Modeling and visualizing dynamic landscape objects and their qualities

Daniël Emanuël van de Vlag

Promotoren: Prof. Dr. Ir. A. Stein
Hoogleraar in de wiskundige en statistische methoden
Wageningen Universiteit

Prof. Dr. M. J. Kraak
Hoogleraar in de nieuwe visualisatietechnieken in de cartografie
Universiteit Utrecht

Promotiecommissie: Prof. Dr. Ir. A. K. Bregt
Wageningen Universiteit, Nederland

Prof. Dr. Ir. R. Jeansoulin
Université de Provence, Marseille, France

Prof. Dr. Ir. M. Molenaar
ITC, Enschede, Nederland

Prof. Dr. F. J. Ormeling
Universiteit Utrecht, Nederland

Dit onderzoek is uitgevoerd binnen de onderzoekschool PE & RC te Wageningen.

Modeling and visualizing dynamic landscape objects and their qualities

Daniël Emanuël van de Vlag

PROEFSCHRIFT

ter verkrijging van de graad van doctor
op gezag van de rector magnificus
van Wageningen Universiteit,
Prof. Dr. M.J. Kropff,
in het openbaar te verdedigen
op donderdag 6 april 2006
de namiddags te drie uur
in het auditorium van het ITC,
te Enschede

Modeling and visualizing dynamic landscape objects and their qualities. 2006. PhD thesis. 200 pages.

ISBN 90-8504-384-0

ITC dissertation number: 132
ITC. Enschede, the Netherlands.

To my father and mother,
Klaas & Ria van de Vlag

To my beloved fiancée,
Elske Ezinga

PRO DEO

Abstract

This thesis focuses on modeling and visualizing dynamic landscape objects and their qualities. It contains ontologies to characterize and model dynamic landscape features using spatial data. It considers their spatial data qualities and visualizes them by explorative methods. In this study, the dynamic landscape features are derived from a coastal movement application within the Netherlands, whereby beaches are subject to nourishment due to severe erosion.

The description and classification of beach objects and their processes essentially grounds on the perception of the coastal landscape. Modeling a landscape is a basic agreement on the conceptualization of these features and processes. The aim is to develop a framework for conceptualization of dynamic beach objects, to understand the physical processes involved and to illustrate decision rules adopted in classification of these objects. Also, quality issues related to beach nourishments are studied, visualized and explored, using new visualization techniques.

A domain-specific ontology can serve as a framework for the conceptualization of beach objects and their processes. The discrimination into product and problem ontology supports the guidance for classification of these objects and to elucidate which data 'fit for use'. Data qualities are assessed using a quality matrix, where ontological features are portrayed against quality elements. Elements of positional, thematic and temporal accuracy and data completeness are considered of high importance for the beach nourishment application.

The problem and product ontology helps to define two scenarios; the first determined by the regulations from the Ministry for Public Works; the second grounded on the abilities from an existing spatial dataset. A comparison between them shows that 72.8% of the objects suitable and non-suitable for nourishment are correctly classified. A higher overlap is found in areas where actual beach nourishments were carried out. Inaccuracies in attributes, i.e. altitude, vegetation and wetness, influence the determination of the objects. A sensitivity analysis applied on altitude shows that determinate boundaries for beach nourishment objects are not reasonable and consequently should be treated as vague objects.

The ontology for beach objects is extended with a spatio-temporal ontology that considers objects to be vague and dynamic. It contains

full membership functions for crisp objects, partial membership functions for fuzzy objects and temporal membership functions for dynamic fuzzy objects. The temporal membership functions include seasonal changes of vegetation and daily changes in wetness. A sensitivity analysis shows that the calculated beach nourishment volumes are practically insensitive in relation to assumptions on the temporal membership functions. A spatio-temporal ontology, as an extent of a spatial ontology, is shown to model dynamic processes in landscape studies in a more realistic way.

To classify a coastal landscape, I also consider the level of scale. Object hierarchy is essential but is often ignored when collecting and classifying landscape features. A fuzzy decision tree considers a hierarchical structure for classification based on decision rules on object attributes. These attributes are defined on the basis of uncertain parameters that may change in space and time. A Bayesian hierarchical model deals with modeling and handling this uncertainty. In the beach management application, Bayesian hierarchical modeling is applied to obtain posterior probability distributions for several boundary regions. The posterior distributions yield lower and upper limits of membership functions describing boundaries between object classes. In this way, a proper fuzzy decision tree is build that includes the inherent dynamic uncertainty.

The spatial information of the application contains large multivariate and multi-temporal datasets. An integrated prototype for visualization and exploration of multivariate spatiotemporal datasets is introduced. It is applied to understand and explain the behaviour of dynamic beach objects and their uncertainties. It consists of the map environment (MAP), a parallel coordinate plot environment (PCP) for visualizing attributes of the dataset, and a temporal ordered space matrix environment (TOSM) for presenting spatio-temporal patterns. The TOSM is a new exploration method and can be seen as a schematized map, whereby the rows in the TOSM environment represent time, the columns represent geographic units, and individual cells are colored according to the value of user defined attributes. The prototype is applied on four case studies. A usability test is performed to test for the differences in the ability to detect patterns in multivariate spatio-temporal datasets for each environment. Test measures are efficiency, effectiveness and user's satisfaction. Results show that the TOSM environment and the integrated prototype have significantly better performances in efficiency and user's satisfaction than the MAP and PCP environment.

Samenvatting

Deze studie richt zich op het modeleren en visualiseren van dynamische landschapsobjecten en hun kwaliteitsaspecten. Het bevat ontologieën voor het karakteriseren en modeleren van dynamische landschapseigenschappen aan de hand van ruimtelijke gegevens. Het beschouwt de kwaliteitsaspecten van de gegevens en visualiseert deze aan de hand van exploratieve methoden. In deze studie zijn de dynamische landschapseigenschappen afkomstig van een toepassing in kustafslag en strandbeheer in Nederland, waarbij stranden onderhevig aan kusterosie worden gesuppleerd.

De beschrijving en classificatie van strandobjecten en hun processen zijn in grote mate afhankelijk van de perceptie van het kustlandschap. Immers, het modeleren van een landschap is een initiële overeenkomst in de conceptualisatie van landschapseigenschappen en -processen. Het doel van deze studie is om een model te ontwikkelen voor conceptualisatie van dynamische strandobjecten, om inzicht te verkrijgen in de betrokken fysische processen en om beslisregels voor classificatie van deze objecten toe te lichten. Daarnaast worden kwaliteitsaspecten van zandsuppleties bestudeerd, gevisualiseerd en onderzocht aan de hand van nieuwe visualisatie technieken.

Een domeinspecifieke ontologie kan dienen als een model voor de conceptualisatie van strandobjecten en hun processen. De classificatie van deze objecten kan worden ondersteund door onderscheid te maken in een product - en een probleem ontologie. Dit illustreert tevens de geschiktheid van de data. De kwaliteit van de data kan verder worden getoetst aan de hand van een kwaliteitsmatrix, waarbij ontologische landschapseigenschappen worden uitgezet tegen kwaliteitselementen. Voor de studie van zandsuppleties spelen de kwaliteitselementen positionele -, thematische -, temporele nauwkeurigheid en de volledigheid van de data een belangrijke rol.

De probleem - en product ontologie leiden tot twee scenario's; de eerste is bepaald door de voorschriften van het Ministerie van Verkeer en Rijkswaterstaat; de tweede is bepaald door de mogelijkheden van bestaande ruimtelijke gegevens. Een vergelijking tussen beide scenario's toont aan dat 72,8% van de strandobjecten geschikt voor én niet-geschikt voor zandsuppleties correct is geclassificeerd. Een betere overeenkomst is gevonden in het gebied waar actuele zandsuppleties worden uitgevoerd. Onnauwkeurigheden in de

attributen hoogte, vegetatie en bodemvochtigheid beïnvloeden de determinatie van strandobjecten. Een gevoeligheidsanalyse op de hoogtemetingen toont aan dat scherpe begrenzing van strandobjecten niet gerechtvaardigd is en dat deze objecten derhalve als vage objecten beschreven dienen te worden.

De ontologie voor strandobjecten is uitgebreid met een ruimtelijk temporele ontologie, waarbij de strandobjecten als vage en dynamische objecten worden gekenmerkt. Hierin worden scherp begrensde objecten beschreven aan de hand van volledige membership functies, vage objecten volgens gedeeltelijke membership functies en dynamische objecten door middel van temporele membership functies. De temporele membership functies omvatten seizoensveranderingen van vegetatie en dagelijkse veranderingen van bodemvocht. Een gevoeligheidsanalyse toont aan dat de aannames in de temporele membership functies nagenoeg geen invloed uitoefenen op de berekende volumes voor zandsuppleties. Een ruimtelijke temporele ontologie als uitbereiding op een ruimtelijke ontologie weerspiegelt dynamische processen in landschapsstudies op een realistischer manier.

Voor de classificatie van een kustlandschap is het uiterst belangrijk om rekening te houden met het schaalniveau van waarneming. De hiërarchie van de objecten is hierin van essentieel belang. Echter, dit wordt veelal over het hoofd gezien bij collectie en classificatie van landschapseigenschappen. Een vage beslisboom neemt die hiërarchische structuur in ogenschouw, waarbij de classificatie van object attributen wordt uitgevoerd door middel van beslisregels. Deze attributen worden gedefinieerd op basis van onzekere parameters die op hun beurt veranderen in ruimte en tijd. Het modeleren van deze onzekerheden kan worden gedaan met een Bayesiaanse hiërarchisch model. In de studie van de zandsuppleties wordt de Bayesiaanse hiërarchische model toegepast om posteriore waarschijnlijkheidsverdelingen van objectgrenzen te verkrijgen. Deze waarschijnlijkheidsverdelingen resulteren in onder- en bovenlimieten van de membership functies die de objectgrenzen beschrijven. Op deze manier kan een vage beslisboom worden verkregen dat impliciet de dynamische onzekerheid behelst.

De ruimtelijke gegevens van deze studie bestaat uit meerdere multivariabele en temporele datasets. Voor visualisatie en exploratie van deze multivariabele ruimtelijke temporele datasets is een geïntegreerd prototype ontwikkeld. Het prototype is toegepast om het gedrag van de dynamische strandobjecten en hun onzekerheden te begrijpen en te verklaren. Het prototype bestaat uit een kaart

omgeving (MAP), een parallel coordinate plot omgeving (PCP) voor visualisatie van de attributen en een temporal ordered space matrix omgeving (TOSM) voor de beschrijving van ruimtelijke temporele patronen. De TOSM is een nieuwe visualisatie techniek en is een soort schematische kaart waarbij de rijen in de matrix de tijd weergeven en de kolommen de geografische eenheden voorstellen. De individuele cellen van de matrix geven vervolgens de waarde van de attributen weer, gedefinieerd door de gebruiker. Het prototype is toegepast op vier studies. Een gebruikerstest is uitgevoerd om te testen voor significante verschillen in efficiëntie, effectiviteit en tevredenheid van elke visualisatieomgeving voor het ontdekken van patronen in multivariabele ruimtelijk temporele datasets. Test resultaten tonen aan dat de TOSM en het geïntegreerde prototype significant efficiënter en tot grotere tevredenheid stellen dan de MAP en PCP omgeving.

Acknowledgements

Although a PhD dissertation can yield at most one doctor's title to at most one person, it has involved the help, interest and support of many. Since I value these people more highly than any amount of writing - my own or others' - it is appropriate to thank them at the very beginning of this thesis.

I would like to thank my supervisors Alfred Stein and Menno-Jan Kraak for their support, supervision and scientific insights. Our discussions and meetings were valuable and definitely helped me to think, write and become an independent researcher.

I am grateful to the REV!GIS-team, who where a source for inspiration. In particular, I like to thank Robert Jeansoulin and Bérengère Vasseur for the opportunity to do some collaborative research. Bérengère, thanks for guiding me around Marseille and showing me the wonderful 'calanques'. It was a pleasure to stay there.

I would like to thank all my colleagues in the GIS and EOS departments. Ton Mank, Willy Cock, Jeroen van den Worm, Wim Feringa, Corné van Elzakker, Barend Köbben, Richard Knippers, thank you for the pleasant times during the coffee breaks. Ton, this will be the end of our jointly celebrated birthdays. We had some good laughs organizing them, and I will definitely continue the 'healthy birthday celebrations'. Norman Kerle, I like to thank you for your positive words of encouragement; they were 'precious' to me. I would like to thank the secretaries Saskia Tempelman and Mireille Meester for all their help and support. Thanks to Loes Colenbrander for being ready for all the assistance that I needed.

There are two M.Sc. students, Prakash T.N. and Adelina Malunda whom I supervised, and I am grateful for their scientific contributions. I would like to thank Henk Kloosterman, Rob Jordans, Henk Koppejan and Jelmer Cleveringa of the Ministry of Public Works for providing me with datasets and information on beach nourishments. I also would like to mention Boudewijn van Leeuwen, Wan Bakx and Gerard Reinink for their geo-technical support and the acquisition of datasets. Willem Nieuwenhuis and Martin Schouwenburg are thanked for all their efforts in developing a new visualization tool; we made it in the end. Eddie Poppe, Trias Aditya, Corné van Elzakker, Harald van der Werff, Ulanbek Turdukulov, and Rasika Ahangama thanks for your

participation in the focus group user test, giving me valuable feedback on the visualization tool I developed for exploration of space-time patterns. Your comments eventuated in many improvements of the tool. I would also like to thank the GFM2 and PhD students who participated in the usability test.

My thanks are great to my ITC friends, my fellow 'AIO-ers', Harald van der Werff, Marleen Noomen, Arko Lucieer, Jelle Ferwerda and Arta Dilo. I will definitely miss our 'subway meetings'. Harald, thanks for your openness; it helped me to put many things in perspective. Marleen, we spent much time together, and your friendship is precious to me. Arko, in many ways you have been my example and you encouraged me to hang on in difficult times. Jelle, thanks for helping me out by reorganizing my enormous database. Arta, you have contributed to this thesis by many valuable discussions and I am very grateful for that.

There are many friends who always believed in me, even in times that contact was sparse. I would like to thank Prisca, Inge, H.J., Feike, Laura, Johan, Marinde, Wout, Carolien, Laurens, Sifra, Ilja, Karen, Wiecher, Mirjam, Bennie, Mark, Robert, Paul, Jedidja and all others that supported me during the last four years. I am grateful to David and Marianne, my brother and sister-in-law, and Ingmar and Ida Ezinga, my future in-laws, for their sincere interest and support. I am greatly indebted to my parents, Klaas and Ria van de Vlag, whose everlasting encouragements and great support helped me to persevere under all circumstances.

Elske Ezinga, my beloved fiancée, there are no words to describe what you mean to me. In every way, my life has changed with you at my side. You had to withstand stressful moments and doubts, but your unconditional love and encouragements have stimulated me to continue and succeed. As 'our day' will approach soon, I am thrilled to be your husband and then you will always be my 'principessa'.

Above all, I am grateful to the Lord Almighty who is my source of strength.

Contents

Chapter 1

Introduction

'Make fun of wisdom,
and you will never find it.
But if you have understanding,
knowledge comes easily'

Solomon's Proverbs 14 versus 6

1.1 Problem description

A landscape is a specific view of the real world as we perceive it with our eyes. It is bounded to its domain (e.g. urban, rural, coastal, etc.) and includes natural and man-made features that are subject to physical and artificial processes. Within a landscape, these features can be observed at different resolutions liable to the scale of perception. Hierarchy therefore plays an important part in the characterization of these features. Also, the processes that act upon these features cause growing, shrinking, disappearance or eventually appearance of new landscape features, eventuate in uncertainties within the spatial, thematic and temporal domain. A proper description of a landscape therefore needs to take into account these aspects.

The description of a landscape is essentially grounded on how we perceive the landscape. For a particular landscape, different interests may exist between for example soil scientists, environmentalists, geographers, ecologists or biologists. Modeling a landscape is therefore a basic agreement on the conceptualization of the features and processes. In that sense, it can be denoted by the term 'ontology', i.e. a formal specification of a conceptualization (Gruber 1993). Ontology has its roots in philosophy, and has been widely adopted in the field of Artificial Intelligence (see also section 1.2). An ontology includes all aspects of natural features (e.g. events, processes, dynamics, qualities) within a specific domain. In this study, ontologies are used as a framework to describe and understand natural features within the domain of environmental modeling.

In environmental modeling the essence is to develop a simplification of these landscapes. Hence, a direct link exists between processes and its portrayed model (see figure 1.1). A model helps to understand and predict these processes, grounded on process theories that are converted into the model. Conversely, landscape processes can be measured, for instance by empirical field data, and help to optimize the model, and as such eventuates in new (physical) theories.

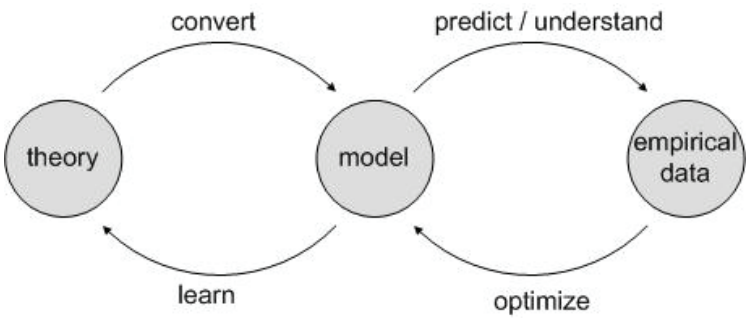


Figure 1.1 *Schematic outline of environmental modeling.*

Modeling landscapes should be analyzed with care, because in real world, natural features often change gradually from one to another (e.g. vegetation, soil types, land use, morphometry, etc.) (Burrough 1996; Fisher *et al.* 2004). Frequently within the modeling stage, this is done in a crisp way (i.e. with clear boundaries). Clearly, crisp objects representing features that change gradually are not satisfactorily. Instead, it is pragmatic to model these features as fuzzy objects, whereby part-of relations between multiple objects are established (see section 1.3). Besides, our world is also a very dynamic place (Peuquet 2002). Fuzzy objects change in space over time and are inseparably intertwining the representation of space and time as they relate to our (experiential) world. Consequently, the indeterminate characteristics of natural features are inextricably connected to the properties of fuzzy objects.

In general, spatial data sets are applied to classify these dynamic natural features. In the last decades, there is an increase use of geographical information systems (GIS) and remote sensing operations for environmental modeling. Frequently, the same data set may be used at various stages during these operations. Therefore, the quality from the input data has an effect on the quality of the end product. The imperfection in spatial data is important to the quality of the final product. Also, decisions rules in classification, as well as the indeterminate characteristics of natural features, should be considered. Hence, quality issues should therefore be related to all stages in the model.

In this thesis, it is my intent to develop a framework for conceptualization of dynamic natural features in a landscape, to understand the physical processes involved and to illustrate decision rules adopted within the classification process. Also, quality issues of these features are studied, visualized and explored, using new visualization techniques. This study has been executed within the pan

European project REV!GIS (= Uncertain Knowledge Maintenance and Revision in Geographic Information Systems) (REV!GIS 2000).

1.2 Ontology

At the beginning of Book IV of his Metaphysics, Aristotle (350 BC) introduces ontology as 'the study of being'. What distinguishes the concept of being is that it is the most general concept. The philosopher Tugendthat (1982) reasoned that for everything and anything one can say that it is. Everything and anything, therefore, can be called being.

Ontology addresses two related questions. What are the categories (or beings) of the world? And what are the laws that govern these categories? An ontology can be related to various sciences, for example, in chemistry we search for chemical elements and the laws which they obey. In physics we try to discover the elementary particles and their laws. Ontology, however, is not a science among sciences (Grossman 1983). It is also not a catalogue of the world, a taxonomy, a terminology or a list of objects, things or whatever else. If anything, an ontology is the general framework (= structure) within which catalogues, taxonomies, terminologies may be given suitable organization. According to Poli (1996), this means that somewhere a boundary must be drawn between ontology and taxonomy.

Ontology represents the 'objective' side of the real world, while the theory of knowledge represents the 'subjective' side. The two sides are obviously dependent, but are not necessarily the same (e.g. compare the front and rear side of a coin). An ontology is not reducible to pure cognitive analysis. It is not an epistemology or a theory of knowledge. In order to conduct ontological analysis, it is necessary to 'neutralize', the cognitive dimension, that is, to reduce it to the default state. Hence, the default state is the descriptive one, where the dimensions of interest are as neutral as possible (= 'natural attitude'). It is possible to modify the default state and construct ontologies of the other cognitive states as well, but this involves modifications of the core of the ontology (Poli 1996).

This study aims to clarify 'ontology' for categorization in the field of environmental modeling. Perceivable, spatial objects form the core of this ontology. Most natural features, however, are not crisp in spatial extent and thematic content and change in time (see section 1.1). Hence, these characteristics should be integrated into the categorization in dynamic landscape objects.

1.3 Fuzzy objects

The focus of this study is on spatial objects with indeterminate boundaries, the so-called fuzzy objects. Zadeh introduced the idea 'fuzzy sets' to deal with inexact concepts in a definable way (Zadeh 1965). Fuzziness is a type of imprecision characterizing objects that for various reasons cannot have, or do not have sharply defined boundaries. It is therefore appropriate to deal with ambiguity, vagueness and ambivalence in mathematical or conceptual models of natural features (Burrough 1996; Fisher 1999; Zhang and Foody 2001).

In fuzzy sets, the grade of membership is expressed in terms of a scale that can vary between 0 and 1. Fuzziness is an admission of the possibility that an object belongs to a set, or that a given statement is true. It is not a probabilistic attribute, in which the grade of membership is linked to a given statistically defined probability function.

Fuzzy sets enable representation of imprecisely defined objects. In general two groups of techniques can be identified. The similarity relation model is data-driven and involves searching for similar patterns in datasets. The semantic import model is user-driven where membership functions are defined by experts. Therefore, an ontology for landscape objects, as described in section 1.2, needs to consider natural features by means of fuzzy sets.

1.4 The issue of spatial data quality

The indeterminate characteristics of natural features are related to the issue of spatial data quality. Quality is defined as "*totality of characteristics of a product that bear on its ability to satisfy stated and implied needs*" (ISO 2003). These product characteristics are defined as internal quality. Next to internal quality, De Bruin (2000) and Chrisman (1984) explicitly refers to "*…fitness-for-use…*". Quality, therefore, is a relative concept dependent on the pursued aims and the considered context. Data quality standards define levels of qualities for data to receive a particular certification, knowing that standards cannot consider all possible uses. Data quality standards are helpful, but the decision regarding its use can only be made in context (Harvey 1998). For this reason, Devillers *et al.* (2005) have criticised spatial data quality standards for providing the user with insufficient information to allow for assessing fitness for use.

Jacobi (1994) describes the general agreement of the European Standardization Organization (CEN) for the definition of spatial data quality standards. He defines:

- *Lineage:* the description of data history, source material, data capture method, date, data producer, and processing method.
- *Accuracy:* the probability of correctly assigning a value.
- *Ability for abstraction:* how well can real world features be defined.
- *Completeness:* the degree of presence of necessary descriptive data to match real world phenomena.
- *Consistency:* the correlations and rules governing a dataset (spatial coherence, duplicate registration, levels of generalization).
- *Reliability:* the consistency of quality parameters and the probability of detecting errors.
- *Currency:* the frequency of updating and changing the map's database.

Guptill (1998) describes the same data quality standards, but redefines accuracy in positional accuracy (i.e. the measures of the horizontal and vertical accuracy of the features in the data) and attribute accuracy (i.e. the accuracy of scalar or nominal values associated with features or relationships contained in the data set).

ISO/TC211 (ISO 2003) – technical committee 211 of the International Standardisation Organisation - has developed a number of international standards for geographic information, including quality principles and quality evaluation principles. The ISO-standards are fairly similar to the CEN standards. ISO, however, makes a distinction in accuracy between positional, thematic and temporal accuracy.

In this study, the following ISO spatial data quality standards will be dealt with regarding dynamic natural features:

- Positional accuracy, i.e. the spatial extent of crisp and fuzzy boundaries.
- Thematic accuracy, i.e. class definitions and separability, as well as accuracy of attributes.
- Temporal accuracy, i.e. changing position and shape of attributes in time as a consequence of time of data capture.
- Data set completeness.

Van Oort (2006) describes in his thesis in great detail the issue of spatial data quality. He pointed out that from the end of the 1980s onwards, interests in spatial data quality have grown strongly within the disciplines of cartography, geography, survey and geodesy. Especially with the emergence of GIS and spatial data from satellites concerns about spatial data quality have increased. With the adoption of GIS in other disciplines, the attention for spatial data quality and its propagation in models has also grown, for example in ecology (Hunsaker *et al.* 2001) and in environmental modeling (Heuvelink 1998; Brimicombe 2003). In the field of statistics, geostatistics emerged (Matheron 1971; Cressie 1991). It is of great importance in spatial data quality research as most variables are spatially or temporally correlated. Geostatistics offers the theory to model both implications of temporal and spatial correlation (Pebesma *et al.* 2005).

Next to the description and quantification of the quality standards, as mentioned above, this study concerns the incorporation of these quality standards in the model stage. Often, spatial data sets are structured in a hierarchical way, meaning that natural features are classified on their parameters and their qualities that in turn are derived from other parameters. As such hierarchical levels in the fuzzy object classification are distinguished.

1.5 Hierarchical Classification

Natural features represented in geographical information systems require classification into objects. How the classification is done depends upon a large extent on the way people perceive the physical environment and the processes that take place in it (see also section 1.1). It is based on decisions on how to represent them.

Hierarchical models help to comprehend and illustrate the decisions and qualities involved. Classification of objects based on a hierarchical model has been widely used in the environmental and life sciences by means of decision tree classification (Tso and Mather 2001). Bayesian hierarchical models are able to address uncertainties in the classification.

The Bayesian hierarchical model provides a unified framework for quantifying uncertainty, in which new data are considered in the context of pre-existing knowledge. A key feature of Bayesian statistics is its ability to incorporate prior and auxiliary information. A Bayesian hierarchical model addresses prior uncertainties in attributes. These attributes are modelled by means of a probability distribution containing parameters that are in turn uncertain as well

(Gelman *et al.* 1996). By modeling both the observed data and these prior uncertainties, the Bayesian approach is able to combine complex data models with external knowledge (Banerjee *et al.* 2004).

A decision tree classifier for landscape objects is a hierarchical structure consisting of several levels, whereby each object is completely described by a set of attributes and a class label. At each level a test is applied to one or more attribute values. The purpose of using a hierarchical structure for classifying objects is to gain more comprehensive understanding of relationships between objects at different scales of observation or at different levels of detail. Decision tree classification however, results into a crisp discretization of natural features. Several authors proposed and applied fuzzy decision tree techniques to classify natural features (Janikow 1998; Li *et al.* 2003; Olaru and Wehenkel 2003; Peng and Flach 2001).

This study adds to the classification of natural features into objects by means of a fuzzy decision tree classifier. Furthermore, a Bayesian hierarchical model is adopted to assess the uncertainty into the classification.

1.6 Geovisualization

The correct interpretation of products, i.e. landscape objects and their qualities, depends largely on its communication and representation from the modeling process to the user. Visualization is an appropriate method for the presentation of patterns and spatial behaviour of these landscape objects.

The objective of visualization is to analyze and explore information about relationships graphically, whereas cartography aims at conveying spatial and temporal relationships in geospatial data. Before the era of geographic information systems (GIS), paper maps and statistics were probably the most prominent tools for researchers to study and visualize their geospatial data, all within the field of cartography. Via GIS, cartography evolved into explorative cartography, creating possibilities for the user to interact with the map and the data sets behind it. This trend in mapping is strongly influenced by developments in other fields (Kraak 2003).

In the 1990s, the field of scientific visualization linked visualization to more specific ways in which modern computer technology can facilitate the process of “making data visible” in real time in order to strengthen knowledge (Kraak 2003). Scientific visualization deals mainly with medical imaging, process model visualization and molecular chemistry. Next to scientific visualization, the field of

information visualization emerged, focusing on non-numerical and multivariate information. Visualization as such can be defined as: *"…the use of computer-supported, interactive, visual representations of data to amplify cognition…"* (Card *et al.* 1999). In the field of exploratory data analysis, visualization tools should assist in an interactive, undirected search for exploring structures and trends, while the maps and graphics should provide a hypothesis.

From the map perspective, this leaded to the field of geovisualization that integrates approaches from scientific visualization, information visualization, explorative cartography, exploratory data analysis and GIS to provide theory, methods and tools for visual exploration, analysis, synthesis and presentation of geospatial data (MacEachren and Kraak, 2001). In a geovisualization environment, maps are used to stimulate visual thinking about geospatial patterns, relationships and trends. An important approach here is to view geospatial data sets in a number of alternative ways, e.g. using multiple representations without constraints set by traditional techniques (Kraak 2003).

The basic aspects of geovisualization are communication by means of new display techniques, while formalisms using new computer technologies can add interaction and dynamics. The purpose of geovisualization is insight in geospatial information. Discovery, decision-making and explanation are the main goals of this insight. Therefore, geovisualization is useful to the extent that it increases our ability to perform cognitive activities.

In this study, the visualization of objects (i.e. landscape objects, geographical units, etc.) into a map will be obvious the first method to comprehend its spatial context, as maps enable the recognition of spatial patterns. New computer technologies as animation and multiple linked views will enable interactive and dynamic exploration of these objects. Spatial data quality elements can be revealed and communicated using multivariate visualization techniques.

1.7 Study area

The whole study is illustrated with a coastal landscape study in the Netherlands. The study area is a 4 by 4 km dynamic coastal area in the north-western part of the Isle Ameland (see figure 1.2).

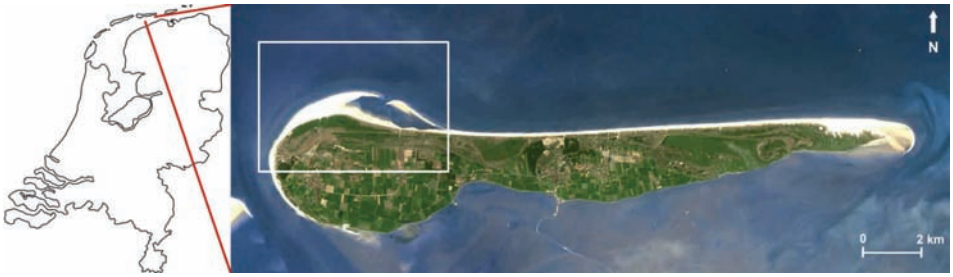


Figure 1.2 *The study area for the beach nourishment application is located in the white box. Image is taken by Aster satellite on 26 July 2001.*

Coastal erosion plays a major role in this area, as a result of strong tidal currents and storm events. To counteract this erosion, beach nourishments are carried out. Beach nourishments are sand suppletions for those beach areas that are under threat to be eroded and are therefore a risk for the safety of the interior. The interest of this study area is to describe, model and visualize beach objects that are suitable for nourishment.

1.8 Research objectives

The general objective of this study is to explore the usefulness of different techniques to describe, model and visualize landscape objects with indeterminate boundaries and their dynamics.

The main objectives of this study are:

- To develop and apply an environmental model with landscape objects using a domain ontology.
- To extend this model, covering the spatio-temporal aspects of dynamic landscape objects using fuzzy set theory and the determination of their spatial, thematic and temporal accuracies.
- To develop and implement a hierarchical structure for classifying landscape objects.
- To develop and apply techniques to visualize these landscape objects and their multivariate data quality elements.

The objectives concerning the coastal landscape study:

- To identify and conceptualize beach objects suitable for nourishment using a domain ontology.

- To apply and implement fuzzy techniques and temporal influences for the conceptualization and modeling stage of beach objects and to assess the effects for the beach nourishment model.
- To apply, implement and validate a fuzzy decision tree classification that classifies beach objects by corresponding decision rules.
- To study and quantitatively assess spatial data quality elements involved in the beach nourishment model.
- To develop, apply and test a visualization tool that illustrate the beach objects and their quality elements.

In figure 1.3 an overview is given of the objectives and the structure of this thesis, outlining the focus of individual chapters.

1.9 Structure of the thesis

The structure of the thesis originates directly from the research objectives (see figure 1.3). Chapter 2 describes in more detail the study area, the data sets and the theory behind beach nourishments. Chapter 3 describes the ontological approach for modeling objects within a beach management application. In this chapter, the spatial ontology is given, assuming no temporal influences. In chapter 4, the ontological approach of the previous chapter is extended for modeling dynamic beach objects with indeterminate boundaries. Here, assumptions are made about the temporal aspects of the objects. Moreover, indeterminacy of the boundaries is assumed. Chapter 5 illustrates an hierarchical model for modeling the beach objects. A fuzzy decision tree approach is applied for classification of these beach objects. Furthermore, a Bayesian hierarchical model is adopted to assess the uncertainty into the classification. Chapter 6 introduces an integrated method to visualize multivariate spatio-temporal data sets. A prototype is developed and tested by means of a focus group and a usability test. It can handle objects (i.e. landscape objects or geographical units) and depict them in time, as well as showing all attributes belonging to each object. It has been applied on four different case studies, including the beach management application. Chapter 7 contains the main conclusions of this study.

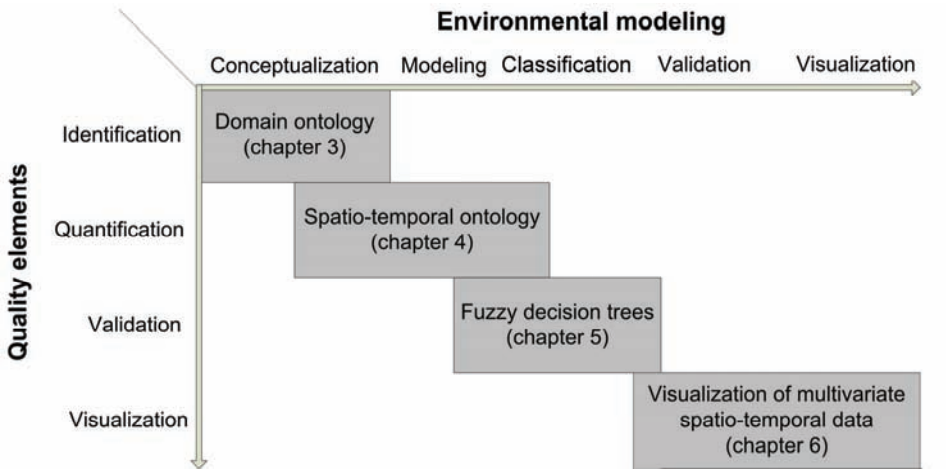


Figure 1.3 *Thesis structure. Stages of the environmental modeling process are depicted along the horizontal axis. The vertical axis illustrates aspects of spatial data quality elements.*

Chapter 2

Ameland case study

'What we observe is not nature itself, but nature exposed to our method of questioning'

Werner Heisenberg (1901-1976)

2.1 Introduction

This chapter focuses on a coastal area in the northern Netherlands. Beach erosion and sedimentation influence its morphology, which in turn has an economic impact on beach management and public security. Previous work in this area considered identification, dynamics and representation of coastal landscape units as foredune, beach and foreshore (Cheng 1999).

2.2 Study area

The study area is located at the north-western part of Ameland, a coastal barrier island on the fringe between the Wadden Sea and the North Sea (Figure 2.1). The Wadden Sea is a shallow tidal sea consisting of tidal flats and channels. It covers virtually all flats at high tide, but reveals them as separated features, sticking out above the water, at low tide. The Wadden Sea has a continuous sediment demand, which is primarily satisfied by transport from ebb-tidal deltas and island coasts.

Geomorphological processes such as erosion, transport and sedimentation of sandy materials are causing major changes the coast of Ameland, in particular at the north-western part of the island, which is a geomorphologically highly dynamic area. A migrating channel from the ebb-tidal delta has reached the western part and erodes the beachplain. Subsequently, the beachplain is extended eastwards and consists of a swashbar with a lagoon behind it (Eleveld 1996). Sand nourishments have to be carried out to counteract beach erosion. To calculate the required volume of sand, the expected erosion, the recurrence interval and the sand reserve have to be determined. Either beach nourishment or underwater sand nourishment is applied (Roelse 2002). This study restricts to beach nourishments, as their effects on the maintenance of the coastline are better known. Beach objects required for nourishment are to be localized, of which the borders, however, are subject to complex coastal dynamics. Identification of beach objects should include both the attributes, and events and processes working on them.

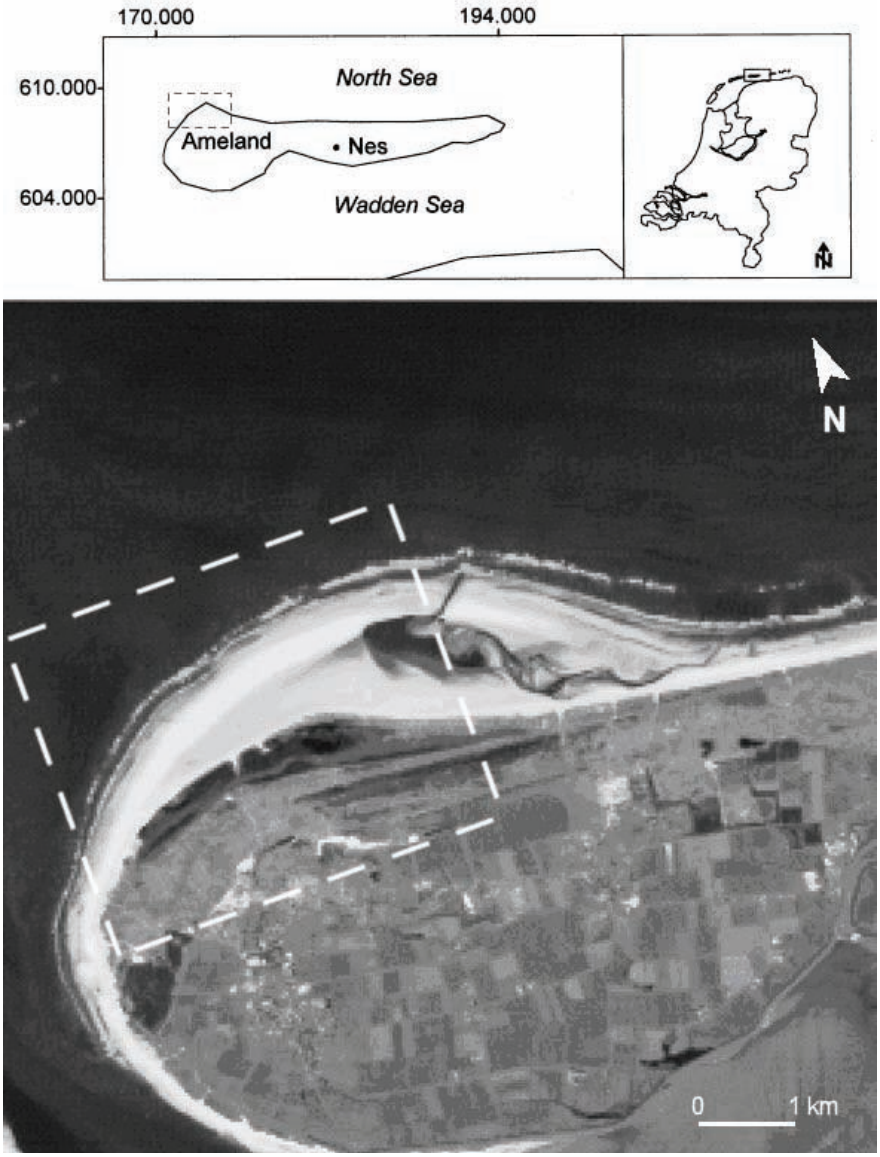


Figure 2.1 *Study area of 4 by 4 km, as seen from the Landsat 7 satellite on 13 may 2000.*

2.3 Datasets

The dataset of the Ameland study area is shown in table 2.1. The digital elevation model of Ameland is derived from the JARKUS data from the DONAR database (Eleveld 1999). The DONAR database contains annual beach and foredune profiles derived from

stereometric analysis of aerial photographs for the dry part of the coastal transect. Transects are 200 to 250 m apart. Elevation is measured at 5 m intervals along a cross-shore line starting at about 800 m seawards to 200 m landwards from the first ridges of dunes. The underwater part of the profile is measured with echosoundings from ships with automatic position-finding systems. The elevation measurements are executed by means of laser altimetry (see figure 2.2). The measurements are oriented within the same reference system and marked by beach posts (De Ruig and Louise 1991). From the point data, these profiles are interpolated towards a 30 m by 30 m grid, using the IDW interpolator of ARCGIS geostatistical analyst. The positional accuracy has a RMSE (root mean square error) of 1.0 m for the X and Y coordinates. To obtain an accuracy value of the Z-value, 5% of the data points are used for cross validation. For the total dataset (1980-2000) the RMSE for Z at $\alpha = 0.05$ fluctuated between 0.17 and 0.39 m.

Table 2.1 *Dataset specifications of satellite images and digital elevation models (DEM).*

Data	Date	Pixel resolution	Remarks	RMSE
Landsat 5 TM	23-05-1989	30 x 30 m	Geometric corrected	1.46
	15-05-1992	30 x 30 m	Geometric corrected	1.63
	07-11-1995	30 x 30 m	Geometric corrected	1.62
Landsat 7 ETM+	30-07-1999	30 x 30 m	Geometric corrected	1.56
	18-10-1999	30 x 30 m	Geometric corrected	1.56
	13-05-2000	30 x 30 m	Geometric corrected	1.61
DEM (Jarkus)	1980-2000	30 x 30 m	Derived from profiles	1.00

Satellite images are derived from Landsat5-TM and Landsat7-ETM+ satellites. Landsat5-TM (Thematic Mapper) registers electro-magnetic radiation in seven wavelengths intervals ranging from 0.45 µm (blue) to 2.35 µm (infrared) and 10.4-12.5 µm and 10.4-12.5 µm (thermal infrared). Landsat7-ETM+ has an additional panchromatic band (0.52 – 0.90 µm). Landsat images contain pixels corresponding to 30 m by 30 m ground surface, with exception for the thermal band with a spatial resolution of 120 m by 120 m and the panchromatic band (Landsat7) with a spatial resolution of 15 m by 15 m.

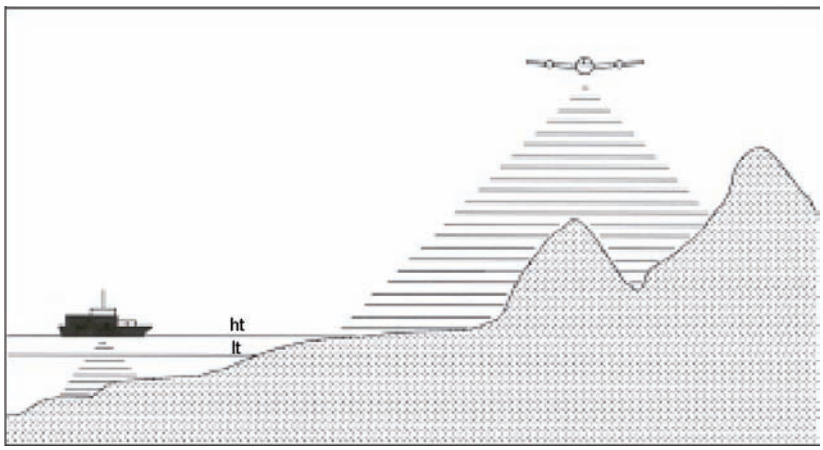


Figure 2.2 *Depth measurements using echosoundings from ships and elevation measurements using airborne laser altimetry.*

2.4 Beach nourishment

According to Dutch policy regulations, beach nourishments are carried out (1) if safety of the interior is at risk, (2) to safeguard dune objects, (3) to stimulate and manage beach recreation, or (4) to reduce the loss of nature areas (Roelse 2002). Reference is made to the basal coastline (*BCL*), being the coastline position on 1-1-1990 as well as to the beach volume at basal coastline (V_b), being the standard for preservation of the coastline. Beach nourishments are carried out when the actual coastline (*ACL*) is below the basal coastline (*BCL*). The volume V_a can be calculated from the surface area using the elevation between the upper and lower boundaries of the *ACL* multiplied with the beach area (figure 2.3).

