INPE-00000-TDI/0000

ALGEBRAIC MODELLING OF SPATIOTEMPORAL OBJECTS: UNDERSTANDING CHANGE IN THE BRAZILIAN AMAZON

Olga Regina Fradico de Oliveira Bittencourt

Doctorate Thesis at the Post-graduation Course in Applied Computing Science, advised by Dr. Gilberto Câmara and Dr. Lúbia Vinhas, approved August 26, 2009.

The original is available at:
<http://urlib.net/sid.inpe.br/iris@........................>

INPE
São José dos Campos
2009

Publicado por:

Instituto Nacional de Pesquisas Espaciais (INPE)
Gabinete do Diretor – (GB)
Serviço de Informação e Documentação (SID)
Caixa Postal 515 – CEP 12.245-970
São José dos Campos – SP – Brasil
Tel.: (012) 3945-6911
Fax: (012) 3945-6919
E-mail: pubtc@sid.inpe.br

Solicita-se intercâmbio
We ask for exchange

Ministério da
Ciência e Tecnologia

BRASIL
UM PAÍS DE TODOS
GOVERNO FEDERAL

INPE-00000-TDI/0000

ALGEBRAIC MODELLING OF SPATIOTEMPORAL OBJECTS: UNDERSTANDING CHANGE IN THE BRAZILIAN AMAZON

Olga Regina Fradico de Oliveira Bittencourt

Doctorate Thesis at the Post-graduation Course in Applied Computing Science,
advised by Dr. Gilberto Câmara and Dr. Lúbia Vinhas,
approved August 26, 2009.

The original is available at:
<http://urlib.net/sid.inpe.br/iris@........................>

INPE
São José dos Campos
2009

Dados Internacionais de Catalogação na Publicação

Bittencourt, Olga R. F. de Oliveira.

Algebraic modelling of spatiotemporal objects: understanding change in the Brazilian Amazon / Olga Regina Fradico de Oliveira Bittencourt. - São José dos Campos: INPE, 2009
98p. ; (INPE-0000 -TDI/00)

Thesis (Doctorate in Applied Computing Science) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2009.
Advisers : Dr. Gilberto Câmara and Dr. Lúbia Vinhas

1. Algebraic modelling. 2. Landscape objects. 3. Spatiotemporal evolution. 4. Deforestation patterns. 5. Modelagem algébrica. 6. Objetos de paisagem. 7. Evolução espaço-temporal. 8. Padrões de desmatamento I. Algebraic modelling of spatiotemporal objects: understanding change in the Brazilian Amazon.

CDU 04.414.23: 423.45

FOLHA DE APROVAÇÃO

“Porque é frágil a memória dos homens e para que, com o tempo, não caiam no esquecimento os feitos dos mortais, nasceu o remédio da escrita para que, por meio dele, os fatos passados se conservem como presentes para o futuro.“

Arenga de 1260 (Viseu, Arquivo do Museu de Grão Vasco, PERG / 08)

Para minha família.

AGRADECIMENTOS

É na hora de agradecer, que se olha para trás e se tem a oportunidade de rapidamente tomar ciência do processo de transformação pelo qual passou o pesquisador. Com essa consciência, os agradecimentos deixam de ser simples obrigado e passam a ter cada vez mais sentido. Assim, depois de tão longa jornada é com prazer que posso ratificar tudo o que já foi dito:

Ao Jorge, que caminha comigo dedicando amor, tempo e compreensão. Você possui o mérito de me fazer ver que sou capaz de aspirações acima de minhas ambições individuais e me ajuda a encontrar forças para superar aquilo que penso ser o meu limite. À Gabriela, que apesar de ter chegado há pouco tempo já trouxe tantas alegrias e se tornou tão essencial em nossas vidas. A presença de vocês me faz dar passos mais largos em direção aos meus sonhos mais verdadeiros... "Sonhos que sonhamos juntos são realidade". A vocês: obrigada por tudo e os meus mais sinceros agradecimentos.

Aos meus pais, pela sólida formação que me permitiram chegar até aqui. A minha mãe, minha família, meus amigos e também, Jorge e Gabi: obrigada a vocês que acreditaram, compartilharam e alimentaram esse ideal, incentivando-me a prosseguir na jornada, fossem quais fossem os obstáculos. Que com uma palavra, um gesto, e até mesmo um olhar me fizeram perceber que era capaz. Muito obrigada por todas as vezes que me ouviram e deram seu apoio. A vocês o meu carinho, gratidão e desculpas pela ausência nos momentos em que minha companhia era importante. Em mais essa conquista há muito de suas presenças.

Em especial aos amigos Adair, Eduilson, Eva, Giovana, Ilka, Joyce, Karla, Missae, Sérgio, aos amigos do GeoPro e a todos com quem convivi nessa jornada: "Durante anos de convívio desfrutamos bons momentos. Superamos dificuldades, enriquecemos nossas vidas. Cada um levará consigo qual seja o seu caminho: a saudade e o carinho, daquele que se tornou... amigo."

Ao meu orientador Gilberto Câmara por seu apoio, inspiração e idéias que possibilitaram a execução deste trabalho. A minha orientadora Lúbia pelo carinho e persistência. Vocês me fizeram novamente enxergar que a pesquisa é um universo quase mágico. Impressiona a capacidade de transformação que ela exerce em nossas vidas. Podemos falar até mesmo de um antes e um depois: um antes de ansiedades e expectativas; um depois de experiências, conhecimento, crescimento e ainda mais expectativas. Obrigada pela dedicação, pela amizade e por orientarem esse caminho de constantes descobertas e transformações.

A Deus, que nos momentos em que me senti mais frágil diante da vida clareaste minha visão e fizeste com que eu enxergasse com nitidez o futuro que meu esforço poderia trazer. Tu me acompanhaste sempre, dando-me forças e iluminando o meu caminho. Agora tenho mais certeza deo que nunca estive sozinha e que a conquista que comemoro é nossa.

Enfim, a todos aqueles que de alguma forma contribuíram para que este trabalho chegasse a bom termo, o meu sincero: Muito Obrigada!

Olga Bittencourt

ABSTRACT

This thesis proposes an algebra to describe how spatiotemporal objects evolve, named *geospatial algebra,* and a model to apply it. This algebra is composed of a set of operations, axioms and rules defined by the application. Specifically, we handle *evolving objects*, which are objects that evolve by changing their boundaries and attributes. These objects appear in cases of land change in rural and urban areas. We also propose operators to track the history of a set of *evolving objects* as well as the individual history of each object in the set. In addition, we developed a system to use the algebra and analyze time series of deforestation objects in three case studies of land use and land cover in the Brazilian Amazon. Our results show that, by tracking the object evolution, we can discover and quantify important issues related to the patterns of deforestation in the Brazilian Amazon.

MODELAGEM ALGÉBRICA DE DADOS ESPAÇO-TEMPORAIS: ENTENDENDO AS MUDANÇAS NA AMAZÔNIA

RESUMO

Essa tese propõe uma álgebra, a *Álgebra GeoEspacial*, para descrever a evolução de objetos espaço-temporais. Ela é composta de um conjunto de operações, axiomas e regras definidas para cada aplicação. Especificamente, nós trabalhamos com *objetos evolutivos*, que são objetos que evoluem alterando suas bordas e seus atributos. Eles ocorrem, por exemplo, em casos de mudança de uso e cobertura do solo em áreas urbanas e rurais. Nós também propomos operadores para tratar histórias de evolução de um conjunto de *objetos evolutivos*, assim como tratar as histórias individuais de cada objeto no conjunto. Nós aplicamos a álgebra para analisar séries temporais de áreas que sofreram mudança de uso e cobertura do solo Amazônia. Nossos resultados mostram que acompanhando a evolução dos objetos, somos capazes de descobrir e quantificar relevantes informações a respeito dos padrões de desflorestamento.

SUMMARY

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

1 INTRODUCTION

The computational modelling of geospatial information continues to be, after decades of research, a challenging problem that defies a definitive solution. Since computer models assign human-conceived geographic entities to data types, the theory associated with assigning types and classes to elements of the geographic world has been a focus of intensive research. Recently, there has been interest in modelling and representing geospatial objects whose properties change (FRANK, 2003; WORBOYS, 2005; GRENON, SMITH, 2003; GALTON, 2004; GOODCHILD et al., 2007). Such interest has a practical motivation. New generations of satellites and mobile devices have enabled new forms of communication and spatial information processing. As the sensor web or environmental monitoring expands, we are becoming overwhelmed with streams of data that provide information about change.

1.1 Problem definition

Representing change in a GIS (Geographic Information Systems) is not only an issue of handling time-varying data. It also concerns how objects gain or lose their identities, how their properties change over long time scales, what changes happen simultaneously, and the laws of nature and the interactions among people that result in change. We consider that finding a unique conceptualization of spatiotemporal models and operators is (arguably) an unsolvable problem, as stated by GRENON and SMITH (2003). PELEKIS et al. (2004) and RODDICK et al. (2004) reviewed the different types of spatiotemporal data models proposed in the literature: there is no consensus on how to model or handle spatiotemporal data. Spatiotemporal models are mostly application-specific: each of them focuses on specific spatiotemporal data aspects. To solve a problem of spatiotemporal data modelling, we need to consider the needs and constraints of the application domain and choose a suitable approach.

We want to model and handle data related to sets of *geospatial objects*. We consider a *geospatial object* as an entity represented in our computer models, with a unique identity, located in geographic space and with well-defined boundaries and attributes. The location, boundaries and attribute values of a *geospatial object* can change during its lifetime. Specifically, we are interested in a geospatial problem that occurs frequently in land information: the problem of *land use* and *land cover change*. The 'change' is defined by SINGH (1989) as the different states that objects adopt in distinct observed timestamps. Suppose there is a natural land cover, such as a tropical forest. Migrant farmers and settlers move into the region, and remove the original land cover and replace the forest with pasture and agriculture. New regions might be deforested and incorporated into existing farms. These new regions can be detected in remote sensing images through various methods for land cover change detection (SINGH, 1989; ROGAN, CHEN, 2004). We adopted the definition of *landscape objects* proposed by SILVA et al. (2008) to identify these regions detected in remote sensing images by image segmentation or visual interpretation.

At INPE (National Institute for Space Research), the PRODES (Program for Deforestation Assessment in the Brazilian Amazon) (INPE, 2009) has used remote sensing imagery to generate information about deforested areas in the Brazilian Amazon since 1988. PRODES methodology yearly detects patches of deforested areas to calculate the annual deforestation rate. Besides this main goal of the program (to calculate the annual deforestation rate), we also get multitemporal snapshots of *landscape objects* in the same region. Despite this, researchers do not fully exploit the temporal aspects of this dataset because they are looking at one snapshot at a time. We propose to exploit this multitemporal dataset by considering *landscape objects* as *geospatial objects* bound to specific locations but whose geometry, topology and properties change. Following this, it is possible to detect and quantify patterns that emerge from processes related to change. Such patterns are observed in remote sensing data but there is no comprehensive theory on how they evolve or on how to recover their history. To do that, we need to develop modelling tools that include needs and constraints specific to the application domain.

1.2 Related work

In recent years there has been a large growth of geographical datasets and associated research in GIS. One of the aims of these studies is the accurate representation of objects and phenomena of the real world. One challenge is to handle the complexity of data that combine space and time (GALTON, 2004; HORNSBY, EGENHOFER, 2000; MEDAK, 2001). The current generation of spatial database management systems (DBMSs) models static spatial data. New technologies such as dynamic and location-based systems require different methods of data modelling, data representation and algorithms. As a result, they motivate research in spatiotemporal data. Handling this data requires extensions of spatial databases (ERWIG, SCHNEIDER, 2002; GÜTING et al., 2004; SELLIS et al., 2003), notions such as *trajectory* (GÜTING, SCHNEIDER, 2005), specialized query methods (SISTLA et al., 1997), spatiotemporal predicates (ERWIG et al., 1999), spatiotemporal relationships (CLARAMUNT, JIANG, 2001), indexing techniques (ŠALTENIS et al., 2000) and spatiotemporal query languages (HORNSBY, EGENHOFER, 2000; HUANG, CLARAMUNT, 2002; ERWIG, SCHNEIDER, 2002a).

ANDRIENKO et al. (2008) review some concepts related to change that we adopted in this work. They define 'movement' as the change in the physical position of an entity about some reference system, geographic space, within which one can assess position. The 'trajectory' is defined as the path made by the moving entity through the space in which it moves. Objects that produce movement, called *moving objects*, are relevant when we monitor the spatial and temporal positions of objects such as planes, storms or cars. Global positioning systems (GPS) technology produces many *moving object* data. The widespread availability of location-based systems motivated the development of *moving* *objects* databases and related work (ERWIG, SCHNEIDER, 2002; GÜTING, SCHNEIDER, 2005).

To integrate *moving objects* in spatial databases, one of the most promising approaches is the design of algebraic data types, such as the one proposed by GÜTING et al. (2003). Their algebra includes data types to model moving points (objects where only the position in space is relevant) and moving regions (objects where the position and the time-dependent extent are relevant). Their operators focus on issues related to the trajectories of objects, so that we can answer questions such as "*Given the trajectories of snowstorms and airplane flights, which flights went through a snowstorm?*" GÜTING et al. (2004) described the implementation of their algebra for moving objects in the SECONDO environment, an extensible and modular DBMS.

Besides moving objects, we cite two other types of *geospatial objects*: *socio-economic objects* and *landscape objects*. Both are distinct from *moving objects*, which are focused on trajectories because they focus on another kind of change. They are bound to specific locations, but their geometry, topology and properties change. We can illustrate typical changes on *landscape objects* with land use and land cover changes. In *socio-economic objects*, typical changes occur, for example, in parcels and roads in an urban cadastre. The main distinction between them is that while changes to *socio-economic units* are individually detailed because they depend on social and political actions and need to be previously known to be modelled, changes to *landscape objects* are automatically detected by extraction methods and need to be processed based on specific constraints to give informative data about the objects and their changes. This distinction is more practical than theoretical, but it is necessary to facilitate the detection and understanding of object changes.

Handling *socio-economic* and *landscape objects* requires tracking the changes that occurred during an object's lifetime, such as creation, splitting and merging (HORNSBY, EGENHOFER, 2000; MEDAK, 2001). These objects focus on identity, life and genealogy concepts. ‘Life’ defines the set of changes that an object experiences during its existence. ‘Identity’ encompasses the traits that distinguish each object during its existence. ‘Genealogy’ is concerned with the interaction between an object and the state of the object at distinct times. MEDAK (2001) explores this idea, proposing an algebra for modelling change in *socio-economic units,* with operations such as *aggregation* and *fusion*, to answer questions such as “*When was this parcel divided?*” HORNSBY and EGENHOFER (2000) propose to model change using a Change Description Language (CDL) that includes operations such as *create*, *destroy*, *eliminate* and *reincarnate*.

Representing change in *landscape objects* is sometimes restricted by available data, which consist of the geometry and attributes of the objects. Handling this data, BECKER (1997) characterizes the main processes of land cover change in the Brazilian Amazon by linking agricultural producers with their different strategies for land use. ESCADA (2003) also associates agents of land use change to the occupation process. She defines a typology that represents the main processes associated with different categories of rural properties existing in a region, concluding that different agents involved in the land use change (e.g. small farmers, farmers, cattle breeders) can be distinguished by their different land use patterns.

SILVA et al. (2008) propose a method to classify *landscape objects* obtained from a remote sensing image database using data mining, digital image processing, and landscape ecology theory. Their results show that analyzing *landscape objects* on images from distinct dates is an effective alternative to identify agents and model land use change. However, they do not discuss how objects evolve in time. Extending this work, MOTA et al. (2009) propose to add to this by developing a method that uses previous cases as well as knowledge elicited from a specialist, a technique called Case-Based Reasoning (CBR), to discover rules to construct the evolution of objects. However, they do not discuss how to model the rules of evolution and the geospatial objects.

1.3 Hypothesis, objectives and contributions

New approaches to handle the evolution of spatiotemporal data can help domain experts increase their knowledge about the processes that result in land use and land cover changes. This thesis takes on this challenge, by proposing an algebra to model geospatial objects. Our specific questions are:

> "*How can we handle geospatial objects that evolve in space and time?*"
>
> "*How can we rebuild the history of landscape objects, given a set of snapshots of the area?*"
>
> *"How can we discover deforestation patterns in the Brazilian Amazon?"*

To address these questions, we have considered the following hypothesis:

1. *Landscape objects* are characterized by different types. A *geospatial algebra* with specific operations handles these types and tracks their evolution over time.

2. Deforestation processes in the Brazilian Amazon generate patterns of change. The analysis of landscape objects acquired from multitemporal remote sensing images allows these patterns to be characterized and quantified.

The main contribution of this thesis is the development of an algebra to model the evolution of geospatial objects. The algebra comprises a set of operations, axioms and rules defined by the application. It also comprises operators to track the history of each individual object in the set. In addition, we developed a system to use the algebra in three case studies of land use and land cover in the Brazilian Amazon. We aimed to discover and quantify patterns of deforestation to corroborate the second hypothesis. The system enables users to assess the patterns of change and their evolution in time, to analyze them, to adjust the rules according to field knowledge about the process and to make new inferences about the patterns of evolution.

The author of this thesis is part of the research group on GeoInformatics at INPE focused on Formal Models for Spatiotemporal Data (BITTENCOURT et al., 2007; COSTA et al., 2006) and Geographical Data Mining (MOTA et al., 2008; MOTA et al., 2009; SILVA et al., 2008).

1.4 Thesis layout

This thesis is structured in four chapters and two annexes:

1. Chapter 2 describes the concepts of *geospatial evolution* with its *geospatial algebra*.

2. Chapter 3 describes the application of the *geospatial algebra* for discovering patterns of land cover change evolution in the Brazilian Amazon, through two real case studies.

3. Chapter 4 presents some conclusions, recommendations and suggestions for future directions.

4. Annex A describes the computational environment used.

5. Annex B presents the road map to handle deforestation objects from PRODES dataset.

2 DESCRIBING HOW LANDSCAPE OBJECTS EVOLVE USING A GEOSPATIAL ALGEBRA

2.1 Introduction

We live in a changing world. Our urban and rural landscapes are being altered like never before. More and more people live in cities, and many traditional rural areas are being altered or destroyed for agricultural production. Change is stronger in the emerging economies of the developing world, such as China, Brazil and India, which are growing faster than the developed economies. This acceleration of change provides a strong motivation for research in GeoInformatics: our tools and methods for representing and handling geospatial data should be capable of dealing with change. The 'change' is defined by SINGH (1989) as the different states that objects adopt in distinct observed timestamps. Indeed, there is much work in the literature about modelling and representing change in geospatial objects (FRANK, 2003; WORBOYS, 2005; GRENON, SMITH, 2003; GALTON, 2004; GOODCHILD et al., 2007). However, we need to progress further, both as regards theories of change and as well as in our methods and computer representations of change.

Representing change in a GIS is not only an issue of handling time-varying data. It also concerns how objects acquire or lose their identity, how their properties change over time, what changes happen simultaneously, and what the laws of nature and the interactions among people that cause change. As stated by GRENON and SMITH (2003), coming up with spatiotemporal models and operators is difficult. They argue that one could build a model of reality by defining geospatial objects as basic entities and then describing the forces that modify them. This is the so-called *endurantist* perspective. Alternatively, one could build a model of reality by defining the forces of change (processes) as basic entities and then describing how these forces modify the objects. This is the *perdurantist* perspective. GRENON and SMITH (2003) state that "*if we* *want to do justice to the whole of reality in non-reductionistic fashion, then we need both types of component.*" Although we agree with this view, we recognize that building computational models that represent both *endurantist* and *perdurantist* views of reality is not always possible. To solve a problem of spatiotemporal data modelling, we need to consider the needs and constraints of the application domain and choose a suitable approach.

In this thesis, we define a *geospatial object* as a spatiotemporal entity represented in our computer models, such that it has a unique identity, is located in geographic space and has well-defined boundaries and attributes. The location, boundaries and attribute values of a *geospatial object* can change during its lifetime. We are interested in *geospatial objects* of three types: *moving, socio-economic* and *landscape objects*. *Socio-economic* and *landscape objects* are bound to specific locations, but their geometry, topology and properties change. We can illustrate typical changes to *landscape objects* with land use and land cover changes. In *socio-economic objects*, typical changes occur for example in parcels and roads in an urban cadastre.

