UNIVERSITÉ LAVAL

Towards Development of Fuzzy Spatial Datacubes:
Fundamental Concepts with Example for Multidimensional Coastal Erosion Risk Assessment and Representation

Thèse

Amaneh Jadidi Mardkheh

Doctorat en Sciences Géomatiques
Philosophiae Doctor (Ph.D.)

Québec, Canada

Résumé

Les systèmes actuels de base de données géodécisionnels (GeoBI) ne tiennent généralement pas compte de l'incertitude liée à l'imprécision et le flou des objets; ils supposent que les objets ont une sémantique, une géométrie et une temporalité bien définies et précises. Un exemple de cela est la représentation des zones à risque par des polygones avec des limites bien définies. Ces polygones sont créés en utilisant des agrégations d'un ensemble d'unités spatiales définies sur soit des intérêts des organismes responsables ou les divisions de recensement national. Malgré la variation spatio-temporelle des multiples critères impliqués dans l'analyse du risque, chaque polygone a une valeur unique de risque attribué de façon homogène sur l'étendue du territoire. En réalité, la valeur du risque change progressivement d'un polygone à l'autre. Le passage d'une zone à l'autre n'est donc pas bien représenté avec les modèles d'objets bien définis (crisp).

Cette thèse propose des concepts fondamentaux pour le développement d'une approche combinant le paradigme GeoBI et le concept flou de considérer la présence de l'incertitude spatiale dans la représentation des zones à risque. En fin de compte, nous supposons cela devrait améliorer l'analyse du risque. Pour ce faire, un cadre conceptuel est développé pour créer un model conceptuel d'une base de donnée multidimensionnelle avec une application pour l'analyse du risque d'érosion côtier. Ensuite, une approche de la représentation des risques fondée sur la logique floue est développée pour traiter l'incertitude spatiale inhérente liée à l'imprécision et le flou des objets. Pour cela, les fonctions d'appartenance floues sont définies en basant sur l'indice de vulnérabilité qui est un composant important du risque. Au lieu de déterminer les limites bien définies entre les zones à risque, l'approche proposée permet une transition en douceur d'une zone à une autre. Les valeurs d'appartenance de plusieurs indicateurs sont ensuite agrégées basées sur la formule des risques et les règles SI-ALORS de la logique floue pour représenter les zones à risque. Ensuite, les éléments clés d'un cube de données spatiales floues sont formalisés en combinant la théorie des ensembles flous et le paradigme de GeoBI. En plus, certains opérateurs d'agrégation spatiale floue sont présentés.

En résumé, la principale contribution de cette thèse se réfère de la combinaison de la théorie des ensembles flous et le paradigme de GeoBI. Cela permet l'extraction de connaissances plus compréhensibles et appropriées avec le raisonnement humain à partir de données spatiales et non-spatiales. Pour ce faire, un cadre conceptuel a été proposé sur la base de paradigme GéoBI afin de développer un cube de données spatiale floue dans le system de Spatial Online Analytical Processing (SOLAP) pour évaluer le risque de l'érosion côtière. Cela nécessite d'abord d'élaborer un cadre pour concevoir le modèle conceptuel basé sur les paramètres de risque, d'autre part, de mettre en œuvre l'objet spatial flou dans une base de données spatiales multidimensionnelle, puis l'agrégation des objets spatiaux flous pour envisager à la représentation multi-

échelle des zones à risque. Pour valider l'approche proposée, elle est appliquée à la région Perce (Est du Québec, Canada) comme une étude de cas.

Abstract

Current Geospatial Business Intelligence (GeoBI) systems typically do not take into account the uncertainty related to vagueness and fuzziness of objects; they assume that the objects have well-defined and exact semantics, geometry, and temporality. Representation of fuzzy zones by polygons with well-defined boundaries is an example of such approximation. This thesis uses an application in Coastal Erosion Risk Analysis (CERA) to illustrate the problems. CERA polygons are created using aggregations of a set of spatial units defined by either the stakeholders' interests or national census divisions. Despite spatiotemporal variation of the multiple criteria involved in estimating the extent of coastal erosion risk, each polygon typically has a unique value of risk attributed homogeneously across its spatial extent. In reality, risk value changes gradually within polygons and when going from one polygon to another. Therefore, the transition from one zone to another is not properly represented with crisp object models.