Next, it is possible to calculate structural erosion (*e*), by plotting the beach volumes from before 1990 against V_b (figure 2.4). A negative trendline indicates erosion, whereas a positive trendline indicates sedimentation.

The volume of sand for beach nourishment (*V*) can be calculated as:

$$V = e \cdot l - R \qquad (1)$$

where:

e = structural erosion, a downward trendline of autonomous erosion for at least 10 years (as shown in figure 2.3) (m^3/y)

l = desired lifetime for nourishment (y)

R = sand reserve, volume of sand surplus after extraction of V_b-level (m^3) (Roelse 2002).

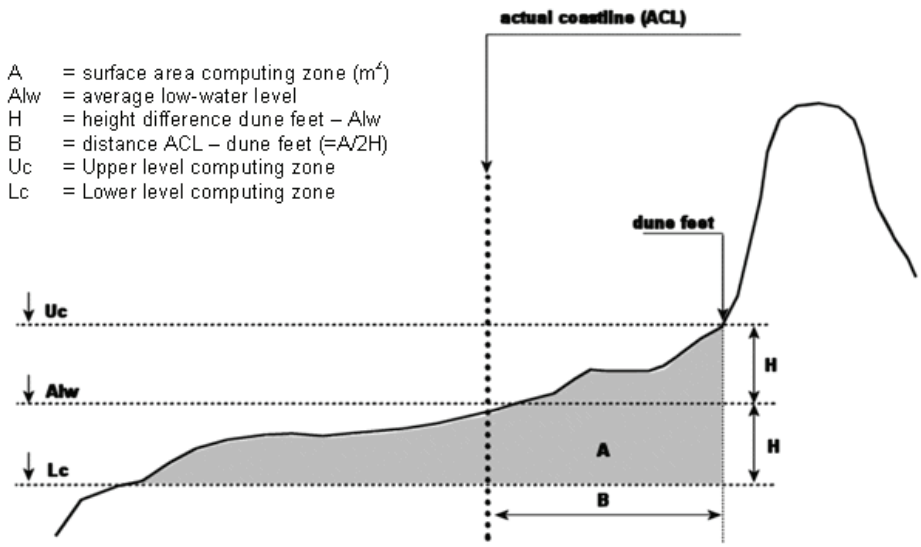


Figure 2.3 *Cross-section of a coastal area illustrating the estimation of the actual coastline (ACL) and the surface area of the beach.*

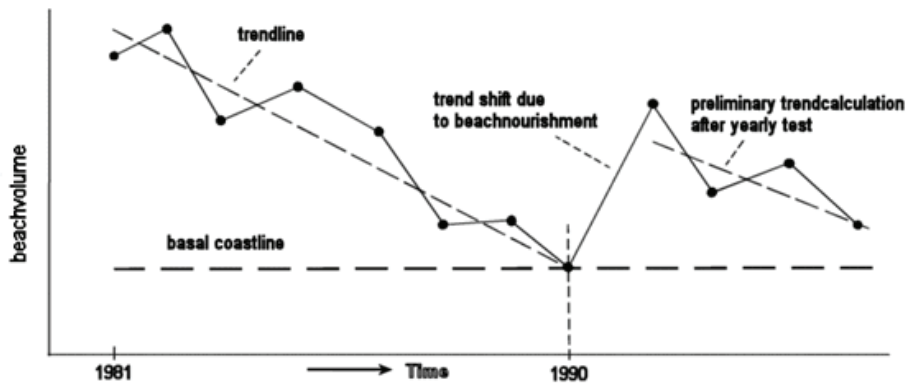


Figure 2.4 *Calculation of erosion trendline by plotting beach volumes against time. The beach volume at basal coastline serves as standard for preservation of the coastline. A trend shift will occur in case of beach nourishment.*

Traditionally, beach nourishments are carried out by dump trucks restricted within coast compartments (figure 2.5). Nowadays, with navigational systems, dump trucks can precisely nourish beach areas and are not restricted to compartment limits.

Figure 2.5 *Execution of beach nourishments in November 2001.*

To decide if a beach area is suitable for nourishment, two constraints need to be fulfilled:
C_1: a beach area shows structural erosion,
C_2: the volume for beach nourishment should exceed 200 x 10^3 m^3.
Constraint C_2 is a soft constraint, as nourishment may be carried out, depending on local and regional policies.

Chapter 3

An application of problem and product ontologies for the revision of beach nourishments

'There is a science which studies being as being'

Aristotle (384-322 BC)

This chapter is based on the following paper:

Van de Vlag, D.E., Vasseur, B., Stein, A., Jeansoulin, R. (2005), An Application of Problem and Product Ontologies for the Revision of Beach Nourishments, *International Journal of Geographical Information Science* ***Vol. 19 (10)****, pp. 1057-1072.*

3.1 Introduction

Geographic Information Systems present a model that reflects observable real world properties (Frank and Mark 1991). Such a model is in particular useful if it systematically corresponds with the observed reality (Frank 1997). To be able to interpret a GIS however, a user needs a reasoning model. Such a reasoning model is called an 'application ontology' if it is based on datasets. 'Ontology' (from Greek 'ὄντως' = 'really') is the study of the nature of reality. An ontology concerns objects, relations, states, events and processes in space (Chandrasekaran *et al.* 1998) (Kuhn 2001). The ontology depends both on the model-maker and on the context and may be different for different users (Guarino and Giaretta 1995). In fact, the degree of conceptualization, which underlies the language used by a particular knowledge base, varies in dependence of their purposes (Noy and McGuinness 2001).

Physical processes within natural systems can serve as a framework for an ontology within a GIS (Burrough and McDonnell 1998). Models describing these processes can be used to understand the system and to predict processes in the future. External factors, such as human influences and random fluctuations, can be taken into account, leading to the possibility to calculate, visualize and compare effects of different scenarios. Besides, models may improve understanding, integrate knowledge and make quantitative predictions for decision-making (Jakeman *et al.* 1995). A specific aim of combining a model with a GIS is to build a complementary quality tool for decision-support.

Coastal environments are among the most dynamic on the earth's surface (Carter 1988). Their dynamics may be the result of circulatory (in space), periodic (in time), possibly even chaotic, processes and interactions. Selection of coastal entities to be represented in a GIS is not always straightforward (Carter 1988). It depends upon user requirements and their qualities. When coastal entities are to be used for decision making, as is described in this study, metadata delivered by the data producers helps in understanding the selection of entities. Metadata alone are not sufficient to answer if the data 'fits for use' (Jeansoulin and Wilson 2002). The European IST-FET project '*REV!GIS*' has developed a methodology for the interpretation of the quality of data into a 'fitness for use' information (Jeansoulin and Wilson 2002). This methodology is based upon an ontological approach as a conceptual framework to capture human and data knowledge in an application. It encloses the user reasoning into a 'virtual dataset', whereby competency questions and rules lead to a list of real world objects.

The comparison between objects and their quality elements in this 'virtual dataset' as well as the 'actual dataset' determines its fitness for use.

Selection of coastal entities can be done, using new, more accurate spatial datasets that are becoming available. Revision or fusion of the newly created datasets with the existing datasets may improve the product quality in the end (Edwards and Jeansoulin 2004). Revision, on the one hand, aims at adding the new data to the existing dataset, by removing as little as possible from the existing data, such that the newly created dataset is consistent (Roy 2001). This implies a direct dependency between the existing and the new dataset. Fusion, on the other hand, combines two or more independent datasets, excluding the possibility of conflicts (Wilson 2002).

This chapter focuses on beach nourishments at Ameland in the northern Netherlands (see chapter 2). The dataset is described in table 2.1. The aim of this study is to build an ontology for natural coastal entities to be analyzed in relation to beach management activities. To do so, digital elevation models and satellite images (landsatTM) from 1980-2000 are combined and interpreted. Here, the abilities of newly available data to revise current beach management practices are utilized. Existing relations between datasets (the product ontology) and user theory (the problem ontology) are analyzed.

3.2 Methods

3.2.1 Ontology

The ontological approach chosen in this chapter is to handle the underlying data management problem as an integration of both data and semantics, within a common reasoning framework (Jeansoulin and Wilson 2002). The ontological approach clarifies the structure of knowledge, and leads to coherent knowledge base. An ontology exists of objects, their attributes and relationships (that may be time dependent), events, processes and states. Ontologies may enable knowledge sharing and reuse for different domains or time intervals. Finally, ontology can represent beliefs, goals, hypotheses and predictions about a domain, in addition to facts (Chandrasekaran *et al.* 1999). The use of ontologies for beach nourishments are proper for the domain of coastal areas in the Netherlands, however it is not necessarily universally valid.

The ontological approach favors a separation between the data as a 'product output', being the producer viewpoint, or as a 'problem input', the user viewpoint. Producers produce maps, not always with a particular problem in mind. A satellite-image is not produced for a particular problem, however it can be used for many purposes. Hence two ontologies apply: a 'problem ontology' and a 'product ontology' (see figure 3.1). To determine if a dataset fits for use, a comparison between the problem ontology and the product ontology need to be made.

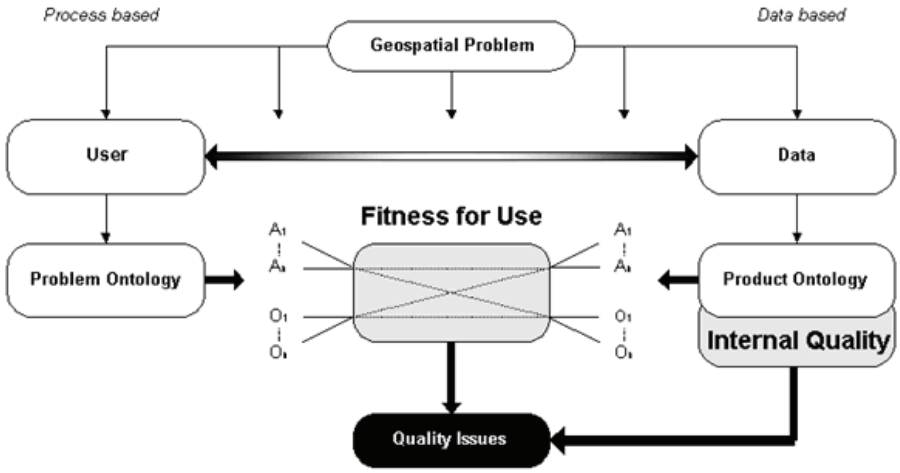


Figure 3.1 *An ontology for a geospatial problem. It separates the expected data as a problem input and the actual data as product output. The fitness-for-use issue is the gap between the expected data en the actual data. The internal quality issue results from data production methods.*

3.2.1.1 Problem ontology

The scope of an ontology is a list of competency questions that a knowledge base based on the ontology should be able to answer (Gruninger and Fox 1995). This leads to the 'problem ontology'. A problem ontology describes a reasoning model to interpret parts of the real world, based on the concepts that the user wants to represent. The problem ontology includes rules on how these concepts are understood to solve the problem. In this application the problems are defined as, (1) how to localize beach objects that require nourishment, both in space and in content and (2) to assist the decision maker to manage the process of nourishment in time.

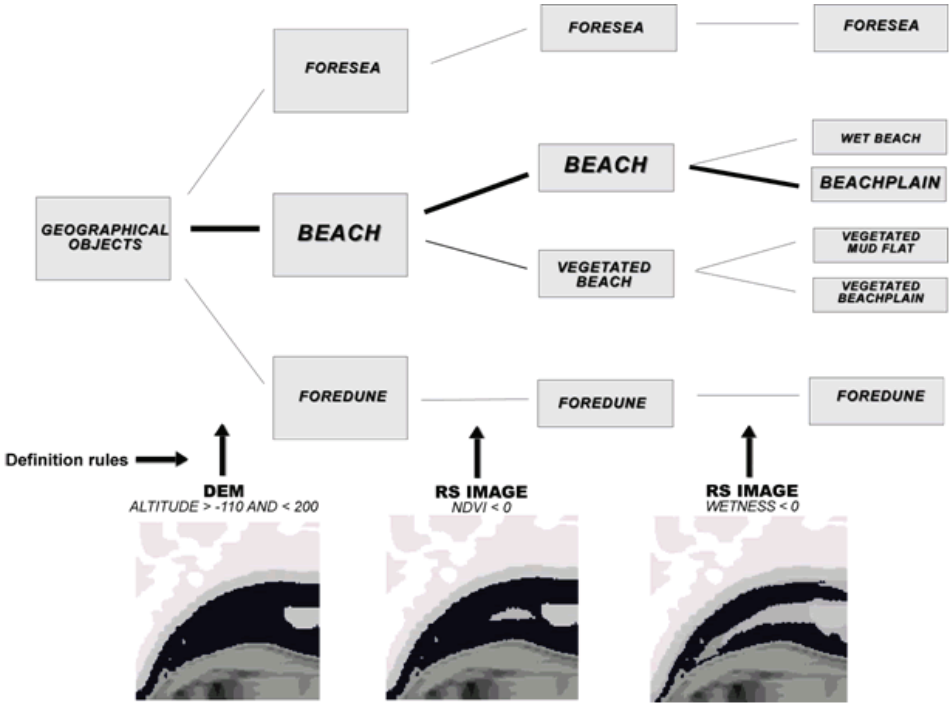


Figure 3.2 *Hierarchical classification method for localizing beach objects (beachplains) suitable for nourishment. Classification for a beachplain in 1995 (in black) is draped over a digital elevation model, with light to dark grayscales for low to high altitudes.*

Consider a set $\{A_i\}_{i=1,\ldots,n}$ of objects that require beach nourishment. To localize A-objects, two scenarios are considered, S_I based upon current nourishment practices and S_{II} – a revised method – whereby nourishment treatment is derived from the abilities of different data sources. For both scenarios A-objects need to satisfy the constraints C_1 and C_2 for beach nourishment. For S_I, A-objects are determined by means of altitude. S_{II} considers altitude, vegetation index and wetness index for object extraction, as shown in figure 3.2. Altitude distinguishes beach objects from foresea and foredune objects. A further distinction can be made on the basis of vegetation cover and wetness to identify those beachplain objects that are of interest for beach nourishment. For example, heavy equipment applied for beach nourishments may be stuck in mudflats or may annihilate pioneer vegetation on vegetated beaches. Therefore, some objects are unsuitable for beach nourishment, despite their altitude. Unlike S_I, structural erosion in S_{II} is not determined for each compartment as an object, but for individual pixels.

Accordingly, a problem ontology can be build, with competency questions, rules and a list of real world objects (see table 3.1). Objects, attributes and relationships derived from the problem ontology are displayed in table 3.2a.

Table 3.1 *Table defines the problem ontology for both scenarios into competency questions, rules, and list of real world objects for a beach management problem.*

	Comp. questions	Rules	List of RW objects
S_I	What is a beach?	Beach: -1.1 < Z < 2 m	Beach
	What is structural erosion?	10 yr downward trendline	Compartment
	What is nourishment?	C_1: Structural Erosion, and C_2: > 200 x 10^3 m^3 beach	A-object (compartment)
	Adjacent compartments?	Compartment boundaries share boundary	A-object (compartment)
S_{II}	What is a beach?	Beach: -1.1 < Z < 2 m Non-veg: NDVI < 0 Dry: WI < 0	Beachplain
	What is structural erosion?	10 yr downward trendline	A-object (pixel based)
	What is nourishment?	C_1: Structural Erosion, and C_2: > 200 x 10^3 m^3 beach	A-object (pixel based)

3.2.1.2 Product ontology

Product ontology is a list of real world objects that can be extracted by scientific and independent methods. The ontology of spatial datasets by itself is the so-called product ontology and is shown in table 3.2b.

Table 3.2a *Objects, attributes, relationships, resolution and date of the problem ontology for both scenarios. (NR=not relevant)*

Problem	Objects	Attributes	Relationships	Resolution	Date
S_I	A-objects (compartment based)	altitude	Adjacency of compartments	30 m	1995
S_{II}	A-objects (pixel based)	altitude vegetation index wetness index	NR	30 m	1995

Table 3.2b *Objects, attributes, relationships, resolution and date of the product ontology for both scenarios. (NR=not relevant)*

Product	Objects	Attributes	Relationships	Resolution	Date
DEM	None (continuous map)	altitude	NR	30 m	1995
Landsat	None (continuous, multispectral image)	vegetation index wetness index	NR	30 m	1995

For S_I, only the digital elevation model for calculating beach objects is required, as altitude is the only attribute. For S_{II} the digital elevation model is combined with Landsat images, as altitude, vegetation index and wetness index are required. For non-vegetated dry beachplains two conditions are applied, based on a normalized vegetation index (< 0) and a wetness index (< 0). Hence, a distinction is made in beachplains subjected to structural erosion or sedimentation. This is done, by calculating a local erosion or sedimentation trend for each pixel within the last 10 years.

3.2.2 Spatial Data Quality Issues

An important issue in ontological studies concerns data quality (Vasseur *et al*. 2003). Two spatial data quality concepts are identified (Aalders and Morrison 1998). Internal quality considers the internal characteristics of a dataset as a result of data production methods, whereas, external data quality is the correspondence between data characteristics and the user needs, also known as 'fitness-for-use' (Chrisman 1983; Juran *et al*. 1974; Veregin 1999). Fitness for use depends upon the objective and is therefore application-dependent. The producer of a data set usually provides internal data quality information, while fitness for use is assessed when evaluating the use of a data set (Brassel *et al*. 1995).

Quality issues are suggested for both internal and external quality by standardization (ISO 2000; CEN/TC-287; FGDC 2000) and in the scientific literature (Bédard and Vallière 1995; Wang and Strong 1996; Guptill and Morrison 1995). In this study, the focus is on thematic accuracy as a critical quality issue for the coastal application. It will be considered in relation to the problem and product ontology.

In the coastal movement application, thematic accuracy is specified by the correctness of beach object classification and the attribute accuracy. Attribute accuracy is the ability to measure attributes accurately for a well-defined feature and can be determined using error matrices for a nominal case, or with mathematical analysis (e.g. Gaussian error model) for interval/ratio data (Goodchild 1995). The correctness of beach object classification depends on the vagueness of the attributes, in existence or in content. Furthermore, classification correctness also depends on the vagueness of classification definitions. These definitions can be extracted from the problem ontology.

3.3 Results

3.3.1 Product Ontology

For S_I, objects suitable for beach nourishment are classified to identify compartments, by calculating the shortest distance to the profiles for each pixel (figure 3.3). For each compartment the overall structural erosion trend within the last 10 years is calculated. Figure 3.4 (left) shows 3 objects suitable for beach nourishment $\{A_i\}_{i=1,\ldots,3}$ based on the compartmental structure. When calculating the volumetric difference between 1995 and 1990 only the bottom left object however has serious erosion, equal to 46 x 10^3 m^3. Consequently, this object does not comply with constraint C_2 and hence beach nourishment will not be carried out.

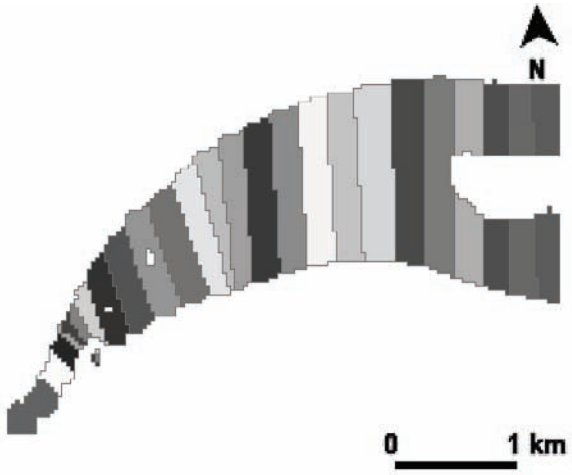


Figure 3.3 *Division in compartments for the beach area in 1995 for S_I.*

For S_{II}, a hierarchical classification method is applied to detect different types of beach objects, according to figure 3.2. Figure 3.5 shows the result of classification for different beach objects for 1995. Figure 3.4 (right) shows 4 objects suitable for beach nourishment $\{\tilde{A}_i\}_{i=1,\ldots,4}$ after classifying beach objects. Only the 2 largest objects however have serious structural erosion: 30 x 10^3 m^3 for the bottom left object and 59 x 10^3 m^3 for the central object. Both objects do not comply with the minimum volume constraint (C_2), and beach nourishment will not be carried out.

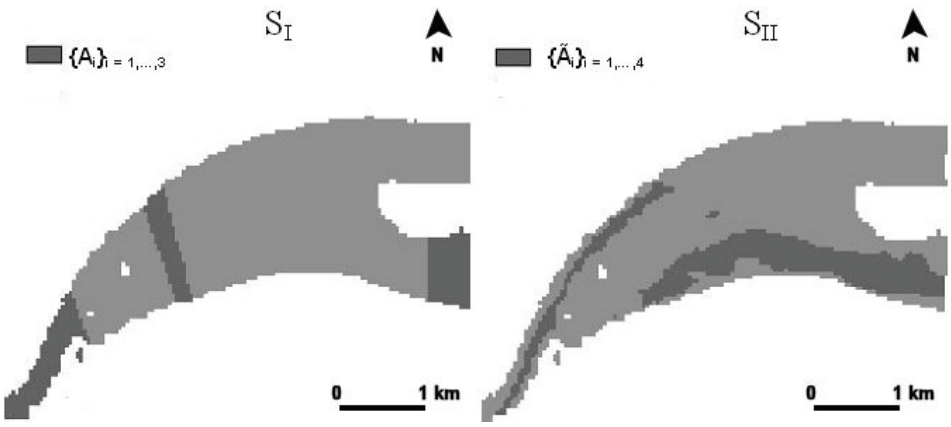


Figure 3.4 *Images visualizing sets of objects suitable for beach nourishment $\{A_i\}_{i=1,\ldots,n}$ and sets of objects non suitable for beach nourishment $\{\neg A_i\}_{i=1,\ldots,m}$ in 1995 for both scenarios.*

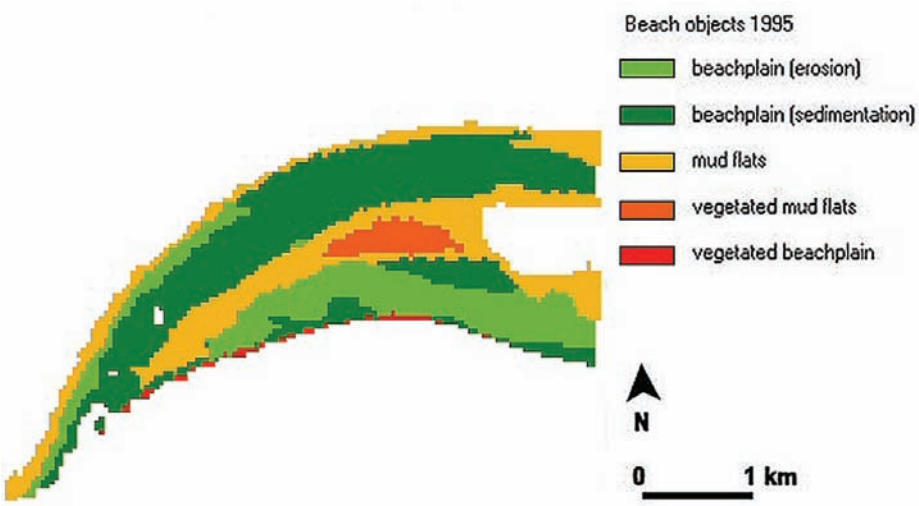


Figure 3.5 *Classification of beach objects in 1995, using methodology from S_{II}.*

3.3.2 Scenario I vs. Scenario II

The two scenarios show different images (figure 3.4). Dissimilarity is a direct consequence of methodological differences. S_I is based on the user theory with compartmental structural erosion, whereas S_{II} is the result of revision, using the benefits of additional data to predict beach objects. Table 3.3 proves the dissimilarity in an error matrix of

pixels classified to beach objects suitable for beach nourishment (A-objects) and objects non suitable for beach nourishment (¬A-objects). Only 5.5% of the total amount of beach pixels (n=237) is in both images classified as A-objects. The overall accuracy of A- and ¬A-objects is 72.8%. From figure 3.4 it appears that the main overlap is in the bottom left part. Actual beach nourishments have been carried out in 1997 for object A_1 in SI (figure 3.4). S_{II} shows another significant object in the central part of the picture. As there may be significant erosion, beach nourishments will not be carried out in that area. No direct threat for safety risks occurs. An onshore movement of a shoal just north of this object exists, provoking sedimentation seawards of this object.

Table 3.3 *Error matrix of classified pixels (n) derived from S_I (row wise) and S_{II} (column wise) for 1995.*

	$\{\tilde{A}_i\}_{i = 1,\ldots,4}$	$\{\neg\tilde{A}_i\}_{i = 1,\ldots,m}$	Total
$\{A_i\}_{i = 1,\ldots,3}$	237	432	669
$\{\neg A_i\}_{i = 1,\ldots,m}$	734	2890	3624
Total	971	3332	4293

3.3.3 Problem vs. Product Ontology and their Quality Issues

Different objects may be identified by S_I and S_{II} as defined by the problem ontology. Discrepancies between problem and product ontologies can be partly explained by spatial data quality issues (table 3.4). The positional accuracy is determined by the absolute measurement accuracy of the X, Y and Z coordinates, the conversion from point measurements to a grid, the geometry of the compartments (for S_I) and the positional accuracy of multiple data sources (for S_{II}). Thematic accuracy is defined by the interpolation method and classification correctness of the attributes. For S_I, this is done by the Z-value, for S_{II}, by a combination of the Z-value, the NDVI and wetness index. Time of data capture of the different attributes affects the temporal accuracy. With time dependent attributes and fusion of multiple data sources, temporal inaccuracies influence calculations by S_{II}. Logical inconsistencies are identified by the different conceptual data structures used for both scenarios: compartments for S_I, and pixels for S_{II}. For both scenarios, data completeness occurs if the grid covers the area of interest. In the foreshore area, this is not the case, as profile measurements do not continue to the outer edge of the study area and some satellite-images have missing values. Model completeness is mainly

determined by the omission of key attributes and the incompleteness of the trendline with multi-temporal beach volumes.

Table 3.4 *Overview of spatial data quality issues for both scenarios.*

Quality Issue	Scenario I	Scenario II
Positional accuracy	X Y Z	X Y Z
	Geometry of compartments	Fusion of data sources
	Grid conversion	Grid conversion
Thematic accuracy	Z interpolation	Z interpolation
	Z classification	Z classification
		Attribute accuracy of NDVI and wetness index
		Classification of NDVI and wetness index
Temporal accuracy	Data capture	Data capture
		Time dependency of NDVI and wetness index
Logical consistency	Compartment based	Pixel based
Completeness	Data completeness (DEM)	Data completeness (DEM and satellite images)
	Completeness of trendline	Completeness of trendline
	Geometry of compartments	

3.3.4 Attribute accuracy assessment

For attribute accuracy, a sensitivity analysis for the interpolated Z value was carried out within the upper and lower limit of the cross-validated RMSE for both scenarios. For 1995, the cross-validated RMSE is 0.28 m for half width of 95 percent confidence interval. A decrease of the Z value corresponds with an increasing overlap of A-objects from 5.5% to 12.5% in both scenarios (table 3.5). With lower altitude, structural erosion levels are more severe, resulting in more objects suitable for beach nourishment. An increase of the Z value (table 3.6) however, shows a decrease in overlap of A-objects down to 1.5 % in both scenarios. As concern nourishment volumes, the decrease of the Z value result in an increase of the volumes of A-objects. For S_I, the volume of $\{A_i\}$ increases to 220 x 10^3 m^3. For SII, the volume of $\{\tilde{A}_i\}$ for the 2 largest objects increases to 78 x 10^3 m^3 for the bottom left object and to 270 x 10^3 m^3 for the central object. Hence, the volumes of the A-object in SI and the central Ã-objects in

S_{II} exceed the constraint for the minimal volume of beach nourishment (C_2). An increase of the Z-value reduces the volumes to 3 x 10^3 m^3 in S_I, and to 48 x 10^3 m^3 for the central object in S_{II}. Altitude therefore affects the geometry and volumes of the different beach objects.

Table 3.5 *Error matrix of classified pixels (n) derived from S_I (row wise) and S_{II} (column wise) for 1995 with Z-value minus RMSE.*

	$\{\tilde{A}_i\}_{i = 1,\ldots,4}$	$\{\neg\tilde{A}_i\}_{i = 1,\ldots,m}$	Total
$\{A_i\}_{i = 1,\ldots,3}$	535	935	1470
$\{\neg A_i\}_{i = 1,\ldots,m}$	746	2077	2823
Total	1281	3012	4293

Table 3.6 *Error matrix of classified pixels (n) derived from S_I (row wise) and S_{II} (column wise) for 1995 with Z-value plus RMSE.*

	$\{\tilde{A}_i\}_{i = 1,\ldots,4}$	$\{\neg\tilde{A}_i\}_{i = 1,\ldots,m}$	Total
$\{A_i\}_{i = 1,\ldots,3}$	64	199	263
$\{\neg A_i\}_{i = 1,\ldots,m}$	439	3591	4030
Total	503	3789	4293

3.4 Discussion

In this chapter 2 ontologies are considered to describe objects suitable for nourishment (A-objects) and objects non-suitable for beach nourishment (¬A-objects). A problem ontology is used as a 'virtual dataset', as it is visualized in the mind of the user (Jeansoulin and Wilson 2002). A product ontology makes use of spatial datasets. To determine issues of data quality, data producers often define quality of a product as being consistent with specifications, while data users define it as meeting or exceeding their expectations (Kahn and Strong 1998). Hence, to solve the 'fitness for use' question, users have to define the level of acceptability for each quality issue for each object and relation. For 'internal quality' attribute inaccuracy is a deciding quality issue. Vagueness of the concept and the definitions for beach nourishment has an effect that A-objects can only be identified with a limited level of certainty. This corresponds with previous research by Cheng (1999 and 2002), who found that most geographical objects are naturally indeterminate or vague by means of their inherent characteristics (e.g. continuity, heterogeneity, dynamics or scale dependency).

By choosing two scenarios, a first approach is made, in assigning each geospatial problem to a continuum between process based ontologies and data based ontologies. Both scenarios deal with assigning geospatial objects suitable for beach nourishment (A-objects). The methodologies to achieve this purpose are different for both scenarios. S_I is process based, based on user requirements and it utilizes only the digital elevation models. S_{II} is data based, rested on the abilities of the dataset, utilizing digital elevation models and satellite images. This is also shown in figure 3.1, where different routes between the geospatial problem and the modeling phase illustrate the different possibilities to reach a final model of a geospatial problem. The discrepancies between the results of S_I and S_{II} are not only caused by spatial data quality issues, but also the result of ontological differences. The results differ due to the context of describing the problem. In a compartment based approach (S_I) each pixel has no direct influence on the outcome, as structural erosion is averaged over the compartment. In a pixel based approach (S_{II}), each pixel and it neighbors have a significant influence on the final results for structural erosion, and on non-vegetated, dry areas. Contextual approaches, e.g. Bayesian methods, can support in illuminating the contribution of neighborhood effects.

The ontology approach and accuracy assessment show elucidating perspectives of the spatial extent of A-objects. First, overlap in A-objects occurs primarily where actual beach nourishments are carried out. Second, sensitivity for accuracy assessment exists, showing that determinate boundaries for A-objects are not logical. Consequently, A-objects may be treated as vague objects, meaning that their spatial extent is not fixed. Accordingly for A-objects, uncertainty plays an important role at three definition levels (after Molenaar 1998):

- the existential uncertainty expresses how sure we are that an A-object really exist,
- the extensional uncertainty reflects the spatial extent of an A-object,
- the geometric uncertainty refers to the precision with which the boundary of an A-object can be determined.

The vagueness of the different attributes as altitude, structural erosion, wetness index and vegetation index may have a major contribution to the extensional uncertainty.

The use of temporal geospatial data implies the possibility for temporal inaccuracies. Temporal accuracy can act upon the accuracy of time measurement of the data, the temporal validity of the data or the temporal consistency of the data. Temporal uncertainty in the beach nourishment application is related to the time of data capture.

An example is the assignment of the normalized vegetation index to vegetated beach objects (NDVI > 0) from satellite images. Here, lower NDVI values are detected in spring, resulting in no evidence of pioneer vegetation at the mudflats. In summertime, this evidence is certainly present. Also, tide effects and weather influences, related to the time of data capture, influence the accuracy of the wetness index significantly. Further research will be carried out to study the effects of temporal uncertainty on beach nourishments.

3.5 Conclusions

An application of ontologies to determine objects for beach nourishment, demonstrates elucidating aspects about the spatial distribution of these objects. First, a distinction into 2 scenarios shows differences between a process-based ontology and a data-based ontology. For 1995, overall accuracy for objects suitable and non-suitable for beach nourishment is 72.8 %, with the largest overlap in a region where actual beach nourishments have occurred. Discrepancies between outcomes can be accredited to ontological differences between both scenarios. Secondly, the ontological approach shows that inaccuracies of the attributes influence object determination. The sensitivity analysis, as applied on the attribute altitude, illustrates a significant increase of objects suitable for nourishment for both scenarios, when altitude is decreased within the lower limit of the cross-validated RMSE for 95 percent confidence interval. Here, the constraint for minimal volume of beach nourishment is exceeded, meaning that Dutch authorities need to act in the near future. Sensitivity of altitude shows that determinate boundaries for beach nourishment objects are not sensible and consequently should be treated as vague objects. The study showed that the use of ontologies, as applied in a beach management application, can differentiate between dataset and user theory. It also showed how it can be helpful in understanding if data will fit for use. To define competency questions, rules and list of real world objects first, appropriate datasets can be selected that are needed to solve the problems.

Chapter 4

Modeling Dynamic Beach Objects Using Spatio-temporal Ontologies

'When one admits that nothing is certain one must also admit that some things are much more nearly certain than others'

Bertrand Russell (1872-1970)

This chapter is based on the following papers:

Van de Vlag, D.E. and Stein, A. (in review). Landsat TM images to define Spatio-Temporal Ontologies for Beach Nourishments. *Journal of Environmental Informatics.*

Van de Vlag, D.E. (2004). Concepts and Representation of Beach Nourishments by Spatio-temporal Ontologies. In: The Netherlands and the North Sea. Dutch Geography 2000-2004, Nederlandse Geografische Studies 325, T. Dietz, P. Hoekstra, F. Thissen (eds). Utrecht, The Netherlands, pp 45-55.

Van de Vlag, D.E., Stein, A., Vasseur, B. (2004). Concepts and Representation of Beach Nourishments by Spatio-temporal Ontologies. ISSDQ symposium, Vienna, 15-17 April 2004. In Proceedings of the ISSDQ '04, *GeoInfo* **28(b),** pp. 353-368.

4.1 Introduction

Landscape objects are often dynamic. Typical examples include beaches and dunes under coastal morphodynamics (Eleveld 1999), agricultural fields in a rural-urban environment changing under population pressure (Marceau *et al.* 2001) and soil and geomorphological objects under shifting cultivation in Cameroon (Yemefack 2005). Spatio-temporal datasets may be helpful for monitoring these dynamic landscape objects. Multi-source geo-information obtained using satellite imagery is increasingly used to provide these sets. Satellite images supply basic measurements of biological and physical characteristics of the landscape objects, such as their position, shape, elevation, color, temperature and moisture content (Wilkie and Finn 1996). A combination of images obtained at several instances may be useful for monitoring purposes. In this respect, remote sensing can assist policy makers, resource users, and resource managers.

A typical aspect of landscape features is that they are frequently vague, both in their definition and in their spatial extent (Fisher *et al.* 2005). The main reasons are that both context and definitions are poor. Also, the objects are often delineated by conceptual ideas rather than by actual and quantifiable spatial extent.

Combination of different data sources can be difficult, because of different resolutions, spectral decomposition and sensor characteristics. To achieve information exchange between different data sources, the study of ontologies may unify different conceptualizations of geographical space into one geographical ontology (Kokla and Kavouras 2001). An ontology can be defined as an explicit specification of a conceptualization (Gruber 1993). To do so, classes of objects are defined, including their relations and functions. Ontologies are hence characteristic for specific domains.

In the past, ontologies have proven to be useful for handling real world features within a geographical information system (Jeansoulin and Wilson 2002; Van de Vlag *et al.* 2005). For each specific domain, ontologies identify and define a set of relevant concepts that characterize a given application domain. In this sense, domain ontologies reduce conceptual and terminological confusion. They also support interoperability and knowledge sharing within various government organizations (Jeansoulin and Wilson 2002). Here, decision makers demand certain data (Foody and Atkinson 2002), whereas uncertainty is inherent in spatial information. In recent scientific research on spatial information, uncertainty – and spatial data quality in general – are emphasized (Shi *et al.* 2003; Frank and

Grum 2004; Van Oort 2006). Ontologies help to understand the role of the quality of the data sources as well as their fitness for use by decision makers (Hunter 2001). So far, ontologies have been primarily applied in the spatial domain. For analyzing and understanding dynamic geographical problems, however, a spatio-temporal ontology, i.e. an ontology representing space and time, is essential (Frank 2003a; Frank 2003b).

The aim of this chapter is to define and use a spatio-temporal ontology for modeling dynamic landscape features. Such an ontology allows us to integrate different multitemporal data sources. Its definition is based on an extension of spatial ontologies with temporal issues, such as processes and events. This spatial-temporal ontology is applied to a beach management problem in the Netherlands, where due to erosion, beach nourishments have to be carried out. The three methods are compared. First, beach areas suitable for nourishment are represented and modeled by crisp objects. Second, they are represented and modeled by fuzzy sets. Third, inclusion of a time series allows us to represent beach processes and events. A sensitivity analysis in the end validates choices in spatial and temporal fuzzification.

4.2 Methodology

To understand characteristics of landscape features in time, an appropriate conceptualization of a spatio-temporal dataset is needed. To do so, landscape features are modeled by monitoring activities (figure 4.1). Five operations are distinguished:

1. An ontological approach, which serves as a guideline for an explicit conceptualization of the landscape features.
2. Implementation of fuzzy rules, to describe spatial and thematic vagueness.
3. Time correction, to correct for temporal variation of the attributes.
4. A quality matrix, to relate ontological features with several quality elements (e.g. positional-, thematic-, and temporal accuracy, completeness, etc.).
5. Model testing, by means of a sensitivity analysis.

In the following sections, a more detail description on each of these five operations is given.

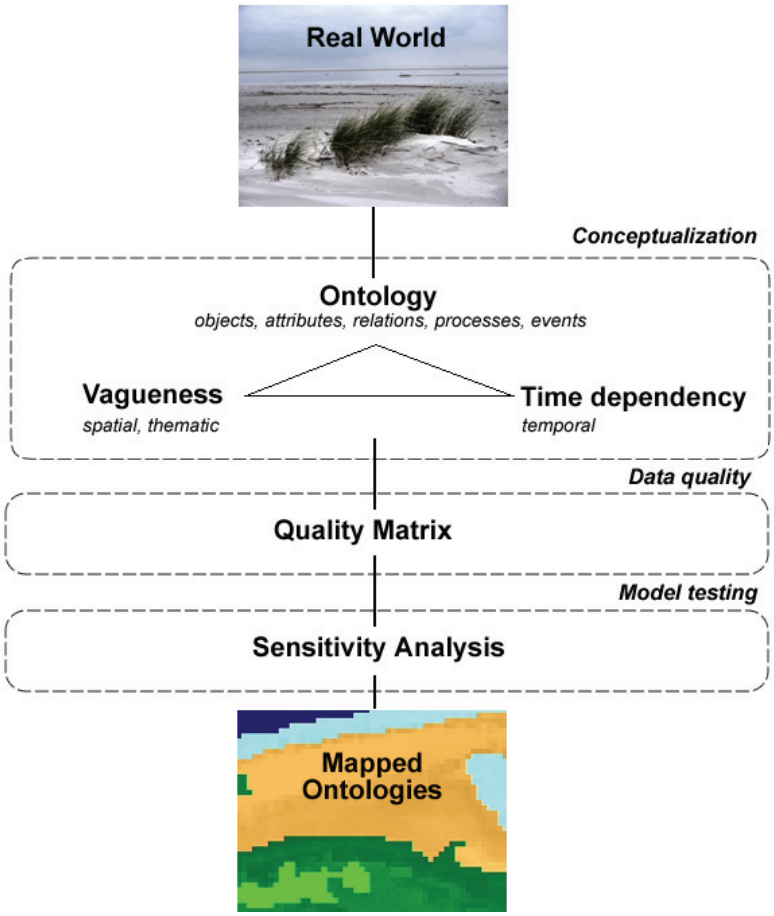


Figure 4.1 *Modeling of geospatial dynamics, from features in the real world to objects in the model.*

4.2.1 Ontology

A common reasoning framework clarifies the structure of knowledge, and leads to coherent knowledge base (Jeansoulin and Wilson 2002). An ontology describes "the metaphysical study of the nature of being and existence" (Frank 2003b). It describes a conceptualization of the real world and is closely related to software engineering activities like conceptual analysis and domain modeling (Guarino 1998). An ontology determines what is independent of an observer. This includes physical reality, as the position of an object in Cartesian space, but also human agreements, e.g. classification rules or social arrangements. At the heart of Aristotle's ontology is a theory of

‘substances’ (things, or bodies) and ‘accidents’ (qualities, events, processes) (Smith 2001).

The management of dynamic landscape features is addressed in an ontological approach as an integration of data with semantics. For modeling dynamic landscape features, a general ontology for objects in the real world is proposed, derived from monitoring activities. Objects are characterized by their ‘substances’ and ‘accidents’, whereby ‘substances’ are attributes (*att*), which are dependent in space (*x,y*), value (*v*) and time (*t*), and have relationships between other objects (*rel*). ‘Accidents’ are related to the spatio-temporal behavior of the datasets, and include events (*ev*) and processes (*proc*). Hence, an ontology exists of objects, their attributes and relationships, events, processes and states (Chandrasekaran *et al.* 1999; Kuhn 2001). For an application in landscape monitoring, it is possible to define:

$$object \in S_{obj}(att_{\alpha}, proc_{\beta}, ev_{\gamma}) \tag{1}$$

where S_{obj} is the landscape object conceptualized by an ontology and where α, β and γ point to an index set.