A *moving object* is distinct because it changes its position continuously. In most applications dealing with *moving objects* (such as urban transportation systems), their geometries does not change. In contrast, *socio-economic* and *landscape objects* change their geometries often, such as when two parcels merge. The distinction between *socio-economic* and *landscape objects,* and *moving objects* is thus based on practical rather than theoretical considerations. This view of distinguishing objects of different types is also supported by PELEKIS et al. (2005), who review different spatiotemporal data models proposed in the literature. It would be possible to build models of object change that handle more than one case (e.g., GOODCHILD et al., 2007). However, our assumption simplifies the task of building land information systems that handle spatiotemporal data.

In the following, we restrict our discussion to *socio-economic* and *landscape objects*. Handling these has some common points: for example, they require tracking the changes that occurred during an object's lifetime, such as creation, splitting and merging. Because of this, we apply some *socio-economic* definitions to model *landscape objects*. Their main distinction is that changes to *socio-economic* units are individually detailed because they depend on social and political actions, and these changes need to be previously known to be modelled. On the other hand, changes to *landscape objects* can be automatically detected by extraction methods and are processed based on application-specific constraints to give information data about the objects and their changes.

We are interested in a problem that occurs frequently in land information: the problem of *land use* and *land cover change*. Suppose there is a natural land cover, such as a tropical forest. Migrant farmers and settlers move into the region. They remove the original land cover and replace the forest with pasture and agriculture. New regions might be deforested and incorporated into existing farms. These new regions can be detected in remote sensing images through various change detection methods, such as those related by SINGH (1989) and ROGAN and CHEN (2004). SILVA et al. (2008) defined these structures detected in remote sensing images by image segmentation or visual interpretation as *landscape objects*.

As a result of the image analysis, we get snapshots of *landscape objects* at different times. This is a common situation in land information systems, where data comes from ground surveys or multitemporal remote sensing images. Then an important question is: "*How can we rebuild the histories of landscape objects, given a set of snapshots of the area*"? The input is the state of the world at discrete times $t_1, t_2, \ldots, t_n$. At this point, we have a set of objects with no history at each time. To build a historical dataset, we need to consider issues such as "*When can an object be considered an evolution of another object in a previous time?*" and "*Which objects at a time t resulted from the evolution of other objects in the same time?*"

There are two main approaches to this problem. The first is to represent the different types of objects and how they interact. We then address questions such as: "*What changes occurred during the object's lifetime?*" and "*When was the object created and who are its ancestors?*" In conceptual terms, we create an ontology of *endurants*, or a SNAP ontology (GRENON, SMITH, 2003). The alternative is to describe change in terms of processes and events: "*What were the causes of change?*", "*What are the different agents of change and how do they interact?*" To address these and similar questions, we would need to build an ontology of *perdurants* (a SPAN ontology (GRENON, SMITH, 2003)).

One example of the process-based approach is agent-based modelling of urban dynamics (BATTY, 2005). In these models, agents modify locations. Agent-based modelling focuses on building planning scenarios, rather than reproducing actual data patterns. Another example of a SPAN ontology is event-based calculus (WORBOYS, 2005) which captures an object's history using events that are external to the objects themselves. To use event-based calculus, we need to identify the possible events; we then set up the possible outcomes of applying these events on objects (KLIPPEL et al., 2007). Event-based techniques have been proven useful in applications such as traffic models. HORNSBY and COLE (2007) propose a model for dealing with the semantics of traffic objects using events such as *departure* and *arrival*.

In most *land information systems*, we are constrained by the available data, which consist of the geometry and attributes of the objects. Information about object types can be extracted from these datasets. However, information about the causes of change is usually not available. The existing proposals to handle *moving objects* do not fit to solve our problem. It is because *moving object* research focuses on the 'trajectories' of objects and does not allow the use of the types of objects as main information to build the process of change. *Socio-economic* approaches have similarities with *landscape objects* but still do not suit for our problem because we usually do not know either detailed individual information or the causes of change that govern the change process.

Hence we developed a *geospatial model* to solve our problem using a SNAP ontology. This consists of an abstract data type ('geospatial object') with subtypes, and a limited number of spatial operations: create, update, merge, and split. Using polymorphism, these operations handle objects of different types. The system builds a historical dataset by storing data on changes in individual objects. We call the model a *geospatial evolution model*.

The remainder of this chapter is organized as follows: Section 2.2 discusses why we need *geospatial objects* of different types to model change. We describe our model of *geospatial evolution* in Section 2.3 and the proposed *geospatial algebra* in Section 2.4. Section 2.5 presents the modelling of a real-life experiment using our algebra. We conclude the Chapter by discussing the potential and the limits of our model.

2.2 Landscape objects

As outlined above, we deal with *geospatial objects*. Specifically, we are interested in *landscape objects*, which we consider to be conceptually different from *moving* and *socio-economic objects*. This section discusses why we need to have objects of different types and how we extended and applied concepts from *socio-economic objects* to describe the evolution of *landscape objects*. This work is an extension to existing proposals of *lifeline models* (HORNSBY, EGENHOFER, 2000; MEDAK, 2001) that handle *socio-economic objects*. *Lifeline models* use three ideas: *identity*, *life*, and *genealogy*. 'Identity' is the characteristic that uniquely distinguishes one object from another. 'Life' is the time period during which an object exists in the model. 'Genealogy' refers to an object's relation to other objects over time. *Lifeline models* include primitives (e.g., split or merge) and operations (e.g., create or destroy) that can be performed on objects.

One existing approach to *lifeline models* is to use a detailed set of operations. For example, HORNSBY and EGENHOFER (1997) consider the case where the province of Quebec separates from Canada. To deal with such cases, they define operations such as secede, reincarnate, splinter, dissolve, continue_existence and continue_nonexistence. Each operation handles a specific situation (HORNSBY, EGENHOFER, 1997). Although this approach is feasible, it leads to a large set of operations. The number of operations is bound to increase as the complexity of the information system grows. In addition, the historical description needs to be done interactively by a domain expert. Such an approach may not work for *land information systems* with lots of *landscape objects*, where manual intervention by experts for each situation would be clumsy.

Handling landscape dynamics information is crucial to understanding and to representing landscape changes. However, representing changes in *landscape objects* is sometimes restricted by available data, which consist of the geometry and attributes of the objects. Handling this data, BECKER (1997) characterizes the main processes of land cover change in the Brazilian Amazon by linking agricultural producers with their different strategies for land use. ESCADA (2003) also associates agents of land use change to the occupation process. She defines a typology that represents the main processes associated with different categories of rural properties in a region, concluding that different agents involved in the land use change (e.g., small farmers, farmers, cattle breeders) can be distinguished by their different land use patterns.

For modelling multitemporal series of *landscape objects* in large-scale land information systems, we need different methods. These methods should be able to elicit object change by automatic means. As the changes follow the temporal line, we refer to them as *evolutions* and we refer to objects that changed as *evolving objects*. Within this evolution research theme, we cite SILVA et al. (2005), who developed a precursor in generating automation methods to handle *landscape objects evolutions*. They proposed a method to classify *landscape objects* obtained from a remote sensing image dataset using data mining, digital image processing, and landscape ecology theory. Their results show that analyzing *landscape objects* on images from distinct dates is an effective alternative to identify agents and model land use change. However, they do not discuss how objects evolve in time. Extending this work, MOTA et al. (2009) proposed to build the evolution by developing a method that obtains the rules for object evolution using previous cases as well as knowledge elicited from a specialist, a technique called Case-Based Reasoning (CBR). These rules are used to describe how geospatial objects can evolve. However, they do not discuss how to model the rules and the objects. We have not come across works where an algebraic modelling of evolution was defined to allow the evolution of *landscape objects* in a simple and automatic way, as we propose in this work.

2.3 Geospatial evolution model

The aim of this thesis is to rebuild the evolution of a set of *landscape objects*, given multitemporal data obtained by ground surveys or from remote sensing images. Data from remote sensing images produce a set of snapshots at discrete times $t_1, t_2 \ldots t_n$. Each snapshot contains *landscape objects* with no history. Our method builds a historical dataset, finding out which objects evolve from those existing previously. To do this, we put the original objects (without history) in an input dataset, called OD *(Objects Dataset)*. We then create a second dataset to store the objects' histories, called EOD *(Evolving Objects Dataset)*. The result is that within the OD we have the set of individual original objects and within the EOD we have the histories composed by the individual objects and linked to the most recent state of the objects.

A summary creation of the EOD works as follows: at the initial time t_0, we retrieve the set of objects from the original OD and insert it into the EOD. For each timestamp from t_1 to t_n, the *geospatial evolution model* combines the objects from the OD in time t_i with objects from the EOD in t_{i-1} to give the objects in the EOD in time t_i. This sequence of steps is illustrated in Figure 2.1.

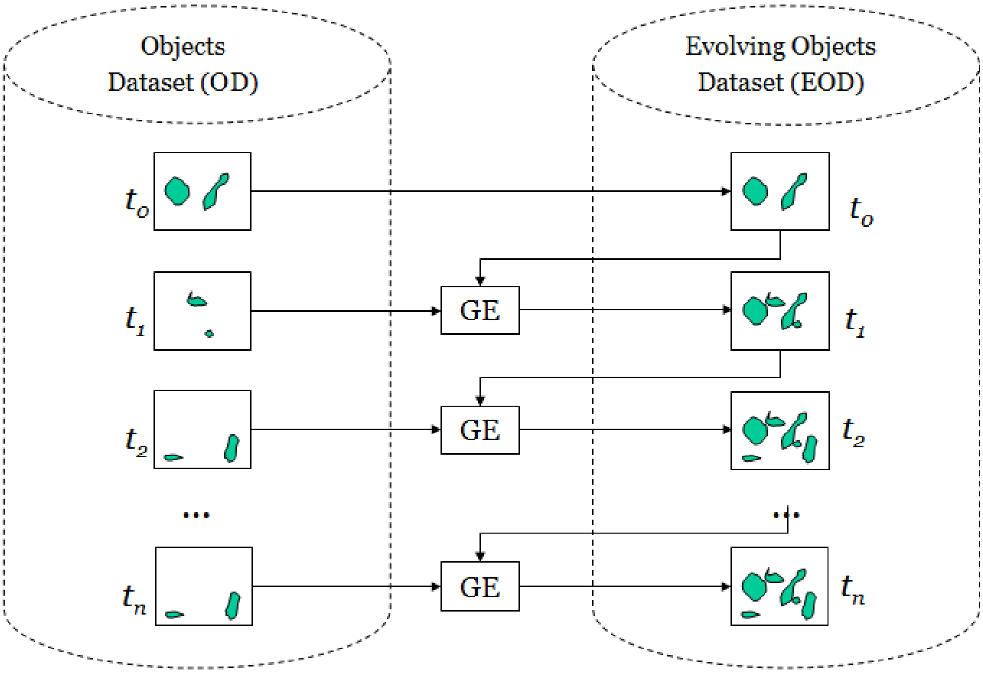


Figure 2.1 - The Geospatial Evolution (GE) model.

Each different application needs to define a suitable set of rules for controlling evolution. We call these 'merge rules' and 'split rules'. A 'merge rule' defines when two *geospatial objects* should be joined, while a 'split rule' determines when a *geospatial object* should be removed from other object. Given these rules, which are pre-conditions to perform 'merge' and 'split' operations, the *geospatial evolution model* works as follows:

1. At time t_o, all objects from OD are assigned an identity in EOD.

2. For all timestamps from t_1 to t_n, perform steps 3 and 4.

3. At time t_i take all new objects from OD (those created between t_{i-1} and t_i). Compare these to all existing objects in EOD (data from time t_{i-1}). For each new object in OD, apply the 'merge' operation:

   a) If there is an allowed pre-condition 'merge_rule' to merge it with an existing object from EOD, join them and create a new object in EOD. This new object in EOD has a new identity, valid from time t_i onwards. The dataset records the identity of its parents.

b) If there is no rule that allows it to merge with any object in EOD, assign it new identity in EOD, valid from time t_i onwards.

4. After all objects from OD have been processed, restart the comparison of objects in OD and EOD to apply the 'split' operation:

    a) If there is an allowed pre-condition 'split_rule' saying that an object *x* should be split from an object *y* in EOD, split it and create a new object in EOD. The new object is the result of the *(y - x)* spatial difference. The other object in EOD, the *y*, will be updated to indicate the evolution because it was the object that previously existed in the same place. Both have new identities, valid from time t_i onwards. Both have the same parents *x* and *y*.

Figure 2.2 shows an example of evolutions from t_0 to t_1. At time t_0, we have the initial set of objects [o_1, o_2, o_3, o_4]. Since this is the first step, they are the same in OD and EOD. At time t_1 we have four new objects in OD [o_5, o_6, o_7, o_8]. In this case, there is one 'merge rule': *if two objects touch, they are joined*. There is one 'split rule': *if an object is created and it overlaps a previous object, they are split*. The model identifies the objects that should be merged. According to the rules, o_1 and o_5 are joined to make o_9, and o_2 and o_6 are joined to make o_{10}. The model identifies the objects that should be split. According to the rules, o_8 split from o_4 resulting in the creation of a new object o_{11}, and the update of the genealogy of o_8 results in the new object o_{12}. This creates the new set of objects [o_3, o_7, o_9, o_{10}, o_{11}, o_{12}], valid for time t_1 in EOD.

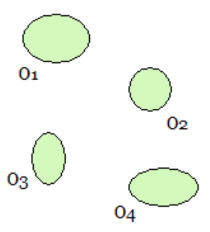
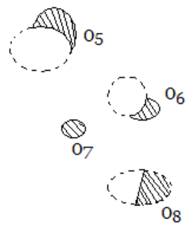
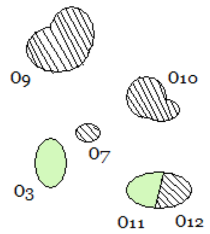
	OD (t_o) = EOD (t_o)	OD (t_1)	EOD (t_1)
Spatial Configuration			
Objects	[o_1,o_2,o_3,o_4]	[o_5,o_6,o_7,o_8]	[o_3,,o_7, o_9, o_{10}, o_{11}, o_{12}]
Genealogy			O_9 = merge (o_1,o_5) O_{10} = merge (o_2,o_6) O_{11} = split (o_4,o_8) O_{12} = split (o_4,o_8)

Figure 2.2 - Example of objects' evolutions.

To store the semantic evolution, we build each object with history in the EOD as a genealogy tree. At the lowermost level of the tree, we have the ancestor objects. As these 'merge' and 'split' with others, the tree grows upwards. As an example, take the objects in Figure 2.3 (a). In the first timestamp (t_o), there is one object (o_1). Then at time t_1, a new object appears (o_5). This is joined with object o_1 to create object o_9.

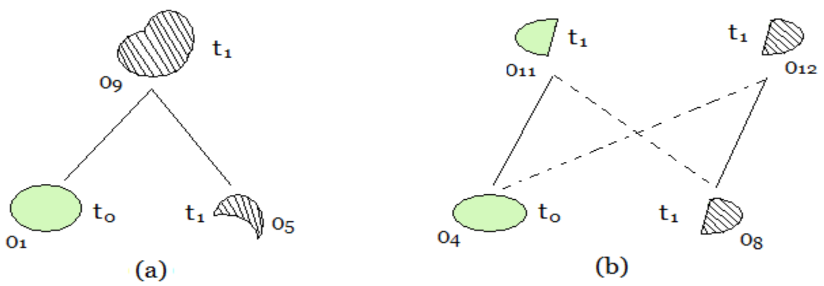


Figure 2.3 - Generation of genealogy trees with objects from two different times: (a) Merge operation tree; (b) Split operation tree.

As a second example, take the objects in Figure 2.3 (b). In the first timestamp, (t_o), there is one object (o_4). Then at time t_1, a new object (o_8) appears, sharing the same place as o_4. Then it is necessary to split object (o_8) from object (o_4)

because there is a distinct object in the same place. In this case, a new object (o_{11}), resulting from the difference between o_4 and o_8, will be created with o_4 as an ancestor and o_8 as a pointer. The object o_8 will be updated (o_{12}) with o_4 as a pointer and o_8 as the ancestor. The genealogy tree contains all of the object's states. Thus, we can recover the full history of the evolution or the snapshot of existing objects at any timestamp.

In the above example, all objects have the same type. This approach can be extended to more complex situations, with objects of different types. In this case, the result of the merge and split operations depends on the types of the input objects. Consider Figure 2.4, taken from (MOTA et al., 2009), where some prototypical *landscape objects* are portrayed. In this case, there are three types of *landscape objects*: *LargeGeometric* (LG), *Linear* (LIN) and *Small Geometric* (SG). For this case, we consider the following rules:

R1. A *SmallGeometric* object that touches a *LargeGeometric* object will merge resulting in a larger *LargeGeometric* object.

R2. Two adjacent *Linear* objects will not merge.

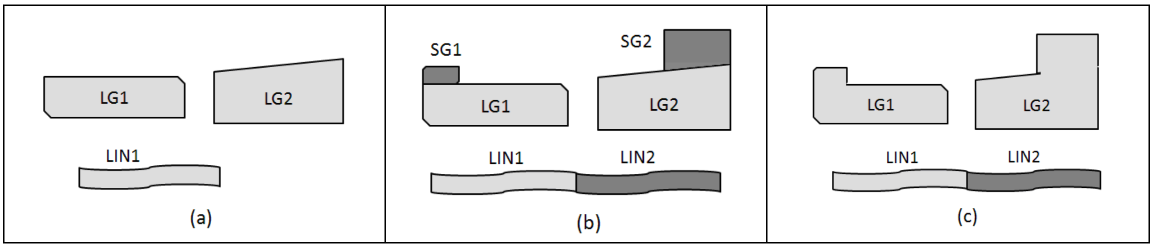


Figure 2.4 - Evolution of prototypical *landscape objects*: (a) Time t_1; (b) Time t_2 before application of rules; (c) Time t_2 after application of rules.
Source: MOTA et al. (2009).

Figure 2.4 (a) illustrates three objects at time t_1. At time t_2, three new objects appear as shown in Figure 2.4 (b). After applying the rules, we have four objects: LG1, LG2, LIN1 and LIN2, shown in Figure 2.4 (c). LG1 and LG2 followed rule R1 and were merged with SG1 and SG2, respectively; LIN1 and LIN2 followed rule R2 and did not merge.

The need to assign a new identity to the *evolving object* (such as o_9 in Figure 2.3 (a)) is necessary to handle cases when a new object drives the evolution of two existing objects (Figure 2.5), or when two new objects drive the evolution of another one (Figure 2.6). By giving a new identity to the resulting object, we can record its evolution using a genealogy tree.

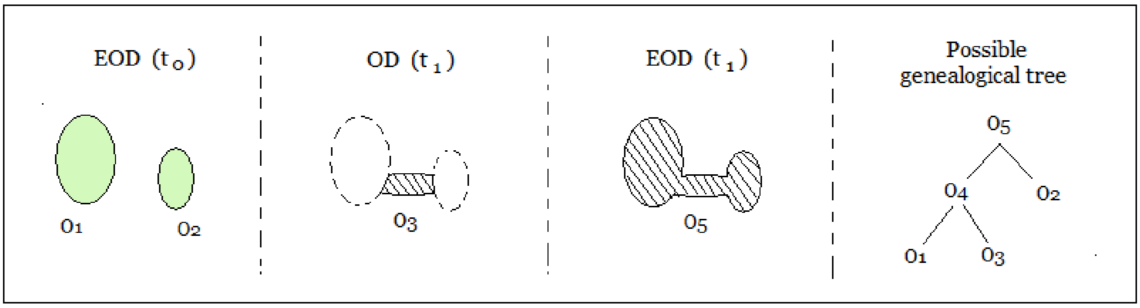


Figure 2.5 - A new object drives the evolution of two existing objects.

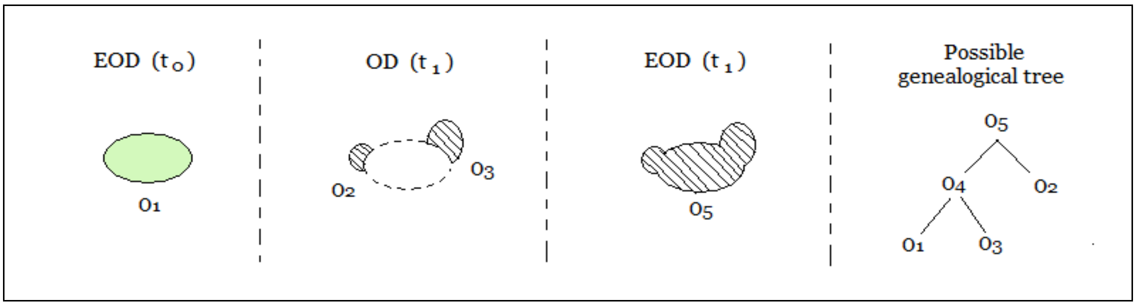


Figure 2.6 - Two new objects drive the evolution of one existing object.

