team leader for the *Étude de faisabilité des activités de mise en place d'un système de télédétection et de mise en cohérence de base des données au niveau national au Burundi* for the Rural Development Programme of the Delegation of the European Commission in Burundi.

In 2006, she contacted the Land-Dynamics Group of the Department of Environmental Sciences of Wageningen University in the Netherlands with the proposal to start her self-financed PhD research based upon a selection of already published project results and further research described in this thesis.

In 2008, she joined the Netherlands' Cadastre, Land Registry and Mapping Agency (Kadaster) where she worked as project leader in the Department of Spatial Development and Consultancy occupying herself with projects in the Netherlands and Turkey. In 2009, she became Co-ordinator Back Office *ad interim* of the Department of Kadaster International dealing with projects worldwide and participating in academic working groups. From 2010 onwards, she is Geodetic Advisor at the Department of Kadaster International working in projects in the Russian Federation, Serbia, Turkey and Vietnam, in addition to participating in the International Alliance on Land Tenure and Administration (IALTA) and international conferences.

I was just guessing at numbers and figures,
Pulling the puzzles apart.
Questions of science, science and progress,
Do not speak as loud as my heart.

Coldplay, The Scientist (2002)

Selection of publications

On systematic analysis of land-cover/land-use relationships:

Cihlar, J., Jansen, L.J.M., 2001. From land cover to land use: a methodology for efficient land-use mapping over large areas. *The Professional Geographer* 53(2): 275-289.

Jansen, L.J.M., Di Gregorio, A., 2003. Land-use data collection using the 'Land-Cover Classification System' (LCCS): results from a case study in Kenya. *Land-Use Policy* 20(2): 131-148.

Jansen, L.J.M., Di Gregorio, A., 2004. Obtaining land-use information from a remotely sensed land-cover map: results from a case study in Lebanon. *International Journal of Applied Earth Observation & Geoinformation* 5: 141-157.

On land-change in Sub-Saharan Africa:

Jansen, L.J.M., Bagnoli, M., Focacci, M., 2008. Analysis of land-cover/use change dynamics in Manica Province in Mozambique in a period of transition (1990-2004). *Forest Ecology & Management* 254: 308-326.

Jansen, L., Diarra, S., 1992. Mali-Sud, étude diachronique des surfaces agricoles. Quantification des superficies agricoles et de la dégradation pour quatre terroirs villageois de 1952 à 1987. Version révisée. Institut d'Economie Rurale/Institut Royal des Tropiques. KIT Publications, Amsterdam. 57pp.

Jansen, L.J.M., Mahamadou, H., Sarfatti, P., 2003. Land-cover change analyses: an application in Niger using the FAO/UNEP Land-Cover Classification System (LCCS). *Journal of Agriculture & Environment for International Development* 97(1-2): 47-68.

Jansen, L.J.M., Ndiaye, D.S., 2006. Land-cover change dynamics 1978-1999 of (peri-) urban agriculture in the Dakar region. *Journal of Agriculture & Environment for International Development* 100(1-2): 29-52.

Urbano, F., Jansen, L.J.M., 2007. Spatially explicit land-cover/use data as input to an integrated assessment of Mozambican forests. In: Zeil, P., Kienberger, S. (Eds.). Geoinformation for development – bridging the divide through partnerships. Wichman Publishers, Heidelberg. Pp. 3-13.

On semantic characterisation of human settlement areas:

Jansen, L.J.M., 2009. Semantic characterisation of human settlement areas: critical issues to be considered. In: Gamba, P., Herold, M. (Eds.). Global mapping of human settlement: experiences, data sets and prospects. Taylor and Francis, UK. Pp. 241-257.

Miscellaneous

Jansen, L.J.M., Karatas, M., Küsek, G., Lemmen, Ch., Wouters, R., 2010. The computerised land re-allotment process in Turkey and the Netherlands in multi-purpose land-consolidation projects. FIG Congress Facing the Challenges – Building the Capacity, 11-16 April 2010, Sydney, Australia. 17pp.

Kok, K., Veldkamp, A., Jansen, L.J.M., 2010. Spatial-scale sensitivity in modelling deforestation patterns: issues for modellers and implications for policy makers. *Ecology and Society* submitted.

Van der Molen, P., Jansen, L.J.M., 2010. Land management in the Netherlands: the role of the State as policy maker and landowner. International Land Management Symposium, 10-11 May 2010, Hannover, Germany. 20pp.

PE&RC Education Certificate

With the educational activities listed below the PhD candidate has complied with the educational requirements set by the C.T. de Wit Graduate School for Production Ecology and Resource Conservation (PE&RC) which comprises of a minimum total of 32 ECTS (= 22 weeks of activities)

Review of literature (6.0 ECTS[43])
- Analysis of land change with parameterised multi-level class sets: general introduction (2006).

Writing of project proposal (4.5 ECTS)
- Analysis of land change with parameterised multi-level class sets: general introduction (2006).

Post-graduate courses (3.2 ECTS)
- Land-use and land-cover harmonisation workshop; IGBP-IHDP LUCC Focus 1; Rome, Italy (2006).
- Expert consultation on strategies for global land-cover mapping and monitoring; UNEP/FAO/IAO/Regione Toscane; Artimino, Italy (2003).
- A classification system for landscape monitoring: Nordic and international perspectives workshop; National Environmental Research Institute (NERI); Tune, Denmark (2001,2002).

Laboratory training and working visits (2.5 ECTS)
- Modelling land-use and land-cover change in Europe and Northern Asia project; IIASA; Laxenburg, Austria (1999).
- CSIR SAC and Institute for Soil, Climate & Water; Pretoria, South Africa (1997, 1999).