Equation 1 shows that each particular object is characterized by a specific set of attributes, processes and events. For continuous attributes with index set α it follows that:

$$att_{\alpha} = f_{att}(x, y, v, t) \tag{2}$$

Equation (2) illustrates a field model, where each point in space and time for different properties can be observed. Ontological features (objects, attributes, relations, processes and events) specify the ontology for an application. Processes are time dependent and require a special set of attributes. The value of the process is not only dependent on the attribute value, but also on situations in the past, trend, expectations and random noise. Events are special processes, i.e. they occur suddenly and may change attribute values and processes.

4.2.2 Implementation of Fuzzy Rules

Landscape units are by nature vague in their content and extent (Frank 2003b; Fisher *et al.* 2005). For vague objects three types of uncertainty exists (Molenaar 1998):

- existential uncertainty expresses how sure we are that an object really exist,

- extensional uncertainty reflects the spatial extent of an object,
- geometric uncertainty refers to the precision with which the boundary of an object can be determined.

In this chapter the focus is on existential and extensional uncertainty. To do so, fuzzy rules are implemented using membership functions to describe the spatial and thematic uncertainty of the variables that describe objects. These membership functions are based on semantic import models for each attribute, with a value 0 as no membership, a value 1 as full membership, and a value in between as partial membership. The field model is adapted for attributes from equation 2 to:

$$att_{\alpha} = mv_{\alpha} \cdot f_{att}(x, y, v, t) \tag{3}$$

For index set a, a membership function mv_a is included to express the spatial (x,y), thematic (v) and temporal (t) uncertainty of the attributes.

4.2.3 Time Correction

Landscapes can be observed in space, time and theme (Peuquet 1994). Since dynamic landscape features can not be monitored in real time, but only in a representation of reality, the temporal effects need to be abstracted to create snapshots of the world. This can be described by the following equation, which Goodchild called 'geographical reality' (Goodchild 1992):

$$att = f_{att}(x, y, v) \tag{4}$$

To analyze and understand dynamic landscape features, the spatial ontologies (see chapter 3) are extended for representing space and time (Claramunt 1997; Frank 2003a; Frank 2003b). By doing so, temporal inaccuracies should be accounted for. Temporal inaccuracy can affect the accuracy of time measurement of the data, the temporal validity of the data or the temporal consistency of the data.

This chapter manages temporal inaccuracies that occur due to temporal variability of the data. For any a at time t, correction factors ($CF_{a,t}$) are implemented to correct for this temporal uncertainty, whereby $CF_{a,t}$ defines the 'degree of certainty' for time t of data capture – with values between 0.5 and 1. This correction factor is applied on the basis of the slope of the fuzzy membership function. If $CF_{a,t}$ = 1, the temporal certainty of the data capture is high, and the membership function will be identical as equation 3. If $CF_{a,t}$ < 1, the

temporal certainty of data capture is lower and the slope of the membership function will become less steep according to:

$$att_{\alpha,t} = CF_{\alpha,t} \cdot mv_{\alpha,slope} \cdot f_{att}(x, y, v) \quad (5)$$

This correction factor only applies on the slopes of the membership functions, i.e. the location where there is no full membership or non-membership. Here, the transition zone described by partial memberships will therefore increase.

4.2.4 Quality Matrix

Fitness for use models the relation between available data and data required to analyze landscape phenomena (Hunter 2001). Herein a quality matrix, where ISO quality elements (ISO 2003) are portrayed against application features derived from the ontology, can be practical. Features and quality elements are defined first, followed by integration with ontological concepts. Each column of the quality matrix indicates an ontological feature, each row the quality elements. The relevant cells of the matrix link – when applicable – a quality value to a feature.

A quality matrix may assist a decision maker, for example to find the best available dataset for an application. It comprehensibly formalizes spatio-temporal problems, taking into account the objects, fuzzy rules and integration of temporal dimensions, and current standards of quality.

4.2.5 Model testing

Establishing the usefulness of models as a means of improving our understanding of predicting geospatial dynamics requires the use of objective measures of performance (Gardner and Urban 2003). Testing occurs throughout the stages of model development by means of measures on the model structure, parameter sensitivity, model adequacy and hypothesis testing. A quantitative model that exhibits large fluctuations in output for relatively small changes in the value of some input parameter is sensitive to the parameter, whereas a model which exhibits small output variation for substantial perturbations is insensitive to the parameter.

The beach nourishment model is tested using a sensitivity analysis, which determines the model response to any realistic set of parameter perturbations (Rose and Swartzman 1981). Therefore, a probabilistic sensitivity analysis is applied that reflects the likely value of an uncertain parameter, based on the probability distribution.

Bounds are selected as a confidence interval for all possible values of the parameter.

4.3 Beach Nourishment Application

4.3.1 Revision of beach nourishment

In this chapter, the dataset for 1995 is used as illustration for modeling beach nourishments (see section 2.3) For beach management purposes, Rijkswaterstaat - the directorate-general of the Ministry of Public Works responsible for the maintenance of the coast - divides the beach area into compartments. Compartments are the regions between two transects (see section 2.4). Each compartment has two boundaries to its adjacent compartment (CL_1 and CL_2), a beach-sea boundary (*BS*) and a beach-dune boundary (*BD*). The beach objects are thus structured into compartments within a higher conceptual level, based on perceivable regions on the beach. Traditionally, beach nourishments are carried out by dump trucks restricted within these compartments.

So far, volumes of sand are calculated as crisp objects on the basis of these compartments. The interest of this application is to take into account the 'fuzzy' nature and the 'dynamics' of these objects. Besides, with assistance of navigational systems, dump trucks can precisely nourish beach areas and are not restricted to compartment limits.

Identification of areas that require beach nourishment depends upon terrain elevation (height around zero), vegetation index (non-vegetated zones) and wetness index (dry zones). Elevations between –1.1 and 2 m at Dutch standard sea-level (NAP) are considered as beach areas. Such areas are derived from digital elevation models (DEMs). Similarly, non-vegetated and dry zones are derived from Landsat TM imagery, using respectively Normalized Difference Vegetation Index (NDVI) (Rouse *et al.* 1974) and wetness index (Crist and Cicone 1984). NDVI is a spectral transformation that is applied to bands 3 and 4 of the Landsat image for the purpose of assessing the health and vigor of vegetated surfaces. NDVI has the advantage of simplicity and common use by remote sensing analysts (Dunham *et al.* 2005). The wetness index is a tasseled cap transformation that provides excellent information for land use applications because it allows the separation of barren (bright) soils from vegetated and wet soils. Non-vegetated zones are selected as areas with negative NDVI values; dry zones are selected as areas with a wetness index lower than zero. The delineation of the object beach should satisfy the

constraints for elevation, non-vegetated and dry zones, e.g. the so-called beachplain (Van de Vlag *et al.* 2005).

4.3.2 Modeling beach nourishment

Beach nourishment is modeled in three different ways. Firstly, a crisp approach considers sharp boundaries between different beach objects corresponding to a traditional approach. Secondly, a fuzzy approach describes compartments as membership functions of three parameters. Thirdly, temporal membership functions are included for modeling temporal processes.

4.3.2.1 Crisp compartmental method (CC)

Identification of beach areas that require beach nourishment requires compartments (*C*) to distinguish zones with sedimentation from those with erosion. These compartments are perpendicular to the coast, 100-200 m wide, and fixed in time. Compartment width is a fixed limit (*CL.geo*), Compartment length on the other hand, being the distance between the beach-sea boundary (*BS.geo*) and the beach-dune boundary (*BD.geo*), is fuzzy, as both boundaries are dynamic (*BS.geo(t)*, *BD.geo(t)*), due to erosion and sedimentation. For the crisp compartmental (CC) approach, these boundaries are assumed to be sharp.

On the basis of trendline calculation the decision-maker decides if sedimentation and erosion occur within a compartment. On account of beach volumes, the amount of erosion and sedimentation is calculated For *np* pixels each of size *ps* the beach volume within a compartment (*C.vol(t)*) is calculated as:

$$C.vol(t) = ps^2 \times \sum_{i=1}^{np} e(i,t) \tag{6}$$

where *e(i,t)* is the elevation of pixel *i* at time *t*, i.e. the difference between the actual elevation measurement and the lower computing boundary (see section 2.4). A compartment is indicated for nourishment, when the beach volume in a compartment shows structural erosion, e.g. obeying the two constraints (C_1 and C_2).

4.3.2.2 Fuzzy compartmental method (FC)

Spatial and thematic uncertainty of the attributes is modeled using fuzzy logic. Beach compartments suitable for nourishment are identified by their memberships to dry, non-vegetated beaches. Hence, a compartment is bound by two static compartment

boundaries (*CL.geo*) and by two fuzzy boundaries: the sea-beach boundary (*BS.geo(t)*) and the beach-dune boundary (*BD.geo(t)*). These boundaries are illustrated in figure 4.2.

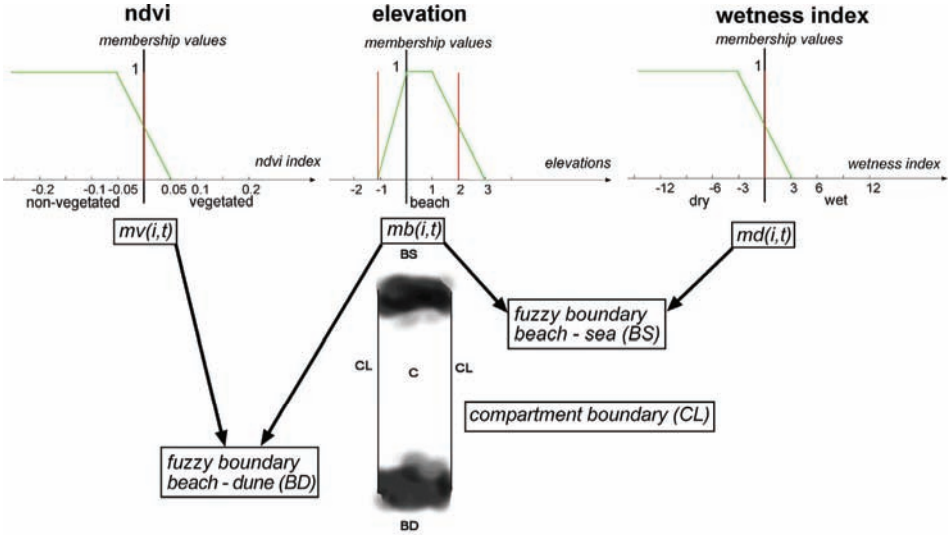


Figure 4.2 *Compartment, boundaries and their various fuzzy membership functions. The lower image visualizes a compartment (C), with two adjacent crisp boundaries (CL) and two fuzzy boundaries (BS) and (BD).*

The sand volume within the fuzzy compartmental (FC) method can be calculated, using:

$$C.vol(t) = ps \times \sum_{i=1}^{np} m(i,t) \times e(i,t) \tag{7}$$

where *m(i,t)* equals the membership value of location *i* in compartment *C* at time *t*. It is calculated as:

$$m(i,t) = \min\{mb(i,t), md(i,t), mv(i,t)\} \tag{8}$$

where *mb(i,t)* is the membership function of the beach object, *md(i,t)* that of dry object and *mv(i,t)* that of a non-vegetated object in which pixel *i* occurs at time *t*. Membership functions are compiled as triangular functions. The *mb(i,t)* equals 1 if elevation ranges from 0 to 1 m amsl, it increases linearly from 0 to 1 between –1.1 to 0 m amsl and decreases linearly from 1 to 0 between 1 and 3 m amsl, and it equals 0 elsewhere. The soil is wet if the wetness index is positive, and dry if it is negative. A membership function *md(i,t)* for fuzzification equals 1 if wetness index is less than -3, and decrease

linearly from 1 to 0 for the wetness index moving from -3 to 3 m, and it equals 0 elsewhere. Land is covered with vegetation if the NDVI is positive and is bare if it is negative. A membership function *mv(i,t)* equals 1 if the NDVI value is less than -0.05, it equals 0 if the NDVI is larger than 0.05 and it decrease linearly from 1 to 0 in between.

4.3.2.3 Temporal fuzzy compartmental method (TFC)

To include temporal uncertainty into the beach nourishment processes, daily fluctuations for the wetness index, monthly fluctuations for the vegetation index and yearly fluctuations for elevation are considered. Three temporal membership functions are introduced to reflect the appropriate time scale for these attributes. In the Netherlands, the growing season for vegetation starts early March and reaches maximal values in June. Hence, in June the certainty to detect healthy fully-grown vegetation and thus large NDVI values is high. The temporal membership function *nv(t)* corresponds to the growing season and equals 1 between 1 June and 1 August, it equals 0.5 between 1 November until 1 March, and it is linear in between. Similarly for soil wetness, the temporal membership function *nd(t)* corresponds to tide fluctuations and equals 1 during flood time and equals 0.5 during low tide, and further follows a sine form, i.e. $nd(t) = 0.75+0.25 \cdot \cos(2\pi \cdot t/12.5)$, with t expressed in hours in relation to high tide. Finally for elevation, the temporal membership function *nb(t)* is constant and describes the actual digital elevation model.

The temporal fuzzy approach (TFC) first calculated correction factors (*CFs*) derived from introducing date and time of data capture into the temporal membership functions. Next, the spatial membership functions originate from the fuzzy compartmental (FC) method *mb(i,t)*, *md(i,t)* and *mv(i,t)* are corrected with this correction factor (figure 4.3). Low *CFs* lead to membership functions that are less steep; *CFs* close to 1 result in membership functions similar to those in *mb(i,t)*, *md(i,t)* and *mv(i,t)*.

The fuzzy- and temporal fuzzy approach differ from each other, as the slope of the membership function is corrected according to the temporal (un)certainty of the vegetation- and wetness index. At high tide, the certainty that a beach is wet or dry is high, while at low tide this temporal certainty is low. Similarly, in summer the certainty about vegetation is high, while during growing season this certainty is lower. For 1995, the capturing date of the Landsat image is 7 November, at 9h30, just after low tide corresponding with a low temporal certainty for vegetation and wetness. The image is

radiometric corrected to correct for atmospheric conditions that may affect NDVI and wetness index values.

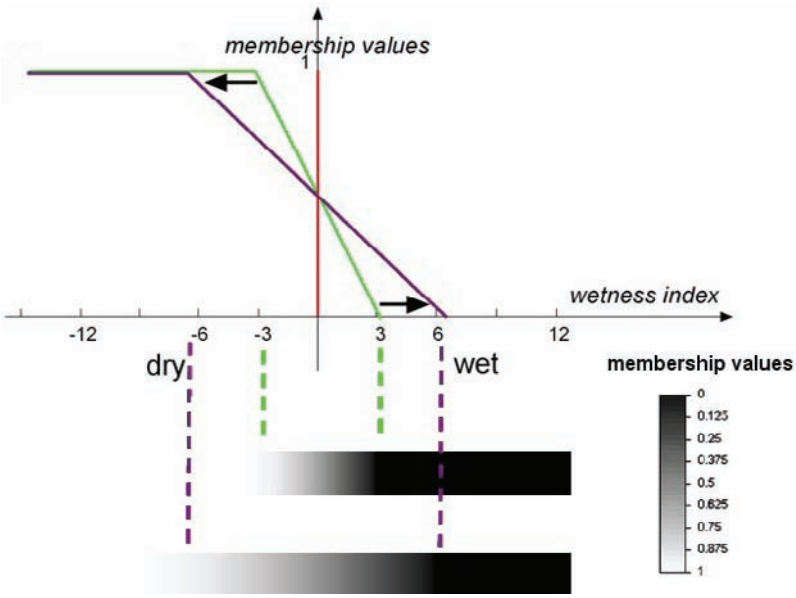


Figure 4.3 *The slope of the membership function for dry beach areas is corrected for tide influences. In this case, the vagueness of the boundary area will increase.*

4.3.3 Quality elements and quality matrix

Before constructing a quality matrix, spatial data quality elements need to be defined for the beach nourishment application. For spatial uncertainty, the following ISO quality elements are considered important (ISO 2003):

- Positional accuracy, whereby sub-elements of interest are: 1) relative or internal positional accuracy, i.e. closeness of the relative positions of objects in a dataset to their respective relative positions accepted as or being true; 2) gridded data position, i.e. closeness of gridded data position values to values accepted as or being true.
- Thematic accuracy, with sub-elements of interest: 1) the accuracy of quantitative attributes, i.e. the correctness of quantitative attributes and of the classifications of objects and their relationships; 2) classification correctness, i.e. comparison of the classes assigned to objects or their attributes to a universe of discourse (e.g. ground truth or reference dataset).

- Temporal uncertainty of the compartments can be recognized by temporal accuracy. In particular the sub-element for accuracy of a time measurement, i.e. correctness of the temporal references of an item (reporting of error in time measurement).
- Completeness; this is determined by sub-element data completeness, i.e. the commission and omission of datasets.

By applying an ontological approach, a quality matrix is constructed, projecting ontological features such as objects, attributes, relationships, processes and events against the ISO quality elements.

4.3.4 Sensitivity Analysis

By means of a sensitivity analysis, the effects of spatial and temporal influences of attributes are determined to the amount of beach nourishment. The sensitivity of calculated beach nourishment volumes is based on statistical properties, i.e. the variance of the attribute elevation, NDVI and wetness. The sensitivity of the output is also investigated on statistical properties of the tide cycle and vegetation growth.

For NDVI and wetness, sensitivity analysis has been executed with 1 standard deviation range from the dividing value 0. A positive NDVI value corresponds, as at least theoretically, with presence of vegetation and a negative value with absence of vegetation, whereas positive and negative wetness values correspond to wet and dry, respectively.

For elevation, the cross-validated RMSE are observed by interpolating the elevation profiles to grid maps. Here, 5% of the data points are used for cross validation, to obtain an accuracy value of the Z value. For 1995, the cross-validated RMSE is 0.28 m for half width of 95% confidence interval. The sensitivity analysis for elevation is carried out within the upper and lower limit of the cross-validated RMSE for the interpolated Z values, i.e. from Z – 0.28 m as the lower limit to Z + 0.28 m as the upper limit.

For the temporal fuzzy compartmental (TFC) method choices in the correction factor (*CF*) of the temporal membership functions were evaluated. These choices represent correction factors values that are in the vicinity of these choices, with minimum of 0.5 and maximum of 1. A value equal to 0 would be highly unlikely. Therefore, the correction factor (*CF*) in the membership function for the tidal cycles is varied between 0.3 and 0.7, whereas its highest value is kept equal

to 1 when perfect timing of data capture is observed. Also, the correction factor in the membership function for the vegetation growth cycle varies between 0.3 and 0.7, with the highest value equal to 1.

4.4 Results

4.4.1 Comparison of the approaches

The methodology used in this chapter construct in three phases a spatio-temporal ontology for the beach nourishment application. The crisp compartmental method (CC) defines beach areas suitable for nourishment by crisp boundaries of the compartment. The object of interest is the compartment (*C.id*), whereby the sand volume calculation (*C.vol*) is grounded on the minimum intersection of the attributes elevation, vegetation and wetness. The fuzzy compartmental method (FC) defines beach areas suitable for nourishments by fuzzy boundaries of the compartment. The sand volume calculation is then based on the minimum intersection of the membership functions of the attributes elevation, vegetation and wetness. For the temporal fuzzy compartmental method (TFC), the sand volume calculation is corrected for temporal uncertainty and is the minimum intersection of the corrected membership functions of the attributes.

Results are presented in figure 4.4, showing the beachplain as the beach area that is dry and non-vegetated. The slope of the membership functions for vegetation is corrected for temporal uncertainty, and is less steep, resulting in a fuzzy beachplain (figure 4.4).

After trendline calculation, two regions are indicated as areas with structural erosion (figure 4.5). The southwest region is of interest, as most actual beach nourishments are actively carried out in that area. Moreover, the threat for safety for the public is higher, as there is only a single row of dunes to protect the hinterland. The calculation for beach volumes, using the compartmental methods, is represented in table 4.1. The fuzzy approach indicates lower volumes, as thematic uncertainty, as well as temporal uncertainty is accounted for. The volumes in table 4.1 are below criteria for nourishment (200,000 m^3), and beach nourishment in 1995 will not be carried out.

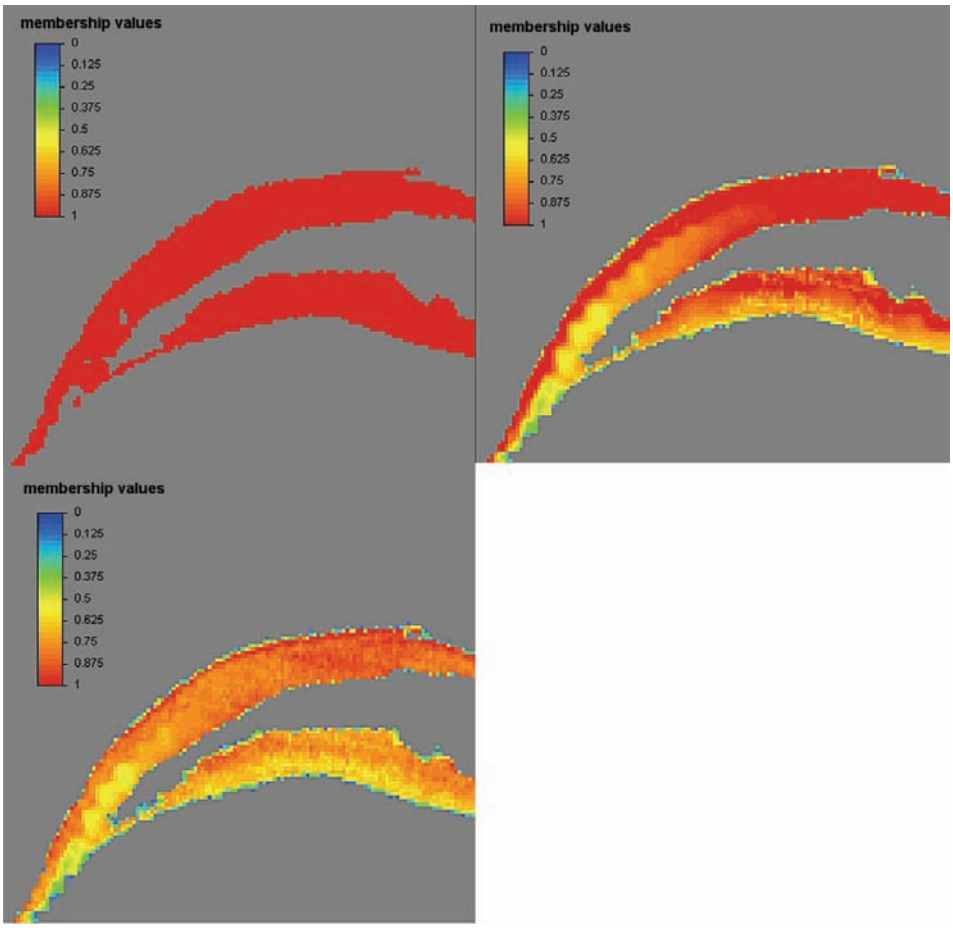


Figure 4.4 *Beachplain classification, using a crisp (top left), fuzzy (top right) and temporal fuzzy (bottom) method.*

Table 4.1 *Beach nourishment volumes using the crisp compartmental (CC), the fuzzy compartmental (FC) and the temporal fuzzy compartmental (TFC) approach.*

	CC	FC	TFC
Volume (m^3)	78,600	58,700	51,600

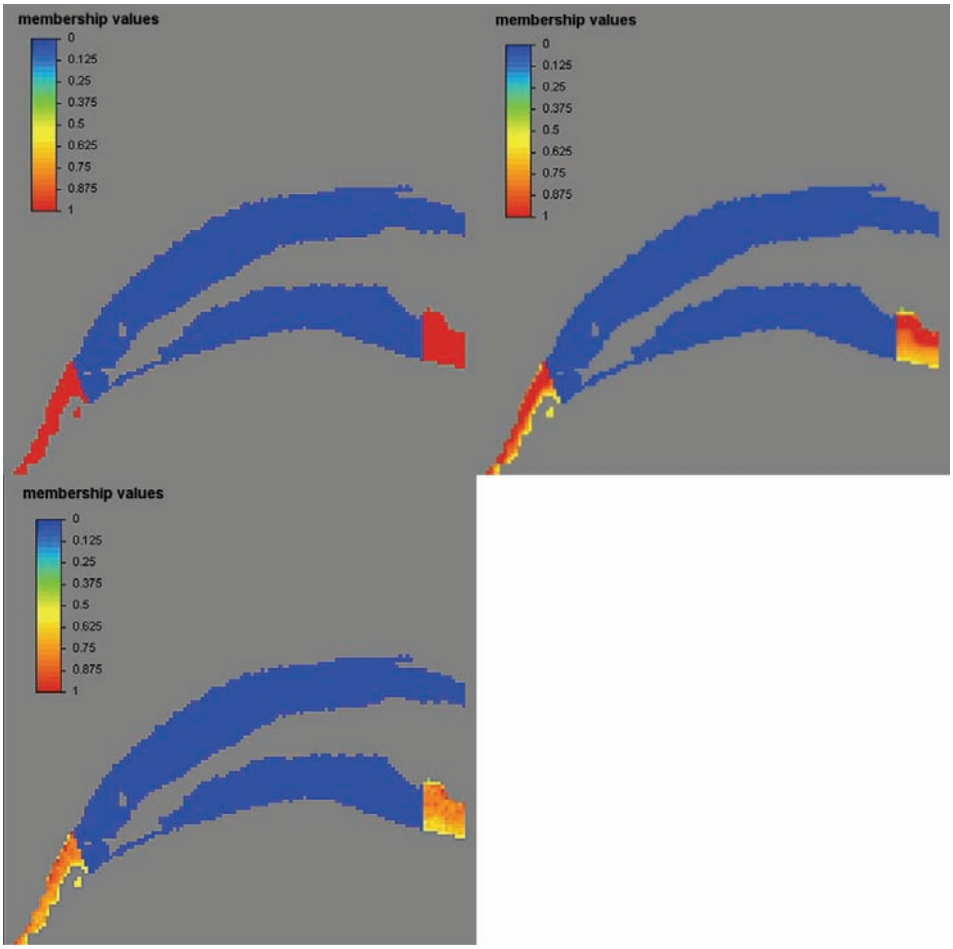


Figure 4.5 *Final result of CC (top left), FC (top right) and TFC (bottom) method. Legend is membership values, with 0 is no membership and 1 is full membership.*

4.4.2 Quality Elements

Table 4.2 describes the quality of objects and attributes that applies to the case study. The objects consist of compartment (*C.id*) and its boundaries (*CL.id*, *BD.id*, *BS.id*). The attributes consider elevation (*BD.z*, *BS.z*), wetness index (*BS.wi*), vegetation index (*BD.ndvi*) and structural erosion (*C.se*, *C.vol90*). The prominent feature of interest is the amount of beach volume, represented by *C.vol* in table 4.2. Row wise, the quality elements are described. Positional accuracy is represented by relative and gridded data position. Different membership functions related to corresponding attributes occur for classification correctness and temporal accuracy. The quantitative

attribute accuracy is only valid for the attributes considering elevation for 1995 (*BD.z*, *BS.z*, *C.vol*). The data completeness is illustrated in the last column and equals 86.7% for the 1995 data set.

Table 4.2 *Quality elements for the ontological features for 1995. Abbreviations: ClasCor = classification correctness, QAA = quantitative attribute accuracy, ATM = accuracy of time measurement, NR = not relevant.*

	Positional Accuracy		Thematic Accuracy		Temporal Accuracy	Completeness
Objects	*Rel.*	*Grid.*	*ClasCor*	*QAA*	*ATM*	*Data*
C.id	48.6 m	30.3 m			< 1 year	86.7%
CL.id	NR	30.3 m			< 1 year	86.7%
BD.id	48.6 m	30.3 m			< 1 year	86.7%
BS.id	48.6 m	30.3 m			< 1 year	86.7%
Attributes						
C.vol	48.6 m	30.3 m	m(i,t)	± 0.28 m	CF•m(i,t)	86.7%
C.vol90	48.6 m	30.3 m	m(i,t)	NR	CF•m(i,t)	NR
BD.ndvi	48.6 m	30.3 m	mv(i,t)	NR	nv(t)	86.7%
BD.z	NR	30.3 m	mb(i,t)	± 0.28 m	nb(t)	86.7%
BS.wi	48.6 m	30.3 m	md(i,t)	NR	nd(t)	86.7%
BS.z	NR	30.3 m	mb(i,t)	± 0.28 m	nb(t)	86.7%
C.se	48.6 m	30.3 m	m(i,t)	NR	CF•m(i,t)	86.7%
…	…	…	…	…	…	…

The three different procedures lead to different quality assessments. The crisp compartment method results in statements on objects that have a positional accuracy. The fuzzy compartment method describes both the positional and the thematic accuracy. Thematic accuracy is described by membership functions that were included from the semantic import model. Finally, the temporal fuzzy compartment method includes statements on positional, thematic and temporal aspects. Note that the temporal membership functions were also based on a semantic import model (see section 4.3.2.3).

4.4.3 Sensitivity Analysis

Beach nourishment volumes are affected by choices for elevation (figure 4.6a). Generally, an increase is observed with increasing elevation, which is most prominent for the CC method, and less prominent for the FC and TFC methods, with values ranging from 30,000 to 100,000 m^3. Hence, interpolation uncertainties may have a large influence on the calculated beach volumes. Similarly, the volumes are sensitive for the choice of the dividing NDVI value, with a sharp jump around the NDVI value of 0 (figure 4.6b). The FC and TFC methods give almost similar results, whereas the jump around the NDVI value of 0 for the CC method is even more prominent. An explanation is that a small change in positive NDVI values signifies

pioneering vegetation as vegetation, decreasing the sand volumes for beach nourishments. Also, the standard deviation is relatively large, as NDVI values equal to 0 for beach and sea, and have positive values for dunes. For wetness, a dependence on the choice of equilibrium (i.e. wetness index equals 0) is observed, with a decrease of the volume of beach nourishment with increasing wetness value (figure 4.6c).

Interestingly, the large sensitivity of the results obtained for the spatial variables does not lead to modifying the assumptions made in the space-time membership function. The volume of beach nourishment shows changes in the range of 51,400 to 51,700 m^3 for modifying assumptions on the tidal membership function, and values from 47,200 to 55,100 m^3 for the assumptions in the vegetation growth cycle (figure 4.6d). It demonstrates that the model is rather insensitive to these assumptions.

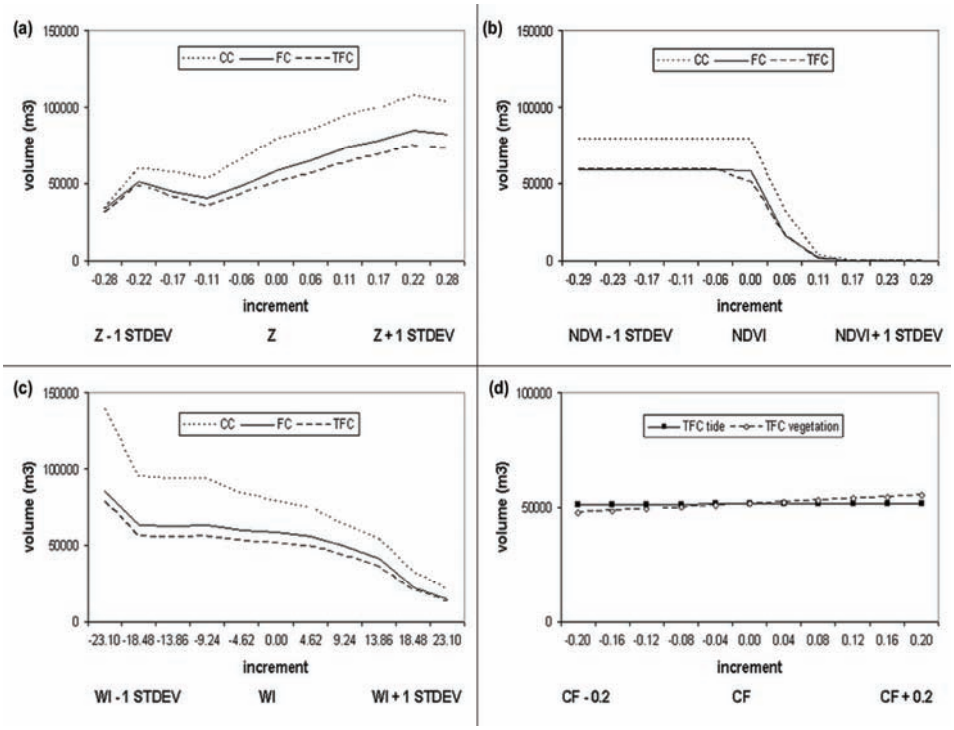


Figure 4.6 *Results of sensitivity analysis using CC, FC and TFC method for (a) elevation, (b) NDVI and (c) wetness index. For temporal variables (d) only the TFC method is applied.*

4.5 Discussion

In this chapter, an ontological approach is used to describe beach objects for nourishment. This approach concerns space, theme and time issues for a dynamic geographic application. The use of identical ontologies over time can help to detect and determine objects that change between 'snapshots'. On the one hand it gives a good understanding of the underlying spatio-temporal problem, i.e. the nourishment of beach areas in space and time, on the other hand the it does not deal with different conceptual levels. With higher resolutions different attributes, processes and events to describe beach objects might be involved, leading to different spatio-temporal ontologies. Also, the selection of quality elements depends upon the spatial scale of the dataset.

The selection of quality elements essential for the study area is done by using previous studies (Cheng *et al.* 2001; Van de Vlag *et al.* 2005). The accuracy to describe objects requiring beach nourishment is important. Positional elements concern elevation, and a choice for boundaries between adjacent objects. The elevation measurements from the DONAR database are accurate and precise and therefore affect positional accuracy to a lesser extent. Positional accuracy is in particular determined by the grid resolution of the DEM and the geometric correction of the satellite imagery.

Thematic accuracy concerns determination of a classification using quantitative figures. Here, it is chosen to apply semantic import models for implementation of fuzzy rules and these are based on expert knowledge, therefore this may be subjective and influenced by spatial variation.

In this chapter, temporal accuracy is modeled by seasonal factors for vegetation, by tidal fluctuations for the wetness index, and by both erosion/sedimentation processes and by incidental changes for elevation. Seasonal factors may be difficult to relate to vegetation, as other factors, such as moisture content and vegetation stress, contribute as well. Also relations between tidal effects and wetness are subject to factors like vegetation and weather. Currently, weather influences are not considered, as these are complicated to observe and difficult to model due to several time dimensions, e.g. influences like large storm events or wind erosion are measured on different temporal scales. Finally, processes on sedimentation and erosion are applicable to larger areas of land and events like severe drought and large storms may interfere.

Causes of uncertainty in objects to determine suitability for beach nourishment lie in natural variation of the object describing variables. Together with a poor definition of context and definition rules for nourishment, this requires objects to be treated as vague objects. Membership functions represent their thematic and temporal uncertainty. In this chapter, temporal uncertainty is treated as a part of thematic uncertainty and their interrelationship is joined inextricable. This is expressed by using temporal (un)certainty by correction of the transition zones of the membership functions for thematic uncertainty.

A realistic geospatial model to fully describe beach features may be complicated. Relationships between ontological features occur at multiple conceptual levels. Quality elements for each of these features contribute to the overall quality of such a model. Therefore, a hierarchical structure may lead to an improved description of ontological features, as well as its quality elements. Scale issues, in time and space, leads to a different conceptualization of a spatio-temporal problem and therefore a different ontology and quality elements. A hierarchical representation of a model – from a knowledge based or Bayesian perspective - can assist a decision maker in selecting the best dataset for the ontology and quality elements in mind.

The sensitivity analysis as carried out here highlights the effects of choices for spatial variables, and much less so for choices in the temporal membership functions. It appears, that at least for this chapter, temporal effects can relatively simply be modeled, and that any subjective choice has much less influence than choices in spatial values.

The approach used in this chapter can be applied to other spatio-temporal datasets, like monitoring areas sensitivity for bushfires. Here, the ontology considers several datasets for conceptualization, whereby these areas are described by weather conditions, flame characteristics, fuel characteristics, elevation, etc. All these attributes concerns their own space, theme and time issues, and have their influence on the quality of the dataset.

In a recent study about the history of coastal protection, Charlier *et al.* (2005) state that beach nourishments is still favored with many instances. Decision makers prefer commonly profile feeding by simple deposition of material on the beach. Beach nourishments may act as both protection and restoration projects and has been implemented in all continents. Accordingly, many lessons have been learned that

encompass decision choices of feeding material, preservation of source material sites, wave climate, etc. Several researchers in the Netherlands examined these decision choices. Van Rijn (1997) studied the sediment transport and budget along the central coastal zone of the Netherlands. He developed a mathematical model to calculate sand volume changes, based on hydrodynamic (waves and currents) and sand transport processes. Van der Wal (2000) examined the side-effect of beach nourishments on the rate of aeolian sand transport in the foredune area of central Ameland and concluded that grain size and adaptation length, which is a measure of the distance over which sediment transport adapts to a new equilibrium condition, both affected the topography of the beach-foredune area. Van Noortwijk and Peerbolte (2000) and Van Vuren *et al.* (2004) have studied extensively the optimal economical aspects with respect to the recurrence interval of beach nourishments in the Netherlands. In the approach as described in this chapter, the ontologies are grounded on geographical choices to identify beach objects suitable for nourishment. The categorization of these objects, as well as the implementation of fuzzy set techniques and landscape dynamics, precedes the actual beach nourishment process.

4.6 Conclusions

This chapter exploits the use of spatio-temporal ontologies for describing beach objects suitable for beach nourishments. The spatio-temporal ontology includes fuzziness to describe vagueness of landscape features. The fuzzy approach is more realistic, as the described objects are more similar to those occurring in decision procedures and to the use in ontologies. The spatio-temporal ontology also includes important temporal aspects by means of temporal membership functions. Hence, the spatio-temporal ontology is therefore more generally applicable to space-time studies.

The case study focuses on beach nourishment at the Isle of Ameland. It is shown that the three methods can be implemented and applied to determine the required amount. The three approaches resulted in different amounts of beach nourishment volumes, ranging from 51600 to 78,600 m^3. By means of a quality matrix, ontological features (i.e. objects, attributes, relationships) are projected against their quality elements. Further steps are identified to quantitatively evaluate different quality elements, such as positional, temporal and thematic accuracy. A sensitivity analysis showed that interpolation uncertainty for elevation and choices in dividing values between absence and presence of vegetation and between wet and dry soils may have a large influence on the calculated beach volumes. The

influence of choices for the temporal membership function however is weaker.

Chapter 5

Incorporating Uncertainty via Hierarchical Classification using Fuzzy Decision Trees

'It is the theory which decides what can be observed'

Albert Einstein (1879-1955)

This chapter is based on the following papers:

Van de Vlag, D.E. and Stein, A. (in review). Incorporating uncertainty via hierarchical classification using fuzzy decision trees. *IEEE Transactions on Geoscience and Remote Sensing.*

Van de Vlag, D.E. and Stein, A. (2005). Uncertainty propagation in beach classification using fuzzy decision trees. ISSDQ '05 Conference Proceedings, Beijing China.

5.1 Introduction

Acquisition of geospatial data by means of remote sensing is based on a specific view of real-world objects. Such objects can be tangible features, like trees or rocks, or they can be more abstract, like a landscape feature or an altitude measurement (Bennett and Armstrong 2001). This view affects the way that geographic data and their uncertainties are collected and how they can be used for further analysis (Peuquet 2002; Fisher 1999). To classify a landscape, with its features and processes, it is of utmost importance to consider the level of scale, e.g. local, regional, national, etc. For instance, a mudflat object that is identified on local scale, is part of a beach object on regional scale and classified as coastal object at national level. This hierarchy is also found in the structure of geographic data, meaning that objects classified on their parameters are in turn derived from other parameters. As such hierarchical levels in object classification are distinguished.

Several classification techniques exist that incorporate uncertainties. Decision trees classify objects on the basis of decisions using a hierarchical structure and are capable of incorporating expert knowledge in the decision rules (Pearl 1988; Debeljak *et al.* 2001). Artificial neural networks (ANNs) classify objects when there are sufficient training data and if little knowledge exists about relationships in the data (Rojas 1991). Dempster-Shafer's theory of evidence (Dempster 1967; Shafer 1976) is a knowledge-driven approach that introduces the logic of 'ignorance' as a means to evaluate and control classification uncertainty (Lein 2003). According to this method, a level of 'belief' will form, based on the evidence available to support a conclusion as objects are assigned to the categories of interest.

In this chapter, uncertainties in decision rules are illustrated in a hierarchical way. Therefore, the use of decision trees for classification is considered. A decision tree classifier has a decision rule at each level, involving several combinations of the attributes. Application of the rule results either into a leaf, allocating an object to a class, or a new decision node, specifying a further decision rule. In this way, a decision tree is suitable for a hierarchical classification. Upon application, a full classification of a series of objects is done by moving down the tree until a leaf is reached (Quinlan 1986).

So far, decision trees result into a crisp discretization. Crisp decisions work well if class boundaries are non-overlapping and clearly defined. For geographic applications one may expect objects to gradually vary in attribute values (Burrough 1996). Geographic objects may have

fuzzy boundaries, indicating within-class heterogeneity (Ricotta and Avena 1999). Such objects require the use of fuzzy decision trees. These are based on fuzzy set theory (Zadeh 1965), follow similar steps as crisp decision tree, but allow to deal with characteristics of geographical objects (Janikow 1998; Suárez and Lutsko 1999).

The aim of this chapter is to define and use fuzzy decision trees for object classification. The study is illustrated with a coastal management application in the northern part of the Netherlands. Here, beach areas are affected by strong tide currents, causing beach erosion that is counteracted by beach nourishments. These areas are therefore dynamic. The objective is to classify these objects taking into account the decision making process of nourishments as well as its uncertainties.

5.2 Concepts and methods

5.2.1 Categorization of geographical objects

Real-world objects represented in geographical information systems require classification. How the classification is done depends to a large extent upon perception of the physical environment and the processes that take place in it. A decision on how to represent them is liable to scale of perception, spatial and semantic uncertainties of the objects and uncertainty of object identification.

Hierarchical models help to comprehend and illustrate the decisions and uncertainties involved. Categorization of objects using a hierarchical model has been widely used in the environmental and life sciences by means of decision tree classification (Tso and Mather 2001). In this study, hierarchical models are used to (1) determine multi-scale objects, (2) calculate uncertainties in attributes and (3) classify data based on decision rules.

5.2.2 Image-objects, segmentation and scale

Satellite imagery serves as a tool to detect real world objects. Perception of an image content can be based upon these objects (Blaschke *et al.* 2000; Hay *et al.* 2003). Once having perceived those, they are linked to larger compositions by means of experience and knowledge. Image analysis may yield meaningful objects after segmentation (Baatz and Schaepe 2000; Gorte 1998; Molenaar 1998). Nature however rarely has sharp boundaries, but it is also not a true continuum, showing sometimes soft transitions between objects. These transitions are also subject to object definition and in this way depend on scale as well.

Scale is crucial in object definition (Benz *et al.* 2004, Murwira 2003). In remote sensing scale is present in pixel resolution. A single object may appear differently at different scales. Also, geographical objects are hierarchically linked to the landscapes in which they occur (Smith and Mark 1999). They are typically complex and have different dimensions. Moreover, a landscape can be dynamic. A classification procedure is thus attractive if it implements these scale and hierarchy issues, simplifies the computations and maintains classification accuracy.