As a practical consideration, evolution rules that depend on topological considerations (such as objects touching each other) may be affected by the geometric matching between two snapshots. In general, the user needs to perform suitable pre-processing operations to ensure that there is good correspondence between data from subsequent time steps. This pre-processing avoids incorrect rules resulting in an analysis generated by distinct computational distances of tolerance values. Since this is a more operational matter, we will not discuss it further.

2.4 Geospatial algebra

In this section, we propose an algebraic model for handling object evolution. We agree with GÜTING et al. (2003) and MEDAK (2001) who consider that algebraic data types provide a clean foundation for representing spatiotemporal objects. FRANK (1999) praises the idea of using algebras as a way to build complex systems from simple and modular components. For *moving objects*, GÜTING et al. (2003) propose an algebra composed of a set of spatiotemporal data types, axioms and their operations. Their algebra provides a foundation for handling *moving objects*. It was implemented using SECONDO (GÜTING et al., 2004), an extensible and modular DBMS environment. Although the above works provide relevant foundations, these algebras do not fit to solve our *geospatial evolution* problem.

In the following, we present an overview of our model. For clarity, many details are left out. We follow the usual conventions for abstract data types:

1. Type definitions include a constructor, which builds an instance from other pre-existing types, and an externally viewable set of functions.

2. Data types, functions, instances and axioms use the monospaced font. Type names use capitals for actual types (e.g., Polygon, String) and lowercase for parameterized types (e.g., type, a, b).

3. A function signature is written as function :: A → B → C. This is a function with two input types (A and B) and an output type (C).

4. We use '[]' for list of objects, and '()' for tuples. We use '=' for attribution, '|' for negation, '==' and '/=' for comparison and '+' and '-' for spatial sum and difference, respectively.

5. When there is more than one synonym, they are separated by '|'. Comments are embedded within {-- --}.

We consider some pre-existing types. These include the basic data types String, Double, Int and Bool and types Point, Line, and Polygon for spatial data. The types Time and Interval are available for temporal data. We use the following definitions as building blocks:

```
type ID = String                              {-- object identifier --}
type AttrName = String                        {-- attribute name --}
type AttrValue = String | Double | Int        {-- attribute value --}
type Attribute = (AttrName, AttrValue)        {-- name-value pair --}
type Spatial = Point | Line | Polygon         {-- spatial types --}
type Interval = (Integer, Integer)            {-- time interval --}
type Time = Interval
```

We use two data types to model the *geospatial evolution*: Object and ObjHist. The Object type is associated with individual entities that have a time stamp, spatial and non-spatial attributes:

```
{-- definition of Object data type --}
data Object type = Object ( ID, type, Time, [Spatial], [Attribute])
class Objects t where
createObj:: type → ID → Time → [Spatial] → [Attribute] → Object type
```

Following our idea of using objects of different types, Object is a type parameterized on a generic type, identified here by type. The operation createObj creates an instance of the type Object type. Thus, we use a single operation to create objects of different types. For clarity, we define type_ to use on the definition of operations and indicate a domain of types specific to each application. For example:

```
data type_ = type1 | type2 | type3
```

As defined above, an Object is a static entity, frozen in a particular time. We need another type to account for a multitemporal object. This role is played by the ObjHist type. An instance of ObjHist is created for each new independent object. It then stores the history of the object and all of its descendants. All spatial operations are performed using object histories as inputs and outputs. We define the type ObjHist as a tree of objects, where each level of the tree represents a state of the object. When two objects are merged or when an object is split, a new level of the tree is created. Given a pre-existing polymorphic type Tree, we define the class ObjHist and its operations below.

```
{-- definition of ObjHist data type --}
data ObjHist type_ = Tree Object type_        {-- a tree of objects --}
{-- operations on ObjHist data type --}
class ObjHists obj where
createHist   :: Object type_ → ObjHist type_
getObj       :: ObjHist type_ → Object type_
merge        :: ObjHist type_ → ObjHist type_ → ObjHist type_
split        :: ObjHist type_ → ObjHist type_ → ObjHist type_
merge_rule   :: ObjHist type_ → ObjHist type_ → Bool
split_rule   :: ObjHist type_ → ObjHist type_ → Bool
```

Function createHist creates the initial genealogy tree of an Object. It puts an instance of an object in a new tree. Function getObj recovers the resulting object in the root of the tree, without genealogy. Operation merge is used when two objects can be merged: since both of them have histories, we merge two object histories. Operation split removes part of an object when it is possible. It creates a new object (and its history) and updates the old object based on the spatial difference between them. Since the result of the merge and split operations depends on the input types, we have to define specific functions for each different combination of types.

As we explained in Section 2.3, we need two additional operations (merge_rule and split_rule) to define the pre-conditions of the merge and split operations.

These rules are applied to each object pair, according to the procedure described above. The axioms followed by each operation are described in Table 2.1 .

Table 2.1 - Axioms of merge, merge_rule, split and split_rule operations.

Operation	Axioms
merge	merge (type1, type2) = type3 iff merge_rule (type1, type2) == true and spatial ObjHist type3 = (spatial ObjHist type1) + (spatial ObjHist type2)
merge_rule	merge_rule (type1, type2) = true iff (type1, type2) ∈ domain (type_) and (ObjHist type1) /= (ObjHist type2) and spatial_constraint (ObjHist type1, ObjHist type2) == true and attribute_constraint (ObjHist type1, ObjHist type2) == true
split	split (type1, type2) == type3 iff split_rule (type1, type2) == true and spatial ObjHist type3 = (spatial ObjHist type1) - (spatial ObjHist type2)
split_rule	split_rule (type1, type2) = true iff (type1, type2) ∈ domain (type_) and ObjHist type1 /= ObjHist type2 and (spatial ObjHist type1) !disjoint (spatial ObjHist type2) and spatial_constraint (ObjHist type1, ObjHist type2) == true and attribute_constraint (ObjHist type1, ObjHist type2) == true

For clarity, many details are left out. In this Table 2.1, spatial indicates the geometry of the object. The optional spatial and attribute validations are indicated by the generic functions spatial_constraint and attribute_constraint. They are used to test for example, attribute values, such as ‘greater area values’, and spatial relationships, such as ‘touches’ and ‘inside’.

Using the ObjHist data type, we can now define the types *Evolving Objects Dataset* (EOD) and *Objects Dataset* (OD) as lists of histories, with associated operations:

```
{-- definition of Objects Dataset data type --}
type EOD     = [ObjHist]
type OD      = [ObjHist]
evolve    :: EOD → OD → EOD
snapshot :: EOD →  Time →  [Object]
history   :: EOD → ID → [ObjHist]
```

Table 2.2 describes the axioms of evolve, snapshot and history operations.

Table 2.2 - Axioms of evolve, snapshot and history operations.

Operation	Axioms

evolve	evolve (OD, EOD) = EOD iff
	foreach (a,b) where ((a ∈ OD) and (b ∈ EOD))
	{
	merge (a,b)
	split (a,b)
	}

snapshot	snapshot (EOD, Time) = [Object] iff foreach (a ∈ EOD) { if (Time (getObject a) <= Time) then (getObject a) }
history	history (EOD, ID) = [ObjHist] iff foreach (a ∈ EOD) { if (getID(a) = ID) then (getHist a) }

Operation evolve takes two sets of objects and applies the merge and split operations to them, generating the new state of *evolving objects*. Operation snapshot recovers a list of all objects valid at a specific timestamp. Operation history recovers the history of an object.

The *geospatial evolution model* described in Section 2.3 is summarized in the following steps:

1. Express the constraints to govern the evolution by merge_rule and split_rule.

2. Take the objects from a time t_0 from the spatiotemporal dataset. These objects will be processed to create instances of Object and ObjHist using the operations createObj and createHist. Insert these objects in the OD and EOD.

3. Take the objects from time t_1 and insert them in the OD using createObj. Process them with the objects in t_{n-1} from the EOD with evolve and use createHist to insert them in the EOD of time t_1. Repeat this step until the last timestamp.

4. At any time, query the dataset using the operations: history and snapshot.

Up to this point, we have presented the model in a generic fashion. To be able to test it, the operations have to be expressed in a programming language. A known benefit of algebraic models is their ease of testing and validation (FRANK, KUHN, 1995), particularly if a suitable programming language is chosen. The closer the programming language is to the algebraic model, the easier it becomes to validate the model. Hence, the algebraic model described in this section was implemented using the Haskell functional language (JONES, 2003). However, our model is generic and can be implemented in most programming languages. Annex A presents details about the used computational environment. In the next section, we show how this model is used in practice, with a real case study.

2.5 Evolution of deforestation in the Brazilian Amazon: a case study

This section presents an example of applying the *geospatial algebra* to study the evolution of *geospatial objects* associated with deforestation patches in the Brazilian Amazon rainforest. We use data from the surveying work done by the Brazilian National Institute for Space Research (INPE). Using remote sensing images, the INPE provides data on deforestation and degradation of the Brazilian Amazon tropical forest: it indicates that more than 37,000,000 ha were cut from 1988 to 2008. Given the extent of devastation, it is necessary to develop methodologies to analyze and derive information about the deforestation process.

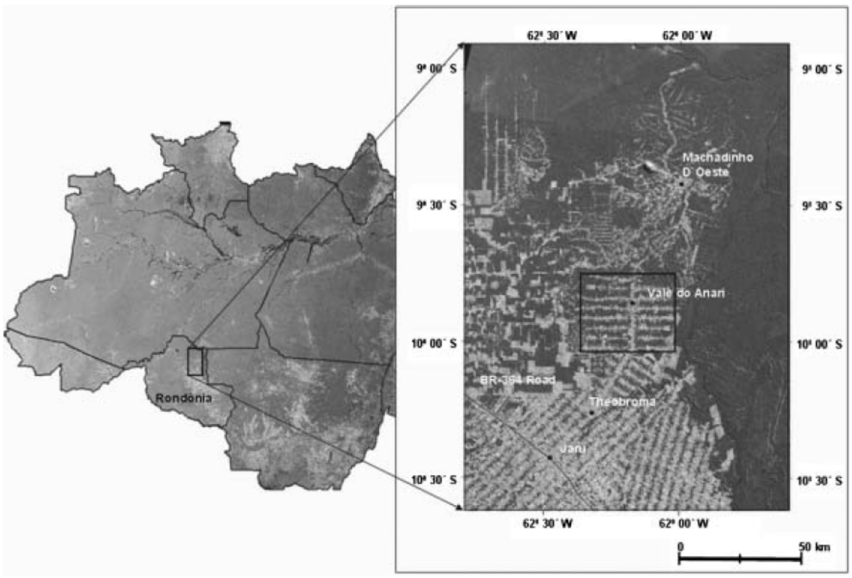


Figure 2.7 – Location of Vale do Anari study area.
Source: SILVA et al. (2008)

To study the evolution of deforestation our input data are a set of deforestation patches. These are our *landscape objects* and will generate *deforestation evolving objects* during the evolution process. By tracking and querying their histories, we can discover and quantify issues related to Brazilian Amazon deforestation patterns. We analyzed deforestation process in the Vale do Anari municipality, Rondônia State (Figure 2.7). This is a 400,000 ha region where occupation started with government-planned rural settlement.

We modelled six sets of *landscape objects*, from 1985 to 2000, with a three-year interval between each set. The data contain 4,070 objects and correspond to a total deforested area of 46,950 ha. Figure 2.8 shows the area deforested in each year and accumulated since the beginning of the analysis. Higher rates of deforestation are found between 1994 and 2000.

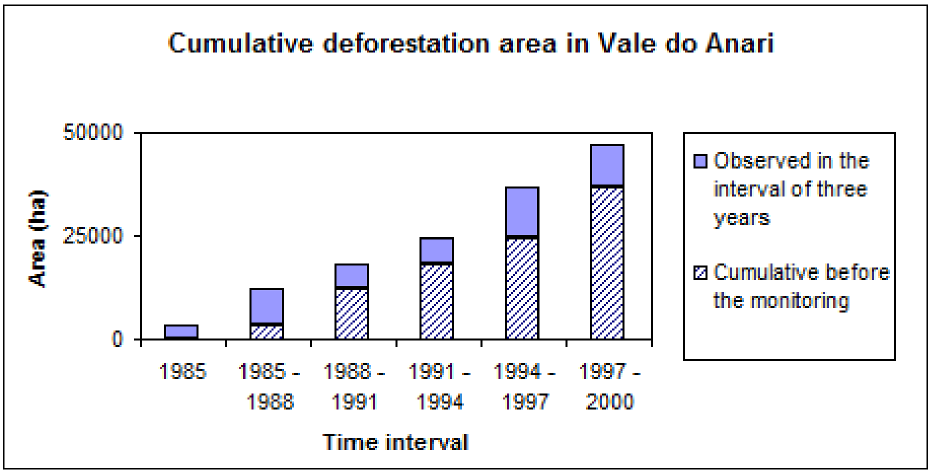


Figure 2.8 - Cumulative deforestation area in Vale do Anari from 1985 to 2000.

SILVA et al. (2008) and MOTA et al. (2009) classified each object of this dataset according to the land change agents acting in this region based on expert knowledge about the area. They defined *AlongRoad*, *Concentration* and *SmallLot landscape objects*. The rules applied to develop the evolution of these objects are based on the objects' sizes and spatial proximities. They were developed by MOTA et al. (2009) and were confirmed by experts on the deforestation domain. They are:

R1. Two adjacent *Concentrations* merge and the new object is a *Concentration*.

R2. Two adjacent *SmallLots* with areas smaller than 50 ha merge and the new object is a *SmallLot*.

R3. A *SmallLot* with an area smaller than 50 ha adjacent to a *Concentration* merge and the new object generated is a *Concentration*.

The constraints on the evolution model are implemented by:

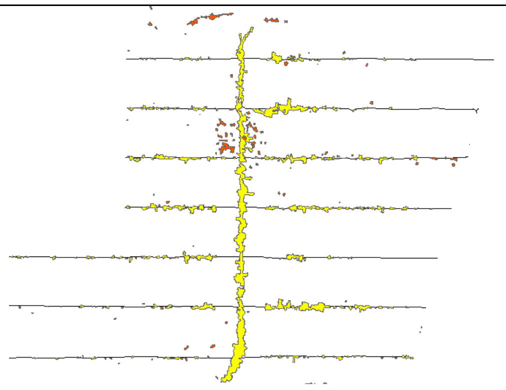
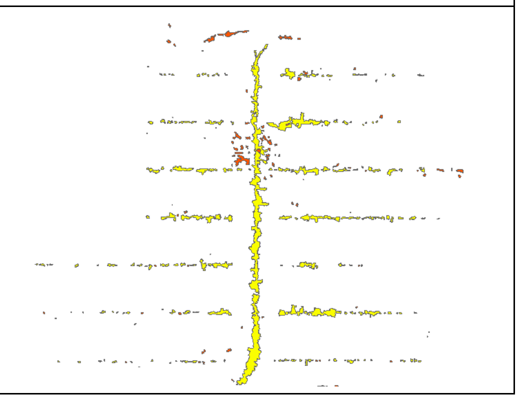
```
merge_rule :: ObjHist Concentration → ObjHist Concentration → true
merge_rule :: ObjHist SmallLot → ObjHist SmallLot →
    if ((area (ObjHist SmallLot1) < 50) and (area (ObjHist SmallLot2) < 50)) = true
    else false
merge_rule :: ObjHist SmallLot → ObjHist Concentration →
    if (area (ObjHist SmallLot) < 50) = true
    else false
```

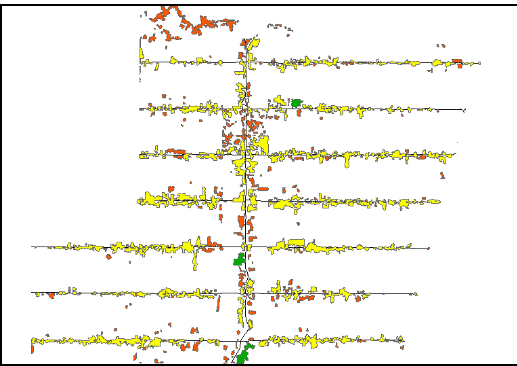
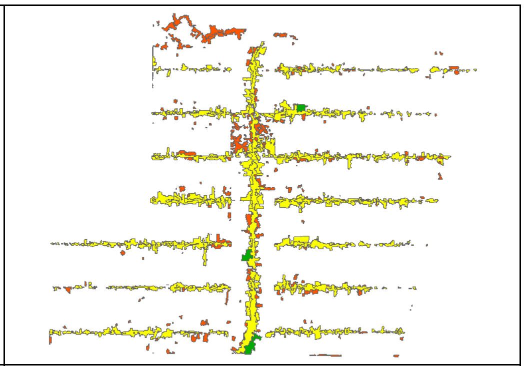
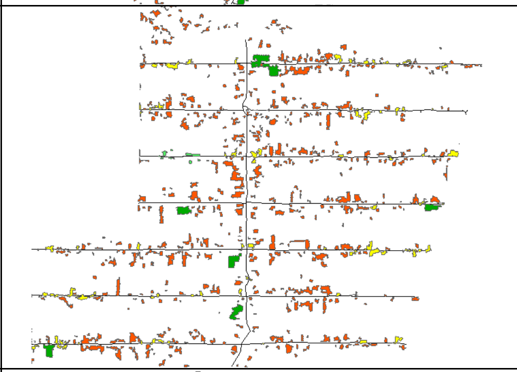
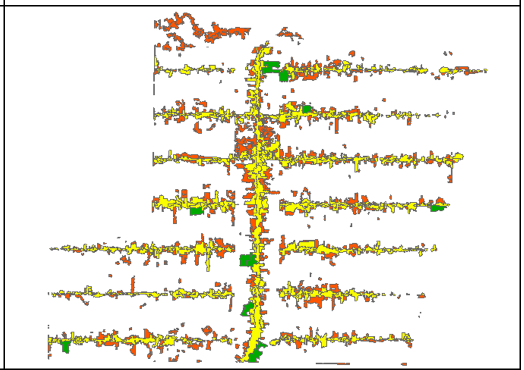
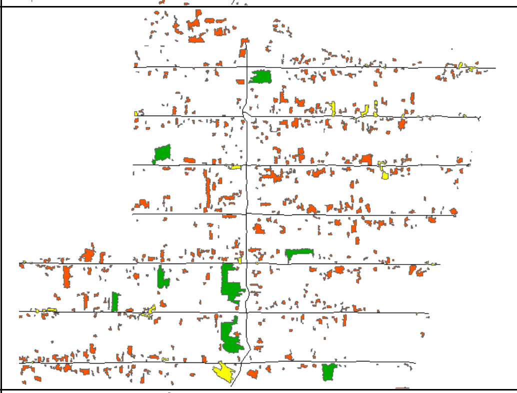
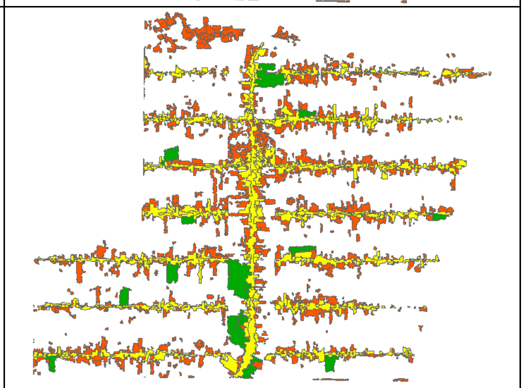
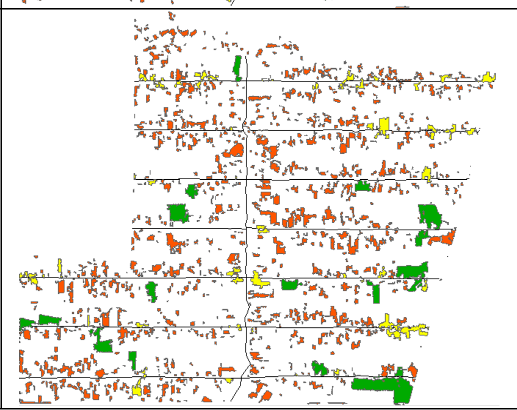
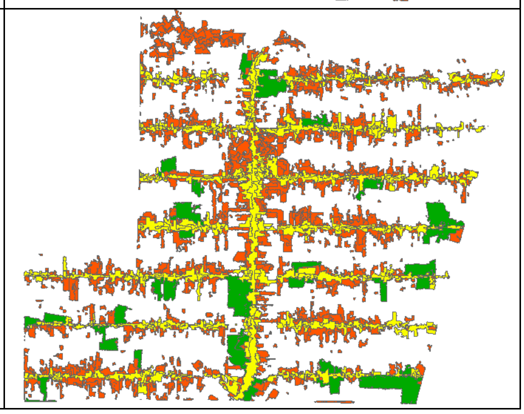
The combinations not defined in the rules will result in ‘false’ values. Other operations are defined by:

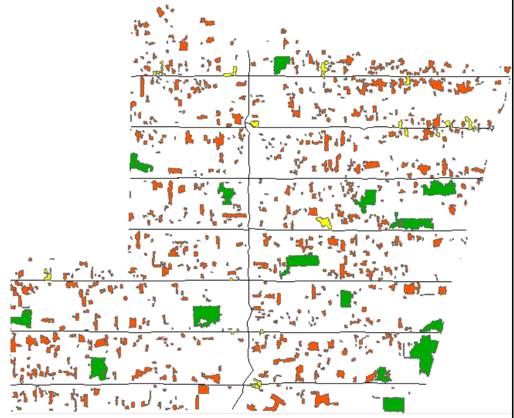
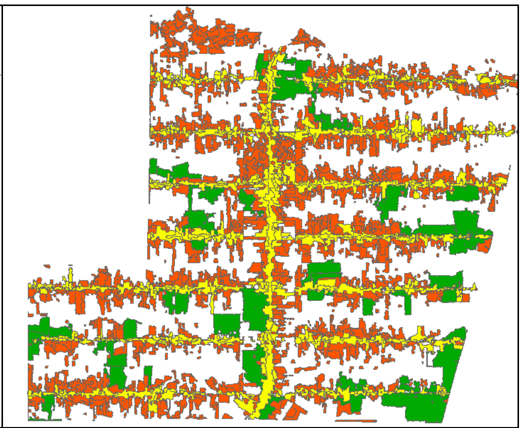
```
merge :: ObjHist Concentration → ObjHist Concentration → ObjHist Concentration
merge :: ObjHist SmallLot → ObjHist Concentration → ObjHist Concentration
merge :: ObjHist SmallLot → ObjHist SmallLot → ObjHist SmallLot
```

We model the change to discover when objects of one type become another type, which we call *evolution*. In this case, this represents discovering when and where the deforestation processes change. Table 2.3 illustrates the state of the *Objects Dataset* (OD) and the *Evolving Objects Dataset* (EOD) generated in each time interval, given a general view of the complete model.