The main objective of the present thesis is to develop a new approach combining GeoBI paradigm and fuzzy concept to consider the presence of the spatial uncertainty in the representation of risk zones. Ultimately, we assume this should improve coastal erosion risk assessment. To do so, a comprehensive GeoBI-based conceptual framework is developed with an application for Coastal Erosion Risk Assessment (CERA). Then, a fuzzy-based risk representation approach is developed to handle the inherent spatial uncertainty related to vagueness and fuzziness of objects. Fuzzy membership functions are defined by an expert-based vulnerability index. Instead of determining well-defined boundaries between risk zones, the proposed approach permits a smooth transition from one zone to another. The membership values of multiple indicators (e.g. slop and elevation of region under study, infrastructures, houses, hydrology network and so on) are then aggregated based on risk formula and Fuzzy IF-THEN rules to represent risk zones. Also, the key elements of a fuzzy spatial datacube are formally defined by combining fuzzy set theory and GeoBI paradigm. In this regard, some operators of fuzzy spatial aggregation are also formally defined.

The main contribution of this study is combining fuzzy set theory and GeoBI. This makes spatial knowledge discovery more understandable with human reasoning and perception. Hence, an analytical conceptual framework was proposed based on GeoBI paradigm to develop a fuzzy spatial datacube within Spatial Online Analytical Processing (SOLAP) to assess coastal erosion risk. This necessitates developing a framework to design a conceptual model based on risk parameters, implementing fuzzy spatial objects in a spatial multi-dimensional database, and aggregating fuzzy spatial objects to deal with multi-scale representation of risk zones. To validate the proposed approach, it is applied to Perce region (Eastern Quebec, Canada) as a case study.

Table of Content

List of Tables

List of Figures

Acronym

BI	Business Intelligence
CER	Coastal Erosion Risk
CERA	Coastal Erosion Risk Assessment
CVI	Coastal Vulnerability Index
CSoVI	Coastal Social Vulnerability Score
DBMS	Database Management Systems
DSS	Decision Support Systems
DTM	Digital Terrain Models
DEM	Digital Elevation Modeling
ETL	Extract-Transfer-Load
GeoBI	Geospatial Business Intelligence
GIS	Geographical Information Systems
HWL	High Water Line
ICZM	Integrated Coastal Zone Management
IC	Integrity Constraints
MHWL	Mean High Water Line
MSL	Mean Sea Level
QMM	Qualified Min-Max
OLAP	Online Analytical Processing
SDSS	Spatial Decision Support Systems
SOLAP	Spatial Online Analytical Processing
SoVI	Social Vulnerability Index
SI	Sensitivity Index
SCI	Sustainable Capacity Index
SLR	Sea Level Rise
SMCM	Spatial Multidimensional Conceptual Model
SQL	Structure Query Language
TL	Tide level

Dedicated to those to whom I always belong
for their love, endless support, and
encouragement
To:
my best friend and husband: Kyarash,
my sweetheart daughter: Amytis,
my beloved parents: Masi and Reza,
my sisters: Malahat, Matine, Mahsa, Rezvan,
and, finally, to the World that respects
SCIENCE!

Acknowledgments

I must begin this section by thanking all of the people whom I have been blessed to work with over the past five years!

To my supervisor, ***Dr. Mir Abolfazl Mostafavi***, I am so happy we started this thesis with lots of doubts and have finished it with strong confidence and assurance. However, the degree of uncertainty exists in everywhere and cannot be completely avoided. I would like to thank you for the five years of key advices kindness, and support you provided despite your tight schedule. I appreciated so much to work with you and having such experience.

Dr. Yvan Bédard, having you as co-supervisor was the best thing that happened during my thesis. Not only I learned tremendous amounts about geospatial domain but also you left me with the legacy of never giving up. This is a lesson of life that will always be with me. Your office was always opened for all my crazy questions and surprises. You have my deepest gratitude, Yvan. Someday, I hope to become someone that you are proud of having had as a student. Thank you!

I would like to thank ***Dr. Bernard Long*** his important role in building this Ph.D. project. I am sure that without his thoughtful advice at the beginning, we could not have realized the complexity of the application under study. A special thank goes to also to ***Dr. Eliane Proppek*** for her advices in risk analysis. Great thanks also belong to ***Dr. Stephan Roche***, for accepting to be "pre-Lecture" of this thesis and later in Jury and to ***Dr. Marc Cocard*** for his support as graduate program director and president of jury.