5.2.3 Bayesian hierarchical model

Hierarchical classifiers based on Bayesian modeling have the ability to deal with uncertainty. Baysian methods work with probabilities rather than absolute values and uncertainties can be explicitly included in the probability distributions and propagated through the model. In addition, the ability to incorporate both data and expert knowledge makes probability-based modeling appealing.

A Bayesian hierarchical model addresses prior probabilities of uncertainties in attributes. These prior probabilities contain parameters that are uncertain as well (Gelman *et al.* 1995). By modeling observed data and unknowns as random variables, this model provides a coherent framework for combining complex data models and expert knowledge (Banerjee *et al.* 2004). Uncertainty is quantified, and newly collected data improve the prior distribution to yield the posterior distribution.

In fact, probability statements are made about unknown vector of parameters θ, given observed data y. We note that the joint probability distribution $p(\theta,y)$ can be written as a product of the prior distribution $p(\theta)$ and the sampling distribution $p(y|\theta)$. Conditioned on the observed values of data y, using Bayes' rule, the posterior distribution equals:

$$p(\theta|y) = \frac{p(\theta,y)}{p(y)} = \frac{p(\theta)p(y|\theta)}{p(y)} \tag{1}$$

$$\propto p(\theta)p(y|\theta) \tag{2}$$

Proportionality in equation (2) arises since the marginal distribution $p(y)$ does not depend on θ. The prior probability $p(\theta)$ can be derived by experimental statistical analysis, ancillary data or assumptions. For normal data with θ equal to the vector (μ,σ), it is fully defined by the prior mean (μ_0) and prior variance (τ_o^2). The sampling distribution

p(y|θ) describes the distribution of the data, given the prior parameters. For normally distributed observations $y = (y_1,\ldots,y_n)$ with known variance (σ^2) this leads to the posterior distribution:

$$p(\theta|y_{1\ldots}, y_n) = p(\theta|\bar{y}) = N(\theta|\mu_n, \tau_n^2) \tag{3}$$

where,

$$\mu_n = \frac{\frac{1}{\tau_0^2}\mu_0 + \frac{n}{\sigma^2}\bar{y}}{\frac{1}{\tau_0^2} + \frac{n}{\sigma^2}} \tag{4}$$

and the precision (reciprocal of the variance) equals:

$$\frac{1}{\tau_n^2} = \frac{1}{\tau_0^2} + \frac{n}{\sigma^2} \tag{5}$$

Here, *n* equals the number of observations, μ_n the posterior mean and τ_n^2 the posterior variance. Hence the posterior distribution of *θ* given *y* is also normal with expectation μ_n and variance τ_n^2.

5.2.4 Hierarchical classification by decision trees

Each geographical object is completely described by a set of attributes and a class label. A decision tree classifier for geographical objects is a hierarchical structure consisting of several levels. At each level a test is applied to one or more attribute values. Attributes can have numerical and nominal values. A decision tree contains arcs, nodes and leaves. Nodes contain splits, to test the value of an expression of the attributes. Arcs from a node to its off-spring are labeled with distinct outcomes of the test. Each leaf has a class label associated with it and no further off-spring.

Decision trees generally proceed top-down. Objects are classified by moving them along the tree until a leaf is reached. The label at that leaf is its class.

Decision trees have proven to be effective for extracting and classifying objects (Quinlan 1988; Byungyong and Landgrebe 1991; Eklund *et al.* 1994, 1998; Lees and Ritman 1991; Teoh and Ma 2003; Xian *et al.* 2002), resulting into a crisp discretization. For various objects this may not suffice, however, as decisions in lower levels of a tree are based on increasingly less data; of which some may not have

any meaning (data fragmentation), and several leafs can represent the same class in case of a high class overlap,. For this reasons soft splits of data have been considered by means of fuzzy decision trees (Janikow 1998; Li *et al.*, 2003; Olaru and Wehenkel 2003; Peng and Flach 2001).

In remote sensing imagery, spectral classes have transition zones and fuzzy boundaries. Heterogeneity may exist within classes. Fuzzy decision trees recognize heterogeneity by partitioning continuous attributes into several fuzzy sets prior to classification. Partition is done heuristically, on the basis of expert experiences and data characteristics. A membership function of a fuzzy set pair is determined according to the characteristics of the attribute data.

5.2.5 Estimation of classification accuracy

Accuracy measures the correspondence between the classification and the class in a real world ontology. In this study, this is done by evaluating error matrices. An error matrix is a square matrix, with columns containing the reference data, and rows the classified data.

Such accuracy is associated with objects in a strictly Boolean fashion, i.e. the classification is either right or wrong. However, fuzzy classifications are performed with fuzzy results. Even so, the membership functions are symmetrical and linear, i.e. the overall possibility for all classes equals 1.0. Hence, the membership values for each class are summed up to create a fuzzy error matrix (Binaghi *et al.* 1999; Xu *et al.* 2005). The accuracy is determined as in the traditional error matrix for hard classification.

5.3 Case study

5.3.1 Decision making in coastal management in the Netherlands

The case addressed in this chapter considers the beach management application (see also chapter 2). Available datasets consist of a digital elevation model, a Landsat7-ETM+ image, as well as digital orthogonal images (table 5.1). All datasets are acquired within a short time period (spring to summer 2003) and are geometric corrected to reduce registration errors.

In this study, the decision is whether certain beach areas need to be nourished or not. This decision is based on the sand balance of eroded beach areas. Therefore, beach areas that require nourishment need to be localized and quantified. The beach area has been divided

into compartments for calculation of sand volumes, by the Ministry of Public Works, i.e. the decision-maker. They are treated as crisp objects, although their physical boundaries are fuzzy and gradual transitions occur between objects.

Table 5.1 *Dataset properties of the beach nourishment application. [a]Abbreviations: Res. = resolution, TDC = time of data capture, RMSE= root mean square error, DOI = digital orthogonal images.*

Dataset	*Database*	*Res.*[a]	*TDC*[a]	*RMSE*[a]
DEM	DONAR	30 m	Spring 2003	30.3 m
L7 ETM	Landsat	30 m	29 May 2003	24.1 m
DOI[a]	DKLN	0.5 m	Summer 2003	<0.5 m

Traditionally, these objects are determined on the basis of altitude, which distinguishes them from sea and fore dune objects. Satellite images make it possible to improve decision making, by including data on vegetation cover using NDVI, and wetness by means of the wetness index. This is relevant, as equipment applied for beach nourishments may get stuck in mudflats or may annihilate pioneer vegetation on vegetated beaches. Beach objects of interest are dry and non-vegetated, the so-called beachplains.

Prerequisite for successful object identification are procedures for extraction algorithms to segment images into a set of objects. Here, it is the aim to take into account the fuzzy nature of objects. In chapter 4 the boundaries between beach/sea and beach/fore dune are described as vague boundaries using altitude, vegetation cover and wetness as imprecise parameters. The aim is to also include these parameters and their uncertainties into the classification. Uncertainty is in turn derived from so-called hyperparameters that can be modeled by a Bayesian hierarchical model.

In 2001, actual beach nourishments have been carried out in the western part of the study area. This affects the altitude near the beach/fore dune boundary, as sand is dumped close to the dunefoot area. A correction is applied on the beach/fore dune boundary by excluding altitude values of boundary observations from this part of the study area. Accordingly, only altitude values of non-nourished beach areas are used to obtain a corrected sampling distribution for the beach/fore dune boundary.

5.3.2 Bayesian modeling

The role of prior uncertainty considers the gradual transitions between altitude, wetness and vegetation classes. Prior distributions for the boundaries of beach objects are given by the slopes of the membership functions described by semantic import models (see chapter 4). They are triangular functions (table 5.2), where $mb_a(x)$ is a beach object, $mnv_n(x)$ a non-vegetated object and $md_w(x)$ a dry object at location x.

Table 5.2 *Fuzzy rule set for the spatial distributions of altitude, ndvi and wetness index for dry, non-vegetated beach objects, i.e. beachplains.*

Dataset	*Attributes*	*Fuzzy rule set*
DEM	altitude (a)	$mb_a(x) = 0$ if $a \in (-\infty, -1.1]$
		$mb_a(x) = 0$ if $a \in [3, \infty)$
		$mb_a(x) = 1 + a/1.1$ if $a \in (-1.1, 0)$
		$mb_a(x) = 3/2 - a/2$ if $a \in (1,3)$
		$mb_a(x) = 1$ if $a \in [0,1]$
Landsat TM	ndvi (n)	$mnv_n(x) = 1$ if $n \in (-\infty, -0.05]$
		$mnv_n(x) = 1/2 + 10n$ if $n \in (-0.05, 0.05)$
		$mnv_n(x) = 0$ if $n \in [0.05, \infty)$
Landsat TM	wetness index (w)	$md_w(x) = 1$ if $w \in (-\infty, -3]$
		$md_w(x) = 1/2 + w/6$ if $w \in (-3,3)$
		$md_w(x) = 0$ if $w \in [3, \infty)$

The prior mean corresponds to the center point of the slope of the membership functions. It is the most likely value for the boundary between objects. Lower and upper limits of the membership slope are derived based on variance. Values are chosen such that all slope values that are physically relevant are between the lower and upper limits with a 99.9% confidence. For example, the definition for beach given by the Ministry of Public Works equals an altitude between -1.1 m and 2.0 m (AMSL) (Roelse 2002). Hence, for the beach/sea boundary the prior mean is set equal to -0.55 m and the prior variance to 0.025 m^2. With these values, the lower limit (ll_0) equals -1.1 m and the upper limit (ul_0) 0.0 m, whereby the upper limit corresponds to the maximum high sea level. For the beach/fore dune boundary the prior mean is thus 2.0 m and the prior variance to 0.09 m^2. The lower limit (ll_0) equals 1.0 m and the upper limit (ul_0) 3.0 m. In the same way, the vegetated/non-vegetated boundary has mean 0 and variance 0.0002, resulting in ll_0 of -0.05 and ul_0 of 0.05. The

wet/dry boundary has mean 0 and variance 0.80, resulting in ll_0 of -3.00 and ul_0 of 3.00.

Multi-resolution segmentation is applied to obtain the sampling distribution for the boundaries of the beach object from the high-resolution digital orthogonal photo (figure 5.1). The boundaries are gridded to 30 m resolution and used as an overlay for the DEM and the satellite images. Altitude, NDVI and wetness samples are collected by extraction of these boundaries from the DEM and satellite images, comprising 174 to 324 samples. Subsequently, probability distributions are drawn from these boundary samples.

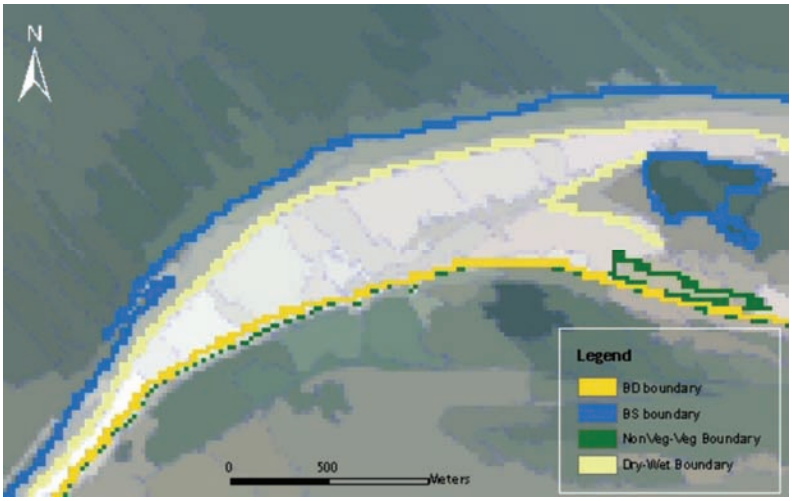


Figure 5.1 *Overlay of boundaries based on multi-resolution segmentation of digital orthogonal photo.*

The posterior distributions are derived from Bayes' Rule (equation 3). Lower (ll_n) and upper limits (ul_n) based on the posterior variance are calculated, again yielding 99.9% posterior confidence interval. Calculations of prior, sampling and posterior distributions were done with the 1stBayes software (O'hagan 1999).

5.3.3 Data processing and validation

Data processing consisted of the following steps. An object-oriented image analysis is performed to obtain segments from the attribute layers. Altitude was derived from the DEM, and wetness and vegetation indices from the Landsat7 ETM image. The segments served as the input into the decision tree. In this study, the hierarchical structure for classification includes 3 levels, i.e. altitude, vegetation cover and wetness. Next, the classification hierarchy is

extended, using posterior probability parameters from the Bayesian hierarchical model, to originate in a fuzzy decision tree. The segments are classified based on this fuzzy decision tree. eCognition™ software environment is used to execute the segmentation and classification processes.

For an estimation of the classification accuracy, a reference data set is collected from field data and a visual classification of the high resolution digital orthogonal image. Next, random points (n = 400) from the reference data in the center of the study area are selected to avoid overrepresentation of sea and foredune areas. Finally, the reference data set is compared with the classification result.

5.4 Results

5.4.1 Bayesian hierarchical modeling

Figure 5.2 shows triplots of the prior distribution, the sampling distribution and the posterior distribution for the three boundary regions. Table 5.3 presents the mean and variance parameters of the prior distributions (μ_0, τ_o^2), the sampling distribution ($\bar{y}, \sigma^2$) and the posterior distribution (μ_n, τ_n^2). Clear differences exist between the parameters of the prior distributions and the sampling distributions. This is particularly the case for the beach/fore dune and the wet/dry boundary. The relative high altitude observations for the beach/fore dune boundary ($\bar{y}$ = 4.24) might be the result of the 2001 beach nourishments. The difference in prior and posterior distribution for wet/dry boundary (μ_0 = 0 against μ_n = -11.76) is due to the inaccurate prior membership function. Clearly, the variance of the prior distributions is much smaller than that of the observations.

Table 5.3 *Input and results of Bayesian hierarchical modeling for the boundary regions of the beach objects.*

boundary	*n*	μ_0	τ_o^2	$\bar{y}$	σ^2	μ_n	τ_n^2
beach/sea	324	-0.55	0.025	-0.95	0.485	-0.90	0.0014
beach/fore dune	174	2.00	0.09	4.24	1.758	4.01	0.0091
veg/nonveg	248	0.00	0.0002	0.038	0.0086	0.032	$3 \cdot 10e^{-5}$
wet/dry	248	0.00	0.80	-23.77	202.60	-11.76	0.40

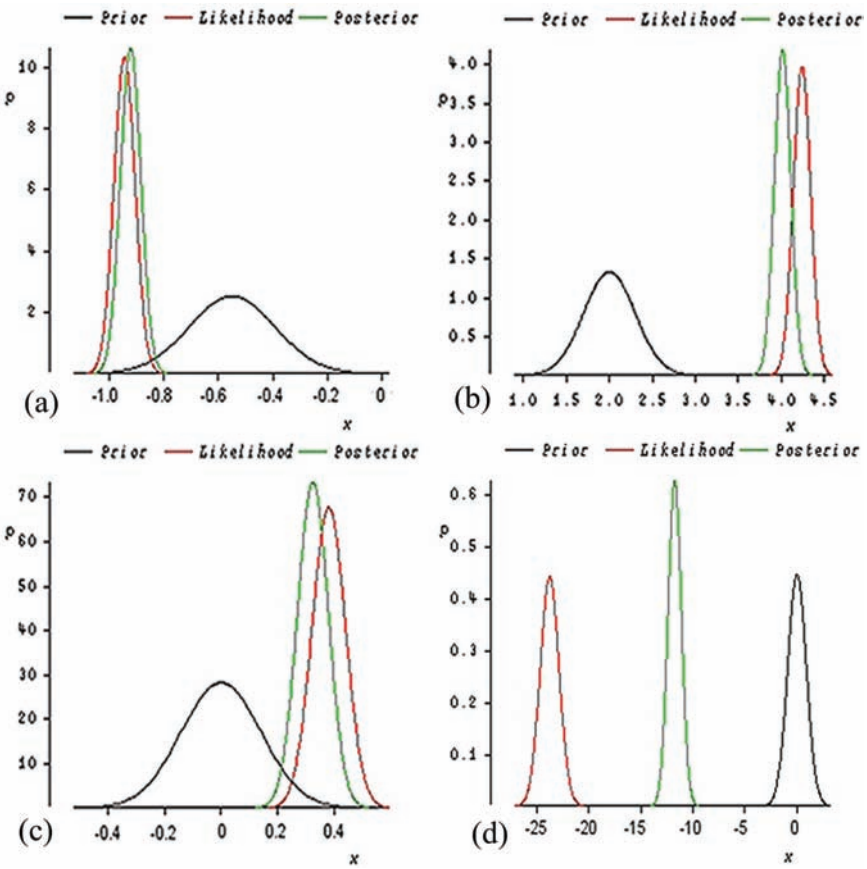


Figure 5.2 *Triplots showing prior, likelihood and posterior distribution for (a) beach/sea boundary, (b) beach/fore dune boundary, (c) vegetated/non-vegetated boundary and (d) wet/dry boundary.*

Table 5.4 shows the lower and upper limits of the prior and posterior distributions. Posterior distributions describe the lower and upper limits more accurately, by means of prior knowledge and data observation. Hence, the domain of natural variation of the boundaries is smaller. Once more, there are apparent differences for the beach/fore dune boundary and the wet/dry boundary.

Table 5.4 *Lower (ll) and upper (ul) limits as set before (prior) and after (posterior) Bayesian modeling.*

boundary	ll_0	ul_0	ll_n	ul_n
beach/sea	-1.10	0	-1.05	-0.80
beach/fore dune	1.00	3.00	3.70	4.33
veg/nonveg	-0.050	0.050	0.015	0.050
wet/dry	-3.00	3.00	-13.86	-9.67

The triplots in figure 5.2 show a difference between the prior distribution and the likelihood of the beach/fore dune boundary samples, as a result of historical beach nourishments. Hence, the beach/fore dune boundary observations in this area are excluded. This resulted in a new triplot (figure 5.3), with a lower limit (ll_n) of 3.13 m and an upper limit (ul_n) of 3.52 m.

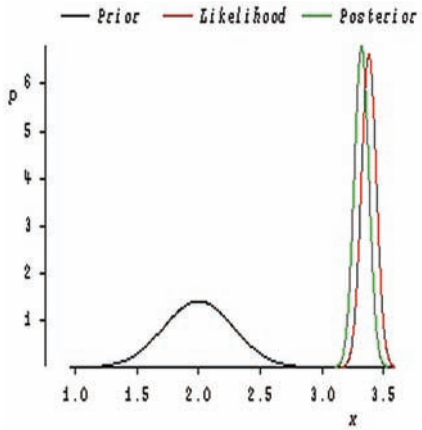


Figure 5.3 *Triplot after beach/fore dune boundary correction.*

5.4.2 Fuzzy decision tree classification

Next, the objects are classified according to a fuzzy decision tree with three decision levels on dry, non-vegetated beach objects (figure 5.4). Final classification is performed on the posterior probability parameters. The posterior lower and upper limits of all boundaries are used to classify the data. For the class “fore dune” a membership value of 1.0 is given to altitudes above 4.33 m, the value 0.0 to altitudes below 3.77 m and a linear membership value between 0.0 and 1.0 for intermediate altitude values.

Figure 5.4 also shows the prior, sampling and posterior mean for all boundaries. Steepness of the membership slopes reflects the variances. Once more it is evident that posterior variances are smaller and therefore have steeper slopes. This means that a higher accuracy for object determination is reached.

Figure 5.5a shows the final classification whereby “beachplains” and “wet beaches” form the major beach objects. There is also “vegetated mudflat” that occurs even after the first dune row. These wet areas are lower areas in a dune region, and are actually no beach objects.

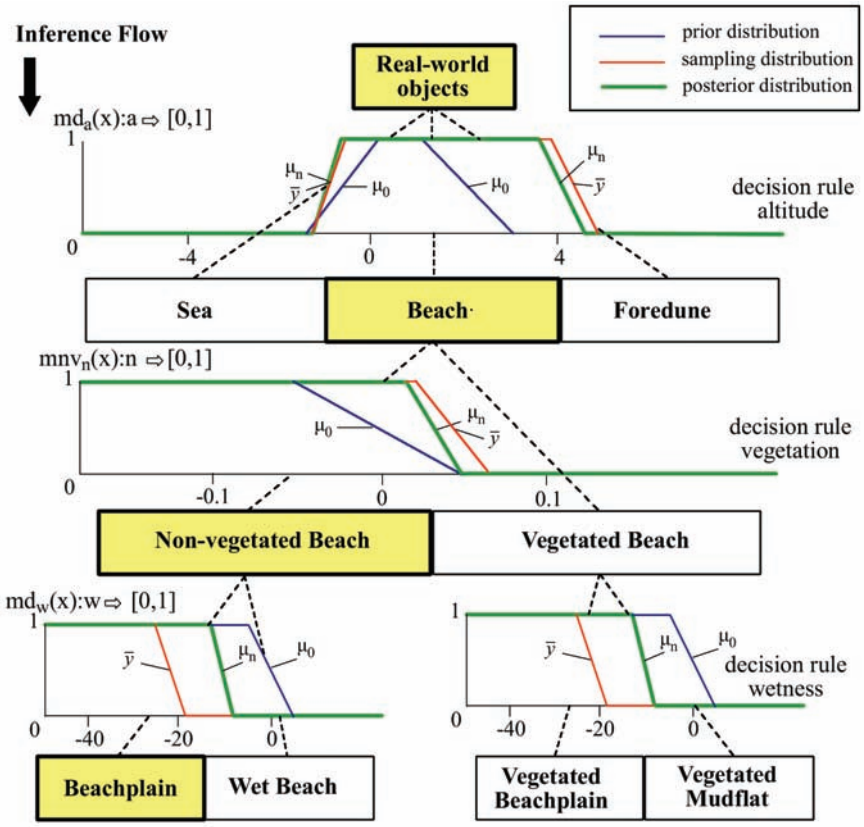


Figure 5.4 *Decision tree for the beach nourishment application. Objects of interest for beach nourishment are the beachplains. Included are the membership functions for each decision rule - based on prior, sampling and posterior distributions - involved for a fuzzy decision tree classification.*

Figure 5.5b shows the final classification after correction for the beach/fore dune boundary. Here, the mutation of the “vegetated mudflat” class into the more likely “fore dune” class is noticed.

Classification results in crisp objects as only the class with the highest membership function is shown.

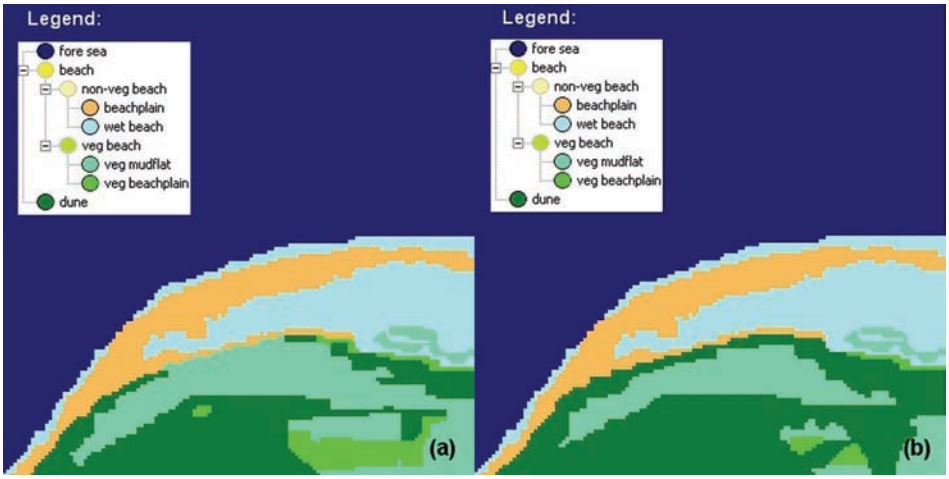


Figure 5.5 *(a) Classification result using eCognition™ (b) Classification result after beach/fore dune boundary correction.*

5.4.3 Classification accuracy

Figure 5.6 shows the outcome for the reference data. Comparison of the reference and the classified data yielded a fuzzy error matrix (table 5.5). Accuracy of the overall classification equals 71.0%, and the κ-statistic equals 0.615. Misclassification between "vegetated mudflats" and "fore dune" as well as confusion between "wet beach" and "beachplain" are accountable for this minimal result. The user's accuracy for "fore dune" is only 40.7%. An improvement is shown in table 5.6. After correction for the beach/fore dune boundary, the user's accuracy for "fore dune" increases to 57.9%. There is a slight increase of the overall classification accuracy to 73.9% and the κ-statistic to 0.651.

Noteworthy is that the sum of all reference samples is less than 400 as membership functions are rounded off during classification and validation.

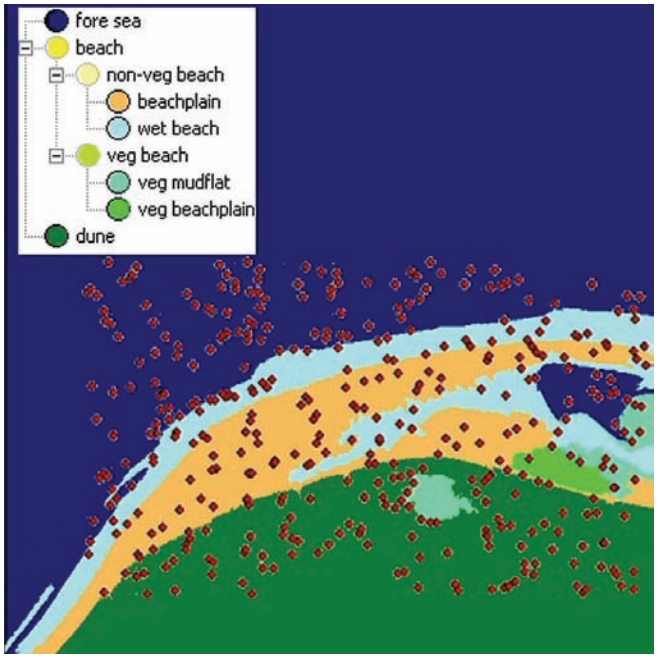


Figure 5.6 *Reference data resulted from visual classification of high resolution digital orthogonal image and field measurements. Superimposed are the random samples used for validation.*

Table 5.5 *Fuzzy error matrix showing classified data (vertical) and reference data (horizontal). Legend: S (sea), B (Beachplain), VB (Vegetated Beachplain), WB (Wet Beach), VM (Vegetated Mudplat), FD (Fore Dune). The overall classification accuracy is 71.0%. The overall κ-statistic equals 0.615*

	S	B	VB	WB	VM	FD	Total
S	153.74	0.00	0.00	9.28	0.00	0.00	163.03
B	0.00	52.82	0.77	23.75	3.64	3.00	83.98
VB	0.00	0.23	0.77	2.86	1.14	0.00	5.00
WB	5.10	13.00	0.00	38.91	0.00	0.00	57.02
VM	0.00	0.00	0.00	7.00	6.00	0.00	13.00
FD	0.00	0.00	1.00	0.00	44.94	31.50	77.44
Total	158.85	66.05	2.55	81.81	55.72	34.50	399.47

Table 5.6 *Fuzzy error matrix for classification after correction for beach/fore dune boundary. Legend: S (sea), B (Beachplain), VB (Vegetated Beachplain), WB (Wet Beach), VM (Vegetated Mudplat), FD (Fore Dune). The overall classification accuracy is 73.9%. The overall κ-statistic equals 0.651*

	S	B	VB	WB	VM	FD	Total
S	153.74	0.00	0.00	9.54	0.00	0.00	163.28
B	0.00	51.59	0.70	23.75	3.64	4.54	84.23
VB	0.00	0.23	0.70	2.86	1.14	0.30	5.23
WB	5.10	13.00	0.00	39.00	0.00	0.00	57.10
VM	0.00	0.00	0.00	7.00	6.00	0.00	13.00
FD	0.00	0.00	0.00	0.00	32.84	45.17	78.00
Total	158.85	64.82	1.41	82.15	43.61	50.01	400.84

5.5 Discussion

This chapter addresses incorporation of uncertainty in beach objects and its exploration using Bayesian hierarchical models and fuzzy decision trees. Incorporation is twofold. First, spatial and thematic uncertainties in attributes arise from the dataset itself. These uncertainties are described by prior expert knowledge and data likelihood. The posterior outcomes result into more accurate uncertainty measures for the beach objects. Second, uncertainties are incorporated during classification and hence during decision making. For beach management, classification of beach objects is decomposed into three decision levels. Each decision level includes posterior uncertainty measures to classify the beach objects.

The Bayesian hierarchical model shows that the prior choice of parameters for the membership functions substantially influence classification. Use of the proposed membership functions as priors is appropriate for the beach/sea and vegetation/non-vegetation boundary. For the beach/fore dune boundary, however, a high misclassification of fore dune objects into vegetated beach objects is noticed. This can be partly corrected by excluding altitude values of those areas where recent beach nourishments have been carried out. Even so, the wet/dry boundary illustrates a mismatch between prior and likelihood. The resulting posterior probability function compromises the two. Apparently the prior knowledge, as derived from earlier studies or experts, seems to be weak and limited regarding wet/dry and beach/fore dune boundary. Information derived from this study can be used to update this knowledge.

The assumption for linear membership functions to describe the boundaries is appropriate, as long as they are gradually changing. The additional benefit of linearity is that one can add the membership values of all data points in the error matrices, yielding the accuracy of the classification. Hence, no uncertainty information is lost.

Fuzzy decision tree classification of beach objects is promising to detect and classify beach areas suitable for nourishments as well as to understand the role of uncertainties in object based classification. Storvik et al. (2005) addresses a Bayesian approach to classify remote sensing data. Their work focuses on the multiresolution aspect of the dataset; as an extension, the current chapter focuses on uncertainty determination and integration.

For beach object classification scale issues are important as small scale relationships between coastal erosion, sediment transportation, beach nourishments and dune forming are opposed to large scale climate change and global sea level rise. Different resolution levels imply different landscape processes and therefore, different classes that can be ordered in a hierarchical way. This approach leads to several classifications at each decision rule on a resolution that is appropriate for the purpose of this chapter. This study might be extended to other applications related to hierarchical classification of fuzzy objects at other scales and need to be further explored.

As an alternative, artificial neural networks (ANN) deal with non-linear dynamic systems, discriminating between actual data and noise and processing previously unencountered patterns (De la Rosa *et al.* 1999). ANNs can be a good approach if there are sufficient data for training. However, it lacks the ability to identify causal relationships and to incorporate expert knowledge and is therefore less suitable for the beach management application.

Dempster-Shafer theory is an alternative for decision-making in environments with limited data and considerable uncertainty, where the goal is to select the best decision from a number of alternatives. It has been applied on coastal zone management studies before (Moore *et al.* 1999, 2001). The strength of the Dempster-Shafer approach is the way that uncertainty is represented (Lein 2003; Shafer 1990; Richards and Jia 1999). The theory relies on the use of belief functions. Belief functions produce an overall expression of certainty. Dempster-Shafer theory introduces the concept of plausibility (Shafer 1990; Richards and Jia 1999). Plausibility defines the degree to which a hypothesis cannot be disbelieved. Belief and

plausibility can be interpreted as lower and upper limits on the possible values of probability in any given situation.

This case study addresses fuzziness in boundaries. Fuzziness is associated with uncertainty born out of a definition of inexactness where an object's membership in a given class is a function of the imprecision surrounding how that class has been defined. Hence, mixed segments may express only partial membership in that class and therefore fuzzy classification methods can be applied to define the membership grade. However, if a given class can be defined by a sufficiently broad range of measures to incorporate an extensive set of possibilities, the evidence theory of Dempster-Shafer may yield good results. In this study, classes are fully described by membership functions derived from a Bayesian approach. Here, a hierarchical knowledge-based classifier incorporating these memberships is more appropriate.

5.6 Conclusions

This study shows the use of fuzzy decision trees for classification of beach objects. A hierarchy in model parameters includes prior probabilities of various attributes. A Bayesian hierarchical model allows one to include the uncertainty into the classification. A combination of prior knowledge with collected data results into posterior distributions and improved classifications. Posterior distributions serve as input for lower and upper limits of membership functions that describe the boundaries between beach objects. A large uncertainty of both the beach/fore dune and the wet/dry boundaries is explained by the weak prior knowledge. Also external factors can play a role, like recent beach nourishments.

Classification results are validated using error matrices that incorporate membership functions. Overall accuracy of the classification equals 71.0%, which increases to 73.9% after when correction for recent beach nourishments.

Chapter 6

Temporal Ordered Space Matrix: Representation of Multivariate Spatio-temporal Data

'To see or not to see, that is the question'

After Hamlet, William Shakespeare (1564-1616)

This chapter is based on the following papers:

Kraak, M.J. and Van de Vlag, D.E. (in prep.) Understanding spatio-temporal patterns: visual ordering of space and time. For Autocarto 2006.

Van de Vlag, D.E., Kraak, M.J., Nieuwenhuis, W. and Schouwenburg, M. (in review). Temporal Ordered Space Matrix: representation of multivariate spatiotemporal data. Submitted to *Computers and Geosciences*.

Van de Vlag, D.E., and Kraak, M.J. (2005). Temporal Ordered Space Matrix: representation of multivariate spatiotemporal data. In GISPLANET 05 Conference Proceedings, S01 Spatial Knowledge. 13p.

Van de Vlag, D.E., and Malunda, A.A. (2005). Interactive Visualization of Geospatial Dynamics. *Geoinformatics* **8(5)**, pp. 28 -31.

Van de Vlag, D.E., and Kraak, M.J. (2004). Multivariate Visualization of Data Quality Elements for Coastal Zone Monitoring, ISPRS Conference Proceedings Commission IV, ISPRS 15-23 July 2004 Istanbul, Turkey, 7p.

6.1 Introduction

Monitoring of dynamic natural phenomena results in large multivariate and multi-temporal data sets. To be able to understand these complex datasets, one needs options to explore these data sets to study trends, correlations and patterns, which should lead to knowledge discovery about the dynamic behaviour of geospatial phenomena. The data can be characterized by three components: space (where), objects (what) and time (when) (Peuquet 1994) that allow changes over time in position or attribute(s) to be registered in both qualitative and quantitative way (Andrienko *et al.* 2003). Next to GIS's analytical functionality, visualization functionality can assist in the exploration of the data sets. Here, the development of the field of geovisualization (MacEachren *et al.* 1999; Kraak 2003; Dykes *et al.* 2005) offers interesting solutions. Geovisualization is a process that involves humans achieving insight by interacting with data through use of manipulable visual displays that provide representations of these data and the operations that can be applied on them. As such it exploits the ability of current computing technology and used methods and techniques from several disciplines like information visualization and geosciences to dynamically analyze and display large amounts of information (Edsall *et al.* 2000).

To visualize the multivariate spatio-temporal data, several techniques exist, however no real integrated solution can be found. From cartography, the approach emphasises the spatio-temporal nature of the data resulting in for instance animated maps. From information visualization the multivariate nature of the data is emphasised and could lead to graphics such as the parallel coordinated plot. In recent years many extensions and new visualization prototypes have been proposed with focus on presentation or exploration of multivariate spatiotemporal data. However, these visualization tools restrict representation of multivariate spatio-temporal data in non-linked windows. Depicting the temporal component of data simultaneously with its spatial information, in an integrated manner, is powerful to support real data analysis tasks.

In this chapter, an integrated method is suggested to visualize multivariate spatio-temporal data sets. The solution proposed here, is a prototype of three dynamically linked views, each representing one of spatial data's components. These are the interactive map view (location), the parallel coordinate plot view (attributes), and the temporal ordered space matrix view (time). Temporal ordered space matrix (TOSM) is a new visualization technique and is a kind of schematized gridded map whereby the rows in the matrix represent time, the columns represent geographic units, and the individual cell

can be color-coded according the value of user defined attributes. A direct link exists between the geographic units in the animated view and the time view. The geographic units are ordered according a linear principle to be explained in section 6.3. The main advantage of the TOSM is that it offers a direct overview of spatial and temporal trends in an ordered space preserving (most) spatial relations.

This chapter concerns the methods and techniques behind visualizing multivariate spatio-temporal data, the environment in which the data is represented (i.e. the interface) and the question "does it work"? (i.e. cognitive aspects). The visualization prototype is applied on four different case studies in The Netherlands. Cognitive aspects of the prototype are described by means of a focus group and usability test.

6.2 Visualization of multivariate spatio-temporal data

In many applications and research areas, more and more multivariate spatio-temporal data become available as a result of powerful data collection and distribution techniques. These multivariate spatio-temporal data contain information on objects located and characterized over time. For analysts, it is a huge challenge to effectively and efficiently detect and understand interesting relationships, patterns, correlations and trends in such data sets (Guo *et al.* 2005). Visualization techniques are powerful tools to detect, analyze and study these changes of characteristics of objects. In particular, the field of geovisualization shows promising results, because geovisualization techniques offer the ability to dynamically analyze and interact with the data.

Representation of objects in space and time has attracted much attention of researchers during the last decade (e.g. Peuquet 1994, 2002; Andrienko and Andrienko 1999; Gahegan 1999; Worboys 2005; MacEachren 1995; MacEachren *et al.* 1999; Edsall and Sidney 2005; Ogao and Kraak 2002; Kraak and MacEachren 1999). Peuquet (1994) described how spatio-temporal data are characterized by three basic components: where (space), what (objects) and when (time). These components are highly interrelated and link to the elementary questions as seen in figure 6.1.

Based on the above basic concepts, Andrienko *et al.* (2003) classify spatio-temporal data according to the kind of changes occurring over time. They recognize existential changes, as appearance and disappearance, which can be related to the "when" component. Besides, they recognize changes in spatial properties, as location,

shape and/or size, orientation, altitude, etc., and this can be related to the "where" component. Last, they identify changes of thematic properties that are expressed through values of the attributes and this can only be partly described by the "what" component. In fact, the "what" component not only indicates the object by itself, but also includes many important characteristics of the object, and these characteristics are described by several attributes. Mennis *et al.* (2000) and Li and Kraak (2005) extended the "where", "what" and "when" perspectives of cognition with a "theme" component to account for multivariate characteristics of objects.

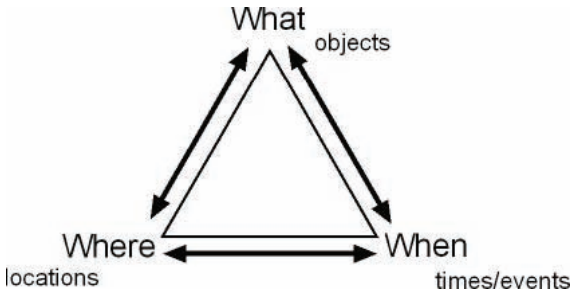


Figure 6.1 *Components of spatio-temporal data after Peuquet (1994).*

The potential for geovisualization and geographic information systems to integrate time into their presentation, both through database and display, is powerful for exploring multivariate spatio-temporal datasets (Edsall and Sidney 2005). According to Peuquet (1994), the overall goal of a temporal GIS is to represent stored spatio-temporal data in a way that conforms to human conceptualizations of the world in space-time. This represents a call for customizable and flexible representation systems to display and analysis of geographic data. In particular, when one deals with multivariate spatio-temporal datasets, several solutions for exploration and visualization are possible, related to the data type and structure. Here, some of these visualization tools and their possibilities and deficiencies are reviewed. An overview of (geo)visualization techniques is given for analyzing and representing multivariate spatio-temporal datasets. First, the focus is on visualization of spatio-temporal data. Second, several tools are portrayed that visualize multivariate datasets. Finally, several existing tools for multivariate spatio-temporal data sets are illustrated.

6.2.1 Spatio-temporal visualization

Spatio-temporal data is often studied to detect and discover changes, regularities or irregularities, forecast trends of development and make decisions. Spatio-temporal data can be visualized using different techniques, such as single map, series of static maps and temporal animation techniques (Kraak 2000; Kraak and Ormeling 2003). The representation of time by a single static map comprises specific graphic variables and symbols that can reveal change, e.g. colour values (tints), arrows, etc. A single map represents a snapshot in time, and together the maps make up an event. Change is perceived by looking at individual maps successively. Also, mapping of spatio-temporal data has been often done using cartographic animations (Acevedo and Masuoka 1997; Kraak and Klomp 1995; Ormeling 1995). Using animation, change is perceived to happen in a single frame by displaying several snapshots after each other. The ability to deal with real world processes as a whole - rather than as instances of time - makes cartographic animations intuitively effective in conveying dynamic phenomena. Cartographic animations reveal interrelations amongst spatio-temporal data components location, attribute and time. Most existing animations are designed as frame-based animations, where there is no link between database and display. For exploration purposes, it is useful if the display is directly linked with the database. This will allow interaction to the user to manipulate represented data as required, using movie playback controls. Kraak *et al.* (1997) define effective and functional operations to both control the graphic representations on display and to assess and interact with the data sets behind them. They distinguish two types of dynamic temporal legends, related to the structure of the data, i.e. linear or cyclic. Typical linear legend types are time sliders and time rulers, while time wheels (or clocks) and free-spinning discs are examples of the cyclic legend type (figure 6.2). Minimum functional operators require options to play with time, e.g. forwards, backwards, slow, fast, pause and stop. Time sliders are often used for linear spatio-temporal data, have their origin in features like media players and shifting time horizontally (figure 6.2a). Time rulers are an alteration on time sliders, whereby time is shifted vertically and therefore has the ability to visualize or illuminate databases (figure 6.2b). Time wheels are suitable for cyclic representations (e.g. monthly fluctuations of vegetation growth over years), but are restricted in their possibility to play backwards (figure 6.2c). Free spinning discs on the other hand have the ability to play forwards and backwards and has been used nowadays in many music media players (figure 6.2d).

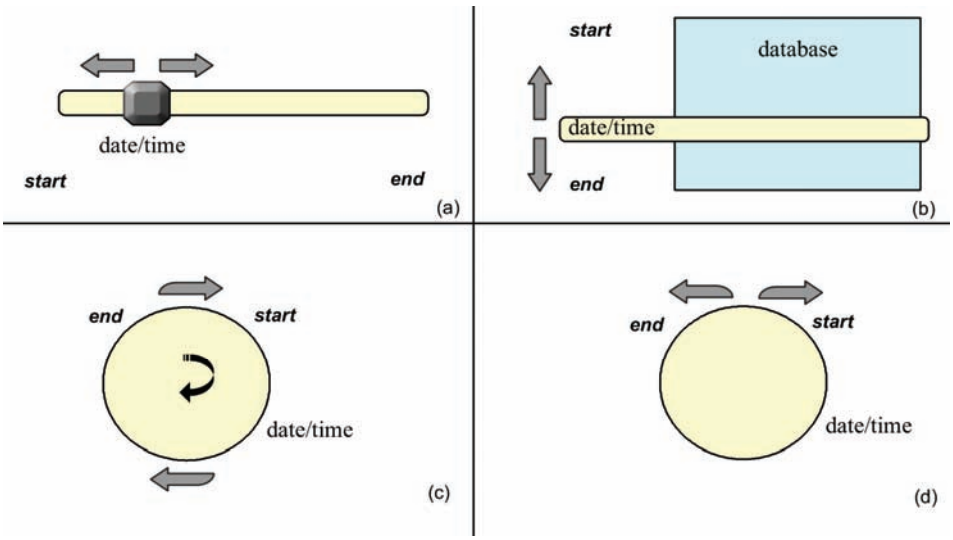


Figure 6.2 *Interactive controlling tools for animation: (a) time slider and (b) time ruler as part of the linear legend type, and (c) time wheel and (d) free-spinning disc as part of the cyclical legend type type (after Kraak et al. 1997).*

Although a number of tools for controlling animation to improve its suitability for analysis are suggested here, cartographic animation lacks the ability to visualize temporal and spatial trends in geospatial dynamics in one glimpse. In particular for identifying and interpretation of features, advanced space-time visualization techniques are or need to be developed to picture trends in geospatial dynamics. Also, visualization of multivariate spatio-temporal dataset using animation is complicated. Blok (1999; 2005) investigated the use of cartographic animation for vegetation dynamics, and identified four categories. These include:

- Changes that can be discovered in spatial domain
- Changes in temporal domain
- Overall spatio-temporal patterns over longer series
- Relative similarities in comparison

In her study, it is evident that monitoring vegetation dynamics requires animation of multivariate spatio-temporal data, especially for the comparison of relative similarities, i.e. investigating patterns of vegetation change in conjunction with others factors like rainfall, etc.