Table 2.3 - OD and EOD datasets of Vale do Anari

Time	*Objects Dataset* (OD)	*Evolving Objects Dataset* (EOD)
1985		

1985 to 1988		
1988 to 1991		
1991 to 1994		
1994 to 1997		

1997 to 2000		
	AlongRoad Concentration SmallLot	

Our method results in 2,342 objects from the initial total of 4,070 objects. Figure 2.9 compares the number of objects per time interval considering the previous EOD plus the current OD in the first column and the number of existing objects in the current EOD after the evolution (second column). During the first and second intervals (up to 1985, and from 1985 to 1988), we detected only new objects, or objects that cannot be considered as an evolution from the previously existing ones. This can be explained by the predominance of *AlongRoad* objects that do not evolve, according to the adopted constraints.

Considering the same interval after this period, the resulting number of objects is always smaller than the number of objects without evolutions. In summary, the only transformation of type is *SmallLots* that evolve to *Concentrations* when they are smaller than 50 ha and are close enough to them. We verified that approximately 20% of the total number of *SmallLot* objects evolved to *Concentrations*. This result shows that many objects were corrected by the rules as evolutions of existing ones. In this case, the real aim is to detect how the land concentration process evolves, and to do this it is need the correct classification of *Concentration* and *SmallLot* objects. This allows to study if the land concentration on this study area is governed by deforestation patterns.

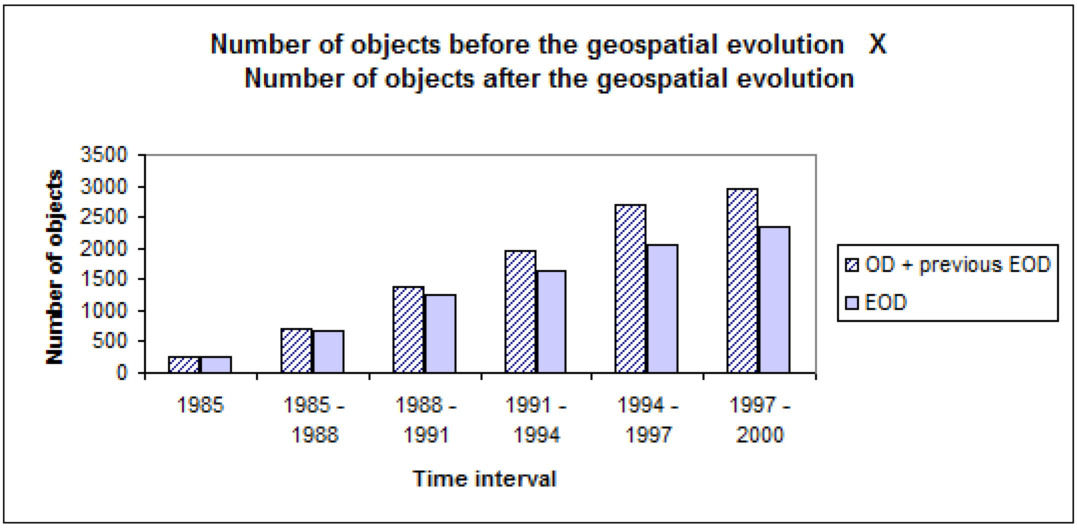


Figure 2.9 - Comparison of the number of objects before and after the *Geospatial evolution* process.

Another way to quantify the evolution of objects is to consider their extension. Table 2.4 compares the extension (sum of areas) of each type of object in the initial ODs from 1985 to 2000 and the resulting EOD on 2000 year. Considering the extent values, around 25% of the area occupied by *SmallLot* objects evolved to *Concentrations*. This value is similar to the 20% of the number of *SmallLot* objects that evolved to *Concentrations*.

Table 2.4 - Consolidated sum of areas (ha) of each type.

Landscape object	*AlongRoad*	*SmallLot*	*Concentration*
Complete ODs from 1985-2000	12,575	9,551	24,812
EOD in 2000 year	12,575	7,191	27,187

Another result is that without considering the evolution, it was classified 47 objects as *Concentrations*. However, in the end of the evolution process the total number of *Concentration* objects was 33. It then shows that do not considering the spatiotemporal evolution can indicates that objects initially considered distinct objects, are, in the reality, extensions of previously one with the same type. In this example, 14 objects, around 30% of the total number of *Concentrations* matches this case. This conclusion is supported by MOTA et al. (2009) using field data. It shows that considering specific rules to evolve objects more accurately defines their real deforestation type.

In addition to set analysis, by modelling the evolution process, we can retrieve information about a particular type or an individual history. Consider a *Concentration* object that started in 1991 as an example. The history of this object can be seen in Figure 2.10. The considered timestamp is the final year of the monitoring interval, and it is possible to identify when and which objects are part of this evolution.

```
( IDENTIFIER, ORIGINAL OBJECT, PARENTS ID, TIME, TYPE )

( 11,  Evolved,          [6,10],  2000, Concentration)
(  6,  ObjectId "598",   [],      2000, SmallLot)
( 10,  Evolved,          [5,9],   2000, Concentration)
(  5,  ObjectId "588",   [],      2000, SmallLot)
(  9,  Evolved,          [1,8],   1997, Concentration)
(  1,  ObjectId "425",   [],      1991, SmallLot)
(  8,  Evolved,          [4,7],   1997, Concentration)
(  4,  ObjectId "753",   [],      1997, Concentration)
(  7,  Evolved,          [2,3],   1994, SmallLot)
(  3,  ObjectId "459",   [],      1994, SmallLot)
(  2,  ObjectId "426",   [],      1991, SmallLot)
```

Figure 2.10 - The history of a *Concentration evolving object*.

Figure 2.11 graphically shows five fragments of snapshots indicating the objects part of the cited *Concentration*. Figure 2.11 (a), (b), (c) e (d) detach each new object that evolved to the resulting *Concentration evolving object* (Figure 2.11 (e)).

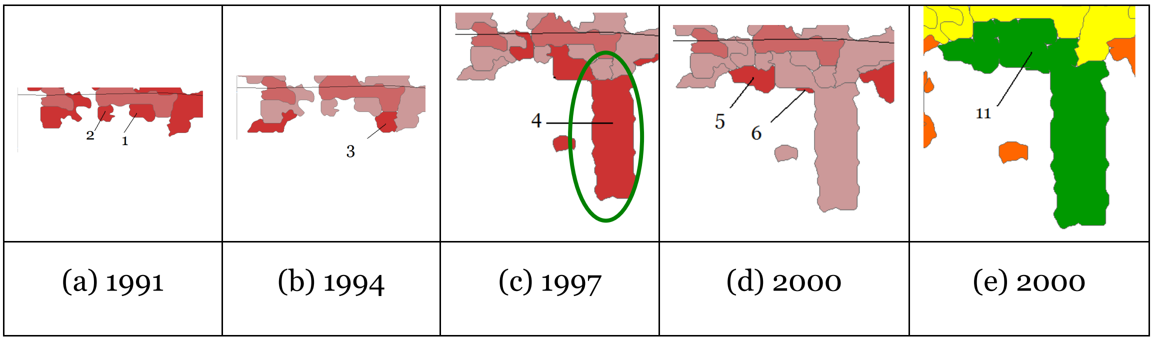


Figure 2.11 - Original objects composing a *Concentration evolving object*: (a) objects 1 and 2; (b) object 3; (c) object 4; (d) objects 5 and 6; (e) object 11: resulting object in the year 2000.

The *resulting evolving object* is composed of six original objects in the year 2000. The *Concentration* object was created in 1994. The five adjacent objects detected by the rules as being part of this *Concentration* object were initially classified as *SmallLots* in 1991, 1994 and 2000. This example shows that the application of the *geospatial evolution model* can improve the accuracy of deforestation detection.

2.6 Preliminary conclusions

The main contribution of this chapter is to propose a *geospatial algebra* to track the evolution history of a set of *evolving objects* as well as the individual history of each object in the set. The algebra describes how *geospatial objects* evolve by changing their boundaries and features. We refer to these objects as *evolving objects* and we consider that they need types to capture their semantics. Our *geospatial algebra* combines distinct types of *landscape objects*, describes and recovers the evolution of objects in a flexible way and considers constraints derived from knowledge about the application domain.

We applied the *geospatial algebra* in the domain of environmental change monitoring using remote sensing images to analyze a time series of deforestation *landscape objects* in the Brazilian Amazon. We identified *landscape objects* as *evolving objects* and were able of evolving them by applying the operations ‘merge’ and ‘split’, which are semantically adaptable to the application.

We can therefore verify the influence of *landscape objects* in close regions, discover patterns associated with the evolution histories and increase our ability to understand the land use changes detectable in remote sensing image datasets. We then reached the goal of identifying deforestation evolution histories, following their changes with time and helping to understand how they evolve. These methods can also be applied in other areas and scenarios.

In the next chapter, we show how this approach is used to model the evolution of deforestation in the other areas of the Brazilian Amazon forest.

3 EVOLUTION OF DEFORESTATION IN THE BRAZILIAN AMAZON

3.1 Introduction

We live in a changing world. Our urban and rural landscapes are being altered like never before. The major difference between Brazil and the rest of the world is that in Brazil we have an important and recognized precedent of making Earth-observation data available (WALDROP, 2008). The acceleration of change provides a strong motivation for research in GeoInformatics. It is important for preserving Earth's ecosystems to know these changes and to understand their impact on our Environment. Besides that, recent advances in satellites and monitoring technologies in the last decades have increased the quality and volume of spatial data. However, access to data is more critical every day. A few satellites can cover the entire globe, but there needs to be a system in place to ensure that their images are readily available to everyone who needs them (WALDROP, 2008).

A series of successful studies exist that try to find and define deforestation patterns using remote sensing in tropical forests. We cite, for example, ZIPPERER (1993) that identifies five deforestation patterns (internal, indentation, cropping, removal and fragmentation), in the state of Maryland, and he evaluates how the different patterns affect the forest through an analysis of land use change. CASEY and CAVIGLIA (2000) propose sustainable agriculture to try to minimize deforestation in tropical forests. They try to identify equalities in the form of land use for agriculture in the state of Campeche, Mexico and in the state of Rondônia, Brazil. LINKIE (2004) maps and analyzes forest loss in areas of the National Park of Kerinci in Sumatra with the goal of identifying areas that are more vulnerable to deforestation. Spatiotemporal analysis in a sequence of images identifies the areas and rates of deforestation for period. He tries to identify patterns in deforested areas near roads and rivers. CHOWDHURY (2006) tries to quantify and analyze the changes that occur in the landscape of the Calakmul Biosphere Reserves in

Mexico. The study evaluates biophysical variables, the socioeconomic context and institutional causes that influence the deforestation in this area. The idea is to identify a behavioural pattern that defines the deforestation in the region.

The Brazilian Amazon covers an area of more than 400 million hectares (ha), being a largely diverse region, in which sub-regions with different rates of change coexist, due the diversity of accessibility, as well as ecological, socioeconomic and political conditions (BECKER, 2001). The Brazilian Amazon contains diverse actors and processes that influence the spatial and temporal patterns of deforestation. In distinct socioeconomic, biophysical and political contexts, multiple actors and institutional arrangements contribute to shape the different trajectories of changes in the region, which translates into diverse rates and patterns of land change in space and time (DE ESPINDOLA et al., 2008).

3.2 The deforestation process in the Brazilian Amazon

The Brazilian National Institute for Space Research (INPE) uses satellite images to provide yearly assessments of the deforestation in the Brazilian Amazon. From the beginning of the Brazilian remote sensing program in 1974, the application of orbital images to detect and monitor the deforestation continues to be a major research theme. The experience acquired from the first research results encouraged a complete survey of deforestation in the Brazilian Amazon, which originally contained 400 million ha of tropical forests. This survey, done with 1:500,000 hardcopy B&W MSS-LandSat images of the year 1977, showed that by that time 2.5% of the original tropical forest cover was deforested. In 1988 INPE launched a deforestation monitoring program, PRODES (INPE, 2009), for the Brazilian Amazon region based on 1: 250,000 TM-Landsat color composites.

PRODES publishes an annual deforestation rate for the Brazilian Amazon and is probably a unique monitoring system worldwide in terms of its temporal and spatial coverage. Data from PRODES show that nearly 37,000,000 ha of forest were cut from 1988 to 2008. Data from remote sensing have become available since 1997 in a large spatial dataset of yearly land cover changes. The PRODES data, methodology, maps and statistics are available with unlimited access. Certainly, it is a successful effort to supply reliable and free data. Beside PRODES, INPE has other monitoring programs, such as DETER (INPE, 2009b) and DEGRAD (INPE, 2009c). Each program has a periodicity and monitors distinct characteristics associated with deforestation and degradation in the Brazilian Amazon. Figure 3.1 depicts some distinctions related to the monitoring time and the level of land use change of each one of these monitoring programs.

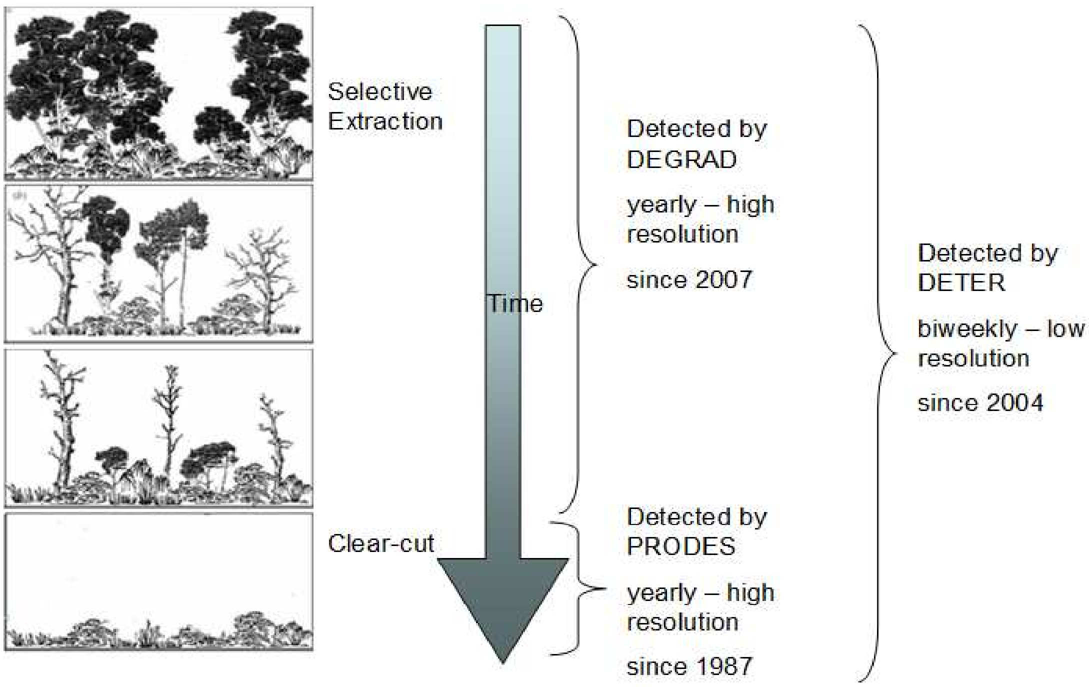


Figure 3.1 - Comparison between the PRODES, DETER and DEGRAD systems.
Source: adapted from INPE (2009d).

In the PRODES method new patches of deforestation are detected yearly, and are used to calculate the annual rate of deforestation. Although this is the main goal of the program, data from PRODES also represent a source of information that can be used to detect and quantify deforestation patterns. There is thus a lack between information produced with the available data and the richness of information that could be produced. Our tools and methods for representing and handling geospatial data should be capable of following these technological advances. To fulfil this demand, it is necessary to develop new methods to handle spatiotemporal data, to deal with change and to help in understanding the deforestation process. Landscape pattern analysis associated with data mining techniques enables the identification and description of different occupation patterns in remote sensing images (BECKER, 1997; MERTENS, LAMBIN, 1997; LAMBIN et al., 2003; ESCADA et al., 2005).

In the Brazilian Amazon, BECKER (1997) identified that the main processes of land cover change are linked to rural producers involved with agriculture, cattle ranching, mining and logging activities and each one employs different strategies for land use. He stated that human occupation forms in areas of expansion of agricultural borders are associated with different processes of land use and land cover changes. LAMBIN et al. (2003) show that different actors involved in land use change, for example, small farmers, farmers and cattle breeders, can be distinguished by their different land use patterns. These patterns evolve in time and new small settlements can emerge and large farms increase their agricultural area at the expense of the forest.

Sharing the same idea about the existence of characteristic deforestation patterns, ESCADA (2003) defined a typology of land use and land cover patterns for the central-north area of Rondônia state in the Brazilian Amazon. This typology forms a synthesis of the main processes associated with the different categories of rural properties established in the region and the different occupation forms. SILVA et al. (2008) expanded this idea treating the problem of detecting land use change patterns starting by establishing a land use and land cover typology. They then extracted semantic information from satellite images using data mining techniques. Following this, they developed a structural classifier to link expert knowledge to patterns detected in the images. They automatically detected and linked actors to characteristics land use formats. One limitation of this approach is that once the object is classified it no longer changes. The pattern is classified for a specific date and it does not consider multitemporal information to classify the objects. Thus, a piece of information that is relevant in the previous history and important to understand the evolution of land use change pattern is lost. MOTA et al. (2009) extended their approach using Case-Based Reasoning, a computational technique, to extract the rules from domain expert knowledge and to allow the use of multitemporal data to classify the objects and govern the evolution. They show the feasibility of modelling the evolution of objects and the importance of recovering their histories. However they do not discuss how to model the rules and the objects.

The objective of this chapter is to apply *geospatial algebra* to answer, in a simple and algebraic way, questions such as *"When can an object can be considered an evolution of another object in a previous time?"* or *"Which did objects at time t resulted from the evolution of other objects in the same time?"*. The next two sections define, respectively, an evolution typology of deforestation and an evolutionary set of rules. The rest of this chapter applies the typology and the rules on the *geospatial evolution model* to analyze two real case studies of deforestation in the Brazilian Amazon.

3.3 Typology for evolution of deforestation

We propose a typology for evolution of deforestation objects. The typology considers individual spatial properties (such as size or shape), spatial topology (such as proximity to other objects) and the state of objects (such as the existent occupation in the first monitoring). It is composed of four types of objects:

- **Linear:** This refers to data available on the first observation and data that contain linear geometric shapes. These indicate consolidated deforested areas, such as cities, and regions close to roads that indicate clearings that allow the traffic of people and commodities. In the whole deforestation process, they evolve in an independent way.