Invited review of (unpublished) journal manuscript (4.0 ECTS)
- Dynamic land-cover response to change in forest resources; Journal of Environmental Management (2007).
- Land-use models and environmental consequences; Landscape & Urban Planning (2004 and 2005).
- Global hybrid land-cover map; IEEE Transactions in Geoscience & Remote Sensing (2000).

Deficiency, refresh, brush-up courses (6.0 ECTS)
- Frontiers of land dynamics (LAD-50806) (2009).

Competence strengthening / skills courses (5.7 ECTS)
- Prince2 (method for project management) course; Van der Molen Project Advice; Arnhem, The Netherlands (2008).
- How to get European funds?; WelcomEurope; Brussels, Belgium (2008).
- Personalizzare ArcInfo con AML and Grid in ArcInfo; ESRI Italia; Rome, Italy (1995).
- Principles of project formulation and appraisal; FAO; Rome, Italy (1994).

Discussion groups / local seminars and other scientific meetings (9.6 ECTS)
- International Geosphere Biosphere Programme (IGBP) and International Human Dimensions of Global Change Programme (IHDP) Land-Use & Land-Cover Change (LUCC) core project workshops, seminars and meetings (1997-2005).
- IGBP Data & Information Systems (DIS) Land-Cover Working Group meetings (1997-1999).

[43] 1 ECTS is equivalent to 28 hours.

- Presentations and participation in the context of land-cover and land-use information in the European union (e.g., EUROSTAT) (1996-2000).
- FAO/UNEP/ITE/WCMC/ITC working group to discuss land-cover and land-use classification approaches, harmonisation and standardisation (1995-1999).

International symposia, workshops and conferences (42.9 ECTS)

- Wicked problems and clumsy solutions: the role of science in the Delta Metropole; Transforum/Studium Generale; Wageningen, The Netherlands (2010).
- International Alliance on Land Tenure and Administration (IALTA) expert meetings; Amsterdam, The Netherlands, and Munich, Germany (2010).
- International Land Management Symposium: Land management strategies for improving urban-rural inter-relationships – best practices and regional solutions; Hanover, Germany (2010).
- Development Policy Review Network Expert Seminar New pressures on land: rethinking policies for development; Utrecht, The Netherlands (2009).
- Academic Panel Sustainable development: liberalisation of land markets and new processes of land grabbing; 15th Annual International Sustainable Development Research Conference; Utrecht, The Netherlands (2009).
- International Conference A wider view on cultural landscape challenges in Europe; Radio Kootwijk, The Netherlands (2008).
- International Conference Policy meets land management: contributions to the achievement of the Millennium Development Goals; Munich, Germany (2008).
- 45th ERSA Congress Modelling land-use change session; Amsterdam, The Netherlands (2005).
- CEReS International Symposium on Remote Sensing Maximization of the use of satellite data for understanding the Earth environment; Chiba University, Chiba, Japan (2005).
- IAO/USAID/FAO/UNEP workshop on Land-cover mapping and change assessment: applications, policies and networks in support of sustainable development; Florence, Italy (2004).
- National Conference: Applications of Geographical Information System in Albania; Albanian Watershed Assessment Project/USDA Forest Service/USAID; Tirana, Albania (2002).
- EUROSTAT International Seminar Land-cover and land-use information systems for European policy needs; EUROSTAT; Luxembourg, Luxembourg (1998).
- EC Concerted action LANES Development of a harmonised framework for multi-purpose land-cover/land-use information systems derived from Earth observation; Luxembourg, Luxembourg, and Rome, Italy (1997).
- Space Applications Institute (SAI) ‘Annual users’ seminar’, and visit to the EC JRC-SAI; Baveno and Ispra, Italy (1996, 1997).
- Earth observation and environmental information conference; Alexandria, Egypt (1997).
- AFRICAGIS’97 Conference; Gaborone, Botswana (1997).
- Presentation of the FAO land-cover classification and AFRICOVER project, Federal Geographic Data Committee (FGDC); USDA and CCRS; Washington, USA, and Ottawa, Canada (1996).
- FAO land-cover classification at the Workshop on legend and classification, AFRICOVER Project Working Group I; FAO; Dakar, Senegal (1996).
- World Overview of Conservation Approaches and Technologies (WOCAT) workshop and steering committee meeting; Sigriswil, Switzerland (1996).
- Atelier régional La gestion rationnelle des eaux et des sols : expériences et perspectives en Afrique de l’Ouest; OSS/GTZ/FAO; Ouagadougou, Burkina Faso (1995).
- 12ieme Journées du réseau érosion; INRA/ORSTOM; Paris, France (1995).

Lecturing / supervision of practical’s / tutorials (12.0 ECTS)

- Assistance to the Paying Agency of Bulgaria with the interpretation of satellite remote sensing data for identification of land use and crops as declared by the farmers who submitted applications for area subsidies; FAO; Sofia, Bulgaria (2008).
- Systematic analysis of land-cover changes in the (peri-) urban agricultural zones for informed decision making and land-use planning; Centre du suivi écologique (CSE); Dakar, Senegal (2003).

- Systematic analysis of land-cover changes for the Land-Cover/Use Inventory by Remote Sensing for the Agricultural Reform project (TCP/ROM/2801); FAO; Bucharest, Romania (2002).
- Systematic land-cover change analysis for natural resources management and food security; Centre AGRHYMET and IAO; Niamey, Niger, and Florence, Italy (2001).