Ogao and Kraak (2002) studied the possibilities for animation of multivariate data by coupling temporal databases and animation functionality, but this is limited to a few variables. Hence, to visualize

spatio-temporal objects a map environment linked to its spatio-temporal database is adopted. The map view is extended with two dynamically linked windows to explore multivariate characteristics of the object (section 6.2.2) and to visualize spatio-temporal trends in one glance (section 6.2.3).

6.2.2 Multivariate visualization

As many spatio-temporal data consist of more than three attributes and therefore do not allow a simple visualization as 2D or 3D displays, more sophisticated visualization tools are needed. Over the past decade, many different multivariate visualization tools have been developed. These visualization techniques empower users to perceive important patterns in a large amount of data, identify areas that need further scrutiny and make sophisticated decisions. But looking at information is only a start. Users also need to manipulate and explore the data, using real-time tools to zoom, filter and relate the information, and undo if they make a mistake. This corresponds to Shneiderman's visual-information-seeking mantra of 'overview first, zoom and filter, then give details on demand' (Shneiderman 1996).

Assessing the benefits of new visualization tools is an important step into geographical analysis. Several authors (Shneiderman 1998; Spence 2001; Keim *et al.* 2005) have tried to classify these techniques. Shneiderman (1998) gave a categorization of visualization tools based on data types and the tasks users wish to perform. He recognizes seven data types, i.e. one-, two-, three-dimensional data; temporal and multidimensional data; and tree and network data. Also, he categorized seven tasks, i.e. overview, zoom, filer, details-on-demand, relate, history, extract. An extension to this categorization is given by Keim *et al.* (2005), who recognizes the following visualization techniques based on data types and interaction techniques: 2D or 3D displays, geometrically-transformed displays, iconic displays, dense pixel displays and stacked displays (figure 6.3).

Geometrically transformed display techniques aim at finding interesting transformations of multi-dimensional data sets. In the field of information visualization, geometrical transformed display techniques are often used and include the scatterplot matrix, a commonly used method in statistics (Andrews, 1972; Cleveland 1993), and the parallel coordinate plot (Inselberg 1985). Star plots (Chambers *et al.* 1983) and Chernoff faces (Chernoff, 1973) are techniques of iconic displays (figure 6.3) that visualize each data item as an icon and the multiple variables as features of the icons. Dense pixel displays visualize each dimension value to a colored pixel and group the pixels belonging to each dimension into adjacent areas

(Keim 2000). By arranging and coloring pixels in an appropriate way, the resulting visualization provides detailed information on correlations, dependencies and hot spots. Stacked displays present data partitioned in a hierarchical fashion. In multidimensional data, the data dimensions to be used for building the hierarchy have to be selected carefully. To obtain a useful visualization, the most important dimensions have to correspond to the first levels of the hierarchy. Examples of stacked displays techniques are dimensional stacking (LeBlanc *et al.* 1990) and Treemaps (Johnson and Shneiderman 1991).

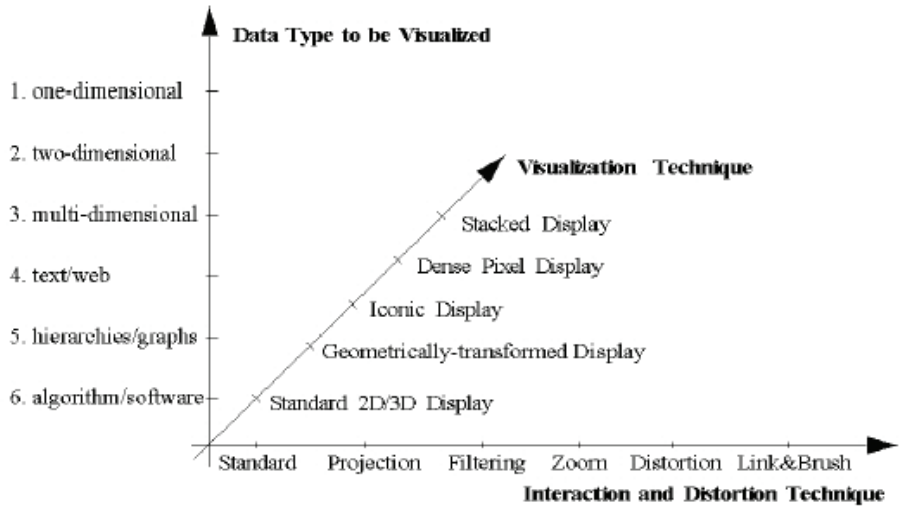


Figure 6.3 *Classification of information visualization techniques (by Keim et al. 2005).*

From all these techniques, the dynamic parallel coordinate plot has been demonstrated as a powerful multivariate visualization technique and has been used in many applications (McGranaghan 1993; Andrienko and Andrienko 2001; Edsall 1999; Spence 2001; Lucieer and Kraak 2004). The dynamic parallel coordinate plot was introduced in 1985 by Inselberg. Parallel coordinates plots map *k*-dimensional space onto the two display dimensions using *k* axes that are parallel to each other, evenly spaced across the display and can be either horizontally or vertically oriented. The individual data values are then marked off for each dimension onto corresponding coordinate with the highest data value as maximum value and the lowest as minimum. The geographic objects are represented as lines connecting the attribute values. From the structure of the resulting display one can draw conclusions for the relationship of the corresponding data

values. A group of lines with a similar gradient can, for example, indicate that their data records correlate positively. Furthermore, outliers in values are easy to detect, and parallel coordinate plots can handle ratio as well as ordinal data (Spence 2001), however they lack the ability for proper representation of nominal data. A possible solution is the mapping of nominal data in an arbitrary order with a user-friendly interface element that allows the rearrangement of the value.

Representation of temporal data using parallel coordinate plots was attempted by Edsall (2003). The easiest way is to represent successive data on each axis, like a time series, however multivariate information is then lost. In his approach, time was mapped as a time stamp to the axis representing data attributes. Another attempt is to map the time dimension as one of the axis in the PCP. However, the dynamic behavior of objects cannot easily been seen by the observer. Hence, the PCP is applied in the prototype, using a separate window for visualization of multivariate characteristics of the objects, but linked with the time stamp of the map view. This facilitates understanding of casual relationships in data- and time dimensions.

6.2.3 Multivariate spatio-temporal visualization

Multivariate spatio-temporal dataset in a geographic context occurs when geoscientists deal with set geospatial data where each geographic space contains its own data array representing location, attribute and temporal dimensions (Monmonier 1990). Hence, there is a need for visualization technique(s) that will provide greater insight towards the multivariate and spatio-temporal characteristics of the dataset. This would involve integration of spatio-temporal and multivariate visualization tools in one visualization environment, taking advantage of both their inherent capabilities. Linking allows interaction between the views and can facilitate exploratory tasks. This was introduced by Monmonier (1989), linking statistical analysis tools such as the scatter plot with maps, adding geographic meaning to the analysis. Recently, many other working examples of linking maps with graphic tools, such as dot plots, parallel coordinate plots, and scatter plots, can be found (Buja *et al.* 1996; Dykes 1997; Adrienko and Adrienko 1999; MacEachren *et al.* 1999; Edsall *et al.* 2001; Anselin *et al.* 2002). For linking of temporal data, a link can be added between an animated map and a graphic tool, using interactive time control mechanisms (see also figure 6.2). Selecting time on these tools would automatically select corresponding elements on map and graph.

There are several multivariate spatio-temporal visualization concepts and products. Concepts specify a single-view representation, while products consist of a visual environment. Often, these visualization concepts and products are appropriate for particular types of datasets. Below, two visualization concepts followed by two products are discussed here.

Visualization of trajectory datasets can be done using the Hägerstrands space-time-cube (Hägerstrand 1970). This concept – originally developed in the sixties of last century – represents spatial data on the x- and y-axis, while time is represented on the z-axis. Nowadays, the appearance of GPS-data resulted in a revival of this concept. Kraak (2003) extended Hägerstrand's space-time-cube to visually deal and explore with spatio-temporal datasets in geovisualization environment by means of interactivity, dynamic visualization and alternative views. These characteristics offer the user full flexibility to view, manipulate and query data. Using the space-time-cube for visualization of multivariate spatio-temporal data, it is clear that it is difficult to represent multivariate data in time, as it is restricted to nominal datasets only. A further extension of the space-time-cube using dynamic parallel coordinate plots by Li and Kraak (2005) improves the representation of multivariate spatio-temporal datasets.

Instead of x- and y-coordinates, the "what" component can be conceptualized as a "theme space" in the same way as the space-time cube (Mennis *et al.* 2000). The integration of the theme space with the space-time cube creates a multidimensional hypercube. Instead of representing spatial coordinates, location and multiple theme values are depicted on the x- and y-axis, while the z-axis represents time. Hypercube based methods are efficient for exploration and analysis of data warehouses and large multivariate spatiotemporal datasets (Marchand *et al.* 2004), but are less useful for visualization purposes of spatial properties.

CommonGIS is an interactive visualization tool to explore spatio-temporal data and it includes a multitude of tools applicable to different data types (Andrienko and Andrienko 2004). The set of tools include animated thematic maps, map series, value flow maps, time graphs and dynamic transformations of the data. Its strengths is in exploration and visualization of dynamic datasets, however it lacks support for multidimensional datasets.

A tool for visualization of multivariate spatio-temporal data is TerraSeer Space-Time Intelligent System (STIS) (TerraSeer 2004),

which is a software package for statistical analysis of census data, and it provides visualizations of data patterns in geographic space and time. TerraSeer (STIS) has functions that allow space-time datasets to be viewed and statistically analyzed by means of animation techniques and several multivariate visualization tools. Also, STIS supports interactive linking and brushing along with time-synchronization of multiple views, which provide simultaneous updates to queries and data selections. The user can visualize attribute data on the PCP while on the map geographic distribution overview is presented. On the temporal slider bar the user can select a data set at a particular time stamp to be visualized on the map and PCP. The software has visualization techniques representing 1-dimension data to n-dimension data, including maps, time plot, table view, parallel coordinate plots, scatter plot, box plot and histogram. Although STIS is effective in visualizing trends and patterns in multivariate spatio-temporal data, it does not provide a visualization tool to explore these data in one glimpse.

Visualization and exploration of multivariate spatio-temporal datasets created its own dilemma, as most existing methods represent multivariate data and spatio-temporal data separately. As a consequence, a novel visualization technique is proposed, whereby spatial and temporal properties of spatio-temporal data sets stay preserved (Van de Vlag and Kraak 2004; Van de Vlag and Kraak 2005). Here, geospatial locations of objects are ordered as linear elements in space and are portrayed against the temporal variations of an attribute. We construct a matrix of squares, named temporal ordered space matrix (TOSM), where each column depicts a geographical unit, each row represents the time and each cell the value attribute. The TOSM is in particular effective for naturally ordered spatial datasets, such as coast and river compartments, traffic data, water level data, etc. For such datasets, a clear directional spatial ordering between the geographical units can be perceived. For instance for coast compartments, a compartment X that is to the west of compartment Y, and a compartment Z that is to the east of compartment Y, the directional spatial ordering for coast compartments from West to East will be X, Y, Z. Also for flooding data, measuring point A is upstream of point B, and point C is downstream of point B, the directional spatial ordering from upstream to downstream will be A, B, C. This is also shown in figure 6.4, where the spatial ordering can be inherited from the location of the measuring points, i.e. upstream and downstream. In the TOSM environment, the position of the geographical unit (i.e. the measuring points) can also be drawn using a linear line. Not in all cases however, there is clear (linear) spatial ordering. The TOSM can be

used for census data with no spatial ordering. In any case, to represent the geographical units in the TOSM, spatial reordering needs to be executed.

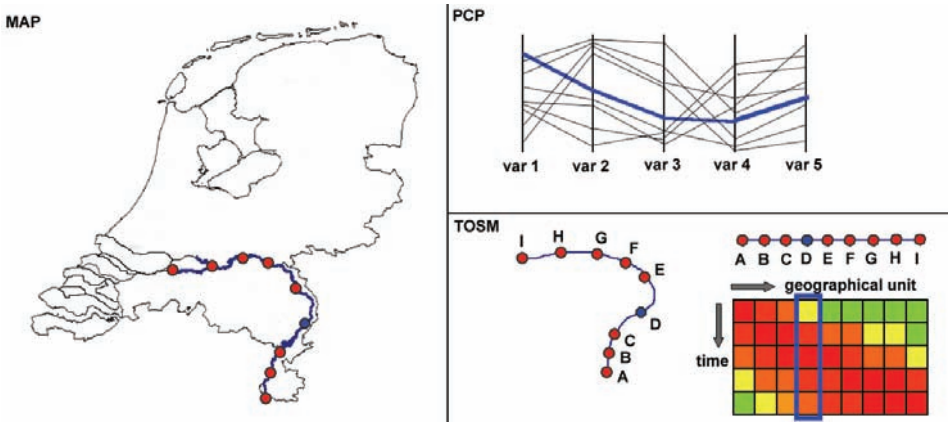


Figure 6.4 *Concept of the TOSM environment for depicting multivariate spatio-temporal data, using a flooding application as illustration.*

6.3 Reordering of spatial data

To visualize location and time into a single two-dimensional graphic, a reduction of location of geographical units into one-dimension is needed. Hence, if geographical units are depicted in a linear direction along the horizontal axis the instant overview from location versus time keeps preserved. This chapter therefore describes the reordering of spatial data needed to obtain this reduction in dimensions. Methods to extract dimensions are strictly related to the spatial neighbourhood of the geographical units. After all, the explicit location and extension of geographical units define implicit relations of its spatial neighbourhood. These spatial relations have been classified in several types (Frank 1996; Papadias and Sellis 1994), including directional relations that describe order in space (Frank 1996; Freska 1992), distance relations that describe proximity in space (Hernández *et al.* 1995) and topological relations that describe neighbourhood and incidence (Egenhofer 1994).

The reordering of spatial data can be easily achieved when datasets are naturally ordered. A directional linearization approach can be applied, whereby the orientation can be inherited from directional relations of the geographical units. When datasets are randomly ordered other (non-linear) methods may be more useful. Here, the

directional linearization approach is discussed first, followed by other linearization approaches.

6.3.1 Spatial ordering by directional linearization

For the TOSM, reordering of objects – while preserving most of its spatial relations – is the strength of this visualization technique. Therefore, the geographical units have to be ordered into linear elements. Usually, to refer to geospatial locations, people use qualitative describers based on their spatial perception and to denote relative directions among geographical units, such as East, West, North East and South West. But this kind of directional reference is imprecise and makes it difficult to identify these objects. There are different spatial models for handling directional terms. A simple one is to divide an area into eight directions based on an angular division of an area into eight equal sectors (N, NE, E, SE, S, SW, W, NW). More precise directions are derived using compass directions in degrees or sketching the initial orientation of the objects.

In case of the TOSM, the centroid of a geographical unit is chosen to locate the object, as shown by an example of reordering 46 municipalities in Overijssel, The Netherlands (see figure 6.5). For line features, the center of line segments is chosen. For point data, the coordinates of the point objects themselves locates the object. For applications with clear linear phenomena - e.g. coastlines, riverbanks, networks, etc. - this selection is straightforward. However, for complex (irregular) thematic data like most census data, spatial reordering of the objects may not perform effectively. The linearization is executed by sketching a directional line with the initial orientation. The initial orientation can be deduced from the structure of the data set, e.g. the SE-NW orientation in figure 6.4, or it can be derived from the initial research objective, e.g. if you want to explore the trend from feature X from North to South. Also, the user can freely rotate the line in each desirable direction, which may lead to new discoveries in pattern and trend recognition. In the final stage, the centroids of the objects are perpendicular projected along the directional line and depicts the order of the objects in the horizontal matrix cells (see figure 6.5).

Obviously, the selection of initial direction is essential on the spatial behaviour of the TOSM. After all, when one draws the directional line not from West to East (see figure 6.5a), but from North to South (see figure 6.5c), geographical units are reordered differently and various other spatial relations and patterns are or may be illuminated. This however can be related to the research question, e.g. if one is

interested in finding spatial or spatio-temporal patterns from West to East, North to South, etc.

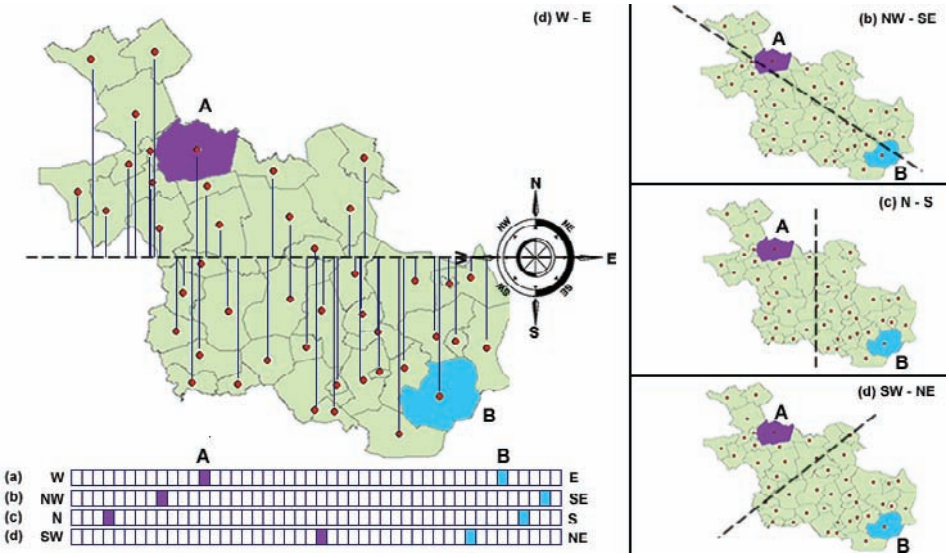


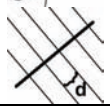
Figure 6.5 *Implementation of initial orientation using a directional line for reordering municipalities in Overijssel, The Netherlands. The red circles are the centroids of the individual objects. Four different directional lines are illustrated: (a) West-East, (b) Northwest – Southeast, (c) North – South, (d) Southwest – Northeast.*

6.3.2 Other linearization methods

Directional linearization of the dataset is uncomplicated when datasets are naturally ordered, like landscape modeling of linear physical features. With most census data, however, this is not the case, as linearity between objects or units are complicated to find or even non-existent. Accordingly, alternative ordering for linearization of the dataset are given to reduce a 2D location on the map into a 1D location in the TOSM. Here, some alternatives are discussed and illustrated in table 6.1.

With some applications there is a clear linear spatial relationship, however the orientation of the objects or units are not straight. Measurement points along a river or a curved road are good examples. Here, freeform drawing of the directional line, thus along the river and the road, may be more suitable.

Table 6.1 *Illustration and applications of spatial ordering methods*

Linearization method	Illustration	Applications
Linear directional line		Linear data, such as coast and river compartments, road paths (straight roads), networks. (applied in TOSM environment for all datasets)
Freehand directional line		Non-linear data, such as rivers, stream channels, road paths (curved roads)
Space filling curves		Raster data
Traveling Salesman Problem		Transportation, logistics data
Plane sweeping algorithms		Linear and non-linear data

Other alternative techniques related to ordering of spatial data are space filling curves, traveling salesman problem and plane sweeping algorithms (Shekhar *et al.* 2004). Space filling curves are optimization techniques, passing through all centroids in particular order, where each centroid lies at an unique distance between each other. Hence, they are effective for spatial ordering of raster data sets. The traveling salesman problem is also an optimization technique that relates the question to find a roundtrip of minimal total length between the centroids, visiting each centroid exactly once. The traveling salesman problem is efficient for transportation and logistics applications; however for detecting spatial patterns in geospatial dynamics they are too rigid, as ordering of attributes is done only by distance. Plane sweeping algorithms reorder data according to the closest distance d between centroids, using an initial direction with a perpendicular plane of distance d. Plane sweeping algorithms, space filling curves and traveling salesman problem, are useful for reordering spatial data sets and can greatly improve the general applicability of the TOSM for other spatio-temporal applications.

6.4 TOSM applications

The prototype is assessed on four different applications to illuminate the strengths and possibilities of the TOSM. Each of these applications are described below, where two applications are naturally ordered (Case study I and II), and two applications have no clear initial orientation (Case study III and IV).

6.4.1 Case study I: River flooding application

Late January 1995, extreme rainfall in Germany, Belgium and the Netherlands caused flooding of the Maas and Rhine River, resulting in evacuation of 250.000 people. The case study involves 16 daily water height measurement points from 25 January to 11 February 1995 along the Maas River in The Netherlands. Water heights of the flooding are standardized to mean water levels. These values are normalized from 0 to 1 for optimal visualization. The Maas River flooding is a single-variate naturally ordered spatio-temporal application. Hence, there is a need for a temporal visualization tool to observe the evolution and movement of the peak flow in time. The TOSM can help for understanding these temporal trends. Visualizing this application in the TOSM, a clear initial orientation can be noticed based on the flow direction of the water.

6.4.2 Case study II: Beach management application

The second case study considers a beach management application in the northern part of the Netherlands (see chapter 2). Here, the beach management application is used as a multivariate naturally ordered spatio-temporal application. It is easy to depict a map with beach compartments suitable for nourishment. However, trends and associations between compartments, changes in time and quality elements involved in the decision making, are more complicated to visualize. Hence, visualizing this application in the TOSM can help to understand and explore these relationships and trends.

6.4.3 Case study III: Housing economics application

The third case study concerns the visualization of yearly housing economics of 45 municipalities in the province Overijssel (East Netherlands) from 1991 – 1997. Five attributes for each municipality are considered, i.e. population, population density, total number of houses, housing density and total number of newly build houses. The data set include a province map with boundaries of all municipalities and temporal databases of the attributes. The objective is to see if there is any spatio-temporal pattern among the variables for the Overijssel province. This study is a multivariate randomly ordered spatio-temporal application. After all, municipalities are not naturally ordered. The TOSM is used to understand temporal trends; however a clear initial orientation can not be noticed from the dataset.

6.4.4 Case study IV: Environmental pollution application

The environmental pollution application includes a multivariate spatio-temporal dataset derived from StatLine (StatLine 2005). The database consists of three pollution sources potassium, nitrogen and

phosphate as total amount of livestock per ha for twelve provinces in the Netherlands from 1995 to 2004. These pollution sources are direct related to farming activities. Mapping of the dataset is done by means of a base map of the Netherlands with boundaries of all provinces. The objective is to see if there is any spatio-temporal pattern among the pollution sources in the Netherlands. As provinces, like municipalities, are randomly ordered, there is no clear initial orientation that can be derived from the dataset.

6.5 The prototype

Integration of different visual methods takes advantage of their inherent capabilities for visualizing spatial distribution of changing phenomena, as well as interrelationships with its change contributing factors, gaining greater insight to information from multivariate spatio-temporal data. Visual methods can perform a range of functions in the visualization process, as described by DiBiase (1990), including exploration to reveal pertinent questions, confirmation of apparent relationships in the data, synthesizing or generalizing findings and presentation of the findings. The visualization process usually begins with exploration and confirmation of relationships in data from the private realm where specialists are engaged in research to public, where the research findings are released to the public for communicating the results. Hence, for implementation of the visualization prototype and its interface, it is essential that besides visualization of obvious relationships in multivariate spatio-temporal data, the prototype should also be able to explore non-obvious relationships.

6.5.1 Functionality of the prototype

The basic aim of an interactive exploratory visualization is to provide a visual environment, which present data to the user in such a way that it promotes the discovery of (inherent) patterns and relationships. Extending this functionality for multivariate spatio-temporal applications will lead to insight concerning the dynamic interactive behaviour of features (i.e. natural objects or geographical units). Insight will be obtained by identification of changes in features using temporal visualization techniques and by comparison to other variables explaining the cause of these changes by means of multivariate visualization techniques. Hence, the visualization prototype should provide multivariate representations, temporal visualization techniques and dynamic interactive functions to allow the user to interact with the data. Three dynamically linked views are selected, i.e. the MAP view to represent the location of the features, a PCP view to visualize multivariate characteristics and a TOSM view to

illustrate spatio-temporal trends. The prototype helps the user to find patterns and relations in multivariate spatiotemporal tasks, such as the relation between vegetation growth and precipitation in time or the temporal changes in population and housing densities in cities.

6.5.2 Implementation of the prototype

Techniques and tools to visualize and explore multivariate spatio-temporal datasets, as described in section 6.5.1, are implemented in a prototype. Implementation was an iterative process that started with the definitions of requirements of the functionality and a priority list for realization within the available time frame. These minimal requirements were the realization of each environment (MAP, PCP and TOSM), dynamically linking of these three environments (see also section 6.5.4), the possibility for drawing a linear directional line in the TOSM view, and interchangeable and scalable axes of the PCP view.

The prototype was developed in JAVATM and implemented in UDIG (User-friendly Desktop Internet GIS), version 1.0.5, developed by Refraction Research Inc (UDIG 2004). UDIG is an open-source platform for GIS applications, mainly to edit and visualize geographic datasets from internet and to develop new tools. It can read files most images files and files from feature and map web servers. With respect to the case studies, shape files of the data are imported. The strength of UDIG is that new tools can easily be developed and added to the user-interface using the databases from each shape file.

After some basic functions had been implemented, a regular test version could be released. Plug-ins with new functions were imported into the plug-in folder of UDIG. Detailed walkthroughs of each version revealed a number of problems, which were all reported to the developers. The main issues were also discussed in regular meetings. Before the prototype was exposed to domain experts (see section 6.6.1), most of these problems of the first evaluation were solved.

6.5.3 User Interface

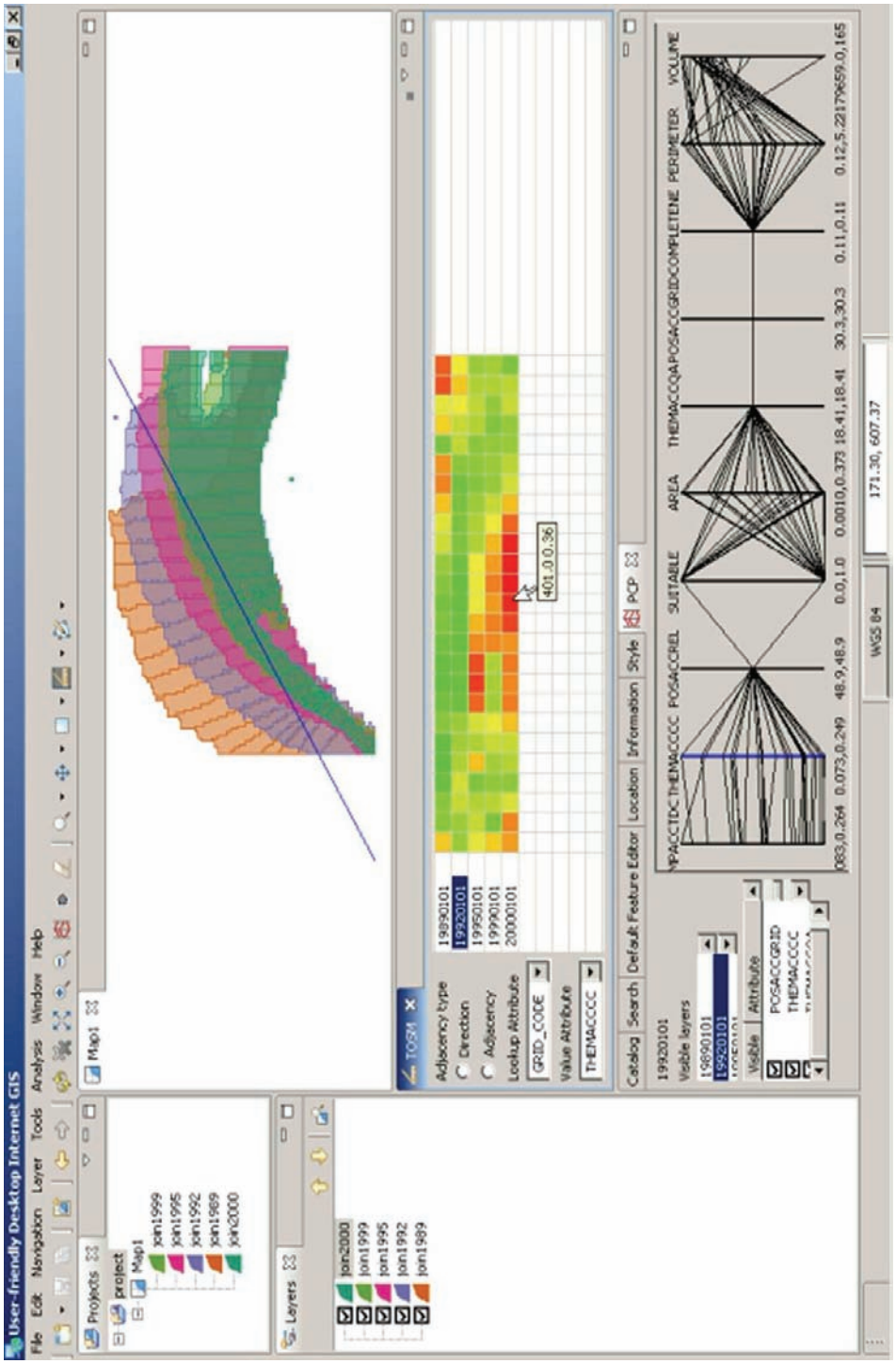
The UDIG user interface has the so-called WIMP style, which stands for Windows, Icons, Menus and Pointers. It contains a series of pull-down menu's, toolbars and buttons to access the main editing and visualization functions, similar to most commercial GIS packages. The general user interface consists of a project menu, where new projects can be opened and stored, and layer menu, that contains all map layers. By default, UDIG renders the layers and open them in the map view.

The prototype for visualization and exploration of multivariate spatio-temporal datasets consists of three visualization environments: a map environment (MAP) - already available in UDIG – and two new developed tools, a parallel coordinate plot (PCP) and the temporal ordered space matrix (TOSM) (see figure 6.6).

The MAP environment displays the area map giving answer to the "where" component. By ticking boxes in the layer menu, the user can decide which map to draw in the MAP environment. Tools for pan, zoom, info, edit and selection are all available in the UDIG tools menu. Value attributes in the MAP environment can only be shown using the info button and selecting a particular object.

The PCP is a geometrically transformed visualization technique, suitable for representing multivariate data (see section 6.2.2). Here, the "what" component is illustrated by the object and its attributes. The prototype is able to generate a PCP using the attribute table of the dataset, showing the attributes as vertical axes of the PCP and each object as a connected line between attribute values. Using a pull-down menu, a user can select the data layer to be seen in the PCP. Besides, a user can select the attribute values to be seen in the PCP by ticking the accompanying boxes. On the bottom of the attribute axis the minimum and maximum values are represented. On mouse-overs near the PCP axes, actual attribute values of objects are shown.

The TOSM is a representation of a matrix whereby ordered space is portrayed against time, illustrating the "when" component. It illustrates schematically location, time and theme of spatio-temporal objects. Here, the user has first to define the initial orientation of the dataset, by drawing a directional line. In the prototype, only the linear directional line has been implemented. Next, the user has to select the lookup attribute (i.e. the object ID) used to link the object to the spatial reordered location, and the attribute to display the theme of interest using pull-down menus. Subsequently, the TOSM is automatically generated. A colour table is added, calculated from the minimum and maximum values of the attribute of interest ranging from green (minimum) to yellow (intermediate) to red (maximum). On mouse-overs on each cell, the object ID number as well as its attribute value is shown.


User-friendly Desktop Internet GIS
File Edit Navigation Layer Tools Analysis Window Help
Projects
project
Map1
jan1999
jan1995
jan1992
jan1989
jan2000
Layers
Map1
TOSM
Adjacency type
Direction
Adjacency
Lookup Attribute
GRID_CODE
Value Attribute
THEMACCCC
19990101
19920101
19950101
19990101
20000101
401.0,0.36
Catalog
Search
Default Feature Editor
Location
Information
Style
PCP
19920101
Visible layers
19990101
19920101
Visible
Attribute
POSACCGRID
THEMACCCC
NPACCTDC THEMACCCC POSACCREL SUITABLE AREA THEMACCQAPOSACCGRIDCOMPLETENE PERIMETER VOLUME
083,0.264 0.073,0.249 48.9,48.9 0.0,1.0 0.0010,0.373 18.41,18.41 30.3,30.3 0.11,0.11 0.12,5.2217965 9.0,165
WGS 84
171.30, 607.37

Previous page

Figure 6.6 *Screenshot of the prototype with beach management application data. Top view depicts the animated map environment; the middle view shows the TOSM; the bottom view illustrates the PCP.*

6.5.4 Interactivity

The user-interface provides a link between the user and the data behind the display. By providing the user with high interactive functions, the represented data can be manipulated depending on user's needs, and is a basic requirement when designing an exploratory visualization prototype. This interactivity can be provided with linking the different views, such that it shows a one-to-one correspondence between objects displayed in one view and those displayed in another view (Cook *et al.* 1997).

In the prototype, the interaction between the three different views (MAP-PCP-TOSM) is based on this dynamic linking of the views. This is achieved by linking the labels of objects as identifiers in each view. A user can select an attribute in the PCP (i.e. a PCP-line) or TOSM (i.e. a TOSM-cell) that corresponds to an object label. Subsequently, the same object will be highlighted in the MAP by blinking. To allow spatio-temporal correlation between the three linked views, linking is also realized by time synchronization of the three views. Attributes selected in the PCP or TOSM for a particular time stamp will be highlighted in the MAP that corresponds with this time stamp. All dynamic linking can be inverted.

6.6 Evaluation test

Evaluation is an important issue in considering the effectiveness of the prototype in improving understanding and exploration of trends, relationships and patterns of multivariate spatio-temporal data sets. Two user tests are applied: a qualitative evaluation using a focus group session and a quantitative evaluation by means of a usability test.

6.6.1 Focus group session

Focus groups are an excellent way of qualitative user testing (Morgan 1998, Kessler 2000). In a focus group session, domain experts evaluate the effectiveness of a prototype in the context of their domain. Harrower *et al* (2000) describe focus groups as a 'cost-effective way to generate qualitative evidence concerning the pros and cons of a geovisualization system'. The main advantage of this method is that it reveals broad patterns of use, misconceptions and

errors; it is also relatively easy, affordable and can be quickly assessed (Faulkner 2000; Robinson *et al* 2005). A possible disadvantage is that the qualitative results may be biased, for example if participants are not freely expressing their opinions and feelings, or if the experimenter lacks control over the session. The main goals of the focus group session were to obtain opinions and reactions on the prototype leading to some improvements and to minimize the potential usability problems in the later usability test.

Six researchers, all with academic background and GIS experience ranging from 4 to 15 years, participated for a 1.5-h focus group session. The session started with a 15-min presentation about the background of the prototype. Then, the participants worked with the prototype with clear directions to focus on the TOSM tool. In this session, only the datasets of the river flooding (case I) and beach management application (case II) were available in the prototype. They filled in a questionnaire for recommendations on the prototype. Finally, their experiences are discussed and recorded in a 40-min group discussion.

They concluded that the prototype seems useful for detecting trends and relationships in multivariate spatio-temporal data. Hence, they emphasize that the prototype is applicable for naturally ordered spatial datasets like case I and case II.

The prototype was not considered very intuitive. Herein, a solution can be found by presetting a default orientation of the directional line. Furthermore, they recommend to add some basic map features, e.g. north arrow, interactive legend and base map, and to improve interactivity by highlighting objects in the map view. For the PCP view, better relationships between variables can be found when the vertical axes of the PCP are interchangeable and scaled.

All these recommendations lead to a list of improvements. Likewise the implementation of the prototype, an iterative process between the list of improvements and a priority list of realization within the available time frame leaded to several alterations of the prototype.

6.6.2 Usability test procedure

Usability evaluation addresses the relationship between tools and their users. In order to be effective, a tool must allow users to accomplish their tasks in the best way possible. The same principle holds for the prototype, where the overall goal is to assist users to explore (hidden) patterns and relationships in multivariate spatiotemporal data.

There are various ways to measure the usability of a tool. International Organization for Standardization (ISO9241-11) identified three usability measures, which include effectives, efficiency and satisfaction (ISO 1993). Nielsen (1993) identified five usability measures and addressed them as 'usability attribute'. These include learnability, efficiency, memorability, errors and satisfaction. The description of each measure is shown in table 6.2.

Table 6.2 *Usability measures from literature (ISO 1993; Nielsen 1993).*

Effectiveness	Measures the ability of the system in meeting the intended goal by looking at how the system assists the user to accurate and correctly complete the tasks.
Efficiency	Observes time and effort required to accomplish particular task by the user.
Satisfaction	Relates directly to users attitude when using the system. Looks on how much the system is acceptable by the user with regard to comfortability felt in using it.
Learnability	How fast can a user who has never seen the user interface before learn it sufficiently well to accomplish basic tasks?
Memorability	If a user has used the system before, can he or she remember enough to use it effectively the next time or does the user have to start over again learning everything?
Errors	This look at how often do users make errors while using the system, how serious are these errors, and how do users recover from these errors?

Usability testing methods involve assessing the tool's ability to meet user's performance (effectiveness and efficiency) and satisfaction objectives. The assessment is usually conducted based on the number of representative user tasks for which a certain number of usability measures are measured.

There are a number of methods for conducting usability testing, including thinking aloud, co-discovery, question-asking and performance measurement methods (Nielsen 1993; Suchan and Brewer 2000). The aim of this usability test is to test the visualization environments (MAP, PCP and TOSM) of the prototype, each separately and combined, for efficiency, effectiveness and user's satisfaction in detecting patterns, trends and relationships in multivariate spatio-temporal datasets. Hence, the question-asking method is adopted whereby the user is asked to write down answers to direct tasks they have to perform.

For the usability test, 17 persons participated, whereby the majority (63 %) has less than 3 years of GIS experience. The usability test group consisted of nine PhD students and nine M.Sc. students in geo-informatics. The test group was divided into three test sessions. Each

session started with a 15-min introduction on the prototype. Next, the participants were divided into three user groups, to promote answering the tasks individually. Each user group started with a different environment (see table 6.3), to counteract learning effects. Each participant received a working sheet with background information on the available environments in UDIG, as well as an explanation of all attributes of the case studies (see appendix A).

Table 6.3 *Framework for usability test.*

User group	MAP	PCP	TOSM	MAP-PCP-TOSM
1	Case I	Case III	Case IV	Case II
2	Case IV	Case I	Case III	Case II
3	Case III	Case IV	Case I	Case II

Five tasks needed to be executed for each working environment (see table 6.4). One task was related to find a multivariate relationship, one task was related to find a temporal trend and the last three tasks were related to find a spatio-temporal pattern in the datasets. The multivariate relationship tasks relate three variables for a particular time period, The multivariate relationship tasks relate three variables for a particular time period, for example in case IV, whereby the relationship is asked between the variables potassium, nitrogen and phosphate for the year 1995. The temporal trend tasks describe temporal changes of one variable, such as the temporal trend for data completeness in case II. The spatio-temporal pattern tasks specify patterns for location of objects and temporal changes in a variable, such as the spatio-temporal pattern of location of measurement points and water height in time for case I. Having four case studies, each participant completed 20 tasks in total, with 4 multivariate tasks, 4 temporal tasks and 12 spatio-temporal tasks (see appendix B).

Table 6.4 *An overview of the five tasks that a participant has to execute during the test for each case study.*

Multivariate relationship	Can you describe and explain the relation between variable *X*, variable *Y* and variable *Z*?
Temporal trend	Can you describe and explain the trend of variable *X* in time?
Spatio-temporal pattern	Can you describe and explain the pattern between location and *X* in time?
Spatio-temporal pattern	Can you describe and explain the pattern between location and *Y* in time?
Spatio-temporal pattern	Can you describe and explain the pattern between location and *Z* in time?

For each task, the participant was asked to record the 'time of thinking'. To test for efficiency, each participant was asked to activate the stopwatch at the beginning of the task. The time was recorded until the moment the participant found the specific relationship, trend or pattern that was asked for in the task. The environment where the user needs less 'time of thinking' for a specific task is obviously the environment with the best efficiency, assuming questions of equal difficulty.

To test for the effectiveness, each answer on the task given by the participant is compared to a predefined answer by the experimenter about this relationship, trend or pattern (see appendix C). For each answer, a correctness score between 0 and 2 was assigned, with 0 is incorrect, 1 is partly correct and 2 is fully correct. Therefore, having five tasks for each visualization environment, the minimum and maximum score equals respectively 0 and 10.

After the actual usability test, the participants were asked to fill in a questionnaire to test for their background and their affinity with the datasets (see appendix D). This information can be used for explaining outliers in the data set and to ensure that the background and dataset knowledge within the test group are rather equal. Furthermore, the questionnaire is used to test the user's satisfaction. Here, each participant was asked to rank the usefulness of the environment for specific tasks, the usefulness for types of dataset (naturally or randomly ordered) and the usefulness of interactive possibilities, using ratings from 1 (very weak) to 5 (very strong).

The results are tested on significance to each other. The hypothesis is that the TOSM separately is more efficient, effective and to the user's satisfaction when compared to the MAP and PCP environments for tasks related to find spatio-temporal patterns. After all, the TOSM is designed to do so. The null hypothesis (H_0) is all samples have been drawn from the same populations, assuming independency. An ANOVA test is applied on all environments to test for efficiency (N=138), effectiveness (N=47) and user's satisfaction (N=47). For testing two environments only, an independent two-sample t-test is applied, whereby equal variances are no longer assumed. A non-parametric Kolgomorov-Smirnov test is applied for two-sample testing effectiveness (N=31) and usefulness (N=30). The principle behind the Kolgomorov-Smirnov test is that independent samples have been drawn from identical populations and that the cumulative frequency distributions for the two samples are essentially similar. Strictly speaking, the Kolgomorov-Smirnov test is effective in the presence of a large number of ties, resulting from the grouping of

data into ordered categories (Blalock 1979, p. 266) as is the case with effectiveness and usefulness scores.

The second hypothesis is that the combined view (MAP-PCP-TOSM) is more effective, efficient and to the user's satisfaction compared to the each separate view for all tasks (multivariate, trend and spatio-temporal). Hence, the same techniques are applied as the previous hypothesis.

