A Process-Oriented Data Model for Fuzzy Spatial Objects

Tao Cheng

Promotor:	Dr. Ir. Martien Molenaar Professor in Geoinformatics and Spatial Data Acquisition
Co-Promotor:	Dr. Wolfgang Kainz Professor in Geoinformatics, Spatial Information Theory and Applied Computer Science

Tao Cheng

A Process-Oriented Data Model for Fuzzy Spatial Objects

Proefschrift
Ter verkrijging van de graad van doctor
Op gezag van de rector magnificus,
Dr. C. M. Karssen,
Van de Landbouwuniversiteit Wageningen,
In het openbaar te verdedigen
Op dinsdag, 18 mei 1999,
Des middags om half twee in de Aula

ITC Publication Series, No. 68
ISBN 90-6164-164-0
The research presented in this thesis was conducted at the
International Institute for Aerospace Survey and Earth Sciences (ITC)
P. O. Box 6,
7500 AA Enschede,
The Netherlands.

CIP-DATA KONINKLIJKE BIBLIOTHEEK, DEN HAAG

ISBN 90-5808-038-2

A Process-Oriented Data Model for Fuzzy Spatial Objects

Thesis Wageningen – With index, ref. – With summary in Dutch.
ISBN 90-5808-038-2

Subject headings: GIS, environmental modeling, spatial-temporal data model, data quality, coastal zone monitoring and management.

Tao Cheng

A Process-Oriented Data Model for Fuzzy Spatial Objects

Thesis
To fulfill the requirements for the degree of doctor
On the authority of the rector magnificus
Dr. C. M. Karssen
To be publicly defended
On Tuesday, 18 May 1999,
At 13:30 hours in the Auditorium
of Wageningen Agricultural University

NN08201, 2612

Propositions

Tao Cheng, 18 May, 1999, A process-oriented data model for fuzzy spatial objects, PhD Dissertation.

1. The concept of objects with fuzzy spatial extent is a generalization of the traditional crisp object concept, unifying the boundary-oriented approach and pixel-oriented approach. The concept also links the object-oriented to the field-oriented characteristics of natural phenomena.
   - This thesis

2. When studying the uncertainty of spatial objects one should not look at the geometric aspects independently of the thematic aspects, because in most cases geometric uncertainty is a consequence of uncertainty of the thematic aspects.
   - This thesis

3. Several measures can be defined mathematically to evaluate the change of the fuzzy area and volume of spatial objects. The semantics of these measures is not trivial, however, and should be studied with care.
   - This thesis

4. The traditional approach in which objects are seen as static is insufficient for the modeling of dynamic environments. Therefore, it is better to consider objects as dynamic and represent them as processes consisting of sequences of states and state transitions.
   - This thesis

5. When using GIS technology, the user should take care not to model maps, but reality.

6. The application of GISs is generally still very primitive since the model provided by GISs is either too simple to describe the relationship between objects in reality or too complicated for the user to understand.

7. Objects are conceptually defined by means of a homogeneity criterion for some of its characteristics. Since homogeneity criteria are seldom crisp, we have to accept uncertainty in the identification of objects.
   - This thesis

8. To understand the world we need to observe, measure and represent it. How we perceive it, how we measure it, and how we represent it depends on the conceptual and technical tools we use. Therefore our image of the world is greatly determined by the metaphors that can be related easily to these tools (after Molenaar, 1998).

9. Humans, not computers, create uncertainty when natural language, with its vagueness, is translated into mathematical symbols.

10. Modeling real world situations always leads to dissatisfaction, because each time a model has been improved we become aware of a higher level of complexity in reality, which our models cannot yet handle.

11. It is easier to learn how to handle new technology than how to use it.

12. It is hard if you have no choice, but it is even harder if you have too many choices.

TO MY PARENTS

Foreword

This thesis is the most important outcome of my research work as a Ph.D. candidate of the International Institute for Aerospace Survey and Earth Sciences (ITC), The Netherlands. Now that the thesis is almost ready, I realize how complicated it is to acquire the title 'Doctor'. It is complicated due to my research topic, which deals with the generalization of current object concepts in GISs, from crisp to fuzzy and from spatial to temporal. It is complicated because I am studying abroad and have to write the thesis in a foreign language. It is also complicated because I am a wife and became a mother last year. And it is complicated because it takes so long, and yet there is still not enough time, so sometimes I had to work on weekends. Dealing with this complex situation has challenged me and enriched my life. Fortunately, I got tremendous amount of help and support from many people and many institutions, whose involvement and assistance reduced the complexity and so contributed to the results reported in this thesis. I wish to extend my sincere thanks to them all.

I should like to express my sincere gratitude to my promoter, Professor Dr. Martien Molenaar, for his effective guidance and supervision. His knowledge of geoinformation science always helped me get through the scientific problems. I am very grateful for his introduction to the formal data structure (FDS) and the related theory proposed by him, and for his helping me to formalize the model proposed in this thesis, which is based upon FDS. I enjoyed our stimulating discussions; he was always very patient with me and gave me constructive suggestions. He taught me how to formalize ideas in formulas and express them in graphs. His invaluable comments on the structure and content of this thesis have considerably improved its quality and readability.

I am also very grateful to my supervisor, Dr. Robert A. van Zuidam, for initiating this research project and arranging the fellowship and funds. He spent a lot of time and effort to teach me coastal geomorphology and management. In order to help me understand and know more about this field, he introduced me to RWS-MD (The Survey Department of the Directorate-General for Public Works and Water Management of the Dutch Ministry of Transport, Public Works and Water Management), sent me to European courses at IHE-Delft (the International Institute for Infrastructural, Hydraulic and Environmental Engineering, Delft) and ITC, and arranged fieldwork on Ameland. Because of his contribution, the research project was able to build a strong application case for the model developed in this thesis. His critical comments helped improve the quality of the thesis greatly.

I owe much to Prof. Dr. Wolfgang Kainz, my co-promoter and supervisor, who taught me much on fuzzy set theory, category theory and database design. Beyond that, he was very patient to help me solve technical issues, such as using Microsoft Word in order to produce a better layout for the thesis. I appreciate very much that he was able to review the thesis despite a very tight schedule. His suggestions and comments contributed greatly to the final shape of this thesis.

I also appreciate very much the advice, assistance and encouragement of Dr. Theo Bouloucos. He helped me figure out the mathematical formulas for combining errors and fuzziness. He recommended a great deal of material and spent a lot of time in reading my thesis, always followed by constructive comments.

I am indebted to Marieke Elveld, my colleague and former office-mate, for her involvement in the collection and preparation of the data for my case study. She put in a lot of effort to help me understanding the geomorphological situation of Ameland, and translated Dutch materials into English for me. I appreciate her time and efforts in arranging the group visit to RWS-MD.

Mrs. Wilma Grootenboer, ITC's nurse, gave me tremendous support and tender care during my stay in ITC, especially during my pregnancy and when my daughter was born. She was always prepared to help me. I acknowledge her for her warm heart.

From time to time I got help from other colleagues at ITC. I appreciate discussions with Dr. Lucas Jassen on dynamics of fuzzy objects, Dr. Rolf De By on spatio-temporal data modeling, Professor Dr. Menno-Jan Kraak and Ms. Connie Blok on visualization, Ms. Yuxian Sun and Ms. Sisi Zlatanova on ArcView, Dr. Ben Gorte on image classification, and Professor Dr. Alfred Stein on statistics. They all provided valuable contributions to my research.

Occasionally I got valuable comments from specialists in GISs, such as Andrew Frank, Michael Worboys, Mei Yuan, and many others, when I participated in international conferences. I thank them all. I appreciate the meeting with Janathan Raper and David Livingstone, at Birkbeck College, University of London, for sharing their experience in spatio-temporal data modeling for coastal zone management and monitoring. I also would like to thank Peter van Oosterom (Dutch Cadaster, Apeldoorn) for stimulating discussions on data modeling and structures for 4-D GISs.

The secretaries of the Department of Geoinformatics and the Department of Earth Resources Survey (ITC) were always glad to lend me a hand when I needed help. I am also grateful to Dr. Elizabeth Kosters for her help and support. Mr. Ard Blenke and Mr. Vasilios Retsios were always helpful in providing hardware and software support. Mrs. Anneke Homan, Ms. Saskia Tempelman, Prof. Jan Nossin, Ms. Annet Pril, Mr. Ad Bakker, Mr. Wan Bakx, Mrs. Riet Allessie, Mrs. Bettine Geerdink-Schukking and the ladies of ITC library are acknowledged for their support and help. I thank Mr. Ian Cressie for editing this thesis within a remarkably short timespan. I also thank Mr. Andries Menning for designing the cover.

I appreciate the friendship with former PhD candidates of ITC, Dr. Yiman Wang, Dr. Wanning Peng, Dr. Morakot Pilouk, Dr. Bin Jiang, Dr. Ali Roshannejad, Dr. Dr. Yaser Bisher and Dr. Christine Pohl. I thank them all for stimulating discussions on different issues that are related to my research.

The research has been funded by ITC. RWS-MD kindly made the data for the Ameland case available. Many people of RWS-MD, such as Jeroen Huising and Ed Vassen, provided very useful information for the case study and I thank them for their kind help.

I should like to take this opportunity to thank my former colleagues in the Department of Photogrammetry and Remote Sensing and at the Research Center of GIS, Wuhan Technical University of Surveying and Mapping (WTUSM), People's Republic of China. I thank Professor Dr. Deren Li for his help and support whether I was at WTUSM or ITC. I also thank Ms. Sun Ping, Ms. Wang Fen, Ms. Zhou Yueqin, Ms. Song Aihong, Mr. Liu Yaolin and Mr. Wu Guofong for their kind help. I also appreciate very much the friendship of the other chinese students at ITC. I enjoyed sharing chinese culture and customs with them by organizing chinese parties in DISH.

I owe a lot to my dear father and mother. Their continuous support and love have encouraged me to overcome many difficulties in my life. I am very grateful to them for taking care of my little daughter. I am indebted to my brothers and sister-in-law. They took upon themselves the sole responsibility of taking care of my parents and they telephoned me regularly with news and encouragement.

Last but not least, I would like to thank my husband, Dr. Donggen Wang, for his constant support, understanding and love. He was the first reader of my papers when this research started and his critical comments improved their readability. I owe a great deal to my daughter, Nancy. She had to go to kindergarten and to stay with my parents because of our tight schedule. I dedicate this thesis to my husband, my daughter and my parents.

Summary

The complexity of the natural environment, its polythetic and dynamic character, requires appropriate new methods to represent it in GISs, if only because in the past there has been a tendency to force reality into sharp and static objects. A more generalized spatio-temporal data model is requirede to deal with fuzziness and dynamics of objects. This need is the motivation behind the research reported in this thesis. In particular, the objective of this research was to develop a spatio-temporal data model for objects with fuzzy spatial extent.

This thesis discusses three aspects related to achieving this objective:

- identification of fuzzy objects,
- detection of dynamic changes in fuzzy objects, and
- representation of objects and their dynamics in a spatio-temporal data model.

For the identification of fuzzy objects, a six-step procedure was proposed to extract objects from field observation data: sampling, interpolation, classification, segmentation, merging and identification. The uncertainties involved in these six steps were investigated and their effect on the mapped objects was analyzed. Three fuzzy object models were proposed to represent fuzzy objects of different application contexts. The concepts of conditional spatial extent, conditional boundary and transition zones of fuzzy objects were put forward and formalized based upon the formal data structure (FDS). In this procedure, uncertainty was transferred from thematic aspects to geometric aspects of objects, i.e. the *existential* uncertainty was converted to *extensional* uncertainty. The spatial effect of uncertainty in thematic aspect was expressed by the relationship between uncertainty of a cell belonging to the spatial extent of an object and the uncertainty of the cell belonging to classes.

To detect dynamic changes in fuzzy objects, a method was proposed to identify objects and their state transitions from fuzzy spatial extents (regions) at different epochs. Similarity indicators of fuzzy regions were calculated based upon overlap between regions at consecutive epochs. Different combinations of indicator values imply different relationships between regions. Regions that were very similar represent the consecutive states of one object. By linking the regions, the historic lifelines of objects are built automatically. Then the relationship between regions became the relationship or interactions between objects, which were expressed in terms of processes, such as shift, merge or split. By comparing the spatial extents of objects at consecutive epochs, the

change of objects was detected. The uncertainty of the change was analyzed by a series of change maps at different certainty levels. These can provide decision makers with more accurate information about change.

For the third, and last, a process-oriented spatio-temporal data model was proposed to represent change and interaction of objects. The model was conceptually designed based upon the formalized representation of state and process of objects and was represented by a star-styled extended entity relationship, which I have called the Star Model. The conceptual design of the Star Model was translated into a relational logical design since many commercial relational database management systems are available. A prototype of the process-oriented spatio-temporal data model was implemented in ArcView based upon the case of Ameland. The user interface and queries of the prototype were developed using Avenue, the programming language of ArcView.

The procedure of identification of fuzzy objects, which extracts fuzzy object data from field observations, unifies the existing field-oriented and object-oriented approaches. Therefore a generalized object concept – object with fuzzy spatial extent – has been developed. This concept links the object-oriented and the field-oriented characteristics of natural phenomena. The objects have conditional boundaries, representing their object characteristics; the interiors of the objects have field properties, representing their gradual and continuous distribution. Furthermore, the concept can handle both fuzzy and crisp objects. In the fuzzy object case, the objects have fuzzy transition or boundary zones, in which conditional boundaries may be defined; whereas crisp objects can be considered as a special case, i.e. there are sharp boundaries for crisp objects. Beyond that, both the boundary-oriented approach and the pixel-oriented approach of object extraction can use this generalized object concept, since the uncertainties of objects are expressed in the formal data structures (FDSs), which is applicable for either approach.

The proposed process-oriented spatio-temporal data model is a general one, from which other models can be derived. It can support analysis and queries of time series data from varying perspectives through location-oriented, time-oriented, feature-oriented and process-oriented queries, in order to understand the behavior of dynamic spatial complexes of natural phenomena. Multi-strands of time can also be generated in this Star Model, each representing the (spatio-temporal) lifeline of an object. The model can represent dynamic processes affecting the spatial and thematic aspects of individual objects and object complexes. Because the model explicitly stores change (process) relative to time, procedures for answering queries relating to temporal relationships, as well as analytical tasks for comparing different sequences of change, are facilitated.

The research findings in this thesis contribute theoretically and practically to the development of spatio-temporal data models for objects with fuzzy spatial extent.

Keywords: object-oriented model, field-oriented model, uncertainty, fuzziness, errors, object identification, dynamics, fuzzy spatial extent, transition zone, conditional boundary, fuzzy object, formal data structure (FDS), pixel-oriented approach, boundary-oriented approach, spatio-temporal data model, process, state, state transition.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background of the research

Natural phenomena are observed in environmental studies with the aim of explaining and predicting their dynamics. In order to facilitate these studies, changes in such phenomena should be detected, analyzed and represented in geographical information systems - GISs. Natural phenomena occur in a physical environment of complex and continuous character, so their representation in GISs requires abstraction and discretization. How the abstraction and discretization is done depends to a large extent on the way people perceive the physical environment and the processes that take place in it.

Conventionally there are two distinct ways of perceiving the physical environment, i.e. the field view and the object view (Goodchild, 1993; Burrough and McDonnel, 1998; Molenaar, 1998a). The field view assumes that the physical environment is a *continuous* field, and each point in the space can be characterized in terms of a set of attributes with values that depend on the location of the observed point. In this view, processes result in changes of attribute values with time, where a value at time t_n depends on the attribute values at the same location and in its neighborhood at time t_{n-1}. Therefore, the field view is a *location*-dependent view. The object view, however, considers that the world is made up of discrete objects with *crisp* boundaries and a well-defined set of attributes that are internally *homogenous*. The description of boundaries (called geometry) and the attributes (called thematic attributes) of the object are linked together through a unique object identifier. In this view, processes are based upon the behavior of interacting objects, while the resulting patterns can be expressed by the state of these objects (Molenaar, 1998a).

These two views lead to two spatial data models used in GISs, i.e. the field model (Couclelis, 1992) and the object model (Nunes, 1991). The field model is commonly used to map attributes (e.g. elevation, rainfall, air pressure) related to locations, which may be represented as isolines (e.g. contours) or modeled by mathematical functions (Goodchild, 1992). The object model is especially related to represent man-made facilities, such as a well, road or a building, which can be considered as point objects, line objects or polygon or area objects.

As most physical theories are spatially and temporally continuous, the field model dominates environmental modeling. But the evaluation of these theories almost always requires the existence of entities that are spatially and temporally discrete (Raper and Livingstone, 1995). For example, in geomorphology, the first step to understanding the processes that form a given landscape is often the identification of 'landforms' or 'features', such as 'beach' or 'coastline'. For this reason, recently the object model has been applied in environmental modeling (Mason *et al.*, 1994; Raper and Livingstone, 1995).

Natural phenomena can often only be sampled sparsely, at a limited number of points, which can then be interpolated to generate a full raster. In such cases, objects can often be identified in two steps: firstly the field is categorized in classes related to the object types to be detected. Secondly, the raster with its classified cells may be segmented to identify the spatial extent of the objects. In the study area (Chapter 2) the beach is defined as the area above the low-water line but beneath the high-water line. The area of beach will be derived from an elevation raster interpolated from a profile measurement along the coastline. Therefore, in this case, objects are derived from field observation data through categorization. In other words, the object representation is derived from field representation when it is used to describe natural phenomena (see also Subsection 3.4.2).

Both the field model and the object model have been implemented in GISs in vector structures (polygon, triangular network, irregular points and contours) and in raster structures (cell grids and regular points) (Kemp, 1992; Molenaar, 1995, 1998a; Burrough and McDonnel, 1998).

The nature of a physical environment is complex and polythetic. The 'continuous' attribute values of the field sometimes show discontinuities, e.g. some environmental gradients can change quite rapidly in space, leading to abrupt spatial vegetation changes (Brown, 1998, p.126). On the other hand, objects are not necessarily internally homogeneous (e.g. inclusions in forests), and the boundaries between different objects are not always crisp and unambiguous (e.g. gradual transition between forest types). Therefore both field and object models are simplistic abstractions of such complex and polythetic phenomena, i.e. they are extremes of the total spectrum of spatial modeling concepts. They have, among others, the following two shortcomings (Burrough and Frank, 1995):

- They are exact models. They do not allow errors or imprecision either in the description or in identification of the entities involved.
- Both models are static descriptions of the reality. They are not dynamic and do not describe spatial changes and/or change with time.

1.1.1 Uncertainty in spatial modeling

Modern literature records considerable interest in issues of uncertainty in geographic data. Research in this area has passed through two phases. In the first phase, most studies deal with uncertainties in measurement or in processing. In the second phase, the studies set out to model the uncertainties due to the complex nature of entities in reality.

In the first phase, stochastic errors in data acquisition and operation were dealt with, based upon the premise that uncertainty exists due to random processes within the spatial system. Hence, statistical and probabilistic models, and error propagation theory, were applied (Altman, 1994).

There are stochastic components in the measurement of geometric aspects of spatial phenomena. This results in uncertainty of location or boundary geometry of objects. The epsilon band method (Dunn *et al.*, 1984) is well known in this context for representing the uncertainty of the geometric aspect, i.e. the uncertainty of a position is usually represented by an epsilon band around a point or a linear feature. For example, Chrisman (1982) used the epsilon model to express cartographic map error; Blakemore (1984) used epsilon bands (applying a point in polygon procedure) to quantify the uncertainty of assigning manufacturing plants to employment office area.

The uncertainty of attribute data is evaluated by different approaches with respect to their types. For categorical data (e.g. land-cover type) accuracy is expressed by an error matrix, which is generated by comparing values on a map layer with ground truth (Lovett, 1995). For numerical attributes, the accuracy, the frequency of misclassification) is estimated by probabilistic methods, e.g. viewshed definitions in Fisher (1992) and flooding risks in Hunter and Goodchild (1995). The propagation of errors in spatial analysis is investigated by tracing the propagation of errors in attribute data to the results of map overlays (Heuvelink *et al.*, 1989) and cartographic classification (Heuvelink and Burrough, 1993).

While some progress has been made in developing error models for digitized lines and areas (Bolstad *et al.*, 1990), it is much more difficult to construct comprehensive models to represent uncertainties in the data source of complex geographical objects and to describe their propagation in the processing (Goodchild, 1989, p. 109). Although the epsilon band method can describe uncertainty in the position of a line or point, it cannot deal with the problem of heterogeneity of attribute in area objects. Nowadays, the study of uncertainty and precision is moving to the second phase, to uncertainties that are inherent and cannot be attributed to randomness, e.g. heterogeneity in mapping units due to inclusions or gradual transitions. Alternative methodologies are required to handle this kind of uncertainty.

Fuzziness is a type of imprecision that characterizes classes that do not have crisply defined boundaries. These inexactly defined classes are called fuzzy sets. They allow partial memberships, which can vary continuously between 0 and 1. Since Zadeh (1965) introduced the idea of 'fuzzy set' dealing with the inexact concepts in a definable way, the theory of fuzzy sets has been developed and applied to many disciplines (Klir and Folgar, 1988).

In the 1970s, the fuzzy set theory was introduced in handling the uncertainty of quantification in human geography (Gale, 1972; Pipkin 1978; Leung, 1979). Later on, fuzzy sets were applied in GISs (Robinson and Strahler, 1984). Robinson (1988) proposed a model for handling inexactness in geographic database that was based on fuzzy logic. Burrough (1989) demonstrated the use of fuzzy sets for soil survey and land evaluation. Wang *et al.* (1990) proposed a fuzzy relational data model for conventional GIS software. Robinson (1990) demonstrated the use of fuzzy sets for representing qualitative linguistic spatial relationships. In remote sensing, the fuzzy set theory has been applied to image classification (Bezedk *et al.*, 1984; Wang, 1990; Inomata and Ogata, 1992). It has also been applied in reasoning in knowledge base systems for GIS and remote sensing (Skidmore *et al.*, 1992). The majority of applications focus on modeling uncertainty in thematic aspects (Usery, 1996; Brown, 1998), i.e. fuzzy classification due to ambiguity in class definition. In these cases thematic attributes (classification result) describing these mapping units are treated as ambiguous or uncertain, while the spatial units are usually represented by crisp boundaries.

Nowadays, fuzzy set theory is being applied to model the uncertainty in geometric aspects of mapping units. These works deal with cases where the locations of boundaries are indeterminate due to uncertain classification or fuzzy definition of the mapping units (Fisher, 1994; Lowell, 1994; Edwards and Lowell, 1996; Usery, 1996; Brown, 1998; Edwards *et al.*, 1998; Molenaar, 1996, 1998b). Among other applications, fuzzy boundaries have been proposed for cartographic maps (Wang and Hall, 1996), fuzzy geometry between objects has been measured (Altman, 1995) and fuzzy topology has been put forward to describe the topologic relationship of fuzzy regions (Dijkmeijer and Hoop, 1996; Clementini and Felice, 1996; Zhan, 1997). Also, people have been investigating source of ambiguity from the cognitive-linguistic viewpoint (Couclelis, 1996; Smith and Varzi, 1997; Plewe, 1997; Coulelis and Gottsegen, 1997). Most advanced progress in research on fuzzy objects can be found in Burrough and Frank (1996) and in Goodchild and Jeansoulin (1998).

In summary, the influence of stochastic errors of imperfect data is discussed in the first phase and the uncertainty due to the fuzzy nature of objects and fields is investigated in the second phase. However, usually these two kinds of uncertainties exist simultaneously so that their combined effect should be taken into account in the modeling process. For example, the locations of boundaries on vegetation maps are not certain, due to the fact that (1) the definition of the mapping units is inexact (some vegetation distributions exhibit naturally continuous variation, with gradual transitions between vegetation types), and (2) data used for mapping are never perfect (errors contained in measurements, which represent either a sample of the vegetation or variables that serve as surrogates for information on vegetation) (Brown, 1998).

However, existing studies do not consider these two aspects together, with the exception of Heuvelink and Burrough (1993) and Goodchild *et al.* (1992). Their discussions were based upon, however, Monto Carlo stimulation, and the analytical relationship of these two aspects were hardly elaborated upon. Furthermore, they only discussed the effect of uncertainty on the classification, not on the geometric aspects (including spatial extent and boundary) of objects, which should be the final result of mapping. Therefore, research to date could not give a clear picture about the relationship of these two uncertainty aspects, nor a complete measurement of their effects on the thematic and geometric aspects of an object. Many researchers just assume that the fuzzy spatial extent is there due to fuzzy classification. It is not clear what the causes of uncertainty are, no what effect they have.

Molenaar (1996) and David *et al.* (1996) proposed conceptual data models to represent indeterminate objects and fields in a computerized information system. The syntax for fuzzy objects has been developed by incorporating uncertainties into a formal syntax model as a generalization of the model used for conventional crisp objects (Molenaar, 1994, 1998a). However, the model still needs to be tested in practice. The procedure to identify a fuzzy object has not been thoroughly investigated. Nevertheless, "indeterminacy is a way of life in many areas and it is sensible to have formal ways for handling it... (Burrough, 1996, p. 27)". Previous work on fuzzy objects could serve as a starting point for this research. Operationalization of these fuzzy object concepts still has a long way to go. Important questions to be solved are: what are the causes of uncertainty; how to identify objects and their uncertainties in the process of modeling physical environments; how do uncertainties propagate in the modeling process; and what is the relationship between errors and fuzziness and their effect on thematic and geometric aspects of objects? These issues constitute the first point to be addressed in this thesis.

1.1.2 Modeling spatial dynamics

The second shortcoming of conventional field and object models is that entities are seen as being frozen in time. However, as many phenomena in the real world are changing with time, GISs should be capable of representing, analyzing and predicting changes of spatial information over time. There are four reasons for needing temporal information and reasoning about time: (1) to understand the rules that govern changes observed in the world; (2) to *explain* the state of the world at some earlier time; (3) to p*redict* the state of the world in the future; (4) to p*lan* a sequence of actions that will lead to a desirable future. These four related points are general enough to be of interest for geographic information systems as well, i.e. "any temporal GIS should provide its users with the means to **understand, explain** and **predict** future events in the world it presents (Al-Taha and Frank, 1993)".

As conventional data models neglect time-induced spatial changes, in recent years, research has emphasized the development of so-called 'spatio-temporal' data models. Current spatio-temporal models are oriented toward the representation of the evolution of spatial entities. There are models proposed to organize temporal information by time-stamping spatial objects, as in the Snapshot Model (Armstrong, 1988), the two-level temporal topological model (Raafat, et al., 1994), the ST-Composites (Langran and Chrisman, 1988) and the ST-Objects (Worboys, 1992) models.

But none of them have been able to portray concepts of transition, motion or process (Yuan, 1996). Others have tried to represent temporal information by events or processes, as in the event-based spatio-temporal data model (ESTDM) (Peuquet and Duan, 1995), the object-oriented geomorphological model (Oogeomorph) (Raper and Livingstone, 1995), the three domain model (Yuan, 1995), and the Triad model (Peuquet and Qian, 1996). The ESTDM and Oogeomorph models could represent changes of point-based data, but they were not designed to handle area data and topological relationships. The three-domain model and the Triad model defined three domains, semantics, time and space, but the links between them were not clearly defined.

In general, most of the existing models treat dynamic changes as differences between states, but none of them provides basic construct types of the underlying knowledge describing a process occurring in the real world. Although event representation and languages have been recently integrated in conceptual spatial models, they do not address the representation of spatio-temporal processes (Claramunt *et al.*, 1997). The processes involved in the development of natural phenomena, i.e. the actual characters of state transitions (such as shift, erode, expand, etc.) are not fully represented in these models. Moreover, the representation of complex changes between spatial objects has hardly been discussed in the literature to date. Most approaches cannot precisely or effectively model the interactions between natural phenomena that involve changes of geometric and thematic aspects of several objects at the same time, such as transition and mutation (e.g. merge and split). Therefore, dynamic interactions between objects are not really represented in these models.

To fully represent transitions, mutation, and movement of processes in space and time the behaviors of natural phenomena need to be considered prior to designing GIS data formats and data structures to represent temporal information (Egenhofer and Golledge, 1998). Semantic analysis of the characteristics and behaviors of natural phenomena is critical for determining a set of high-level spatio-temporal constructs (processes) to be modeled in a temporal GIS (Yuan, 1996).

However, the literature to date hardly discusses the dynamics of objects (particularly spatial changes) in a generic way. In fact, these dynamics are the basis for designing a proper data model to represent and to predict the dynamics of natural phenomena. Similarly, little research has been reported on the dynamic behavior of fuzzy objects with indeterminate boundaries. Furthermore, as a consequence of the fact that these objects are fuzzy, the detected change will also be uncertain. It is necessary to analyze uncertainty propagation in the change detection procedure to provide correct information to decision makers. To achieve these aims, the dynamics of fuzzy objects should be studied. Therefore, detection of the dynamics of fuzzy objects and measuring the uncertainty of change are the second point to be discussed in this thesis.

After discussion of the dynamics of fuzzy objects and determining the spatio-temporal constructs, a new spatio-temporal data model is required to support the representation of the dynamics of objects and their interactions. The description of spatio-temporal processes that act on the natural phenomena is emphasized with the aim of providing users with the means to understand, explain and predict the event in the world it represents. Therefore representation of fuzzy objects and their dynamics in a spatio-temporal data model is the third, and last, point to be investigated here.

1.2 Research questions

The main objective of this research is directly related to the issues discussed in the previous section. Given the complexity of natural phenomena and the limitations of conventional data models, the aim of the research was to develop a new data model for monitoring and representing the dynamics of the natural environment. In other words, the aim was to develop concepts for the identification of objects with indeterminate boundaries and their dynamics from time sequence data, and to represent them in a spatio-temporal data model, to support environmental studies. To achieve this aim, answers to the following research questions were sought:

1. How can procedures to identify objects and their dynamics be developed?
2. What types of uncertainties occur in source data and how are they propagated through such procedures? What is their influence on the mapping result, especially the geometric aspects (e.g. spatial extents and boundaries) of the objects?
3. How to derive and represent fuzzy objects, their spatial extent and their boundaries within different user contexts?
4. How to detect and describe the dynamics of these fuzzy objects?

5 How to represent the dynamics of fuzzy objects in a spatio-temporal data model? Which kind of queries should the model be able to handle to support environmental modeling?

1.3 Research approach

The approach I have taken to provide answers to my research questions is described below.

In a conventional GIS objects are generally considered to be crisp, with determined boundaries. Normally the boundaries of objects are identified first and then their spatial extent is mapped out, i.e. the spatial extent of objects is approached through the boundary. That's why most of existing research models and represents the uncertainty of object boundary, e.g. accuracy models for the coordinates of points of boundaries (Dunn, 1994). In the field-oriented approach the uncertainty is associated to locations (Burrough, 1998) but the concept of object is not available in this approach.

This thesis describes research on the uncertainties of the object, especially related to its spatial extent. Since there are no crisp boundaries between objects in the natural environment, spatial extents of objects are extracted from field data that change gradually and continuously over space. In such cases, the spatial extents of objects are fuzzy and can only be defined based upon certain criteria. Conditional boundaries can then be formed after identification of the fuzzy spatial extents of the objects. Therefore, the geometry uncertainty of object is not just a problem of geometry, but it is rather a problem of object definition or thematic vagueness. In such cases, uncertainty in thematic aspects is propagated to the geometric aspects of the objects.

In this research I have used the generalized data model and the formal data structure (FDS) (Molenaar, 1991, 1998a) to represent the uncertainties of objects. It differs from the cognitive linguistic approach (Couclelis, 1996; Smith and Varzi, 1997; Plewe, 1997; Couclelis and Gottsegen, 1997), which is far from ready for implementation in a computer system.

With respect to the first point listed for studying, the procedure to identify objects and their dynamics from field observation data was examined, since object representation of natural phenomena is usually derived from field representation. After examining this procedure, the causes of uncertainties and their propagation in the procedure were investigated. Special attention was given to the combination of errors and fuzziness in the classification. Then my research focused on the identification of fuzzy objects, in particular to identify their spatial extent and boundaries because of uncertain classification result. The fuzzy objects were derived based upon the application context. It made it clear where the uncertainties originated from and how they propagate to object description and change detection.

With respect to the second point listed, the dynamics of objects were studied, based upon the result of the first point. As spatial extent of an object at an epoch represents a state of the object, the dynamics (i.e. the state transition and processes) of the object can be determined by comparing its states at successive epochs. The dynamic interactions between objects will also be represent by processes such as splitting and merging.

To study the third, and last, point a spatio-temporal data model was designed based upon requirements for representing fuzzy objects and their dynamics. A process-oriented spatio-temporal data model was developed to support analysis and queries of time series data from varying perspectives, for example, through location-oriented, time-oriented, object-oriented and process-oriented queries, and to analyze the behavior of dynamic spatial complexes of natural phenomena.

1.4 Organization of the thesis

The thesis comprises nine consecutive chapters. Chapter 1, this chapter, introduces the research. Chapter 2 presents the case that will be used to illustrate the approaches in the thesis. The case-study area (Ameland, The Netherlands) is introduced, and collection and pre-processing of data are reported. The necessity for discussing fuzzy objects (in Chapter 3) and links between fuzziness and errors (in Chapter 4) are put forward.

Chapter 3 introduces basic principles of fuzzy objects. Issues regarding set and fuzzy set theory, object concepts in GIS, category theory, and approaches to object extraction are discussed. In addition, attempts at using fuzzy set theory to develop fuzzy object models are reviewed and assessed. A unified formal syntax schema to represent fuzzy objects is introduced. This chapter will help readers who are not familiar with fuzzy object concepts to understand the research findings presented later in the thesis.

As uncertainties of fuzziness and errors are treated differently in most existing approaches, Chapter 4 presents a formalized mathematical framework for describing uncertainties that combine these two aspects. The framework built is based upon the principles of dealing with fuzziness by fuzzy set theory and errors by probability theory. The relationship between error and fuzziness is discussed in terms of probability and possibility, and their combination is derived and described by analytical mathematical formula. The interaction between the two uncertainty aspects is demonstrated through the case. In addition, merits and drawbacks of the approach are discussed.

Chapter 5 presents an approach to identify the spatial extent and conditional boundary of fuzzy objects. It starts with the introduction of a procedure for identification of crisp objects from field observation data. Then, three fuzzy object models are introduced to represent fuzzy objects under different situations. Next, procedures to map the spatial extent and conditional boundary of the objects are discussed. They are formalized based upon the formal syntactic schema presented in Chapter 3. The propagation of uncertainties in the procedure of segmentation and merging are investigated. The approach is explained using case study as an example. Modeling results are described and discussed.

Chapter 6 develops an approach for detecting the dynamics of fuzzy objects. The concepts of state and process that are used to describe dynamics are introduced. Based upon these concepts and the assumption that natural phenomena change continuously and gradually, the approach to detect state transition of fuzzy objects is delineated and advocated. It is followed by an investigation of uncertainty of change detection. The case study is used to portray the approach. Advantages and potential limitations of the approach are discussed.

Chapter 7 describes the conceptual design of a process-oriented spatio-temporal model to support the representation of dynamic fuzzy objects. Firstly, it reviews the progress and deficiencies of existing spatio-temporal data models. Next, the dynamics of objects are characterized by formalizing the state and process of objects in mathematical equations. Then, the extended entity relationship (EER) model – the Star Model is presented by showing the classification of abstract object types and their relationships. The Star Model is modified into a process-oriented data model based upon the formalized description of dynamics of objects. Finally, the characteristics of the Star Model are discussed.

Chapter 8 presents the logical design and implementation of the process-oriented spatio-temporal data model. The logical design uses a relational approach for its relatively easy implementation, although the conceptual design in Chapter 7 is an object-oriented frame. An example based upon the case of Ameland is given to demonstrate the whole structure. The user-interface and multi-perspective queries based upon ArcView are presented. Performance of the model is reported.

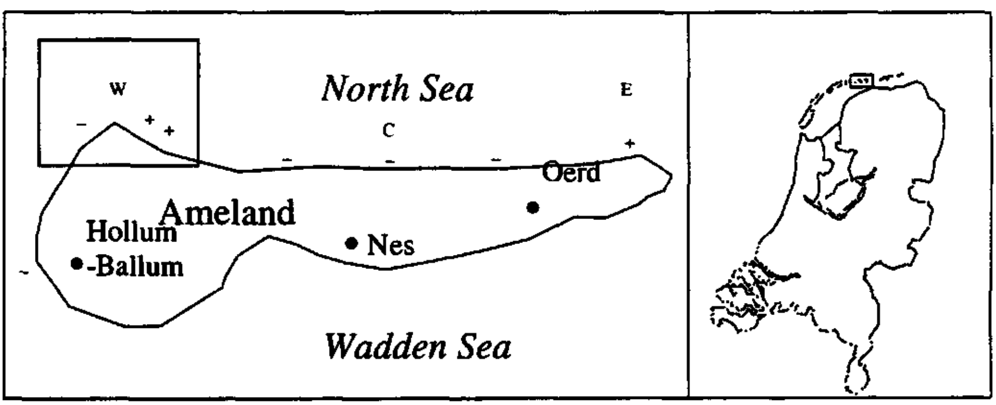
Finally, Chapter 9 concludes the thesis with a summary of research findings and contribution to the theory and practical application of modeling dynamic objects with fuzzy spatial extent. Specifically, the merits and drawbacks of the modeling approaches developed during this study are discussed. Furthermore, fruitful directions for future researches are suggested.

Chapter 2

The Case – Changing Beach of Ameland

2.1 General description of the study area

Ameland, one of the six Dutch barrier islands, was chosen as our test area, in order to demonstrate the practical usefulness of the developed methods and theories. It is situated north of the coast of Friesland, The Netherlands (see Figure 2.1). Its length is approximately 24 km, and its width varies from 1.5 to 4 km. It is made up of three dune complexes: the Hollum-Ballum complex in the west, the Nes-Buren complex in the center, and the Oerderduinen complex in the east. These three dune complexes were originally independent, being separated from each other by tidal inlets. Over the last two centuries sand dikes has been constructed to connect them and today they form one island today (Zuidam *et al.*, 1994).



(- coastal erosion, + accretion)

Figure 2.1 Test site – Ameland, The Netherlands.

2.2 The problem

The coastal zone is a transitional area between land and sea. Such areas are very important for living conditions, fishery, agriculture, tourism, etc. The growing concentration of population and socio-economic activities puts increasing pressure on coastal ecological systems, which at same time are threatened by inundation, coastal erosion, increased flooding, and loss of freshwater reserves and arable land, particular due to rising sea-levels. To sustain development and to minimize loss from possible natural disasters in these areas, it is necessary to guide and monitor the ongoing developments and their consequences.

Ameland is still developing. At certain locations in the middle and southern parts of the western end of the island, severe erosion occurs (due to shifting inlets by marine current), while at other places in the northwest accretion or accumulation occurs. To be able to predict future development, it is necessary to understand the various processes, their interaction and their effect on the development of the island. Such information is quite important for optimizing coastal defense works, e.g. beach nourishment or planting of grass, which both require high investments (Eleveld *et al.*, 1995).

The processes that influence the landscape units may be divided in two types: erosion and accumulation. The aim of the coastal geomorphology study described in this thesis was to understand know how these processes affect landscape units through changing their properties (volume and position). In other words, the problem was to investigate how to operate available data in order to understand current changes and to predict changes in future. The results of the processes can be measured qualitatively and quantitatively, i.e. the qualitative result can be identified by the erosion or accumulation of the landscapes, while the quantity can be measured by estimating (or calculating) the volume change.

For the main part, this study focused on sediment transport at the land– sea interface as a result of erosion and accumulation. Therefore, the most morphodynamically active area in the northwest portion (test window) of the island (Figure 2.1) was selected as the test site.

2.3 Existing approach to the problem

Traditionally, the effect of erosion and accretion is estimated using annual measurements in the form of coastal profiles (see Figure 2.4). Erosion and accumulation are identified by comparing the same profile for two different time horizons (see Figure 2.2). Changes in volume of sand sediments are calculated using software developed by the coast management agency of Rijkswaterstaat (RWS-MD*).

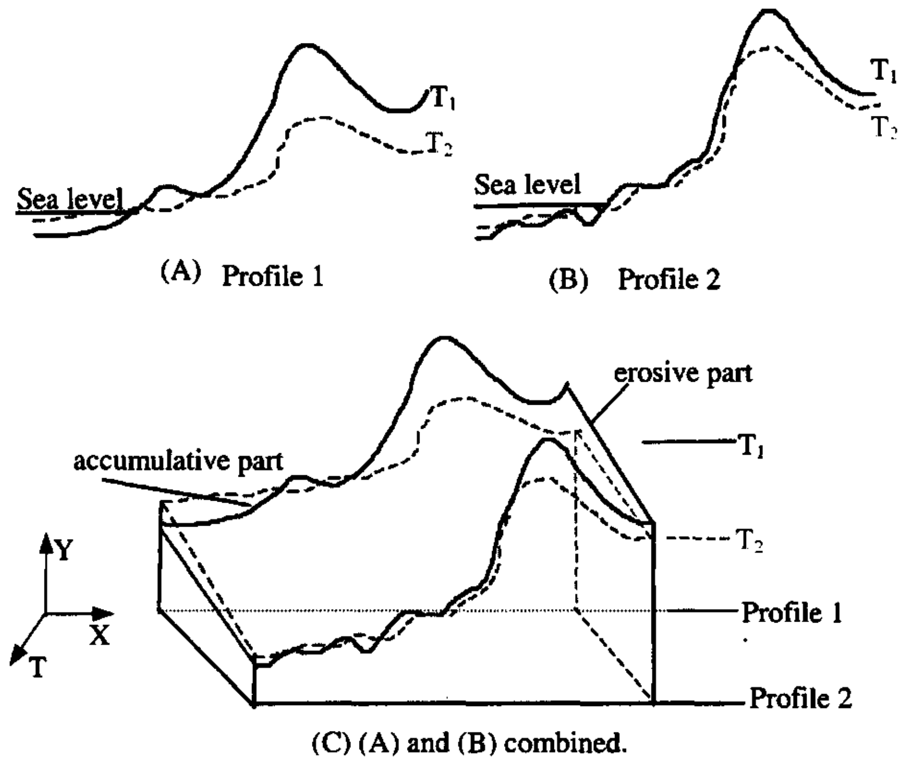


Figure 2.2 Erosion, accumulation and volume changes of a coastal beach zone for two time horizons.