- **Alone:** This refers to objects that are not close enough to have any spatial relationship with other objects. An alternative when defining these objects is to consider the influence proximity to introduce some values of tolerance in the distance between objects.
- **Related:** This refers to objects that have some spatial intersection relationship with other objects.
- **Expanded:** This refers to a type that emerges after the process of evolution. For example, an *alone object* at a time t can evolve if a *related* object touches it at a later time. The *resulting evolving object* will have this *expanded* type.

Unlike the case when the occupation is driven by government settlement, like that presented in the previous chapter, occupation in the Brazilian Amazon also proceeds in waves: (a) taking hold of public land where the forest is still intact, creating unofficial ownership; (b) selling the land; and (c) consolidating appropriation and expanding the property. This process of occupation can be captured in two patterns detectable by considering the evolution of typed objects, the objects with defined type. We refer to these occupation patterns as *evolving patterns* and we define them as:

- **Isolated:** *Alone* objects dominate the land use change process, or the sum of the areas of *alone* objects has high values related to their medium values in the historical series in the region. This pattern can indicate the creation of new ownerships.
- **Expansion:** The land use change process suffers a strong influence from close objects and is marked by incrementing previous deforested areas. This pattern can indicate the consolidation and expansion of properties.

Figure 3.2 exemplifies three fragments of Objects Datasets and their *evolving patterns* indication based on the dominant object's type.

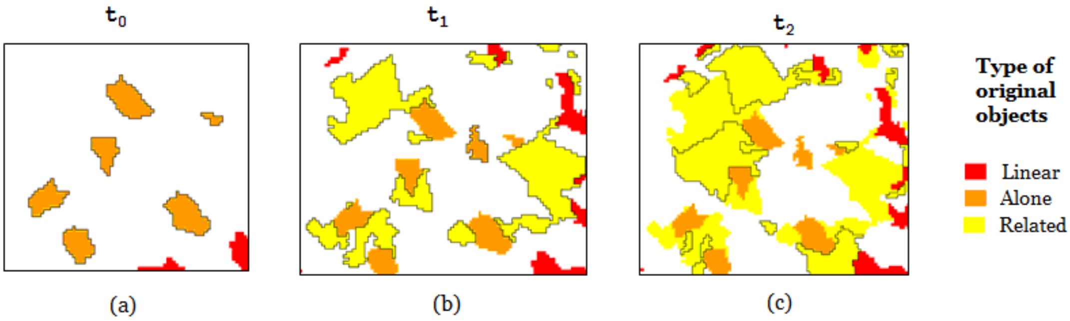


Figure 3.2 - Example of an *Isolated pattern* in time t_0 that evolved to *Expansion pattern* in times t_1 and t_2; (a) *Isolated pattern*, (b) and (c) *Expansion pattern*.

Besides the detection of *evolving patterns*, it is important to elicit their histories. To match this point, each object that merges or splits with another generates a *typed evolving object* that tells all of its history. Figure 3.3 presents the formation process of an *expanded evolving object* that started with an *alone* object (id 1) in the year 2001. The first evolution was a 'merge' with an object (id 2) in 2002. The second evolution was a 'merge' with an object (id 4) in 2003, resulting in the object identified by id 5.

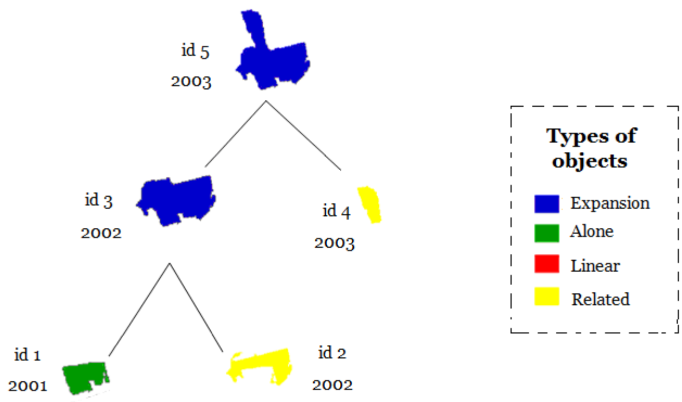


Figure 3.3 - Formation of an *evolving object*.

Each object has identifiers that can be queried. For example, the query "history 5" will recover information about the formation process of the object with identifier 5. The answer, in a sugar format is:

```
(id 5, 2003) = ( merge of (id 4, 2003) with (id 3, 2002))
(id 4, 2003) = original object
(id 3, 2002) = ( merge of (id 2, 2002) with (id 1, 2001))
(id 2, 2002) = original object
(id 1, 2001) = original object
```

A related study based on the histories of *evolving objects* is the evaluation of the contributions of *alone* objects to the whole process. Besides that, we expect to distinguish typical histories such as:

1. The *expanded evolving object* starts with an *alone* object and it continues to be expanded in later times. This can indicate the continuous expansions of specific regions.
2. The *expanded evolving object* is not formed by *alone* objects. The deforestation process contains only *related* objects. This can indicate clearing followed by a quick occupation process.
3. There are successions between *expansion* and *isolated* patterns in the historical land use change in the region.

Analyzing the evolution of deforestation, we expect to be able to point to, for example, regions where the occupation process is starting, is in advanced process of deforestation or has greater chances of expanding quickly.

3.4 Modelling the geospatial evolution

The *geospatial model* is developed by the *geospatial algebra*. Informally, the *geospatial algebra* is a set of functions to model the changes to objects. To build the model, we apply the *geospatial algebra* in the input the object's datasets to produces a new dataset of *evolving objects*. In this algebra, we follow domain application rules that govern the evolution. Based on INPE's expertise, we propose the following informal rules to govern the evolution of our deforestation objects:

1. When a *related*, *alone* or *expanded* object touches a *related* object, they evolve by 'merge' and the result is an *expanded* object.

2. When a *related* or *expanded* object touches an *alone* object, they evolve by 'merge' and the result is an *expanded* object.

3. When a *linear* object touches another *linear* object, they evolve by 'merge', and the result maintains the *linear* type.

4. When a *linear* object touches a *related* or an *expanded* object, they do not evolve. This allows separating *linear* objects from other *evolving objects*. Studying them allows, for example, us to characterize whether later deforestation is driven by roadside and consolidated clearing.

The implementation of these constraints, defined in Chapter 2, is done by:

```
merge_rule :: ObjHist related → ObjHist related → true
merge_rule :: ObjHist related → ObjHist alone → true
merge_rule :: ObjHist related → ObjHist expanded → true
merge_rule :: ObjHist alone → ObjHist expanded → true
merge_rule :: ObjHist linear → ObjHist linear → true
```

The implementation of the 'merge', also defined in Chapter 2, is done by:

```
merge_rule :: ObjHist related → ObjHist related → ObjHist expanded
merge_rule :: ObjHist related → ObjHist alone → ObjHist expanded
merge_rule :: ObjHist related → ObjHist expanded → ObjHist expanded
merge_rule :: ObjHist alone → ObjHist expanded → ObjHist expanded
merge_rule :: ObjHist linear → ObjHist linear → ObjHist linear
```

Figure 3.4 illustrates the geospatial modelling of objects in Figure 3.3 and the transformations.

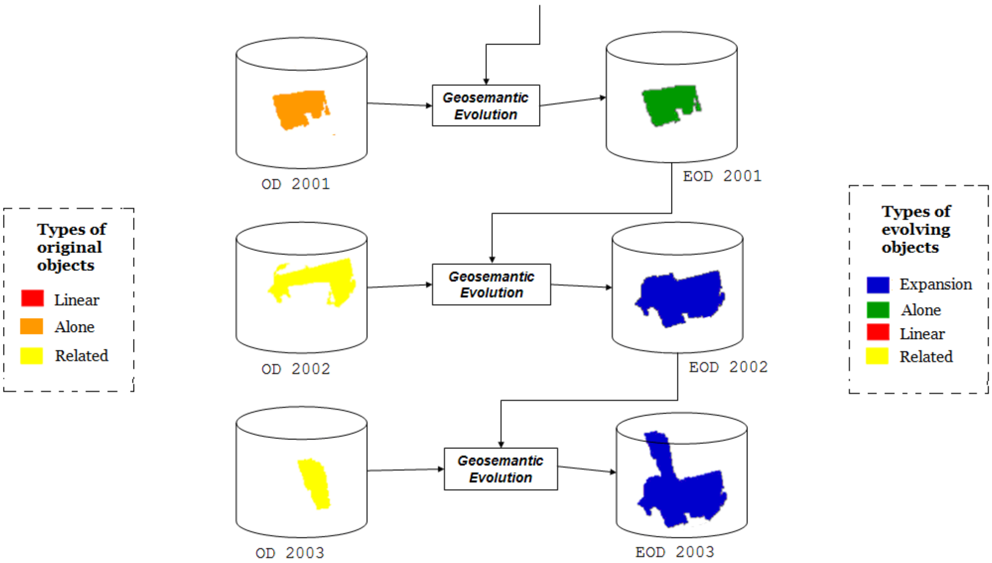


Figure 3.4 - The status of an *evolving object* in the OD and EOD datasets.

After the evolution, we can query the *evolving objects* to retrieve snapshots and the histories of objects at any time. These functions provide information about individual and sets of objects and allows us to extract patterns of deforestation.

3.5 Studied deforestation areas

The next sections shows the deforestation modelling of the Terra do Meio and Novo Progresso case studies, both located in Pará State (Figure 3.5). Annex B presents details about the analyzed data.

For each one, we analyzed multitemporal series composed of ten sets of data from 1997 to 2008. These regions were chosen because they are part of the set of the ten largest deforested regions in the latest years (INPE, 2009d). The deforestation process has been increasing the deforested area until 2005. Deforestation processes in these areas present large values, indicating intense processes of clearing and expansion and drawing the attention of the government and population.

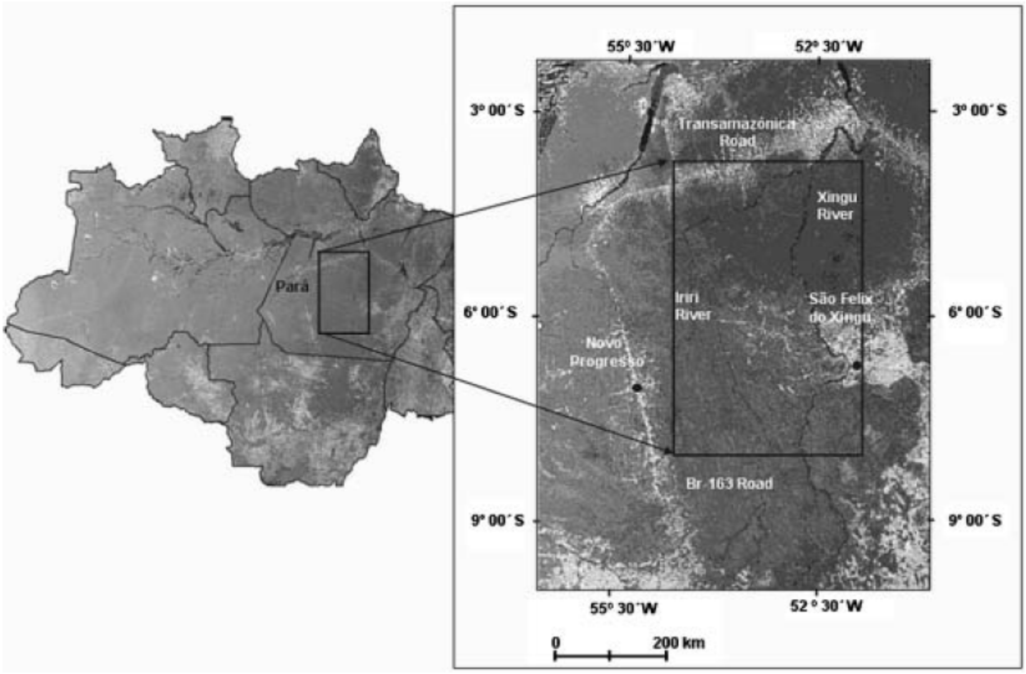


Figure 3.5 - Location of the Terra do Meio and Novo Progresso study areas. Source: SILVA et al., (2008).

In 2004, the PPCDAM, or Plan of Action for the Prevention and Control of Deforestation in the Brazilian Amazon, was developed and executed (CASA CIVIL, 2004). This is a government plan to control the deforestation process, and it is based on three main actions: inspection, land regularization of illegal regions and definition of alternatives to sustainable forest use. This plan was created in March 2004, and the results can be seen in the deforestation rate decrease since 2005. One important part of this plan was the creation of the DETER alert monitoring system by INPE. These alerts are sent to inspection departments that, with the information 'on the fly' about areas being deforested, can inspect and punish illegal deforestation and help to control and avoid new expansion. However, the decrease in the rate of deforestation may have more reasons (MARQUESINI et al., 2008): part of the decrease may have resulted from changes in currency and commodity markets that reduced the profitability of agricultural expansion.

The large volumes of deforested areas and the indications in the current year by DETER, which is outside the scope of this work, show that the deforestation volumes is increasing again, and illustrate the importance of studying these areas to understand and avoid new deforestation in similar areas.

3.6 Terra do Meio case study

Terra do Meio is a large area, around 15,000,000 ha, located in southeastern Pará State following the road PA-279. It includes areas from the Altamira and São Felix do Xingu municipalities. São Felix do Xingu is the municipality with the largest deforestation rates in recent years (INPE, 2009d). We are specifically interested in the area between the left margin of the Xingu River and its tributary, the Iriri River. As it is circled by Indian territories on its north, south and eastern borders, the occupation proceeded more slowly than in Novo Progresso.

The occupation process started in the beginning of the XXth century, but due to difficulties in accessing the region and the decline of rubber occupation, the region experienced economic stagnation with small groups living by subsistence agriculture. This situation prevailed until the Canopus Mining Company opened the so-called 'Canopus road' in the beginning of the 1980s to support cassiterite mining. Migrant families and mahogany loggers then used the road to invade the region (SILVA et al., 2008). The Land Institute of Pará State encouraged the occupation by giving out land parcels of 100 ha to colonists, up to 10 km from the Canopus road. In the early 1990s, some villages started to emerge along this road. Mahogany logging lasted until the end of the 1990s, when all supplies had been exploited. In the last 10 years, a land concentration process occurred with farmers and cattle ranchers entering in the area using the dense road network opened by loggers (AMARAL et al., 2006).

Following this deforestation process, farmers bought land parcels from the original settlers and created large areas for extensive cattle-raising. This activity needs large areas that can be isolated within the forest. The isolation is possible because small roads are not easily to detect; some large farms have small airports for personal transport and the cattle are commonly transported within the forest by 'comitivas', a kind of ropers that entourage cow for large distances, which is cheap and typical. ESCADA et al. (2005), AMARAL et al. (2006) and SILVA et al. (2008) present detailed analyses of the occupation process in the area. Our objective is to expand these works by supplying a simple multitemporal analysis based on the algebraic modelling proposed in Chapter 2.

This case study contains 10,736 original objects from 1997 to 2008. Figure 3.6 presents the incremental and cumulative values of deforested areas. Peaks of annual deforestation were reached in 2002 and 2004, a period related to the expansion of cattle ranching in the region. After this period, we see a decrease in the incremental values, until 2008, which shows the smallest value since 2000.

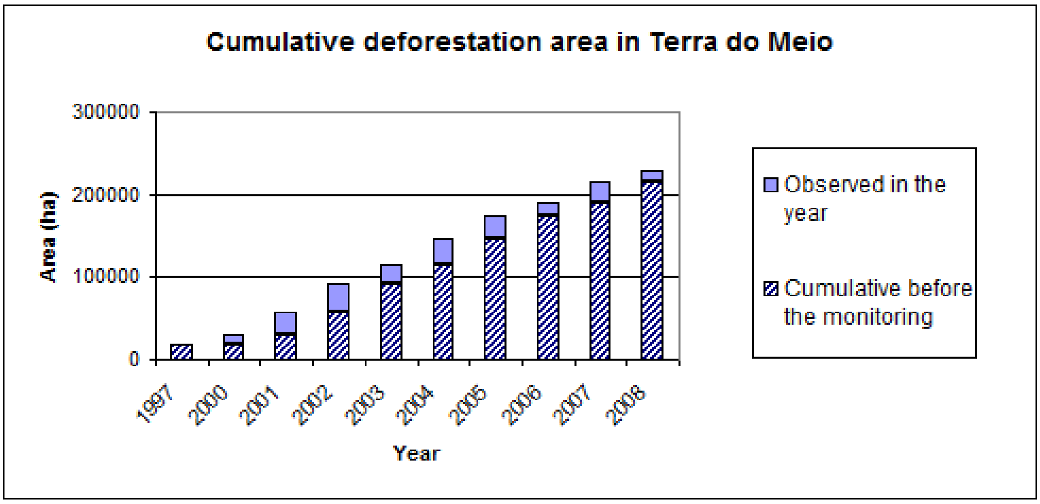


Figure 3.6 - Cumulative deforestation area in Terra do Meio.

We start the evolution process by classifying the data based on our typology. The original classified objects datasets from 1997 to 2008 are illustrated in Figure 3.7.

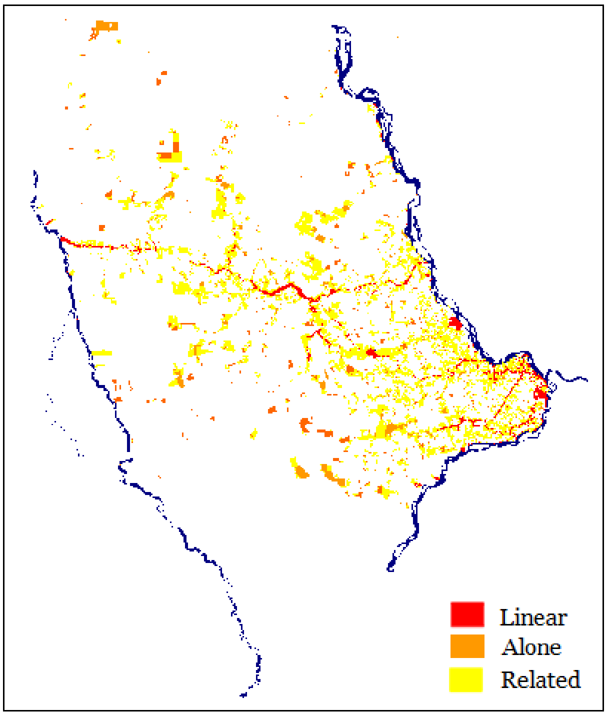


Figure 3.7 - Complete OD from 1997 to 2008 in Terra do Meio classified by type: *related*, *alone* and *linear*.

Figure 3.8 shows the sum of areas of each year separated by type. Analyzing the predominance of these types, done before the evolution process, we distinguish three patterns. An *isolated* pattern dominates the period from 1997 to 2001, when the area of *alone* objects dominated the deforestation process. This matches to many clearings on the area. From 2001 to 2004, we have a dominant *expansion* pattern, with the area being dominated by *related* area values and large areas of *alone* objects. In the interval from 2005 to 2008, the dominant pattern is the *expansion* pattern with a decrease in the search for new areas, while the amount of deforested areas is continually increasing by *expansion*.

After pointing to the *expansion* and *isolated* occupation patterns, we need to build the histories of evolutions. These will help us to answer queries about what happens after the creation of a deforestation *evolving object*, if it evolves, and, mainly, how it evolves.

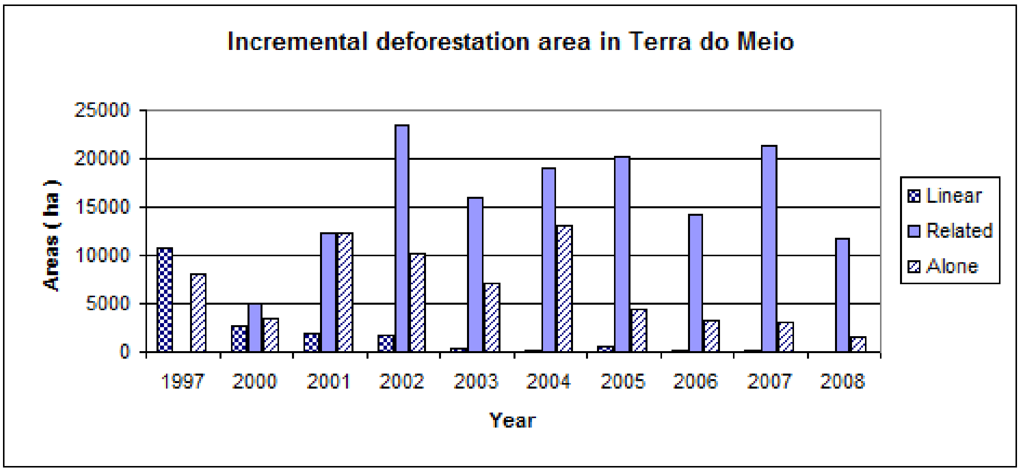


Figure 3.8 - Distribution of incremental deforestation area in the Terra do Meio before the evolution.