The third hypothesis is that in naturally ordered spatial datasets, like case study I and II, it is easier to convey multivariate spatio-temporal patterns, compared to applications that have no clear initial orientation (case study III and IV). Hence, H_0 is that case study I and II are more effective and efficient compared to case study III and IV. An ANOVA test is applied for all case studies. While testing two case studies separately, an independent two-sample t-test is applied for efficiency and non-parametric Kolgomorov-Smirnov test is applied for testing of effectiveness.

6.6.3 Usability test results

An overview of all test results – i.e. recorded time (in seconds), correctness and questionnaire scores - is given is appendix E. The results of the recorded time for spatio-temporal tasks are represented in box plots in figure 6.7. The box plots show that the TOSM performs better in efficiency, i.e. it takes less time to find a spatio-temporal pattern. Also, the TOSM performs better in effectiveness and usefulness. As there are three spatio-temporal tasks, the maximum score for effectiveness will be 6. When looking at the mean values for MAP, PCP and TOSM (see table 6.5) it is clear that the TOSM performs better for all usability measures. Obviously, the TOSM is a visualization technique designed for detecting patterns in spatio-temporal data and a better performance is expected. However, this performance of the TOSM is less distinct for effectiveness, where most scores ranges between 3 and 6.

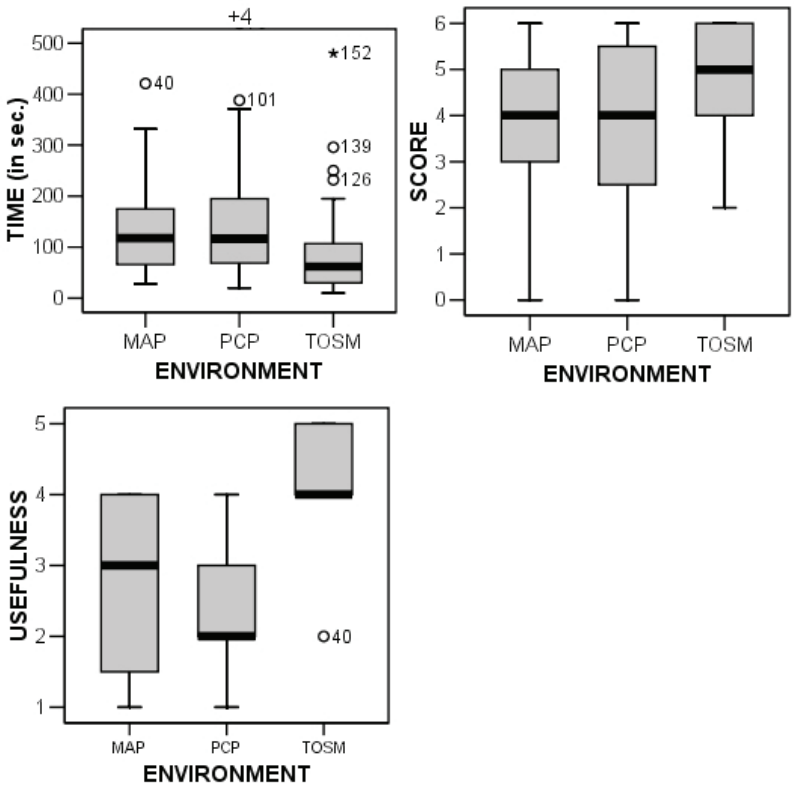


Figure 6.7 *Box plots for the separate environments MAP, PCP and TOSM for efficiency (left), effectiveness (middle) and user's satisfaction (right) for spatio-temporal tasks. Outliers are expressed above the box plot by means of a +-sign.*

Table 6.5 *Mean values for efficiency (in seconds), effectiveness (max score is 6) and usefulness using spatio-temporal tasks only.*

	Efficiency	Effectiveness	Usefulness
MAP	137.50	3.80	2.60
PCP	174.38	3.88	2.50
TOSM	84.86	4.63	4.19

Testing all separate environments and the quality measures for spatio-temporal tasks, efficiency and usefulness show significant results at a 0.05 level (table 6.6, first row). Also, when two environments are compared, significant results are observed for efficiency and user's satisfaction for MAP-TOSM and PCP-TOSM (see table 6.6). Comparing the MAP environment with the PCP

environment, no significant differences for efficiency, effectiveness and usefulness are noticed. This means that H_0 can be rejected, and that the time of recording for the TOSM, as well as the user's satisfaction about the TOSM, significantly differs from other environments for spatio-temporal tasks. However, H_0 can not be rejected for effectiveness of the TOSM against the other environments.

Table 6.6 *Test of significance (p-values) between separate environments for three usability measures for spatio-temporal tasks. [1]Testing procedure: independent two-sample t-test. [2]Testing procedure: non-parametric two-sample Kolgomorov-Smirnov test.*

	Efficiency	Effectiveness	Usefulness
ANOVA	< 0.05	0.273	<0.05
MAP-PCP	0.219[1]	0.993[2]	0.904[2]
MAP-TOSM	< 0.05[1]	0.582[2]	<0.05[2]
PCP-TOSM	< 0.05[1]	0.699[2]	<0.05[2]

Next, the combined environment (MAP-PCP-TOSM) is considered against each separate environment for all tasks. Again, the same results are observed. The box plots in figure 6.8 and the mean values in table 6.7 show that the combined environment towards the separate environments is performing better in efficiency, effectiveness and usefulness. Nevertheless, these results are less obvious for the combined environment against the TOSM environment for usability measures.

Table 6.7 *Mean values for efficiency (in seconds), effectiveness (max. score is 10) and usefulness using all tasks.*

	Efficiency	Effectiveness	Usefulness
MAP	166.64	7.13	2.75
PCP	166.83	6.81	3.19
TOSM	109.43	7.35	3.72
COMBI	93.94	8.19	4.03

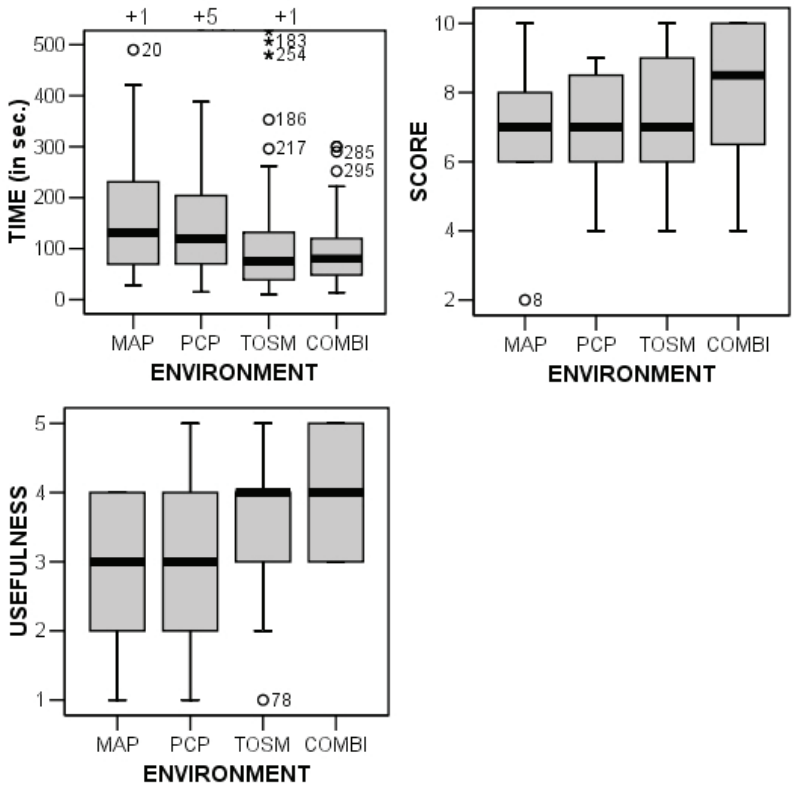


Figure 6.8 *Box plots for the separate environments MAP, PCP and TOSM for efficiency (left), effectiveness (middle) and user's satisfaction (right) for all tasks. Outliers are expressed above the box plot by means of a +-sign.*

When testing for the combined environment against all separate environments, the combined environment shows significant differences for efficiency and usefulness at a 0.05 level for all tasks (table 6.8, first row). Also, significant differences are observed in two-sample tests between the combined environment against the MAP or PCP environment for the same usability measures (table 6.8). H_0 can therefore be rejected for efficiency and user's satisfaction regarding the combined environment against all separate environments, as well as for the combined environment against the MAP or PCP environment. H_0 can not be rejected however for effectiveness regarding all environments, as well as when comparing the combined environment against the TOSM environment for all usability measures.

Table 6.8 *Test of significance (p-values) between the combined environment against separate environments for three usability measures for all tasks. [1]Testing procedure: independent two-sample t-test. [2]Testing procedure: non-parametric two-sample Kolgomorov-Smirnov test.*

	Efficiency	Effectiveness	Usefulness
ANOVA	< 0.05	0.177	<0.05
MAP-COMBI	< 0.05[1]	0.453[2]	<0.05[2]
PCP-COMBI	< 0.05[1]	0.211[2]	<0.05[2]
TOSM-COMBI	0.285[1]	0.382[2]	0.964[2]

Last, all data sets (Case study I, II, III and IV) are considered with respect to all tasks. The box plots in figure 6.9 and the mean values in table 6.9 show the results for efficiency and effectiveness. The usefulness is not measured. Clearly, case study II and IV (respectively beach management and environmental pollution application) perform somewhat better in efficiency with respect to case study I and III (respectively river flooding and housing economics application). As regards to effectiveness, case study II shows a higher average score. The increased performance for case study II can be explained, as for this case study only, the combined environment (MAP-PCP-TOSM) are utilized (see also table 6.3).

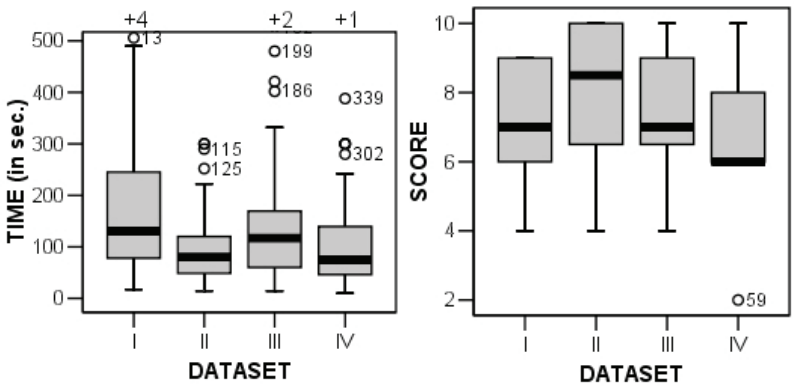


Figure 6.9 *Box plots showing efficiency (left) and effectiveness (right) of the different case studies for all tasks. Outliers are expressed above the box plot by means of a +-sign.*

Testing the case studies for all tasks, only the results for efficiency significantly differ (table 6.10, first row). In the two-sample tests significant differences for efficiency are observed for all case studies, with an exception for case study II and IV and case study III and IV

(table 6.10). There are no significant differences for effectiveness. H_0 can therefore be rejected for efficiency regarding testing all case studies. H_0 can not be rejected however for effectiveness.

Table 6.9 *Mean values for efficiency (in seconds) and effectiveness (max. score is 10) for the different case studies using all tasks.*

	Efficiency	Effectiveness
CASE I	188.42	6.94
CASE II	93.94	8.19
CASE III	136.76	7.44
CASE IV	109.17	6.93

Table 6.10 *Test of significance (p-values) between the case studies for efficiency and effectiveness for all tasks. [1]Testing procedure: independent two-sample t-test. [2]Testing procedure: non-parametric two-sample Kolgomorov-Smirnov test.*

	Efficiency	Effectiveness
ANOVA	< 0.05	0.172
CASE I-CASE II	< 0.05[1]	0.197[2]
CASE I-CASE III	< 0.05[1]	0.958[2]
CASE I-CASE IV	< 0.05[1]	1.000[2]
CASE II-CASE III	< 0.05[1]	0.415[2]
CASE II-CASE IV	0.290[1]	0.453[2]
CASE III-CASE IV	0.133[1]	0.563[2]

6.7 Discussion

Results for efficiency, effectiveness and user's satisfaction in table 6.11 show that the TOSM has high scores for spatio-temporal tasks. The PCP environment has a good score on multivariate tasks for all test measures. The MAP environment is in general the weakest environment for most tasks, directly followed by the PCP environment. The combined environment performs the best, and is very much capable for recognition of relationships, trends and patterns in multivariate spatio-temporal data. Some prudence on these conclusions needs to be considered, as our objective in this chapter is to test the environments within the prototype, and not directly to test our prototype against other (multivariate) spatio-temporal tools. Some prudence on these conclusions needs to be considered, as the objective in this chapter is to test the environments within the prototype, and not directly to test the

prototype against other (multivariate) spatio-temporal tools. Therefore, the usability test results are only justified with respect to the prototype and its separate environments. In this attempt, the test surroundings are carefully judges to optimize the test. The test participants are most young - almost non-experienced - GIS-users to prevent influences derived from a-priori knowledge from GIS. Besides, all computer equipment was identical, no participant from the same user group was sitting next to each other and participants from the focus group were excluded because of affinity with prototype and dataset. However, there are still non-provable factors, like the IQ of the participant (in particular in relation to pattern recognition using a GIS tool), the speed of handling of a participant and the way a participant is accustomed to rate the usefulness of a tool.

Table 6.11 *Ranking of the MAP, PCP, TOSM and COMBINED environments based on mean values with respect to the tasks executed (with: -- = lowest, - = low, + = high, ++ = highest, n.a. = not available, * equal score).*

Test measure	*Task*	*MAP*	*PCP*	*TOSM*	*COMBI*
Efficiency	Multivariate relationship	--	+	-	++
	Temporal trend	--	-	+	++
	Spatio-temporal pattern	-	--	+	++
	All tasks	-	--	+	++
Effectiveness	Multivariate relationship	--	++*	-	++*
	Temporal trend	++	--	-	+
	Spatio-temporal pattern	--	-	+	++
	All tasks	-	--	+	++
User's satisfaction	Multivariate relationship	--	+	-	++
	Temporal trend	n.a.	n.a.	n.a.	n.a.
	Spatio-temporal pattern	-	--	++	+
	All tasks	--	-	+	++

The 'time of thinking' as quantitative measure for efficiency needs some cautions. First, time recording is subjective to the user interpretation, and although most users did perform this with most care, a full check on each user's actions was not performed in this test. Besides, it is clear that all tasks are case specific and therefore not completely identical. Therefore, it can only be said that the 'time of thinking' for each task is an indication in how efficient a visualization environment is in portraying relationships, trends and patterns.

The test for effectiveness needs some consideration, as the experimenter predefines answers for all tasks and rates the participant's answers by correctness scores. Often, participants answered the tasks not completely or gave only an indication where patterns might be found. The experimenter checked each answer thoroughly, and gave only the maximum correctness score in case the answer was complete, or if the participant found an existing pattern not detected by the experiment. To assign correctness score will always be subjective to the scoring method, as well as the experimenter's judging. Besides, correctness scores can easily fluctuate, partly due to the affinity of a participant with pattern recognition in (multivariate) spatio-temporal datasets. Hence, this may give an explanation to the test results, where effectiveness not significant differs for (multivariate) spatio-in any environment.

In the test procedure, paired environments, like MAP-PCP, PCP-TOSM or MAP-TOSM, are not tested against other separate environments or the combined environment (MAP-PCP-TOSM). The foremost reason is that with using a paired environment it is more difficult to trace where the participant finds its patterns exactly, leading to distortion of the test results. With respect to the MAP environment, it has to be said that no legend was available at time of testing. UDIG has no standard legend for attributes in the MAP environment. Using the info tool, the user has the ability to tackle this, meanwhile getting information from all variables. Hence, for multivariate spatio-temporal tasks, this might be an advantage; with spatio-temporal tasks only, this may earlier lead to a disadvantage. Hence, some of the low performance results for the MAP environment can be explained by the lack of a legend.

The focus group clearly raised, that the prototype is more useful for naturally ordered spatial datasets than datasets where no clear spatial ordering can be found. This holds the truth, as the prototype was mainly developed with the beach management application in mind. However, general applicability can be obtained by other – even non-linear - spatial reordering techniques, as have been described in section 6.3.2. In the prototype, the linear directional line was only developed, and this may lead to falsely detected patterns when applied on randomly ordered spatial applications. In the test, two applications – the environmental pollution and the housing economics application – have no clear spatial ordering. To counteract divergent conclusions on patterns, the participants are asked to draw the directional line in a preset orientation, whereby with certainty patterns can be found.

When testing for significance between the data sets (section 6.4), differences are observed in efficiency for most case studies. The prototype, however, does not perform different on efficiency and effectiveness between naturally ordered spatial applications (case study I and II) and randomly ordered spatial applications (case study III and IV). Case study I shows large 'time of thinking' with respect to other case studies, and this might be mainly the result of the learning effect, as case study I was the first case study for all users. Case study II is the only case study that has been applied on the combined environment, leading to higher efficiency and effectiveness scores. Case study III and IV show somewhat intermediate scores. Hence, proper conclusions on the effect of different dataset types, i.e. naturally or randomly ordered, can therefore not be drawn.

The prototype – particularly towards general applicability and detecting trends and patterns in multivariate spatio-temporal data – can be extended by adding topological or proximity relationships to the geospatial units. After all, units next to each other often show similarity in attributes. A topological relation as adjacency can be used to order the units as their boundaries connect to each other. Likewise for geospatial proximity relations - i.e. the geographical distances among geospatial units - where units close to each other show similarities in attributes. Euclidean distance measures can be used to determine the distance between the units in Cartesian space. Even though the prototype has been tested, many improvements can be made to enhance common-use of the tool for multivariate spatio-temporal datasets.

6.8 Conclusions

An integrated method to visualize multivariate spatio-temporal data sets is introduced. A prototype is developed consisting of three dynamically linked views, each representing one of spatial data's components. The MAP environment gives answer to the "where" question, the PCP environment to the "what" question and the TOSM environment to the "when" question. The TOSM (Temporal Ordered Space Matrix) is new visualization technique and is a sort of schematized map whereby the rows in the matrix represent time, the columns represent geographic units, and the individual cell can be colored according the value of user-defined attributes. Datasets need to be reordered spatially for representing the geographical units from a Cartesian space into a 1D linear space. In the prototype, spatial reordering of geographical units is done by directional linearization.

A usability test is performed, consisting of 17 participants, whereby each environment is tested on significance between each other. The

assumption is that the TOSM separately is more efficient, effective and to the user's satisfaction compared to the MAP and PCP environments for tasks related to find spatio-temporal patterns. The null hypothesis is that all samples are independent and have been drawn from identical populations. Significant differences ($p < 0.05$) are observed for efficiency and user's satisfaction between the MAP, PCP and TOSM environments and for two-sample tests between MAP against TOSM and between PCP against TOSM.

The prototype is tested also against all separate environments, where the combined environment shows a significant higher efficiency and usefulness for multivariate spatio-temporal tasks. Even so, significant differences are observed in two-sample tests between the combined environment against the MAP and PCP environment for the same usability measures. Only for effectiveness, no significance differences can be found in any environment.

Although the tests results are promising, there are still many improvements that can be made to enhance general applicability of the prototype, especially in relation to spatial reordering of geographical units and towards applicability for randomly ordered spatial datasets.

Chapter 7

Conclusions

'The important thing in science is not so much to obtain new facts as to discover new ways of thinking about them'

Sir William Bragg (1862-1942)

In this study, modeling and visualizing dynamic landscape objects and their qualities have been proposed. It embraces landscape modeling with natural and artificial processes involved (Chapter 3 & 4), it illustrates decision rules within a classification (Chapter 5) and it represents results by interactive visualization techniques (Chapter 6). The developed methods are applied to beach management at the isle of Ameland, The Netherlands.

7.1 First research objective

The first research objective is to develop and apply an environmental model with landscape objects using a domain ontology. Chapter 3 describes an ontological approach for modeling objects within a beach management application. Here, beach objects suitable for nourishment are identified and conceptualized. The main conclusions are:

- The use of a product and problem ontology is able to differentiate between characteristic datasets and user-defined problem. The problem ontology describes an user-defined reasoning model that interprets natural features and processes of the real world, and is domain specific. The product ontology describes datasets characteristics, i.e. it describes the real world objects as they can be extracted from scientific and independent methods.
- Differences between a product and problem ontology define the qualities of the expected results. Internal data qualities are inherited from the dataset characteristics; 'fitness-for-use' is the result of the discrepancy between the problem and the product ontology.
- As a consequence, product and problem ontology are helpful in understanding if data will fit for use. When competence questions, rules and lists of real world objects are defined, appropriate datasets can be selected that fit these competences.
- As an example, the use of ontologies show that it is possible to manage the beach modeling process in an orderly framework, by defining objects, attributes and relationships. It demonstrates differences in spatial distribution of the objects that can be accredited to its ontological differences.
- The ontological approach illustrates that inaccuracies of the attributes influence object determination. The sensitivity analysis proved that - when applied on the uncertainty of the attribute altitude - a significant increase of objects suitable for nourishment is observed.

7.2 Second research objective

The second research objective is to extend the environmental model, covering the spatio-temporal aspects of dynamic landscape objects using fuzzy set theory and the determination of their spatial, thematic and temporal accuracies. In chapter 4, fuzzy techniques and temporal influences in the conceptualization of beach objects are applied and implemented using a spatio-temporal ontology. The effects for beach nourishments are assessed and quantified. The main conclusions are:

- The spatio-temporal ontology is able to embrace fuzziness to describe vagueness of natural phenomena. The fuzzy approach is more realistic, as the described objects are more similar to those occurring in decision procedures and to the use in the problem ontology.
- The spatio-temporal ontology also includes important temporal aspects by means of temporal membership functions. The use of temporal membership functions improve the quality assessment and as a consequence may be more generally applicable to space-time studies.
- The spatio-temporal ontology includes objects, attributes, relationships, processes and events. Insight into the qualities of these ontological features are given when plot against one another in a quality matrix. This helps the user in defining critical quality elements and can lead to different quality assessments.
- For the beach management application, it is shown that three different ontologies (i.e. a crisp spatial ontology, a fuzzy spatial ontology and a temporal fuzzy spatial ontology) can be implemented and applied to determine the beach objects suitable for beach nourishments. It is also shown that by implementing indeterminacy of the boundaries and temporal uncertainty the amount for beach nourishments decreases.
- Sensitivity analysis shows that all ontologies are susceptible for high sensitivity of attribute uncertainties, and are less sensitive to temporal uncertainties.

7.3 Third research objective

The third research objective is to develop and implement a hierarchical structure for classifying landscape objects. In chapter 5 a fuzzy decision tree classification is applied and validated that categorize beach objects by corresponding decision rules. A Bayesian hierarchical model is applied and implemented to incorporate uncertainties. The main conclusions are:

- A Bayesian hierarchical model allows one to include the uncertainty into the classification. A combination of prior knowledge with collected data results into posterior distributions, with less uncertainty.
- Posterior distributions can serve as an input for classification methods and significantly improve classification results. This method is applied for classifying the indeterminacy of the boundaries between beach objects and demonstrates better results.
- When dealing with weak prior knowledge, the Bayesian hierarchical model exhibits poor results, resulting in high uncertainty. External factors can play a significant role in one's certainty about the prior knowledge.
- Fuzzy decision tree classification portrays the classification and modeling process in a more natural way. It illustrates the decision rules and their uncertainties when propagated through the model.
- Validation by a high resolution image using fuzzy error matrices show promising results for the beach management application.

7.4 Fourth research objective

The fourth research objective is to develop and apply new techniques to visualize dynamic landscape objects and their multivariate data quality elements. Chapter 6 introduces an integrated method to visualize multivariate spatio-temporal datasets. This method is developed, applied and tested to dynamic beach objects and their quality elements. It has also been applied and tested on case studies in river flooding, housing economics and environmental pollution. The main conclusions are:

- Visualization and exploration of multivariate spatio-temporal datasets is difficult, as existing methods represent multivariate data and spatio-temporal data separately. An integrated environment is proposed and developed for the exploration and visualization of multivariate spatio-temporal data. It consists of three dynamically linked views, each representing one of spatial data's components. The MAP environment gives answer to the "where" question, the PCP environment to the "what" question and the TOSM environment to the "when" question.
- The TOSM (Temporal Ordered Space Matrix) is new and promising visualization technique for visualization of

spatio-temporal datasets. It is a kind of schematized map whereby the rows in the matrix represent time, the columns represent geographic units, and the individual cell are colored according the value of user defined attributes. For visualization purposes, datasets have to be reordered spatially for representing the geographical units from a Cartesian space into a 1D linear space.

- Relationships, trends and spatio-temporal patterns disclosed in multivariate spatio-temporal datasets are able to be explored and visualized by means of a prototype of the integrated method, as developed and explained in chapter 6.
- The MAP environment is able to represent the location of dynamic objects, but lacks capability for exploration and visualization of relationships, trends and patterns in multivariate spatio-temporal datasets.
- The PCP environment proved to be efficient, effective and useful for detecting multivariate relationships in attributes of dynamic objects.
- The TOSM environment proved to be efficient, effective and useful for finding spatio-temporal patterns in dynamic objects.
- The integrated prototype is most efficient, effective and useful to discover multivariate spatio-temporal patterns in dynamic objects.
- In the usability test, the prototype does not favour naturally ordered spatial datasets above randomly ordered spatial datasets. Some awareness towards the interpretation of the patterns in randomly ordered spatial datasets need to be considered, as the reordering of these datasets are not straightforward and may lead to ambiguous conclusions.
- The usability test shows that for all tasks the prototype and TOSM environment is more effective and useful compared to MAP and PCP environments based on mean values. However, the prototype and TOSM are not significantly more effective.
- The visualization prototype illuminates relationships, trends and patterns in the beach management application that may be of help for decision-makers. It can be easily extended and applied to other case studies.

7.5 Overall conclusions

The general objective of this study has been to explore the usefulness of new methodologies to describe, model and visualize landscape objects with indeterminate boundaries and their dynamics. To do so, several methods and techniques are compared, applied and discussed within an integrated approach to elucidate the broad scope of modeling dynamic landscape objects. The process of environmental modeling, as explained in figure 1.3, considers several stages, i.e. conceptualization, modeling, classification, validation and visualization. It is a cycle, with spatial datasets and a specific problem as starting point and object visualization and decision-making as the end of the cycle. In fact, the modelling approach aims to better understand and use what one apparently sees as beach objects on satellites images.

Conceptualization of beach objects is supported by different ontologies. Both the user's and the producer's point of view are addressed, leading to a problem and product ontology. The problem ontology serves as a reasoning model for the beach nourishment problem. Differences between the problem ontology and the product ontology indicate which data fit for use. Incorporation of spatial and thematic uncertainty of beach objects is achieved by extending the problem ontology into a spatio-temporal ontology. Here, indeterminate and dynamic aspects are defined. This is further refined by including a hierarchical classification dealing with beach objects at different scale levels and to incorporate their position in a landscape hierarchy as well as their uncertainties. The classified beach objects are visualized in a temporal ordered space matrix to enhance exploration for decision-makers.

The methodology has been applied to landscape objects on the isle of Ameland. It appears to be generally applicable as well to other studies. This is not strictly related to landscape objects only. In the visualization chapter (chapter 6) it is shown that a temporally ordered space matrix can be used for other multivariate spatio-temporal datasets. However, the integrated approach is developed and applied for environmental modeling purposes. Other case studies need to be analyzed to further illustrate the methodologies.

Appendix A

Usability test – Working sheet

Aim of usability test

The aim of the usability test is to test the visualization environments of the prototype (each separately and combined) for efficiency, usefulness and effectiveness in detecting patterns, trends and relationships in multivariate datasets.

Getting started and loading the data

- Open uDig by clicking on Windows START followed by the Run button.
- Enter: "D:\udig\1.0.5\eclipse\udig.exe" –clean
  The uDig program will start.
- Open the workbench environment by clicking on the arrow (top right corner).
- The work environment is now loaded. Add the data by using the pull-down menu layer/add, and to click on files.

The prototype enviroment (UDIG)

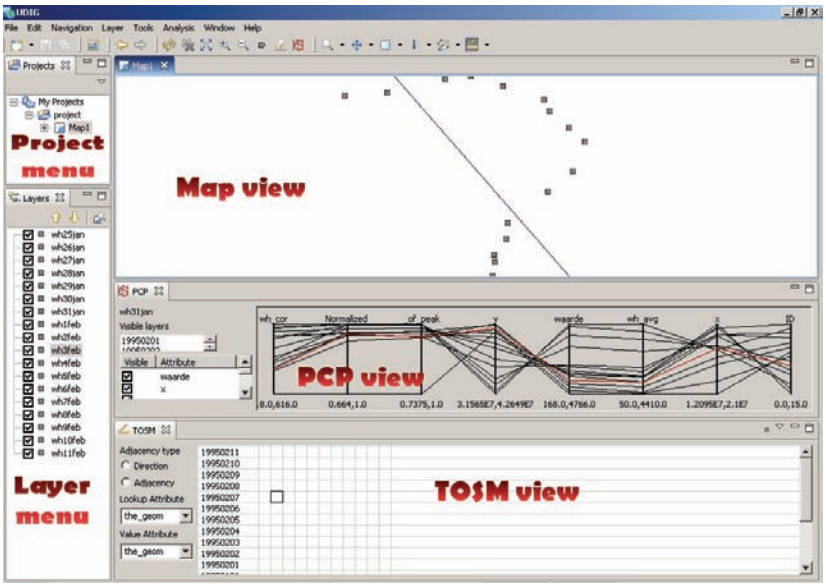


Figure A.1 *The prototype environment.*

MAP-environment

- You can draw layers by ticking the boxes in the Layer menu.
- You can redraw the dataset using the full extent button.
- Use the zoom options to zoom in or zoom out.
- Use the information tool to get information from feature attributes.

Feature attributes are variables that belong to each object, describing its characteristics. After clicking the information tool, select an object in the map and open the information catalogue. Information about this object is represented in this window (see figure A.2).

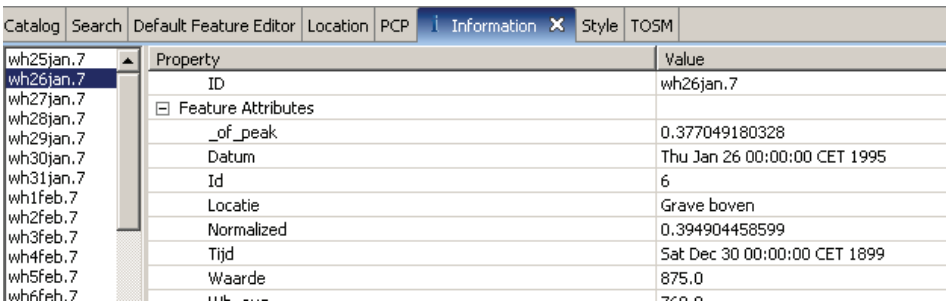


Figure A.2 *Information catalogue*

PCP-environment

To open the PCP-view, use the PCP-button on the toolbar.

- In the PCP-view you can select a line. Each line represents an object in one of the temporal maps (see figure A.3).
- In the PCP-view you can select which layer you like to see (visible layers) (see figure A.3).
- In the PCP-view you can also select which attribute you like to see (see figure A.3).
- The vertical axes in the PCP-view show the variables, showing the minimum and maximum values as legend on the bottom of the axes. The value of each line can be revealed using the mouse pointer near the vertical axis.

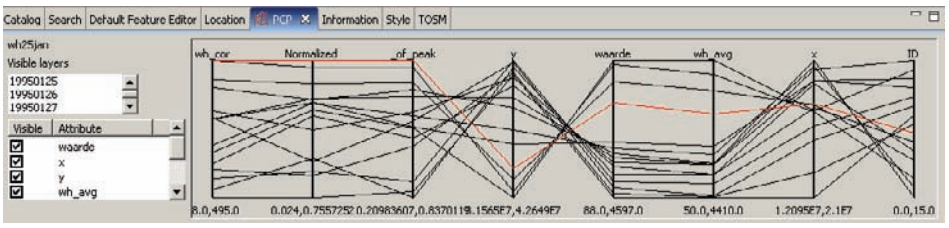


Figure A.3 *PCP-environment*

TOSM-environment

- To open the TOSM view, use the TOSM button underneath the info button and draw a directional line in the map view (as shown in figure 2 – MAP view).
- You can make your selection for the TOSM view using Lookup Attribute to select object-ID numbers of the data set and selecting one of the feature attributes as Value Attribute. The TOSM will now be drawn (see figure A.4).
- You can click on a cell. Each cell represents an object in one of the temporal maps.
- Using the mouse pointer above a cell, information of Lookup Attribute and Value Attribute is shown.

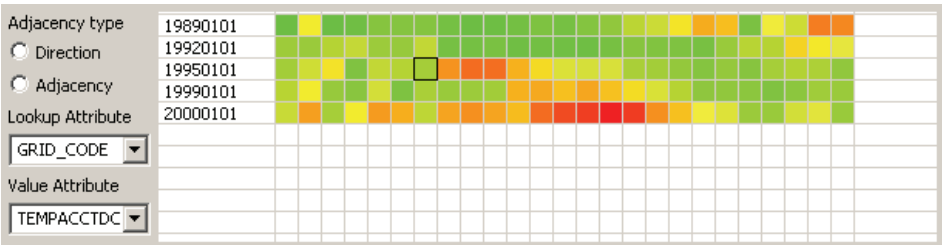


Figure A.4 *TOSM-environment*

Case study I: River flooding application

The data set consists of 16 measurement points along the river Maas, the Netherlands (see figure A.5). For each measurement point, water levels are measured on daily basis from 25 January 1995 to 11 February 1995.

The dataset consists of several attributes:

ID =	id number of measurement object (there is no particular order in ID number)
x =	x-coordinate
y =	y-coordinate
province =	provinces of the Netherlands
major_water=	rivers and other major water objects in the Netherlands
waarde =	water level at measurement point in meters above Dutch Sea Level.
wh_avg =	mean average water level at measurement point in meters above Dutch Sea Level.
wh_cor =	corrected water levels in meters. *wh-cor* = *waarde-wh_avg*.
%_of_peak =	percentage of corrected water level in relation to maximum peak water level. % *_of_peak* = (wh_cor/wh_cor_{peak}). Values between -0.03 and 1.00
Normalized =	normalized corrected water level per measurement point in relation to peak water level. *Normalized* = ($wh\_cor-wh\_cor_{min})/(wh\_cor_{peak}-wh\_cor_{min})$. A value between 0 and 1, with 0 is lowest water level (wh_cor_{min}) and 1 is peak water level (wh_cor_{peak}).

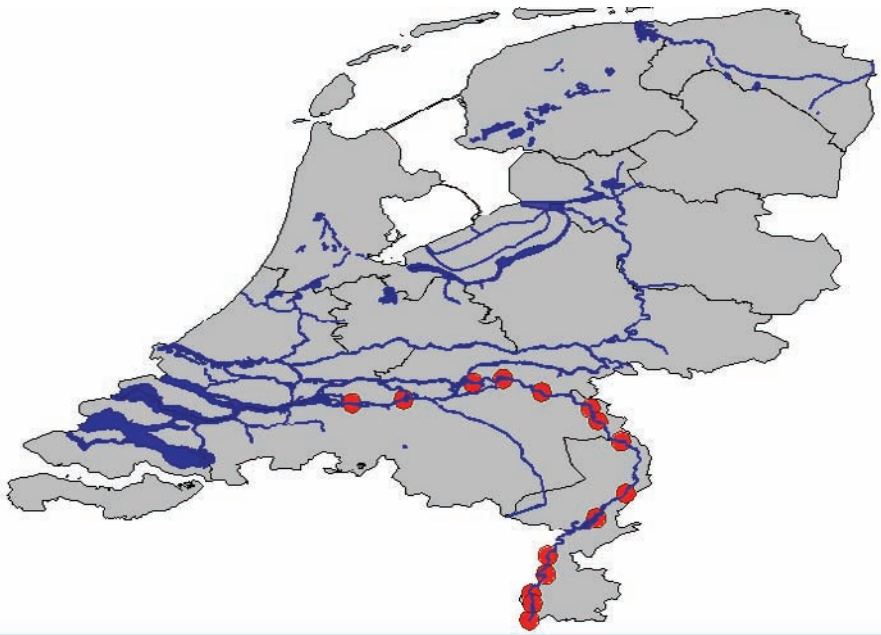


Figure A.5 *Case I, river flooding application, The Netherlands.*

Case study II: Beach management application

The data set consists of 26 beach compartments along the dynamic coast of Ameland (a barrier island) in The Netherlands (see figure A.6). The dataset consist of data from 1989, 1992, 1995, 1999 and 2000. The beach compartments are classified based on altitude, vegetation and wetness. Beach compartments should be non-vegetated, dry beaches. Besides, for each compartment several spatial data quality elements are represented.

The dataset consists of several attributes:

Grid_code = Grid-code of compartment (ID number)

Posaccrel = Relative positional accuracy. Positional accuracy in comparison to the map (in meters per data set).

Posaccgrid= Gridded positional accuracy. Positional accuracy related to grid conversions (in meters per data set).

Themacccc = Thematic accuracy – classification correctness. Uncertainty of compartments classified as non-vegetated, dry beach (0 = certain, 1 = uncertain)

Themaccqa = Thematic accuracy – attribute accuracy. Accuracy of elevation measurements (in centimeter per data set)

Tempacctdc = Temporal accuracy – time of data capture. Correctness of compartments classiefied as non-vegetated, dry beach in relation to time of data capture (tide influences, vegetation cycle) (0 = certain, 1 = uncertain)

Completene = Completeness of the data set. (0 = complete, 1 = incomplete)

Volume = Amount of eroded beach volume (in m^3, negative means erosion, positive means sedimentation)

Suitable = Suitability for beach nourishments per compartment (0 = not suitable, 1 = suitable)

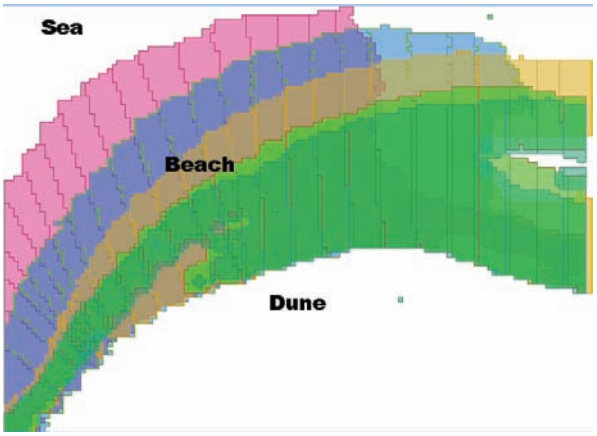


Figure A.6 *Case II, beach management application, Ameland, The Netherlands.*

Case study III: Housing economics application

The data set consists of 45 municipalities in province Overijssel, The Netherlands (see figure A.7). For each municipality several housing economics are represented for the period of 1991-1997.
Keep in mind that municipalities with cities have high number of inhabitants and population density.

The dataset consists of several attributes:

Objectid =	Object ID number.
MU-name =	Municipality name
Pop =	Population (= number of inhabitants)
Popdens =	Population per square kilometre (no/km^2)
Houses =	Number of households
Hodens =	Households per square kilometre (no/km^2)
Newho =	Number of newly build houses

Figure A.7 *Case III, housing economics application, Overijssel, The Netherlands.*

Case study IV: Environmental pollution application

The data set consists of 12 provinces in The Netherlands (see figure A.8). For each province several environmental pollution sources are represented for the period of 1995-2004. The pollution has strong relationships with farming activities.

The dataset consists of several attributes:

Province = Province number (ID number)
Provname = Name of province
Phosphate = Quantity of phosphate
Potassium = Quantity of potassium
Nitrogen = Quantity of nitrogen

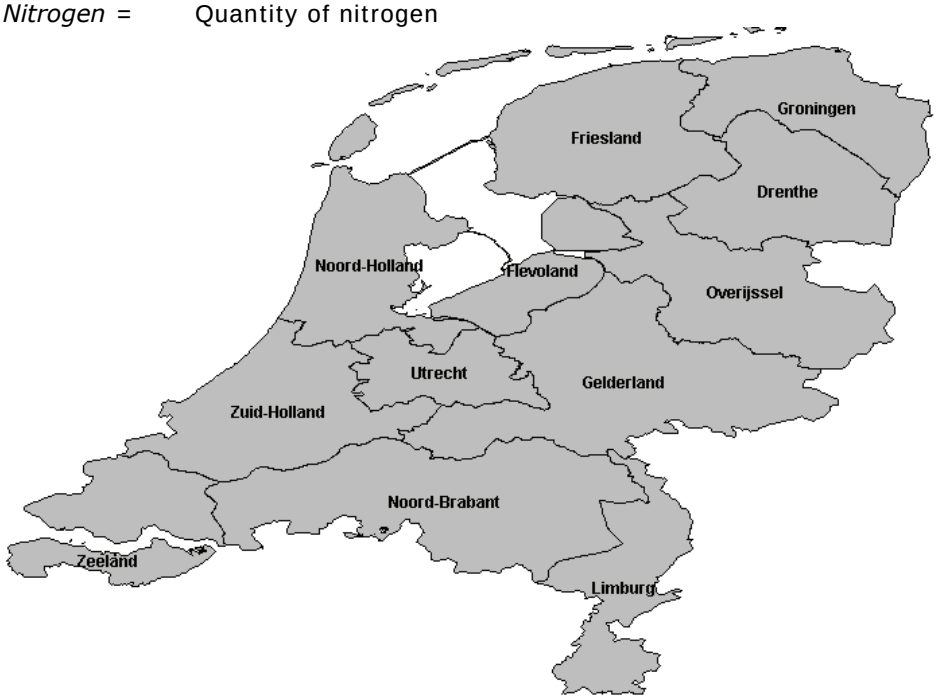


Figure A.8 Case IV, environmental pollution application, The Netherlands.

Appendix B

Exercises – User Group 1

CASE STUDY I in MAP Environment only

- ☐ Start Udig using the Usability Test working sheet: "Getting started and loading the data"
- ☐ Open Case study I. Find your data in D:\maasdata and load all files by selecting them using SHIFT & LEFT mouse button.
- ☐ Read carefully about Case study I using the Usability Test working sheet: "Case study I: River flooding application"
- ☐ Organize the data set, with completely on top *wh11feb* and *wh25jan* as the bottom layer.
- ☐ Now, you are going to try to find trends, relationships and patterns using the MAP environment only. Read about the MAP environment using the Usability Test working sheet: "MAP-environment".

!!!! After reading the question, start the stopwatch.
!!!! Stop the stopwatch as soon as you have found a pattern or relationship.

Question 1.1
Can you describe and explain the relation between variable wh_cor and variable %_of_peak and variable Normalized at wh1feb (= 1 February 1995)?

Time: min sec

Answer:

Question 1.2
Can you describe and explain the trend of variable Normalized in time?

Time: min sec

Answer:

Question 1.3
Can you describe and explain the pattern between location and variable Normalized in time?

Time: min sec

Answer:

Question 1.4
Can you describe and explain the pattern between location and variable wh_avg and in time?

Time: min sec

Answer:

Question 1.5
Can you describe and explain the pattern between location and variable wh_cor in time?