From these calculations, inferences of the changes in profiles over the years can be obtained. Some geomorphologists, however, try to analyze the development trend of the landscape units. To do so, the geomorphologic processes, particularly the erosion and accumulation of sediments, should be distinguished through interpretation of changes in the landscape units, i.e. the foreshore, beach and foredune areas. Unfortunately, this information cannot be obtained as these landscape units have not been defined and recorded in the RWS-MD database. Therefore, to monitor these geomorphologic processes it is necessary to identify these landscape units and trace their changes by field observations.

*Survey Department of the Directorate-General for Public Works and Water Management of the Dutch Ministry of Transport, Public Works and Water Management.

The landscape units – foreshore, beach and foredune – have specific characteristics in terms of the following items:

- altitude;
- slope;
- roughness;
- size (small, intermediate, big, or large);
- material (grain size of sand);
- composition of mineral elements;
- compaction (loose or compact);
- humidity;
- vegetation/land cover.

The definitions of these landscape units usually differ from surveyor to surveyor, from case to case and from time to time. Among other ways, the landscape units may be defined based upon altitude of terrain surface according to different water lines. Heuvel and Hillen (1994) considered that the area beneath the high-tide line (HT) and above the low-tide line (LT) is foreshore; the area beneath the very high-tide (VHT) and above the HT is beach, and the area above the VHT and below the foot of dune is foredune. Others, however, consider that the foreshore is the area above the closure depth (Ruessink and Kroon, 1994) and beneath the low-water line (Graaff, 1977), that beach is the area above the low-water line and beneath dune foot (Reineck, 1980), and that the foredune is the first row of dunes inland from dune foot (see Table 2-1). Furthermore, the values for these water lines are not fixed. Ruessink and Kroon (1994) used –6 m to represent the closure depth in the years 1965 to 1984 and in year 1989, and used –8 m to represent the closure depth in the years 1985 to 1988 and 1990 to 1993. Ruig and Louisse (1991) used –6 m to represent it in all these years. Therefore, there is no invariable and fixed definition of the landscape units. Figure 2.3 illustrates one set of definitions of the landscape units.

Since the landscape units are defined based upon the altitude, their changes are mapped based upon profile measurements and by checking the shift of their position along the profiles. Insight into quantitative change in the sand sediment is gained by calculating trends in the yearly volumes of sand stored within the defined areas. An advantage of this method is that yearly fluctuation and small incidental errors of measurement are smoothed out (Ruig and Louisse, 1991).

Nowdays, Digital Elevation Models (DTMs) derived from profile measurements are used for geomorphology studies, and DTM data is also used in this study; the following section explains the data set and the process used to derive DTM from the profile measurement.

Table 2-1 Definitions of coastal landscape units.

Lanscape Unit	Definition (1)		Definition (2)	
	b_1	b_2	b_1	b_2
Foreshore	Closure depth	Low-water line	Sea bottom	Low-tide line
Beach	Low-water line	Dune foot	Low-tide line	High-tide line
Foredune	Dune foot	First row of dune area	High-tide line	First row of dune area

Note: b_1 and b_2 represent the lower and upper boundaries of the landscape units.

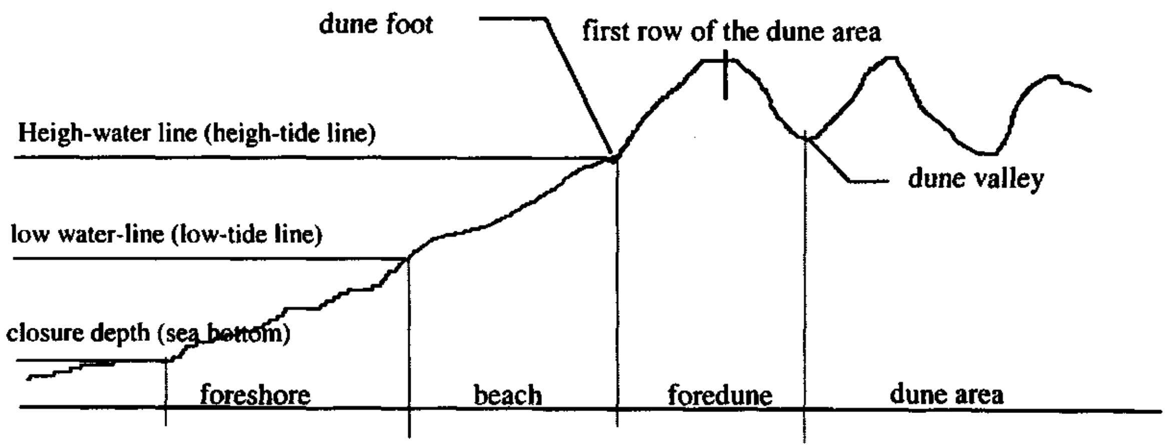


Figure 2.3 Landscape units are defined by the height of the depth of closure depth (–6 m), low water line (–1.1 m) and dune foot (+2 m).

2.4 Data

2.4.1 Coastline monitoring

Since the mid – 1800s, the location of the foot of dune, the high-water line and low-water line along the Dutch coast have been measured each year. These measurements are carried out along defined sections, each demarcated by beach posts. These posts are encountered on the beach along the entire North Sea coast, with distances of 200 – 250 m between each of them. Since 1963, the coastal profile has been measured every year in each section. This includes that the heights/depths are determined up to a distance of about 800 m to seaward of the posts, and up to some 200 m landwards of the first line of dunes. Once every three years the profiles are extended up to 2 to 3 km to seaward (Heuvel and Hillen, 1994). Figure 2.4 illustrates the positions of the profiles along the coastline of Ameland.

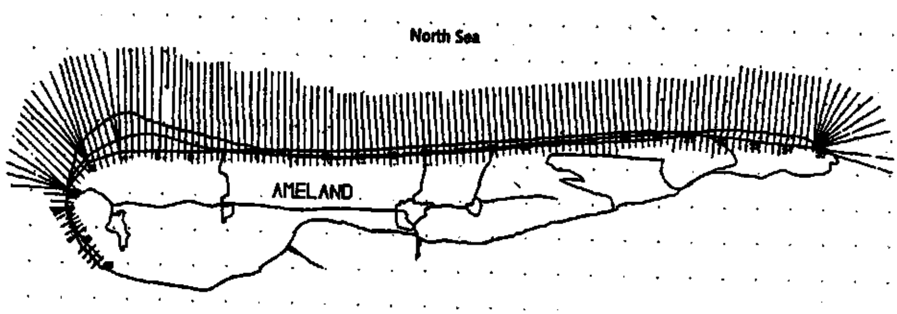


Figure 2.4 Profiles along the coastaline of Ameland (Heuvel and Hillen, 1994).

2.4.2 Data sets

As mentioned in Subsection 2.4.1, profile measurements comprise two types of data that have been brought together in one profile. The dry part of the coastal profile (beach and dune area) was originally measured by levelling, but from 1976 onwards aerial photograph and photogrammtry have been used. The underwarter part of the profile has always been measured with echosounders on ships with automatic position-finding systems. The heights and depths are uniquely related to a fixed Dutch Ordnance-Datum ('N.A.P.', Normal Amsterdam Level, which is approximately mean sea level) (Ruig and Louisse, 1991). Tidal corrections are made afterwards. The mean overall standard deviation of a single measurement of depth is about 20 cm (Nanninga, 1985; quoted in Ruig and Louisse (1991)). Since 1995, heights are also measured by laser scanning (laser-altimetry, Huising *et al.*, 1995). The annual coastal measurements are interpolated along the profile with 10 – 20 m intervals. The interpolated height values are stored in JARKUS database (Heuvel and Hillen, 1994). An overview of the data offered by RWS-MD is given in Table 2-2.

Table 2-2 Available data sets of annual height measurements.

Data type	Time span	Characteristics	Coverage
Sonar	Since 1993	Original measurements, sparsely distributed	Part
Laser-altimetry	Since 1995	Original measurements, sparsely distributed	Part
Photogrammetry	Since 1994	Original measurements, sparsely distributed	Part
JARKUS database	Since 1965	Interpolated data, along the profile	Complete

The amount and frequency of the JARKUS data is very useful for analyzing temporal (yearly) change in the landscape units. Therefore, data obtained between 1989 – 1995 were used for this study.

2.5 Data pre-processing*

As the JARKUS data set is stored in a unique format, conversion is needed to obtain the data into an exchangeable format such as (x, y, z). Since the data are distributed along the coastal profile, they are to be interpolated to obtain a complete coverage of the test site. This step is finished by using the function of *Surfer Access System* (Golden Software Inc.). Finally a height raster of $(60\,\mathrm{m} \times 60\,\mathrm{m})$ grids are derived. The procedure of data pre-processing is illustrated in Figure 2.5.

*Marieke Eleveld (ITC, Enschede, The Netheralnds) prepared the DTM data for the case study.

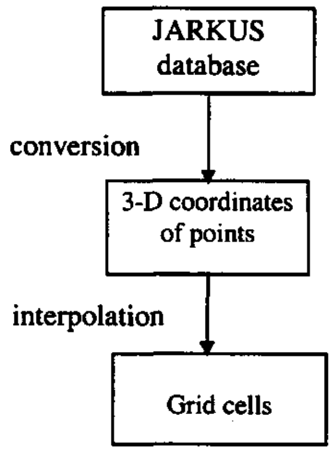


Figure 2.5 Flow diagram of data pre-processing.

Chapter 3

Basic Concepts of Fuzzy Objects

3.1 Introduction

Spatial objects that are represented in a conventional GIS are generally considered to be crisp, i.e. they have determined boundaries. For example, land parcels in cadastral systems are differentiated and identified by sharp boundaries. The basic assumption is that the classification of landscape units is crisp and spatial objects within these classes can be clearly determined. A second assumption is that objects are internally homogeneous and can be differentiated by crisp boundaries. Under the first assumption the threshold values or criteria for classification are sharply defined. Classes do not overlap, so that each object will be assigned to only one class. Under the second assumption the spatial extent of each object can be defined unambiguously and it will not contain unidentified inclusions of areas not belonging to the object. The determination of the spatial extent of geo-objects is then generally approached through the boundaries, or, more precisely, through the position of the boundary points. The analysis of the geometric uncertainty of the objects is therefore often based on accuracy models for the coordinates of these points (Dunn *et al.*, 1990).

These assumptions, however, are not valid when the spatial extents of objects are extracted from field data that change gradually and continuously over space so that no crisp boundaries can be identified. The boundary between the grassland and woodland may be gradual through a transition zone rather than a crisp boundary. In the case of the coast of Ameland (see Chapter 2), the boundary between beach and foreshore is also not crisp. In such cases, the geometric uncertainty of objects can not be expressed through the positional accuracy of boundary points, so a new strategy is needed to handle this kind of uncertainty.

Fuzzy set theory can handle uncertainty in class definition. By adapting fuzzy set theory in GISs, the uncertainty in object class definition and object description can be represented. This chapter will explain the basics of fuzzy set theory and its application in GIS. The concepts introduced here will also provide fundamentals of concepts to be developed in later chapters.

This chapter is organized as follows. The next section (3.2) introduces the basics of set theory. Then in Section 3.3 the principles of fuzzy set theory are discussed. Approaches to develop the fuzzy membership functions to handle uncertainty in object class definition are discussed in detail. Next, in Section 3.4, the concepts of objects are presented in terms of their definition and extraction. The concepts of fuzzy objects are introduced in the following three sections. Section 3.5 investigates the uncertainties in defining and extracting objects. Section 3.6 reviews approaches to handle these uncertainties. Section 3.7 discusses the representation of fuzzy objects. The state-of-the-art of research in this field is reviewed as well. To conclude, Section 3.8 summarizes the chapter and indicates the discussion to be continued in the following chapters.

3.2 Set theory

This section briefly reviews the basic concepts of set theory upon which the theoretical basis of fuzzy set theory is built.

3.2.1 Set

A set can be described either by listing all its members (the list method) or by specifying some well-defined properties satisfied by all members of the set (the predicate method). The list method, however, can only be used for finite sets. According to Klir and Folger (1988), a set A that has members of $a_1, a_2, \ldots, a_n$ can usually be written as

$$A = \{a_1, a_2, \ldots, a_n\},$$

and a set B whose members satisfy the properties $P_1, P_2, \ldots, P_n$ is usually written as

$$B = \{b \mid b \text{ has properties } P_1, P_2, \ldots, P_n\},$$

where the symbol | denotes the phrase 'such that'.

A simplest set can be defined based on one variable, which requires one (one boundary) or two parameters (the lower and upper boundaries) (see Table 3-1). For example, a set A with lower (b_1) and upper (b_2) boundary can be described as

$$A = \{z \mid b_1 \leq z \leq b_2\}$$

3.2.2 Boolean membership function

The fact that an individual observation z is a member of a set can be represented by a membership function (MF) which takes the value 1; it will take the value 0 otherwise. If an individual observation z is a member or an element of a set A, we write $z \in A$ ($MF = 1$); if not we write $z \notin A$ ($MF = 0$). Hence, we call it a *Boolean* membership and represent it as (MF^B). Regarding the properties of a set A, the membership function can be expressed as follows,

$$
\begin{aligned}
&MF^{B}(z)=1\text{, if } b_1 \le z \le b_2 \qquad (z \in A)\\
&MF^{B}(z)=0\text{, if } z < b_1 \text{ or } z > b_2 \qquad (z \notin A)
\end{aligned}
\tag{3-1}
$$

where b_1 and b_2 define the exact boundaries of set A.

3.2.3 Operations of set

Three operations of sets are used in this thesis. They are union, intersection and complement.

Union: the union of sets A and B is the set containing all the elements that either belong to set A alone, to set B alone, or to both set A and set B.

$$A \cup B = \{z \mid z \in A \text{ or } z \in B\}$$

Intersection: the intersection of sets A and B is the set containing all the elements belonging to both set A and set B.

$$A \cap B = \{z \mid z \in A \text{ and } z \in B\}$$

Complement: the complement is the set of elements that belong to the universal set Z, but do not belong to set A.

$$\overline{A} = \{z \mid z \in Z \text{ and } z \notin A\}$$

3.3 Fuzzy set theory

The theory and application of fuzzy sets has been well documented in the literature. A number of texts are available on this topic, to which readers are referred for more detail, e.g. Kandel (1986), Klir and Folger (1988), and Zimmerman (1984).

3.3.1 Fuzziness and fuzzy sets

Fuzziness is an admission of the possibility that an individual is a member of a set, or that a given statement is true. The assessment of the possibility can be based on subjective or 'expert' knowledge or preferences (see Subsection 3.3.2.1), but it can also be related to uncertainties that have a basis in probability theory (see Subsection 3.3.2.2).

Fuzzy sets are sets or classes that for various reasons cannot, or do not, have sharply defined boundaries, e.g. the 'class of all real numbers which are much greater than 1', or 'the class of beautiful women', or 'the class of tall men'. In such cases, for example, if a man is 1.7 m tall, one cannot say whether he belongs to the class of tall men or not.

If Z denotes a space of objects, then the fuzzy set A in Z is the set of ordered pairs

$$A = \{ (z, MF_A^F(z)) \},\ z \in Z \tag{3-2}$$

where the membership function $MF_A^F(z)$ represents the 'grade of membership of z in A' and $z \in Z$ means that z is contained in Z. Usually $MF_A^F(z)$ is a number in the range [0, 1], with 0 representing non-membership and 1 representing full membership of the set.

3.3.2 Fuzzy membership functions

The membership function of a fuzzy set defines how the grade of membership of z in A is determined. Contrary to a Boolean membership of a crisp set, partial membership is allowed in a fuzzy set. It means that class overlap is permitted and the membership function should be defined in a form to reveal the overlap or transition between the classes. Usually there are two ways to define the membership function, either on the basis of expert knowledge or by using methods of numerical taxonomy. Here we explain these two ways through the approaches most often used, i.e. the semantic import model and the fuzzy *C*-means.

3.3.2.1 semantic import model (SIM)

The semantic import model (*SIM*) is used when users have a more or less clear idea to group the data in a qualitative way, i.e. the central concept of the class is clear, but for various reasons the exact boundary can only be approximated. The fuzzy membership function is defined by adapting a crisp classification, e.g. extending the crisp boundaries into a transition zone. In general there are three simple types of transitions between fuzzy sets, analogous with the three types of crisp sets (Table 3-1). Therefore, fuzzy sets can be characterized by the boundaries (b_1 and b_2) plus the transition zones (d_1 and d_2). For mathematical description, the fuzzy membership function can be a linear, a curved, or an *S*-shaped function (Hootsmans, 1996).

Therefore, the *SIM* approach uses an a-priori-imposed membership function with which individual objects or attribute values can be assigned to a class or a set with a membership grade. It is no trivial task to define a proper membership function for different application cases. Firstly, a proper shape or type of transition has to be selected to represent the transitions between classes. Secondly, the parameters (e.g. the width of class kernel and the width of the transition zones) to describe the function have to be defined.

For example, Burrough (1989) used this approach for soil evaluation. A symmetric membership function, defined below, was chosen to distinguish 'deep' soil from 'shallow' and from 'very deep' soils.

$$MF_a(x) = (1 + a(x-c)^2)^{-1} \quad for\, 0 \le x \le 130$$

Ranges of 0-200 cm were considered to be the universal of soil depth, $c = 100$ cm was chosen as the ideal center and the lower and upper boundary were set as 50 cm and 150 cm, respectively, for the dispersion index $a = 0.004$. However, they used an asymmetric (left range) membership function to represent if the soil was deep enough for a given purpose – if it exceeded that depth by 5 cm or by 200 cm was immaterial. For 'sufficient soil depth' as 80 cm or more, the following membership function was applied

$$MF_a(x) = 1 \quad for\, 130 \ge x \ge c$$

$$MF_a(x) = (1 + a(x-c)^2)^{-1} \quad for\, 0 \le x < c$$

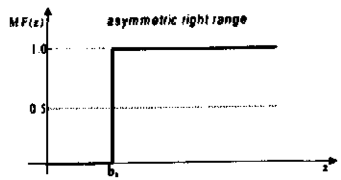
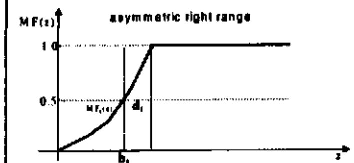
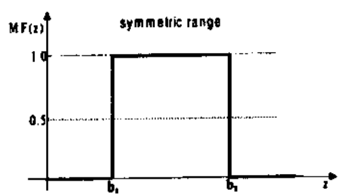
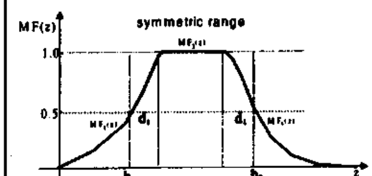
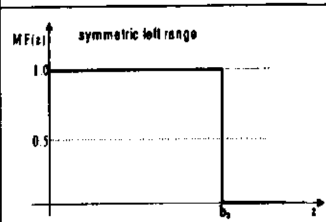
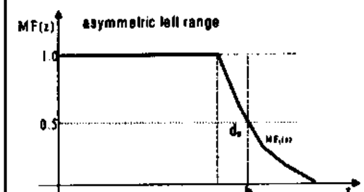
Here c was the left boundary and was set as 80 cm.

Other application of *SIM* in GISs can be widely found in literature, such as the definition of sharpness of boundaries in Wang and Hall (1996), and an air pollution danger zone around a city in Usery (1996). I have adapted this approach for this case study (see Subsection 4.5.1).

3.3.2.2 fuzzy C-means (FCM)

Opposite to the subjective approach of *SIM*, the fuzzy c-means (*FCM*) approach tends to be an objective approach (Burrough and McDonnel, 1998). It is analogous to cluster analysis and numerical taxonomy in that the value of the membership function is computed from a set of attribute data. In such a way, an individual sample may have memberships of multi-classes (illustrated by Figure 3.1).

Table 3-1 Crisp and fuzzy sets.

Type	Crisp		Fuzzy	
	Graph	Function	Graph	Function*
Asymmetric right range		$MF^B(z)=1$ $(z \geq b_1)$ $MF^B(z)=0$ $(z < b_1)$		$MF^F(z)=MF_1^F(z)$ $(z < b_1 + d_1)$ $MF^F(z)=1$ $(z > b_1 + d_1)$
Symmetric range		$MF^B(z)=1$ $(b_1 \leq z \leq b_2)$ $MF^B(z)=0$ $(z < b_1 \text{ or } z > b_2)$		$MF^F(z)=MF_1^F(z)$ $(z < b_1 + d_1)$ $MF^F(z)=1$ $(b_1 + d_1 \leq z \leq b_2 - d_2)$ $MF^F(z)=MF_3^F(z)$ $(z > b_2 - d_2)$
Asymmetric left range		$MF^B(z)=1$ $(z \leq b_2)$ $MF^B(z)=0$ $(z > b_2)$		$MF^F(z)=1$ $(z \leq b_2 - d_2)$ $MF^F(z)=MF_3^F(z)$ $(z > b_2 - d_2)$

* where $MF^F(z)$ is the value of the continuous membership function corresponding to the attribute value z. If parameters d_1 and d_2 are zero, equation yields the Boolean membership function.

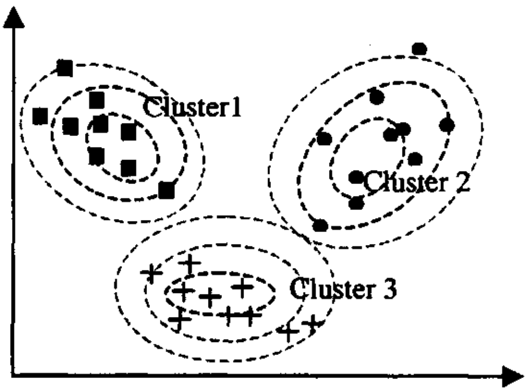


Figure 3.1 Fuzzy C-mean approach (Hootsmans, 1996).

The most commonly used algorithm for computing the membership values is (Bezdek, 1981):

$$MF_{ic} = \left[(d_{ic})^2\right]^{-1/(q-1)} / \sum_{c'=1}^{k} \left[(d_{ic'})^2\right]^{-1/(q-1)} \tag{3-3}$$

in which MF_{ic} expresses the possibility (membership value) of the ith object belonging to the cth cluster, with d the distance measure used for similarity, and the fuzzy exponent q determining the amount of fuzziness. After assigning an object to a class, the class mean is recomputed for the similarity index. The membership of an individual with respect to a class in multivariate space is computed and it is assigned to the class for which the d_{ic} is minimal. The procedure is repeated interactively for each individual until the class stabilizes.

FCM approach was originally designed for unsupervised fuzzy image classification (Bezdek *et al.*, 1984). But it is difficult to determine *c*, the number of clusters, at the initial stage of interactions. Secondly, every pixel contributes to all clusters to some degree, so that the entire data set must be used for pixel similarity derivation. Wang (1990, 1992) proposed to use the posterior probability in a maximum likelihood classifier (*MLC*) as the fuzzy membership. There are two advantages in using *MLC* instead of *FCM*. First, the choice of the number of classes is supported by the analyst's knowledge of the image under a supervised setting. Second, *MLC* takes into account the geometric shape of class hyper-ellipsoids in computing pixel membership. It is therefore deemed to yield the most accurate classification provided there is a near-normal multivariate distribution of pixel values in the classes (Jin and Jensen, 1996).

However, the accuracy level is dictated by a training stage, where individual class signatures are still constructed and evaluated in a platonic manner, i.e. a pixel makes a total contribution to the forming of its enclosing class signature and zero contribution to that of other signatures. In order to overcome the uncertainty in training samples (error or mixed pixels), Ji and Jensen (1996) took a fuzzy training approach in the supervised classification. Instead of supplying a prior the fuzzy membership to the training class statistics, these values were generated as part of the process through an interactive evaluation of sample pixels via *FCM*. A fuzzy parameter estimator was developed with a modified Bezdek's *FCM* engine in order to improve the classification performance. In addition, Inomata and Ugata (1992) proposed to use histograms as membership function for both supervised and unsupervised fuzzy classification.

To conclude, fuzzy function values can be computed through an imposed 'expert' model (*SIM* approach), or by a data driven multivariate procedure (*FCM* approach). In both cases the methods allow class overlap.

3.3.3 Operations of fuzzy set

The three basic operations on fuzzy sets are similar to those that can be used for crisp (or Boolean) ones.

Union: The union of two fuzzy sets *A* and *B* with individual membership functions $MF_a(x)$ and $MF_b(x)$ results in a new fuzzy set *C*, whose membership function can be defined as $MF_c(x) = \max\{MF_a(x), MF_b(x)\}$. The membership function of a fuzzy set *C* that is the union of several fuzzy sets X_1, X_2, X_3, ..., X_n can be defined as

$$MF_c = \max_i(MF_{x_i})$$

where MF_{x_i} is the membership function of fuzzy set X_i (i = 1, ..., n).

Intersection: The intersection of two fuzzy sets *A* and *B* with individual membership functions $MF_a(x)$ and $MF_b(x)$ results in a new fuzzy set *C* and the membership function of *C* can be defined as $MF_c(x) = \min\{MF_a(x), MF_b(x)\}$. For the intersection of several fuzzy sets X_1, X_2, ..., X_n, the following equation applies:

$$MF_c = \min_i(MF_{x_i})$$

where MF_c is the membership function of the fuzzy set C resulting from the union; MF_{x_i} is the membership function of fuzzy set X_i (i = 1, ..., n).

Complement: The complement of a fuzzy set A results in a new fuzzy set C, whose membership function is defined as:

$$MF_c = 1 - MF_a(x)$$

For other operations of fuzzy sets the readers should refer to (Zadeh, 1965).

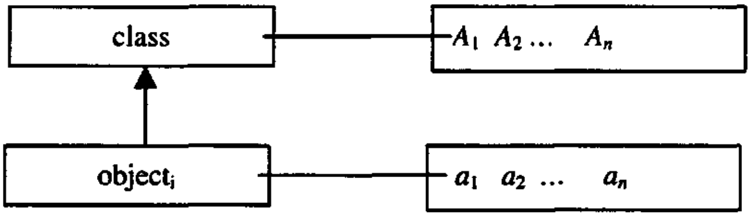
3.4 Concepts of objects in GIS

3.4.1 Definition of objects

As argued in Chapter 1, objects in GISs are abstract descriptions of entities in reality. The abstraction is based upon the principle that the entities possess common attributes, common motor movements, and similar shape and functions (Rosch, 1978, quoted in (Usery, 1996)). Different levels of abstraction form different level of categories, which could make up a classification schema. For example, road, railroad, and canal are basic level geographic objects in a superordinate category of transportation with subordinate elements state road, interstate highway, standard gauge railroad, narrow gauge railroad, shipping canal, and irrigation canal (Usery, 1996). These roads all have functionality of transportation and they all have the attribute of name, segments, destinations, etc. These common attributes define the description structure of that class of objects. In the relational database model, this would mean that a table, containing columns of these attributes, could be defined for each object class. For example, for a building class, it contains information about owner, name, construction time, etc. The value of the attribute differs from object to object. The relationship between objects, class and attribute is illustrated by Figure 3.2.

For study of coastal geomorphology, 'foreshore', 'beach' and 'foredune' can be considered as superordinate categories, while the special forms related to these landscape units (such as bar on foreshore, swash bar, ripple and little dunes on beach, and blowout and washovers on foredunes) are lower level categories.

In addition to the attribute description, the object needs geometric data to represent its location, size and shape (where it is and how big it is). For example, geometric descriptors of a building would be location and area. The complete description of an object in GISs then consists of the thematic component expressed by attribute values and a geometric component either in raster or vector format. These two components are linked through a unique identifier of the object (see Figure 3.3). Two objects are distinguished if their descriptions are not equal.



a_i is the value of attribute A_i for object i (i = 1,...,n)

Figure 3.2 Relationship between object, class and attributes (Molenaar, 1998a).

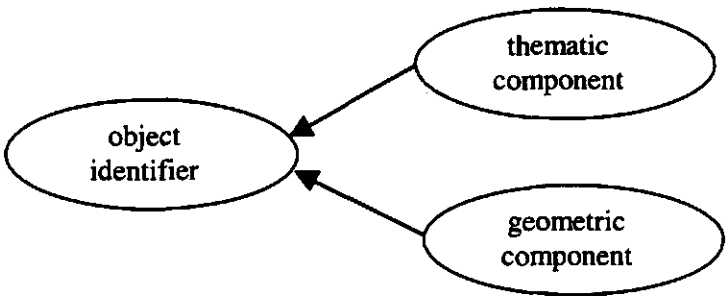


Figure 3.3 Descriptions of objects are linked through the object identifier.

3.4.2 Extraction of objects

Objects are abstractions representing entities in reality. To get entities in reality to objects in GISs, i.e. from raw observed data to the objects represented in GISs, several series of data transformation are required. There are several ways of observing reality and to extract objects from the observation data. Here four ways of object extraction are presented.

(1) land surveying

Land surveying appears to be the most direct approach of obtaining objects. All it requires is the surveying of the location of boundaries of the objects (illustrated by Figure 3.4). In such cases, the coordinates of the boundary points are recorded. However, the surveyor actually performs two tasks in one step: i.e. identification of the boundary of the objects, and the measurement of the boundary location.

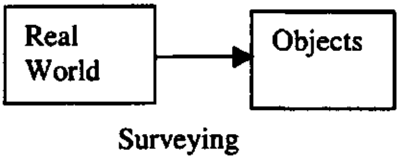


Figure 3.4 Extraction of objects through land surveying.

(2) Image Interpretation

In image interpretation, objects are obtained in three steps (see Figure 3.6). In the first step reality is captured by a 'snapshot' image. In the second step the boundaries of objects are determined through interpretation. In the third step the boundaries of objects are digitized to derive the geometry of the object.

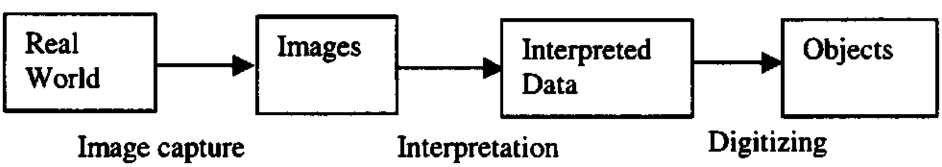


Figure 3.5 Extraction of objects through image interpretation.

(3) Image classification

In image classification, objects can be obtained from entities in reality through three steps (illustrated by Figure 3.4). The first is image capture, meaning reality is captured by a 'snapshot' image. The second step is image classification, i. e. from image data to thematic class data. There are many ways to classify images, ranging from supervised to unsupervised, from single band to multiple bands, and from single to multiple temporal image classification. The third step is image segmentation, to identify a region consisting of pixels that belong to one class. Each region is then assumed to represent the geometry of an object.

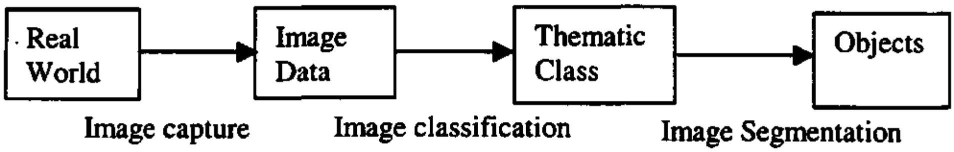


Figure 3.6 Extraction of objects through image classification (after Gahegan and Ehlers, 1997, p.66).

(4) Field Sampling

Four major steps are required to derive objects from reality by means of field sampling (Figure 3.5, Cheng *et al.*, 1997). The real world is seen as a field with attributes that have position-dependent values. In the first step these values are sampled at specific points. The second step is the interpolation of sample values to derive data for the whole area. Either a triangulated irregular network (TIN), or contours or a raster cover of the whole area is created. The third step is categorization, i.e. to assign the cells to different classes of object types. The thematic data is transformed into thematic class data. The last step is object formation, which is similar to image segmentation in the image classification approach.

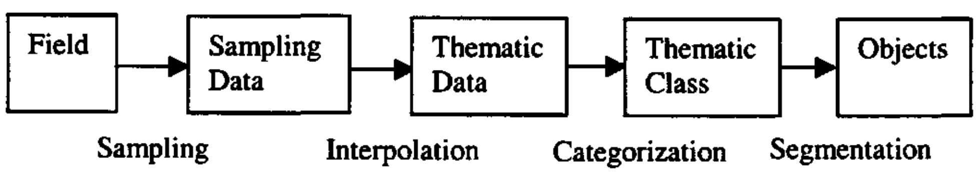


Figure 3.7 Extraction of objects through field sampling.

The procedures of 'image classification' and 'field sampling' are pixel-oriented (field-oriented) approaches. 'Image interpretation' and 'land surveying' are boundary-oriented (object-oriented) approaches. These two approaches of object formation result in two ways of handling uncertainties, which will be discussed in Section 3.8.

3.4.3 Syntactic representation of objects

The syntactic representation of objects proposed by Molenaar (1994, 1998a) is based upon a formal data structure (FDS) (Molenaar, 1991). I will not discuss the complete syntax here, but rather summarize the representation of boundary and spatial extent of objects, because these cases are directly related to this study.

Assume that the geometry of a spatial database has the structure of a planar graph. Each edge will always have one face on its left-hand side and one on its right-hand side. The relationship can be expressed by the functions (Molenaar, 1998a):

Edge e_i has face f_a on its left-hand side => $Le[e_i, f_a] = 1$, otherwise = 0
Edge e_i has face f_a on its right-hand side => $\mathrm{R}i[e_i, f_a] = 1$, otherwise = 0.

With these functions, we can define another function:

$$B[e_i, f_a] = Le[e_i, f_a] + \mathrm{R}i[e_i, f_a].$$

If the value of this function is equal to 0, e_i is not related to f_a; if the value is equal to 2, then the edge has the face on both sides and it is thus inside the face; if the value is equal to 1, then the edge has the face on one side only, so that it must be part of the boundary. The boundary of a face f_a is then defined as:

$$\partial f_a = \{N_a, E_a\} \tag{3-4}$$

where $E_a = \{e_i \mid B[e_i, f_a] = 1\}$ means that all the edges have the face f_a on one side only; and $N_a = \{n_i \mid n_i \in e_i, e_i \in E_a\}$ represents the nodes of the edges of E_a.

The relationship between a face and an area object can be represented by $Part[f_a, O_a]$. If it takes a value equal to 1, this implies that the face belongs to the object. If it takes a value equal to 0, this implies that the face doesn't belong to the object.

The relationship between an edge and an object can then be expressed as

$Le[e_i, O_a] = \mathrm{Le}[e_i, f_a] * \mathrm{Part}[f_a, O_a]$
$\mathrm{R}i[e_i, O_a] = \mathrm{Ri}[e_i, f_a] * \mathrm{Part}[f_a, O_a]$.

Similarly, a function can be defined as

$$B[e_i, O_a] = Le[e_i, O_a] + \mathrm{R}i[e_i, O_a]. \tag{3-5}$$

If the value of this function is equal to 0, the edge is not related to the object; if the value is equal to 2, then the edge is inside the object; if the value is equal to 1, the edge is part of the boundary of the objects that is

$$\partial O_a = \{e \mid B[e, O_a] = 1\}. \tag{3-6}$$

This syntactic representation of the relationship between edges, boundaries, faces and objects can be applied in both vector and raster structures. In the vector geometry the face represents a polygon and in the raster geometry it represents a raster cell (Molenaar, 1994, 1998a, see also Figure 3.8).

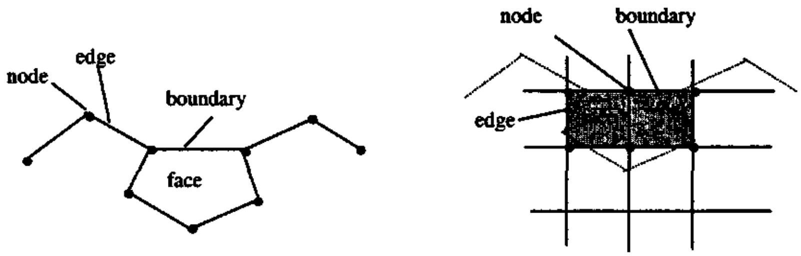


Figure 3.8 The relationship between face, boundary, edge and node (after Molenaar, 1998a).

3.5 Uncertainties in object extraction

3.5.1 Uncertainties in category theory

The definition of objects in GISs is related to category theory. Therefore, inherent uncertainty of category implies that there is uncertainty in the description of objects. There are four aspects will cause uncertainty in the category definition:

- fuzziness

A category is usually considered as a set of instances in which have a high inter-category correlation of appearance, attribute, function, and so on, and a high intra-category variation in those features. Although they have been assumed historically to be clearly definable (thus crisp thresholds have been used to formalize them), linguists and psychologists have shown that internal meanings associated with common categories and words (including those with geographical content) are very vague. The term *fuzzy category* has been used to refer to these cases (Plewe, 1997; Gray, 1997). For example, the description of a bar on the foreshore is something like it has an elongated shape and it is distributed continuously in one direction; the greatest difference between its highest point to its lowest point is greater than or equal to 50 cm (Rwessink and Kroon, 1994).

According to the prototype theory (Rosch (1978), quoted in Plewe (1997, p.126)), some of the instances belonging to a category will have a much higher correlation with characteristics common to the category than others; these 'central' instance are called prototypes. To determine whether an object is a member of a category usually the local condition is compared with the prototype in a gestalt, rather than analytical, way. If they are 'sufficiently' close, the object can be considered to be in the category. Obviously this approach produces categories that tend to be fuzzy in threshold, depending on the personal judgement of the observer. For example, an area of mixed trees and brush could be considered either a member of the 'forest' category or a member of the 'shrubland' category, but not as strongly as unmixed areas.

- multiple criteria

Sometimes categories used to delineate geographic entities are distinguished using more than one criterion, which may not coincide with each other (Plewe, 1997). For example, some consider foreshore to be the area above the low-tide and below high-tide marks, yet others define foreshore to be the area above closure depth and below the low-water line.

- spatially incomplete definitions

Sometimes, definitions have meaning only for parts of fields to which they have been applied. Thus, some areas in the field are 'undefined'. The definition is inapplicability – the field is expected to cover a space that is not really present at some point (Plewe, 1997). For example, the definition of coastal geomorphologic landscape units on Ameland may not always be appropriate for other sites.

- time incoherence

In many cases, definitions only have meaning for a specific period, i.e. the values or relevance of information depends on time. In the case of vegetation, there has been considerable discussion about its coherence over time (Gray, 1997, p.143). For the study described in Chapter 2, the definition of closure depth is –6 m for the years 1965 to 1984, and in 1989, yet it is set at –8 m in the years1990 to 1993.

For these reasons definition of category might be uncertain. If one of the reasons is valid, the object class should be defined as a fuzzy category.

3.5.2 Uncertainties in measurements

Uncertainty in measurement refers to the errors due to the limitations of instruments and human ability. In capturing an image of the real world, data values may be influenced by measurement error and quantization effects that are determined by sensor design and operation, or a change in spatial location may be forced by grid processes or relief distortion (Gahen and Elhers, 1997). Almost always, we are only able to measure a small part of the object of interest. This causes sampling error. Indeed, there are always measurement errors in surveying. The digital coordinates may differ from the location of entities in the real world due to digitizing error. All these uncertainties will influence the accuracy and precision of the object extracted. Sampling error and measurement error will be discussed in detail in Section 4.2.

3.6 Handling uncertainties in object extraction

Two kinds of uncertainty models, namely fuzzy models and stochastic models, are relevant for the further development in this thesis.

3.6.1 Fuzzy models

Fuzzy models use fuzzy membership functions to describe vagueness in class definition. As discussed in Section 3.2, I assume the fact that a class is fuzzy if there are no crisp criteria to decide whether an object belongs to the class or not. This situation will occur when grid cells are to be assigned to height classes related to foreshore, beach and foredune.

3.6.2 Stochastic models

As argued in Chapter 1, stochastic models describe errors in measurement. This might refer to measurement for the evaluation of thematic attributes, but it could also refer to the determination of coordinates of boundary points of objects. Since the determination of the spatial extent of objects is generally approached through boundaries or, more precisely, through the position of boundary points. The analysis of the geometric uncertainty of the objects is therefore often based on accuracy models for the coordinates of these points. The epsilon band method is usually used to represent the locational accuracy of the boundary (Dunn *et al.*, 1990; Chrisman, 1991; Shi, 1994).

3.6.3 Handling uncertainties in object extraction

An explanation follows of how the two kinds of uncertainties mentioned above are propagated through the process of object extraction and how these uncertainties can be handled.

(1) Land surveying

In land surveying, boundary identification and measurement are generally combined into one process, but it still leaves the fact that these are two operations. The uncertainty about geometric position of a boundary is often expressed by means of epsilon bands (Chrisman, 1991; Shi, 1994). But the thematic uncertainty in such case is usually ignored.