We then build the evolution of 10,736 original objects using the rules described in Section 3.4. Table 3.1 shows the number of original objects for each set of each timestamp (representing the *Objects Dataset* – OD) and the number of *evolving objects* after the evolution (representing the final *Evolving Objects Dataset* – EOD). The relevance of considering the evolution process to analyze the deforestation and to relate the objects is exemplified by the comparison between 10,736 original objects and 2,932 *evolving objects* after the final evolution. Therefore, around 70% of original objects have strong relations with their neighbourhoods. Each one of them has a distinct history that can be recovered from its genealogy, its changes in time and its final spatial configuration.

Table 3.1 - Number of objects in the OD and EOD datasets by year.

Year	1997	2000	2001	2002	2003	2004	2005	2006	2007	2008
OD	570	677	1307	992	1053	1360	1367	1049	1313	1048
EOD	570	1047	1702	1988	2222	2622	2766	2853	2814	2932

Table 3.2 presents numerical results of the final evolution in 2008 separated by type. The most relevant information is that although the number of *expanded* objects is not the larger value, the volume of deforested areas represents 83% of

the total deforested area. This shows the relevance of continuous expansions in the deforestation process.

Table 3.2 - Consolidated results in the final evolution in 2008 year separated by types: *expanded*, *alone*, *linear* or *related*.

Type	*Expanded*	*Alone*	*Linear*	*Related*
Number of Objects	931	1,049	646	306
Sum of Area (ha)	379,975	34,781	37,159	4,804
Percentage of Occupied Area	83%	8%	8%	1%

Figure 3.9 shows the percentage of area of each type on the OD and the percentage on the resulting EOD on the final evolution. The 29% of objects that appear as *alone* objects decreased to 8% in the final evolution process. This shows that 72% of *alone* objects expanded in the considered period of time. Therefore, it is important for government surveillance to take care of initial *alone* deforested areas to avoid their later expansion.

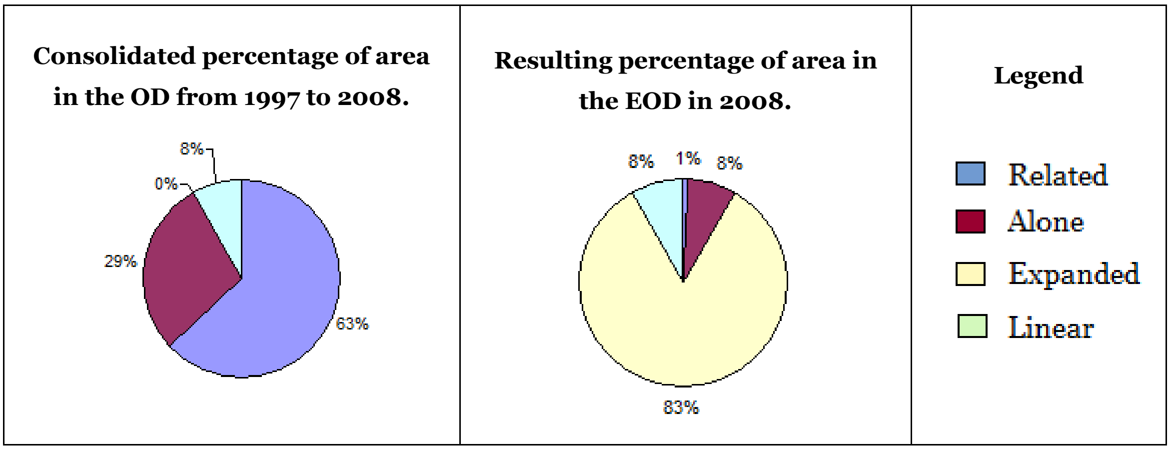


Figure 3.9 - Percentage of area occupied by each type: *related*, *alone*, *expansion* and *linear* before and after the geospatial evolution process.

Table 3.3 shows the area percentages occupied by *alone* objects for the ten ODs. This represents the behavior of the numbers and the extents of new clearings in this period. The creation of new deforested areas is smaller in the last years. This shows the current dominance of *expansion* patterns.

Table 3.3 - Area percentages occupied by *alone* objects in the OD.

Year	1997	2000	2001	2002	2003	2004	2005	2006	2007	2008
Area Percent	43%	30%	47%	29%	30%	41%	18%	19%	13%	11%

Table 3.4 graphically depicts the differences in the percentages of occupied area when considering the types of original objects (OD) and the types resulting by the evolution process (EOD) for 2008 year. It was expected that the sum of area values in the OD for a time t and the area values in the EOD for a previous timestamp generates the value of the EOD for a time t. For example, the EOD value in 2000 should be the value of the OD in 2000 year (677 ha) plus the value of the EOD in 1997 (570 ha). However, this does not happen because some objects (48%) that were *alone* in 1997 changed to *expansion*: the last line represents the area percentage of objects that experienced this evolution.

Table 3.4 - Total area of *alone* objects in the OD and EOD datasets separated by year.

Year	1997	2000	2001	2002	2003	2004	2005	2006	2007	2008
OD (ha)	8092	3373	12318	10242	7136	13059	4472	3250	3081	1499
EOD (ha)	8092	7577	14464	15102	16091	22188	19032	19179	16878	17391
Change percent	-	48%	72%	66%	41%	43%	34%	16%	28%	6%

In addition to the cumulative analysis presented above, after the evolution process, we can retrieve information related to individual histories of evolution. Figure 3.10 illustrates the final result of the evolution process classified by type and detaches three sets of objects. It contains 16 *evolving objects* with 10 *expanded* and 6 *alone evolving* objects.

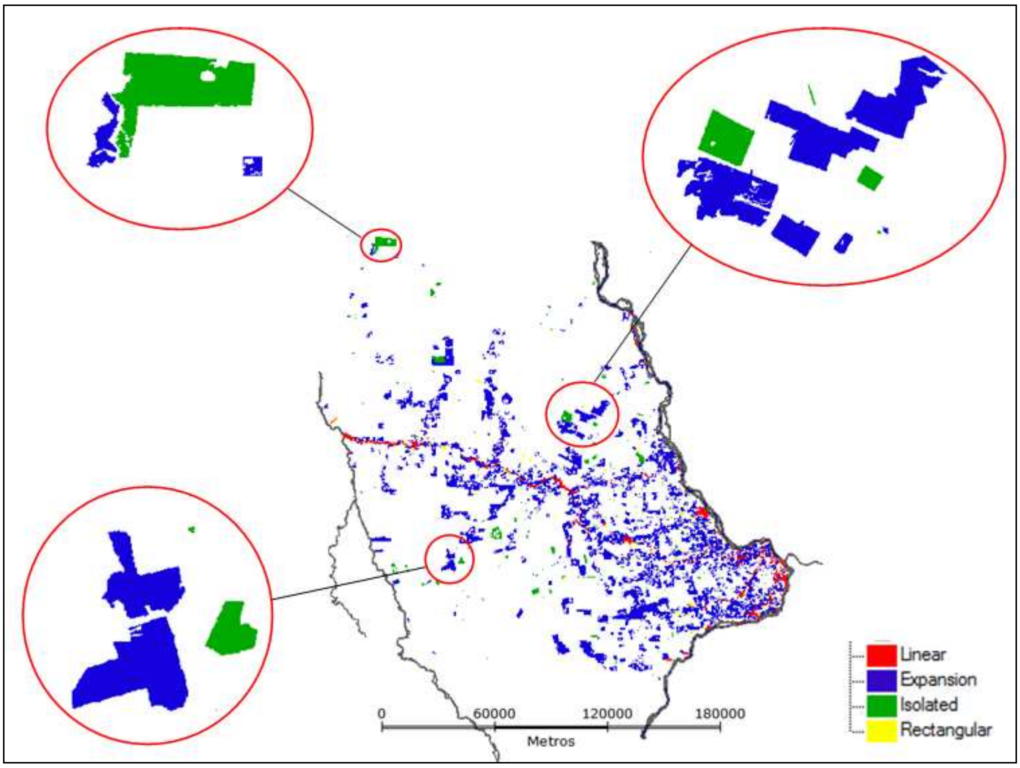


Figure 3.10 - Result of the evolutions in the Terra do Meio in 2008.

Some details are illustrated in Figure 3.11. The interval between lines (year axis) points out the first and last years of change in each *evolving object*. The first column points out the total number of original objects composing the *evolving* objects. The second column points out the number of *alone* objects (evolution length axis). The number of objects that compose the *evolving objects* varies depending on the location in the studied area. The number of *alone* objects belonging to *expanded evolving objects* is between one and three in the most cases. The exceptions are objects south of the Xingu River, which are closer to large consolidated areas: these tend to have larger numbers of objects in the total and also larger numbers of *alone* objects. These characteristics were confirmed for 915 other *expanded evolving objects* from the complete EOD in this case study.

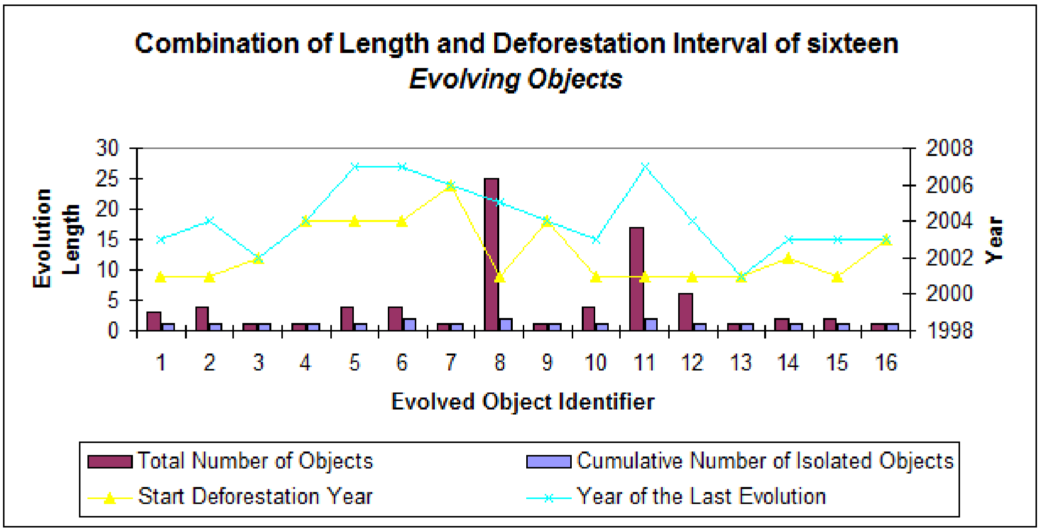


Figure 3.11 - Individual histories of sixteen *evolving objects* in 2008.

Figure 3.12 shows the objects that formed three of the sixteen *evolving objects* cited above.

Figure 3.12 - Formation process of three *evolving objects*.

Analyzing each *expanded evolving* object in the EOD, we observe that the *alone* original object is present in 783 of 915 existing *expanded evolving* objects and represents around 85% of the total deforested area. We also discovered that 95% of *alone* objects were created before the creation of *related* objects that form part of the same *evolving object*. This is an indication that objects really expand from *alone* objects and continue the expansion in later timestamps.

3.7 Novo Progresso case study

The second case study concerns to Novo Progresso municipality, Pará State. This region is crossed by an important road, called BR-163, which connects Cuiabá and Santarém, two strategic municipalities in the Brazilian Amazon. This region is close BR-230, the Transamazônica highway, an important hub of commodities and people that influenced the occupation process in the Brazilian Amazon during the 1970s and 1980s. The occupation process started with logging and the last decade presented an intense process of expansion and mechanization of soybean and rice agriculture.

Figure 3.13 presents the incremental and cumulative values of deforested areas. Peaks of annual deforestation values were reached in 2002 and 2004. After this period, we see that incremental values of area are very similar, including the most recent monitoring in year 2008.

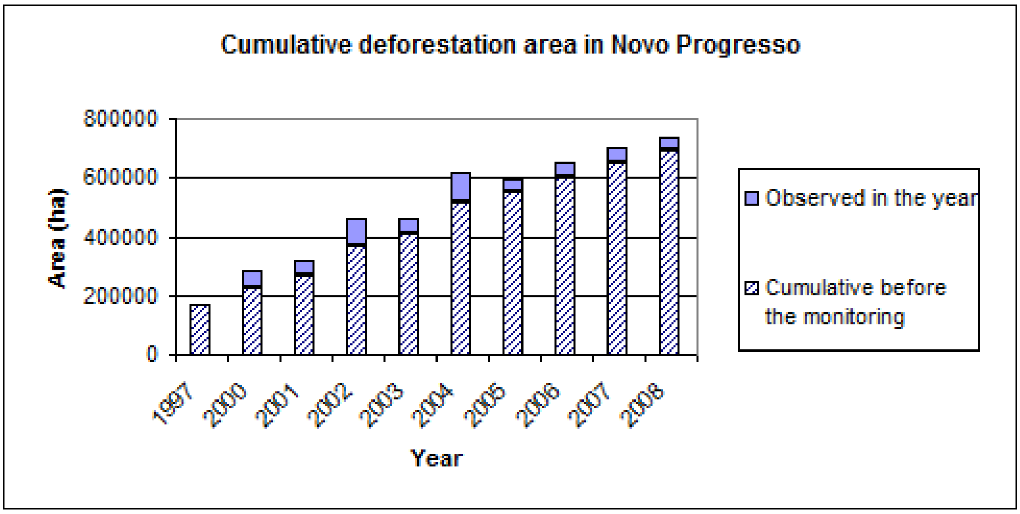


Figure 3.13 - Deforested area values in Novo Progresso.

To establish how deforestation process evolved in the area, first, we classified the data from 1997 to 2008 based on our typology. Figure 3.14 shows the sum of areas of each year, separated by types. We show incremental values starting in year 2000 because the occupied area by linear in 1997 is very large and it does not allow a good graphical visualization of distinct types on later time. It is because this region includes the well-established Novo Progresso municipality that already occupied a large area in the first monitoring in year 1997. All this area was then classified as *linear* and a relevant piece of posterior deforestation was incremental expansions of these areas. In spite of that, the evolution employs this 1997 data.

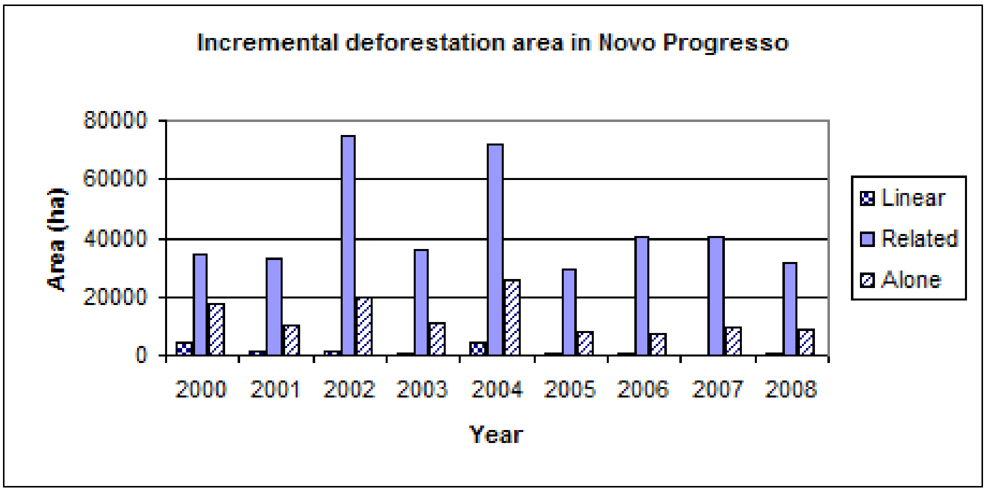


Figure 3.14 - Distribution of incremental deforestation area in Novo Progresso before the evolution.

In the original *Objects Dataset* (OD) (Figure 3.14) *related* objects have predominant role in the sum of deforested areas. It is related to expansions on agriculture areas from the large amount of pre-existent occupied areas. Despite of that, we can distinguish two intervals, if we consider the medium area values of *alone* objects in each year. With this in mind, *isolated* pattern have a relevant contribution to the period from 1997 to 2001, where their values are close *related* values. In the interval from 2002 to 2004 the area followed the deforestation behaviour that occurred in the Brazilian Amazon with a large amount of expansions and new clearings. In the interval from 2005 to 2008 we have the consolidation of the *Expansion* pattern. It shows the decrease in the search for clearings and their expansions. The large sum of deforested areas from 2002 to 2004 intensified the surveillance on the area with the implantation of PPCDAM on 2005. Besides that, due to the agriculture main economic activity in the area, the decrease is also related to minor commodities price on international market. To better understand the histories of change, we applied the rules to establish how deforestation evolved on this area. Figure 3.15 depicts the resulting *evolving objects dataset* (EOD) on 2008, after applying the *deforestation evolution model*.

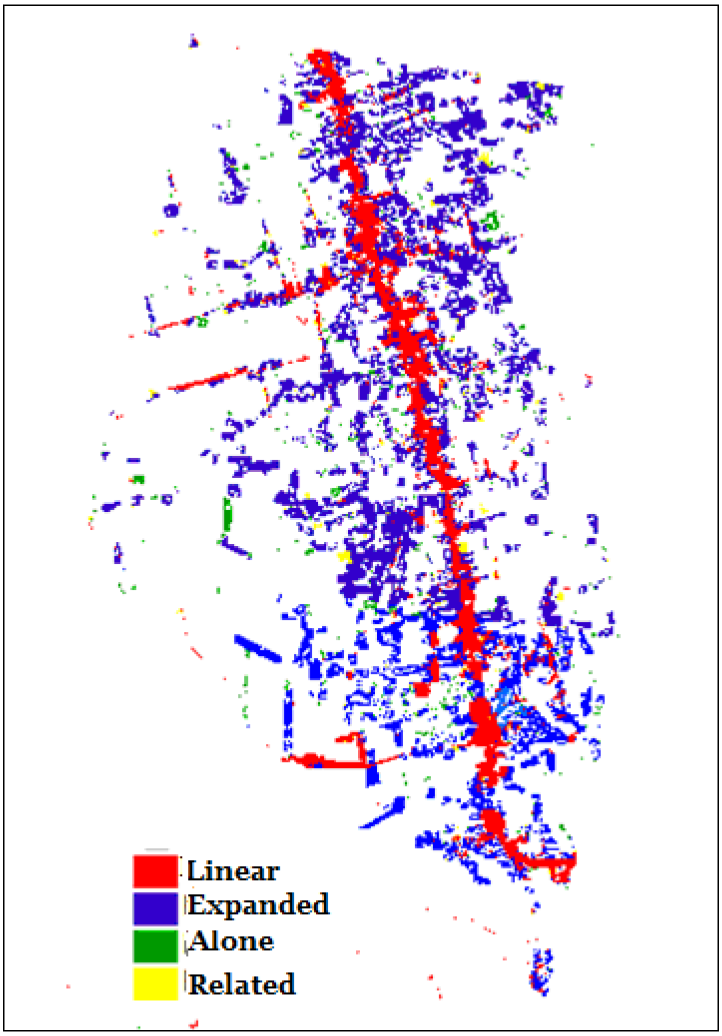


Figure 3.15 - Resulting EOD in year 2008 in Novo Progresso.

Cumulative results are shown in Figure 3.16.