Time: min sec

Answer:

CASE STUDY III in PCP Environment only

- ☐ Start Udig using the Usability Test working sheet: "Getting started and loading the data"
- ☐ Open Case study III. Find your data in D:\overijsseldata and load all files by selecting them using SHIFT & LEFT mouse button.
- ☐ Read carefully about Case study III using the Usability Test working sheet: "Case study III: Housing economics application."
- ☐ Organize the data set in correct temporal order, with *1997* as top layer and *1990* as bottom layer, using arrow up and down.
- ☐ Now, you are going to try to find trends, relationships and patterns using the PCP environment only. Read about the PCP environment using the Usability Test working sheet: "PCP-environment".

!!!! After reading the question, start the stopwatch.
!!!! Stop the stopwatch as soon as you have found a pattern or relationship.

Question 2.1
Can you describe and explain the relation between variable Pop_97 and variable Newho_97 and variable Houses_97 at 1997?

Time: min sec

Answer:

Question 2.2
Can you describe and explain the trend of variable Newho in time for cities? NOTE: Cities have high population and density.

Time: min sec

Answer:

Question 2.3
Can you describe and explain the pattern between location and variable Popdens in time?

Time: min sec

Answer:

Question 2.4
Can you describe and explain the pattern between location and variable Houses in time?

Time: min sec

Answer:

Question 2.5
Can you describe and explain the pattern between location and variable Hodens in time?

Time: min sec

Answer:

CASE IV in TOSM Environment only

- ☐ Start Udig using the Usability Test working sheet: "Getting started and loading the data"
- ☐ Open Case study IV. Find your data in D:\environmentdata and load all files by selecting them using SHIFT & LEFT mouse button.
- ☐ Read carefully about Case study IV using the Usability Test working sheet: "Case study IV: Environmental pollution application"
- ☐ Organize the data set in correct temporal order, with *environ04* as top layer and *environ95* as bottom layer, using arrow up and down.
- ☐ Now, you are going to try to find trends, relationships and patterns using the TOSM environment only. Read about the TOSM environment using the Usability Test working sheet: "TOSM-environment".
- ☐ Create a directional line from Southeast to Northwest (see Usability Test working sheet: "TOSM-environment").

!!!! After reading the question, start the stopwatch.
!!!! Stop the stopwatch as soon as you have found a pattern or relationship.

Question 3.1
Can you describe and explain the relation between variable POTASSIUM and variable NITROGEN and variable PHOSPHATE for 1995?

Time: min sec

Answer:

Question 3.2
Can you describe and explain the trend of variable NITROGEN in time?

Time: min sec

Answer:

Question 3.3
Can you describe and explain the pattern between location and variable NITROGEN in time?

Time: min sec

Answer:

Question 3.4
Can you describe and explain the pattern between location and variable POTASSIUM in time?

Time: min sec

Answer:

Question 3.5
Can you describe and explain the pattern between location and variable PHOSPHATE in time?

Time: min sec

Answer:

CASE II in MAP-TOSM-PCP Environment

- ☐ Start Udig using the Usability Test working sheet: "Getting started and loading the data"
- ☐ Open Case study II. Find your data in D:\amelanddata and load all files by selecting them using SHIFT & LEFT mouse button.
- ☐ Read carefully about Case study II using the Usability Test working sheet: "Case study II: Beach management application"
- ☐ Organize the data set in correct temporal order, with *join2000* as top layer and *join1989* as bottom layer, using arrow up and down.
- ☐ Now, you are going to try to find trends, relationships and patterns using the ALL environments.

!!!! After reading the question, start the stopwatch.
!!!! Stop the stopwatch as soon as you have found a pattern or relationship.

Question 4.1
Can you describe and explain the relation between variable Volume (eroded beach volume) and variable TEMPACCTDC (temporal accuracy) and variable THEMACCC (thematic accuracy) at 2000?

Time: min sec

Answer:

Question 4.2
Can you describe and explain the trend of variable COMPLETENE in time?

Time: min sec

Answer:

Question 4.3
Can you describe and explain the pattern between location and variable THEMACCC in time?

Time: min sec

Answer:

Question 4.4
Can you describe and explain the pattern between location and variable TEMPACCTDC in time?

Time: min sec

Answer:

Question 4.5
Can you describe and explain the pattern between location and variable suitable in time?

Time: min sec

Answer:

Appendix C

Answer sheet for tasks - User group 1.

Case study I: River flooding application

question	*relation/trend/pattern*
1.1	Strong positive relation between variables *wh_cor*, *%_of_peak* and *Normalized*. All variables are high at 1 February 1995, as it was the peak flow of the flooding.
1.2	The trend of variable Normalized is composed of low values at early days of flooding, high values during peak flow, and low values at last days of flooding. This is the results of the peak flow that migrates through the river.
1.3	The spatio-temporal pattern of variable *Normalized* shows a migrating pattern of the peak flow from southeast to northwest, equal to the direction of flow.
1.4	The spatio-temporal pattern of variable *wh_avg* shows a high values from southeast to northwest, as the average water level is based on the altitude.
1.5	The spatio-temporal pattern of variable *wh_cor* shows no prominent pattern between location and the corrected water height. There is no clear pattern in the height of the peak flow from southeast to northwest.

Case study III: Housing economics application

question	*relation/trend/pattern*
2.1	There is a strong positive relation between variables *Pop_97*, *Newho_97* and *Houses_97* in 1997. Cities have high population and also the most houses and newly build houses.
2.2	The trend in *Newho* is slowly increasing as a result of small growth on the housing market.
2.3	The spatio-temporal pattern of variable *Popdens* is more or less constant. There is no real growth in population.
2.4	The spatio-temporal pattern of variable *Houses* is slightly increasing, as a result of growth on the housing market.
2.5	The spatio-temporal pattern of variable *Hodens* shows a slight increase in house density for cities.

Case study IV: Environmental pollution application

question	*relation/trend/pattern*
3.1	There is a strong positive relation between variables *POTASSIUM*, *NITROGEN* and *PHOSPHATE,* as a consequence of occurrence of farming activities.
3.2	The trend in *NITROGEN* is decreasing, due to less harmful farming activities.
3.3	The spatio-temporal pattern of variable *NITROGEN* shows high values for the eastern and south eastern provinces, as there are the most farming activities.
3.4	The spatio-temporal pattern of variable *POTASSIUM* shows high values for the eastern and south eastern provinces, as there are the most farming activities.
3.5	The spatio-temporal pattern of variable *PHOSPHATE* shows high values, though less distinctive, for the eastern and south eastern provinces, as there are the most farming activities.

Case study II: Beach management application

question	*relation/trend/pattern*
4.1	There is a strong positive relation between thematic and temporal inaccuracies and erosion. Erosion influence thematic and temporal accuracies of the beach objects.
4.2	The trend in *completeness* is decreasing. The dataset is more complete for the last three monitoring events.
4.3	The spatio-temporal pattern of variable *THEMACCC* shows a peak for the last three monitoring events in central part of the study area, due to erosion.
4.4	The spatio-temporal pattern of variable *TEMPACCTDC* shows a peak for the last three monitoring events in central part of the study area, due to erosion.
4.5	The spatio-temporal pattern of variable *suitable* shows suitability for beach nourishment from west to central part in the last three monitoring events.

Appendix D

Questionnaire – Usability Test

In order to be able to understand and recognize the ways you explain and describe trends, relationships and patterns you discovered in the different studies, I would like you to briefly answer a few questions on your background and experience. Besides, I want to know about your affinity with the data set and what you think about the usefulness of the different visualization environments.

General information

1) Name ----------

2) Age ----------

3) Nationality ----------

4) How is your English?
(Please tick)
- ☐ Good, it is my mother tongue
- ☐ Yes, I speak English well, although it is not my mother tongue
- ☐ Yes, I speak English reasonable well
- ☐ No, I cannot speak English

Geography education, background and knowledge
(If you have no job, please go to question 7)

5) What is your profession?

---

6) What are your core activities?

---

7) Are you currently a student?
(Please tick)
- ☐ Yes
- ☐ No

8) If so, what educational programme are you following?

---

9) Do you have any experience with Geographic Information Systems (GIS)?
(Please tick)

- ☐ Yes, a lot (> 3 years)
- ☐ Yes, a little (< 3 years)
- ☐ No, not at all (*go to question 11*)

10) Which GIS tools do you often use? *(in case you use several, please name them)*

---

11) Do you have any experience with (geo)visualization tools (maps, animations, PCP, Chernoff faces, scattermatrix, etc, etc).?
(Please tick)

- ☐ Yes, a lot (> 3 years)
- ☐ Yes, a little (< 3 years)
- ☐ No, not at all (*go to question 14*)

12) Which visualization tools do you often use? *(in case you use several, please name them)*

---

13) Do you know what 'brushing' is? So yes, please describe?
(Please tick)

- ☐ Yes, ----------
- ☐ No

14) Have you ever worked with UDIG before?
(Please tick)

- ☐ Yes, often
- ☐ Yes, occasionally
- ☐ Yes, but only once or twice
- ☐ Never

15) Please describe you own 'background' in words:

Affinity with the data sets

16) Do you have any affinity with Maas data set (river flooding, water levels)
(1 = very weak, 2 = weak, 3 = not weak/not strong, 4 = strong, 5 = very strong) *(Please tick)*

1	2	3	4	5
O	O	O	O	O

If you choose 4 or 5, in which town enters the Maas River the Netherlands?

---

17) Do you have any affinity with Overijssel data set (housing economics, population)
(1 = very weak, 2 = weak, 3 = not weak/not strong, 4 = strong, 5 = very strong) *(Please tick)*

1	2	3	4	5
O	O	O	O	O

If you choose 4 or 5, in which city would you expect the highest population?

---

18) Do you have any affinity with Environmental data set (pollution)
(1 = very weak, 2 = weak, 3 = not weak/not strong, 4 = strong, 5 = very strong) *(Please tick)*

1	2	3	4	5
O	O	O	O	O

If you choose 4 or 5, why can you expect higher pollution of farming activities in East Netherlands?

---

19) Do you have any affinity with Ameland data set (coastal erosion, beach nourishments, quality elements)
(1 = very weak, 2 = weak, 3 = not weak/not strong, 4 = strong, 5 = very strong)*(Please tick)*

1	2	3	4	5
O	O	O	O	O

If you choose 4 or 5, what is the cause of this severe coastal erosion at Ameland?

Usefulness

20) How would you rank the usefulness of the MAP-view for detecting multivariate relationships in the data sets?
(1 = very weak, 2 = weak, 3 = not weak/not strong, 4 = strong, 5 = very strong)*(Please tick)*

1	2	3	4	5
O	O	O	O	O

21) How would you rank the usefulness of the MAP-view for detecting spatio-temporal patterns in the data sets?
(1 = very weak, 2 = weak, 3 = not weak/not strong, 4 = strong, 5 = very strong) *(Please tick)*

1	2	3	4	5
O	O	O	O	O

22) How would you rank the usefulness of the PCP-view for detecting multivariate relationships in the data sets?
(1 = very weak, 2 = weak, 3 = not weak/not strong, 4 = strong, 5 = very strong) *(Please tick)*

1	2	3	4	5
O	O	O	O	O

23) How would you rank the usefulness of the PCP-view for detecting spatio-temporal relationships in the data sets?
(1 = very weak, 2 = weak, 3 = not weak/not strong, 4 = strong, 5 = very strong) *(Please tick)*

1	2	3	4	5
O	O	O	O	O

24) How would you rank the usefulness of the TOSM-view for detecting multivariate relationships in the data sets?
(1 = very weak, 2 = weak, 3 = not weak/not strong, 4 = strong, 5 = very strong) *(Please tick)*

1	2	3	4	5
O	O	O	O	O

25) How would you rank the usefulness of the TOSM-view for detecting spatio-temporal relationships in the data sets?
(1 = very weak, 2 = weak, 3 = not weak/not strong, 4 = strong, 5 = very strong) *(Please tick)*

1	2	3	4	5
O	O	O	O	O

26) How would you rank the usefulness of the combined-views for detecting multivariate relationships in the data sets?
(1 = very weak, 2 = weak, 3 = not weak/not strong, 4 = strong, 5 = very strong)*(Please tick)*

1	2	3	4	5
O	O	O	O	O

27) How would you rank the usefulness of the combined-views for detecting spatio-temporal relationships in the data sets?

(1 = very weak, 2 = weak, 3 = not weak/not strong, 4 = strong, 5 = very strong) *(Please tick)*

1	2	3	4	5
O	O	O	O	O

28) How would you rank the usefulness of the combined-views for detecting multivariate spatio-temporal relationships in the data sets?
(1 = very weak, 2 = weak, 3 = not weak/not strong, 4 = strong, 5 = very strong) *(Please tick)*

1	2	3	4	5
O	O	O	O	O

29) The Maas data set and the Ameland data set are characterised by linear features, respectively points along a river and compartments along the coast. These data sets are so-called naturally ordered spatial datasets.
How would you rank the usefulness of the prototype for detecting patterns and relationships for these naturally ordered spatial datasets?
(1 = very weak, 2 = weak, 3 = not weak/not strong, 4 = strong, 5 = very strong) *(Please tick)*

1	2	3	4	5
O	O	O	O	O

30) On the other hand, the housing economics data set and the environmental data set are randomly ordered spatial dataset.
How would you rank the usefulness of the prototype for detecting patterns and relationships for these randomly ordered spatial datasets?
(1 = very weak, 2 = weak, 3 = not weak/not strong, 4 = strong, 5 = very strong) *(Please tick)*

1	2	3	4	5
O	O	O	O	O

31) How would you rank the usefulness of the interactivity (i.e. linking of the views, highlighting/blinking, etc.) between the combined views?
(1 = very weak, 2 = weak, 3 = not weak/not strong, 4 = strong, 5 = very strong) *(Please tick)*

1	2	3	4	5
O	O	O	O	O

32) Do you have any final remarks, comments or recommendations?

Thank you for completing this questionnaire!

Appendix E

Table E.1 *Results of efficiency (in seconds) for each environment (MAP-PCP-TOSM-COMBINED). With: # = number of participant, UG = user group, n.a. = not available. Tx.x = task, with T1.x = river flooding, T2.x = housing economics, T3.x = environmental pollution, T4.x = beach management.*

#	UG	MAP					PCP					TOSM					COMBINED				
		T1.1	*T1.2*	*T1.3*	*T1.4*	*T1.5*	*T2.1*	*T2.2*	*T2.3*	*T2.4*	*T2.5*	*T3.1*	*T3.2*	*T3.3*	*T3.4*	*T3.5*	*T4.1*	*T4.2*	*T4.3*	*T4.4*	*T4.5*
1	1	375	158	267	245	240	60	154	99	270	252	63	17	75	61	30	70	39	75	80	40
2	1	47	78	66	63	68	186	118	147	239	166	145	132	98	30	90	170	205	134	100	159
3	1	250	490	216	225	n.a.	204	120	68	123	110	n.a.	80	13	20	10	130	50	130	74	85
4	1	416	292	163	119	123	15	69	185	48	n.a.	33	150	11	122	62	173	78	222	87	97
5	1	237	175	60	128	114	150	173	74	155	20	180	10	20	10	15	150	20	76	15	135
6	1	39	213	164	143	101	52	204	117	35	34	55	84	71	60	48	129	90	252	90	36
		T3.1	*T3.2*	*T3.3*	*T3.4*	*T3.5*	*T1.1*	*T1.2*	*T1.3*	*T1.4*	*T1.5*	*T2.1*	*T2.2*	*T2.3*	*T2.4*	*T2.5*	*T4.1*	*T4.2*	*T4.3*	*T4.4*	*T4.5*
7	2	134	35	127	63	65	87	n.a.	129	77	75	132	32	41	42	107	119	19	107	80	76
8	2	78	49	30	n.a	n.a	297	91	205	42	54	98	41	14	24	39	50	61	61	23	25
9	2	195	195	182	72	61	310	209	169	106	174	149	176	195	81	159	80	45	156	78	102
10	2	300	60	300	120	300	288	610	840	90	60	120	60	60	60	60	300	120	120	60	120
11	2	n.a.	n.a.	n.a.	n.a.	n.a.	80	134	165	116	220	195	70	76	20	124	n.a.	n.a.	n.a.	n.a.	n.a.
12	2	147	242	175	170	82	65	130	539	371	692	526	480	232	126	77	84	14	33	27	91
		T2.1	*T2.2*	*T2.3*	*T2.4*	*T2.5*	*T3.1*	*T3.2*	*T3.3*	*T3.4*	*T3.5*	*T1.1*	*T1.2*	*T1.3*	*T1.4*	*T1.5*	*T4.1*	*T4.2*	*T4.3*	*T4.4*	*T4.5*
13	3	298	192	114	137	71	n.a	221	280	121	108	506	212	296	n.a.	n.a.	n.a	290	208	88	119
14	3	137	124	117	64	41	48	90	75	44	23	134	17	27	67	71	60	79	133	106	89
15	3	74	122	78	53	28	210	114	70	22	52	82	17	130	49	67	107	53	23	47	33
16	3	402	669	421	332	67	31	130	637	110	388	353	115	250	110	480	60	23	106	25	57
17	3	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	262	89	89	31	108	300	47	65	16	45

Table E.2 *Results of effectiveness (scores) for each environment (MAP-PCP-TOSM-COMBINED). With: # = number of participant, UG = user group, n.a. = not available. Cx.x = correctness score, with C1.x = river flooding, C2.x = housing economics, C3.x = environmental pollution, C4.x = beach management. Scores: 0 = incorrect, 1 = partly correct and 2 = correct.*

#	UG	MAP						PCP						TOSM						COMBINED					
		C1.1	*C1.2*	*C1.3*	*C1.4*	*C1.5*		*C2.1*	*C2.2*	*C2.3*	*C2.4*	*C2.5*		*C3.1*	*C3.2*	*C3.3*	*C3.4*	*C3.5*		*C4.1*	*C4.2*	*C4.3*	*C4.4*	*C4.5*
1	1	0	2	2	0	2		2	1	2	2	2		0	0	2	2	2		0	0	2	2	2
2	1	2	2	0	2	1		2	2	2	0	0		2	2	2	2	2		1	2	2	2	2
3	1	2	2	0	2	1		2	1	2	2	2		1	2	2	2	2		2	2	2	2	2
4	1	2	2	1	2	0		1	2	1	1	2		1	2	2	2	1		1	2	2	2	1
5	1	2	2	2	2	1		2	1	2	1	1		2	2	2	1	2		2	2	2	2	2
6	1	1	2	2	2	1		0	2	1	1	0		1	2	1	1	1		2	1	1	1	2
		C3.1	*C3.2*	*C3.3*	*C3.4*	*C3.5*		*C1.1*	*C1.2*	*C1.3*	*C1.4*	*C1.5*		*C2.1*	*C2.2*	*C2.3*	*C2.4*	*C2.5*		*C4.1*	*C4.2*	*C4.3*	*C4.4*	*C4.5*
7	2	1	2	1	1	1		2	n.a.	1	1	0		2	1	2	1	1		2	2	2	2	2
8	2	0	2	0	0	0		2	2	0	1	2		2	2	1	0	1		2	2	2	2	1
9	2	1	2	1	1	1		2	2	0	0	0		0	1	2	2	2		2	1	0	0	1
10	2	2	2	2	2	0		2	1	2	2	2		2	1	2	1	1		2	2	2	2	2
11	2	n.a.	n.a.	n.a.	n.a.	n.a.		2	0	2	1	1		1	1	2	1	1		n.a.	n.a.	n.a.	n.a.	n.a.
12	2	2	2	0	0	2		2	2	2	2	1		1	2	2	2	2		2	2	1	1	2
		C2.1	*C2.2*	*C2.3*	*C2.4*	*C2.5*		*C3.1*	*C3.2*	*C3.3*	*C3.4*	*C3.5*		*C1.1*	*C1.2*	*C1.3*	*C1.4*	*C1.5*		*C4.1*	*C4.2*	*C4.3*	*C4.4*	*C4.5*
13	3	2	2	0	2	2		n.a.	2	2	1	1		2	1	1	n.a.	n.a.		n.a.	2	2	2	2
14	3	0	2	2	2	2		2	2	2	0	2		2	2	1	1	1		1	2	2	0	1
15	3	2	2	2	2	2		0	0	2	2	2		2	0	1	2	1		2	2	1	1	0
16	3	1	2	2	2	2		2	2	2	0	2		2	2	2	2	1		2	2	2	2	2
17	3	n.a.	n.a.	n.a.	n.a.	n.a.		n.a.	n.a.	n.a.	n.a.	n.a.		2	2	2	2	1		2	2	2	2	2

Appendix E

Table E.3 *Results of user's satisfaction (scores) for each environment (MAP-PCP-TOSM-COMBI) derived from the questionnaire. With: # = number of participant, INTERACT = interactivity, multi = multivariate relationships, temp = spatio-temporal patterns, mst = multivariate spatio-temporal patterns, natural = naturally ordered spatial datasets, random = randomly ordered spatial datasets, Qxx = question number, n.a. = not available. Scores: 1 = very weak, 2 = weak, 3 = not weak/not strong, 4 = strong, 5 = very strong.*

#	MAP		PCP		TOSM		COMBI			DATASET		INTERACT
	multi	temp	multi	temp	multi	temp	multi	temp	mst	natural	random	
	Q20	*Q21*	*Q22*	*Q23*	*Q24*	*Q25*	*Q26*	*Q27*	*Q28*	*Q29*	*Q30*	*Q31*
1	1	1	3	2	5	5	3	3	4	4	4	5
2	2	4	5	2	4	4	4	4	5	4	3	5
3	3	3	2	2	4	4	4	4	4	3	2	4
4	2	2	4	3	4	4	5	5	5	4	3	4
5	2	1	4	2	2	4	3	3	3	3	3	4
6	3	3	4	2	3	4	5	5	5	4	5	5
7	3	3	3	2	3	4	4	3	3	3	3	3
8	4	4	3	4	2	2	5	5	4	4	3	4
9	4	4	4	4	4	4	4	4	4	4	4	4
10	4	4	4	4	4	4	4	4	4	3	4	3
11	3	1	4	2	3	5	5	5	4	4	3	4
12	4	4	3	3	4	4	4	4	4	4	3	4
13	4	2	5	2	3	5	3	3	4	5	4	5
14	3	1	5	1	1	5	5	5	4	4	3	5
15	1	2	5	3	3	5	4	4	5	4	3	4
16	3	3	4	2	3	4	3	3	4	2	n.a.	3
AVG	2.9	2.6	3.9	2.5	3.3	4.2	4.1	4.0	4.1	3.7	3.3	4.1

Bibliography

Aalders, H.J.G.L. and Morrison, J. (1998). Spatial Data Quality for GIS. Geographic Information Research: Trans-Atlantic Perspectives, Eds. M. Craglia and H. Onsrud, Taylor & Francis, London/Bristol, pp. 463-475.

Acevedo, W. and Masuoka, P. (1997) Time-series animation techniques for visualizing urban growth. *Computers & Geosciences*, **23**, pp. 423-435.

Andrews, D.F. (1972) Plots of high-dimensionality data, *Biometrics*, **29**, pp 125-136.

Andrienko, G. and Andrienko N. (1999). Interactive maps for visual data exploration. *International Journal Geographical Information Science*, **13(4)**, pp 355-374.

Andrienko, G. and Andrienko N. (2001). Exploring spatial data with dominant attribute map and parallel coordinates. *Computers, Environment and Urban Systems*, **25(1)**, pp 5-15.

Andrienko, N. and Andrienko, G. (2004) Interactive visual tools to explore spatiotemporal variation. In: (Ed.) Proceedings of the Working Conference on Advanced Visual Interfaces AVI 2004, ed. by Coastabile, M.F., Gallipoli, May 2004, (ACM Press, New York 2004) pp. 417-420.

Andrienko, N., Andrienko, G. and Gatalsky, P. (2003) Exploratory spatio-temporal visualization: an analytical review. *Journal of Visual Languages & Computing* **14(6):**503-541.

Anselin, L. Syabri, I. and Smirnov, O. (2002). Visualizing multivariate spatial correlation with dynamically linked windows. Proceedings of the SCISS Specialist Meeting 'New tools for spatial data analysis', Santa Barbara, USA.

Aristotle, (350 BC). Metaphysics: Book IV. Translated by W. D. Ross. Available online at http://classics.mit.edu/Aristotle/metaphysics.4.iv.html (Accessed 8 May 2002).

Baatz, M. and Schaepe, A. (2000). "Multiresolution Segmentation – an optimization approach for high quality multi-scale image segmentation." In: Strobl/Baschke/Griesebner (eds.): Angewandte Geographische Informationsverarbeitung XII, Wichmann-Verlag, Heidelberg, pp 12-23.

Banerjee, S., Carlin, B. P. and Gelfand, A. E. (2004) Hierarchical Modeling and Analysis for Spatial Data. Monographs on Statistics and Applied Probability 101. Chapman&Hall/CRC. Washington DC, USA, 452 p.

Bédard, Y. and Vallière, D. (1995). Qualité des données à référence spatiale dans un contexte gouvernemental. Research report, Université Laval, Québec, Canada, 55p.

Bennett, D.A. and Armstrong M.P. (2001). Fundamentals of geographic information systems (GIS). In Manual of Geospatial Science and Technology, edited by J. Bossler, London: Taylor and Francis, 2001, pp. 411-430.

Benz, U.C., Hofmann, P., Willhauck, G., Lingenfelder, I., and Heynen, M. (2004). Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information, *ISPRS Journal of Photogrammetry & Remote Sensing*, **58**, 2004, pp. 239-258.

Binaghi, E., Brivio, P.A., Ghezzi, P. and Rampini, A. (1999). A fuzzy set-based accuracy assessment of soft classification. *Patter Recognition Letters*, **20**, pp. 935-948.

Blalock, H. M. (1979). Social Statistics (Revised 2nd Edition). McGraw-Hill, Singapore, 625 p.

Blaschke, T., Lang, S., Lorup, E., Strobl J. and Zeil, P. (2000) Object-oriented image processing in an integrated GIS/remote sensing environment and perspectives for environment applications." In: Cremers, A., Greve, K. (eds): Environmental Information for Planning, Politics and the Public, Metropolis-Verlag, Marburg, Volume II, pp. 555-570.

Blok, C.A., B. Kobben, Cheng, T and Kuterema, A.A. (1999). Visualization of relationships between spatial patterns in tie by cartographic animation. *Cartography and Geografic Information Systems* **26(2)**, pp 139-151.

Blok, C.A. (2005). Dynamic visualization variables in animation to support monitoring of spatial phenomena. Utrecht, Enschede, Universiteit Utrecht, ITC, 2005. Nederlandse Geografische Studies = Netherlands Geographical Studies 328, ITC Dissertation 119, 188 p.

Brassel, K., Bucher, F., Stephan, E. and Vckovski, A. (1995). Completeness. In: Elements of Spatial Data Quality, Guptill, S.C. and Morrison, J.L. eds., Elsevier Science Ltd, Exeter, UK, pp. 81-108.

Bruin, S. de (2000), Geographical Information Modelling for Land Resource Survey. Thesis Wageningen University.

Brimicombe, A.J. (2003). GIS environmental modelling and engineering. London: Taylor and Francis, 312 p.

Buja, A., Cook, D, and Swayne D. F. (1996). Interactive high-dimensional data visualization. *Journal of Computational and Graphical Statistics* **5**. pp 78-99.

Burrough, P.A. (1996). Natural Objects with Indeterminate Boundaries. In: Geographic Objects with Indeterminate Boundaries. Eds., Peter A. Burrough, and Andrew U. Frank. Bristol, PA: Taylor & Francis Inc., pp 3-28.

Burrough, P.A. and McDonnell, R.A. (1998). Principles of Geographical Information Systems, London: Oxford University Press, 333 pp.

Byungyong, K. and Landgrebe, D.A. (1991). Hierarchical decision tree classifiers in high-dimensional and large class data. *IEEE Transactions on the Geosciences and Remote Sensing*, **29(4)**, pp. 518-528.

Card, S.K., MacKinley, J.D. & Schneiderman, B. (1999), Readings in Information Visualizations, Using Vision to Think. Morgan Kaufmann Publishers, Inc. San Francisco, California.

Carter, R.W.G. (1988). Coastal environments. An introduction to the physical, ecological and cultural systems of coastlines. Academic Press, London, 617 pp.

CEN/TC-287, (1994/1995). WG 2, Data description: Quality. Working paper N. 15, August 1994. PT05, Draft Quality Model for Geographic Information, Working paper D3, January 1995.

Chambers, J.M., Cleveland, W.S., Kleiner, B., & Tukey, P.A. (1983) Graphical Methods for Data Analysis. Belmont, CA: Wadsworth.

Chandrasekan, B., Josephson, J.R. and Benjamins, V.R. (1999). What are Ontologies, and why do we need them? *IEEE Intelligent Systems*, **14 (1)**, pp. 20-26.

Chandrasekaran, B., Josephson, J.R. and Benjamins, V.R. (1998). Ontology of Tasks and Methods, Proceedings of KAW'98, Eleventh Workshop on Knowledge Acquisition, Modeling and Management, Inn, Banff, Alberta, Canada, April, 1998. Available online at http://ksi.cpsc.ucalgary.ca/KAW/KAW98/chandra/index.html (Accessed 30 June 2003).

Charlier, R.H., Chaineux, M.C.P. and Morcos, S. (2005). Panorama of the history of coastal protection. *Journal of Coastal Research* **21(1)**, pp 79-111.

Cheng, T. (1999). A process-oriented data model for fuzzy spatial objects. Ph.D.Thesis, Enschede/Wageningen, 163 pp.

Cheng, T. (2002). Fuzzy Objects: Their Changes and Uncertainties. *Photogrammetric Engineering and Remote Sensing*, **68**, pp. 41-49.

Cheng T., Molenaar M. and Hui L. (2001). Formalizing fuzzy objects from uncertain classification results. *International Journal of Geographical Information Science,* **15 (1)**, pp. 27-42.

Chernoff, H. (1973) The use of faces to represent points in k-dimensional space graphically. *Journal of the American Statistical Association*, **68**, pp 361-368.

Chrisman, N.R. (1983). The Role of Quality in the Long Term Functioning of a Geographical Information System. International Symposium on Automated Cartography (Auto Carto 6), Ottawa, Ontario, Canada.

Chrisman, N. R. (1984), The Role of Quality Information in the Long-term Functioning of a Geographical Information System. *Cartographica* **21**, pp. 79-87.

Cleveland, W. S. (1993) Visualizing Data. Summit, NJ: Hobart Press.

Cook, D., Symanzik, J., Majure, J.J., and Cressie, N. (1997). Dynamic graphics in a GIS: More examples using linked software. *Computers and geosciences* **23(4)**, pp 371-385.

Cressie, N. A. C. (1991). Statistics for spatial data. New York: Wiley, 900 p.

Crist, H.P. and Cicone R.C. (1984). Application of the Tasseled Cap Concept to Simulated Thematic Mapper Data. *Photogrammetric Engineering and Remote Sensing* **50(3)**, pp 343-352.

Debeljak, M., Dzeroski, S. Jerina, K., Kobler, A. and Adamic, M. (2001). Habitat suitability modelling for red deer (Cervus elaphus L.) in South-central Slovenia with classification trees, *Ecological Modelling* **138(1-3)**, pp. 321 – 330.

De la Rosa, D., Mayol, F., Moreno, J.A., Bonson, T. and S. Lozano (1999). An expert-system neural-network model (Impelero) for evaluating agricultural soil-erosion in Andalusia region, Southern Spain. *Agric. Ecosys. Environ.* **73**, pp 211–226.

Dempster, A.P. (1967). Upper and Lower Probabilities Induced by a Multivalued Mapping. *Ann. Math. Stat.* **38**, pp 325-339.

De Ruig, H.J.M. and Louise, C.J. (1991). Sand budget trends and changes along the Holland coast. *Journal of Coastal Research* **7(4)**, pp. 1013-1026.

Devillers, R., Bedard, Y. and Jeansoulin, R. (2005). Multidimensional management of geospatial data quality information for its dynamic use within GIS. *Photogrammetric Engineering & Remote Sensing* **71(2)**, pp. 205-215.

DiBiase, D. W. (1990). Scientific Visualization in the earth sciences. *Earth and Mineral Sciences* **59(2)**, pp 12-18.

Dunham, S., Fonstad, M.A., Anderson, S.A., Czajkowski, K. (2005). Using multi-temporal satellite imagery to monitor the response of vegetation to drought in the Great Lakes region, *GIScience and Remote Sensing* **42(3)**, pp. 185-201.

Dykes, J.A. (1997). Exploring spatial data representation with dynamic graphics. *Computers and geosciences* **23(4)**. pp 345-370.

Dykes, J.A., MacEachren, A.M. and Kraak, M.J. (2005). Exploring geovisualization. Elsevier, Amsterdam, The Netherlands 710 p.

Edsall, R. M. (1999) The dynamic parallel coordinate plot: visualizing multivariate geographic data, Proceedings 19th International Cartographic Association Conference, Ottawa, pp. 89-97.

Edsall, R. M. (2003). The parallel coordinate plot in action: design and use for geographic visualization. *Computational Statistics and Data Analysis*, **43**, pp 605-619.

Edsall, R. M. and Sidney, L.R (2005) Applications of a cognitively informed framework for the design of interactive spatio-temporal representations. In: Dykes, J., MacEachren, A.M. and Kraak, M.J., (eds.), Exploring Geovisualization. Elsevier Ltd., Amsterdam, pp 577-589.

Edsall, R. M., Harrower M. and Mennis J. (2000) Visualizing properties of spatial and temporal periodicity in geographic data*. Computers and Geosciences* **26(1)**, pp 109-118.

Edsall, R. M., MacEachren, A. M. and Pickle, L. J. (2001). Case study: design and assessment of an enhanced geographic information system for exploration of multivariate health statistics. IEEE Symposium on Information Visualization 2001 (INFOVIS'01), San Diego, USA, pp 159-162.

Edwards, G., and Jeansoulin, R. (2004). Data Fusion - from a logic perspective with a view to implementation, In: *International Journal of Geographical Information Science* **18(4)**, Guest Editorial. pp. 303-307.

Egenhofer, M. (1994). Deriving the composition of binary topological relations. *Journal of Visual Languages and Computing*, **5**, pp 133-149.

Eklund, P. W., Kirkby, S. D. and Salim, A. (1994). A framework for incremental knowledge base update from additional data coverages." Proceedings of the 7th Australasian Remote Sensing Conference, pp. 367-374.

Eklund, P. W., Kirkby, S. D. and Salim, A. (1998). Data Mining and Soil Salinity Analysis. *International Journal of Geographical Information Science*, **12(3)**, pp. 247-268.

Eleveld, M.A. (1996). Remote sensing for process oriented geomorphological mapping of Ameland, a coastal barrier island in the North of the Netherlands. In: Taussik, J., & Mitchell, J. (eds). Partnership in coastal zone management. Samare Publ., Cardigan, pp. 491-498.

Fairbairn, D., Adrienko, G.L. Adrienko, N.V., Buziek, G. and Dykes, J. (2001) Representation and its relationship with cartographic visualization: a research agenda. *Cartography and Geographic Information Systems*, **28(1)**, pp 13-28.

Faulkner, X. (2000). Usability engineering. Palgrave, New York, 244 p.

FGDC (2000). Federal Geographic Data Committee, FGDC-STD-001-1998. Content Standard for Digital Geospatial Metadata Workbook version 2.0. Federal Geographic Metadata Committee, Washington DC.

Fisher P.F. (1999). Models of Uncertainty in Spatial Data. In: Geographic Information Systems: Principles and Technical Issues. Goodchild M. F., Longley P. A., Maguire D. J. and Rhind D. W. (Eds.), Second ed., New York: John Wiley & Sons, Inc.,1999, pp. 191-203.

Fisher, P. F., Cheng, T. and Wood, J. (2004). Where is Helvellyn? Multiscale morphometry and the mountains of the English Lake District, *Transactions of the Institute of British Geographers* **29**, pp. 106-128.

Fisher, P. F., Wood, J., and Cheng, T. (2005). Fuzziness and ambiguity in multi-scale analysis of landscape morphometry. In: Fuzzy modeling with spatial information for geographic problems. Eds. F.E. Petry., V.B. Robinson, M.A. Cobb. Springer-Verlag, Berlijn, ch.10, pp. 209-232.

Foody, G. M. and Atkinson, P. M., (2002). Current status of uncertainty issues in remote sensing and GIS. In: Foody G.M. and Atkinson P.M. (eds) Uncertainty in Remote Sensing and GIS. John Wiley & Sons Ltd.

Frank A.U. (1996): Qualitative Spatial Reasoning: Cardinal Directions as an Example. *International Journal of Geographical Information Science* **10(3)**, pp. 269-290.

Frank, A.U. (1997). Spatial Ontology: A Geographical Point of View. In Spatial and Temporal Reasoning. (Stock, O., ed.), Dordrecht, The Netherlands, Kluwer Academic Publishers, pp: 135-153.

Frank, A.U. (2003a). A linguistically justified proposal for a spatio-temporal ontology. COSIT 2003 conference. Url: www.comp.leeds.ac.uk/brandon/cosit03ontology/position_papers/Frank.doc. (Accessed 28 June 2003).

Frank, A.U. (2003b). Ontology for spatio-temporal databases. In Spatio-Temporal Databases: The Chorochronos Approach. (al., T.S.e., ed.), Lecture Notes in Computer Science, Berlin, Springer-Verlag, pp 10-79.

Frank, A.U. and Mark, D.M. (1991). Language Issues for Geographical Information Systems. In Geographic Information Systems: Principles and Applications, ed. Maguire, D. Rhind, D., and Goodchild M. London: Longman Co., pp 129-150.

Frank, A.U., & Grum, E. (2004). Proceedings of the ISSDQ '04 Volume 1 and 2. (Frank, A.U., & Grum, E., eds.), Vol. 28a and b, GeoInfo Yellow Series, Vienna, Department of Geoinformation and Cartography.

Freska, C. (1992). Using Orientation Information for Qualitative Spatial Reasoning. In Frank, A. U., Campari, I., and Formentini, U., editors, Theories and Methods of Spatial-Temporal Reasoning in Geographic Space, Springer, Berlin, pp 162-178.

Gahegan, M. (1999) Four barriers to the development of effective exploratory visualisation tools for the geosciences. *International Journal of Geographical Information Science* **13(4)**: pp 289-309.

Gardner, R.H., and Urban D.L. (2003). Model validation and testing: Past lessons, present concerns, future prospects. In: Models in Ecosystem Science, eds Canham C.D., Cole J.C. and Lauenroth W.K. Princeton University Press, Princeton, NJ. 456 p.

Gelman, A., Carlin, J. B. and Stern, H. S. (1995). Bayesian Data Analysis. Chapman and Hall, London, 526 p.

Goodchild, M.F. (1995). Attribute Accuracy. In: Elements of Spatial Data Quality, Guptill, S.C. and Morrison, J.L. eds., Elsevier Science Ltd, Exeter, UK, pp.81-108.

Goodchild. M. (1992). Geographical data modeling. *Computers and Geosciences*, **18(4)**, pp. 401-408.

Gorte, B. (1998). Probabilistic Segmentation of Remotely Sensed Images. PhD-dissertation, In: ITC Publication Series No 63, 1998, 143 p.

Grossmann, R. (1983). The categorial structure of the world*,* Bloomington, Indiana University Press, pp. 3-5.

Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition,* **5(2)**, pp. 199-220.

Gruninger, M. and Fox, M.S. (1995). Methodology for the Design and Evaluation of Ontologies, Workshop on Basic Ontological Issues in Knowledge Sharing, IJCAI-95, Montreal.

Guarino, N. (1998). Formal ontology and information systems. In N. Guarino (ed). Formal Ontology and Information Systems, Proceedings of FOIS'98, Trento, Italy, 6-8 June 1998, pp 3-15. IOS Press, Amsterdam.

Guarino, N. and Giaretta, P. (1995). Ontologies and Knowledge Bases: Towards a Terminological Clarification. In N. Mars (ed.) Towards Very Large Knowledge Bases. IOS Press, Amsterdam: 25-32.

Guo D., Gahegan, M., MacEachren, A.M. and Zhou, B. (2005) Multivariate analysis and geovisualization with an integrated geographic knowledge

discovery approach. *Cartography and Geographic Information Science*, **32(2)**, pp. 113-132.

Guptill S.C. and Morrison, J.L. (1995). Elements of Spatial Data Quality, Elsevier Science Ltd, Exeter, UK 202 pp.

Guptill, S.C. (1998), Building a Geospatial Data Framework - Finding the "Best Available" Data. In M. Goodchilds & R. Jeansoulin (eds.): Data Quality in Geographic Information. From Error to Uncertainty. Hermes, Paris. pp. 31-36.

Hägerstrand, T. (1970) What about people in regional science? *Papers of the Regional Science Association*, **24** pp. 7-21.

Harrower, M., MacEachren, A.M. and Griffin, A. (2000). Developing a geographic visualization tool to support Earth science learning. *Cartography and Geographic Information Science*, **27**, pp 279-294.

Hay, G.J., Blaschke T., Marceau D. J. and Bouchard, A. (2003). A comparison of three image-object methods for the multiscale analysis of landscape structure. *ISPRS Journal of Photogrammetry & Remote Sensing*, **57**, pp 327-345.

Hernandez, D., Clementini, E.,and Di Felice, P. (1995) Qualitative Distances. In A. Frank and W. Kuhn (eds.), Spatial Information Theory-A Theoretical Basis for GIS. International Conference COSIT '95 Semmering, Austria, *Lecture Notes in Computer Science* **988**, Berlin: Springer-Verlag, pp. 45-58.

Heuvelink, G.B.M. (1998). Error propagation in environmental modelling with GIS. London: Taylor & Francis, 127 p.

Hunsaker, C.T., Goodchild, M.F., Friedl, M.A. and Case T.J. (eds.) (2001). Spatial uncertainty in ecology: implications for remote sensing and GIS applications. New York, Springer-Verlag, 402 p.

Hunter, G. J. (2001). Spatial Data Quality Revisited. Proceedings of GeoInfo 2001, Rio de Janeiro, Brazil, 4-5th October, pp. 1-7.

Inselberg, A. (1985) The plane with parallel coordinates. *The Visual Computer* **1(2)**: pp 69-91.

ISO (1993). ISO CD 9241-11: Guidelines for specifying and measuring usability.

ISO, (2000). ISO/DIS 19115 Geographic information.

ISO (2003). ISO/TC 211 Geographic information/Geomatics - 19113, 1914, Geographic information quality, and evaluation procedures, as sent to the ISO Central Secretariat for publication. 68 pp.

Jacobi, O. (1994), Data Quality in GIS. *Proceedings Eurocarto 12*, contribution XIV, Copenhagen.

Jakeman, A.J., Beck, M.B. and McAleer, M.J. (1995). Preface. In: Jakeman, A.J., Beck, M.B. & McAleer, M.J. (eds). Modelling Change in Environmental Systems. Wiley, Chichester, pp.xvii-xxi.

Janikow C.Z. (1998). Fuzzy Decision Trees: Issues and Methods. *IEEE Transactions on Systems, Man, and Cybernetics*, **28(1)**, pp. 1-14.

Jeansoulin, R. and Wilson, N. (2002). Quality of Geographic Information: Ontological approach and Artificial Intelligence Tools in the REV!GIS project, 8th EC-GI&GIS Workshop, Dublin 3-5 July, 2002.