(2) Image interpretation

In image interpretation, decisions about the spatial extent of the object are often separated from the measurement of the geometry of the object boundary, which are often digitized after the interpretation has been made. When objects are fuzzy, the location of the repeatedly interpreted boundary will be uncertain (Middlekoop, 1990; Ewards, 1994) since errors occur in digitizing. The positional uncertainty in manually digitized map data is discussed (Bolstad *et al.*, 1990). We will discuss uncertainty in interpretation in detail in Section 3.8 in order to derive a unified syntactic representation of fuzzy objects.

(3) Remote sensing image classification

Gahen and Elhers (1997) discussed the uncertainties in the procedure of image classification. They used four models to represent data in four different forms: Field Model, Image Model, Thematic Model and Object Model (Gahegan, 1996). A data set in each model may be described by a number of abstract properties, i.e. data value, spatial extent, temporal extent, and lineage information, and uncertainties can be associated with these aspects (measurement and labeling errors). The uncertainty associated with the mapping result is the uncertainty that already presents in one data form plus the uncertainty associated with the data transformation. For example, through image capture, a data value may change due to measurement error or quantization effects, a change in spatial location may be forced by grid processes or relief distortion (Gahen and Elhers, 1997). Uncertainty of images is related to these aspects. Therefore, the uncertainty of objects is comprises uncertainty in image plus uncertainty associated with the transformation of classification and segmentation, if objects are extracted by image classification (cf. Subsection 3.4.2).

The work of Gahen and Elhers (1997) provides a general conceptual framework with which to understand and trace uncertainty in the procedure for extracting objects by image classification, where it originates from and what affect it has. However, the actual effects of transformation have not been analytically calculated. So the practical application is not available. Furthermore, their model does not describe the inter-relationships between the various types of uncertainty.

(4) Field sampling

Measurement and sampling errors occur in the data obtained by field sampling (cf. Section 4.1), and these errors propagate to the interpolated thematic data. There is vagueness in the class definition of objects. These two kinds of uncertainties combine in the categorization step, leading to uncertain classification, which propagates to objects as fuzzy spatial extent and conditional boundaries of objects through segmentation. As this approach has been used to extract the landscape units on Ameland, I develop a method to deal with the uncertainties of this procedure in detail in Chapters 4 and 5.

3.7 Syntactic representation of fuzzy objects

This section discusses the fuzzy aspects of objects and then describes the syntactic representation of fuzzy objects.

3.7.1 Fuzzy aspects of objects

The basic structure of the description of a fuzzy object should be similar to that of a conventional, crisp object. The uncertainties corresponding to the descriptions should be represented in terms of three aspects, i.e. the assignment of an object to the object classes, the assignment of attribute values to an object, and the assignment of a spatial description to an object (Molennar, 1998a).

- Fuzzy object classes

For classification, if an object belongs to an class then $MF[O,C] = 1$, otherwise $MF[O,C] = 0$. If the classification does not give definite (crisp) results then the assignment of the object to class C will be uncertain, which will be expressed by a fuzzy class membership function that will take a value between 0 and 1, such as

$$CLASS(O) = \{C \mid M[O,C] > 0\}$$

- Fuzzy attribute values

The domain of the attributes may contain values or value classes that are not clearly defined. In other cases the value of the domain may be clearly defined but that there is no sufficiently reliable measuring procedure for attaining values with high enough accuracy. The evaluation of an attribute A for an object O should then be expresses by a pair $A[O]=(a,m_u)$, where a is the estimated attribute value and m_u is the measure of the uncertainty that a is a correct value.

- Fuzzy geometry

The geometric description consists of the topology of the spatial objects, their shape and their position. There are two ways to handle uncertainty in geometry of fuzzy objects, corresponding to two approaches of object extraction. This will be explained in Subsection 3.7.2.

3.7.2 Pixel-oriented and boundary-oriented approaches

In the pixel-oriented approach, analysis starts from cell- (or pixel) based information, as for example in remote sensing classification. In remote sensing applications the spectral classification of an image results in a class label and a likelihood value per pixel (Lillesand and Kiefer, 1994). Under the maximum likelihood criterion, each pixel will be assigned to the class C_k for which the likelihood function has a maximum value. A region with cover type C_k will then be a segment that consists of mutually adjacent pixels that have been assigned to that class. This region represents the spatial extent of an object. As the pixels do not belong to a class with the likelihood value of 1, the spatial extent of the object will be uncertain (Molenaar, 1998b).

In the boundary-oriented approach, analysis of uncertainties is directly linked to boundaries of objects, i.e. the locations of the boundaries are fuzzy and the fuzziness of the objects is then a consequence of the fuzzy boundaries. The example of image interpretation may illustrate this. Image interpretation generally results in the delineation of the boundaries of spatial objects. In many applications, such as vegetation or land use mapping, these objects are fuzzy, so that repeated interpretation lead to varying object boundaries. Middlekoop (1990) and Edwards and Lowell (1996) analyzed the uncertainty of such interpretation through an overlay of several repetition results. The spatial uncertainties of the objects are evaluated by overlaying the multiple interpretation. The spatial variability of the boundaries of coinciding polygons can then be evaluated in a geometric sense (Edwards 1994; Edwards and Lowell, 1996).

3.7.3 An unified syntactic representation of fuzzy objects

The syntactic approach described in Section 3.4.3 demonstrates that vector geometry and raster geometry each have similarly expressive power. This implies that the handling of spatial uncertainty should in principle also be the same for both geometric structures. Therefore, it must be possible to combine or even unify pixel-oriented and boundary-oriented approaches (Molenaar, 1998b). We named such an approach as a unified approach.

As discussed above, in the pixel-oriented approach the classification of an image results in a class label and a likelihood value $L[P_{ij}, C_k]$ per pixel, which can be used to denote the uncertainty of each pixel belonging to the cover class. If each pixel is assigned to a class for which the likelihood value is maximum, the spatial extent of objects will be segmented by clustering adjacent pixels belonging to the same classes. The uncertainty of the relationship between a pixel and the object is due to the function $L[P, C]$. Thus, if $P_{ij} \in O_a$, then $Part[P_{ij}, O_a] = L[P_{ij}, C_k]$.

In the boundary-oriented approach, Edwards and Lowell (1996) overlaid the polygons that resulted from repeated interpretation. The location of an average boundary was estimated for each pair of adjacent regions. Middelkoop (1990) proposed evaluating the uncertainty in photo interpretation by rastering the interpretation results. He evaluated the functions *Part*[], relating the cells of the overlay to the identified regions through frequency count. A similar approach can be applied to the vector structure of the boundaries (Molenaar, 1998b). The repeated interpretations are overlaid and for each face of this overlay the frequency can be counted of its assignment to each of the objects. If these numbers are divided by the number of interpretations, we can obtain relative frequencies that indicate how certain we are that a face belongs to some particular object.

Therefore, through function *Part*[] the pixel-oriented approach and the boundary-oriented approach is unified. The identification of the object boundary is uncertain because the spatial extent of the object is uncertain in the sense that the value of the function $Part[f, O]$ (cf. Subsection 3.4.3), relating the face to the object, not just 0 or 1, but varies between 0 and 1.

If the object is fuzzy in the sense that its spatial extent is uncertain, the relationship of the face and the object is uncertain, which can be expressed by assigning a fuzzy face set to an object (Molenaar, 1998a)

$$Face(O_a) = \{f_j \mid Part[f_j, O_a] > 0\} \tag{3-7}$$

For raster geometry, the cell size should be chosen so that the value of this function per cell is homogeneous. In the vector geometry the faces should be defined so that the function is homogeneous per face.

A boundary function relating an edge to a fuzzy object can be defined according to Equation 3-5. But the value of the function will vary from 0 to 2. If the value is equal to 0, the edge is not related to the object; if the value is greater than 0 and less than 2, the edge is an indeterminate boundary of the objects; if the value is equal to 2, the edge is inside the object. Therefore, the indeterminate boundary of an object is defined as (Cheng *et al.*, 1998)

$$B\partial O_n = \{e \mid 0 < B[e, O_n] < 2\} \quad (3\text{-}8)^*$$

Based upon certain criteria, the conditional spatial extent of objects can be identified, so the indeterminate boundary becomes a conditional boundary (to be discussed further in Subsection 5.4.3).

3.8 Summary and discussion

This chapter has introduced the basic concepts of sets, fuzzy sets and fuzzy objects. For fuzzy objects, the uncertainty aspects in procedures of object extraction were explained, and the approaches to handle these uncertainties were reviewed. A unified syntactic representation of fuzzy objects was presented.

Discussions found in the literature generally emphasize one kind of uncertainty of objects, either in data measurement or in class definition. The stochastic model is often used to represent uncertainty in measurement of geometric locations and thematic attributes; fuzzy models are mainly used to describe uncertainty in class definition.

However, as shown in Subsection 3.7.2, the uncertainty of geometry is not only a problem of boundary location, but is also related to vagueness of object definition. For example, vagueness in class definition will result in fuzzy spatial extent and indeterminate boundary of object in image interpretation. Therefore, fuzzy models should be integrated into the description of uncertainty in geometric aspects. Furthermore, errors in measurement data might result in misclassification or in 'wrong' shape of objects. For example, an incorrect data value for a pixel may lead to mis-classification, and position errors in image data may result in an object being assigned the 'wrong' shape, i.e. wrong extent. Therefore, the stochastic model should be combined with the fuzzy model (to be discussed in Chapter 4).

---

* The concept of a crisp boundary of an object is only valid when the face belongs to the object (*Part*[*f*, *O*] = 1) and the edge has the face on either the left or right side. Therefore, Equation 3-8 is valid only for fuzzy objects.

The procedure for identifying fuzzy objects from field sampling data has not been thoroughly investigated and formalized yet. In Chapter 5 I describe my investigation of the procedure and the tracking of the propagation of uncertainties. Furthermore, existing approaches do not describe the inter-relationships between various fuzzy aspects of objects (e.g. the effect of uncertainties in thematic aspects on geometric aspects). This aspect is also discussed in the Chapter 5.

Chapter 4

Combination of Errors and Fuzziness*

4.1 Introduction

Uncertainty refers to our incomplete and inexact knowledge of the world, which has a direct effect on the predictions we want to make in environmental studies. Any serious evaluation and prediction should give estimates of uncertainties in the results.

As explained in Chapter 3, we can distinguish two types of uncertainty. The first results from errors in measurement, and the second is related to fuzziness in class definition. Although these two types of uncertainty are often discussed separately, in this chapter they will be discussed in combination. Section 4.2 introduces data uncertainty. Section 4.3 discusses the similarity and differences between errors and fuzziness in terms of probability and possibility. After giving a clear definition of these two kinds of uncertainties, Section 4.4 explains how they can be combined mathematically. Section 4.5 and Section 4.6 present the case study. Finally, the major findings are summarized in Section 4.7.

4.2 Data uncertainty

Data uncertainty means that the true value of a parameter is unknown (or unknowable). It has to do with our observations of nature or society: we are unsure of what exactly we are observing or measuring. Two sources of data uncertainty are sampling and measurement errors.

- Sampling error

Many geographical studies are concerned with the spatial distribution of a variable, such as altitude or rainfall, which varies continuously from place to place. For a large area it would be impossible to measure the value of this

* This chapter is based upon Cheng *et al.* (1997).

variable at every point. Almost always, we are only able to measure a small part of the object of interest. We must make (sometimes unjustified or un-testable) assumptions about our sampling strategy. This causes sampling error. The type and magnitude of these errors can be determined by repeated sub-sampling or by more exhaustive sampling. An example of sampling error is that we can only measure the height of sampling points along a coastal profile, not all the points in the test area.

- Measurement error

When the sampling is repeated for a variable at the same location, the observations are different. For example, repeated measurement of a distance in land surveying usually gives different results. Sometimes some variables should satisfy a relationship theoretically, but the actual observations do not give agreement with the rules. For example, the total value of three angles of a planar triangle should be equal to 180^0, but in practice it is not always true. Another example of measurement error is measuring the height of the sampling points along the coastal profiles: the observation value of a point will be different if we repeat the measurement several times.

The difference between repeated observations and between an observation and its theoretical value is due to errors associated with observations. These errors are inevitable due to the limits of human ability, instruments and the circumstances under which observation takes place. These errors are usually random. They can be determined from the characteristics of the device and by repeated sampling. They are stochastic, therefore one may apply probability theory (e.g. normal distribution) to describe these 'random errors', in terms of mathematical expectation and variance (Hogg and Craig, 1970).

4.3 Probability vs. possibility

Errors and fuzziness are different both conceptually and theoretically. Error is associated with data uncertainty and can be described by probability theory. Fuzziness is associated with class definition and can be handled by possibility (fuzzy set) theory. Probability expresses the likelihood that a certain event will occur. Possibility describes event ambiguity. It measures the degree to which an event occurs. In simple words, whether an event occurs is 'random'. To what degree it occurs is 'fuzzy'. For example, If you ask the question "Is there an apple in the refrigerator?" the answer will be related to probability. Now, suppose you know that there is a half-eaten apple in the fridge. If you ask, "To what degree is there an apple in the refrigerator?" then this is a question with a fuzzy answer, even if on this occasion it is still 0.5 (Openshaw and Openshaw, 1997). More detailed discussion about probability and possibility can be found in Zimmerman (1984).

In reality, errors and fuzziness often occur simultaneously and should be dealt with together. As discussed in Chapter 2, the definition of landscape units is generally not crisp, therefore different surveyors may judge a real world situation differently, from case to case and from time to time, when mapping it. In fact, all the unit definitions are vague as no crisp boundary can be identified to differentiate the landscape units. If an area has a height value -7.0 m, it is difficult to decide whether it should be classified as offshore or foreshore. In practice, it may be 50% offshore and 50% foreshore, indicating that these landscape units, which form a continuum in space, cannot be defined in a traditional, crisp way. For that reason, the concept of fuzzy objects was introduced in Chapter 3. Furthermore, errors in profile measurements should not be ignored in the subsequent analysis. Therefore, I discuss the combination of errors and fuzziness in classification.

4.4 Uncertainties combining errors and fuzziness

Uncertain reasoning models usually include three components (Molenaar, 1998a): (1) uncertainty of the data value, (2) uncertainty of primitive evidence and (3) uncertainty of the class definition or intention. To elaborate these three uncertainty aspects, let us take an example of classification, such as "if $b_1 \le z \le b_2$ then z belongs to Class A". The three uncertainty aspects may be:

(1) uncertainty of the data is the uncertainty of the observation of z, which can be represented by standard deviation of measurement.
(2) uncertainty of the class refers to the definition of Class A in terms of observation z not being clear, as the landscape units are not crisp. In this case, the crisp boundary (b_1, b_2) of the class may be converted into a transition zone and use a fuzzy membership function to describe the uncertainty in the rule (cf. section 3.3.2).
(3) uncertainty of evidence refers to the uncertainty of evidence used for inference. For example, when we use a remote sensing image to identify the distribution of vegetation, we usually don't have enough evidence to assign a grid cell to either a grassland or to a maize crop if we only use the spectral information. In this study, however, there is no evidence uncertainty and therefore it will not be discussed here.

Therefore, the uncertainty of the classification should be approached from uncertainty in data and uncertainty in class definition. If there is no uncertainty in the observation, the uncertainty of the classification will be the uncertainty of the class. But if there are uncertainties in the observation and in the class definition, the calculation of uncertainty of classification is a relatively complicated procedure.

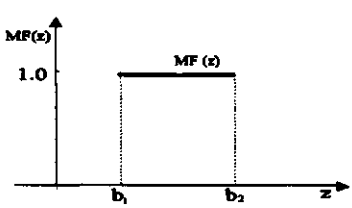
Heuvelink and Burrough (1993) discussed error propagation in cartographic modeling using crisp and fuzzy classifications. They illustrated several situations that might arise when a thematic value containing error is placed in a class defined by boundaries. When the distribution of observation data is well within the class boundaries, the classification is certain, but when the observation distribution straddles the boundaries, the classification result is uncertain. They used the Monte Carlo method to obtain the possible class and corresponding membership value for these uncertain cases. However, the influence of error on the classification result cannot be calibrated in an explicitly analytical form in the Monte Carlo method. I analyzed the error propagation in an analytical approach by generalizing the possible situations of classification into the four cases indicated in Table 4-1.

Table 4-1 Combination of errors and fuzziness.

observation / classification	*Error-free*	*Containing errors*
Crisp	(1) Boolean operation	(3) Boolean operation with probability
Fuzzy	(2) Fuzzy operation	(4) Fuzzy operation with probability

Case 1: the definition of the class is certain (crisp) and the value of observation is error-free.

Assuming a class is defined by crisp boundaries represented as (b_1, b_2), the classification result only depends on the value (z) of the observation. If the value is in the interval of (b_1, b_2), then the observation belongs to the class with full membership = 1. Otherwise, it does not belong to the class or has membership = 0. This case can be represented by a Boolean membership function



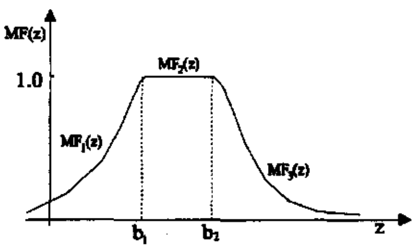
$$MF(z)=\begin{cases}0 & z<b_1\\ 1 & b_1\le z\le b_2\\ 0 & z>b_2\end{cases} \tag{4-1}$$

where $MF(z)$ is the value of the membership function for z.

Figure 4.1 Combination of errors and fuzziness – Case 1.

Case 2: the definition of the class is fuzzy, the value of observation is error-free.

In contrast with crisp class boundaries, fuzzy class boundaries are to be applied and the Boolean membership function is replaced by a fuzzy membership function, as shown in Figure 4.2. In general, I assume that the membership function can be represented by three compound membership functions as follows,



$$MF(z)=\begin{cases} MF_1(z) & (z<b_1) \\ MF_2(z) & (b_1\le z\le b_2) \\ MF_3(z) & (z>b_2) \end{cases} \qquad (4\text{-}2)$$

Figure 4.2 Combination of errors and fuzziness – Case 2.

where $MF(z)$ is the value of the fuzzy membership function corresponding to the observation value z, $MF_i(z)$ represents different membership functions depending on the value of z. Normally, $MF_2(z) = 1$.

Case 3: the definition of the class is certain, the value of observation contains random error.

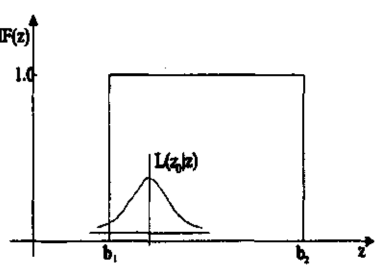


Figure 4.3 Combination of errors and fuzziness – Case 3.

The problem now is to evaluate the value of the membership function that a point (P_i) belongs to a crisp class (h_c) when it has an observed value z that contains random error. It is assumed that z has a normal distribution $f(z|z_0)$ with mean z_0 and a standard deviation σ. The likelihood that the real height of P_i is z_0 can be represented as $L(z_0|z)$.

The membership function $MF(z)$ can be calculated by the convolution of the membership function $MF(z_0)$ (defined Equation 4-1) and the likelihood function $L(z_0|z)$,

$$\begin{aligned} MF(z) &= \int_{-\infty}^{+\infty} MF(z_0)\cdot L(z-z_0\,|\,z)\cdot dz_0 \\ &= 0\cdot\int_{-\infty}^{b_1} L(z-z_0)\cdot dz_0 + 1\cdot\int_{b_1}^{b_2} L(z-z_0)\cdot dz_0 + 0\cdot\int_{b_2}^{+\infty} L(z-z_0)\cdot dz_0 \\ &= L(b_1 \le z_0 \le b_2\,|\,z\} = L(z_0 < b_2\,|\,z) - L(z_0 < b_1\,|\,z) \end{aligned} \qquad (4\text{-}3)$$

The certainty of the classification of z_0 is determined by the likelihood of z_0 falling in the interval (b_1, b_2). It implies that the probability is transferred into the fuzziness of classification.

Case 4: the definition of the class is fuzzy, the value of the observation containing error.

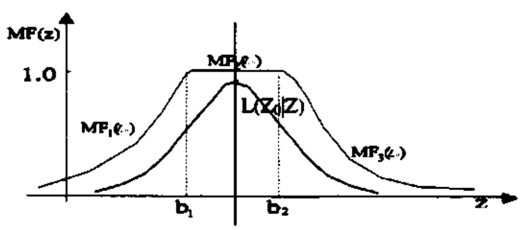


Figure 4.4 Combination of errors and fuzziness – Case 4.

This occurs when fuzziness (Case 2) and errors (Case 3) are combined. However, we can treat this as a general case of Case 3 by replacing its crisp classification by the fuzzy classification consisting of three sub-parts (defined in Equation 4-2). So the value of the membership function $MF(z)$ can be evaluated as

$$\begin{aligned} MF(z) &= \int_{-\infty}^{b_1} MF_1(z_0)\cdot L(z-z_0)dz_0 + \int_{b_1}^{b_2} MF_2(z_0)\cdot L(z-z_0)dz_0 \\ &\quad + \int_{b_2}^{+\infty} MF_3(z_0)\cdot L(z-z_0)dz_0 \\ &= P_1[MF_1(z_0)] + P_2[MF_2(z_0)] + P_3[MF_3(z_0)] \end{aligned} \qquad (4\text{-}4)$$

$MF(z)$ is the value of the membership function of the classification corresponding to the observation value z, and $P_i[MF_i(z_0)]$ is the probability value for each sub-part of the membership function defined in Equation 4-2.

Equation 4-4 is a general representation of the membership function value for cases of fuzziness combining probabilities. It can be adapted to other situations in which the zones of $(-\infty, b_1), (b_2, +\infty)$ are further divided into small intervals with different membership functions (see, for example, Section 4.5, Case study).

4.5 Case study

(1) fuzzy membership function for height classes

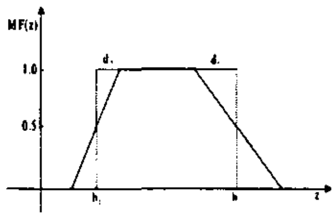
We use the fuzzy set approach to represent vagueness in the definition of the landscape units discussed in Chapter 2. A fuzzy membership function is used instead of crisp classification criteria. According to (Burrough, 1989), the membership function should ensure that the degree of the membership is 1 at the center of the set and that it falls off in an appropriate way through the fuzzy boundaries to the regions outside the set where it takes the value 0. The point where the membership value is 0.5 is called the cross-over point. I have chosen the boundaries b_1 and b_2 (Table 2-1) to be the cross-over points and have defined the transition zones d_1 and d_2. For example, the height range of the transition of foreshore to beach is 2 m, as some experts set –6 m to be the closure depth while others set it at –8 m. Therefore, the crisp definitions for coastal landscape units in Table 2-1 have been modified into the fuzzy definitions in Table 4-2.

Table 4-2 Fuzzy definition for coastal landscape units.

Class Code	Landscape Unit	b_1 (m)	b_2 (m)	d_1 (m)	d_2 (m)
1	Foreshore	-6.0	-1.1	2.0	0.5
2	Beach	-1.1	2.0	0.5	0.5
3	Foredune	2.0	25.0	0.5	3.0

Note: b_1 and b_2 represent the cross-over points of the landscape units; d_1 and d_2 represent the half-width of transition zones.

A trapezoidal membership function (Figure 4.5) is defined to represent the fuzzy definition of the landscape units, where z is the height value of the grid cell.



$$MF(z) = \begin{cases} 0 \quad (z < b_1 - d_1) \\ \dfrac{1}{2d_1}(z + d_1 - b_1)\,(b_1 - d_1 \le z \le b_1 + d_1) \\ 1 \quad (b_1 + d_1 \le z \le b_2 - d_2) \\ -\dfrac{1}{2d_2}(z - d_2 - b_2)(b_2 - d_2 < z < b_2 + d_2) \\ 0 \quad (z > b_2 + d_2) \end{cases} \tag{4-5}$$

Figure 4.5 Fuzzy membership function of classification of landscape units.

(2) errors in the interpolated height values

The accuracy of the interpolated height values of each grid cell depends on two aspects. One is the observation error of the sample points, which is related to many factors of measurement, such as instruments and atmospheric conditions. The other aspect is error of interpolation, depending on the validity of the interpolation functions. Functions used for height interpolation assume a certain degree of smoothness of the surface. Actual deviation of the real topographic surface from the interpolated surface is considered as random. The interpolation accuracy is determined from the difference between the interpolated values and the values obtained with high accuracy (which can be considered as a field-check method). This implies that data of higher accuracy are considered as the mean of the interpolated values. Assuming that the height observations are independent and normally distributed, the variance can be calculated by summing the squares of the difference between the interpolated value and the mean. This approach has been adopted by RWS-MD to access the accuracy of their digital elevation model (Huising *et al.*, 1996). As I use the same kind of sample data and the same interpolation method, I assume their experimental result ($\sigma = 0.15\,\text{m}$) represents the uncertainty of the height value of each grid cell.

(3) fuzzy classification of stochastic height values

As the height value of the grid cells is stochastic, the fuzzy membership value of a grid cell to each class may be calculated according to Equation 4-4 by adopting the fuzzy membership function as defined in Equation 4-5. Equation 4-6 is evaluated for the interpolated height value z of each grid cell with respect to the different types of landscape units.

$$MR(z)=\int_{-\infty}^{b_1-d_1}0\cdot L(z-z_0)\cdot dz_0+\int_{b_1-d_1}^{b_1+d_1}\frac{1}{2d_1}(z+d_1-b_1)\cdot L(z-z_0)\cdot dz_0+\int_{b_1+d_1}^{b_2-d_2}1\cdot L(z-z_0)\cdot dz_0$$
$$+\int_{b_2-d_2}^{b_2+d_2}-\frac{1}{2d_2}(z-d_2-b_2)\cdot L(z-z_0)\cdot dz_0+\int_{b_2+d_2}^{+\infty}0\cdot L(z-z_0)\cdot dz_0 \tag{4-6}$$
$$=\frac{1}{2d_1}\int_{b_1-d_1}^{b_1+d_1}z\cdot L(z-z_0)dz_0-\frac{1}{2d_2}\int_{b_2-d_2}^{b_2+d_2}z\cdot L(z-z_0)dz_0+\frac{d_1-b_1}{2d_1}\int_{b_1-d_1}^{b_1+d_1}L(z-z_0)dz_0$$
$$+\int_{b_1+d_1}^{b_2-d_2}L(z-z_0)dz+\frac{b_2+d_2}{2d_2}\int_{b_2-d_2}^{b_2+d_2}L(z-z_0)dz_0$$

4.6 Results of the case study

4.6.1 Error influence on crisp classification

The crisp object definition defined in Chapter 2 (see Table 2-1) is used to discuss the effect of errors. The effect of the stochastic observations was evaluated for $\sigma = 0.15$ m (Series 1), 0.30 m (Series 2), 0.45 m (Series 3), 0.60 m (Series 4) and 2.0 m (Series 5), respectively. The membership values for each grid cell belonging to the height classes were calculated using Equation 4-3. Figure 4.6 shows that the membership value of each grid cell to each class decreases with increase in σ (where Series 1 to Series 5 represents these five error levels). There is a slight difference between the first four cases; for $\sigma = 2.0$ m, however, these differences become significantly larger.

The identification of fuzzy objects based upon these classification results will be discussed in Chapter 5. Here we present mapping results based upon FC-object model (cf. Section 5.3). In the first four cases, the objects identified had the same spatial extents and boundaries. But when $\sigma = 2.0$ m, 65 grid cells changed their class type and the extents of the objects identified were different from the first four series. On checking the height values of these changed grid cells, it was noted that these values are around the class boundaries (-1.1 m and 2.0 m). The maximum membership values of these grid cells were less than 0.5, implying that when standard deviation grows too large (2.0 m), the objects mapped out will be different. Figure 4.8 presents the modeling results for $\sigma = 0.15$ m and $\sigma = 2.0$ m; the results of other three series were very similar to those of $\sigma = 0.15$ m.

4.6.2 Error influence on fuzzy classification

By assigning σ to be 0.15 m, 0.30 m, 0.45 m, 0.60 m and 2.0 m and calculating, respectively, fuzzy membership values for each grid cell according to Equation 4-6, similar conclusions as those for crisp classification were reached, i.e. there is no obvious difference between the first four cases (Figure 4.7), but there are obvious differences between these four cases and that of $\sigma = 2.0\,\text{m}$. Indeed, 59 grid cells changed their classes (Figure 4.9E).

4.6.3 Discussion

In the case studied there is no obvious difference between the influences of error on crisp and on fuzzy classifications. When the observation data values are close to those of class boundaries, and when the standard deviation is too large (e.g. $\sigma = 2.0\,\text{m}$), the class type of grid cells may change significantly.

From the experiment, I conclude that thematic (classification) uncertainty will influence the uncertainty of the spatial extent in the sense that the membership value of cells representing the spatial extent of the landscape units is lower when the accuracy of thematic data (interpolated height values) decreases.

4.7 Summary

This chapter discusses the combined effect of uncertainties in data and in classification criteria on classification results. Mathematical expression have been explained to evaluate the uncertainty of classification results for four different situations:

1. error free observation data and crisp classes
2. error free observation data and fuzzy classes
3. stochastic observation data and crisp classes
4. stochastic observation data and fuzzy classes

The differences between Situation 3 and 4 were illustrated by describing a case. This reveals that the stochastic errors had a similar influence on both crisp and fuzzy classification. It is also revealed that thematic uncertainty will definitely influence the uncertainty of the spatial extent of the landscape units. Further analysis of the relationship between thematic aspects with geometric aspect is discussed in Chapter 5.

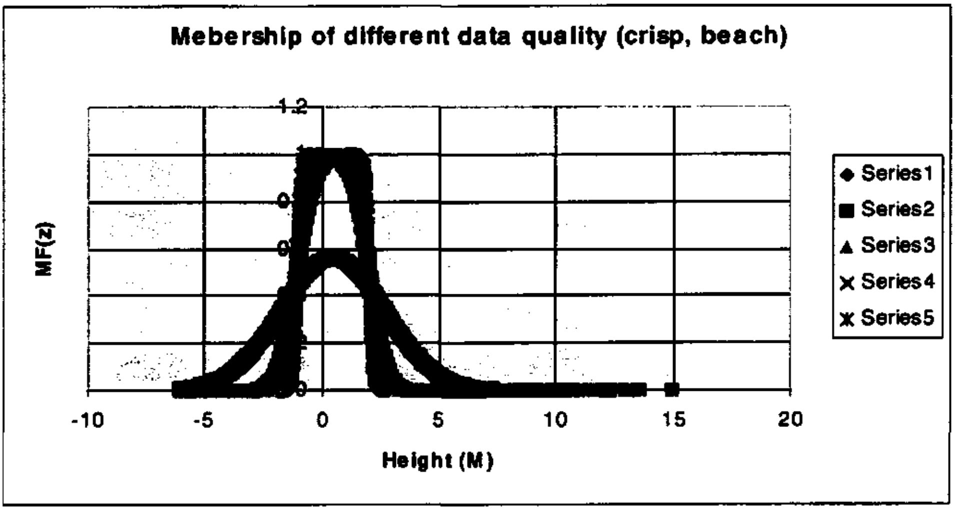


Figure 4.6 Influence of errors on crisp classification.
(Series 1 $\sigma = 0.15$ m, Series 2 $\sigma = 0.30$ m, Series 3 $\sigma = 0.45$ m, Series 4 $\sigma = 0.60$ m, Series 5 $\sigma = 2.0$ m)

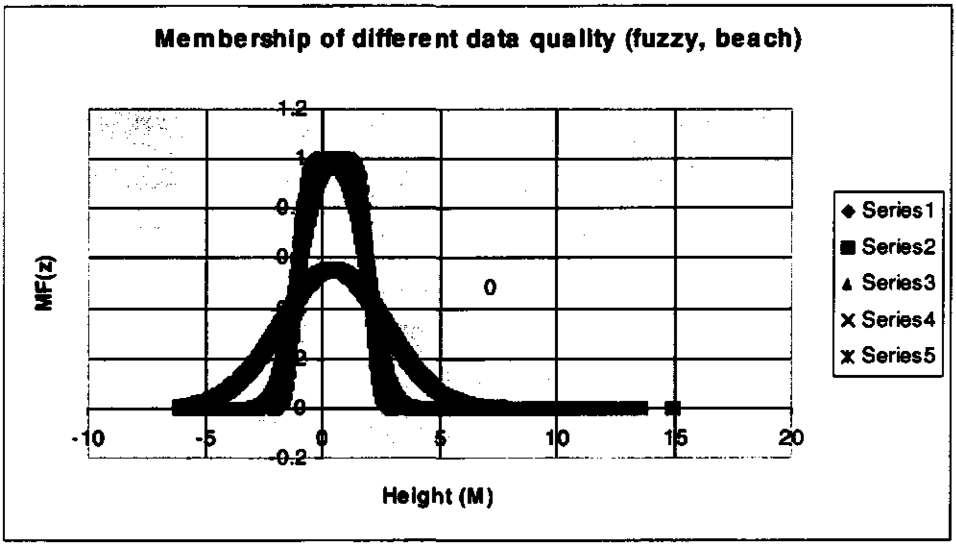


Figure 4.7 Influence of errors on fuzzy classification.
(Series 1 $\sigma = 0.15$ m, Series 2 $\sigma = 0.30$ m, Series 3 $\sigma = 0.45$ m, Series 4 $\sigma = 0.60$ m, Series 5 $\sigma = 2.0$ m)

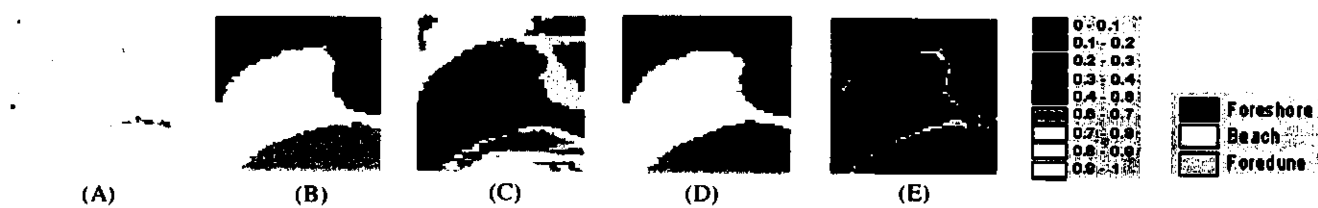


Figure 4.8 Crisp classification results.
(A) Maximum membership function value and (B) class type ($\sigma = 0.15$ m)
(C) Maximum membership function value and (D) class type ($\sigma = 2.0$ m)
(E) Difference between (B) and (D) (black means no change)

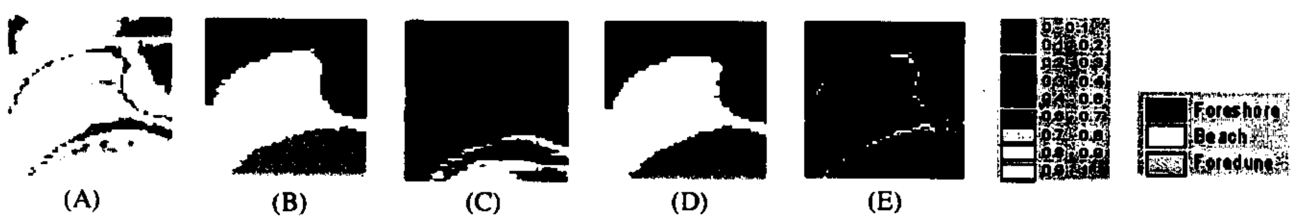


Figure 4.9 Fuzzy classification results.
(A) Maximum membership function value and (B) class type ($\sigma = 0.15$ m)
(c) Maximum membership function value and (D) class type ($\sigma = 2.0$ m)

Chapter 5

Identification of Spatial Extent of Fuzzy Objects*

5.1 Introduction

This chapter discusses the procedure for identifying fuzzy objects from the fuzzy classification results. It firstly investigates the procedure for extracting objects from field observation data, for which three fuzzy object models are proposed. The procedure to identify the spatial extents of objects in different models is discussed in detail. The propagation of uncertainties in the identification procedure is investigated thoroughly in order to have a complete understanding of the transformation of uncertainties from field model to object model. Finally, the last section summarizes the major findings in the chapter.

5.2 A procedure for extracting objects from field observation data

As mentioned in Subsection 3.4.2, there are four approaches to extract objects from field observational data. Here I use the second approach of field sampling, in which field characteristics are sampled at specific points. These sparsely distributed sample points are then interpolated to generate data for the whole area sampled. In this study, the height values of points on coastal profiles had been measured and an elevation raster covering the whole test area was created after interpolation. Then the raster was classified and segmented to yield regions that represent the spatial extent of objects. Three types of objects were then extracted from these height data: foreshore, beach and foredune areas, each related to a height interval. Therefore, in general, objects were extracted from field observation data in a six-step procedure:

1. **Sampling** data values at specific sample points.

2. **Interpolation** of the observed data to generate a complete raster covering the observed area.

* This chapter is based upon Cheng *et al.* (1998).

3. **Classification** of all grid cells into pre-defined classes. Each grid cell is assigned to a class interval that can be related to one of the natural units.

4. **Segmentation** of the classified raster into regions. Each contiguous set of grid cells belonging to one class will form a region that represents the spatial extent of a particular natural unit.

5. **Merging** regions that are smaller than a pre-defined lower threshold for mapping units with an adjacent region. Traditional merging methods, such as "window filtering", "nibbling", "dropping the longest shared boundary", and "maximum area merging" (Ma and Zhao, 1995) can be used.

6. **Identification** of objects represented by the regions, i.e. to identify the actual objects represented by the regions (after merging).

In the case study, Steps 1 to 5 result in regions of the height classes. The identification of the actual objects represented by these regions needs further analysis. For example, a region of a sand bar on a foreshore may have the same height as a beach region. After classification, this region may be initially identified as beach. If such a region is fully contained in a foreshore area, however, it should be identified as a sand bar instead of as beach. Therefore, it is necessary to analyze the spatial topological relationships between the regions in order to identify the objects whose spatial extents are represented by these regions.

If the dynamics of natural phenomena are to be monitored, time series data will be used. In such a case, the regions extracted at one epoch represent a snapshot (a state) of the dynamic objects, and one region at that epoch represents a particular state of one object. In this chapter attention is paid to the situation of a particular epoch; the dynamic situation is discussed in Chapter 6.

The procedure to extract objects from field observation data is illustrated in Figure 5.1, by a crisp example. The steps presented in this figure start with the interpolated grid cells after Step 2. In Step 3 the grid cells are classified into three elevation classes: 'S' (foreshore) ranging from –6 m to –1.1m, 'B' (beach) ranging from –1.1 m to 2 m and 'D' (foredune) ranging from 2 m to 23 m. The segmentation of Step 4 identifies three regions. In Step 5, Region 3 has been merged into Region 2 because it was initially identified as a foreshore area, but it appeared to be fully contained in a beach area. So finally two objects – 'S' (foreshore area) and 'B' (beach area) are identified in Step 6.

In the procedure, data is converted from a low-level semantic form (field sampling) to a high-level semantic form (distinct objects), mainly through interpolation, classification, segmentation and merging. Uncertainty propagation in the interpolation and classification has been discussed in

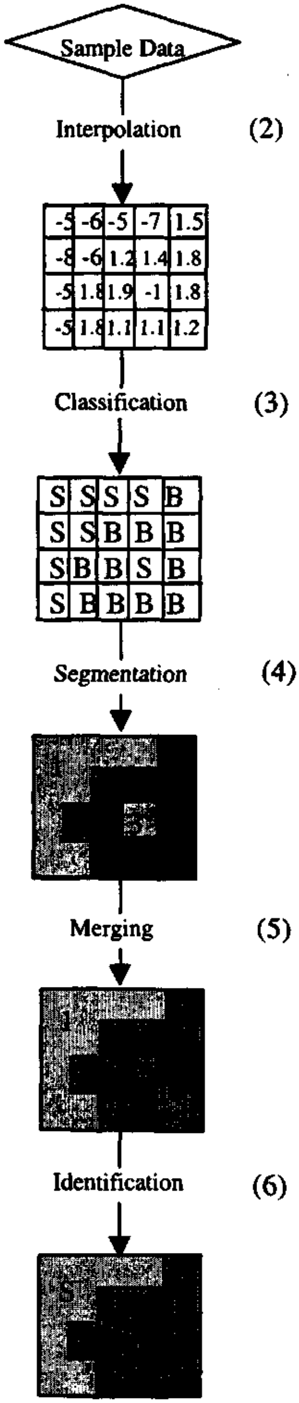


Figure 5.1 Procedure for object identification.

Chapter 4. Here uncertainty propagation in the steps of segmentation and merging are mainly discussed.

Two semantically different situations can be considered: either the objects might overlap or they are considered to be spatially disjoint. This implies different types of relationship between fuzzy spatial extents of objects. Therefore, fuzzy object models for different situations are discussed. Following that, uncertainty transfer in segmentation and merging is discussed.

5.3 Fuzzy object models

After classification, each grid cell is assigned a membership vector consisting of the elements $0 \le MF[P_{ij}, C_k] \le 1$ ($k=1, N$). Here $MF[P, C]$ represents the membership of grid cell P belonging to object class C (see Section 4.5), and N is the total number of object class types.