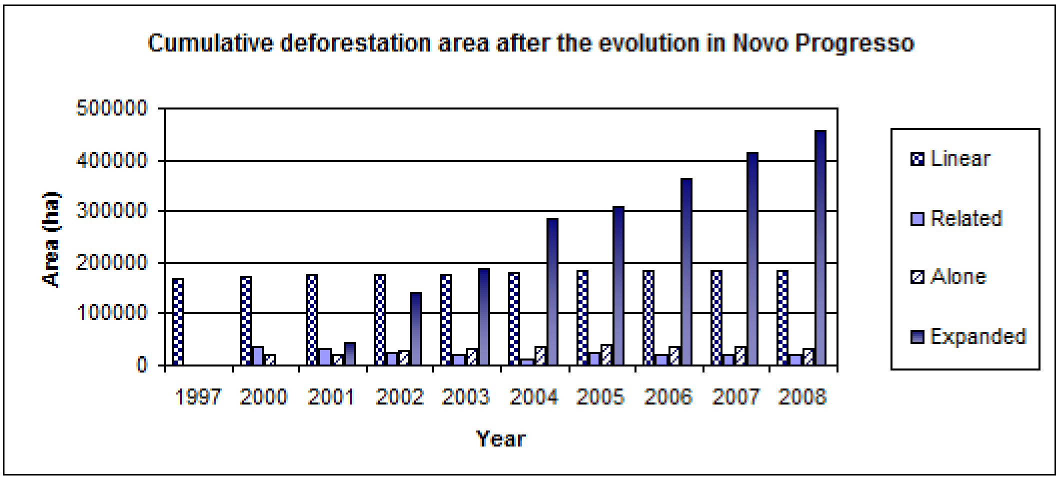


Figure 3.16 - Distribution of cumulative deforestation area values in Novo Progresso separated by year.

Table 3.5 presents summarized numeric results of deforested area in the final evolution in year 2008. The main rate of evolution (66%) is generated by *Expansions*. In this case, the consolidated area (*linear* objects) on the first monitoring was very large (around 26%) and, confirming results in Terra do Meio case study, the *expansions* are responsible by the majority of deforestation process.

Table 3.5 - Consolidated results in the final evolution in year 2008 year.

Patch type	*Expanded*	*Alone*	*Linear*	*Related*
Number of Objects	939	1,094	912	1,210
Sum of Area (ha)	456,780	32,762	183,153	19,810
Percentage of Occupied Area	66%	4%	26%	3%

Figure 3.17 shows a summary of the area difference between the *Original Datasets* (ODs) from 1997 to 2000 and the resulting evolution in year 2008. This result shows objects that started as new clearings and expanded to larger deforested areas. *Alone* objects initially occupied 22% of deforested areas and just 6% stayed as *alone* object after the evolution process. It shows that 72% of *alone* objects expanded from 1997 to 2008.

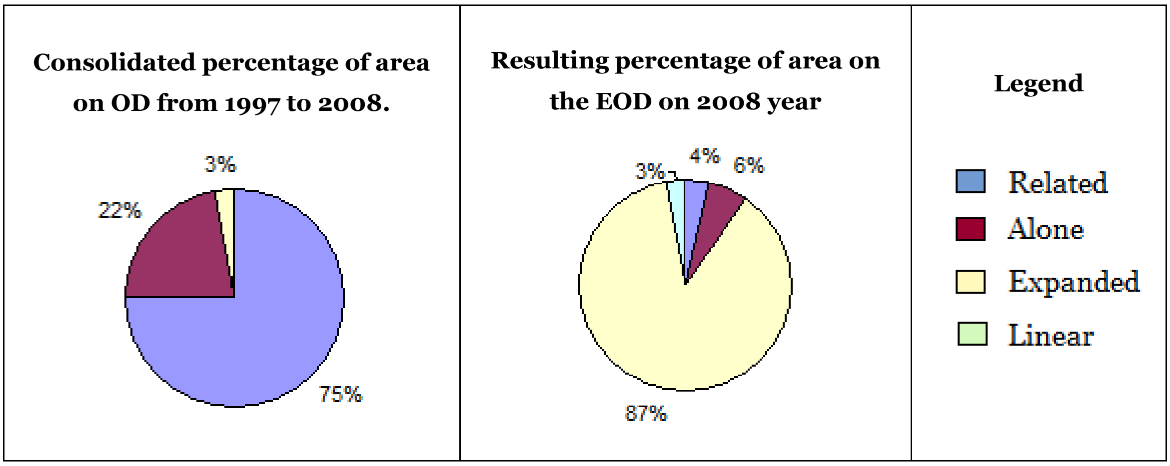


Figure 3.17 - Percentage of area occupied by each patch types: *related*, *alone*, *expansion* and *linear*, before and after the process of evolution.

Individual histories analysis shows that the number of objects composing the resulting *evolving objects* presents large variation depending on the localization. It is the same characteristic discovered on Terra do Meio. Analyzing each 939 *evolving object*, we saw that *alone* objects are presented in 581 *expansions*. It shows that *alone* objects make part of around 62% of total *expanded evolving objects*. This rate is smaller than the previous case study. One explanation is the large area of already deforested areas that influenced new deforestations, in addition to the agriculture main activity and geographic characteristics of the region. Another result is that around 90% of *alone objects* were created before the creation of *related* objects that form part of the same *evolving object*. As we saw in Terra do Meio analyses, this is an indication that *related* objects expand from *alone* objects and continue the expansion in later timestamps.

4 CONCLUSIONS

The main contribution of this thesis is the development of an algebra to model the evolution of spatiotemporal objects. The algebra comprises a set of operations, axioms and rules defined by the application. It also comprises operators to track the history of each individual object in the set.

We also developed a system to use the algebra in three case studies of land use and land cover in the Brazilian Amazon. We aimed to discover and quantify patterns of deforestation. The system enables users to assess patterns of change and their evolution in time, to analyze them, to adjust the rules according to field knowledge about the process and to make new inferences about the patterns of evolution.

We applied the *geospatial algebra* in the domain of environmental change monitoring using remote sensing images to analyze a time series of deforestation patches in the Brazilian Amazon. We identified land-change patches as *evolving objects* and were able to evolve them by applying the operations 'merge' and 'split', which are adaptable to the application. Using *geospatial algebra* we combine distinct types of land-change patches. This work is one more step towards more detailed studies on the development of general theories to discover patterns in the Brazilian Amazon.

We proposed a typology to describe the evolution of deforestation. The proposed typology, as we see in the results, allows us to study the evolution patterns from a broader level. Therefore, we can describe set or individual histories of evolution, verify their influence on nearby regions, discover patterns associated with the evolution histories and increase the ability to understand the land use changes that are detectable in remote-sensing image datasets. Following this vision, our method was able to extract patterns of expansion and distinguish between consolidated and newly cleared areas. We thus reached the aim of identifying general deforestation patterns, following their changes in time and helping to understand how they evolve. In addition, our methods can be applied in other areas without extensive changes or specific knowledge about the deforestation process in the area.

Advances can be done to improve the application of *geospatial algebra* in the environmental domain and to use it to better support economics and policy making in the Brazilian Amazon. The evolution of objects provides insight into the broader scope and complementary perspectives. This may help us to answer questions such as: *Are there general patterns that describe deforestation in the Brazilian Amazon? How can we discover them? There are specific patterns related to socio-economic activities, such as soybean agriculture or cattle raising?* The *geospatial algebra* we propose contributes to the efforts to answer these complex questions. We consider that similar applications of *geospatial algebra* could be applied to many other change situations and other parts of the world. We also believe that this methodology is flexible enough to be used for other types of applications, for example, urban cadastral data.

Future studies can be carried out on other areas and scenarios. The next steps may include the development of a complete system of *evolving objects* as well as new operations to advance our algebra. We also propose two major research directions that use *fuzzy logic* to study the influence of proximity on deforestation patterns and the development of evolving characteristics and metrics to find relations and discover new patterns.

BIBLIOGRAPHICAL REFERENCES

AMARAL, S.; MONTEIRO, A. M. V.; CAMARA, G.; ESCADA, M. I. S.; AGUIAR, A. P. D. Redes e conectividades na estruturação da frente de ocupação do Xingu/Iriri-Pará. **Geografia**, v. 31, n. 3, p. 655-675, December 2006.

ANDRIENKO, N.; ANDRIENKO, G.; PELEKIS, N.; SPACCAPIETRA, S. Basic concepts of movement data. In: SPRINGER (Ed.). **Mobility, Data Mining and Privacy**. Springer Berlin Heidelberg, v. 1, p. 15-38, 2008.

BATTY, M. Agents, cells, and cities: new representational models for simulating multiscale urban dynamics. **Environment and Planning A**, v. 37, n. 8, p. 1373-1394, 2005.

BECKER, B. K. **Amazônia**. São Paulo: Ática, 1997. 112 p.

BECKER, B. K. Revisão das políticas de ocupação da Amazônia: é possível identificar modelos para projetar cenários? **Parcerias Estratégicas**, v. 12, p. 135-159, 2001.

BITTENCOURT, O. O.; CÂMARA, G.; VINHAS, L.; MOTA, J. S. Rule-based evolution of typed spatiotemporal objects. In: SIMPÓSIO BRASILEIRO DE GEOINFORMÁTICA, 9. (GEOINFO), 2007, Campos do Jordão. **Anais**... São José dos Campos: INPE, 2007.

CÂMARA, G.; FREITAS, U. M.; CASANOVA, M. A. Fields and objects algebras for GIS operations. In: SIMPOSIO BRASIALEIRO DE GEOPROCESSAMENTO, 3., 1995, Sao Paulo. **Anais**..., p. 407-424, 1995. (INPE-6446-PRE/2485).

CAMARA, G.; VINHAS, L.; QUEIROZ, G. R.; FERREIRA, K. R.; MONTEIRO, A. M.; CARVALHO, M.; CASANOVA, M. TerraLib: an open-source GIS library for large-scale environmental and socio-economic applications. In: HALL, B.; LEAHY, M. (Ed.). **Open Source Approaches to Spatial Data Handling**. Berlin: Springer-Verlag, p. 247-270, 2008.

CASA CIVIL DA PRESIDÊNCIA DA REPÚBLICA DO BRASIL (CASA CIVIL). **PPCDAM - Plano de Ação para Prevenção e Controle do Desmatamento na Amazônia Legal**. Brasília, DF. 2004. Available at < http://www.planalto.gov.br/ casacivil/desmat.pdf>. Last acess: 29th July 2009.

CASEY, J. F.; CAVIGLIA, J. L. Deforestation and agroforestry adoption in tropical forests: can we generalize? Western Agricultural Economics Association Annual Meetings, Vancouver, British Columbia. **Proceedings...** 29th July 2000.

CHAKRAVARTY, M. **The Haskell 98 foreign function interface 1.0: an addendum to the Haskell 98 report.** 2003. Available at: <http://www.cse.unsw.edu.au/~chak/haskell/ffi/>. Last access: 12th July 2009. 37p.

CHOWDHURY, R. R. Landscape change in the Calakmul Biosphere Reserve, Mexico: modelling the driving forces of smallholder deforestation in land parcels. **Applied Geography**, v. 26, n. 2, p. 129-152, 2006.

CLARAMUNT, C.; JIANG, B. An integrated representation of spatial and temporal relationships between evolving regions. **Journal of Geographical Systems**, v. 3, n. 4, p. 411-428, 2001.

COSTA, S. S.; CÂMARA, G.; PALOMO, D. TerraHS: Integration of functional programming and spatial datasets for GIS application development. In: SIMPÓSIO BRASILEIRO DE GEOINFORMÁTICA, 8. (GEOINFO), 2006, Campos do Jordão. **Anais**... São José dos Campos: INPE, 2006, p. 109-127.

COSTA, S. S.; AGUIAR, A. P.; CÂMARA, G. Prototyping GIS application in functional programming. In: Simpósio de Geotecnologias no Pantanal, 2, GeoPantanal, Corumbá, MS. **Proceedings**... Accepted, 2009.

DE ESPINDOLA, G.; AGUIAR, A. D.; CAMARA, G. Modelling deforestation and its intraregional interactions in the Brazilian Amazon: market pressure versus public policies. In: Conferência Amazônia em Perspectiva: Ciência Integrada para um Futuro Sustentável, Manaus, AM. **Proceedings...**, 2008.

ERWIG, M.; GÜTING, R. H.; SCHNEIDER, M.; VAZIRGIANNIS, M. Spatiotemporal data types: an approach to modelling and querying moving objects in datasets. **GeoInformatica**, v. 3, n. 3, p. 269-296, 1999.

ERWIG, M.; SCHNEIDER, M. Spatio-Temporal Predicates. **IEEE Transactions on Knowledge and Data Engineering**, v. 14, n. 4, p. 881-901, 2002.

ERWIG, M.; SCHNEIDER, M. STQL: A spatiotemporal query language. In: LADNER, R.; SHAW, K.; ABDELGUERFI, M. (Ed.). **Mining spatiotemporal information systems**. Dordrecht: Kluwer Academic, 2002a, p. 105-126.

ESCADA, M. I. S. **Evolução de padrões de uso e cobertura da terra na região Centro-Norte de Rondônia**. 2003. 264 p. (INPE-10209-TDI/899). Doctorate thesis at post-graduation course in Remote Sensing. Instituto Nacional de Pesquisas Espaciais, São José dos Campos. 2003.

ESCADA, M. I. S.; MONTEIRO, A. M. V.; AGUIAR, A. P. D.; CARNEIRO, T. G. S.; CÂMARA, G. Análise de padrões e processos de ocupação para a construção de modelos na Amazônia: experimentos em Rondônia. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 12. (SBSR), 2005, Goiânia. **Anais**... São José dos Campos: INPE, 2005. p. 2973-2984.

FRANK, A. One step up the abstraction ladder: combining algebras - from functional pieces to a whole. COSIT - Conference on spatial information theory. **Proceedings**... Springer-Verlag, p. 95-108, 1999.

FRANK, A. Ontology for Spatio-temporal Datasets. In: KOUBARAKIS, M.; SELLIS, T. (Ed.). **Spatio-temporal datasets: the chorochronos approach**. Berlin: Springer-Verlag, 2003, p. 9-78. Lecture Notes in Computer Science, v 2520.

FRANK, A.; KUHN, W. Specifying Open GIS with functional languages. In: EGENHOFER, M.; HERRING, J. (Ed.). **Advances in Spatial Datasets—4th International Symposium, SSD '95**, Portland, ME. Berlin: Springer-Verlag, 1995, p. 184-195. Lecture Notes in Computer Science, v. 951.

GALTON, A. Fields and Objects in Space, Time, and Space-time. **Spatial Cognition and Computation**, v. 4, n. 1, p. 39-68, 2004.

GOODCHILD, M.; YUAN, M.; COVA, T. Towards a general theory of geographic representation in GIS. **International Journal of Geographical Information Science**, v. 21, n. 3, p. 239 - 260, 2007.

GRENON, P.; SMITH, B. SNAP and SPAN: Towards Dynamic Spatial Ontology. **Spatial Cognition & Computation**, v. 4, n. 1, p. 69-104, 2003.

GÜTING, H.; BEHR, T.; ALMEIDA, V. T. D.; DING, Z.; HOFFMANN, F.; SPIEKERMANN, M. **SECONDO: An Extensible DBMS Architecture and Prototype**. Report. Fernuniversität Hagen, 2004.

GÜTING, R. H.; BOHLEN, M. H.; ERWIG, M.; JENSEN, C. S.; LORENTZOS, N.; NARDELLI, E.; SCHNEIDER, M.; VIQUEIRA, J. R. R. Spatiotemporal models and languages: an approach based on data types. In: KOUBARAKIS, M.; SELLIS, T. (Ed.). **Spatio-Temporal Datasets**. Berlin: Springer-Verlag, 2003. Lecture Notes in Computer Science v2520, p. 117-176.

GÜTING, R. H.; SCHNEIDER, M. **Moving objects datasets**. San Francisco, CA, USA: Morgan Kaufmann Publishers, 2005. ISBSN 0-12-088799-1.

HORNSBY, K.; EGENHOFER, M. Qualitative representation of change, in spatial information theory: a theoretical basis for GIS. In: **International Conference COSIT '97**, Springer-Verlag, 1997. Lecture Notes in Computer Science 1329, p. 15-33.

HORNSBY, K.; EGENHOFER, M. Identity-based change: a foundation for spatiotemporal knowledge representation. **International Journal of Geographical Information Science**, v. 14, n. 3, p. 207-224, 2000.

HORNSBY, K. S.; COLE, S. Modelling moving geospatial objects from an event-based perspective. **Transactions in GIS**, v. 11, n. 4, p. 555-573, 2007.

HUANG, B.; CLARAMUNT, C. STOQL: An ODMG-based spatiotemporal object model and query language. In: RICHARDSON, D.; OOSTEROM, P. (EDs), **Proceedings of the 9th Spatial Data Handling Symposium**, Ottawa, Canada, Springer, p. 225-237, 2002.

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS (INPE). **PRODES - Program for deforestation assesment in the Brazilian Legal Amazonia**. São José dos Campos, Brazil: INPE, 2009. Available at: www.obt.inpe.br/prodes. Last access: 29th March 2009.

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS (INPE). **DETER - Real time deforestation detecting**. São José dos Campos, Brazil: INPE, 2009b. Available at: www.obt.inpe.br/deter. Last access: 29th July 2009.

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS (INPE). **Forest cover monitoring by sattelites: PRODES, DETER, DEGRAD and QUEIMADAS systems 2007-2008**. São José dos Campos, Brazil, 2009c. 47 p. Available at: <http://www.obt.inpe.br/prodes/Relatorio_Prodes2008.pdf>. Last access: 29th July 2009.

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS (INPE). **DEGRAD - Program for mapping forest degradation in the Brazilian Legal Amazonia**. São José dos Campos, Brazil: INPE, 2009d. Available at: www.obt.inpe.br/degrad. Last access: 29th March 2009.

JONES, S. P. (Ed.). **Haskell 98 language and libraries: the revised report**. Cambridge: Cambridge University Press, 2003. ISBN 0-521-826144.

KLIPPEL, A.; WORBOYS, M.; DUCKHAM, M. Identifying factors of geographic event conceptualisation. **International Journal of Geographical Information Science**, v. 22, n. 2, p. 183-204, 2007.

LAMBIN, E. F.; GEIST, H. J.; LEPERS, E. Dynamics of land use and land-cover change in tropical regions. **Annual Review of Environment and Resources**, v. 28, p. 205-241, 2003.

LINKIE, M.; SMITH, R. J.; LEADER-WILLIAMS, N. Mapping and predicting deforestation patterns in the lowlands of Sumatra. **Biodiversity and Conservation**, v. 13, p. 1809-1818, 2004.

MARQUESINI, M.; ADARIO, P.; LEITÃO, S.; FURTADO, M.; DONATTI, C.; GOYET, C. Desmatamento na Amazônia: O leão acordou - uma análise do Plano de Ação para a Prevenção e Controle do Desmatamento na Amazônia Legal. ed. **Greenpeace**, March 2008. Available at: <http://www.greenpeace.org/raw/content/brasil/documentos/amazonia/desmatamento-na-amaz-nia-o-l.pdf>. Last access: 29th July 2009.

MEDAK, D. Lifestyles. In: FRANK, A. U., RAPER, J., & CHEYLAN, J.-P. (Ed.). **Life and motion of socio-economic units**. London: Taylor & Francis, 2001. ESF Series.

MERTENS, B.; LAMBIN, E. Spatial modelling of deforestation in Southern Cameroon: spatial disaggregation of diverse deforestation processes. **Applied Geography**, v. 17, p. 143-162, 1997.

MOTA, J. S.; CAMARA, G.; ESCADA, M. I.; BITTENCOURT, O. O.; FONSECA, L. M. G.; VINHAS, L. Case-Based Reasoning eliciting the evolution of geospatial objects. **Proceedings of the Conference on Spatial Information theory: COSIT'09**. Aber Wrac'h, France, p. 405-420, 2009.

MOTA, J. S.; CAMARA, G.; OLIVEIRA, O. F.; FONSECA, L. M. G.; ESCADA, M. I. S. Applying case-based reasoning in the evolution of deforestation patterns in the Brazilian Amazonia. In: 23RD ANNUAL ACM SYMPOSIUM ON APPLIED COMPUTING (SAC'08), Fortaleza, 2008. **Proceedings**... New York, NY, USA: ACM, v. 1, p. 1683-1687, 2008.

OGC. **The OpenGIS® guide: introduction to interoperable geoprocessing and the OpenGIS specification**. Boston: Open GIS Consortium, Inc, 1996.

PELEKIS, N.; THEODOULIDIS, B.; KOPANAKIS, I.; THEODORIDIS, Y. Literature review of spatio-temporal dataset models. **The Knowledge Engineering Review**, v. 19, n. 3, p. 235-274, 2004.

RODDICK, J.; EGENHOFER, M.; HOEL, E; PAPADIAS, D; SALZBERG, B. Spatial, Temporal and Spatiotemporal Datasets Hot Issues and Directions for PhD Research. SIGMOD '93, **SIGMOD Record**, v. 33, n. 2, p. 126-131, 2004.

ROGAN, J.; CHEN, D. M. Remote sensing technology for mapping and monitoring land-cover and land use change. **Progress in Planning**, v. 61, n.4, p. 301-325, 2004.