Johnson, B. and Shneiderman, B. (1991) Tree-maps: a space-filling approach to the visualization of hierarchical information structures. Proceedings IEEE Visualization '91, Piscataway, NJ: IEEE, pp. 284-291.

Juran, J. M., Gryna, F. M. J., and Bingham, R. S. (1974). Quality Control Handbook, McGraw-Hill, New York.

Kahn, B. K. and Strong, D. M. (1998). Product and Service Performance Model for Information Quality: An Update. Proceedings of the 1998 Conference on Information Quality, pp 102-115.

Keim, D.A. (2000). Designing pixel-oriented visualization techniques: theory and applications. *IEEE Transactions and Visualization and Computer Graphics*, **6(1)**, pp 59-78.

Keim, D.A., Panse, C. and Sips, M., (2005). Information Visualization: Scope, Techniques and Opportunities for Geovisualization. In: Dykes, J., MacEachren, A.M., and Kraak, M.J., (eds.), Exploring geovisualization. Elsevier Ltd., Amsterdam.

Kessler, F.C. (2000). Focus group as means of qualitatively assessing the U-boat narrative. *Cartographica*, **37**, pp 33-60.

Kokla, M and Kavouras, M. (2001). Fusion of Top-level and Geographic Domain Ontologies Based on Context Formation and Complementarily, *International Journal of Geographical Information Science*, **15(7)**, pp. 679-687.

Kraak, M. J. (2000). Visualization of time dimension. Time in GIS: Issues in spatio-temporal modelling. L. Heres, *Netherlands Geodetic Commission: Publication on Geodesy*, **47**, pp. 27-35.

Kraak, M.J. (2003) Geovisualization illustrated. *ISPRS Journal of Photogrammetry and Remote Sensing*, **57(1)** pp. 1-10.

Kraak, M.J. and MacEachren, A.M. (1999) Visualization for exploration of spatial data. In: *International Journal of Geographical Information Science*, **13(4)**, pp. 285-287.

Kraak, M.J. and Ormeling, F.J. (2003) Cartography: visualization of geospatial data. Harlow, Addison Wesley, Longman, 2003. 205 p.

Kraak, M.J. and Klomp, A. (1995) A classification of cartographic animations: towards a tool for the design of dynamic maps in a GIS environment. In: Proceedings of the seminar on teaching animated cartography : Madrid, Spain August 30 - September 1, 1995. pp. 29-35.

Kraak, M.J., Edsall, R.M. and MacEachren, A.M. (1997) Cartographic animation and legends for temporal maps : exploration and or interaction. In: Proceedings of the 18th ICA conference ICC 1997, Sweden, 23-27 June, Vol. I. 8 p.

Kuhn, W. (2001) Ontologies in Support of Activities in Geographical Space. *International Journal of Geographical Information Science*, **15(7)**, pp. 613-631.

LeBlanc, J., Ward, M. O. and Wittels, N. (1990) Exploring N-dimensional databases, Proceedings Visualization '90, San Francisco, CA, pp.230-239.

Lein, J.K. (2003). Applying evidential reasoning methods to agricultural land cover classification. *International Journal of Remote Sensing* **24(21)**, pp 4161-4180.

Lees, B. G. and Ritman, K. (1991). Decision tree and rule induction approach to integration of remotely sensed and GIS data in mapping vegetation in disturbed or hilly environments. *Environmental Management*, **15(6)**, pp. 823-831.

Li, P., Haskell, R. E. and Hanna, D.M (2003). Optimizing Fuzzy Decision Tree Using Genetic Algorithms. Proceedings of the International Conference on Artificial Intelligence, June 2003, Las Vegas, USA, pp. 469-474.

Li, X. and Kraak, M.J. (2005). New views on multivariable spatio-temporal data: the space time cube expanded. In Proceedings of International Symposium on Spatio-temporal Modeling, Spatial Reasoning, Spatial Analysis, Data Mining and Data Fusion, 27-29 August, 2005, Beijing, China, pp. 199-202.

Lucieer, A. and Kraak, M.J. (2004). Interactive and visual fuzzy classification of remotely sensed imagery for exploration of uncertainty. *International Journal of Geographical Information Science*, **18(5)**, 2004, pp. 491-512.

MacEachren, A.M. (1995). How maps work: representation, visualization and design. New York, The Guildford Press. 513 p.

MacEachren, A.M., Wachowicz, M., Edsall, R., Haug, D., Masters, R., (1999). Constructing a knowledge from multivariate spatio-temporal data: integrating geovisualization (GVis) with knowledge discovery in databases (KDD). *International Journal of Geographic Information Science*, **13(4)**, pp 311- 334.

Marceau, D.J., Guindon L., Bruel M., and Marois C.,(2001). Building temporal topology in a GIS database to study the land-use changes in a rural-urban environment. *The Professional Geographer*, **53(4)**, pp. 546-558.

Marchand, P., Brisebois, A., Bedard, Y. and Edwards, G. (2004). Implementation and evaluation of a hypercube-based method for spatiotemporal exploration and analysis. *ISPRS Journal of Photogrammetry and Remote Sensing*, **59**, pp. 6-20.

Matheron, G. (1971). The theory of regionalised variables and its applications, Cahier No. 5, Centre de Morphologie Mathématique de Fontainebleau, 211 p.

McGranaghan, M. (1993). A Cartographic View of Spatial Data Quality, *Cartographica***, 30(2 & 3)**, pp. 8-19.

Mennis, J.L., Peuquet, D.J. and Qian, L. (2000) A Conceptual Framework for Incorporating Cognitive Principles into Geographical Database Representation. *International Journal of Geographic Information Science,* **14(6)**, pp. 501-520.

Molenaar, M. (1998). An Introduction to the Theory of Spatial Object Modelling for GIS. Taylor & Francis, London, 246 pp.

Moore, A., Jones, A., Sims, P. and G. Blackwell (2001). Intelligent Metadata Extraction for Integrated Coastal Zone Management. Proceedings of GeoComputation, University of Queensland, Brisbane, Australia, CD-ROM: ISBN 1864995637.

Moore, A., Jones, A., Sims, P. and G. Blackwell (2001). Integrated Coastal Zone Management's Holistic Agency: An Ontology of Geography and GeoComputation. In: Proceedings of Thirteenth Annual Colloquium of the Spatial Information Research Centre. P.A. Whigham (ed). 2 - 5 Dec 2001, Dunedin, New Zealand. University of Otago, pp.211-222.

Morgan, D.L. (1998). The Focus Group Guidebook – Focus Group Kit 1 (Newbury Park, Ca: Sage).

Murwira, A. (2003). Scale matters! A new approach to quantify spatial heterogeneity for predicting the distribution of wildlife. PhD-dissertation Wageningen University, In: ITC Publication Series No 106, 195 p.

Nielsen, J. (1993). Usability Engineering, San Diego, Academic Press. 358 p.

Noy, N. and McGuinness, D.L. (2001). Ontology Development 101: A Guide to Creating Your First Ontology. Stanford Medical Informatics Technical Report #SMI-2001-0880. Available online at: http://www.ksl.stanford.edu/people/dlm/papers/ontology-tutorial-noy-mcguinness.pdf (accessed 30 June 2003).

O'Hagan, A. (1999), 1stBayes Software. Available online at http://www.shef.ac.uk/~st1ao/1b.html (accessed 12 July 2005).

Ogao, P.J. (2002). Exploratory visualization of temporal geospatial data using animation. ITC Dissertation No.89, The Netherlands.

Ogao, P.J. and Kraak, M.J. (2002) Defining visualization operations for temporal cartographic animation design. In*: International Journal of Applied Earth Observation and Geoinformation*, **1**, pp. 23-31.

Olaru, C. and Wehenkel, L. (2003) A complete fuzzy decision tree technique. *Fuzzy Sets and Systems* **138(2)**, pp. 221-254.

Ormeling, F.J. (1995) Teaching animation cartography. In: Proceedings of the Seminar on Teaching Animated Cartography. Madrid. Spain.

Papadias D. and Sellis D. (1994) Qualitative Representation of Spatial Knowledge in Two Dimensional Space, *VLDB Journal, Special Issue on Spatial Databases*, **3(4)**, pp. 479-516.

Pearl, J. (1988). Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann Publishers, San Mateo, CA, U.S.A.

Pebesma, E.J., Duin, R.N.M. and Burrough P.A. (2005). Mapping sea bird densities over the North Sea: spatially aggregated estimates and temporal changes. *Environmentrics*, **16(6)**, pp. 573-587.

Peng, Y. and Flach, P. (2001). Soft Discretization to Enhance the Continuous Decision Tree Induction. In: Integrating Aspects of Data Mining, Decision Support and Meta-Learning, Christophe Giraud-Carrier, Nada Lavrac and Steve Moyle, editors, ECML/PKDD'01 workshop notes, September 2001, pp 109-118.

Peuquet, D.J. (1994). It's about time: A conceptual framework for the representation of spatio-temporal dynamics in geographic information systems. *Annals of the Association of American Geographers* **84**, pp. 441-461.

Peuquet, D.J. (2002). Representations of Space and Time, The Guilford Press, New York. 380 p.

Poli, R. (1996). Ontology for knowledge organization, In R. Green (ed.), Knowledge organization and change, Indeks, Frankfurt, pp. 313-319.

Quinlan, J. R. (1988). Decision Trees and Multi-valued Attributes. *Machine Intelligence*, **11**, pp 305-318.

Quinlan, J.R. (1986). Induction of decision trees, *Machine Learning*, **1(1)**, pp. 81-106.

REV!GIS (2000). Available online at http://www.lsis.org/REVIGIS/.

Robinson, A. C., Chen J., Lengerich G., Meyer H., and MacEachren, A.M. (2005). Combining usability techniques to design geovisualization tools for epidemiology. Proceedings of the Auto-Carto Conference, Las Vegas, NV, March 18-23.

Richards J.A. and Jia, X. (1999). Remote Sensing Digital Image Analysis. Third edition. Springer-Verlag Berlin, Germany, 363 p.

Ricotta C. and Avena G.C. (1999). The influence of fuzzy set theory on the areal extent of thematic map classes. *International Journal of Remote Sensing*, **20(1)**, pp. 201-205.

Roelse, P. (2002). Water en Zand in Balans. Evaluatie zandsuppleties na 1990; een morfologische beschouwing. Internal Report Rijksinsituut voor Kust en Zee, 2002.003, Middelburg, The Netherlands, 108 pp.

Rojas, R. (1991). Neural networks; a systematic approach. Springer-Verlag, Berlin, 501p.

Rose, K.A. and Swartzman, G.L. (1981). A review of parameter sensitivity methods applicable to ecosystem models. NUREG/CR-2016. Washington DC: US Buclear Regulatory Commission.

Rouse, J. W., Jr., Haas, R. H., Deering, D. W., Schell, J. A. and Harlan, J. C. (1974). Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation, NASA/GSFC type III final report: Greenbelt, Maryland, NASA, 371p.

Roy, A.J.O. (2001). A Survey of Knowledge Representation Formalisms Involving Uncertainty, Internal Technical Report, REV!GIS project, IST-1999-14189, 40 pp.

Shafer, G. (1976). A Mathematical Theory of Evidence. Princeton, NJ: Princeton University Press.

Shafer, G. (1990) Perspectives on the Theory and Practice of Belief functions. *International Journal of Approximate Reasoning* **4**, pp. 323-362.

Shekhar, S., Zhang, P., Huang, Y., Vatsavai, R. (2004): Trends in Spatial Data Mining In: Kargupta, H., et al. (eds.): Data Mining: Next Generation Challenges and Future Directions. AAAI Press. pp. 357 – 380.

Shi, W., Goodchild M.F., and Fisher P.F. (2003). Proceedings of the Second International Symposium on Spatial Data Quality. Hong Kong: Hong Kong Polytechnic University.

Shneiderman, B. (1996) The eyes have it: A task by data type taxonomy for information visualizations, In Proceedings IEEE Visual Languages, Boulder, CO, Sept 1996, pp 336-343.

Shneiderman, B. (1998). Designing the User Interface. Strategies for Effective Human-Computer Interaction. Addison Wesley Longman, Inc., third edition, 637 p.

Smith, B. (2001). Objects and Their Environments: From Aristotle to Ecological Psychology. In Andrew Frank (ed.) The Life and Motion of Socioeconomic Units (GISDATA 8), London: Taylor and Francis, pp 79-97.

Smith B. and Mark. D. (1999). Ontology with human subjects testing: An empirical investigation of geographic categories. *American Journal of Economics and Sociology,* **58(2)**, pp 245-272.

Spence, R. (2001). Information Visualization, Addison Wesley / ACM Press Books, Harlow.

StatLine (2005). http://statline.cbs.nl (accessed 10 October 2005).

Storvik, G., Fjørtoft R. and Schistad Solberg, A.H. (2005). A Bayesian Approach to Classification of Multiresolution Remote Sensing Data. *IEEE Transactions on Geoscience and Remote Sensing*, **43(3)**, pp 539-547.

Suárez, A. and Lutsko, J.F. (1999) Globally optimal fuzzy decision trees for classification and regression, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **21(12)**, pp. 1297-1311.

Suchan, T. A. and C. A. Brewer (2000). Qualitative methods for research on mapmaking and map use. *Professional Geographer* **52(1)**: 145-154.

Teoh S. T. and Ma, K. (2003). PaintingClass: Interactive Construction, Visualization and Exploration of Decision Trees." In Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-03), New York, August 24–27 2003. ACM Press, pp 667-672.

TerraSeer (2004). The TerraSeer Space-Time Intelligence System. Available online at: http://www.terraseer.com (accessed 12 July 2005).

Tso B.K.C. and Mather P.M. (2001). Classification Methods for Remotely Sensed Data. London: Taylor & Francis Ltd, 2001, 272 p.

Tugendthat, E. (1982). Traditional and analytical philosophy. Lecture on the philosophy of language - Translated by P. A. Gorner - Cambridge, Cambridge University Press, pp. 21.

UDIG (2004). User-friendly Desktop Internet GIS. Available online at: http://udig.refractions.net (accessed 8 August 2005).

Van de Vlag, D.E. (2004). Concepts and Representation of Beach Nourishments by Spatio-temporal Ontologies. In: The Netherlands and the North Sea. Dutch Geography 2000-2004, Nederlandse Geografische Studies 325, T. Dietz, P Hoekstra, F. Thissen (eds.), Utrecht, The Netherlands, pp 45-55.

Van de Vlag, D.E. and Kraak, M.J. (2004) Multivariate Visualization of Data Quality Elements for Coastal Zone Monitoring, ISPRS Conference Proceedings Commission IV, ISPRS 15-23 July 2004 Istanbul, Turkey,7p.

Van de Vlag, D.E. and Kraak, M.J. (2005) Temporal Ordered Space Matrix: representation of multivariate spatiotemporal data. In GISPLANET'05 Conference Proceedings, S01 Spatial Knowledge, 13p.

Van de Vlag, D.E., Vasseur, B., Stein, A., Jeansoulin, R. (2005), An Application of Problem and Product Ontologies for the Revision of Beach Nourishments, *International Journal of Geographical Information Science* **19 (10)**, pp. 1057-1072.

Van der Wal, D. (2000). Modelling aeolian sand transport and morphological development in two beach nourishment areas. *Earth Surface Processes and Landforms* **25**, pp 77-92.

Van Noortwijk, J.M. and Peerbolte, E.B. (2000). Optimal sand nourishment decisions. *Journal of waterway, port, coastal, and ocean engineering*, **126(1)**, pp 30-38.

Van Oort, P.A.J. (2006). Spatial data quality: from description to application. PhD Thesis, Wageningen University, The Netherlands. 124 p.

Van Rijn, L.C. (1997). Sediment transport and budget of the central coastal zone of Holland. *Coastal Engineering* **32(1)**, pp. 61-90.

Van Vuren, S., Kok, M. and Jorissen, R.E. (2004). Coastal defence and societal activities in the coastal zone: compatible or conflicting interests?, *Journal of Coastal Research*, **20(2)**, pp 550-561.

Vasseur, B., Devillers, R. and Jeansoulin, R. (2003). Ontological approach of the fitness of use of geospatial dataset. In: Proceedings of 6th AGILE Conference on Geographic Information Science, (Gould, M., Robert, L., & Stéphane, C., eds.), Published by Presses Polytechniques et Universitaires Romandes, Lyon, France, pp. 497-504.

Vasseur, B., Vlag, D.E. van de, Stein, A., Jeansoulin, R., Dilo, A. (in review). Quality-aware ontologies for dynamic spatial objects with indeterminate boundaries. *Applied Ontology*.

Veregin, H. (1999). Data quality parameters. In: Geographical Information Systems, P. A. Longley, M. F. Goodchild, D. J. Maguire, and D. W. Rhind, eds., John Wiley & Sons, Inc., pp. 177-189.

Wang, R. Y. and Strong, D. M. (1996). Beyond Accuracy: What Data Quality Means to Data Consumers. *Journal of Management Information Systems*, **12(4)**, pp. 5-34.

Wilkie, D. S. and Finn, J. T. (1996). Remote Sensing Imagery for Natural Resources Monitoring. Colombia University Press, New York.

Wilson, N. (2002). A Survey of Numerical Uncertainty Formalisms with Reference to GIS Applications, Internal Technical Report, REV!GIS project, IST-1999-14189, 26 pp.

Worboys, M.F. (2005) Event-oriented approaches to geographic phenomena. *International Journal of Geographical Information Science* **19(1)**: 1-28.

Xian, G., Zhu, Z., Hoppus, M. and Fleming M. (2002). Application of decision-tree techniques to forest group and basal area mapping using satellite imagery and forest inventory data. Pecora 15/Land Satellite Information IV/ISPRS Commission I/FIEOS 2002 Conference Proceedings, 10-15 November 2002, Denver, CO USA.

Xu, M., Watanachaturaporn, P., Varshney, P.K. and Arora, M.K.(2005). Decision tree regression for soft classification of remote sensing data. *Remote Sensing of Environment*, **97**, pp. 322-336.

Yemefack, M. (2005). Modelling and monitoring soil and land use dynamics within shifting agricultural landscape mosaic systems. Enschede, ITC, 2005. ITC Dissertation 121, 194 p. ISBN: 90-6164-233-7.

Zadeh, L.A. (1965). Fuzzy Sets. *Information and Control* **8**, pp. 338-353.

Zhang, J. and Foody, G.M. (2001). Fully-fuzzy supervised classification of sub-urban land cover from remotely sensed imagery: Statistical and artificial neural network approaches, *International Journal of Remote Sensing* **22(4)**, pp 615-628.

ITC Dissertation List

1. Akinyede, J.O. (1990), Highway cost modelling and route selection using a geotechnical information system
2. Pan, P.H. (1990), 90-9003-757-8, Spatial structure theory in machine vision and applications to structural and textural analysis of remotely sensed images
3. Bocco Verdinelli, G.H.R. (1990), Gully erosion analysis using remote sensing and geographic information systems: a case study in Central Mexico
4. Sharif, M. (1991), Composite sampling optimization for DTM in the context of GIS
5. Drummond, J.E. (1991), Determining and processing quality parameters in geographic information systems
6. Groten, S. (1991), Satellite monitoring of agro-ecosystems in the Sahel
7. Sharifi, A. (1991), 90-6164-074-1, Development of an appropriate resource information system to support agricultural management at farm enterprise level
8. Zee, D. van der (1991), 90-6164-075-X, Recreation studied from above: Air photo interpretation as input into land evaluation for recreation
9. Mannaerts, C. (1991), 90-6164-085-7, Assessment of the transferability of laboratory rainfall-runoff and rainfall - soil loss relationships to field and catchment scales: a study in the Cape Verde Islands
10. Wang, Z.S. (1991), 90-393-0333-9, An expert system for cartographic symbol design
11. Yunxian, Z. (1991), 90-6164-081-4, Application of Radon transforms to the processing of airborne geophysical data
12. Zuviria, M. de (1992), 90-6164-077-6, Mapping agro-topoclimates by integrating topographic, meteorological and land ecological data in a geographic information system: a case study of the Lom Sak area, North Central Thailand
13. Westen, C. van (1993), 90-6164-078-4, Application of Geographic Information Systems to landslide hazard zonation
14. Shi, W. (1994), 90-6164-099-7, Modelling positional and thematic uncertainties in integration of remote sensing and geographic information systems
15. Javelosa, R. (1994), 90-6164-086-5, Active Quaternary environments in the Philippine mobile belt
16. Lo, K.C. (1994), 90-9006526-1, High Quality Automatic DEM, Digital Elevation Model Generation from Multiple Imagery
17. Wokabi, S.M. (1994), 90-6164-102-0, Quantified land evaluation for maize yield gap analysis at three sites on the eastern slope of Mt. Kenya
18. Rodriguez, O.S. (1995), Land Use conflicts and planning strategies in urban fringes: a case study of Western Caracas, Venezuela
19. Meer, F.D. van der (1995), 90-5485-385-9, Imaging spectrometry & the Ronda peridotites

20. Kufoniyi, O. (1995), 90-6164-105-5, Spatial coincidence: automated database updating and data consistency in vector GIS
21. Zambezi, P. (1995), Geochemistry of the Nkombwa Hill carbonatite complex of Isoka District, north-east Zambia, with special emphasis on economic minerals
22. Woldai, T. (1995), The application of remote sensing to the study of the geology and structure of the Carboniferous in the Calañas area, pyrite belt, SW Spain
23. Verweij, P.A. (1995), 90-6164-109-8, Spatial and temporal modelling of vegetation patterns: burning and grazing in the Paramo of Los Nevados National Park, Colombia
24. Pohl, C. (1996), 90-6164-121-7, Geometric Aspects of Multisensor Image Fusion for Topographic Map Updating in the Humid Tropics
25. Bin, J. (1996), 90-6266-128-9, Fuzzy overlay analysis and visualization in GIS
26. Metternicht, G.I. (1996), 90-6164-118-7, Detecting and monitoring land degradation features and processes in the Cochabamba Valleys, Bolivia. A synergistic approach
27. Chu. T.H. (1996), 90-6164-120-9, Development of a Computerized Aid to Integrated Land Use Planning (CAILUP) at regional level in irrigated areas: a case study for the Quan Lo Phung Hiep region in the Mekong Delta, Vietnam
28. Roshannejad, A. (1996), 90-9009284-6, The management of spatio-temporal data in a national geographic information system
29. Terlien, M.T..J. (1996), 90-6164-115-2, Modelling Spatial and Temporal Variations in Rainfall-Triggered Landslides: the integration of hydrologic models, slope stability models and GIS for the hazard zonation of rainfall-triggered landslides with examples from Manizales, Colombia
30. Mahavir, J. (1996), 90-6164-117-9, Modelling settlement patterns for metropolitan regions: inputs from remote sensing
31. Al-Amir, S. (1996), 90-6164-116-0, Modern spatial planning practice as supported by the multi-applicable tools of remote sensing and GIS: the Syrian case
32. Pilouk, M. (1996), 90-6164-122-5, Integrated modelling for 3D GIS
33. Duan, Z. (1996), 90-6164-123-3, Optimization modelling of a river-aquifer system with technical interventions: a case study for the Huangshui river and the coastal aquifer, Shandong, China
34. Man, W.H. de (1996), 90-9009-775-9, Surveys: informatie als norm: een verkenning van de institutionalisering van dorp - surveys in Thailand en op de Filippijnen
35. Vekerdy, Z. (1996), 90-6164-119-5, GIS-based hydrological modelling of alluvial regions: using the example of the Kisaföld, Hungary
36. Gomes Pereira, L.M. (1996), 90-407-1385-5, A Robust and Adaptive Matching Procedure for Automatic Modelling of Terrain Relief
37. Fandino Lozano, M.T. (1996), 90-6164-129-2, A Framework of Ecological Evaluation oriented at the Establishment and Management of Protected Areas: a case study of the Santuario de Iguaque, Colombia
38. Toxopeus, B. (1996), 90-6164-126-8, ISM: an Interactive Spatial and temporal Modelling system as a tool in ecosystem management: with

two case studies: Cibodas biosphere reserve, West Java Indonesia: Amboseli biosphere reserve, Kajiado district, Central Southern Kenya

39. Wang, Y. (1997), 90-6164-131-4, Satellite SAR imagery for topographic mapping of tidal flat areas in the Dutch Wadden Sea
40. Saldana-Lopez, A. (1997), 90-6164-133-0, Complexity of soils and Soilscape patterns on the southern slopes of the Ayllon Range, central Spain: a GIS assisted modelling approach
41. Ceccarelli, T. (1997), 90-6164-135-7, Towards a planning support system for communal areas in the Zambezi valley, Zimbabwe; a multi-criteria evaluation linking farm household analysis, land evaluation and geographic information systems
42. Peng, W. (1997), 90-6164-134-9, Automated generalization in GIS
43. Lawas, M.C. (1997), 90-6164-137-3, The Resource Users' Knowledge, the neglected input in Land resource management: the case of the Kankanaey farmers in Benguet, Philippines
44. Bijker, W. (1997), 90-6164-139-X, Radar for rain forest: A monitoring system for land cover Change in the Colombian Amazon
45. Farshad, A. (1997), 90-6164-142-X, Analysis of integrated land and water management practices within different agricultural systems under semi-arid conditions of Iran and evaluation of their sustainability
46. Orlic, B. (1997), 90-6164-140-3, Predicting subsurface conditions for geotechnical modelling
47. Bishr, Y. (1997), 90-6164-141-1, Semantic Aspects of Interoperable GIS
48. Zhang, X. (1998), 90-6164-144-6, Coal fires in Northwest China: detection, monitoring and prediction using remote sensing data
49. Gens, R. (1998), 90-6164-155-1, Quality assessment of SAR interferometric data
50. Turkstra, J. (1998), 90-6164-147-0, Urban development and geographical information: spatial and temporal patterns of urban development and land values using integrated geo-data, Villaviciencia, Colombia
51. Cassells, C.J.S. (1998), 90-6164-234-5, Thermal modelling of underground coal fires in northern China
52. Naseri, M.Y. (1998), 90-6164-195-0, Characterization of Salt-affected Soils for Modelling Sustainable Land Management in Semi-arid Environment: a case study in the Gorgan Region, Northeast, Iran
53. Gorte B.G.H. (1998), 90-6164-157-8, Probabilistic Segmentation of Remotely Sensed Images
54. Tegaye, T.A. (1998), 90-6164-158-6, The hydrological system of the lake district basin, central main Ethiopian rift
55. Wang, D. (1998), 90-6864-551-7, Conjoint approaches to developing activity-based models
56. Bastidas de Calderon, M. (1998), 90-6164-193-4, Environmental fragility and vulnerability of Amazonian landscapes and ecosystems in the middle Orinoco river basin, Venezuela
57. Moameni, A. (1999), Soil quality changes under long-term wheat cultivation in the Marvdasht plain, South-Central Iran
58. Groenigen, J.W. van (1999), 90-6164-156-X, Constrained optimisation of spatial sampling: a geostatistical approach

59. Cheng, T. (1999), 90-6164-164-0, A process-oriented data model for fuzzy spatial objects
60. Wolski, P. (1999), 90-6164-165-9, Application of reservoir modelling to hydrotopes identified by remote sensing
61. Acharya, B. (1999), 90-6164-168-3, Forest biodiversity assessment: A spatial analysis of tree species diversity in Nepal
62. Abkar, A.A. (1999), 90-6164-169-1, Likelihood-based segmentation and classification of remotely sensed images
63. Yanuariadi, T. (1999), 90-5808-082-X, Sustainable Land Allocation: GIS-based decision support for industrial forest plantation development in Indonesia
64. Abu Bakr, M. (1999), 90-6164-170-5, An Integrated Agro-Economic and Agro-Ecological Framework for Land Use Planning and Policy Analysis
65. Eleveld, M.A. (1999), 90-6461-166-7, Exploring coastal morphodynamics of Ameland (The Netherlands) with remote sensing monitoring techniques and dynamic modelling in GIS
66. Hong, Y. (1999), 90-6164-172-1, Imaging Spectrometry for Hydrocarbon Microseepage
67. Mainam, F. (1999), 90-6164-179-9, Modelling soil erodibility in the semiarid zone of Cameroon
68. Bakr, M.I. (2000), 90-6164-176-4, A Stochastic Inverse-Management Approach to Groundwater Quality
69. Zlatanova, Z. (2000), 90-6164-178-0, 3D GIS for Urban Development
70. Ottichilo, W.K. (2000), 90-5808-197-4, Wildlife Dynamics: An Analysis of Change in the Masai Mara Ecosystem
71. Kaymakci, N. (2000), 90-6164-181-0, Tectono-stratigraphical Evolution of the Cankori Basin (Central Anatolia, Turkey)
72. Gonzalez, R. (2000), 90-5808-246-6, Platforms and Terraces: Bridging participation and GIS in joint-learning for watershed management with the Ifugaos of the Philippines
73. Schetselaar, E. (2000), 90-6164-180-2, Integrated analyses of granite-gneiss terrain from field and multisource remotely sensed data. A case study from the Canadian Shield
74. Mesgari, M.S. (2000), 90-3651-511-4, Topological Cell-Tuple Structure for Three-Dimensional Spatial Data
75. Bie, C.A.J.M. de (2000), 90-5808-253-9, Comparative Performance Analysis of Agro-Ecosystems
76. Khaemba, W.M. (2000), 90-5808-280-6, Spatial Statistics for Natural Resource Management
77. Shrestha, D. (2000), 90-6164-189-6, Aspects of erosion and sedimentation in the Nepalese Himalaya: highland-lowland relations
78. Asadi Haroni, H. (2000), 90-6164-185-3, The Zarshuran Gold Deposit Model Applied in a Mineral Exploration GIS in Iran
79. Raza, A. (2001), 90-3651-540-8, Object-oriented Temporal GIS for Urban Applications
80. Farah, H. (2001), 90-5808-331-4, Estimation of regional evaporation under different weather conditions from satellite and meteorological data. A case study in the Naivasha Basin, Kenya

81. Zheng, D. (2001), 90-6164-190-X, A Neural - Fuzzy Approach to Linguistic Knowledge Acquisition and Assessment in Spatial Decision Making
82. Sahu, B.K. (2001), Aeromagnetics of continental areas flanking the Indian Ocean; with implications for geological correlation and reassembly of Central Gondwana
83. Alfestawi, Y.A.M. (2001), 90-6164-198-5, The structural, paleogeographical and hydrocarbon systems analysis of the Ghadamis and Murzuq Basins, West Libya, with emphasis on their relation to the intervening Al Qarqaf Arch
84. Liu, X. (2001), 90-5808-496-5, Mapping and Modelling the Habitat of Giant Pandas in Foping Nature Reserve, China
85. Oindo, B.O. (2001), 90-5808-495-7, Spatial Patterns of Species Diversity in Kenya
86. Carranza, E.J.M. (2002), 90-6164-203-5, Geologically-constrained Mineral Potential Mapping
87. Rugege, D. (2002), 90-5808-584-8, Regional Analysis of Maize-Based Land Use Systems for Early Warning Applications
88. Liu, Y. (2002), 90-5808-648-8, Categorical Database Generalization in GIS
89. Ogao, P. (2002), 90-6164-206-X, Exploratory Visualization of Temporal Geospatial Data using Animation
90. Abadi, A.M. (2002), 90-6164-205-1, Tectonics of the Sirt Basin – Inferences from tectonic subsidence analysis, stress inversion and gravity modelling
91. Geneletti, D. (2002), 90-5383-831-7, Ecological Evaluation for Environmental Impact Assessment
92. Sedogo, L.G. (2002), 90-5808-751-4, Integration of Participatory Local and Regional Planning for Resources Management using Remote Sensing and GIS
93. Montoya, L. (2002), 90-6164-208-6, Urban Disaster Management: a case study of earthquake risk assessment in Carthago, Costa Rica
94. Mobin-ud-Din, A. (2002), 90-5808-761-1, Estimation of Net Groundwater Use in Irrigated River Basins using Geo-information Techniques: A case study in Rechna Doab, Pakistan
95. Said, M.Y. (2003), 90-5808-794-8, Multiscale perspectives of species richness in East Africa
96. Schmidt, K.S. (2003), 90-5808-830-8, Hyperspectral Remote Sensing of Vegetation Species Distribution in a Saltmarsh
97. Lopez Binnquist, C. (2003), 90-3651-900-4, The Endurance of Mexican Amate Paper: Exploring Additional Dimensions to the Sustainable Development Concept
98. Huang, Z. (2003), 90-6164-211-6, Data Integration for Urban Transport Planning
99. Cheng, J. (2003), 90-6164-212-4, Modelling Spatial and Temporal Urban Growth
100. Campos dos Santos, J.L. (2003), 90-6164-214-0, A Biodiversity Information System in an Open Data/Metadatabase Architecture
101. Hengl, T. (2003), 90-5808-896-0, PEDOMETRIC MAPPING, Bridging the gaps between conventional and pedometric approaches

102. Barrera Bassols, N. (2003), 90-6164-217-5, Symbolism, Knowledge and management of Soil and Land Resources in Indigenous Communities: Ethnopedology at Global, Regional and Local Scales
103. Zhan, Q. (2003), 90-5808-917-7, A Hierarchical Object-Based Approach for Urban Land-Use Classification from Remote Sensing Data
104. Daag, A.S. (2003), 90-6164-218-3, Modelling the Erosion of Pyroclastic Flow Deposits and the Occurrences of Lahars at Mt. Pinatubo, Philippines
105. Bacic, I.L.Z. (2003), 90-5808-902-9, Demand-driven Land Evaluation with case studies in Santa Catarina, Brazil
106. Murwira, A. (2003), 90-5808-951-7, Scale matters! A new approach to quantify spatial heterogeneity for predicting the distribution of wildlife
107. Mazvimavi, D. (2003), 90-5808-950-9, Estimation of Flow Characteristics of Ungauged Catchments. A case study in Zimbabwe
108. Tang, X. (2004), 90-6164-220-5, Spatial Object Modelling in Fuzzy Topological Spaces with Applications to Land Cover Change
109. Kariuki, P. (2004), 90-6164-221-3, Spectroscopy and Swelling Soils; an integrated approach
110. Morales, J. (2004), 90-6164-222-1, Model Driven Methodology for the Design of Geo-information Services
111. Mutanga, O. (2004), 90-5808-981-9, Hyperspectral Remote Sensing of Tropical Grass Quality and Quantity
112. Šliužas, R.V. (2004), 90-6164-223-X, Managing Informal Settlements: a study using geo-information in Dar es Salaam, Tanzania
113. Lucieer, A. (2004), 90-6164-225-6, Uncertainties in Segmentation and their Visualisation
114. Corsi, F. (2004), 90-8504-090-6, Applications of existing biodiversity information: Capacity to support decision-making
115. Tuladhar, A. (2004), 90-6164-224-8, Parcel-based Geo-information System: Concepts and Guidelines
116. Elzakker, C. van (2004), 90-6809-365-7, The use of maps in the exploration of geographic data
117. Nidumolu, U.B. (2004), 90-8504-138-4, Integrating Geo-information models with participatory approaches: applications in land use analysis
118. Koua, E.L. (2005), 90-6164-229-9, Computational and Visual Support for Exploratory Geovisualization and Knowledge Construction
119. Blok, C.A. (2005), Dynamic visualization variables in animation to support monitoring of spatial phenomena
120. Meratnia, N. (2005), 90-365-2152-1, Towards Database Support for Moving Object Data
121. Yemefack, M. (2005), 90-6164-233-7, Modelling and monitoring Soil and Land Use Dynamics within Shifting Agricultural Landscape Mosaic Systems
122. Kheirkhah, M. (2005), 90-8504-256-9, Decision support system for floodwater spreading site selection in Iran
123. Nangendo, G. (2005), 90-8504-200-3, Changing forest-woodland-savanna mosaics in Uganda: with implications for conservation
124. Mohamed, Y.A. (2005), 04-15-38483-4, The Nile Hydroclimatology: impact of the Sudd wetland (Distinction)
125. Duker, A.A. (2005), 90-8504-243-7, Spatial analysis of factors implicated in mycobacterium ulcerans infection in Ghana

126. Ferwerda, J.G., (2005), 90-8504-209-7, Charting the Quality of Forage: Measuring and mapping the variation of chemical components in foliage with hyperspectral remote sensing
127. Martinez, J. (2005), 90-6164-235-3, Monitoring intra-urban inequalities with GIS-based indicators. With a case study in Rosario, Argentina
128. Saavedra, C. (2005), 90-8504-289-5, Estimating spatial patterns of soil erosion and deposition in the Andean region using Geo-information techniques. A case study in Cochabamba, Bolivia
129. Vaiphasa, C. (2006), 90-8504-353-0, Remote Sensing Techniques for Mangrove Mapping
130. Porwal, A. (2006), 90-6164-240-X, Mineral Potential Mapping with Mathematical Geological Models
131. Werff, H. van der, (2006), 90-6164-238-8, Knowledge-based remote sensing of complex objects: recognition of spectral and spatial patterns resulting from natural hydrocarbon seepages
132. Vlag, D.E. van de (2006), 90-8504-384-0, Modeling and visualizing dynamic landscape objects and their qualities

Publications of the author

Kraak, M.J. and **Van de Vlag, D.E.** (in prep.) Understanding spatio-temporal patterns: visual ordering of space and time. For Autocarto 2006.

Van de Vlag, D.E., Kraak, M.J., Nieuwenhuis, W. and Schouwenburg, M. (in review). Temporal Ordered Space Matrix: representation of multivariate spatiotemporal data, *Computers & Geosciences*.

Van de Vlag, D.E. and Stein, A. (in review). Fuzzy decision tree approach for coastal object detection and its quality elements. *IEEE Transactions on Geoscience and Remote Sensing*.

Van de Vlag, D.E. and Stein, A. (in review). Landsat TM images to define Spatio-Temporal Ontologies for Beach Nourishments, *Journal of Environmental Informatics*.

Vasseur, B., **Van de Vlag, D.E.**, Stein, A, Jeansoulin, R., Dilo, A. (in review). Quality-aware ontologies for dynamic spatial objects with indeterminate boundaries. *Applied Ontology*.

Jetten, V., Poesen, J. Nachtergaele, J. and **Van de Vlag, D.E.** (in press). Spatial modelling of ephemeral gully incision, a combined empirical and physical approach. In: Soil Erosion and Sediment Redistribution in River Catchments. Eds. P. Owens and A. Collins. CAB International.

Van de Vlag, D.E., Vasseur, B., Stein, A., Jeansoulin, R. (2005). An Application of Problem and Product Ontologies for the Revision of Beach Nourishments, *International Journal of Geographical Information Science* **19(10)**, pp. 1057-1072.

Van de Vlag, D.E., and Malunda, A.A. (2005). Interactive Visualization of Geospatial Dynamics. *Geoinformatics* **8**, July/August 2005, pp 28-31.

Van de Vlag, D.E. and Kraak, M.J. (2005). Temporal Ordered Space Matrix: representation of multivariate spatiotemporal data. In: GISPLANET 05 Conference Proceedings, S01 Spatial Knowledge, 13p.

Van de Vlag, D.E. (2004). Concepts and Representation of Beach Nourishments by Spatio-temporal Ontologies. In: The Netherlands and the North Sea. Dutch Geography 2000-2004, Nederlandse Geografische Studies 325, T. Dietz, P. Hoekstra, F. Thissen (eds). Utrecht, The Netherlands, pp 45-55.

Van de Vlag, D.E. and Kraak, M.J., (2004). Multivariate Visualization of Data Quality Elements for Coastal Zone Monitoring, ISPRS Conference Proceedings Commission IV, ISPRS 15-23 July 2004 Istanbul, Turkey.

Vasseur, B., **Van de Vlag, D.E.**, Stein, A., Jeansoulin, R., Dilo, A. (2004). Spatio-temporal Ontology for Defining the Quality of an Application. In: Proceedings of the ISSDQ '04, *GeoInfo* **28(a)**, Vienna, Austria, pp. 67-81.

Stein, A., Dilo, A., Lucieer, A., **Van de Vlag, D.E.** (2004). Definition, Identification and Decision Ontologies for Revision of Vague Spatial Objects. In: Proceedings of the ISSDQ '04, *GeoInfo* **28(a)**, Vienna, Austria, pp. 99-115.

Van de Vlag, D.E., Stein A., Vasseur, B. (2004). Concepts and Representation of Beach Nourishments by Spatio-temporal Ontologies. In: Proceedings of the ISSDQ '04, *GeoInfo* **28(b)**, Vienna, Austria, pp. 353-368.

Van de Vlag, D.E., Stein A., Vasseur, B. (2004). Concepten en Weergave van Zandsuppleties op Ameland door Spatio-temporele Ontologieën. Uitgebreide Samenvatting, GIN dag.

Jetten, V., Poesen, J. Nachtergaele, J. and **Van de Vlag, D.E.** (2004). Spatial modelling of ephemeral gully incision, a combined empirical and physical approach. Symposium: "Soil erosion and sediment redistribution in river catchments: measurement, modelling and management in the 21st Century". September 2003, National Soil Resources Institute (NSRI), Cranfield university, Silsoe, UK.

Van de Vlag, D.E. (2003). Application with Undeterminate Boundaries and Dynamic Monitoring. Internal Report REV!GIS project, R323. pp. 16.

Van de Vlag, D.E. (2002). Hierarchical Modelling and its Visualization of Multi-temporal Geospatial Data. Geovisualization 2002, Albufeira, Portugal, extended abstract.

Stein, A., and **Van de Vlag, D.E.** (2002), Application with Undeterminate Boundaries and Dynamic Monitoring. Internal Report REV!GIS project, R223. pp. 12.

Van de Vlag, D.E., Jetten, V., Nachtergaele, J. and Poesen. J. (2000). Event based modelling of gully incision and development in the Belgian loess belt. International Symposium on Gully erosion under global change, April 2000, Leuven, Belgium, extended abstract.

Biography

Daniël Emanuël van de Vlag was born on the 13th of June 1975 in Apeldoorn, The Netherlands. From 1993 to 1998, he studied Physical Geography at Utrecht University, with a specialization in land degradation and geographical information systems. He wrote his Masters thesis on modeling the soil water balance for a river catchment in France. From 1999 to 2000, he worked as a research officer for the European MWISED project on modeling gully erosion in the Belgian loess belt. At the end of 2000, Daniël went to Australia and New Zealand to travel around for 10 months. In October 2001, he started as a Ph.D. researcher at the International Institute for Geo-Information Science and Earth Observation (ITC) in Enschede, The Netherlands. His research focused on developing methods and techniques to model dynamic landscape objects. His research interests include ontologies of spatial data, remote sensing, fuzzy classification, spatial data quality and visualization. The developed methods are applied on a case study on beach nourishments in the north-western part of the island Ameland, The Netherlands. During his Ph.D. research, he did collaborative research as a visiting scientist at the University of Provence in Marseille, France. He has presented his work at several international conferences and has published in international peer-refereed journals. Additionally, he has taught lectures on ontology, fuzzy classification and spatial data quality at ITC.

The work in this thesis was funded by:

International Institute for Geo-Information Science and Earth Observation (ITC), Enschede, The Netherlands

The work was carried out according to the PhD regulations of:

Wageningen University, The Netherlands

This research was funded by the European Community project REV!GIS (IST-1999-14189).