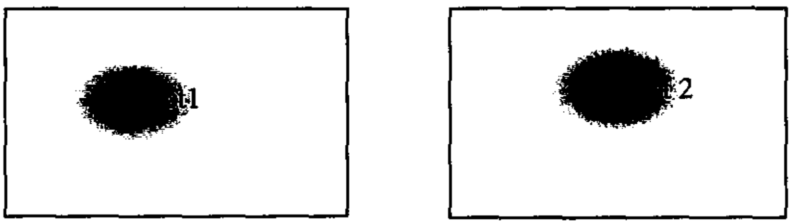
The objects of one class can then be represented as a layer of the raster, so that N layers of objects will be formed. A comparison of the N layer classification results with the conventional crisp classification shows that each layer consists of fuzzy regions with fuzzy boundaries. If each region represents the spatial extent of an object, the object is called a fuzzy-fuzzy object (FF-Object), where the first 'fuzzy' means that its boundary is fuzzy and the second 'fuzzy' means that its interior is fuzzy, because it contains cells that have been assigned to the region with a certainty less than 1 (see Figure 5.2A). Then the traditional crisp object can be called crisp-crisp (CC-Object), which means that the boundary and the interior of the object are crisp.

There are two ways of interpreting FF-objects when we discuss objects defined in different contexts. In some applications, fuzzy spatial overlaps among objects are permitted. In this notation, objects are defined in different contexts, say in soil map and the other in a land use map. Then it is quite possible that they overlap, i.e. their semantic definition does in general not forbid their overlap. In this case, the cells in the overlap region may belong to multiple objects.

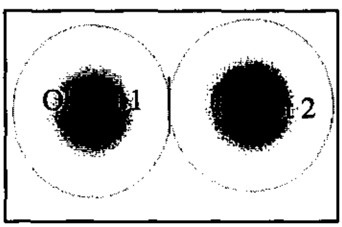
In other applications area objects are defined as being spatially disjoint in space (in single context), i.e. they do not overlap such that each grid cell belongs in principle to one object. If the objects form a spatial partition then each cell belongs to exactly one object, as in the case study, where foreshore, beach and foredune are considered to be spatially disjoint objects. Although the boundary between beach and foredune is not very crisp, a specific location should either belong to beach or foredune, but not to both. In this case it is necessary to combine the objects of different layers into one layer and to form a complete spatial partition of the area, which can be further differentiated into two cases. The first is that a boundary has to be set to define explicitly the spatial extent of objects and assign each grid cell exactly to one object; the second is that a clear boundary cannot be defined, but that there are transition zones between the objects. In the transition zones, no decision is made about which object the grid cells might belong to.

To differentiate these two situations, I call objects in the first view crisp-fuzzy objects (CF-Object, see Figure 5.2B), which means that the conditional boundaries between objects are crisp but the interiors of the objects are fuzzy. I call objects in the second view fuzzy-crisp objects (FC-Object, see Figure 5.2C), where fuzzy means that their boundaries (transition zones) are fuzzy and crisp means that their interiors (cores) are certain.

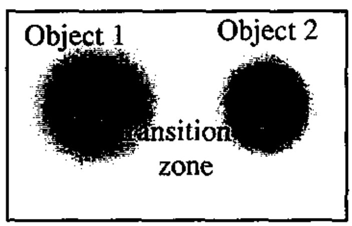
Figure 5.2 illustrates the three fuzzy object models. It can be seen that the FF-object has a fuzzy spatial extent, the FC-object has a crisp internal core but a fuzzy boundary (transition zone) and the CF-object has a crisp conditional boundary but a fuzzy interior.



(A) FF-object model



(B) CF-object model



(C) FC-object model

Figure 5.2 Fuzzy object models.

5.4 Propagation of uncertainty in segmentation

Section 5.3 discussed three fuzzy object models. For each fuzzy object models, different criteria should be applied to select the grid cells that represent the spatial extent of objects. In this section I discuss these criteria for each model.

As argued in Section 5.1, classification and segmentation are essential to extract objects from field observation data. The classification is uncertain in the sense that the membership function value varies per grid cell and is less than one. To identify spatial extent, first the grid cells have to be assigned to classes and then segment the raster into regions with grid cells of the same classes, which represent spatial extents of fuzzy objets. Therefore, I first discuss the criteria for assigning the grid cells to classes based upon their fuzzy classification result, then the criteria for segmentation. Since the spatial extents of the objects are fuzzy, only conditional boundaries can be defined. This is discussed in Subsection 5.4.3 (The discussion is based upon the syntactic representation of fuzzy objects (see Subsection 3.4.3 and Section 3.8) in order to formalize the three fuzzy object models).

Let $MF[P_{ij}, C_k]$ represent the membership of grid cell P_{ij} belonging to class C_k, and $D[P_{ij}, C_k]$ represent the decision function assigning P_{ij} to a region of C_k.

5.4.1 Assigning grid cells to classes

5.4.1.1 fuzzy-fuzzy object model

If $(MF[P_{ij}, C_k] \geq Threshold)$ then let $D[P_{ij}, C_k] = MF[P_{ij}, C_k]$;
(means belonging to a region of C_k)
If $(MF[P_{ij}, C_k] < Threshold))$ then let $D[P_{ij}, C_k] = 0$.
(means not belonging to a region of C_k)

5.4.1.2 fuzzy-crisp object model

If $(MF[P_{ij}, C_k] = 1)$ then let $D[P_{ij}, C_k] = MF[P_{ij}, C_k] = 1$;
(means belonging to the core of an object of C_k)
If $(0 < MF[P_{ij}, C_k] < 1)$ then let $D[P_{ij}, C_k] = MF[P_{ij}, C_i]$;
(means belonging to the transition zone)
If $(MF[P_{ij}, C_k] = 0))$ then let $D[P_{ij}, C_k] = 0$.
(means not belonging to an object of C_k)

5.4.1.3 Grid cells in fuzzy regions for crisp-fuzzy objects

Let $NM[P_{ij}, C_k] = 1 - MF[P_{ij}, C_k]$ represent no-membership, i.e. the certainty that p_{ij} does not belong to class C_k; and let $XM[P_{ij}, C_k]$ express the membership that p_{ij} belongs exclusively to C_k and not to any other classes C_l for any $l \neq k$. Because $XM[P_{ij}, C_k]$ expresses that the grid cell belongs to class C_k <u>and</u> not to any other classes, it can be derived by applying minimum operations as

$$XM[P_{ij}, C_k] = MIN(MF[P_{ij}, C_k], MIN_{l \neq k}(NM[P_{ij}, C_l])). \qquad (5\text{-}1)$$

As P_{ij} can only belong to one class, it requires only one class for which the function $XM[]$ has maximum value for P_{ij}. If there are more classes with the same maximum values then additional evidence is required to be able to select a unique class. It can be represented as

if $XM[P_{ij}, C_k] = MAX_{c_l}(XM[P_{ij}, C_l]) \quad (l = 1, \cdots, N)$
then let $D[P_{ij}, C_k] = XM[P_{ij}, C_k]$
otherwise $D[P_{ij}, C_k] = 0$. (5-2)

For example, in our case a grid cell has a membership vector

$$MF[P, C] = \begin{Bmatrix} 0.2 \\ 0.7 \\ 0.1 \end{Bmatrix}$$

where C_1 is foreshore class, C_2 is beach class and C_3 is foredune class.

So

$$NM[P, C] = \begin{Bmatrix} 1-0.2 \\ 1-0.7 \\ 1-0.1 \end{Bmatrix} = \begin{Bmatrix} 0.8 \\ 0.3 \\ 0.9 \end{Bmatrix}$$

$$XM[P, C] = \begin{Bmatrix} MIN(0.2, MIN(0.3, 0.9)) \\ MIN(0.7, MIN(0.8, 0.9)) \\ MIN(0.1, MIN(0.8, 0.3)) \end{Bmatrix} = \begin{Bmatrix} 0.2 \\ 0.7 \\ 0.1 \end{Bmatrix}$$

As $XM[P, C_2] = MAX \begin{Bmatrix} 0.2 \\ 0.7 \\ 0.1 \end{Bmatrix} = 0.7$, therefore $D[P, C_2]$=0.7.

It means that this cell is assigned to class C_2 (beach class) with certainty 0.7.

5.4.2 Segmentation

After assigning the grid cells to classes, regions of different class types can be formed.

Let $P_{kl} \in S_a$ where S_a is a region of class type C_k and let $Part[P_{ij}, S_a] \equiv D[P_{ij}, C_k]$.

If a grid cell P_{ij} satisfies the following two conditions:

$$ADJACENT[P_{kl}, P_{ij}] = 1 \text{ and } D[P_{ij}, C_k] \neq 0,$$

then $P_{ij} \in S_a$ and we note $Part[P_{ij}, S_a] = D[P_{ij}, C_k]$
to be the membership function of P_{ij} to S_a.

If the two conditions are not satisfied, $P_{ij} \notin S_a$, i.e. $Part[P_{ij}, S_a] = 0$. (5-3)

Since S_a represents the spatial extent of an object O_a of C_k, the relationship between P_{ij} and O_a (see Subsections 3.4.3 and 3.8.3) can be defined as

$$Part[P_{ij}, O_a] = Part[P_{ij}, S_a] = D[P_{ij}, C_k]. \tag{5-4}$$

Therefore the relationship between P_{ij} and O_a for objects in the CF-object model can be written as

$$Part[P_{ij}, O_a] = D[P_{ij}, C_k] = XM[P_{ij}, C_k]. \tag{5-5}$$

The relationship between P_{ij} and O_a for objects in the FF-object model and the FC-object model can be written as

$$Part[P_{ij}, O_a] = D[P_{ij}, C_k] = MF[P_{ij}, C_k]. \tag{5-6}$$

Equations 5-5 and 5-6 express the relationship between uncertainty of the cell belonging to the spatial extent of an object and the uncertainty of a grid cell belonging to classes, i.e. the relationship of uncertainties between geometric aspects and thematic aspect. It means that uncertainty of thematic aspects (fuzzy membership function value of grid cells) is converted to geometric aspects of the objects (fuzzy spatial extent and fuzzy boundary) through segmentation.

5.4.3 Boundaries and adjacency relationship between fuzzy objects

Since spatial extent is fuzzy, there is no crisp boundary between objects. After identification of the spatial extent of objects, i.e. assigning grid cells to regions, boundaries are formed, however. We can call them conditional boundaries because they are different from crisp boundaries. In this subsection I discuss the syntactic representation of the conditional boundaries for each fuzzy object model (Cheng *et al.*, 1998).

According to Equation 3-6 of Subsection 3.4.3, the boundary of an object consists of edges that have the object on one side. To check if an edge has an object on one side, the relationship of the edge and the faces belonging to the object should be checked.

5.4.3.1 FF-object model

As defined in Subsection 4.2.1, the faces of a FF-object should satisfy

$$Face(O_a) = Cell(O_a) = \{P_{ij} \mid Part[P_{ij}, O_a] > C\} \qquad (5\text{-}7)$$

where O_a is a FF-object, C is the value of threshold and $Part[P_{ij}, O_a]$ is defined in Equation 5-6.

With this function the conditional function can be defined:

$$Part[P_{ij}, O_a \mid c] = 1$$

where c is the value of the threshold.

And for the relationship between edges, faces and objects

$$Le[e_i, O_a \mid c] = Le[e_i, f_a] * Part[f_a, O_a \mid c]$$
$$Ri[e_i, O_a \mid c] = Ri[e_i, f_a] * Part[f_a, O_a \mid c].$$

They are crisp functions that take the value of either 1 or 0.

With the conditional function

$$B[e_i, O_a \mid c] = Le[e_i, O_a \mid c] + Ri[e_i, O_a \mid c],$$

the conditional boundary of O_a at certainty level c can be found:

$$\partial_c O_a = \{e \mid B[e, O_a \mid c] = 1\}. \qquad (5\text{-}8)$$

This means that the boundary of FF-objects is a threshold-cut boundary, consisting of the edges of outermost grid cells of the object.

5.4.3.2 FC-object model

Transition zones between FC-objects can be defined as those cells that do not belong to any core of the FC-objects but do belong to the intersection of the indeterminate boundaries. The transition zones between two FC-objects comprise the cells that satisfy

$$\begin{aligned} Face(O_a, O_b) &= Cell(O_a, O_b) \\ &= \{P_{ij} \mid 0 < Part[P_{ij}, O_a] < 1 \; and \; 0 < Part[P_{ij}, O_b] < 1\} \end{aligned} \tag{5-9}$$

where O_a and O_b are two FC-objects, $Part[P_{ij}, O_a]$ and $Part[P_{ij}, O_b]$ is the uncertainty of cell P_{ij} belonging to O_a and O_b, respectively, and is defined in Equation 5-6. For more than two objects the formulas still apply.

Here we define the 'confusion index' (CI) for each grid cell as one minus the difference between the maximal and second maximal membership values (Burrough, 1996),

$$CI = 1 - (Max1(L[P_{ij}, C_k]) - Max2(L[P_{ij}, C_l])) \; (k, l = 1, \ldots, N; \; k \neq l) \tag{5-10}$$

where N is the total number of the class types, $Max1$ is a maximum operation among the likelihood vectors, and $Max2$ is an operation to find the second largest value from the likelihood vectors. When the likelihood function values of adjacent grid cells are very similar, the zones of confusion divide regions with relatively homogenous membership values. In fact, these confusion zones indicate the presence of conditional boundaries; they are transition zones between FC-objects.

5.4.3.3 CF-object model

For CF-objects, the conditional boundary between two objects is the transition boundary between two classes. A set of faces of two objects should satisfy

$$Face(O_a) = Cell(O_a) = \{P_{ij} \mid Part[P_{ij}, O_a] > Part[P_{ij}, O_b]\}$$

and (5-11)

$$Face(O_b) = Cell(O_b) = \{P_{ij} \mid Part[P_{ij}, O_b] > Part[P_{ij}, O_a]\}.$$

Then the transition boundary consists of edges that have simultaneously the cells of O_a on the left side and the cells of O_b on the right side, or the cells of O_a on the right side and the cells of O_b on the left side. Therefore, the edges of the boundary should satisfy

$$E_{a,b} = \{e_i \mid B[e_i, f_a] = 1 \; and \; B[e_i, f_b] = 1 \; and \; f_a \in Face(O_a) \; and \; f_b \in Face(O_b)\}$$

Then the transition boundary is

$$\partial(O_a, O_b) = \{N_{a,b}, E_{a,b}\} \text{ and } N_{a,b} = \{n_i \mid n_i \in e_i, e_i \in E_{a,b}\}.$$

5.4.4 Summary

Here I have discussed the uncertainty propagation from classification to segmentation, i.e. from thematic data to geometric aspects of objects. The uncertainty of a cell belonging to classes is propagated during the spatial segmentation to the uncertainty of the cell belonging to a region.

5.5 Propagation of uncertainty in merging

Merging is a process whereby objects are combined to build a composite element; after the process the original entities ceases to exist (Molenaar, 1998a). In a merging process, small regions are resolved into other larger regions, which will change the uncertainty of the cells belonging to the regions and the spatial topologic relationship among these regions.

The merging process has to satisfy two kind criteria. One is a thematic criterion, which specifies the (sub) classes and the geometric criterion (e.g. size) of the regions that can be merged. The other is a topologic criterion, which requires that the regions to be merged are adjacent (Molenaar, 1998a). Here the following steps are used to implement these criteria:

(1) Calculate the area for each region, indicating the regions smaller than the minimum-mapping unit.

(2) check the topological relationships of the indicated region with other regions.

If it is not adjacent to another region (e.g. regions of FF-objects), it is merged out; stop;
If it is adjacent to only one region, it is merged into the adjacent region; stop;
If it is adjacent to many regions, go to Step 3.

(3) Calculate the sharpness of boundary (edges of the outmost cells) of the region to be merged.

Wang & Hall (1996) defined the sharpness of a boundary for the polygon map of nominal properties based upon the purity of polygons. Here I adopt the concept of sharpness for the boundary of the regions. If an edge e has grid cell $P_{i,j-1}$ and $P_{i,j}$ on its left and right side, the sharpness of the edge can be defined as

$$Sharpness(e) = \frac{1}{N}\sum_{k=1}^{N}\left|XL[P_{i,j-1}, C_k] - XL[P_{ij}, C_k]\right| \tag{5-12}$$

where $XL[P, C]$ refers to the membership of grid cell P exclusively belonging to class C, N is the total number of the class types of features.

(4) Compare the sharpness of the edges of the boundary, resolving the edge that has the minimum value of sharpness. The class type of the cell is changed to the class type of the region that it is merged into. The uncertainty of the cell belonging to the new region is calculated based upon Equation 5-5 or Equation 5-6.

(5) Repeat Steps 3 to 4 until all the grid cells of the small regions have been merged to other regions.

An example of calculating the sharpness of edges of a boundary and resolving edges follows. As illustrated in Figure 5.3, there is a small region containing only one grid cell. The fuzzy membership values of the grid cell to three classes are {0.1, 0.2, 0.7}. Its four neighbors, grid cells 1, 2, 3 and 4 have the following fuzzy membership values to the three classes: Cell 1{0.8, 0.0, 0.2}, Cell 2{0.9, 0.0, 0.1}, Cell 3{0.0, 0.9, 0.1} and Cell 4{0.0, 0.8, 0.2}. The sharpness of the four edges can be calculated as:

$$Sharpness_1 = \frac{|0.1-0.8|+|0.2-0.0|+|0.7-0.2|}{3} = 0.46$$

$$Sharpness_2 = \frac{|0.1-0.9|+|0.2-0.0|+|0.7-0.1|}{3} = 0.53$$

$$Sharpness_3 = \frac{|0.1-0.0|+|0.2-0.9|+|0.7-0.1|}{3} = 0.46$$

$$Sharpness_4 = \frac{|0.1-0.0|+|0.2-0.8|+|0.7-0.2|}{3} = 0.40$$

Therefore, the sharpness of Edge 4 is lowest and will be resolved. Hence, the cell will be changed to Class B with certainty 0.2.

(A) (B) (C)

Figure 5.3 Merging process.
(A) Classification result (B) Four edges of the small region C
(C) Merging result

5.6 Case study

This section describes the identification of the fuzzy spatial extent of objects in our case study; i.e. the identification of spatial extent of foreshore, beach, foredune in the coastal region of Ameland.

5.6.1 Modeling by Crisp-Crisp Object Model

First, I mapped the spatial extent of these landscape units based upon the crisp definition summarized in Table 2-1. After interpolation, the grid cells were classified into three classes of landscape units (applying Equation 4-1). There are grid cells that have height values of (-6.07 m) and (-6.15 m) that do not fall into any of the specified classes. They are classified as unknown. After segmentation, five regions were formed, including one region of unknown class. Checking the area of each region, two regions, including the one of unknown class, were smaller than the predefined minimum size of the mapping units. They were merged with adjacent areas. Therefore, three regions were finally formed. Figure 5.4 presents the modeling result.

According to the fuzzy definition of these landscape units (Table 3-2), each grid cell has three membership values to the three classes, as shown in Figure 5.5 A, B and C.

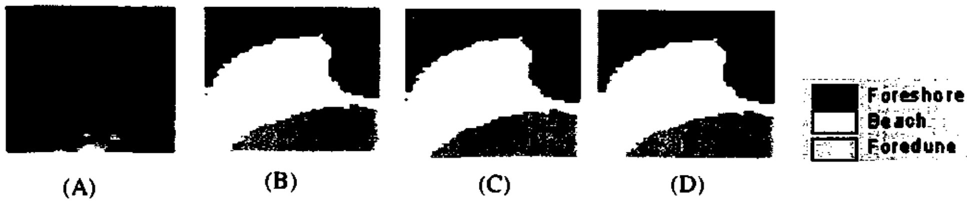


Figure 5.4 Results of CC-Object Model.
(A) Original height values (B) Map of classes
(C) Segmented regions (D) Regions after merging

Figure 5.5 Fuzzy classification results.
(A) Membership value of belonging to foreshore
(B) Membership value of belong to beach
(C) Membership value of belonging to foredune

Here I derive the objects for different fuzzy object models.

5.6.2 Modeling by Fuzzy-Fuzzy Object Model

FF-objects are formed according to the rules of Subsection 5.4.1.1; the results are shown in Figure 5.6. The edges of the outmost grid cells of each object are the conditional boundaries, with a threshold of 0.2. Figure 5.6 A, B and C each represent a layer with objects of one type. When these layers are overlaid, it is clear that these regions overlap. The fuzzy spatial extent of the objects is shown in Figure 5.6 D to F.

5.6.3 Modeling by Fuzzy-Crisp Object Model

Figure 5.7 A represents the core of the FC-objects. These objects have been formed according to the rules of Subsection 5.4.1.2. Figure 5.7 B represents the 'confusion index (*CI*)' as defined in Subsection 5.4.3.2. According to the definition of *CI*, the cells with values approximately equal to 1 represent transition zones among FC-objects. By combining Figures 5.7 A and B, FC-objects can be shown as in C.

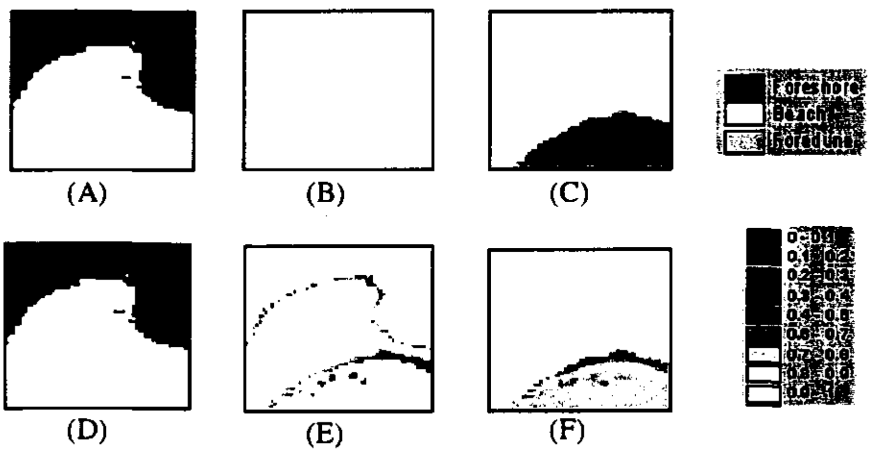


Figure 5.6 Results of FF-Object model.
(A) FF-objects belong to foreshore* (B) FF-objects belong to beach*
(C) FF-objects belong to foredune*
(D) – (F) FF-objects of (A) – (C) with fuzziness
(* threshold = 0.2)

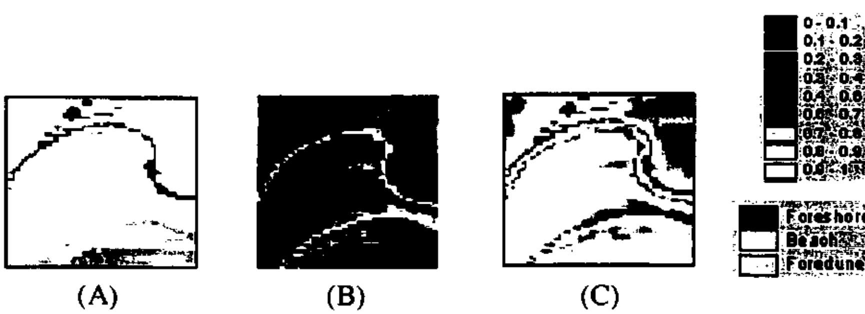


Figure 5.7 Results of FC-Object model.
(A) Cores of FC-objects (B) Confusion index (CI) (C) FC-objects with fuzziness

5.6.4 Modeling by Crisp-Fuzzy Object Model

The modeling result of the CF-object model is shown in Figure 5.8. Figure A shows the spatial extent of CF-objects. These objects were formed according to the rules of Subsection 5.4.1.3. Figure 5.8 B represents the uncertainty of cells belonging to the objects. Combining the uncertainty with the object gives the result shown in C. D shows the transition boundaries among these three objects (belonging to three classes).

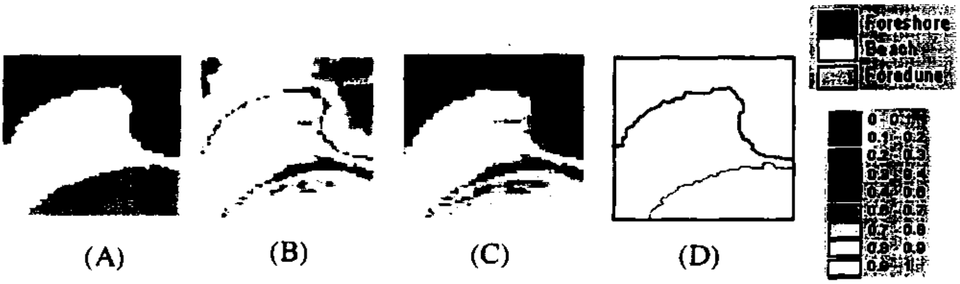


Figure 5.8 Results of CF-Object model.
(A) CF-object Model (B) Uncertainty of cells belonging to objects
(C) Objects with uncertainty (D) Conditional boundaries between regions
(darker means greater uncertainty)

5.7 Discussion

Since a beach should be adjacent to a foreshore area <u>and</u> to a foredune area, the following topological relationships between objects should be satisfied after merging.

- a beach area only adjacent to foreshore must be reclassified as a sand bar
- a beach area only adjacent to foredune area must be reclassified as 'blowout' or 'dune valley'.

These criteria can be used to guide object identification, as illustrated in Section 5.2. However, the temporal information can also be used (see Chapter 6) to identify objects.

5.8 Summary

This chapter presents an approach for identifying spatial extent of fuzzy objects from field observation data. Conventionally, the boundary of objects will be identified first and then their spatial extent. However, when there are no crisp boundaries between objects, the fuzzy spatial extent of objects should be identified first. Conditional boundaries can then be formed.

The propagation of uncertainties in the procedure of segmentation and merging is analyzed in this chapter. Uncertainty of thematic aspects is converted to geometric aspects of the objects through these two steps.

This chapter proposes three fuzzy object models to represent objects with fuzzy spatial extents for different situations. The FF-object model is useful to model fuzzy objects that spatially overlap each other, and the FC-object model and CF-objects model are suitable for modeling fuzzy objects that are spatially disjoint. The identification of objects by these models is illustrated by a case from coastal geomorphology.

So far, I have discussed uncertainty in object extraction through sampling categorization. The next chapter (Chapter 6) explains how the dynamics can be evaluated for the objects obtained by this method.

Chapter 6

Dynamics of Fuzzy Objects*

6.1 Introduction

Fuzzy regions can be extracted from field observation data (Chapter 5), but a further step is needed to identify the objects that are represented by these regions. Conventionally, this step is based on interpretation by domain experts or by a field check. Afterwards, changes in objects are detected by comparing their states at different epochs. The processes the objects have undergone are then analyzed by the experts.

This chapter describes a method for identifying the objects and detecting their dynamics automatically, based upon the regions obtained at different epochs. The chapter is structured as follows. Section 6.2 introduces the concept of state and process to describe the dynamics of natural phenomena. Section 6.3 develops a method to identify the objects and their dynamics, based upon the fuzzy spatial extent (regions) extracted at different epochs. In particular, the assumption of the method, the spatial relationship between the regions, and state transitions of objects are discussed. The practical usefulness of the advocated method is illustrated in Section 6.4 through a case study. Since these objects are fuzzy, their changes will be fuzzy. Therefore, Section 6.5 discusses changes of these fuzzy objects and the uncertainties of those changes. The final section (6.6) summarizes the major findings of the chapter.

6.2 State and process

In previous chapters, natural phenomena have been interpreted as objects. Their thematic descriptions can be expressed through a predefined attribute structure (cf. Subsection 3.4.1), whereas their geometric descriptions can be expressed through their spatial extent. Therefore, the dynamics of the objects can be represented as changes of their thematic attribute values and spatial extent.

* This chapter is based upon Cheng and Molenaar (1997, 1998b, 1998c).

Sometimes, the spatial extent of an object might change (e.g. the boundary lines identifying a land parcel may vary over time), while its thematic aspect (the owner of the parcel) does not. Sometimes, thematic attributes associated with the object may vary over time without changing its geometry (e.g. the vegetation type of a specific parcel might be different for different years). But for natural phenomena, the geometric aspects of an object generally change due to a change in thematic aspects. If the values of the thematic attributes at some locations change beyond a specific threshold, this may cause changes in the spatial extent of natural phenomena. Thus, I mainly consider changes of geometry, such as shift, merge and split, that are due to changes of the thematic aspects of the phenomena

The terms 'state' and 'process' represent the dynamics of natural phenomena. 'State' defines the thematic and geometric attribute values of natural phenomena at a specific time, and 'process' describes the physical change or transition between states. State is usually monitored at a fixed temporal interval, such as one day, one month, one year or ten years. Process is analyzed by comparing states with the intent of detecting change.

6.3 Identification of state transitions

The dynamics of objects are monitored through time series data (Figure 6.1). At a fixed temporal interval, the states of objects are monitored by field observation and are represented by the spatial extents as they appear at different epochs. Then links between regions at two consecutive epochs should be built to form a historical line of an object. For example, Region 1 is linked with Region 4, and 4 with 7, representing the lifeline of Object 1; Region 2 is linked with Region 5, representing the lifeline of Object 2; Region 3 with Region 6, representing the lifeline of Object 3; Region 5 with 8 and region 6 with 8, representing the lifeline of Object 4. This procedure is usually done by domain experts. Afterwards, the processes the objects involved in such as a shift, are also analyzed. The temporal relationships between objects are also derived, e.g. Object 2 and Object 3 have merged into a new Object 4. This section proposes an approach for analyzing the relationships of regions and for identifying objects and their processes automatically.

6.3.1 Assumption of the method

The procedure in Section 5.3 is used to derive the regions that represent the spatial extents of objects at each epoch. The regions at different epochs can be linked to form the lifeline of an object. When this is done, the object can actually be identified. This can be realized based on the assumption that from a long-term point of view natural phenomena change gradually. Especially the change of the coastal zone in the case study of Chapter 2 can be regarded as continuous (Galton, 1997) and slow, so that the objects can be considered to be rather stable. The approach developed in this section is designed for such cases.

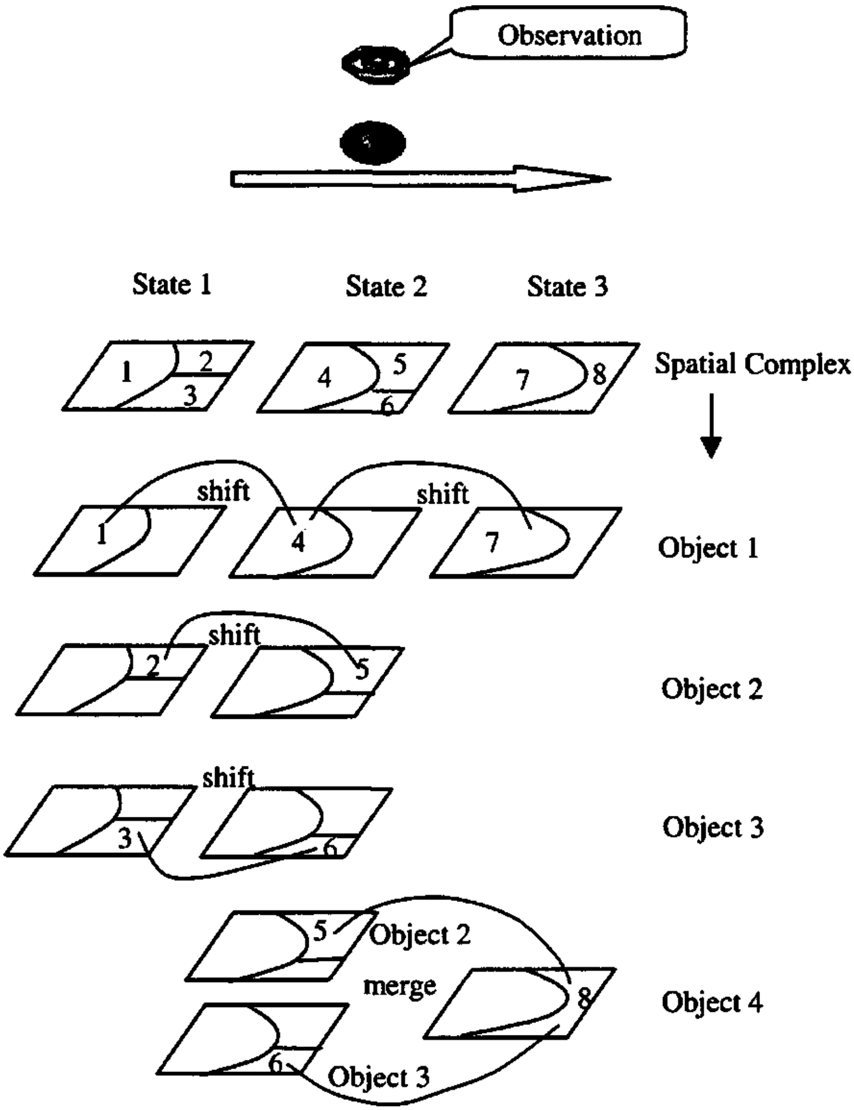


Figure 6.1 States and processes.

Under this assumption, two regions that are the spatial extents at different epochs of one and the same object should have an overlap that is larger than their overlaps with the spatial extent of any other object. Therefore, we can find the successor of a region at epoch t_n by calculating its spatial overlaps with all regions that appeared at epoch t_{n+1}. The one with maximum overlap can be identified as the successor.

6.3.2 Similarity indicators

The overlap of two regions S_a and S_b can be found through the intersection of their cell sets. It is a very simple raster-based operation.

$$Overlap[S_a, S_b] = Cells(S_a) \cap Cells(S_b) \tag{6-1}$$

where $Cells(S_a)$ and $Cells(S_b)$ represents the sets of grid cells belonging to region S_a and S_b, respectively.

As the regions are uncertain, the spatial overlap measured between two regions should take their fuzziness into account. The possibility of a grid cell being part of the overlap of two fuzzy regions can be defined as (Dijkmeijer and Hoop, 1996):

$$Overlap[S_a, S_b \mid P_{ij}] = MIN\{Part[P_{ij}, S_a], Part[P_{ij}, S_b]\} \tag{6-2}$$

where $Part[P_{ij}, S_a]$ and $Part[P_{ij}, S_b]$, which represents the certainty of the grid cell P_{ij} belonging to S_a and to S_b, respectively, are defined as in Equation 5-4. The unit size of P_{ij} is considered to be 1 here, so that the size of a fuzzy region S is defined as

$$Size(S) = \sum_{P_{ij}} Part[P_{ij}, S] \text{ where } P_{ij} \in Cells(S). \tag{6-3}$$

The size of the overlap of two fuzzy regions is then

$$SOverlap(S_a, S_b) = \sum_{P_{ij}} Overlap[S_a, S_b \mid P_{ij}] \tag{6-4}$$

where $P_{ij} \in Cells(S_a) \cap Cells(S_b)$.

From the spatial overlap between regions, we can match regions that are spatially related. Let R_i be the set of regions at epoch T_i, and let $S_a \in R_1$ and $S_b \in R_2$. The following indicators can be used to evaluate the types of relationship between regions at two epochs.
The relative fuzzy overlap between two regions can be defined as

$$ROverlap(S_b \mid S_a) = SOverlap(S_a, S_b) / Size(S_a) \tag{6-5}$$

$$ROverlap(S_a \mid S_b) = SOverlap(S_a, S_b) / Size(S_b) \tag{6-6}$$

where $ROverlap(S_b|S_a)$ represents the ratio of the overlap to the size of S_a (relative fuzzy overlap to S_a); $ROverlap(S_b|S_a)$ is the ratio of the overlap to the size of S_b (relative fuzzy overlap to S_b).

The similarity of two fuzzy regions can be defined as

$$Similarity(S_a, S_b) = \frac{SOverlap(S_a, S_b)}{\sqrt{Size(S_a) \cdot Size(S_b)}} \tag{6-7}$$

6.3.3 State transitions

Using indicators of relative fuzzy overlap and similarity of two fuzzy regions, object state transitions can be identified between two epochs. Seven fundamental cases are shown in

Table 6-1. The combinations of indicator functions behave differently for these seven cases. The state transition of objects can be identified by the following steps.

For all $S_b \in R_2$ compute Size (S_b)
For all $S_a \in R_1$ do
> compute Size (S_a)
For all $S_b \in R_2$
> compute *SOverlap*(S_a, S_b)
> compute *Roverlap*($S_b|S_a$), *Roverlap*($S_a|S_b$), *Similarity*(S_a,S_b)
> evaluate *shift*(S_a; S_b), *expand*(S_a;S_b), *shrink*(S_a;S_b)
> evaluate *split*(S_a; ...S_b, ...), *appear*(S_b)
> evaluate *merge*(..., S_a, ...; S_b), *disappear*(S_a)

Table 6-1 Identification and presentation of state transition.

Regions at T_1	Regions at T_2	Overlay	Indicators: $Rovelap(S_b \mid S_a)$ $Rovelap(S_c \mid S_a)$	Indicators: $Rovelap(S_a \mid S_b)$ $/Rovelap(S_a \mid S_c)$	Indicators: Similarity	State Transition	Symbol
A	B		Large -	Large -	High -	shift(S_a;S_b)	
A	B C		Small Small	Large Large	Low Low	split(S_a;S_b,S_c)	
B C	A		Small Small	Large Large	Low Low	merge(S_b,S_c; S_a)	
A	B		Large -	Small -	Low -	expand(S_a;S_b)	
A	B		Small -	Large -	Low -	shrink(S_a;S_b)	
	B	-	- -	- -	- -	appear(S_b)	
A	-	-	- -	- -	- -	disappear(S_a)	

6.3.4 Dynamic objects

The procedure of the previous subsection can be used to identify relationships between regions, which represent possible dynamic relationships (state transitions) between objects at two different epochs. A complete lifeline for each object can be built by linking all its states. Regions thus related can be linked to form lifelines of objects that may have 'shifted', 'expanded' or 'shrunk' in two successive epochs. Regions that appear at a specific moment represent new objects, and regions that vanish at some moment represent disappearing objects. Furthermore, 'merging' and 'splitting' objects can be identified. The lifelines constructed represent dynamic objects. The procedure will be applied to the case study in the next section.

6.4 Case study

6.4.1 Identified fuzzy regions in different years

As discussed in Chapter 5, Section 5.2, five steps including sampling, interpolation, classification, segmentation and merging were implemented to identify fuzzy regions for each year. The regions formed are based upon the CF-object models, illustrated by Figure 6.2, where the numbers in the figure are the identifiers of the regions used in Table 6-2.

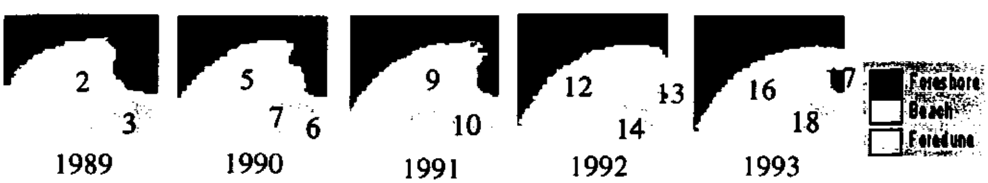


Figure 6.2 Classified regions.

6.4.2 Identified fuzzy objects and their dynamics

Regions are extracted per year, whereas objects might exist several years, i.e. their lifelines are represented by a sequence of regions, each of which represents the geometric state of an object in one year.

Lifelines of objects were constructed by evaluating the relationships between regions for two successive years, i.e. by identifying the type of state transition between S_a and S_b based upon the indicators according to the situations indicated in Table 6-1. To assign indicators to the high or low levels requires intuition, and therefore, expert knowledge. Further research is required to establish threshold for these indicators.

The fuzzy sizes of these regions and the fuzzy overlap of regions of successive years are shown in Table 6-2. The indicators of Subsection 6.3.2 can now be evaluated; with these we can link the regions by several lines (as shown in Table 6-2), which indicates that the regions connected by these lines are most likely the representations of the spatial extent of an object in successive years. For example, Region 1 is linked with Region 4, Region 4 with 8, 8 with 11; Region 3 is linked with Region 6, 6 with 10, 10 with 14. In 1990 we notice that a new region had been formed (Region 7). By checking the overlap of Region 7 with regions in 1989 and 1991, we identified overlap with Regions 3 and 10. Therefore these three regions are also linked by lines.

Table 6-2 Fuzzy overlaps and links between fuzzy regions.

Year	Region	Area	Overlap with regions in next year				Class Type
1989	1	1108.1	937.5	81.8	0.0	0.0	Foreshore
	2	1246.8	106.3	1104.8	9.2	0.0	Beach
	3	644.3	0.0	12.7	572.5	27.5	Foredune
1990	4	1138.7	975.0	76.0	0.0		Foreshore
	5	1229.7	76.0	1129.5	2.6		Beach
	6	586.8	0.0	0.0	564.3		Foredune
	7	28.0	0.0	0.0	26.3		Beach
1991	8	1101.3	862.7	116.9	6.4	0.0	Foreshore
	9	1260.1	87.3	1146.6	0.0	0.5	Beach
	10	609.8	0.0	3.3	0.0	605.7	Foredune
1992	11	1004.9	751.5	6.8	0.0	0.0	Foreshore
	12	1288.7	119.3	1101.1	38.9	2.8	Beach
	13	6.4	0.0	1.6	4.6	0.0	Foreshore
	14	625.7	0.0	2.7	0.0	604.4	Foredune

For example, for Region 1 in 1989 (S_1) and Region 4 in 1990 (S_4), $Soverl(S_1,S_4)$ = 937.531, and $Size(S_1)$ = 1108.1, $Size(S_4)$ = 1138.7. So

$$Rovelap(S_1|S_4) = 937.5/1138.7 = 0.823$$
$$Rovelap(S_4|S_1) = 937.5/1108.1 = 0.846$$
$$Similarity(S_1,S_4) = 0.835$$

Therefore, these two regions are very similar and can be considered as instances of the same object (here called Object 1) at two epochs. As there are differences between the boundaries of these two regions, Object 1 is considered to have shifted from Region 1 in 1989 to Region 4 in 1990.