ŠALTENIS, S.; JENSEN, C. S.; LEUTENEGGER, S. T.; LOPEZ, M. A. Indexing the positions of continuously moving objects. **Proceedings of the ACM SIGMOD International Conference on Management of data**, v. 29, n. 2, p. 331-342, 2000.

SELLIS, T.; FRANK, A. U.; GRUMBACH, S.; GÜTING, R. H.; KOUBARAKIS, M. (Ed.). **Spatio-Temporal Datasets: The Chorochronos Approach.** Berlin: Springer, 2003. Lecture Notes in Computer Science, v.2520. ISBN: 978-3-540-40552-8.

SILVA, M. P. S.; CÂMARA, G.; SOUZA, R. C. M.; VALERIANO, D. M.; ESCADA, M. I. S. Mining patterns of change in remote sensing image datasets. In: WORKSHOP DOS CURSOS DE COMPUTAÇÃO APLICADA DO INPE, 5.

(WORCAP), 2005, São José dos Campos. **Anais**... São José dos Campos: INPE, 2005.

SILVA, M. P. S.; CÂMARA, G.; ESCADA, M. I. S.; SOUZA, R. C. M. Remote sensing image mining: detecting agents of land use change in tropical forest areas. **International Journal of Remote Sensing**, v. 29, n. 16, p. 4803-4822, 2008.

SINGH, A. Digital change detection techniques using remotely-sensed data. **International Journal of Remote Sensing**, v. 10, p. 989-1003, 1989.

SISTLA, A. P.; WOLFSON, O.; CHAMBERLAIN, S.; DAO, S. Modelling and Querying Moving Objects. **Proceedings of the 13th International Conference on Data Engineering (ICDE'97)**, p. 422-432, April 1997.

WALDROP, M. M.; Markets can save forests. In: CAMPBELL, P.; CAMPBELL N.; MILLER, L. (Eds.). **Nature**, n. 452, p. 127-128, 2008. Available at <http://www.nature.com/nature/journal/v452/n7184/full/452127b.html>. Last access: 29th July 2009.

WORBOYS, M. Event-oriented approaches to geographic phenomena. **International Journal of Geographical Information Science**, v. 19, n. 1, p. 1-28, 2005.

ZIPPERER, W. C. Deforestation patterns and their effects on forest patches. **Landscape Ecology**, v. 8, p. 177-184, 1993.

ANNEX A – COMPUTATIONAL ENVIRONMENT

To implement the *geospatial algebra* and model the evolution of deforestation we used TerraLib library, TerraHS environment and TerraView application. They were developed as part of Brazilian National Institute for Space Research (INPE) researches on Geographic Information Systems (GIS). TerraLib is a geographic library that contains functions to allow the development of creation of GIS tools. TerraHS is a geographic environment to allow prototyping and performing algebras. The *geospatial algebra* was built within the TerraHS to extend this environment to handle the evolution of *geospatial objects*. TerraView application implements the database format proposed on TerraLib and was used to visualize our case studies. Next sections present some details about them.

A.1 TerraLib

TerraLib (CÂMARA et al., 2008) is an open-source GIS software library developed at INPE to support large-scale applications using socio-economic, cadastral and environmental data. Its core development team includes the Image Processing Division (DPI/INPE), the Computer Graphics Technology Group of the Catholic University of Rio de Janeiro (TECGRAF-PUC-RIO) and GIS Division on Research Foundation for Space Technology (GEO/FUNCATE).

The most relevant feature is to extend object-relational Database Management System (DBMS) technology to support spatiotemporal models, spatial analysis, image datasets, dynamic modelling and to allow spatial, temporal and attribute queries on the dataset. Further, this library provides the ease of customization and upward compatibility to the OpenGIS Consortium (OGC, 1996).

TerraLib is developed in C++ programming language and provides independence of DBMS, efficiency and extensibility. Therefore, TerraLib allows a collaborative environment for the development of multiple GIS applications and the use of technological advances on spatial databases.

A.2 TerraView

TerraView is an open source geographic application that employs TerraLib library. This provides basic functions of data conversion, display, exploratory spatial data analysis, spatial statistical modelling, and spatial and non-spatial queries. Several Brazilian public institutions use TerraView for public policy making, developing, for example, studies in spatial epidemiology and crime analysis. Figure A.1 shows the user interface for the TerraView application.

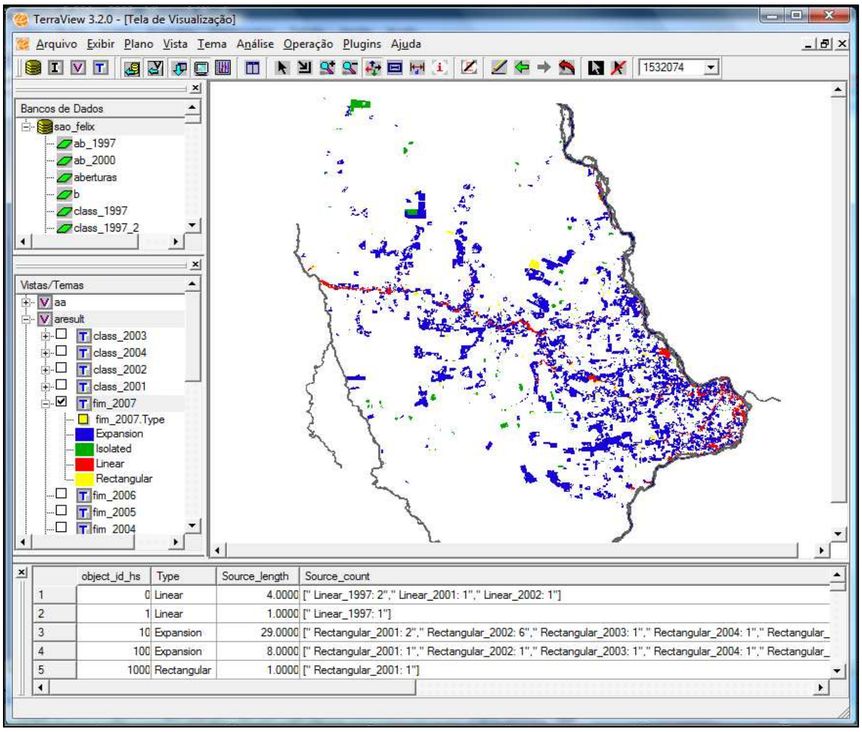


Figure A.1 - User interface for the TerraView application.

A.3 TerraHS

TerraHS (COSTA et al., 2006) is an environment that uses the data handling capabilities of TerraLib to enable the development of geographic applications in Haskell functional language. TerraHS takes the advantages of using functional languages to express algebraic theories: rapid prototyping and the simple, complete and extensible definition. TerraHS has methods to handle geo-objects and geo-fields data types, as defined by CÂMARA et al. (1995). Geo-objects represent phenomena that may have one or more graphical representations and correspond to the geo-referenced set of coordinates that describe the location of the objects. Geo-fields represent continuous geographical variables over some region of the Earth. Further, TerraHS allows the development of specific data types, the extensibility that we used to define the *geospatial algebra*.

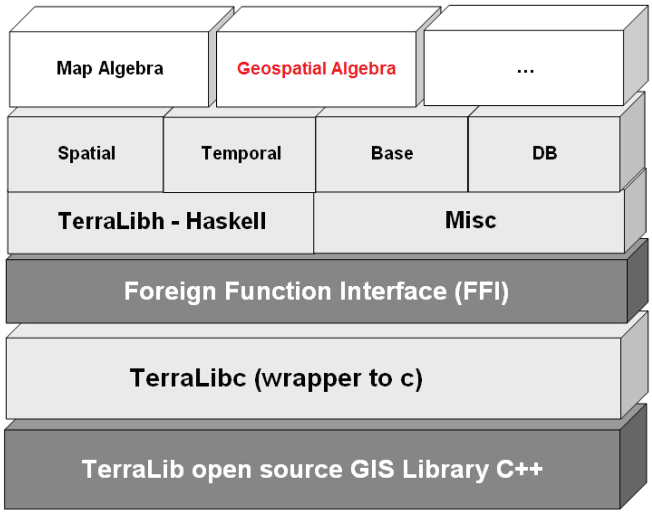


Figure A.2 - Archictecture of TerraHS.
Source: adapted from COSTA et al. (2006).

Now, we present a brief description of TerraHS adapted from (COSTA et al., 2009). TerraHS creates the link to TerraLib library by using the *Foreign Function Interface* (FFI) (CHAKRAVARTY, 2003) with additional code written in C (TerraLibC), which maps the FFI to TerraLib methods. Figure A.2 illustrates the architecture of TerraHS separated by the components and its relationships. Lighter colors represent the parts provided by TerraHS and darker colors represent the existing components.

Lower layers provide basic services over which upper layer services are implemented. In the bottom layer, TerraLib supports different spatial dataset management and many spatial algorithms. In the second layer, TerraLibC maps the Terralib C++ methods to the Haskell FFI. In the third layer, the FFI enables calling the TerraLibC functions from Haskell. In the fourth layer, TerraLibH contains the modules that map TerraLib C++ classes to Haskell data types and functions, TeGeometry.hs, TeDataset.hs and so on. Misc holds the modules that provide auxiliary functions to TerraHS, such as string and generic functions.

The fifth layer contains data types and services to support new specific algebras in the last layer. This is the level that we added the *geospatial algebra* on the TerraHS. They describe algebraic abstract data types for spatial, temporal, dataset and base data types. To deal with specific data types in a generic way, TerraHS contains two classes: the type class ModelConvert to map between TerraHS data types and specific data types, and the type class ModelPersistence to provide generic functions to store and to retrieve specific data types from a spatial dataset. Figure A. 3 shows this approach.

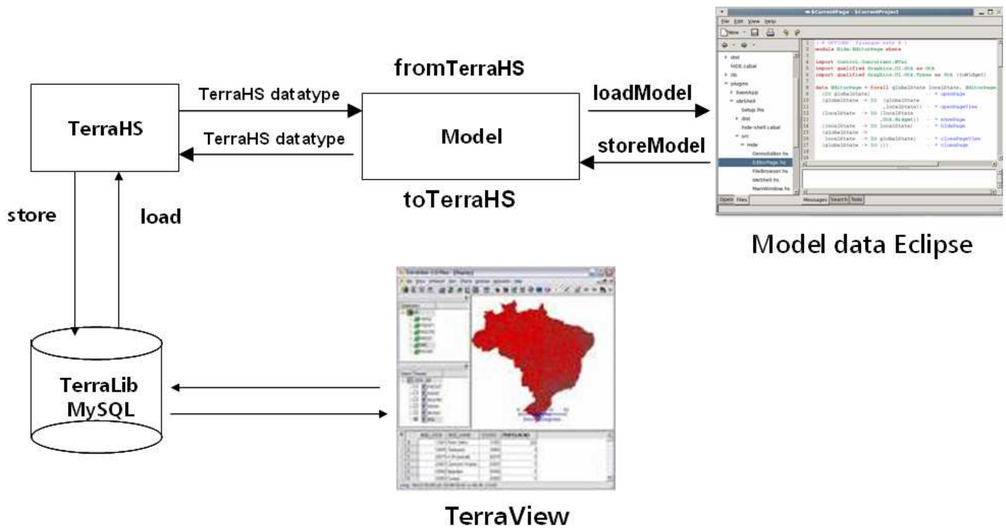


Figure A. 3 - Retrieving and storing specific data type from spatial database.
Source: COSTA et al.(2009).

Geospatial algebra was developed by adding the type class EvolvingObject that implement the functions. Further, we adapted our objects and some functions to the geo-object available data type. The application is done on a main Haskell program accessing the operations defined on the algebra.

ANNEX B– USING THE PROGRAM FOR DEFORESTATION ASSESMENT IN THE BRAZILIAN LEGAL AMAZONIA (PRODES) DATA: A FUNCTIONAL ROAD MAP

The Brazilian National Institute for Space Research (INPE) uses satellite images to provide yearly assessments of the deforestation in the Brazilian Amazon. Data from PRODES, the deforestation monitoring program, show that nearly 37,000,000 ha of forest were cut from 1988 to 2008. PRODES uses images from TM and CCD sensors. The TM, aboard of Landsat satellite from NASA, has spatial resolution of 30 meters and covers Brazil every sixteen days. The CCD, carried on the CBERS satellite from INPE, presents a spatial resolution of 20 meters and covers Brazil every twenty six days. These sensors are precise enough to correctly indicate deforestation areas greater than 6,25 ha. Using this precision, PRODES calculates the annual deforestation rate based on the area of patches, regions where clear cut deforestation was detected and that point out the forest cover was completely removed. Figure B.1 exemplifies clear-cut detected regions.

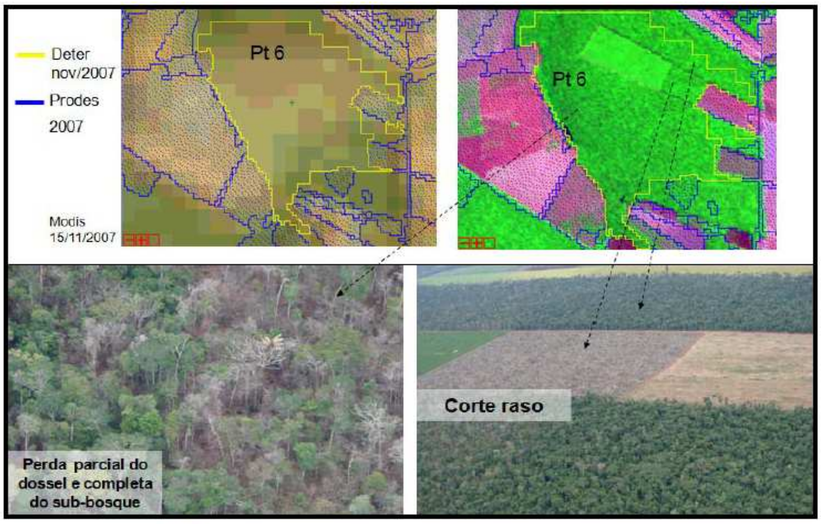


Figure B.1 - Complete loss of land cover by clear-cut.
Source: INPE (2009).

The aim of this annex is to show the used data and how we prepared these data to our experiments in the thesis.

B.1 Identification of landscape objects

We used scenes in vectorial maps on shapefile format available at PRODES (INPE, 2009). To recover deforestation data on these scenes, it is necessary to fit the PRODES data to describe real deforested areas because some objects are separated in distinct patches. This is due to operational limitations imposed by the evaluation method within PRODES system. To solve this problem we defined an identification process composed of a sequence of two steps: *Semantic Cleaning* and *Temporal Consolidation*, illustrated in Figure B.2.

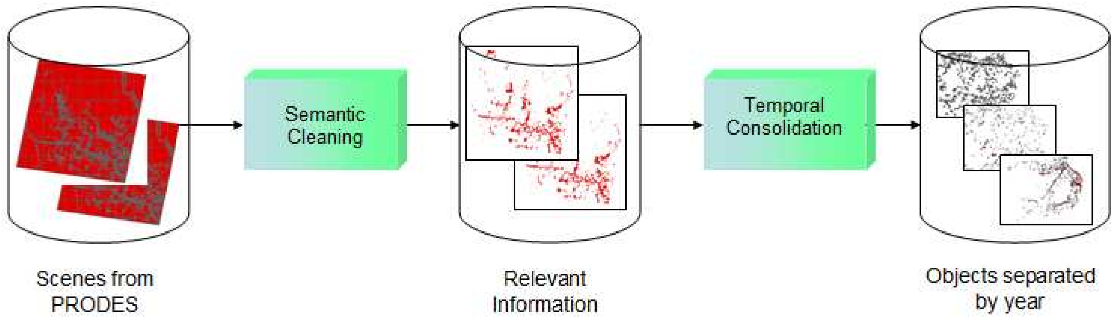


Figure B.2 - Identification of *landscape objects* from PRODES monitoring system.

Then, the first step, the *Semantic Cleaning*, corresponds to separate the patches to obtain a dataset that contains only deforested regions. The dataset is classified based on its semantics, such as 'deforestation', 'clouds' or 'hydrograph', and the main idea is to separate the relevant information to evolution: the 'deforestation' objects.

The annual monitoring is done by increments and the dataset of each year contains all information about previous monitoring since 1997. The second step, then, corresponds to separate the *landscape objects* of each year, the *Temporal Consolidation*. We applied the *geospatial algebra* to rebuild objects that are separated in parts. In other words, we model the evolution with the following rule: *if a patch of a timestamp touches another patch within the same timestamp, they merge because they are the same landscape object*. Figure B.3

depicts a real *landscape object* from a specific timestamp separated in seven distinct patches. If this object is not consolidated before evolution, the evolution generates 'false transformations' because this division will evolve considering seven distinct objects.

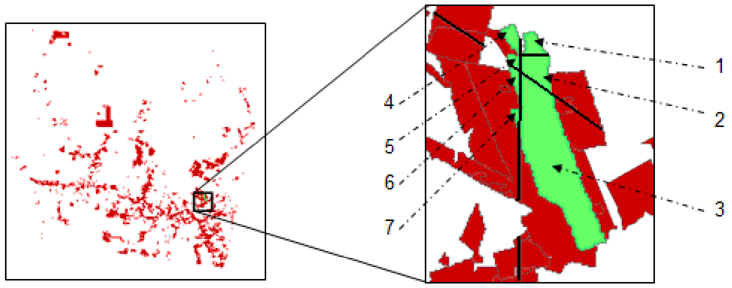


Figure B.3 - Real example of an object separated in seven parts.

After the *Temporal Consolidation*, we have *Objects Datasets* (OD) with precise spatial configuration to be examined or evolved.

B.2 Studied deforestation areas

This Section presents the original PRODES data used in the case studies in this thesis and the objects resulting after applying the *Semantic Cleaning* and *Temporal Consolidation* processes. The Novo Progresso and Terra do Meio studied areas are located in Pará State and concern to areas with intense deforestation process in the last 10 years. For each case study, we analyzed temporal series of ten timestamps in the time interval from 1997 to 2008, the most recent data produced by PRODES that become available on July 2009.

Figure B.4 illustrates the PRODES interface with the snapshot of LandSat satellite scenes in the Novo Progresso municipality. In this case study we use data from scenes 227/65 and 227/66 that cover an area of around 8,014,000 ha and contain 22,613 original patches. The studied area in Terra do Meio matches to scenes 226/64, 225/64, 226/65 and 225/65 of TM/LandSat satellite that cover around 7.600.000 ha with 51,487 original patches.

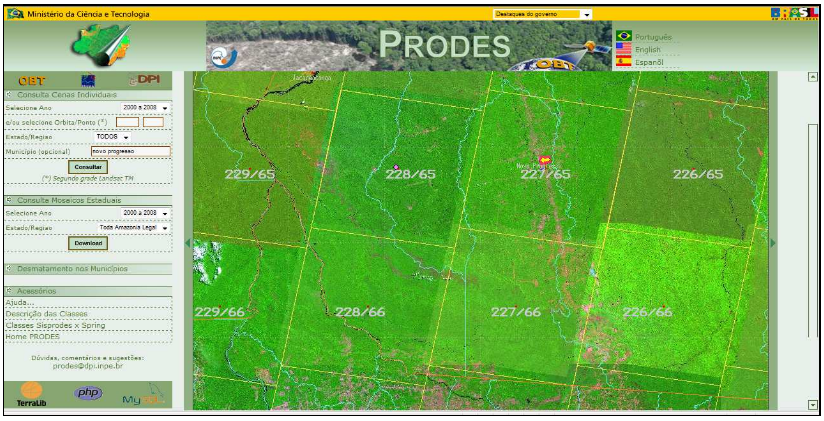


Figure B.4 - PRODES snapshot of TM/LandSat scenes in the Novo Progresso municipality.

Table B.1 presents the number of objects used in both case studies. First column contains the number of original objects and the second contains the number of objects after the *Semantic Cleaning*. The percentage of objects that indicate deforestation is showed in the third column. In the Novo Progresso case study around 67% from the original patches corresponds to deforestation objects. In Terra do Meio, this value corresponds to 75%. These results shows the high number of objects do not used in deforestation evolution analysis. The last column shows the number of objects after *Temporal Consolidation*. Around 20% of objects in Novo Progresso were merged to correspond to the real regions of deforestation. Terra do Meio presents a very distinct value because we separated the objects that are not part of the specific interesting studied region after the *Semantic Cleaning*.

Table B.1 - Number objects in the Novo Progresso and Terra do Meio case studies.

Scene	Original Patches	Objects after *Semantic Cleaning*	Percentage of Objects of Deforestation	Objects after *Temporal Consolidation*
Novo Progresso	22,613	15,309	67,70%	12,328
Terra do Meio	51,487	38,790	75,34%	11,681