A similar conclusion is valid for Region 3 in 1989 (S_3) and Region 6 in 1990 (S_6), where:

$ROvelap(S_3|S_6) = 572.5/644.3 = 0.819$
$ROvelap(S_6|S_3) = 572.5/586.8 = 0.976$
$Similarity(S_3,S_6) = 0.894.$

Therefore, these two regions, too, are instances of the same object (Object 3) at two epochs.

We also calculated the indicators for Region 3 (S_3) and Region 7 (S_7),

$ROvelap(S_7|S_3) = 27.5/644.3 = 0.043$
$ROvelap(S_3|S_7) = 27.5/28.0 = 0.982$
$Similarity(S_3,S_7) = 0.205.$

I conclude that these two regions are not similar to each other, but that Region 7 is more or less contained in Region 3. It can be identified as a new object appearing in 1990, and is split from Object 3 (Region 3 represents its spatial extent in 1989). By analyzing the overlap between regions of 1990 and 1991, I found that Region 7 disappeared in 1991: it had merged into Object 3 (Region 10 in 1991). Using the above approach, the objects and the processes the objects had evolved into were identified, as illustrated by Figure 6.3. The icons represent the regions (states) of objects at different times. The symbols represent the types of state transition. The figure shows that Object 4 split off from Object 3 between 1989 and 1990; it merged again with Object 3 between 1990 and 1991; Object 5 split off from Object 1 between 1991 and 1992.

Figure 6.4A presents the identified objects, whereas Figure 6.4B presents the uncertainties of their spatial extent by means of grey scale. This figure shows that the uncertainties of the assignment of grid cells to regions change from year to year together with the spatial extent of the objects.

6.4.3 Discussion

This case study explains the procedure of identification of fuzzy objects belonging to three object types. These types have been defined so that the objects always form a spatial partition of the mapped region. The dynamics of such a spatial complex can be determined by comparison of the states of the complex at successive epochs.

The processing time for the evaluation of indicators to identify the state transition of objects will increase with the number of regions appearing at each epoch. However, a two-step approach can be used to solve this problem. The first step compares the spatial overlaps of regions per object type. If region S_b at time t_1 is of the same type as region S_a at time t_0, and if they have a maximal similarity, then S_b can be considered as the successor of S_a. In this step most regions can be identified as spatial extents of objects at different epochs. There may, however, be regions (e.g. Region 7 in the case study) that have no spatial overlaps with regions of the same type at previous and/or later epochs. In this case, a further step should be taken.

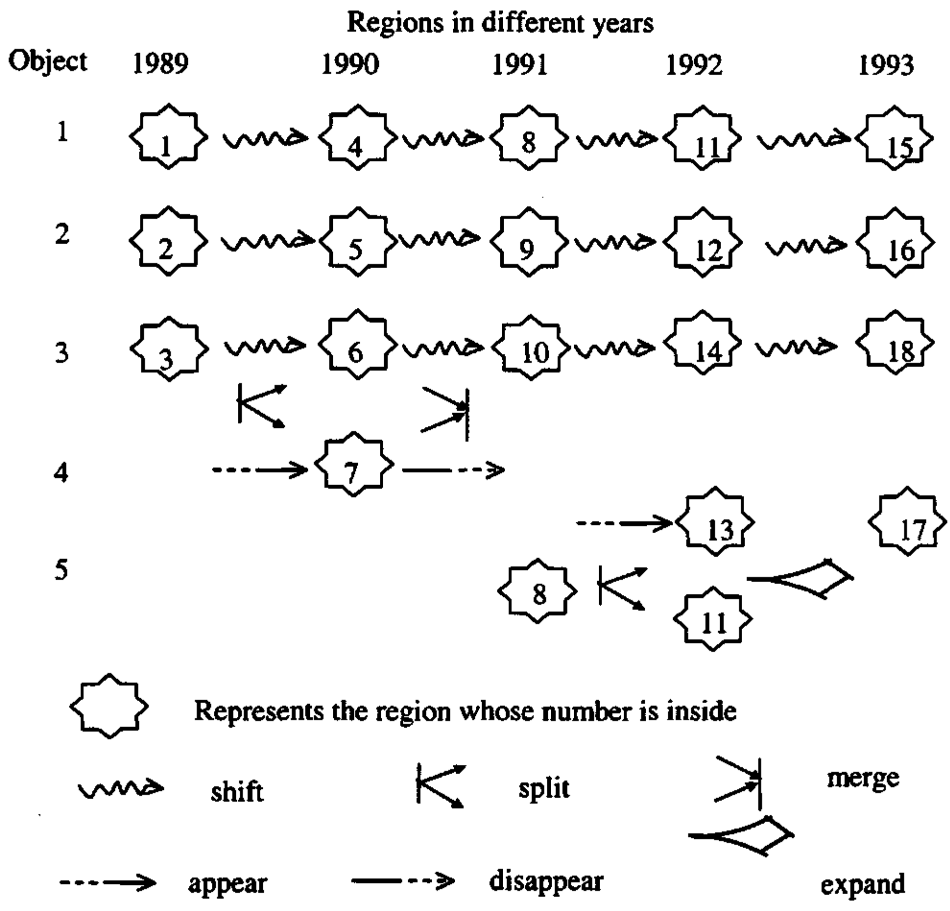
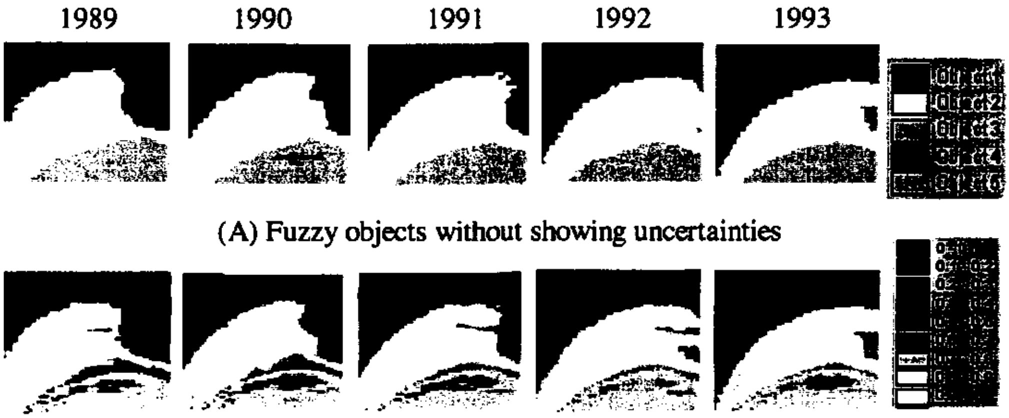


Figure 6.3 Identified fuzzy objects and processes.



(A) Fuzzy objects without showing uncertainties

(B) Fuzzy objects showing uncertainties

Figure 6.4 States of identified fuzzy objects.

We can compare the spatial overlap of this region with the regions belonging to other object types and check whether the region has been misclassified. According to the classification result, Region 7 belonged to the beach. Checking its spatial relationship with regions in other years, however, showed that it is only related to regions belonging to foredune and, so it was assigned to the foredune instead of beach. In fact, it was a lower part of the foredune region (dune valley). Therefore, by using the temporal information of the objects, the misclassified regions could be identified. This can be considered one of the advantages of the proposed approach, i.e. to adjust classification results. This situation has been discussed in Section 5.2 and Section 5.5, where the topological relationship between the regions should be used for object identification. For example, Region 7 (Figure 6.4 (1990)) was initially classified to beach area, but it is adjacent to only foredune, so it should be reclassified and identified as a dune valley. However, this situation was detected automatically by the method just described above.

6.5 Change of fuzzy objects and its uncertainty

This section discusses the change of the fuzzy objects, especially their change of shape, area and volume.

6.5.1 Change of shape and its uncertainty

By comparing the spatial extents of an object at two successive years we can derive the change of shape. This can be done through a simple spatial overlay operation. The uncertainty of change can be derived based upon the uncertainty of the grid cells belonging to object's extent for each year

$$MF_{S_a:S_b} = MIN(Part[P_{ij}, S_a]_{t_1}, Part[P_{ij}, S_b]_{t_2}) \quad (a \neq b) \tag{6-8}$$

where S_a is the spatial extent of object O_a at t_1, S_b is the spatial extent of object O_b at t_2; $Part[P_{ij}, S_a]_{t_1}$ represents the uncertainty of grid cell P_{ij} belonging to S_a at time t_1; $Part[P_{ij}, S_a]_{t_2}$ represents the uncertainty of P_{ij} belonging to S_b at time t_2.

For example, a grid cell belonged to foreshore (Object O_a) in year 1989 with certainty value $Part[P_{ij}, S_a]_{t_1} = 0.7$. It belonged to beach (Object O_b) in year 1990 with certainty value $Part[P_{ij}, S_b]_{t_2} = 0.8$. Therefore, the uncertainty of change of this cell according to Equation 6-8 is then,

$$MF_{S_a:S_b} = MIN(0.7, 0.8) = 0.7.$$

The changes of extent of these landscape units between 1989–1990, and between 1990–1991 are presented in Figure 6.5. It can be seen that the foreshore and beach were very active, but the foredune was quite stable. The changes of the foreshore and the beach were normally opposite to each other. It was also found that the certainties of change of the cells close to the center of change area were higher than that close to the edge of the change area. This is due to the fact that the cells close to the edge of the change area are close to one of the edges of the two objects, which are less certain than the cells closer to the center of object.

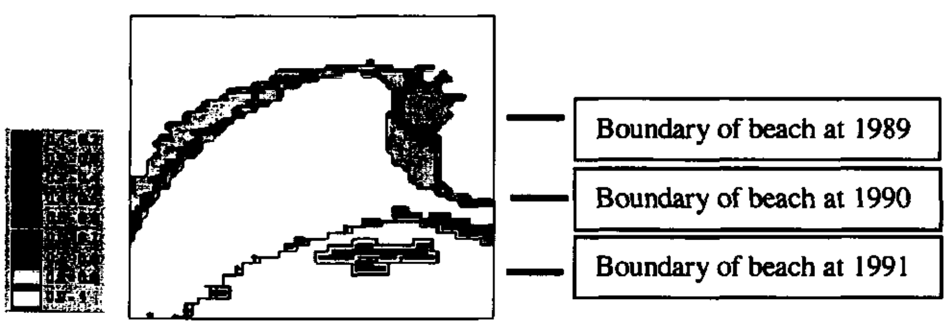


Figure 6.5 Shape changes of fuzzy objects.

Based upon this analysis, the developing trends of these landscape units can be analyzed qualitatively. Moreover, based upon this result, the changes of the landscape units can be calculated at different certainty levels. The change of foreshore and beach (1989 – 1990) at different certainty levels is reported in Table 6-3. The number of pixels falls with the increasing level of certainty. It implies that only sure about change from foreshore to beach (accumulation) for 25 pixels and beach to foreshore (erosion) for 25 pixels.

Table 6-3 Change between foreshore and beach at different certainty levels.

Certainty level ≥	1.0	0.9	0.8	0.7	0.6	0.5
Foreshore to beach	25*	58	67	75	87	94
Beach to foreshore	25	77	92	102	110	121

(*Number of pixels)

Based upon the change map of Figure 6.5, a series of change maps for different certainty levels was derived (Figure 6.6).

6.5.2 Change of area and volume

The area of a fuzzy object is the size of its spatial extent, which has been defined in Equation 6-3. Therefore, area of a fuzzy object O_a is then defined as

$$Size(O_a) = \sum_{P_{ij}} Part[P_{ij}, S_a] * Size(Cell) \tag{6-9}$$

where $P_{ij} \in S_a$, S_a is the fuzzy spatial extent of O_a, $Part[P_{ij}, S_a]$ is the uncertainty that grid cell P_{ij} belongs to S_a, area of grid cell = 60×60 (m^2).

Calculating the volume of a fuzzy object is similar to calculating its area. In both case uncertainties of the grid cell belonging to the object have to be taken into account.

$$Volume(S) = \sum_{P_{ij}} Part[P_{ij}, S] * h_{P_{ij}} * Size(Cell) \tag{6-10}$$

where $h_{P_{ij}}$ is the height of the grid cell with respect to a reference level and it is –20 m in our case, since some points on the test area are lower than sea level, e.g., –16 m. Other symbols refer to Equation 6-9.

The area and volume of the landscape units are presented in Figure 6.7, which shows that the fuzzy area of the whole region is not constant. This is due to the certainty of spatial extents of the landscape units varying from year to year. The total volume of sediment in the test field is decreasing which indicates general erosion, which can be used to guide the coastal defense works such as beach nourishment that needs high investments. The fuzzy area of foreshore and beach is shown in Table 6-4. The change of foreshore is derived from Table 6-4 by comparing values for two consecutive years. It is reported in the second row of Table 6-6.

Table 6-4 Fuzzy area of foreshore and beach (60 m × 60 m).

Year	1989	1990	1991	1992	1993
Foreshore	1108.1	1138.7	1101.3	1011.3	953.9
Beach	1246.8	1229.7	1260.1	1288.7	1137.3

This is one way to measure the change of area. Another way is to calculate the change of area based upon the change map (see Figure 6.5) by considering the change uncertainty of per pixel. The change between the foreshore and beach is reported in Table 6-5.

Table 6-5 Fuzzy change between foreshore and beach.

Year	89-90	90-91	91-92	92-93
Foreshore to beach	81.8	76.0	116.9	8.4
Beach to foreshore	106.3	76.0	87.3	158.2

As foreshore only interacts with beach, the change of foreshore is only due to exchange with beach. Therefore, the change of foreshore can be derived by subtracting the change of beach to foreshore from change of foreshore to beach. The final result of area change of foreshore is reported in the first row of Table 6-6. From the comparison of this result with that based upon Table 6-4 (the second row of Table 6-6), the differences become obvious, especially in the year 1992 – 1993. Checking in Table 6-5, it can be seen that the beach erodes on one side (the value is 158.2 in Table 6-5) but little accumulation can be found on the other side (the value is 8.4 in Table 6-5). So the net change of the foreshore is accumulation (+149.8). According to this value, the area of foreshore should be enlarged. However, checking with Table 6-6, the total area of the foreshore decreased (the value is -57.4 (year 92-93) in Table 6-6). These different are discussed in the next section.

Table 6-6 Fuzzy change of foreshore (60 m × 60 m).

Year	89-90	90-91	91-92	92-93
Based upon Table 6-5	+24.5	0.0	-29.6	+149.8
Based upon Table 6-4	+30.6	-37.4	-90.0	-57.4

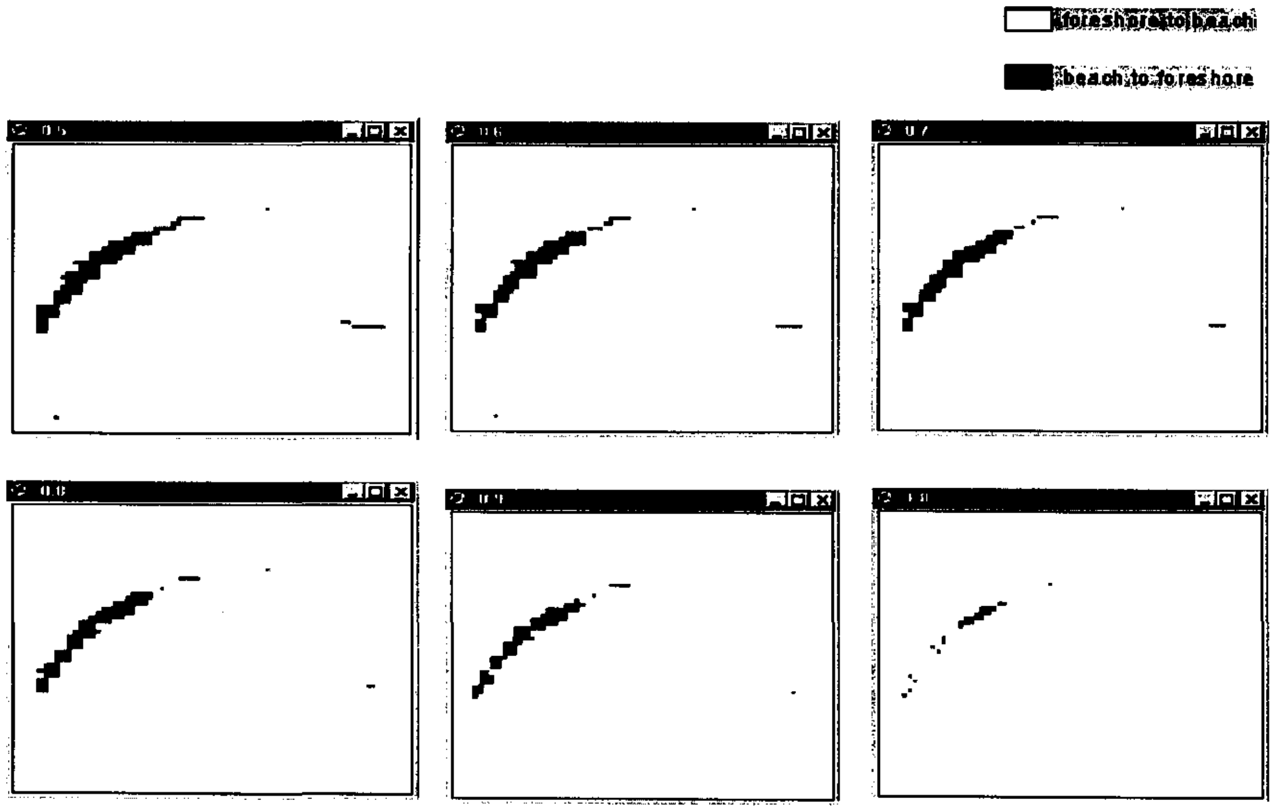
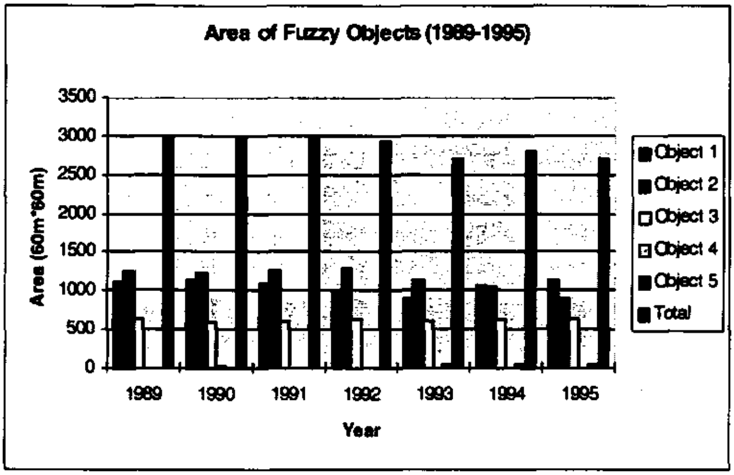
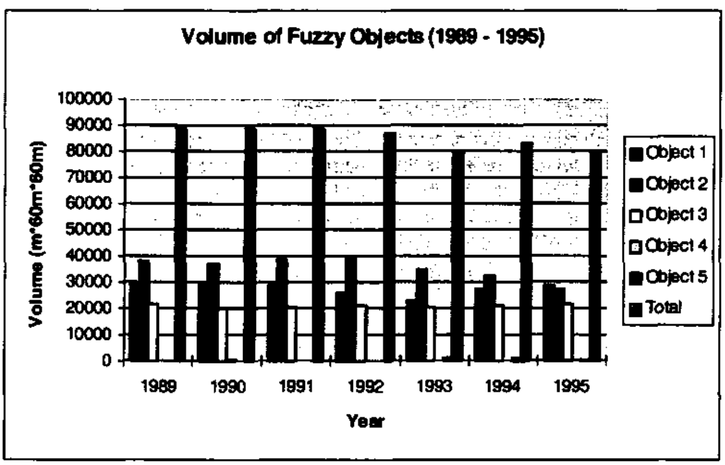


Figure 6.6 Change between foreshore and beach at different certainty levels (1989 – 1990).



(A)



(B)

Figure 6.7 Dynamic changes of area and volume of fuzzy objects.
*Object 1 and Object 5 are foreshore areas;
Object 2 is a beach area;
Object 3 and Object 4 are foredune areas.

6.5.3 Discussion

I proposed two ways of calculating the change of area. In the first approach, the cells of the whole area of foreshore, including the unchanged part, are considered for the calculation. The change area of foreshore is:

$$Change1 = \sum_{cells \in F_{1990}} Part[P_{ij}, F]_{1990} - \sum_{cells \in F_{1989}} Part[P_{ij}, F]_{1989}$$

$$= \sum_{1} Part[P_{ij}, F]_{1990} + \sum_{3} Part[P_{ij}, F]_{1990} - \sum_{2} Part[P_{ij}, F]_{1989} - \sum_{3} Part[P_{ij}, F]_{1989}$$

$$= \left(\sum_{1} Part\ [P_{ij}, F\ ]_{1990} - \sum_{2} Part\ [P_{ij}, F\ ]_{1989} \right) + \left(\sum_{3} Part\ [P_{ij}, F\ ]_{1990} - \sum_{3} Part\ [P_{ij}, F\ ]_{1989} \right)$$

where F means foreshore, the numbers 1, 2 and 3 represents the area indicated in Figure 6.8, respectively, and $Part[P_{ij}, F]$ represents the membership of the grid cell P_{ij} belonging to the foreshore area.

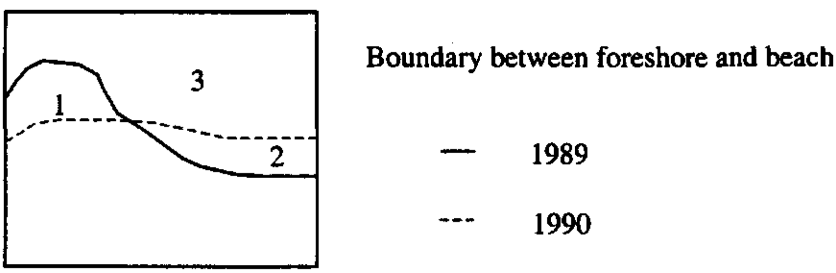


Figure 6.8 Calculation of change of foreshore.

Therefore, in this case, the calculation is related to cells in the changed areas and to their certainties of belonging to the areas (Area 1 and Area 2). It is also related to the change of certainty in the unchanged area (Area 3).

In the second approach, only cells in the changed area are considered for the calculation. The change area of foreshore is calculated as:

$$Change2 = \sum_{cells \in B_{1989} \cap cells \in F_{1990}} Min(Part[P_{ij}, B]_{1989}, Part[P_{ij}, F]_{1990})$$

$$- \sum_{cells \in F_{1989} \cap cells \in B_{1990}} Min(Part[P_{ij}, F]_{1989}, Part[P_{ij}, B]_{1990}$$

$$= \sum_{1} Min(Part[P_{ij}, B]_{1989}, Part[P_{ij}, F]_{1990}) - \sum_{2} Min(Part[P_{ij}, F]_{1989}, Part[P_{ij}, B]_{1990})$$

Therefore, in this case study, the calculation is related to the certainty of the cells in the changed areas (Area 1 and Area 2), which is related to both the certainty of the cells belonging to foreshore and to beach.

Comparing these two approaches of calculation makes clear that the first approach is more related to change of certainty of cells belonging to foreshore in two consecutive years, while the second approach is more related to the area of the change.

If

$$\forall P_{ij} \in Area1 \;\; Part[P_{ij}, B]_{1989} \geq Part[P_{ij}, F]_{1990} \text{ and}$$

$$\forall P_{ij} \in Area2 \;\; Part[P_{ij}, F]_{1989} \leq Part[P_{ij}, B]_{1990}$$

then

$$Change2 = \sum_{1} Part[P_{ij}, F]_{1990} - \sum_{2} Part[P_{ij}, F]_{1989} .$$

In this case, the difference between *Change*1 and *Change*2 is mainly due to the change of certainty of the cells in Area 3.

The difference due to these two approaches is very large in the year 1992-1993 (Table 6-6). This is mainly due to the fact that the assignment of cells to foreshore lower in the year 1993 than the certainty in the year 1992 (see also Figure 6.4).

Which method should be chosen depends on the specific case. It should be evaluated for a given landscape unit, situation and application to gain a better understanding of how to proceed with representation and analysis (Plewe, 1997). It is important to understand and analyze the uncertainty behind the calculation in order to provide accurate information to decision makers. For the first case the area is predefined, the uncertainty is related to the whole area. The change of area also considers the uncertainty of the whole area. For the second case, uncertainty is considered only for the area of change. Therefore, when we want to measure the change related to a landscape unit, the first approach should be taken. When we want to measure the change as an interaction between two objects, the second approach should be taken. Generally speaking, Figure 6.5 provided a more accurate and efficient way of representing the change, since the maps in Figure 6.6 could be derived from it, as can the result of *Change*2.

6.6 Summary

This chapter presents a method for identifying objects and their dynamics for fuzzy regions obtained at different epochs. Their dynamics are revealed through the spatial extents (states of objects) at different epochs; the objects are determined by comparing the relationships of these spatial extents. Simultaneously, the processes through which these objects evolve have been identified. The methodology has been demonstrated by an empirical example in a coastal geomorphological study of Ameland.

Chapter 7

A Process-Oriented Spatio-Temporal Data Model*

7.1 Introduction

The previous two chapters discussed the identification of fuzzy objects and their dynamics. The main subject of this chapter is the design of a database to support environmental modeling.

It has been explained in the previous chapter (Section 6.2, Figure 6.1) that the process of information extraction fellows several steps, i.e. from observation to states (spatial extents) of objects, then to objects and their processes. The database model will be designed so that information about objects, their state and processes and the relationships between them can be extracted from the data stored in the database. The model should support analysis of data in time series from varying perspectives, through location-oriented, time-oriented, object-oriented and process-oriented queries in order to understand the behavior of natural phenomena.

The chapter is organized as follows. The next section, Section 7.2, gives a detailed review of existing spatio-temporal data models. The limitations of existing models are summarized as a starting point for the design of the data model. The design is demonstrated for the case presented in Chapter 6 by formalizing, in Section 7.3, the representation of the dynamics of natural phenomena. Section 7.4 proposes a process-oriented spatio-temporal data model based upon the formalized descriptions. Multi-perspective queries are discussed in this section. Conclusions and future work are summarized in the last section, Section 7.5.

* This chapter is based upon Cheng and Molenaar (1998a).

7.2 Review of spatio-temporal data models

7.2.1 Basics of spatio-temporal data models

Since 1980 there has been increasing academic interest in the development of data models that incorporate time. Until recently, the semantics of the time domain, including its structure, dimension and indeterminacy has been intensively discussed (Snodgrass, 1992). Time has been added to many data models: entity-relationship models, semantic data models, knowledge-based data models, deductive data models, and object-oriented models. However, most of the literature on temporal databases is based on the relational data model, in which, the change of an entity is represented as a new version of the entity. For these revisions, three representations can be distinguished: version of a table, version of a tuple, and version of an attribute. These three approaches have different features. The first approach is simple in concept but results in high data redundancy. The third approach is compact but stores multiple attribute versions within a tuple, which cannot be handled by standard relational algebra. The second approach lies between the other two and is mostly used for representation and implementation in time modeling (Langran, 1992; Tansel *et al.*, 1994).

For incorporating temporal data models into GISs, four basic spatio-temporal data models have been proposed at conceptual level. In these models, time is considered as the fourth dimension, which is orthogonal to the other three spatial dimensions in a 4-D space-time.

The ***Space-Time cube*** is presented in GIS literature as a three-dimension cube that represents one time and two space dimensions. In this model time is considered to be similar to other two spatial dimensions, orthogonal to each of them. In the space-time cube, spatio-temporal objects are treated as hypothetical solids. To access information from this cube would require referencing a point, tracing a vector, slicing a cross-section, or trimming a smaller cube within the whole (larger) cube. Each of these operations would be progressively more complex, and the difficulty would increase with data volume (Langran, 1992).

Alternatively to the three approach of table, tuple and attribute 'version' in temporal databases, temporal information has been incorporated into spatial GISs in three other data models by time-stamping layers (Snapshot models, Armstrong, 1988), attributes (Space-Time Composites, Langran and Chrisman, 1988), and Spatiotemporal Objects (Worboys, 1992).

The ***Snapshot*** is an intuitively appealing space-time model. "They have roots in traditional mapping and mimic the progressive nature of a slow motion video (Langran, 1992, pp.38)." It shows the states of geographic spatial complexes at different times without implications about whether changes occur within the time step between two layers, i.e. each snapshot describes what exist at time T_i, but there is no record of how the state of T_{i+1} changes from that of T_i.

The major drawback of the snapshot model is the large amount of duplication of data with unchanged properties and the risk of data inconsistency (Langran, 1992; Yuan, 1996).

The ***Space-Time Composites*** (ST-composites, Langran and Christman, 1988) can be derived from temporal overlays of time-stamped layers (snapshots). A ST-composite describes the change of a spatial object through a period of time. Therefore, the model represents the world as a set of spatially homogenous and temporally uniform objects in a 2D space. The model permits users to treat space atemporally and time aspatially (Langran, 1992). But it fails to capture temporality between attributes across space (i.e. motion or movement) (Yuan, 1996). Since each change causes a portion of the coverage to break from its parent object to become a discrete object with its own distinct history, updating of the database requires reconstruction of ST-composites, which leads to updating both spatial objects and attribute tables.

The ***Spatiotemporal object model*** was developed by introducing an object-oriented approach (Worboys *et al.*, 1990, 1994b) in design of a temporal GIS database. Worboys (1992, 1994a) proposed that objects in geographic space could be referenced as spatiotemporal objects (ST-objects) and non-spatiotemporal objects. ST-objects represent the world as a finite collection of disjoint objects (ST-atoms). An ST-atom is a right prism, represented by an ordered pair $<S, T>$, where S is the base of the prism, representing its spatial extent, and T is the height of the prism representing its temporal extent (Worboys, 1992, 1994a; Jang *et al.*, 1994). The ST-object model is able to record changes in attributes of a ST-object in both spatial and temporal dimensions, together or separately, by projecting its ST-atoms to the spatial and/or temporal space. However, gradual changes in space through time cannot be represented in the ST-objects model since its ST-atoms are discrete.

7.2.2 Progress in spatio-temporal data models

Four basic conceptual models described in Subsection 7.2.1 are adapted to different cases.

- ***4-D hyper-cube***

The space-time cube is expanded from 2-D spatial dimensions to 3-D spatial dimensions as a 4-D hyper-cube (Mason *et al.*, 1994). Although is difficult to display the 4-D cube in a normal 3-D space, the spatio-temporal changes can be shown by projecting the data to different planes by fixing one dimension of (x, y, z or t). For example, the change of attribute values such as temperature is demonstrated by linking the voxels along the temporal line. This model facilitates interpolation along the time dimension, but extracting and tracking of objects is quite difficult.

- ***two-level temporal topological model***

The tuple-based time-stamping approach is modified to record the spatial or thematic state of the entity in a vector-based data model (Raafat *et al.*, 1994). The states of an entity are represented as the contents of relations, while the time duration of an entity is represented as tuples' intervals. The spatial states and thematic states of the entity are stored separately. Raafat *et al.* introduced the concept of a Geographical Entity Identifier (GID) into this data model. A GID was an identity attribute for a geographical entity type whose values uniquely identify each geographical entity. For example, the GID for the entity type 'highway network' may be the attribute Highway-ID. The attributes of an entity are classified into a GID relation (Master relation) and a non-GID relation (Slave relation). Time is embedded only in a master relation in order to reduce redundancy of time intervals and simplify the operation of the database management system (DBMS), and time can propagate to slave relations through the joining the slave relation with the master relation. It is claimed that there are many advantages to these two level topologies: for example, efficient data sharing, data integrity and data security, etc. For more details refer to (Raafat *et al.*, 1994).

- ***OOgeomorph***

Spatiotemporal object concepts are used in the 4-D geomorphic information system OOgeomorph (Raper and Livingstone, 1995). It is a geomorpho_info module that represents the dynamics of a coastal geomorphologic system and includes the associated geomorphologic theories. Every geomorphologic phenomenon is represented by a set of form, process, and material objects, and each object is, in turn, represented by a set of attribute objects (mform, mprocess and mmaterial). Three-dimensional location (x, y, z) and one-dimensional time (t) are referenced to objects in attributes class (att) (see Figure 7.1). OOgeomorph can handle point-based information well, but it has difficulty in processing area data and topological relationships.

OOgeomorph is similar to the 4-D hyper cube in the sense that the object in OOgeomorph is also 4-D referenced, as in 4-D hyper cube. However, in 4-D hyper cube the object concept is not easy to maintain. The object concept in Oogeomorph is different from that of the ST-object model, where the ST-atom is formed by its spatio-temporal associations. But the object in OOgeomorpho is associated by its relationship defined in geomorphologic theories, where the data and objects are structured according to physical processes and theories. This kind of structure should facilitate environmental modeling in GISs.

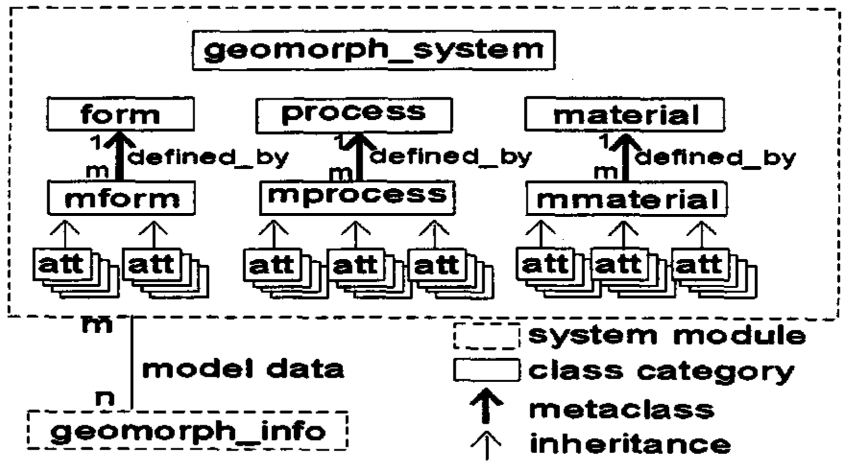


Figure 7.1 The structure of OOgeomorph (Raper and Livingstone, 1995).

- ***ESTDM***

Neither location-based nor feature-based representation is well suited for analysis of overall temporal relationships of events and patterns of changes through time, so an Event-based Spatio-Temporal Data Model (ESTDM) was proposed to facilitate this kind of analysis (Peuquet and Duan, 1995). Location in time becomes the primary organizational basis for recording change and it is represented as an 'event-list', which consists of a base state and events (changes in states). Each event is associated with several components (Figure 7.2 B), and each component stores all cells that have changed to the same value, regardless of location or previous value of the cells. A separate component structure is defined to record the cells of the component by using a run-length-encoding technique (Figure 7.2 A), in order to reduce the volume of storage space required for recording cells. For example, the cells that have changed to value 15 are recorded by (4, 16, 19), (5, 16, 20) and (6, 15, 20), which means that the cells in Row 4 from Columns 16 to 19, Row 5 from Columns 16 to 20, and Row 6 from Columns 15 to 20 change to value 15. Compared with the snapshot model, this approach is a compact representation of raster-based spatio-temporal information. The *changes* are stored in relation to a previous state rather than a snapshot of an instance. The most significant capability of ESTDM lies in its ability to perform temporal manipulations on data (e.g., temporal scale change) and temporally based comparisons in a sequential manner. However, ESTDM is raster-based and the changes are related to grid cells, therefore tracking changes that occur to pre-defined entities is quite difficult.

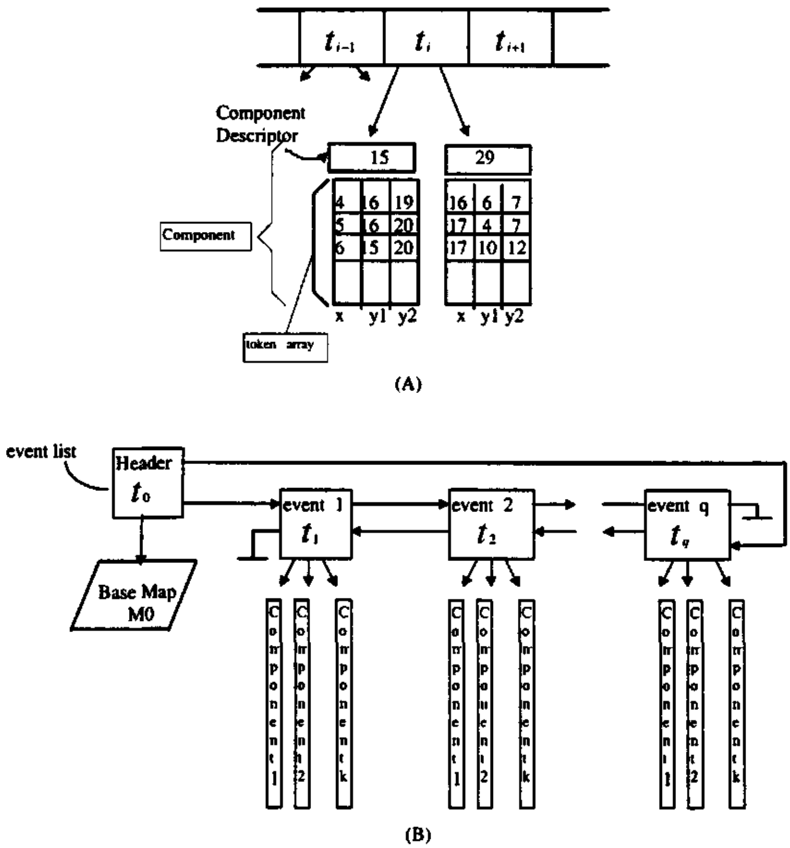


Figure 7.2 Event-based Spatio-Temporal Data Model (ESTDM) (Peuquet and Duan, 1995).

- ***Triad and Three Domain Models***

In order to track pre-defined entities and to provide a full view of changes related to different aspects of spatio-temporal objects, a *Triad Model* (Peuquet 1995; Peuquet and Qian, 1996) was proposed. The Triad Model represents three views of spatio-temporal changes: feature-based, event-based and location-based (Figure 7.3).

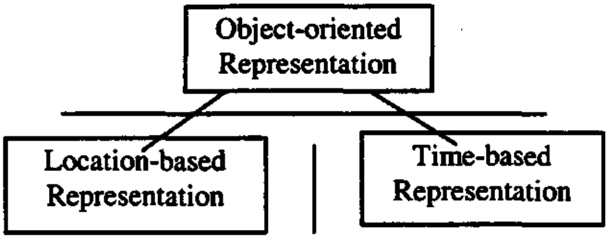


Figure 7.3 Framework of the Triad Model (adopted from Peuquet, 1995).

The feature-based representation includes four types of attributes, i.e. the generalized locational indicator, temporal intervals, nonspatio-temporal data and higher-level knowledge about known spatio-temporal phenomena. It will support modeling changes related to individual features or their components. In the event-based model, an event is stored as an observation in the time-based view. Each event and the attributes describing it are stored in their chronological order of occurrence. The attributes might include time of change, locations of change and types of objects. The location-based representation is a raster-based snapshot, which represents each pixel in the gridded array of a locational list, and supports modeling changes of individual locations.

Although the Triad model seems a great step forward in handling spatio-temporal data, the question of complexity of the querying process has not been fully answered. Furthermore, the Triad model pays attention to changes in feature, location and event, but it doesn't tell the user which process is contained in the changes nor the relationship between the features involved.

The *Three Domain* Model of semantics, time and space is quite similar to the Triad model (Figure 7.4). It was proposed to manage wildfire information (Yuan, 1995). The semantic domain consists of a wildfire's concrete or abstract concepts of aspatial and atemporal properties, such as name of individual fire event, fire intensity, fire type, or forest stand. The temporal domain consists of temporal objects of points and lines, which represents instance time and time intervals, respectively. The temporal domain supports analysis of reasoning about fire frequency or fire cycles. The spatial domain is composed of spatial objects of points, lines, polygons, cells and volumes. Each of them represents zero-, one-, or three- dimensional spatial units. It has been suggested that each domain could have its own DBMS for data storage, maintenance, and operation (Yuan, 1995). However, ways of defining and linking the semantic object, temporal objects and spatial objects, and ways of reducing the redundancy of the (unchanged and changed) data were not discussed.

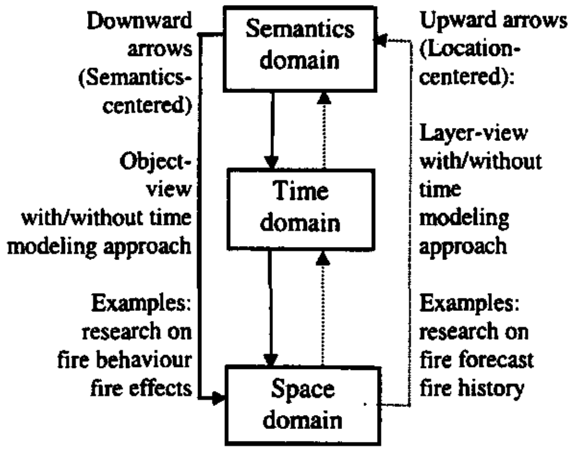


Figure 7.4 Three domain model (Yuan, 1995).

- ***RSTDM***

As some objects at different locations may share the same spatial and thematic attributes, a Rich Spatio-Temporal Data Model (*RSTDM*) was proposed by Roshannejad (1996) to reduce the data redundancy; a object is uniquely defined by its identifier (ID) and is independent from its spatial and attribute representations. The relationships between objects and their spatial and nonspatial attributes are recorded as another type of object. In this way, multiple objects can share representations, so that a considerable amount of space in the database can be saved. However, only a few man-made objects have been implemented in the *RSTDM*, e.g. buildings, roads and wells, which change suddenly and obviously, and for which there are a lot of common representations of different objects. How to adapt the model for natural phenomena, with irregular shapes, gradual changes and few common representations, has not been discussed.

7.2.3 Summary

The development of temporal data models in computer science has influenced the trend of temporal modeling in GIS. In general, most existing models treat dynamics as states through time (e.g. the Snapshot Model and two-level temporal topological model) or as differences between states, e. g. the ST-composites and ST-objects models can represent states and state changes at locations.

However, the processes involved in the development of natural phenomena, i.e. the actual characteristics of state transitions (shift, erode, expand, etc.) are not fully represented in these models, although the process-involved approach has been used in the development of the ESTDM, OOgeomorph, the Triad and Three Domain Models with a certain amount of success. None of them can precisely nor effectively model transitions, mutation and movement.

Furthermore, most approaches cannot precisely or effectively model interactions between natural phenomena, such as merge and split, which involve changes of geometric and thematic aspects of several objects at the same time. Therefore, interactions between dynamic objects are not really represented in these models.

To fully represent the process of transitions, mutation, and movement a top-down approach is needed, because behaviors of natural phenomena need to be considered prior to the design of GIS data formats and data structures for temporal information. Semantic analysis of the characteristics and behavior of natural phenomena is critical to determining a set of high-level spatio-temporal constructs (processes) to be modeled in a temporal GIS. Therefore, I follow this top-down approach and first determine the spatio-temporal constructs based upon the case presented in Chapter 6. Then I incorporate these spatio-temporal constructs into the model.

7.3 Formalizing the representation of objects and their dynamics

7.3.1 Introduction

As argued in Section 6.2, the dynamic characteristics of natural phenomena can be represented by their states and processes. Therefore, the most important spatio-temporal constructs to be modeled in the data model are 'state' and 'process' (cf. Section 6.2 for their description). In the case study in Section 6.4, the states and processes of coastal geomorphology were analyzed. Here I use them to derive a formalized representation of state and process.

Term 'object' (Smith, 1995; Hornsby and Egenhofer, 1997) is used to represent natural phenomenon such as foreshore, beach and foredune, which exist as physical entities in a specific application domain. Temporal representations in spatial databases concern the state and state transitions (process) of (complexes of) spatial objects.

An object has three aspects to represent its spatio-temporal state in terms of 'where' (geometric), 'when' (temporal) and 'what' (thematic) aspects (Cheng *et al.*, 1995; Shi and Zhang, 1996; Usery, 1996; Tang *et al.*, 1996). These three aspects constitute a complete image for the state of an object.

Normally the geometric aspects consist of topological relationships, the location of the boundary and the shape of the object. In the case study they are revealed by spatial extent of the objects (the regions which represent the spatial extent of objects at a particular epoch). The thematic description of objects has two components. The first is a set of attributes that relates to the object as such, e.g. object type and owner. The second represents an internal field, e.g. height may vary within a beach area or dune area. This can be represented by a height raster describing this internal variation. The certainty that a raster cell belongs to a beach area or to the foreshore depends mainly on the height.

Therefore, the thematic attributes of an object can be differentiated in two types: one is homogeneous for the whole area (no location dependent variability within the area); the other is non-homogenous (location dependent variability within area). The non-homogenous thematic attributes are location dependent and can be represented in a raster format, linked to spatial extent of the object.

For example, in the Ameland case (Section 6.4), the state of the beach in year 1989 is represented by the Region 2 (representing where it is) and the thematic data associated with each grid cell of this area.

As argued in Chapter 6, state transition of an object implies a process that the object is involved in. Generally speaking, three types of processes exist: (1) a pure spatial change process – affecting only geometric aspects of the object (2) a pure thematic change process – affecting only the thematic aspects; and (3) a combined spatial and thematic change process – affecting both geometric and the thematic aspects. These processes can be analyzed per object, but it is also possible to study processes in dynamic spatial complexes made up of several objects. This last case is a fourth type of process, such as split and merge.

The process related to change of homogenous thematic attributes can be represented in the form of a pure thematic change process, such as

$$change\_owner(ObjectID, OwnerID_1, OwnerID_2).$$

Because spatial extent of natural phenomena is extracted from field observation data, change of spatial extent is due to change in (the non-homogenous) thematic attributes. Therefore, for a natural phenomenon there is no pure spatial change that does not change its thematic attributes. Since the non-homogenous attributes are location related, processes that change both geometric and thematic aspects can be represented as those affecting only geometric aspects. I discuss the representation of these processes in Subsection 7.3.3.

7.3.2 Formalizing the representation of state

Molenaar (1998a) introduced a descriptor for objects that includes their geometric and thematic attributes.

$$DESCR(O_i) = \{ID(O_i), Geom(O_i), Them(O_i)\} \tag{7-1}$$

where $ID(O_i)$ is the identifier of the object, $Geom(O_i)$ is the geometric description of the object, $Them(O_i)$ is the thematic description of the object.

I use the same syntax to represent the state of an object as follows:

$$DESCR(O_{i_j}) = \{ID(O_i), Geom(O_{i_j}), Them(O_{i_j}), Time(O_{i_j})\} \tag{7-2}$$

where $ID(O_i)$ represents the identifier of the object, j is the j^{th} state of the object, and $Geom(O_{i_j})$, $Them(O_{i_j})$ and $Time(O_{i_j})$ represent the geometric, thematic and temporal components of the object O_i at state j, respectively.

For example, the state of beach in year 1989 can be represented as (see Figures 6.2 and 6.4):

$$\begin{aligned} DESCR(O_{2_{1989}}) &= DESCR(Beach_{1989}) \\ &= \{ID(O_2), \mathrm{Re}\,gion(2), Them(O_2), Time(O_2) = 1989\}. \end{aligned}$$

If we use time as the key dimension to represent the states of the object, Equation 7-2 can be modified to

$$DESCR(O_i \mid t_j) = \{ID(O_i), Geom(O_i \mid t_j), Them(O_i \mid t_j)\}. \tag{7-3}$$

7.3.3 Formalizing the representation of process

The current state of an object can be considered as a transition from its previous state (Figure 7.5), which can be described as

$$DESCR(O_i \mid t) = P_{t-1,t}(O_i) * DESCR(O_i \mid t-1) \tag{7-4}$$

where $DESCR(O_i|t)$ represents the state of object (O) at time t, $DESCR(O_i|t\text{-}1)$ represents the state of object (O) at time (t-1), and $P_{t-1,t}$ represents a state transition operator.

The four types of processes (see Subsection 7.3.1) are presented in Subsection 7.3.3.1 – 7.3.3.4.

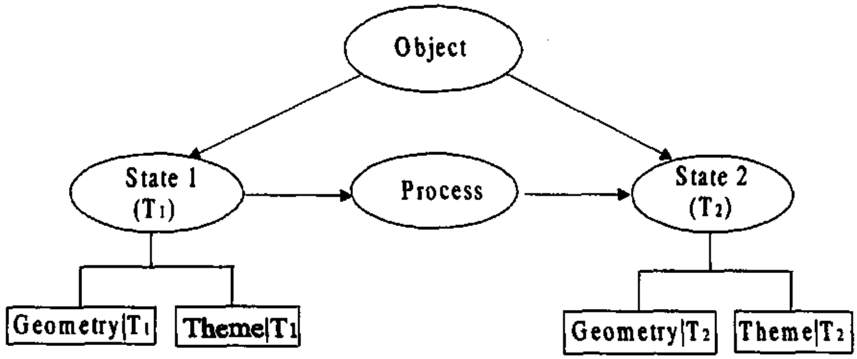


Figure 7.5 State transition of an object through process.*

Object: representing a natural phenomenon, such as foreshore, beach and foredune;

State: state of an object at a specific time, including geometric (Geometry) and thematic aspects (Theme);

Geometry: geometric aspects of an object;

Theme: thematic aspects of an object;

Process: process involving the state transition;

Time: representing the time at which the state of an object is measured (Valid), or recorded (Transaction). Both of them can be represented as point time. For the processes, Time is a period, it can be represent as a line time.

7.3.3.1 Pure spatial change process

This is the change affecting only the geometric aspects of an object; it can be described as

$$P_{t_2,t_2}(O_i) = f\{Geom(O_i \mid t_1)\} \tag{7-5}$$
$$Geom(O_i \mid t_2) = P_{t_1,t_2}(O_i) * Geom(O_i \mid t_1)$$

For example, when an object O_i is moved from $Geom(t_1)$ to $Geom(t_2)$, this process can be represented as

$$Move_{t_1,t_2}(O_i) = Move\{Geom(O_i \mid t_1), Geom(O_i \mid t_2)\}.$$

Therefore

* These definitions are also applied to Figure 7.6 to Figure 7.11.

$$\begin{aligned}
&DESCR(O_i \mid t_2) = P_{t_1,t_2}(O_i) * DESCR(O_i \mid t_1) \\
&= \{ID(O_i), Move_{t_1,t_2}(O_i) * Geom(O_i \mid t_1), Them(O_i \mid t_1)\} \\
&= \{ID(O_i), Geom(O_i \mid t_2), Them(O_i \mid t_1)\}.
\end{aligned}$$

Thus, in this case $Them(O_i \mid t_2) = Them(O_i \mid t_1)$.

7.3.3.2 Pure thematic change process

This is the change affecting only the thematic aspects of an object. It can be represented as

$$\begin{aligned}
&P_{t_1,t_2}(O_i) = f\{THEM(O_i \mid t_1)\} \\
&Them(O_i \mid t_2) = P_{t_1,t_2}(O_i) * Them(O_i \mid t_1). \qquad (7\text{-}6)
\end{aligned}$$

For example, the change of owner of a parcel of land or change of material of a road can be represented as

$$Change_{t_1,t_2}(O_i) = Change\{Them(O_i \mid t_1), Them(O_i \mid t_2)\}.$$

Therefore

$$\begin{aligned}
&DESCR(O_i \mid t_2) = P_{t_1,t_2}(O_i) * DESCR(O_i \mid t_1) \\
&= \{ID(O_i), Geom(O_i \mid t_1), Change_{t_1,t_2}(O_i) * Them(O_i \mid t_1)\} \\
&= \{ID(O_i), Geom(O_i \mid t_1), Them(O_i \mid t_2)\}.
\end{aligned}$$

Thus, in this case $Geom(O_i \mid t_2) = Geom(O_i \mid t_1)$.

7.3.3.3 Combined spatial and thematic change process

This is the change affecting both geometric and thematic aspects of an object. It can be formulized as

$$P_{t_1,t_2}(O_i) = f\{Geom(O_i \mid t_1), Them(O_i \mid t_1)\}. \qquad (7\text{-}7)$$

For example, the height of grid cells of a beach changes with sediment erosion or accumulation, which results in a shift of boundary of the beach. This can be represented as a combination of the processes discussed above as

$$DESCR(O_i \mid t_2) = P_{t_1,t_2}(O_i) * DESCR(O_i \mid t_1)$$
$$= \{ID(O_i), Shift_{t_1,t_2}(O_i) * Geom(O_i \mid t_1), Change_{t_1,t_2} * Them(O_i \mid t_1)\}$$
$$= \{ID(O_i), Geom(O_i \mid t_2), Them(O_i \mid t_2)\}.$$

7.3.3.4 multiple-object involved process

The processes discussed above are only related to a single object. Sometimes, the process might involve several objects, which can be described as

$$P_{t_1,t_2}(O_1, O_2, \cdots, O_n) = f\{DESCR(O_1 \mid t_1), DESCR(O_2 \mid t_1), \cdots, DESCR(O_n \mid t_1)\}. \tag{7-8}$$

For example, the case where two objects, O_1 and O_2, are merged into another object O_3 can be represented as

$$DESCR(O_3 \mid t_2) = Merge_{t_1,t_2}(O_1, O_2, O_3) * \{DESCR(O_1 \mid t_1), DESCR(O_2 \mid t_1)\}$$
$$= Merge_{t_1,t_2}\{(ID(O_1), ID(O_2), ID(O_3); Geom(O_1 \mid t_1), Geom(O_2 \mid t_1),$$
$$Geom(O_3 \mid t_2); Them(O_1 \mid t_1), Them(O_2 \mid t_1), Them(O_3 \mid t_2)\}$$
$$= \{ID(O_3), Geom(O_3 \mid t_2), Them(O_3 \mid t_2)\}.$$

Equation 7-8 can be adapted further to express processes involved in different spatial and thematic aspects.

As argued above, there is no pure spatial change, so that a process affecting both geometric and thematic aspects can be simplified as affecting only geometric aspects. For example, the merge process of two objects can be represented as

$$P_{t_1,t_2}(O_1, O_2, O_3) = Merge\{ID(O_1), ID(O_2), ID(O_3);$$
$$Geom(O_1 \mid t_1), Geom(O_2 \mid t_1), Geom(O_3 \mid t_3)\}.$$

The thematic aspects, i.e. internal height variations within the area, which are associated with the geometric aspects, do not have to be explicitly expressed in the process. In this way the description form of the process is simplified.

7.3.3.5 Discussions

The semantic descriptions of spatio-temporal processes were discussed by Claramunt and Thérlault (1996). Three main classes of basic spatio-temporal processes were proposed to distinguish evolving and mutating objects:

(1) *Evolution* of a single entity represents basic changes, transformation and movements of an entity, which includes processes of Appearance, Disappearance, Stability, Expansion and Displacement.
(2) The *Functional* process represents change relating to several entities. It includes the processes of Succession, Permutation, Production, Reproduction, and Transmission.
(3) The *Restructuring* process represents spatial topologic change involving several entities. It includes the processes of Split, Union and Reallocation.

The taxonomy listed above are catalogued according to spatial aspects of the objects. The evolution process can be considered as change related to the geometric aspects of a single object and can be represented by the formalism of Subsection 7.3.2.1. The functional and restructuring processes can be considered as multiple entities involved processes. Both situations can be represented in the formalism of Subsection 7.3.2.4. However, thematic change, which might occur as well, is not described in this taxonomy. Therefore the syntaxes proposed in Subsections 7.3.3.1 to 7.3.3.4 are general representations of processes.

Al-Taha and Barrera (1994) proposed temporal constructs to track object identifiers. The processes between objects are described by eleven operators, such as CREATE and DESTROY for creating of a new identity and for disposing, forever, of the essence of an object. Other operations are KILL, REINCARNATE, IDENTIFY, EVOLUTION, FUSION, FISSION, SPWAN, AGGREGATE and DISAGGREGATE. The work of Claramunt and Thérlault (1996) and Al-Taha and Barrera (1994) provides approaches for describing the changes of objects in processes in an object-oriented way. They were unable, however, to explain how to link the processes with other data in the spatio-temporal database.

7.4 A process-oriented spatio-temporal data model – the Star Model

The formalized description of the dynamics of objects suggests that five spatio-temporal constructs should be defined in the spatio-temporal data model to represent the objects and their dynamics. They are '*Object*', '*Time*', '*Process*', '*Geometry*' and '*Theme*'. The extended entity relationships (EER) (Hughes, 1991) between these five entity types are illustrated by Figure 7.6. As the figure can be generalized as a star style model, I have called it the '*Star M*odel'.

These five entity types are defined as follows (with primary key attributes underlined):

1. Entity type *Object* with attribute <u>Oid</u> (a unique identifier).
2. Entity type *Time* with attributes <u>Tid</u> (a unique identifier) and Value (a

unique value, e.g. year 1998). It has two subclasses, valid time and transaction time. Both of them can be a line time (period) or a point time. In my case, I only consider the valid time.

3. Entity type *Process* with attributes Pid (a unique identifier) and Ptype (type of process, e.g. as shift, merge).
4. Entity type *Theme* with attributes Ttype (type of the attribute, e.g. owner, material), TtypeNum (the number of a class in a Ttype) and Tvalue (e.g. the name of the owner).
5. Entity type *Geometry* with attributes Gid (a unique identifier). It has two subclasses, *Simplex* and *Complex. Simplex* can be further classified as Grid-element, Node, Arc, Triangle and Voxel. *Complex* can be further classified as Point, Line, Surface and Body. Here we only consider the Grid and Surface (i.e. Region) for the Geometry. The description of Region and Grid is different and is discussed in Chapter 8.

The relationships between the entities are as follows:

1. *EXIST* relationship

1.1 The N:1 relationship *EXIST* between *Object* and *Time*. It includes two parts, 'Start_at' and 'End_at'. The membership class of this relationship is mandatory for *Object*.

1.2 The N:1 relationship *EXIST* between *Geometry* and *Time*. The membership class of this relationship is mandatory for *Geometry* if visualizing the spatial distribution of objects at different epochs is required.

2. *ENVOLVED_IN* relationship

2.1 The M:N relationship *ENVOLVED_IN* between *Object* and *Process*.

2.2 The M:N relationship *ENVOLVED_IN* between *Geometry* and *Process*.

2.3 The M:N relationship *ENVOLVED_IN* between *Theme* and *Process*.

3. *HAS* relationship

3.1 The M:N relationship *HAS* between *Object* and *Theme*.

3.2 The M:N relationship *HAS* between *Object* and *Geometry* (*Complex*).

4. The N:1 relationship *HAPPEN_AT* between *Process* and *Time*. The membership class of this relationship is mandatory for *Process*.

5. *HAS* relationship

5.1 The N:1 relationship *HAS* between *Geomtery* (*Simplex*) and *Theme* (non-homogenous), e.g. the relationship between the grid cell and height value.

5.2 The N:1 relationship *HAS* between *Geomtery* (*Complex*) and *Theme* (homogenous), e.g. the relationship between region and vegetation type.

5.3 The N:1 relationship *HAS* between *Complex* and *Simplex*.

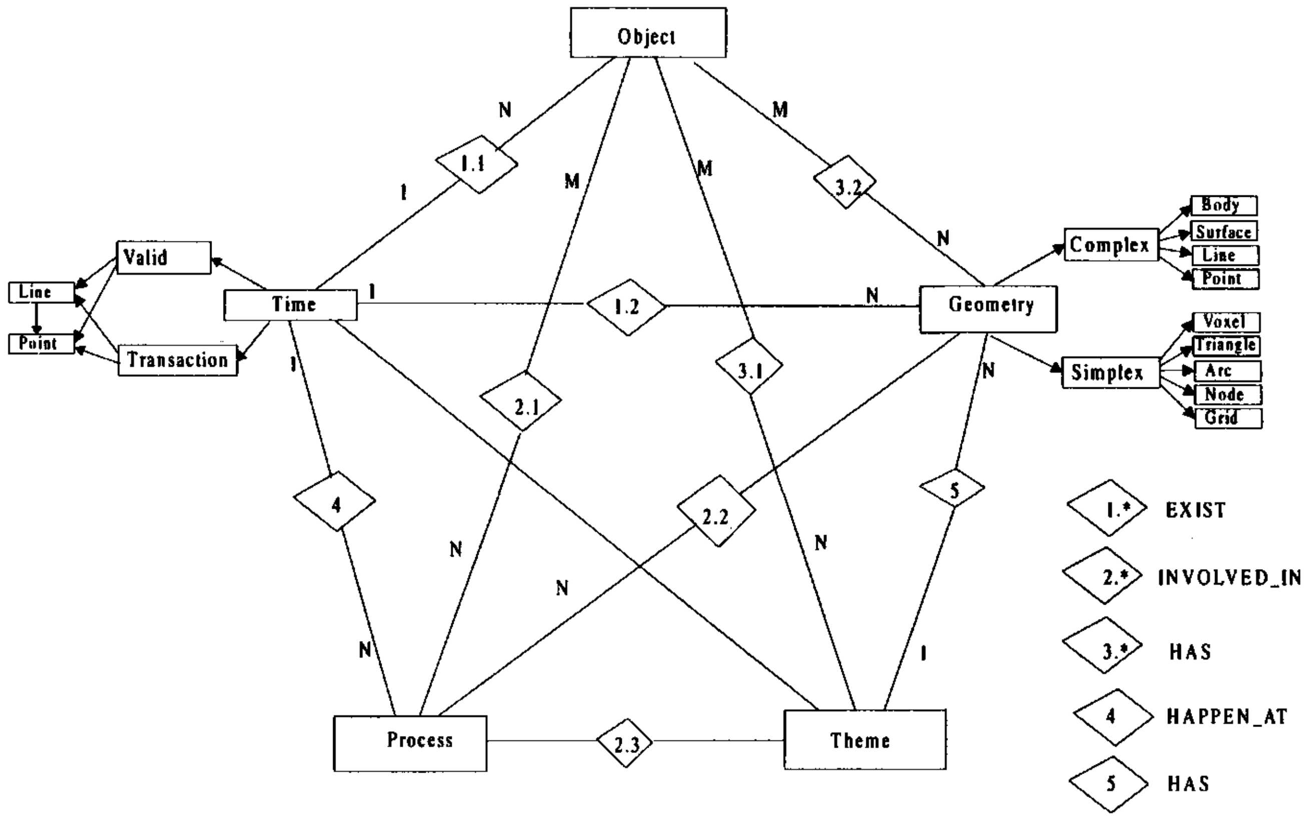


Figure 7.6 Entity Relationship Model – the Star Model

Simplex: the primitive element of an object; *Complex*: combination of simplex to compose complex object.
For other symbols refer to Figure 7.5.

Each entity type in the Star Model is connected to the other four. Such a kind of structure can support queries and analysis based on different perspectives, such as time-oriented, object-oriented, location-oriented and process-oriented. However, as formalized in Section 7.3, the process conveys the geometric and/or thematic information of the object that participates in the process, which represents a particular state of the object. Thus, the states of objects can be revealed through the processes i.e. explicit links of objects to its states (cf. Figure 7.5) are not necessary, but can be built through the processes (see Figure 7.7). The basic advantage of this approach is that such a structure represents the relationships (interaction) between the objects. Since the process is explicitly represented, the knowledge can facilitate physical environmental modeling.

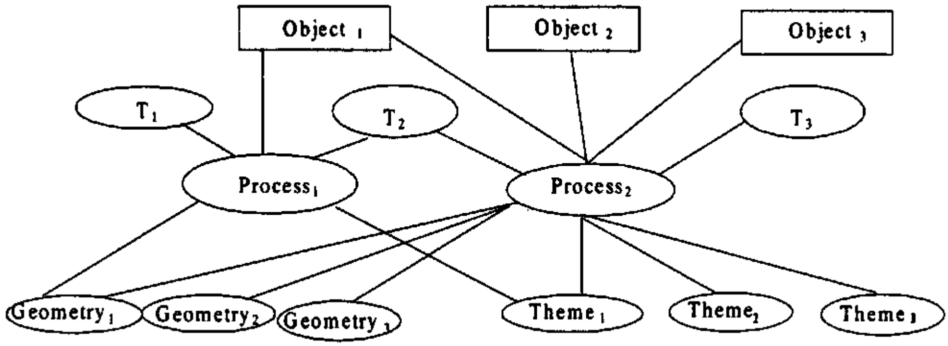


Figure 7.7 Objects are linked to their states through processes.

Because the processes can serve as a central link between an object and its (dynamic) states, the Star Model can be restructured as Figure 7.8, which shows that the *Object* is not explicitly linked to *Geometry* and *Theme* as in Figure 7.6, but through *Process*. Thus this model can be seen as a process-oriented model, as all other entity types can be connected through the process. The process can describe when, where and what thematic values change for which objects (see Figure 7.9). Furthermore, the interaction of objects can be described if two or more objects are linked by one process.

By distinguishing two types of the thematic attributes, Figure 7.9 can be further modified into Figure 7.10, which shows that changes of homogenous thematic attributes are linked to objects through processes; the non-homogenous thematic attributes are represented based upon raster elements and linked to an object through geometry and process.

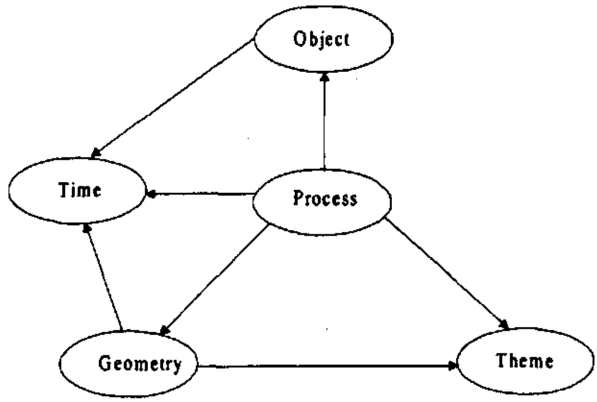


Figure 7.8 Restructuring of the Star Model – a process-oriented view.

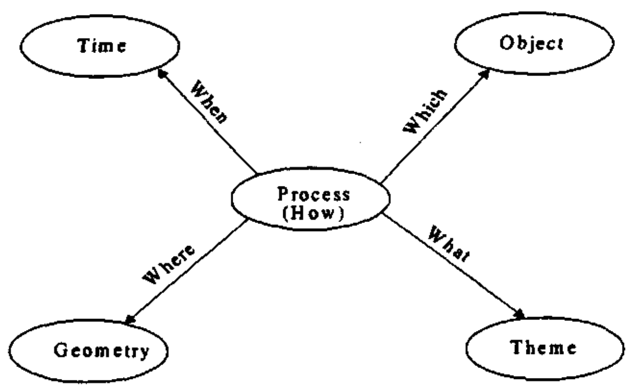


Figure 7.9 Links of process to other components

7.5 Summary and discussion

This chapter describes a new type of spatio-temporal data model, called the process-oriented spatio-temporal data model (the Star Model). Unlike existing approaches used in GIS, the Star Model is designed to explicitly represent change over space relative to time. Furthermore, the interactions between objects are also represented as processes. The model can represent dynamic processes affecting the spatial and thematic aspects of individual objects and object complexes. Because the Star Model explicitly stores change (process) relative to time, procedures for answering queries relating to temporal relationships, as well as analytical tasks for comparing different sequences of change, are greatly facilitated.

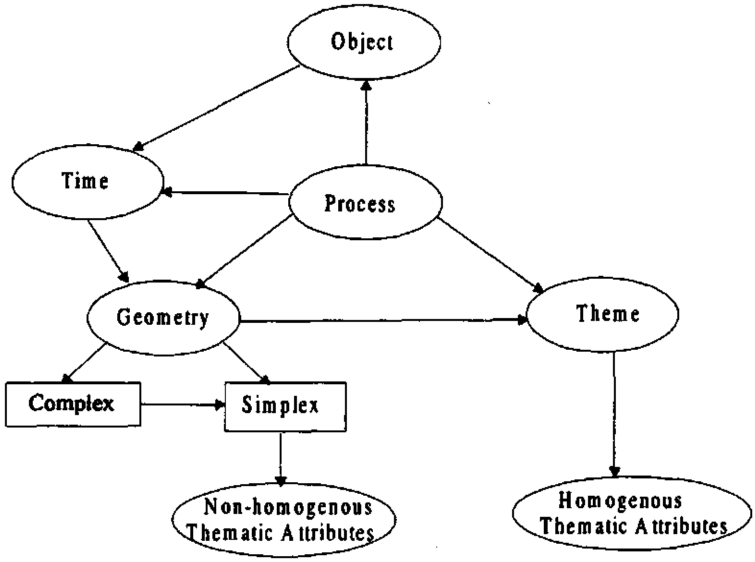


Figure 7.10 A process-oriented spatio-temporal data model.

The Star Model could be considered as a *general* model that is able to represent reality from object-based, location-based, time-based, and process-based perspectives. Other models can be derived from the model, i.e. they can be seen as special cases of this model. For example, the Triad Model (Peuquet and Qian, 1996) and the Three Domain Model (Yuan, 1995) can be considered as special cases of Star Model when we only consider the object, spatial and temporal aspects. The event-based model ESTDM can be considered as an event-based view of the Star Model.

Multi-strands of time can be generated in the Star Model, each representing the (spatio-temporal) lifeline of an object. A life line may split in two when an object splits, as occurred in years 1989–1990 for Objects 3 and 4 and in years 1991–1992 for Objects 1 and 5 (Figure 6.3B). Lifelines may also merge when objects merge, as in years 1990–1991 for Objects 3 and 4 in Figure 6.3B, see also Figure 7.11.

The logical design and implementation of the Star Model is discussed in Chapter 8.

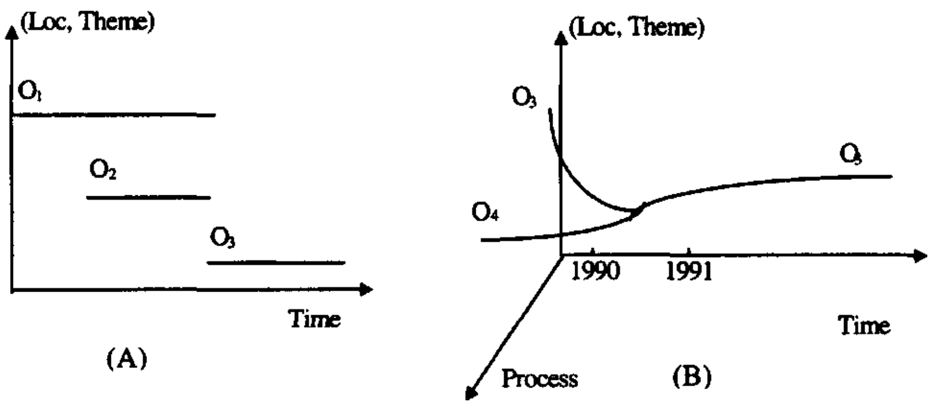


Figure 7.11 Lifelines of spatio-temporal objects.
(A) The lifelines of objects are usually parallel in other data models.
(B) The lifelines of objects in the Star Model can merge and split.

Chapter 8

Logic Design and Implementation of the Star Model

8.1 Introduction

In the previous chapter, the conceptual model – the Star Model – was derived from the extended entity relationship (EER) diagram. This chapter explains the logical design and implementation.

Although I have taken an object-oriented approach for the conceptual design of the Star Model, here I use a relational approach for its logic design and implementation since many commercial database management system (DBMS) based on relational models are widely available. Therefore, I will discuss the process of translating the EER diagrams to a relational schema.

This chapter is organized as follows. Section 8.2 explains the process of translating the EER diagram to a relational scheme. This process is illustrated using the Ameland case (cf. Section 6.4). Section 8.3 presents the database of the coastal geomorphology landscape units. Section 8.4 discusses the multi-perspective queries supported by the proposed data model. Section 8.5 illustrates the user-interface designed for the multi-perspective queries based upon ArcView, and Section 8.6, the last section, summarizes the chapter and discusses further issues related to the logical design and implementation.

8.2 Translation of EER to relational structure

8.2.1 Translation of entity types

Each entity type is mapped to a relational scheme. There are five basic relational schemes, which contains the basic description of the entity types.

- *Object* (<u>Oid</u>), where Oid is the identifier of an object in *Object* type;
- *Time* (<u>Tid</u>, Value), where Tid and Value are the identifier and the value of a time object of *Time* type;

- *Process* (Pid, Ptype), where Pid and Ptype are the identifier and the type of a process in *Process* type;
- *Geometry* (Gid), where Gis is the identifier of a geometry of *Geometry* type;

  The *Geometry* entity type has two subclasses: *Simplex* and *Complex*. In our case, *Simplex* is the grid cell and *Complex* is the region, which represents the spatial extent of the objects. The relation schemes of these two subclasses are:

  Simplex(Gridid)
  Complex(Rid, Area, Volume)

  where Gridid is the identifier of a grid cell and Rid is the identifier of a region.

- *Theme* (Ttype, TtypeNum, Tvalue), where Ttype is the type, TtypeNum is the number and Tvalue is the Value of a theme object of Theme type.

  The entity type *Theme* has two subclasses:

  The homogeneous *Theme*, *HTheme*(HTtype, HTtypeNum, HTvalue). The HType can be the owner, material, or land cover type. Httype is the number of a class belonging to a HTtype and Htvalue is its value. For example, in the Ameland case, only land cover type (Ctype) is considered, so HttypeNum has three numbers (1, 2, 3) with Htvalue of (foreshore, beach and foredune). Therefore, the description of H*Theme* is simplified as (CtypeNum, Tvlaue) for the case (cf. Table 8-3).

  The non-homogeneous *Theme*, *ITheme*(ITtype, ITtypeNum, ITvalue). For the case described in Chapter 2 (Ameland), the ITtype can be height or fuzzy membership function value of a grid cell belonging to a class. Different ITvalues will have different ITtypeNums. Since the ITvalues are location-related, I record them according to Grid cell, which implies that ITtypeNum is converted to Gridid, TTvalue is converted to Height and MFvalue in the description of *Simplex*. Therefore, the entity type of *ITheme* is not recorded explicitly, but in the description of *Simplex* (Relation 5.2 of Figure 7.6, see also the translation in Subsection 8.2.2).

8.2.2 Translation of entity relationships

The translation of entity types (see Figure 7.6) obeys the following guidelines:

(1) cardinality N:1 – define a foreign key in the relation schema corresponding to the entity type at the n-side.

- Translate Relation 1.1 between *Object* and *Time* by adding foreign keys in the relation scheme of *Object*,
  Object(Oid, Tid$_1$ (Start_time), Tid$_2$ (End_Time)).

- Translate part of Relation 3.1 between *Object* and (invariant, homogeneous) *Theme* by adding foreign keys in the relation scheme of *Object*,
  Object(Oid, CtypeNum, Tid1 (Start_Time), Tid2 (End_Time)).

- Translate Relation 1.2 between *Geometry* and *Time* by adding foreign keys in the relation scheme of *Geometry*,
  Geometry(Gid, Tid (Exist_Time)).

- Translate Relation 4 between *Process* and *Time* by adding foreign keys in the relation scheme of *Process*,
  Process(Pid, Ptype, Tid$_1$ (Start_Time), Tid$_2$ (End_Time)).

- Translate Relation 5.1 between *Complex* and *Theme* by adding foreign keys in the relation scheme of *Complex*,
  Complex(Rid, Tid, HTtypeNum, Area, Volume).

- Translate Relation 5.2 between *Simplex* and *Theme* by adding foreign keys in the relation scheme of *Simplex*,
  Simplex(Gridid, Tid, Height, MFvalue).

- Translate Relation 5.3 between *Complex* and *Simplex* by adding foreign keys in the relation scheme of *Simplex*,
  Simplex(Gridid, Tid, Height, MFvalue, Rid).

(2) cardinality M:N

The relations of Figure 7.6 that have not been translated are the M:N cardinality Relations 2.1, 2.2, 2.3, 3.1 and 3.2. All these relations are related to *Process*. So I use a foreign key in the description of *Process* to represent a relation. This foreign key is a list of *Object*s (Relation 2.1), or a list of *Complex*es (Relation 2.2) or a list of *Theme*s (Relation 2.3).

Therefore, Relations 2.1, 2.2 and 2.3 are mapped as follows:

2.1 O*Process*(Pid, Ptype, <u>OidList</u>, Start_Time, End_Time)
2.2 R*Process*(Pid, Ptype, <u>RidList</u>, Start_Time, End_Time)
2.3 T*Process*(Pid, Ptype, <u>HThemeList</u>, Start_Time, End_Time).

Relation 3.1 can be represented by combining Relations 2.1 and 2.3, as
*Process*1(Pid, Ptype, OidList, HThemelist, Start_Time, End_Time).

Relation 3.2 can be represented by combining Relations 2.1 and 2.2, as
*Process*2(Pid, Ptype, OidList, RidList, Start_Time, End_Time).

The scheme listed above provides a general framework for translating the Star Model into a relational scheme. With regarding to the case discussed in Chapter 6, some modification of the structure of the scheme has been made.

Firstly, according to the procedure of data processing (cf. Figures 5.1 and 6.1), HTtypeNum of *Complex* (regions) was created by clustering the cells belonging to a same class into regions; CtypeNum of *Object* was then created after identification of the object. Therefore, CtypeNum was obtained from HTtypeNum, which was originally created for the cells after classification. After identification of the objects, HTtypeNum = CtypeNum. In order to avoid data redundancy, HTtypeNum of *Complex* and CtypeNum of *Object* are removed from the schemes and converted to CtypeNum of *Simplex*, because CtypeNum was created first for grid cell and both the segmentation and identification need this information of the grid cells.

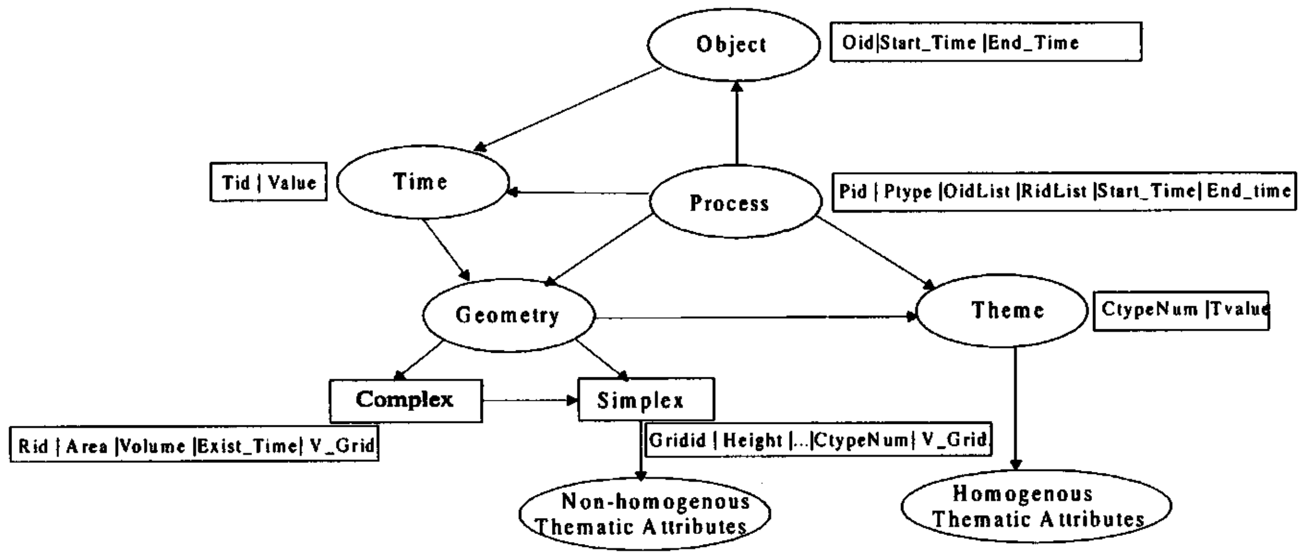
Secondly, the link between *Simplex* and *Complex*, Rid of *Simplex*, is replaced by V_Grid of *Simplex* and V_Grid of *Complex*. V_Grid is the label number of the region created each year after segmentation but before merging (see Tables 8-4, 8-5), which is only unique for regions created in the same year.

Thirdly, *Simplex* is recorded annually. So Tid of *Simplex* is removed.

Fourthly, only the second form of *Process* (Process2) is available in our case, so it alone will be discussed. It can be seen from the Table 8-6 that each object in the OidList and each region in the RidList is stored as one record in the table (cf. Subsection 8.2.2).

Finally, the following relational schemes were built for the case (see Figure 8.1):

- *Time*(<u>Tid</u>, Value)
  As for *Time*, I only define the valid time, i.e. the time at which the natural phenomena were measured. Each *Time* object will have an identifier (Tid) and a value.



Oid: Object identifier
Pid: Process identifier
Tid: Time identifier
Rid: Region identifier
Gid: Grid element identifier
Start_Time: Starting time
End_Time:Finishing time
Exist_Time: Existing time

Ptype: Process type such as shift, erode, merge, etc.
Ctype: Class type of regions such as foreshore, beach and foredune
CtypeNum: Number of class type
Tvalue: Name of the class
OidList: Objects participate in the process
RidList: Regions participating in the process
Height: Interpolated height value of a grid cell
V_Grid: Label of a region

Figure 8.1 Data structure of the Star Model.

- H*Theme*(CtypeNum, TValue)

  HTheme records the information about the land cover types. It includes a type number (CtypeNum) and its value (Tvalue) (cf. Table 8-3).

- *Object*(Oid, Tid_1(Start_Time), Tid_2(End_Time))

  Object should contain the information on its identifier (Oid), its starting time (coming into existence) and its ending time (disappearing time). The starting time (Start_Time) and ending time (End_Time) will be updated.

- *Simplex*(Gridid, Height, MFvalue, CtypeNum, V_Grid)

  Simplex represents information related to the grid cells. It records the non-homogeneous thematic attributes, such as height value and uncertainty of belonging to *Complex* (MFvalue). As explain above, the class type of a cell is also recorded. The links of *Simplex* and *Complex* are kept through V_Grid.

- *Complex*(Rid, Area, Volume, Exist_Time, V_Grid)

  Complex describes the spatial extent of natural phenomena, i.e. regions extracted in different years. The identifier (Rid), its fuzzy area (Area), volume (Volume), the existing time (the measuring time, Exist_Time) and its link to *Simplex* are recorded through V_Grid.

- *Process*(Pid, Ptype, OidList, RidList, Start_Time, End_Time)

  Only the *process* related to Objects and Regions are discussed in the case. The description contains the *process* identifier (Pid), its type (Ptype), the list of Objects (OidList) and Regions (RidList), the starting time (Start_Time) and ending time (End_Time). The number of objects in the OidList and regions in RidList should be the same, in order to built a unique link (one-to-one) between an object and a region.

  For example, the *process*es illustrated in Figure 6.2 can be represented as

Process(1, 'Shift', OidList(O_1); RidList(Rid_1, Rid_2*); t_1, t_2)*
Process(2, 'Merge', OidList(O_1, O_2, O_3); RidList(Rid_1, Rid_2, Rid_3); t_1, t_2)
Process(3, 'Split',OidList(O_1, O_2, O_3); RidList(Rid_1, Rid_2, Rid_3); t_1, t_2)
Process(4, 'Expand',OidList(O_i); RidList(Rid_1, Rid_2); t_1, t_2)
Process(5, 'Shrink',OidList(Oi); RidList(Rid_1, Rid_2); t_1, t_2)
Process(6, 'Appear', OidList(O_i); RidList(Rid_1); t_1, t_2)
Process(7, 'Disappear', OidList(O_i); RidList(Rid_1); t_1, t_2)

where *O* represents an object identifier, *Rid* is the region representing spatial extent of the object at different states, t_1 is the starting time and t_2 is the ending time of the *process*. For example, the split *process* of Object 3 from 1989 to 1990 can be represented as

Process(4, 'Split', OidList(3, 3, 4); *RidList*(3, 6,7); 1989, 1990);

which means Object 3 (foredune, represented by Region 3 in 1989) was split into Object 3 (represented by Region 6 in 1990) and Object 4 (a special form of foredune, represented by Region 7 in 1990) during 1989 to 1990.

8.3 Database of the Ameland case

The case presented in Chapter 6 was applied to test the Star Model. Seven year data, from years 1989 to 1995, are available for the case study. The following tables were built based upon the case. The table of *Time* is shown in Table 8-1.

Table 8-1 Table of *Time*.

Tid	Value
1	1/1/89
2	1/1/90
3	1/1/91
4	1/1/92
5	1/1/93
6	1/1/94
7	1/1/95

There are five objects in total from years 1989 to 1995 (Table 8-2). Their existing times are listed in the column of (Start_Time) and (End_Time). The values of these two columns correspond to the values of time in Table 8-1 (*Time*).

The homogenous thematic attributes for regions and objects are the class types; three classes of landscape units are stored in Table 8-3.

The non-homogeneous thematic attributes are stored for each year in a table of *Simplex* (e.g. Table 8-4). Table 8-4 shows that each grid cell contains its identifier, the height value (Height), the membership function value (MFvalue) of the cell belonging to *complex*, and the label of the corresponding *complex* (V_Grid) (see also Table 8-5).

Table 8-2 Table of *Objects*.

[illegible]	[illegible]	[illegible]
1	1	7
2	1	7
3	1	7
4	2	3
5	4	7

Table 8-3 Table of *HTheme*.

[illegible]	[illegible]
1	Foreshore
2	Beach
3	Foredune

Table 8-4 Part of the table of *Simplex* in 1989.

[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
1	-4.5273	0.8682	1	1
2	-4.4366	0.8908	1	1
3	-5.4110	0.6473	1	1
4	-5.2175	0.6956	1	1
5	-4.9571	0.7607	1	1
6	-4.7322	0.8170	1	1
7	-4.5367	0.8658	1	1
8	-4.2969	0.9254	1	1
9	-4.1167	0.9661	1	1
...	...	...	...	...
1280	-0.8566	0.8790	2	2
...	...	...	...	...
3140	2.862	0.9870	3	4

There are in total 18 regions appearing in years 1989 to 1995. Their information is stored in Table 8-5. CtypeNum corresponds to Table 8-3, where 1 means foreshore, 2 means beach and 3 means foredune. V_Grid corresponds to Table 8-4.

The *processes* are stored in Table 8-6; the interaction between objects and development of the objects are described.

All the tables are stored in Microsoft Access format.

Table 8-5 Part of the table of *Complex*.

[illegible]	[illegible]	Volume	[illegible]	[illegible]
1	1,108.10	29,144.20	1	1
2	1,264.79	38,156.10	1	2
3	644.26	21,920.50	1	4
4	1,138.71	30,178.00	2	1
5	1,229.66	37,631.10	2	2
6	586.79	19,883.90	2	3
7	27.95	889.64	2	4
8	1,101.31	29,052.00	3	1
9	1,260.11	38,560.20	3	2
10	609.83	20,852.40	3	4
11	1,004.85	26,202.80	4	1
12	1,288.68	39,455.00	4	2
13	6.42	184.87	4	7
14	625.70	21,286.40	4	8
15	901.79	2,971.20	5	1
16	1,137.31	34,995.90	5	2
17	52.10	1,423.16	5	3
18	610.86	20,777.80	5	4

8.4 Multi-perspective queries

As argued in the previous chapter (Subsection 7.2.3), to be able to analyze the behavior of natural phenomena, the spatio-temporal model should support analysis of data in time series from varying perspectives, through location-oriented, time-oriented, object-oriented and process-oriented queries. Under the structure defined in Tables 8-1 to 8-6, such multi-perspective queries are possible.

- *time-oriented* view. This supports the analysis and queries according to time sequence. For example:

  "Which objects existed at time 1/1/91?"

  SELECT Oid
  FROM *object*
  WHERE ((*object*.S_Time)<#1/1/91#)
  AND((*object*.E_Time)>#1/1/91#));

Table 8-6 Part of the table of *Process*.

Pid	Ptype	Oid1	Oid2	Oid3	Rid1	Rid2	Rid3	Start Time	End Time
1	Shift	1			1	4		1	2
2	Shift	2			2	5		1	2
3	Shift	3			3	6		1	2
4	Split	3	3	4	3	6	7	1	2
5	Appear		4			7		1	2
6	Shift	1			4	8		2	3
7	Shift	2			5	9		2	3
8	Shift	3			6	10		2	3
9	Disappear	4			7			2	3
10	Merge	3	4	3	6	7	10	2	3
11	Shift	1			8	11		3	4
12	Shift	2			9	12		3	4
13	Appear		5			13		3	4
14	Shift	3			10	14		3	4
15	Split	1	1	5	8	11	13	3	4
16	Shift	1			11	15		4	5
17	Shift	2			12	16		4	5
18	Shift	3			14	18		4	5
19	Expand	5			13	17		4	5

"Which spatial objects existed at time 1/1/91?"

SELECT Rid
FROM *complex*
WHERE (*complex*.Time)=#1/1/91#;

"Which processes occured during a time period (#1/1/91#, #1/1/92#)

SELECT *
FROM *process*
WHERE ((*process*.Start_Time)=#1/1/91#) AND ((*process*.End_Time)=#1/1/92#));

- *object-oriented* view. This supports the analysis based on the objects, including their spatial extent, the processes, their thematic values, and the time span of their existence. For example:

"When did Object 4 exist?"

SELECT Start_Time, End_Time
FROM *object*
WHERE (*object*.Oid)=4;

"What is the development of Object 3 during time period (#1/1/90#, #1/1/1992)?"

SELECT *
FROM *process*
WHERE (((*process*.Oid_1)=3) OR ((*process*.Oid_2)=3) OR ((*process*.Oid_3)=3)) AND (((*process*.Start_Time)=#1/1/90#) OR ((*process*.End_Time)=#1/1/92#));

"Where is Object 3 at time t?"

For this query, several steps have been implemented:

(1) find the region that is the spatial extent of Object 3 at time #1/1/90#;
SELECT Rid_1, Rid_2, Rid_3
FROM *process*
WHERE ((*process*.Oid_1)=3) OR ((*process*.Oid_2)=3) OR ((*process*.Oid_3)=3)) AND ((*process*.Start_Time)=#1/1/90#));

(2) find the grid cells that belong to this region;
SELECT Gridid
FROM *process*, *simplex*, *complex*
WHERE ((*process*.Rid_1=*complex*.Rid) AND ((*complex*.V_ID)=*simplex*.V_ID);

(here I assume *process*.Oid_1=3)
(3) visualize the query result

"What's the change of the object at time t_1 to t_2 (at different certainty level)?"

For this query, several steps have been implemented:
(1) find out where the object is at time t_1 and time t_2, respectively;
(2) calculate their change (difference);
(3) calculate the uncertainty of change;
(4) (optional: calculate the change area at different certainty level);
(5) visualize the change with certainty.

- *process-oriented* view. This is an event-based view, which tells users what happened/changed during which time period.
  a) "Where/When did process *P* occur from t_1 to t_2?"
  b) "Which object *O* was involved in a process *P* (from t_1 to t_2)?"
  c) "Which process was followed by the process of 'merge'? "

For example:

"When did the *process* 'split' happen?"

SELECT Start_time, End_Time
FROM *process*
WHERE (*process*.Ptype) = 'split';

"Which object was involved in the process 'merge'?"

SELECT Oid
FROM *process*
WHERE (*process*.Ptype) = 'merge';

"What happened three years after the process 'split' occurred?"

For this query, two steps are required:

(1) find the time at which the process 'split' occurred;
SELECT End_Time
FROM *process*
WHERE (*process*.Ptype) = 'split';

(2) find which processes occurred within three years of the found in Query (1);
SELECT *

FROM *process*
WHERE ((*process*.Start_Time) ≥ *t*)
and (*process*.End_time) ≤ *t* +3);
(here I assume the query result of Step 1 is *t*)

- *location*-oriented view. This provides analysis and queries related to locations,

  "What is the height value of a given location (when)"
  SELECT Height
  FROM *simplex*
  WHERE Time = #1/1/91;

- Combined queries

  "Which spatial object is split into other spatial objects at time 1/1/91?"
  SELECT Rid
  FROM *complex, process*
  WHERE ((*process*.Start_Time) = #1/01/91#)
  AND ((*process*.Ptype) = 'split)
  AND ((*complex*.Rid) = (*process*.Rid_1));

  "Which object is split into other objects at time 1/1/91?"
  SELECT *
  FROM *object, process*
  WHERE ((*process*.Start_Time) = #1/01/91#)
  AND ((*process*.Ptype) = 'split')
  AND ((*object*.Oid) = (*process*.Oid_1));

8.5 User interface

To simplify the operation for users to query and to visualize the query result, a window-based query prototype was implemented based upon the functions of Spatial Analyst and Avenue provided by ArcView.

Since the data stored in Microsoft Access format are compatible with ArcView, they can be manipulated directly in ArcView. The main emphasis of the implementation of the user interface was organizing the queries according to the different viewpoints described in Section 8.4. All query and visualization functions were implemented by Avenue, the programming language of ArcView.

Figure 8.2 shows the user interface of multi-perspective queries. The user interface of an object-oriented query is shown in Figure 8.3. Figure 8.4 shows the query result of the change of Object 1 (foreshore) between two years (1989 to 1990). Figure 8.4 shows how the change of the fuzzy object can be visualized by displaying its changed spatial extent and its fuzziness. If changes of an object are displayed for many years, the developing trend of the object can be generated. This feature facilitates the predication of landscape units in the future.

Figures 8.5 and 8.6 illustrate the querying result of the process Object 3 underwent from year 1989 to 1991. It can be seen from Figure 8.4 that Object 3 (foredune, represented by Area 3 in 1989) split into itself (represented by Area 6 in 1990) and a new Object 4 (a special form of foredune, represented by Area 7 in 1990) during 1989 to 1990. Figure 8.6 presents the description of the processes.

8.6 Summary and discussion

This chapter presents the logical design and implementation of the Star Model. The process for translating the EER diagram to relational tables has been explained in detail. The structures of the tables were presented for the Ameland case. A description of the implementation of a prototype of multi-perspective queries based upon ArcView is given. This prototype supports queries from multi-perspectives, for example, object-oriented, time-oriented, process-oriented and location-oriented. This feature facilitates the spatio-temporal modeling so that the user can analyze natural phenomena and their changes in multiple ways.

There are several issues related to the logical design and implementation for the Star Model:

(1) the non-normalization of tables

I have not normalized the table of process (Table 8-6) because I consider the number of objects that participate in the *process*es to differ from case to case and from domain to domain. This structure is flexible for other applications.

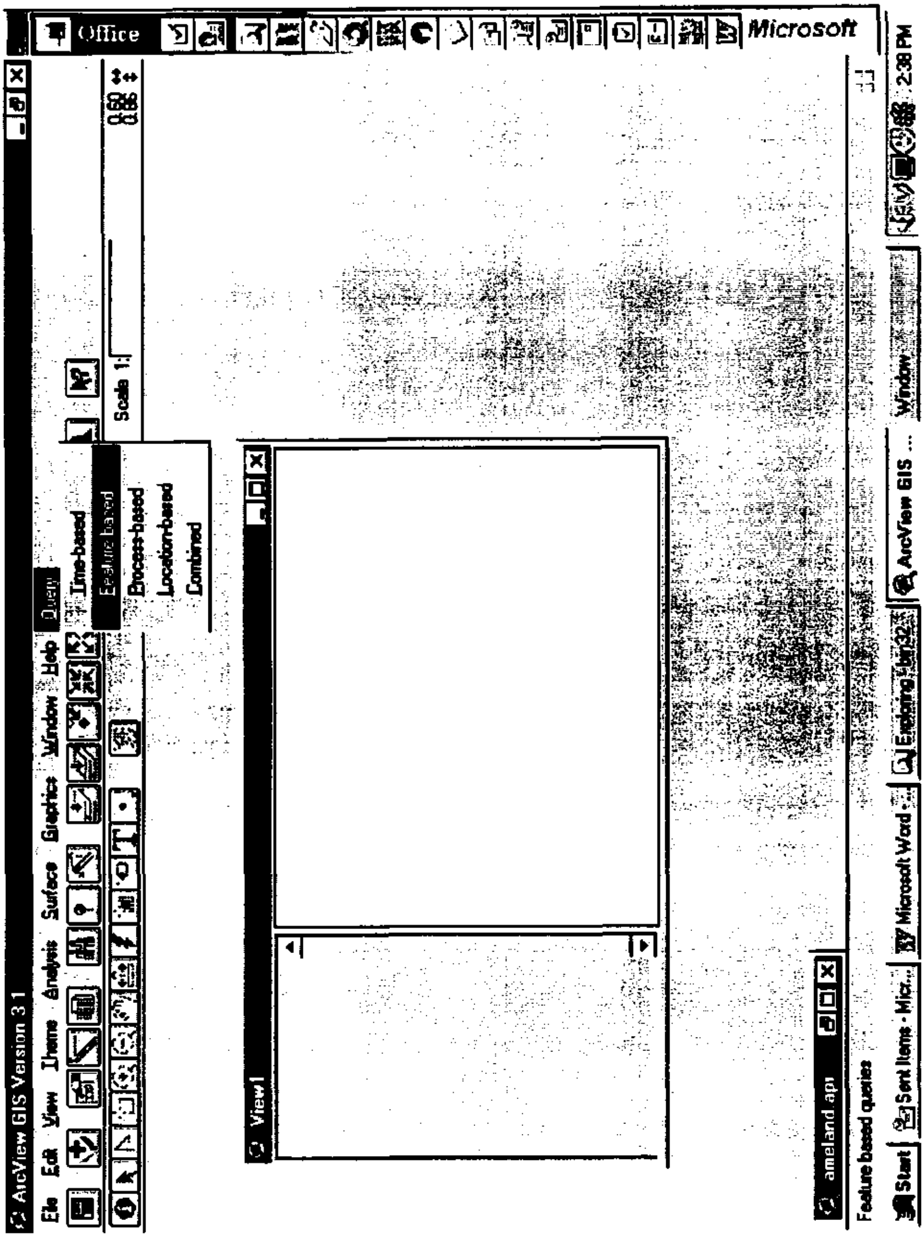


Figure 8.2 User interface of multi-perspective queries.

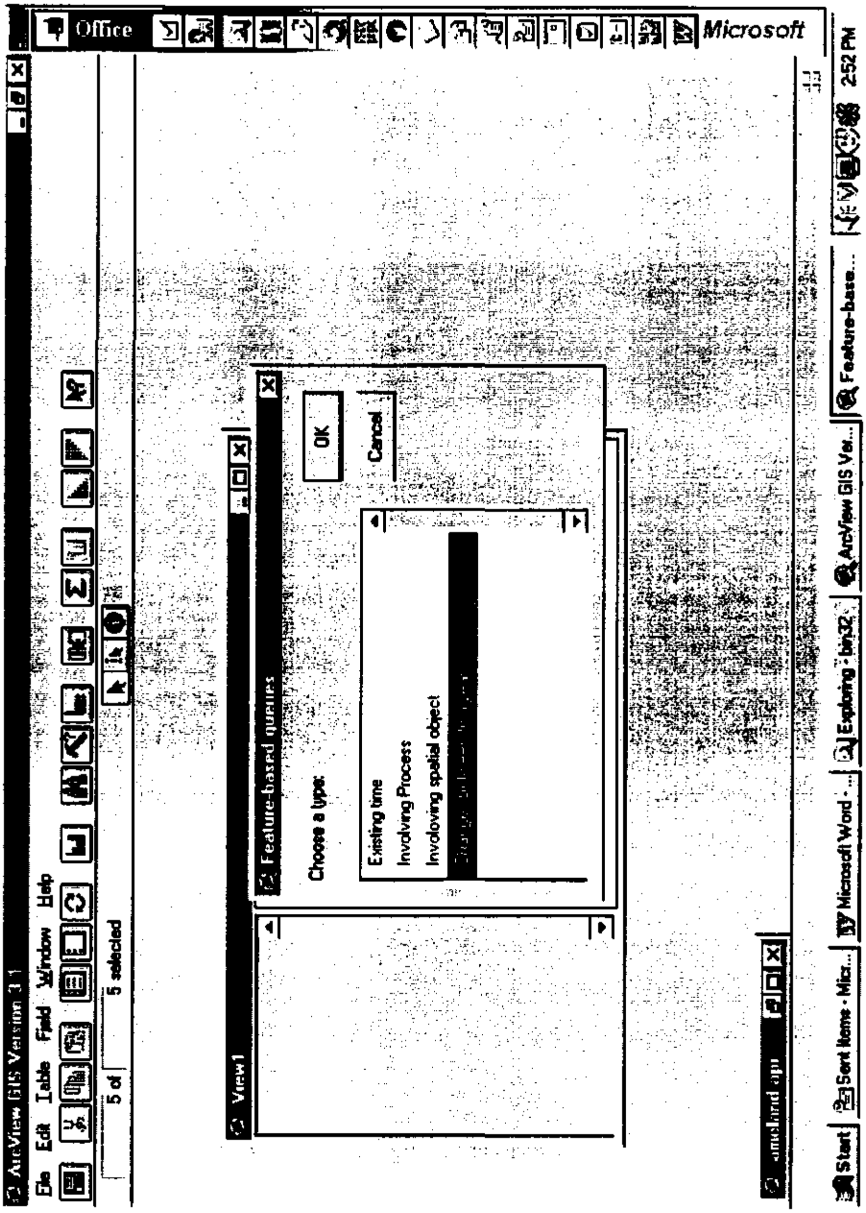


Figure 8.3 User interface of the object-oriented query.

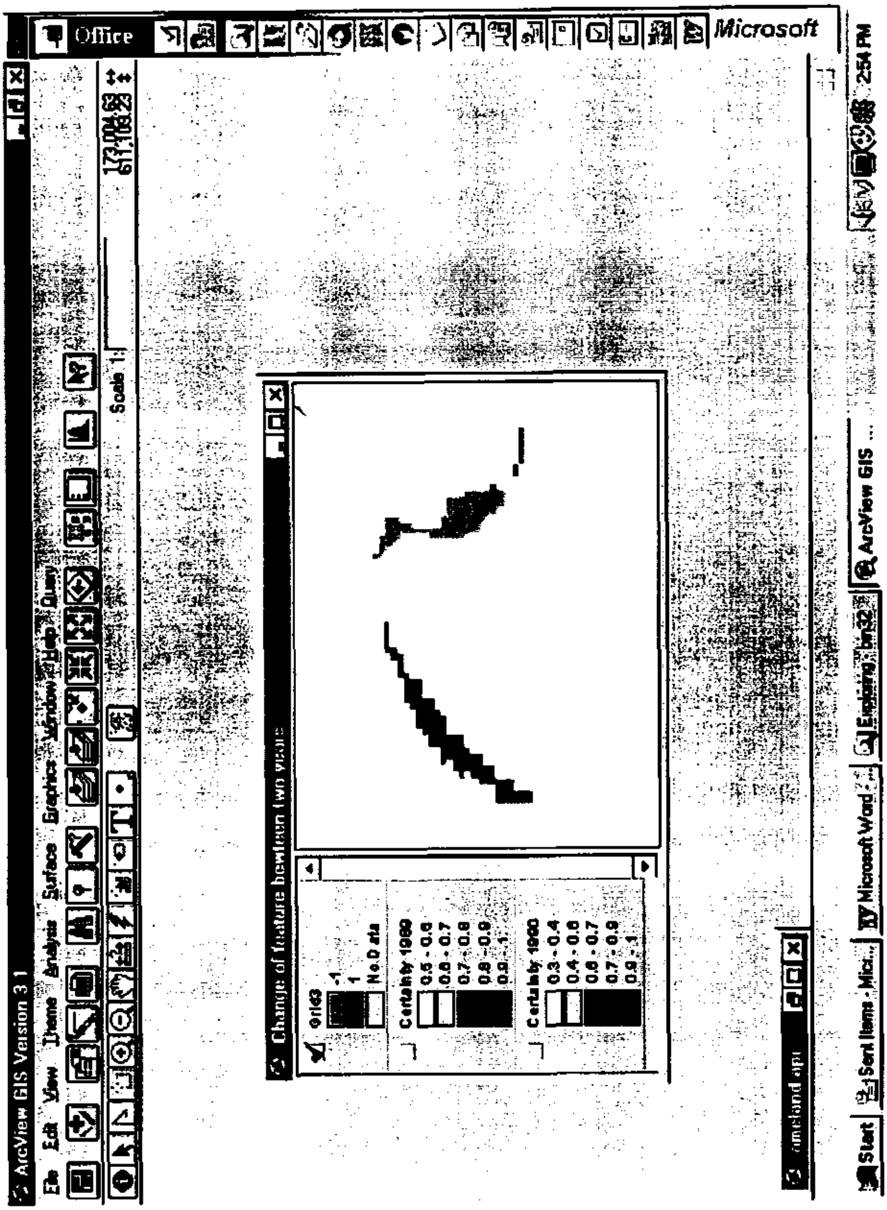


Figure 8.4 Query result of the change of Object 1 from year 1989 to 1990.

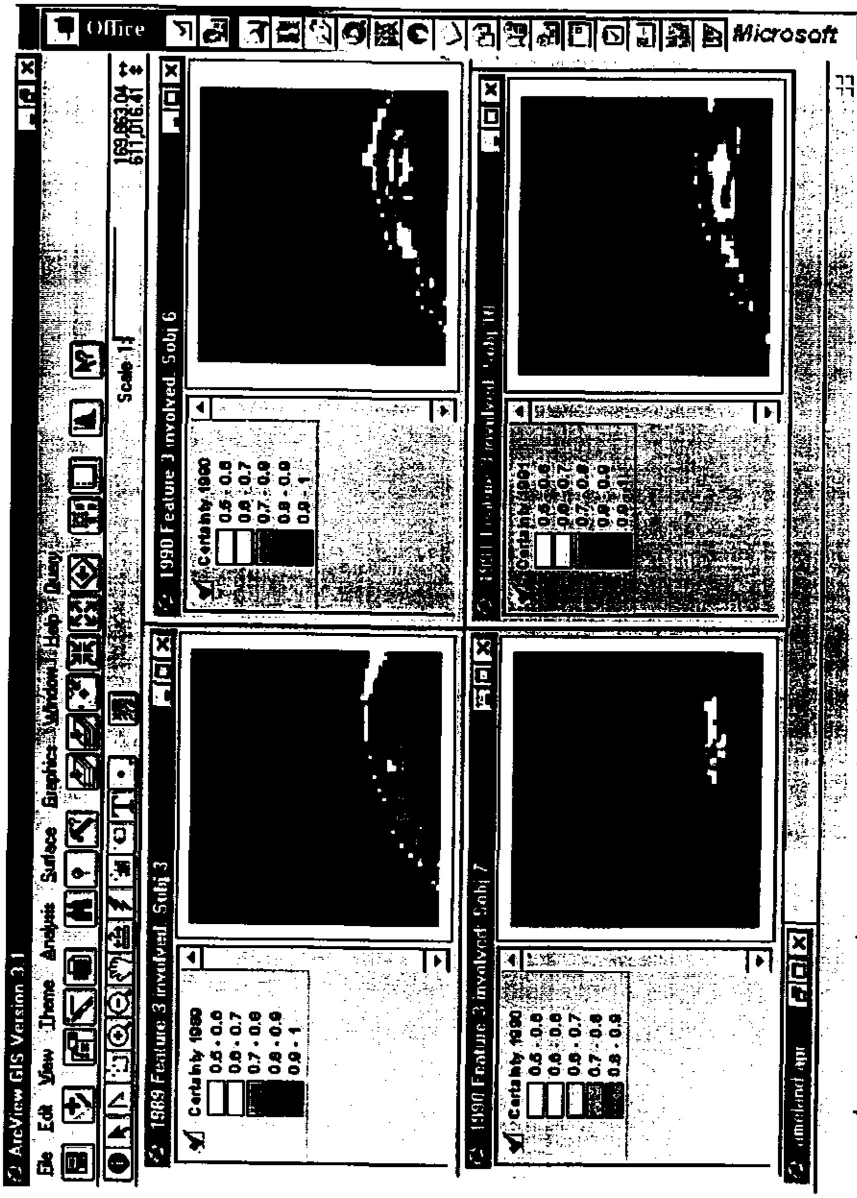


Figure 8.5 Querying result (A) of processes that Object 3 underwent from year 1898 to 1991.

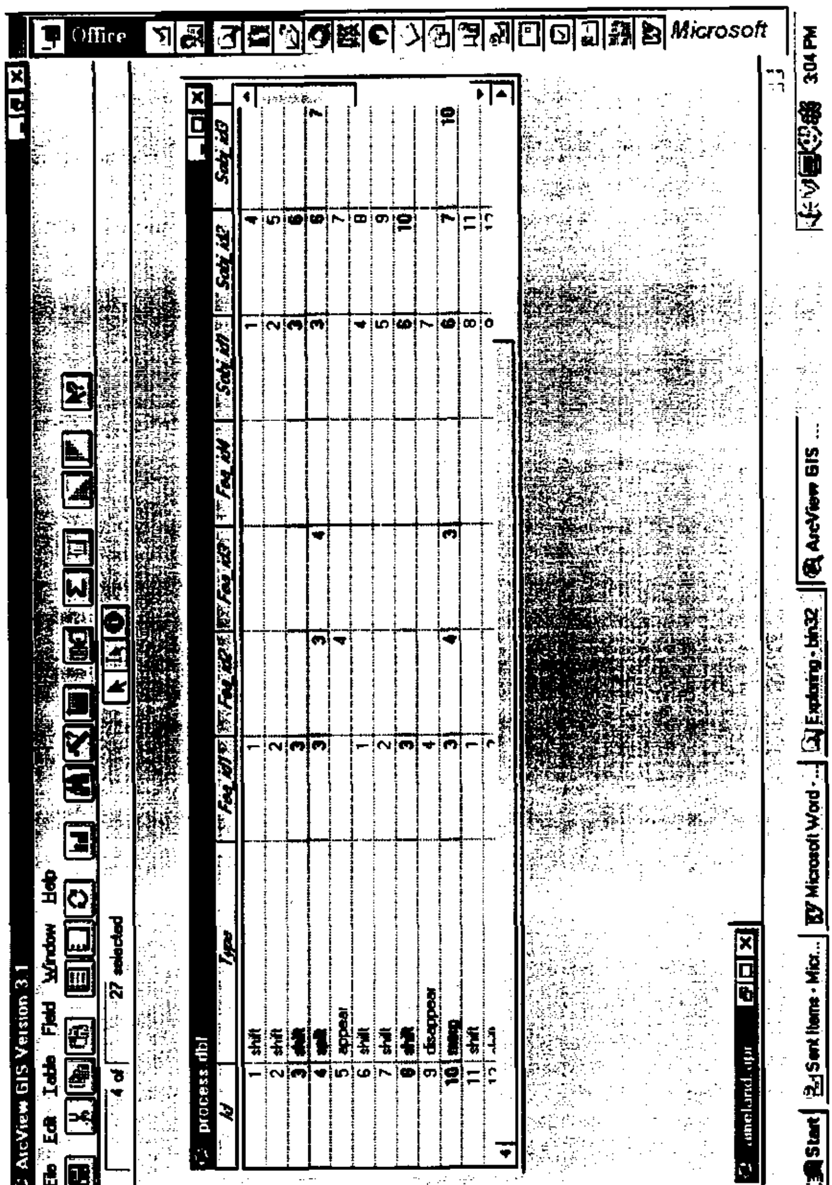


Figure 8.6 Querying result (B) of processes that Object 3 underwent from year 1989 to 1991.

However, there are two ways to normalize the table. One is to separate the processes of different type into different tables. It requires that the number of objects involved in any one type of process be fixed. The data of processes containing the Pid, Ptype, Start_time, and End_time are stored in one table (*Process*). Then the objects and regions involved in the process are stored in tables of different types. For example Table 8-6 can be restructured as:

Table of *Process*
Process(Pid, Ptype, Start_Time, End_tme)

Table of Process 'Shift'
Shift(Pid, Oid, Rid_1, Rid_2)

Table of process 'Expand'
Expand(Pid, Oid, Rid_1, Rid_2)

Table of process 'Shrink'
Shrink(Pid, Oid, Rid_1, Rid_2)

Table of process 'Merge'
Merge(Pid, Oid_1, Oid_2, Oid_3, Rid_1, Rid_2, Rid_3)

Table of process 'Split'
Split(Pid, Oid_1, Oid_2, Oid_3, Rid_1, Rid_2, Rid_3)

Table of process 'Appear'
Appear(Pid, Oid, *Rid)*

Table of process 'Disappear'
Expand(Pid, Oid, Rid)

where *Pid* represents the process identifier, *Oid* represents an object identifier, *Rid* represents the identifier of the region, *Start_Time* is the starting time and *End_Time* is the ending time of the process.

The other way is to add two tables to describe *OidList* and *RidList* (cf. Subsection 8.2.2) in the relational scheme of *Process*, instead of treating each element in the list as one record in the table of process. The relation scheme of *Process* would be

Process(Pid, Ptype, OidListID, RidListID, Start_T, End_T).

The structures of Table of *OidList* and Table of *RidList* would be:

OidList(<u>OidListId</u>, Oid)
RidList(<u>RidListId</u>, Rid)

(2) dynamic visualization of process

The processes of object can only be visualized by either states of regions or by the changes. A dynamic visualization of the processes would be very helpful to explore the dynamics of the natural phenomena and to predict situations in future. This is a topic for further research, which is discussed in the last chapter of this thesis.

Chapter 9

Conclusions and Discussion

9.1 Introduction

Efforts to integrate GIS and environmental modeling have recently led to the consideration of new spatio-temporal data models to represent natural phenomena and their dynamics. This has posed two challenges to GIS researchers.

The first challenge comes from the fact that natural phenomena occur in a physical environment with a complex and continuous character, and their representation in GISs requires abstraction and discretization. However, conventional data models (the field model and the object model) are simple abstractions of such complexity. They are crisp models, unable to handle uncertainties either in values of attributes or in identification of the entities involved. This has recently generated interest in developing fuzzy object concepts and models.

The second challenge stems from the fact that, in order to explain and predict the dynamics of natural phenomena, it is necessary to present and model the dynamic changes. However, conventional data models are static and are unable to describe spatial changes. Consequently, research has concentrated on developing what are known as spatio-temporal data models.

It was against this background that the study reported in this thesis was initiated and conducted. One particular objective was to develop a spatio-temporal data model for objects with fuzzy spatial extent. To meet this objective, the study concentrated on three aspects:

1. identification of fuzzy objects,
2. detection of dynamic changes in fuzzy objects, and
3. representation of objects and their dynamics in a spatio-temporal data model.

In this final chapter of the thesis, I summarize and discuss the research findings reported in the previous chapters. In addition, issues for future research are identified.

9.2 Summary

The spatio-temporal data model developed in this research followed a three-stage approach. Each stage covered one of the aspects mentioned above (Section 9.1) in the introduction to this chapter.

1. Identification of fuzzy object.

In this first stage, the information on object states was extracted from raw data. Issues related to this stage of data processing were solved as follows.

- A procedure to extract an object from field observation data was established.

  The procedure for extracting objects from field observation data consisted of six steps: sampling, interpolation, classification, segmentation, merging and identification. As raw data were normally sampled at specific points, interpolation was needed to obtain the value of points covering the whole area. Then, certain classification criteria were applied to assign the cells of the interpolated raster into classes of object types. After segmentation of the classified raster and the merging of small regions, spatial objects (regions) were formed to represent the spatial extent of natural phenomena (called objects) at different states (epochs) (Chapter5).

  Through this procedure, the representation of real-world phenomena was converted from a field-oriented approach to an object-oriented approach, leading to a generalized object concept for representing natural phenomena. I discuss this point in Section 9.3.

- The propagation of uncertainties in the procedure was investigated and formalized.

  Uncertainty propagation in the whole procedure from field observation data to identified spatial extent and boundaries of objects was analyzed.

  - Propagation of uncertainties in classification. Since two kinds of uncertainties exist namely 'errors' in measurement and 'fuzziness' in object class definition. Four combinations (cases) of these uncertainties have been discussed in Chapter 4. The stochastic aspects of the observed data and the uncertainty of object class definition were combined through a convolution operation. The case study showed that stochastic errors have similar influence on both crisp and fuzzy classification.

- Propagation of uncertainties in segmentation. Uncertainties of classification were propagated through the segmentation to the spatial extent of objects, i.e. the spatial extent of objects were fuzzy due to fuzzy classification. As a consequence, the concepts of conditional spatial extent and conditional boundaries have been developed, which could only be identified under certain criteria. The propagation of uncertainty from classification to fuzzy spatial extent was expressed explicitly by the relationship between uncertainty of a cell belonging to the spatial extent of an object and the uncertainty of a pixel belonging to classes. This showed that thematic uncertainty definitely influences the uncertainty of spatial extent and conditional boundaries or transition zones of the objects (Chapter 5).

- Propagation of uncertainties in merging. A fuzzy merging approach was proposed for merging regions smaller than pre-defined mapping units. In this approach, sharpness of edges of the boundaries between the regions (to be merged) and its neighbors were compared and the edge that was less sharp was merged out, so it is a more accurate merging method (Chapter 5).

- Three fuzzy object models were proposed to represent fuzzy objects.

  Fuzzy object models were proposed to represent objects for different situations. The CC-objects model (crisp boundary and crisp interior) can be used to model crisp objects. The FF-object model (fuzzy boundary and fuzzy interior) is suitable for modeling fuzzy objects that spatially overlap. The FC-object model (fuzzy boundary and crisp conditional interior) and CF-object model (crisp conditional boundary and fuzzy interior) are useful for modeling fuzzy objects that are spatially disjoint. The representation of spatial extent and conditional boundaries (or transition zones) of objects for these three fuzzy object models were formalized in mathematical formulas.

2. Detection of dynamic changes in fuzzy objects

Since the spatial extents (regions) extracted at different epochs represent the states of objects, the historic lifelines of objects were built by linking the regions at consecutive epochs. The processes through which the objects had evolved were also identified by comparing the state transition of the objects.

- A method was proposed to identify objects and their state transitions from fuzzy spatial extents at different epochs.

  The relationships between regions at consecutive epochs were described through similarity indicators derived from their spatial overlap. After evaluating similarity between regions, these regions that represent the consecutive states of one object were linked. In this way, the lifelines of objects could be viewed, and thus the objects determined. Since the objects were identified through the regions representing their states, the state transitions of the object were identified as well. The state transitions were described as processes such as shift and merge (Chapter 6).

- Uncertainty of detected change was discussed.

  By comparing the spatial extents of objects at consecutive epochs, the change of objects was detected. The uncertainty of the change was analyzed from a series of change maps at different certainty levels. Such maps can provide decision makers with more accurate information about change (Chapter 6).

3. A spatio-temporal model was designed to represent fuzzy objects and their dynamics.

   - A process-oriented spatio-temporal data model was proposed.

     A process-oriented spatio-temporal data model was proposed that explicitly represents changes of objects over space with time. The model can also represent the dynamic interactions between objects. Therefore, the process-oriented data model can represents dynamic processes affecting the spatial and thematic aspects of individual objects and object complexes. Because the model explicitly stores change (process) relative to time, procedures for answering queries relating to temporal relationships, as well as analytical tasks for comparing different sequences of change, were facilitated (Chapter 7).

   - The prototype of the process-oriented spatio-temporal data model was implemented in ArcView.

     The conceptual design of the process-oriented spatio-temporal data model was converted into a logical design by converting the extended entity relationship (EER) model into a relational schema (Chapter 8). A prototype of the data model was applied in a study of the coastal geomorphology of Ameland. The user-interface of the multi-perspective queries was developed based upon ArcView and its programming language, Avenue.

9.3 Discussion and conclusions

The previous section summarized the concepts, methods and model developed in the thesis. I will evaluate the contribution of the thesis to the state-of-the art in fuzzy object modeling and the spatio-temporal data models.

9.3.1 Discussion of the fuzzy object modeling approach

- A generalized object concept – object with fuzzy spatial extent has been developed.

  - The research uses a formal data model approach to represent uncertainties of objects and their description. The boundary-oriented approach and the pixel-oriented approach were unified in such a formal data model, thus unifying the boundary-oriented and pixel-oriented approach for object extraction and expression (cf. Section 3.8). The objects can be approached either by identifying the boundaries or by identifying their spatial extents.

  - The model can handle both fuzzy and crisp objects, with the latter can be considered a special case of the former. In the fuzzy object case, objects have fuzzy transition or boundary zones in which conditional boundaries may be defined; for crisp objects the boundaries are sharp.

  - This concept unifies the field and object approach by defining a procedure for extracting (fuzzy) object data from field observations. The concept of object with fuzzy spatial extent links the object-oriented to the field-oriented characteristics of natural phenomena. The objects have conditional boundaries, representing their object characteristics; the interiors of the objects have field properties, representing their gradual and continuous distribution.

  - This approach has been implemented by raster geometry. Since both vector and raster geometry can be represented by the formal data structure (FDS), the proposed approach can also be implemented by vector geometry.

- The procedure of object identification makes the spatial effect of uncertainty of thematic field data on object geometry explicit.

  In this modeling approach, uncertainties of objects were tracked from field observation through the process of identifying their spatial extent. This model expresses the relationship between uncertainty of a cell belonging to the spatial extent of an object and the uncertainty of the cell belonging to classes, i.e. the relationship of uncertainties between the geometric aspects and the thematic aspect. It means that uncertainty is transferred from thematic aspects to geometric aspects of objects during spatial clustering, i.e. *existential* uncertainty is converted to *extensional* uncertainty.

- A method has been developed that identifies the objects and detects their dynamics of these objects automatically.

  Conventionally, object identification is realized through image interpretation by domain experts or by checking in the field. Afterwards, the changes of objects are detected by comparing the states of objects at different epochs. The processes through which the objects evolved are then analyzed by the experts. This thesis, however, describes a method that can identify objects and detect their dynamics automatically. The objects and processes are determined by comparison of object states at successive epochs. In this way, the historic lifelines of objects are built automatically. The relationship, or interactions, between objects are identified.

  Beyond that, by using temporal information on the objects, the misclassified regions can be detected and corrected, another advantage of the proposed method (cf. Subsection 6.4.3).

  As demonstrated by the case in Section 6.5, this method enables automated identification of objects and their dynamics, albeit under the assumption that the changes of the objects are continuous and slow from long-term point view. Sudden changes, such as overnight storms, were not considered. It also requires that the temporal data be obtained during the season, at a more or less fixed temporal and spatial resolution.

  No thorough analysis has yet been made as to how threshold values should be established for similarity indicators used to evaluate the relationship between regions and the state transitions of objects. The values of these thresholds are related to the application domain and the definition of the process. Such a relationship implies that the objects and processes identified are uncertain, this will be discussed in Section 9.4.

9.3.2 Discussion about the Star Model

The process-oriented spatio-temporal data model proposed can support analysis and queries of time series data from varying perspectives through location-oriented, time-oriented, feature-oriented and process-oriented queries, in order to understand the behavior of dynamic spatial complexes of natural phenomena.

The Star model is a general model from which other models can be derived. For example, the Triad Model (Peuqeut and Qian, 1996) and the Three Domain Model (Yuan, 1995) can be considered as special cases of the Star Model, in which only the object, spatial and temporal aspects are considered. The event-based model (ESTDM) can be considered as an event-based view of the Star Model.

The prototype of Star model has been implemented and applied to a coastal geomorphology case to demonstrate the practical usefulness of the model.

9.4 Future research

Although the present study has contributed to the development of spatio-temporal data models for fuzzy objects, there are several aspects of the model developed that maybe improved in future research.

- Object identification procedure.

Several issues related to the procedure of object identification need further research. The first issue is the modeling of uncertainties in the object class definition due to time incoherence. In this approach, time incoherence in class definition was partly expressed by the fuzzy membership function. The handling of time-incoherent uncertainty needs further investigation, though, since fuzzy set theory basically represents uncertainty in class definition at a specific time. Present theory cannot yet handle time-dependent class definitions.

The second issue is that the segmentation in this procedure is still primitive. Only the class type, which is derived based upon thematic data, is used in segmentation. After segmentation, there are some regions that are smaller than pre-defined mapping units or that do not obey certain topographic criteria between objects in reality (e.g. Region 7 was classified as beach area but was contained in a foredure area, cf. Figure 6.2). Re-segmentation might be possible by adding extra evidence (e.g. environmental evidence as a topological relationship). The topological relationships between objects can also be used as information for the final identification of objects. How to integrate the topological information in the procedure, either in segmentation or merging, will be another interesting point for future research.

The third issue is the contiguity of the spatial extent $Face(O_i)$ and the connectivity of $Face(O_i|Threshold)$ (Threshold $\in [0,1]$). If the threshold is higher, the connectivity of the faces belonging to an object will change so that beyond a certain threshold value islands may occur in the regions of an object. How to handle this situation needs further analysis.

- Uncertainty of identification of dynamics of objects.

As pointed out in Subsection 9.3.1, the thresholds of similarity indicators used to evaluate the relationship between regions and to identify the state transitions of objects have not yet been clearly established; there is uncertainty in the identification of dynamics of objects. Although relationships between existential uncertainty, extensional uncertainty and geometric uncertainty have been discussed, the uncertainty of the lifeline of an object due to extensional uncertainty has not been investigated.

Furthermore, the uncertainty of processes identified from state transitions has not discussed. To derive the uncertainty of identified objects and the processes involved needs further study.

- Process-oriented data model.

In the process-oriented spatio-temporal model, location related information was stored in raster-based layers according to time sequence. This leads to data redundancy for consecutive years. This is not a big issue in the prototype, in which only several years data were stored, but that may be different for real monitoring projects of large areas. Therefore, other structures should be developed for this raster-based data to reduce data redundancy.

Another issue is the modification of this model for three- dimensional spatial objects (e.g. underground coal fires). The amount of data required tends to grow significantly for three-dimensional applications.

- Visualization of dynamic process.

In this study processes were visualized by displaying the spatial extent of objects involved in them or by the symbols used in Figure 6.3. However, these are static representations. There are several alternatives for visualizing dynamic processes, such as morphing and cartographic animation (Blok *et al.*, 1999). But further study on this aspect is necessary to give a live view of a process or to explore the data for prediction. A complicating factor in this is that dynamic objects require the visualization of uncertainty aspects and change.

- Prediction of the development of landscape units.

  It is a valuable exercise to investigate the change patterns of the landscape units, and the change pattern of their fuzziness, as well, in order to find out about possible spatio-temporal relationships between them. Prediction of potential development of the landscape units based upon detected changes is another point for further work.

- More case studies and other approaches for data observation.

  Due to the limitations of time and data, this thesis used the case of the coastal geomorphology of Ameland to illustrate the approach and test the data model. It would be very useful to test the model developed on other cases in which objects have fuzzy spatial extent, such as in monitoring temporal change in vegetation. It may also be necessary to use this approach to investigate the uncertainties of objects extracted from data acquired from other sources, such as remote sensing images or scanned cartographic maps.

References

Al-Taha, K., and Berrera, R., 1994, Identities through time, *Proceedings of ISPRS Working Group II/2, Workshop on the requirements for Integrated Geographic Information Systems*, New Orleans.

Al-Taha, K. and Frank, A.U., 1993, What a temporal GIS can do for cadastral systems, GISA'93 in Sharjah, pp. 13-1–13-17.

Altman, D., 1994, Fuzzy set theoretic approaches for handling imprecision in spatial analysis, *International Journal of Geographical Information Systems*, Vol. **8**, No. 3, 271–289.

Armstrong, M. P., 1988, Temporality in spatial databases, *Proceedings of GIS/LIS'88*, Vol. **2**, 880–889.

Bezdek, J. C., 1981, *Pattern Recognition with Fuzzy Objective Function Algorithms*, New York: Plenum.

Bezdek, J. C., Enrlick, R. and Wang, F., 1984, FCM: The fuzzy c-means clustering algorithm. *Computer & Geosciences*, **10**, 191–203.

Blakmore, M., 1984, Generalisation and error in spatial database, *Cartographica*, **21**: 131–139.

Blok, C., Kobben, B., Cheng, T. and Kuterema, A. A., 1999, Visualization of relationships between spatial patterns in time by cartographic animation, accepted by *Cartography and Geographic Information System.*

Bolstad, P.V., Gessler, P. and Lillesand, T. M., 1990, Positional uncertainty in manually digitized map data, *International Journal of Geographical Information Science*, Vol. **4**, No. 4, 399–412.

Brown, D., G., 1998, Classification and boundary vagueness in mapping presettlement forest types, *International Journal of Geographical Information Science*, Vol. **12**, No. 2, 105–129.

Burrough, P. A., 1989, Fuzzy mathematical methods for soil survey and land evaluation, *Journal of Soil Science*, **40**, 477–492.

Burrough, P. A., 1996, Natural objects with indeterminate boundaries, *Geographic Objects with Indeterminate Boundaries*, edited by P. A. Burrough and A. U. Frank, London: Taylor & Francis, pp. 3–28.

Burrough P. A. and Frank, A. U., 1995, Concepts and paradigms in spatial information: are current geographical information systems truly generic? *International Journal of Geographical Information System*, Vol. **9**, No. 2, pp. 101–116.

Burrough, P. A, and Frank, A.U., 1996, *Geographic Objects with Indeterminate Boundaries*, London: Taylor & Francis.

Burrough, P.A., Kijn, R. van and Rikken, M., 1996, Spatial data quality and error analysis issues: GIS function and environmental modeling, *GIS and Environmental Modeling: Progress and Research Issue*, edited by M. F. Goodchild, L. T. Steyart, B. O. Parks, C. Johnstone, D. Maidmend, M. Crane, and S. Glendinning, Fort Collins: GIS World, pp. 29-34.

Burrough, P. A. and McDonnel, R.A., 1998, *Principles of Geographical Information Systems*, London: Oxford University Press.

Cheng, T., Molenaar, M. and Bouloucos, T., 1997, Identification of fuzzy objects from field observation data, *Spatial Information Theory: A Theoretical Basis for GIS, Lecture Notes in Computer Sciences*, edited by S. C. Hirtle and A. U. Frank, Berlin: Spring-Verlag, **Vol. 1329**, pp. 241–259.

Cheng, T. and Molenaar, M., 1997, Dynamics of fuzzy objects, *Proceedinsg of the International Workshop on Dynamic and Multi-dimensional GIS*, edited by Y. C. Lee and Z. Li, Hong Kong Polytechnic University, pp.49–63.

Cheng, T. and Molenaar, M., 1998a, A process-oriented spatio-temporal data model to support physical environmental modeling, *Proceedings of the 8th International Symposium on Spatial Data Handling*, edited by T. K. Poiker and N. Chrisman, pp. 418–430.

Cheng, T. and Molenaar, M., 1998b, The identification and monitoring of objects with fuzzy spatial extent, *International Archives of Photogrammetry and Remote Sensing (ISPRS)*, Vol. **32**, Part 3, pp. 207–212.

Cheng, T. and Molenaar, M., 1998c, Change of fuzzy Objects, paper presented at *First Conference of Association of Geographic Information Laboratories in Europe (AGILE)*, Enschede, The Netherlands, April 23-25.

Cheng, T., Molenaar, M. and Bouloucos, T., 1998, Fuzzy object models and their application, *Proceedings of International Conference of Spatial Information Science, Technology and its Applications* (SIST'98), Dec. 13-16, Wuhan, pp. 408–416.

Cheng, T., Zuidam, R. A. van, and Kainz, W., 1995, A unified spatio-temporal data model for 4-D GIS, *Proceedings of GIS/LIS'95 Annual Conference* and Exposition, pp. 967–976.

Bolstad, P., Gessler, P. and Lillesand, T. M., 1990, Positional uncertainty in manually digitized map data, *International Journal of Geographical Information Systems*, Vol. **4**, No. 4, 399–412.

Chrisman, N. R., 1982, A theory of cartographic error and its measurement in digital database, *Proceedings of Auto-Carto 5*, Crystal city, VA, pp. 159–168.

Chrisman, N.R., 1991, The error component in spatial data, *Geographical Information Systems*, edited by D. J. Maguire, M. F. Goodchild and D. W. Rhind, Harlow : Longman, pp. 165–174.

Claramunt, C., and Thériault, M., 1996, Toward semantics for modelling spatio-temporal processes within GIS, *Advances in GIS Researches II, Proceedings of 7th International Symposium on Spatial Data Handling (SDH'96)*, edited by M. J. Kraak and M. Molenaar, London: Taylor & Francis, pp. 47–63.

Claramunt, C., Parent, C., and Thériault, M., 1997, Design patterns for spatio-temporal processes, *Searching for Semantics: Data Mining, Reverse Engineering*, edited by S. Spaccapictra and F. Maryanski, pp. 415–428.

Couclelis, H., 1992, People manipulate objects (but cultivate fields): beyond the raster-vector debate, *Theories and methods of spatio-temporal reasoning in geographic space, Lecture Notes in Computer Science*, edited by A. U. Frank, I. Campari and U. Formentini, Berlin: Springer-Verlag, Vol. **639**, pp. 65–77.

Couclelis, H., 1996, Toward an operational typology of geographic entities with ill-defined boundaries, *Geographic Objects with Indeterminate Boundaries*, edited by P. A. Burrough and A. U. Frank, London: Taylor & Francis, pp. 45–56.

Couclelis, H. and Gottsegen, J., 1997, What maps mean to people: denotation, connotation, and geographic visualization in land-use debates, *Spatial Information Theory: A Theoretical Basis for GIS, Lecture Notes in Computer Sciences*, Vol. 1329, edited by S. C. Hirtle and A. U. Frank, Berlin: Spring-Verlag, pp. 151–162.

Clementini E. and Felice, P. D., 1996, An algebraic model for spatial objects with indeterminate boundaries, *Geographic Objects with Indeterminate Boundaries*, edited by P. A. Burrough and A. U. Frank, London: Taylor & Francis, pp. 155–170.

David, B. and Herrewegen, M., van den, Salgé, F., 1996, Conceptual models for geometry and quality of geographic information, *Geographic Objects with Indeterminate Boundaries*, edited by P. A. Burrough and A. U. Frank, London: Taylor & Francis, pp. 193–206.

Dijkmeijer, J. and Hoop, S. de, 1996, Topologic relationships between fuzzy area objects, *Advances in GIS Researches II, Proceedings of the 7th International Symposium on Spatial Data Handling (SDH'96)*, edited by M. J. Kraak and M. Molenaar, London: Taylor & Francis, pp. 377–394.

Dunn, R, Harrison, A. R. and White J. C., 1990, Positional accuracy and measurement error in digital databases of land use and empirical study, *International Journal of Geographic Information Systems*, Vol. **4**, No. 4, 385–398.

Dutton, G., 1992, Handling positional uncertainty in spatial databases, *Proceedings of the 5th International Symposium on Spatial Data Handling*, Charleston, South Carolina, pp. 460–469.

Eastman, J.R., Kyem, P. A. K., Toledano, J. and Jin, W., 1993, *Explorations in Geographic Information Systems,* Volume 4: *GIS and Decision Making*, Geneva: United Nations Institute for Training and Research (UNITAR).

Ebdon, D., 1985, *Statistics in Geography* (Second Edition), Oxford: Basil Blackwell.

Edwards, G., 1994, Aggregation and desegregation of fuzzy polygon for spatial-temporal modelling, *Advanced Geographic Data Modeling*, edited by M. Molenaar and Hoop, S. de, The Netherlands Geodetic Commission, New Series, Nr. 40, pp.141–154.

Edwards, G., and Lowell, K. E., 1996, Modeling uncertainty in photointerpreted boundaries, *Photogrammetric Engineering and Remote Sensing*, **62**, 337–391.

Edwards, G., Fortin, M., Thonson, K., Aubert, E. and Lowell, K., 1998, The propagation of boundary uncertainty from maps to models, *Data Quality in Geographic Information – From Error to Uncertainty*, edited by M. F. Goodchild, and R. Jeansoulin, Paris: Hermes, pp.135–150.

Eleveld, M.A., Cheng, T. and Zuidam, R. A. van, 1995, Towards a decision support system for coastal zone management by applying morphodynamic modelling with remote sensing data input in a 4-D GIS environment, *Proceedings of Third Thematic Conference of Remote Sensing for Marine and Coastal Environments, Technology And Applications*, Seattle, USA, Vol. **1**, pp. 256–265.

Egenhofer, M. J. and Golledge, R. G., 1998, *Spatial and Temporal Reasoning in Geographic Information Systems*, Oxford: Oxford University Press.

Fisher, P., 1992, First experiments in viewshed uncertainty: the accuracy of the viewshed area, *Photogrammetric Engineering and Remote Sensing*, **57**, 1321–1327.

Fisher, P., 1994, Problem and fuzzy models of viewshed operation, *Innovations in GIS 1,* edited by M. Worboys, London: Taylor & Francis, pp. 162–175.

Fisher, P., 1996, Boolean and Fuzzy Regions, *Geographic Objects with Indeterminate Boundaries*, edited by P. A. Burrough and A. U. Frank, London: Taylor & Francis, pp. 87–94.

Gale, S., 1972, Inexactness, fuzzy sets, and approach to the representation of behavioral geography, *Geographical Analysis*, **4**, 337–349.

Galton, A., 1997, Continuous change in spatial regions, *Spatial Information Theory: A Theoretical Basis for GIS, Lecture Notes in Computer Sciences*, edited by S. C. Hirtle, and A. U. Frank, Berlin: Spring-Verlag, Vol. **1329**, pp.1–13.

Gahegan, M. N., 1996, Specifying the transformations within and between geographic data models, *Transactions in GIS*, Vol. 1, No. 2, pp. 137–152.

Gahegan, M. N. and Ehlers, M., 1997, A framework for modeling of uncertainty in an integrated geographic information system, *Proceeding of the International Workshop on Dynamic and Multi-dimensional GIS*, edited by Y.C. Lee and Z. Li, Hongkong Polytechnic University, pp.64–79.

Golden Software Inc., 1990, *Manual of Surfer Access System.*

Goodchild, 1989, Modeling error in objects and fields, *Accuracy of Spatial Databases*, edited by M. F. Goodchild and S. Gopal, London: Taylor & Francis, pp. 107 –113.

Goodchild, M. F., 1992, Geographical data modeling, *Computer & Geosciences*, Vol. **18**, pp. 400–408.

Goodchild, M. F., 1993, The state of GIS for environmental problem-solving, *Environmental Modeling with GIS*, edited by M. F. Goodchild, B O. Parks, and L.T. Steyaert, New York: Oxford University Press, pp. 8–15.

Goodchild, M. F. and Gopal, S., 1989, *Accuracy of Spatial Databases*, London: Taylor & Francis.

Goodchild, M. F., Guoqing, S. and Shiren, Y., 1992, Development and test of an error model for categorical data, *International Journal of Geographical Information Systems*, Vol.**6**, No. 2, 87–104.

Goodchild, M. F. and Jeansoulin, R., 1998, *Data Quality in Geographic Information – From Error to Uncertainty*, Paris: Hermes.

Goodchild, M.F., Parks, B. O., and Steyaert, L.T., 1993, *Environmental modeling with GIS: 1st International Conference on Integrating Geographic Information Systems and Environmental Modeling*, New York: Oxford University Press.

Goodchild, M.F., Steyart, L. T., Parks, B. O., Johnstone, C., Maidmend, D., Crane, M., and Glendinning, S., 1996, *GIS and Environmental Modeling: Progress and Research Issue*, Fort Collins: GIS World, - 486 p.

Graaff, L.W. S. de, 1977, Het Strand: De Relatie Tussen Processen, Materialen en Vormen, *en Een Proeve van Terminologie-Gebruik, Koninklijl Nederlands Aardrijkskundig Genootschap Geografisch Tijdschrift*, Nieuwereeks **XI** 1977, Nr. 1, pp.47–67.

Gray, M. V., 1997, Classification as an impediment to the reliable and valid use of spatial information: a disaggregate approach, *Spatial Information Theory: A Theoretical Basis for GIS*, *Lecture Notes in Computer Sciences*, edited by S. C. Hirtle, and A. U. Frank, Berlin: Spring-Verlag, Vol. **1329**, pp. 241–259.

Heuvel, T. van and Hillen, R., 1994, *Coastline Management*, internal publication of Directorate-Generla for Public Works and Water

Management, the Dutch Ministry of Transport, Public Works and Water Management.

Heuvelink, G. B. M. and Burrough, P. A., 1993, Error propagation in cartographic modeling using Boolean and continuous classification, *International Journal of Geographical Information Systems*, Vol. 7, No. 3, 231–246.

Heuvelink, G. B. M., Burrough, P. A. and Stein, A., 1989, Propagation of errors in spatial modelling with GIS, *International Journal of Geographical Information Systems*, Vol. **3**, No.4, 303–322.

Hogg, R. V. and Craig, A. T., 1970, *Introduction to Mathematical Statistics* (Third edition), New York: Macmillan Publishing.

Hornsby, K. and Egenhofer, M. J., 1997, Qualitative representation of change, *Spatial Information Theory: A Theoretical Basis for GIS, Lecture Notes in Computer Sciences*, edited by S. C. Hirtle, and A. U. Frank, Berlin: Spring-Verlag, Vol. **1329**, pp. 15–33.

Hootsmans, 1996, R., *Fuzzy Sets and Series Analysis for Visual Decision support in Spatial Data Exploration,* Ph.D thesis, Utrecht, The Netherlands.

Hughes, J. G., 1991, *Object-Oriented Databases*, Prentice Hall.

Huising, E.J., Jordans, R.W.L., et al., 1996, *Framework Digital Elevation Models Dutch Coast - Various Techniques for Topographic Measurement Tested on a common Site*, The Netherlands Remote Sensing Board (BCRS) Programme Bureau, Rijkswaterstaat, Survey Department.

Hunter G. J. and Goodchild, M. F., 1995, Dealing with error in spatial databases: a simple case study, *Photogrammetric Engineering & Remote Sensing*, Vol. **61**, No. 5, pp. 529–537.

Inomata, Y. and Ogata, S., 1992, Supervised, unsupervised fuzzy classifications using histograms as membership functions, *Proceedings of 6th Australasian Remote Sensing Conference*, Wellington, Vol. **2**, pp. 434–441.

Jang, Y.P. and Johnson, R.G., 1994, Evolutions of object states in temporal object-oriented databases, *Proceedings of 22nd Annual Computer Science Conference (ACM),* pp.304-311.

Ji, M. and Jensen, J. R., Fuzzy training in supervised image classification, *Geographic Information Science*, (A Journal of the Association of Chinese Professionals in Geographic Information Systems (Abroad)), Vol. **2**, No. 1-2, 1996, pp. 1–11.

Kandel, A., 1986, Fuzzy mathematical techniques with applications, MA: Addison-Wesley.

Kemp, K. K., 1992, Environmental Modeling with GIS: a strategy for dealing with spatial continuity, *Proceedings of GIS/LIS'92 Annual Conference and Exposition*, Nov. 10-12, San Jose, California, Vol.1, pp.397–406.

Klir, G. J. and Folger, T. A., 1988, *Fuzzy Sets, Uncertainty, and Information*, Englewood Cliffs: Prentic Hall.

Langran, G., 1992, *Time in Geographic Information system*, London: Taylor & Francis.

Langran, G. and Chrisman, N. R., 1988, A framework for temporal geographic information. *Cartographica*, **25**, 3, 1–14.

Leung, Y. C., 1979, Locational choice: a fuzzy set approach, *Geographical Bulletin*, **15**, 28–34.

Lovett, A. A., 1995, Spatial Analysis, *Geographic Information Material for a Post-Graduate Course, Vol.2: SGIS Technology*, edited by A.U. Frank, Department of Geoinformation, Technical University Vienna, pp. 453–488.

Lowell, K.E., 1994, Probabilistic temporal GIS modeling involving more than two map classes, *International Journal of Geographical Information Systems*, **8**, 73–93.

Lillesand, T. M. and Kiefer, R. W., 1994, *Remote sensing and image interpretation* (Third Edition), Chichester: John Wiley.

Ma, Z. and Zhao, S., 1995, Using an object and rule-based merging algorithm to eliminate small regions less than minimum mapping unit, *Geographic Information Sciences*, (A Journal of the Association of Chinese Professionals in Geographic Information Systems (Abroad)), Vol.**1**, No.2, pp.110–117.

Mason, D. C., O'Conaill, M. A. and Bell, S. B. M., 1994, Handling four-dimensional geo-referenced data in environmental GIS, *International Journal of Geographical Information System*, Vol. **8**, No. 2, pp.191–215.

Middelkoop, H., 1990, Uncertainty in a GIS: A test for quantifying interpretation output, *ITC Journal*, 225–233.

Molenaar, M., 1991, Formal data structures and query spaces, *Konzeption und Einsatz von Umweltinfornmationssytemen*, edited by O. Günther, H. Kuhn, R. Mayer-Foll, and F. J. Radenmacher, Berlin: Springer Verleg. pp. 340–364.

Molenaar, M., 1994, A syntax for the representation of fuzzy spatial objects, *Advanced Geographic Data Modelling*, edited by M. Molenaar and S. de Hoop, Netherlands Geodetic Commission, New Series, Nr. 40, Delft, pp. 155–169.

Molenaar, M. 1995, Spatial concepts as implemented in GIS, *Geographic Information Material for a Post-Graduate Course*, Vol. **1**: Spatial Information, edited by A. U. Frank, Department of Geoinformation, Technical University Vienna, pp. 91–154.

Molenaar, M., 1996, A Syntactic Approach for handling the semantics of fuzzy spatial objects, *Geographic Objects with Indeterminate Boundaries*, edited by P. A. Burrough and A. U. Frank, London: Taylor & Francis, pp.207–224.

Molenaar, M., 1998a, *An Introduction to the Theory of Spatial Object Modeling*, London: Taylor & Francis.

Molenaar, M., 1998b, The fuzzy spatial extent of objects identified with remote sensing and photo-interpretation, *Data Quality in Geographic Information – From Error to Uncertainty*, edited by M. F. Goodchild and R. Jeansoulin, Paris: Hermes, pp.127–134.

Nanninga, M., 1985, The accuracy of echo sounding description in a mathematical model, *Thesis*, University of Technology Delft.

Nunes, T.L., 1991, Geographic space as a set of concrete geographical entities, *Cognitive and Linguistic Aspects of Geographic Space*, edited by D. Mark and A. U. Frank, Dordrecht: Kluwer, 9–23.

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

Peuquet, D. J., 1995. It's about time: a conceptual framework for the representation of temporal dynamics in geographic information systems, *Temporal Data in Geographic Information Systems*, edited by A. U. Frank, W. Kuhn and P. Haunold, pp.149–170.

Peuquet, D. J. and Duan, N., 1995, An event-based spatiotemporal data model (ESTDM) for temporal analysis of geographical data, *International Journal of Geographical Information Systems*, Vol. **9**, No. 1, pp. 7–24.

Peuquet, D. J. and Qian, L., 1996, An integrated database design for temporal GIS, *Advances in GIS Researches II, Proceedings of 7th International Symposium on Spatial Data Handling*, (*SDH'96*), edited by M. J. Kraak and M. Molenaar, London: Taylor & Francis, pp. 21–32.

Pipkin, J. S., 1978, Fuzzy sets and spatial choice, *Annals of the Association of American Geographers*, 68, 196–204.

Plewe, B, 1997, A representation-oriented taxonomy of gradation, *Spatial Information Theory: A Theoretical Basis for GIS, Lecture Notes in Computer Sciences*, edited by S. C. Hirtle and A.U. Frank, Berlin: Spring-Verlag, Vol. **1329**, pp. 121–136.

Poulter, M. A., 1997, On the integration of earth observation data: defining landscape boundaries to a GIS, *Geographic Objects with Indeterminate Boundaries*, edited by P.A. Burrough and A. Frank, London: Taylor & Francis, pp. 287–298.

Raafat, H., Yang, Z. and Gauthier D., 1994, Relational spatial topologies for historical geographical information, *International Journal of Geographical Information System*, Vol. **8**, No. 2, pp.163–173.

Raper, J. and Livingstone, D., 1995, Development of a geomorphological spatial model using object-oriented design, *International Journal of Geographical Information Systems*, Vol. **9**, No. 4, pp.359–383.

Reineck, H.E. and Singh, I.B., 1980, *Depositional Sedimentary Environments* (Second edition), Berlin: Springer Verleg.

Robinson, V. B. and Strahler, A. H., 1984, Issues in designing geographic information systems under conditions of inexactness, *Proceedings of* *the 10th International Symposium on Machine Processing of Remote Sensed Data*, Purdue University, Lafayette, pp.198–204.

Robison, V. B., 1988, Some implications of fuzzy set theory applied to geographic databases, *Computers, Environment and Urban systems*, **12**, 89–98.

Robison, V. B., 1990, Interactive machine acquisition of a fuzzy spatial relation, *Computer & Geosciences*, **16**, 857–872.

Roshannejad, A. A., 1996, *The Management of Spatio-Temporal Data in a National Geographic Information System*, Ph.D. Thesis, ITC, The Netherlands.

Rosch, E., 1978, Principles of categorization, *Cognition and Categorization*, edited by E. Rosch and B. B. Lloyd, Hillsdale, NJ: Erlbaum, pp.27–48.

Ruig, J. H. M. de and Louisse, C. J., 1991, Sand budget trends and changes along the Holland coast, *Journal of Coastal Research*, Vol. **7**, No. 4, pp. 1013–1026.

Ruessink, B. G. and Kroon, A., 1994, The behaviour of a multiple bar system in the nearshore zone of Terschelling, The Netherlands: 1965 – 1993, *International Journal of Marine Geology, Geochemistry and Geophysics*, No. 121, pp. 187–197.

Shi, W., 1994, *Modelling Positional and Thematic Uncertainty in Integration of Remote Sensing and GIS*, Ph. D. thesis, ITC, The Netherlands.

Shi, W. and Zhang, M. 1995, Object-oriented approach for spatial, temporal, and attribute data modeling. *Proceedings of GIS/LIS*, Bethesda: ACSM/ASPRS, AAG, URISA, AM/FM, **2**: 903–912.

Skidmore, A. K., Baang, J., and Luckananurug, P., 1992, Knowledge based methods in remote sensing and GIS, *Proceedings of 6th Australasian Remote Sensing Conference*, Wellington, Vol. **2**, pp. 394–403.

Smith, B., 1995, On drawing lines on a map, *Spatial Information Theory: A Theoritical Basis for GIS,Lecture notes in computer science*, edited by A. Frank and W. Kuhn, Berlin: Spring-Verlag, Vol. **998**, pp. 475–484.

Smith, B. and Varzi, A. C., 1997, Fiat and bona fide boundaries: towards an ontology of spatially extended objects, *Spatial Information Theory: A Theoretical Basis for GIS, Lecture Notes in Computer Sciences*, edited by S. C. Hirtle and A. U. Frank, Berlin: Spring-Verlag, Vol. **1329**, pp. 103–120.

Snodgrass, R. T., 1992, Temporal database, *Theories and Methods of Spatio-Temporal Reasoning in Geographic Space, Lecture Notes in Computer Science*, edited by A. U. Frank, I. Campari and U. Formentini, Berlin: Springer-Verlag, Vol. **639**, pp. 22–64.

Tansel, A. U., Clifford, J., Gadia, S., Jajodia, S., Segev, A. and Snodgrass, R., 1993, *Temporal Databases: Theory, Design, and Implementation*, Reading, MA: The Benjamin /Cummings Publishing Company.

Tang, A. G., Adams, T. M. and Usery, E. L., 1996, A spatial data model design for feature-based geographical information systems, *International*

Journal of Geogaphical Information System, Vol. **10**, No. 5, pp.643–659.

Usery, E. L., 1996, A conceptual framework and fuzzy set implementation for geographic feature, *Geographic Objects with Indeterminate Boundaries*, edited by P. A. Burrough and A. U. Frank, London: Taylor & Francis, pp. 71–85.

Wang, F., 1990, Fuzzy supervised classification of remote sensing images, *I.E.E.E. Transaction on Geoscience and Remote Sensing*, **28**, 194–201.

Wang, F., 1992, Improving remote sensing image analysis through fuzzy information representation, *Photogrammetric Engineering & Remote sensing*, **56** (8), 1163–1168.

Wang, F. and Hall, G. B., 1996, Fuzzy representation of geographical boundaries in GIS, *International Journal of Geographical Information Systems*, Vol. **10**, No.5, 573–590.

Wang, F., Hall, G. B., and Subaryono, 1990, Fuzzy information representation and processing in conventional GIS software: database design and application, *International Journal of Geographic Information systems*, **4**, 261–283.

Worboys, F. M., 1992, Object-Oriented Model of Spatiotemporal Information, *Proceedings of GIS/LIS '92*, San Jose, California, Vol. **2**, pp.825–834.

Worboys, F. M., 1994a, A unified model for spatial and temporal information, *The Computer Journal*, Vol. **37**, No.1, pp.26–34.

Worboys, F. M., 1994b, Object-Oriented approaches to geo-referenced information, *International Journal of Geographical Information Systems*, Vol. **8**, No.4, pp.385–399.

Worboys, F.M., Hernshaw, H., M. and Maguire, D. J., 1990, Object-oriented data modeling for spatial databases, *International Journal of Geographical Information Systems*, Vol. **4**, No.4, pp.369–383.

Yuan, M., 1995, Wildfire conceptual modeling for building GIS space-time model, *Temporal Data in Geographic Information Systems*, complied by A. U. Frank, W. Kuhn and P. Haunold, pp.47–56.

Yuan, M., 1996, Temporal GIS and spatio-temporal modeling, *Proceedings of the Third International Conference/Workshop on Integrating Geographic Information Systems and Environmental Modeling*, Santa_fa, New Mexico, U.S.A.

Zadeh, L. A., 1965, Fuzzy sets, *Information and Control*, **8**, 338–353.

Zhan, F.B., 1997, Approximation of topological relationship between fuzzy regions satisfying a linguistically described query, *Spatial Information Theory: A Theoretical Basis for GIS, Lecture Notes in Computer Sciences*, edited by S. C. Hirtle and A. U. Frank, Berlin: Spring-Verlag, Vol. **1329**, pp.509–510.

Zimmerman, H. J., 1984, *Fuzzy Set Theory and its Applications*, Hingham, MA: Kluwer Academic.

Zuidam, R. A. van, Ploh, C. and Genderen, J. L. van, 1994, Synergy of remotely sensed data for coastal environmental studies: the Ameland-Waddensea example, northern Netherlands, *Proceedingsof the Second Thematic Conferences on Remote Sensing for Marine and Coastal Environments*, New Orleans, Louisiana, USA, pp. I-323–334.

Publications which cover the research in this thesis

Cheng, T. and Molenaar, M, 1999, Diachronic analysis of fuzzy objects, under review of *Journal* of *GeoInformatica.*

Cheng, T. and Molenaar, M, 1999, Objects with fuzzy spatial extent, accepted by *Photogrammetric Engineering and Remote Sensing.*

Blok, C., Kobben, B., Cheng, T. and Kuterema, A. A., 1999, The visualization of relationships between spatial patterns in time by cartographic animation, accepted by *Cartography and Geographic Information Systems.*

Molenaar, M., and Cheng, T., 1999, Fuzzy spatial objects and their dynamics, under review of the *International Journal of Geographical Information Systems.*

Zuidam, R. A. van, Farifteh, J., Eleveld, M. and Cheng, T., 1998, Research in remote sensing, dynamic models and GIS applications for integrated coastal zone management, *Journal of Coastal Conservation*, No. 4.

Cheng, T. and Molenaar, M., 1998a, A process-oriented spatio-temporal data model to support physical environmental modeling, *Proceedings of the 8th International Symposium on Spatial Data Handling*, edited by T. K. Poiker and N. Chrisman, pp. 418–430.

Cheng, T., Molenaar, M. and Bouloucos, T., 1997, Identification of fuzzy objects from field observation data", *Spatial Information Theory: A Theoretical Basis for GIS, Lecture Notes in Computer Sciences*, edited by S. C. Hirtle and A. U. Frank, Berlin: Spring-Verlag, Vol. **1329**, pp. 241–259.

Cheng, T. and Molenaar, M., 1999, Syntactic representation of three fuzzy object models, to be presented in the *International Symposium on Spatial Data Quality*, Hong Kong Polytechnic University, 18th – 20th July, Hong Kong.

Cheng, T., Molenaar, M. and Bouloucos, T., 1998, Fuzzy object models and their application, *Proceedings of International Conference of Spatial Information Science, Technology and its Applications* (SIST'98), Dec. 13-16, Wuhan, pp. 408–416.

Cheng, T. and Molenaar, M., 1998b, The identification and monitoring of objects with fuzzy spatial extent, *International Archives of Photogrammetry and Remote Sensing (ISPRS)*, Vol. **32**, Part 3, pp. 207–212.

Cheng, T. and Molenaar, M., 1998c, Change of fuzzy objects, paper presented at *First Conference of Association of Geographic Information*

Laboratories in Europe (AGILE), Enschede, The Netherlands, April 23-25.

Molenaar, M., and Cheng, T., 1998, Fuzzy spatial objects and their dynamics, *International Archives of Photogrammetry and Remote Sensing (ISPRS)*, Vol. **32**, Part 4, pp.389–394.

Cheng, T. and Molenaar, M., 1997, Dynamics of fuzzy objects, *Proceeding of the International Workshop on Dynamic and Multi-dimensional GIS*, edited by Y. C. Lee and Z. Li, Hong Kong Polytechnic University, pp.49–63.

Cheng, T., Kainz, W. and Zuidam, R. A. van, 1996, Coupling GIS and environmental modelling: the implications for spatio-temporal data modeling, *International Archives of Photogrammetry and Remote Sensing (ISPRS)*, Vol. **III**, pp.849–856.

Cheng, T., 1995, Developing a spatio-temporal GIS shell to support coastal environmental modeling, *COSIT'95 Doctoral Consortium*, compiled by W. Kuhn and S. Timpf, pp.7–14.

Cheng, T., Eleveld, M. A. and Zuidam, R. A. van, 1995, Creation of a 4-D GIS working platform for coastal environmental monitoring and management, *Proceedings of Third Thematic Conference of Remote Sensing for Marine and Coastal Environments, Technology And Applications*, Seattle, Vol. **2**, pp. 260–272.

Cheng, T., Zuidam, R.A. van and Kainz, W., 1995, A unified spatio-temporal data model for 4-D GIS, *Proceedings of GIS/LIS'95 Annual Conference and Exposition*, Nov. 12-16, Nashville, USA, pp. 967–976.

Eleveld, M. A., Cheng, T. and Zuidam, R. A. van, 1995, Towards a decision support system for coastal zone management by applying morphodynamic modelling with remote sensing data input in a 4-D GIS environment, *Proceedings of Third Thematic Conference of Remote Sensing for Marine and Coastal Environments, Technology And Applications*, Seattle, Vol. **1**, pp. 256–265.

Samenvatting

De natuurlijke omgeving is complex doordat ze een polithematisch en dynamisch karakter heeft. De representatie ervan in een GIS vereist een nieuwe benadering die meer mogelijkheden biedt dan de gangbare weergave in termen van crispe en statische objecten. Dit vereist een ruimtelijk datamodel waarin het objectbegrip verder is uitgewerkt zodat ook vage en dynamische objecten gerepresenteerd kunnen worden. Hierin ligt de motivatie van het onderzoek waarvan in deze dissertatie verslag wordt gedaan. Het hoofddoel van dit onderzoek was om een spatio-temporeel datamodel te ontwikkelen voor de representatie van objecten met een vage ruimtelijke extensie. Drie onderwerpen zullen in dit verband worden behandeld:

1. De identificatie van vage objecten,
2. De detectie van veranderingen en dynamiek van vage objecten,
3. De representatie van vage dynamische objecten in een spatio-temporeel datamodel.

Een procedure bestaande uit zes stappen wordt geformuleerd voor de identificatie van vage objecten uit geregionaliseerde veld (raster) data: bemonstering, interpolatie (regionalisatie), classificatie, segmentatie, samenvoeging en identificatie. De onzekerheidsaspecten van deze stappen worden geanalyseerd en het effect ervan op de eindresultaten wordt geëvalueerd. Er worden drie vage object modellen beschreven die voor verschillende gebruikscontexten geschikt zijn. Verder worden de begrippen "conditionele ruimtelijke extensie", "conditionele grens" en "overgangszone tussen vage objecten" gedefinieerd en geformaliseerd op basis van de Formele Data Structuur. Hierbij wordt uitgelegd hoe thematische onzekerheden doorwerken op de geometrie van ruimtelijke objecten, waarbij duidelijk wordt hoe de begrippen *"existentiële onzekerheid"* en *"extensionele onzekerheid"* gerelateerd zijn. Het ruimtelijk effect van thematische onzekerheid wordt zichtbaar doordat de onzekerheid of een rastercel tot een object hoort afhangt van de classificatieonzekerheid van de cel.

De methode voor detectie van veranderingen van objecten door de tijd is gebaseerd op de identificatie van de objecten en de toestandsveranderingen van hun vage ruimtelijke extensies op verschillende waarnemingstijdstippen. Indicatoren voor de vaststelling van de similariteit van objecten worden geëvalueerd op basis van de overlap van hun ruimtelijke extensie op opeenvolgende waarnemingstijdstippen. Verschillende combinaties van

waarden voor deze indicatoren duiden op verschillende typen van tijdsrelaties tussen de objecten. Ruimtelijke segmenten van twee opeenvolgende tijdstippen die elkaar sterk overlappen en dus een grote similariteit vertonen worden dan geacht de opeenvolgende toestanden van een en hetzelfde object te representeren. De historische levenslijnen van de objecten worden gevormd door deze segmenten door de tijd te verbinden, daarmee worden ook veranderingen van objecten vastgesteld. Vanuit de relaties tussen de segmenten kunnen interacties tussen objecten achterhaald worden. Deze worden beschreven door processen als, verschuiving, samenvoeging, splitsing, etc. De onzekerheden aangaande de gedetecteerde veranderingen worden geëvalueerd in de vorm van een serie veranderingskaartjes op verschillende onzekerheidsniveaus.

Als derde onderdeel van dit onderzoek werd een spatio-temporeel datamodel ontwikkeld waarin de verandering van en interactie tussen objecten gerepresenteerd kan worden. Dit model is gebaseerd op de geformaliseerde beschrijving van de toestanden van objecten en de processen waaraan ze deelhebben. Het model is een uitwerking van het entity-relationship model en heeft een sterstructuur, het wordt daarom het Star Model genoemd. Dit model werd omgezet in een relationeel model, zodat het implementeerbaar is in gangbare commerciële databases. A prototype van dit spatio-temporele datamodel is geïmplementeerd in Arc View en toont het voorbeeld van Ameland. Het gebruikersinterface en de queries zijn geprogrammeerd in Avenue, de programmeertaal van Arc View.

De procedure waarmee de objecten geïdentificeerd worden combineert de huidige veld- en objectbenaderingen. Daarom is het object begrip veralgemeniseerd tot objecten met een vage extensie (fuzzy spatial extent), dit concept verbindt de object -en veldkarakteristieken van natuurlijke terreinkenmerken. Deze vage objecten hebben conditionele grenzen; daarbinnen vertonen ze veldkenmerken, d.w.z. hun oppervlak vertoont terreinkenmerken met een continue verloop. De gekozen representatie is geschikt voor zowel crispe als vage objecten. Vage objecten vertonen overgangszones, waarin eventueel conditionele grenzen gedefinieerd kunnen worden. De crispe objecten vormen dan een speciaal geval hiervan, doordat ze scherpe grenzen vertonen. Daarenboven kunnen voor de objectidentificatie zowel de grens- als de pixelrepresentatie in dit veralgemeniseerde object begrip gehanteerd worden, omdat de onzekerheden van de objecten in de formele data structuur uitgedrukt kunnen worden.

Het voorgestelde datamodel heeft een algemeen karakter, waaruit andere modellen afgeleid kunnen worden. Het maakt de analyse van gegevens over tijdseries mogelijk vanuit verschillende gezichtspunten, door locatiegerichte, tijdgerichte, objectgerichte en procesgerichte bevragingen. Hiermee kan het gedrag van dynamische ruimtelijke complexen bestudeerd worden. Meerdere tijdseries kunnen worden gegenereerd die de levenslijnen van objecten weergeven. Daarmee geeft het model processen weer die aangrijpen op de ruimtelijke en thematische aspecten van de individuele objecten en van ruimtelijke complexen. Omdat het model veranderingen in de tijd expliciet opslaat maakt het bevragingen mogelijk over tijdsrelaties tussen objecten, zowel als analyses van verschillende tijdseries.

De resultaten van dit onderzoek vormen een theoretische en een praktische bijdrage aan de ontwikkeling van spatio-temporele datamodellen voor objecten met een vage ruimtelijke extensie.

Curriculum Vitae

Tao Cheng received her Bachelor of Science degree (1989) in Surveying Engineering and her Master of Science degree (1992) in Photogrammetry and Remote Sensing from Wuhan Technical University of Surveying and Mapping (WTUSM), P. R. China. She was a researcher and lecturer at the Research Center for GIS of WTUSM from 1992 to 1994. She started her PhD research project at ITC in 1995.

Tao Cheng has been involved in various scientific research projects covering topics in intelligent GISs and image processing. Her current research interests span uncertainty in geographic information, spatio-temporal data models, and the integration of GIS, remote sensing and decision support systems for coastal environmental monitoring and management. She is the author and co-author of a number of scientific articles.