Thanking ***Ms. Carmen Couture*** is really difficult. She served as more than an Educational Agent on my dossier. She became my second family in Quebec. I did not feel lonely when she was around at tough times. Carmen, you have earned the deepest gratitude of my heart. You changed my life when I was nearly lost in this white Snow-Land.

Once I asked **my father** what treasure I would inherit from him. He answered me: "You"! For a long time, I was baffled by his unexpected response. Now, I am able to see that indeed the richest inheritance from my parents is to be ME! **Mom** and **Dad**, thank you for your sacrifices and unending support. You are always with me, even when we are separated by thousands of kilometers. Your hearts are always with me and mine is yours! Please accept this manuscript as a token of my appreciation of all your sacrifices.

My little sisters, ***Malahat, Matine, Mahsa, Rezvan***: I know I was not there for you every time you needed me and I could not fulfill all of my sisterly duties. However, you dwell in my heart forever and nothing is more important than you. Thank you for your years of patience and encouragement.

To you my best friend and husband, ***Kyarash***: since our eyes met in that beautiful Alpine valley of Grenoble, you assured me of your everlasting support, love, honesty, and confidence. You taught me how much the real life is fuzzy and how should we be prepared for an optimum solution. We had many tough periods all these years, but your positive belief in my aptitude always drives me to improve. Saying only thank you is so cliché. I prefer to say that--without you--it would have been impossible to make my dreams come true! I will do my best to always be the ONE you've searched for. ☺

Now, some words for my sweetie, ***Amytis***: I know that I sacrificed precious time with you to work on this thesis and I keenly felt your absence. I hope that someday all the hard work I put into it and hereafter in my career makes you proud of your Mommy. I will be so relieved when that day comes.

To my dearest friends: ***Atiyeh, Jessie*** and ***Alborz:*** Merely saying only "thank you" is inadequate for all you've done for me over these five years. Your encouragement, your presence, and your support have been remarkable. Also, I would like to thank all my colleagues at Center for Research in Geomatics: ***Eve, Hedia, Mathieu, Abbas, Marc, Elodie, Vincent, Guillaume, Tania, Eric, Sonia,*** and ***Eveline***. Well-deserved gratitude goes to Ms. **Danielle Goulet** for her incredible work to streamline financial procedure. In closing, I would like to thank all the people in my life who have taught me new things. Each of those lessons has contributed to the achievement of this lifelong goal.

Preface

The results of this research project have been published and communicated in three international scientific journals and five international and national conferences on Geospatial Information Science. These papers have been written by the author of this thesis under supervision of Dr. Mostafavi and Dr. Bedard and wit the collaboration of Dr. Long, Dr. Shahriari and Ms. Grenier.

Journal Papers:

Jadidi A., Bédard Y., Mostafavi M.A., Fuzzy Spatial Datacube for Multi-scale Coastal Risk Assessment: Towards Fuzzy Spatial Aggregation to Support Geospatial Decision Models, International Journal of Geographical Information Science, Submitted.

Jadidi A., Mostafavi M.A., Bédard Y., Shahriari K., Spatial Representation of Coastal Risk: A Fuzzy Approach to Deal with Uncertainty, ISPRS International; Journal of Geo-Information, In process.

Jadidi A., Mostafavi M.A., Bédard Y., Long B., Grenier E., Using Geospatial Business Intelligence Paradigm to Design a Multidimensional Conceptual Model for Efficient Coastal Erosion Risk Assessment, International Journal of Coastal Conservation, 17 (3), 527-543, 2013.

Conference Papers and Communications:

Jadidi A., Mostafavi M.A., Bédard Y., Fuzzy-based Spatial Knowledge Discovery in Service of Efficient Coastal Erosion Risk Assessment, Spatial Knowledge Information Conference, Banff, 7-9 Feb 2014.

Jadidi A., Mostafavi M.A., Bédard Y., Dealing with Uncertainty in Coastal Risk Assessment: Fuzzy Representation of Coastal Risk Zones, GSDI 13 Geospatial Conference, Quebec, 14-17 May 2012

Jadidi A., Mostafavi M.A., Bédard Y., Long B., Toward an Integrated Spatial Decision Support System to Improve Coastal Erosion Risk Assessment: Modeling and Representation of Risk Zones, FIG Working Week 2012, Rome, Italy, 6-10 May 2012

Jadidi A., Mostafavi M.A., Bédard Y., Long B., Développement d'un outil géo-décisionnel pour améliorer l'évaluation du risque d'érosion côtière, Géomatique 2011, 12-13 oct. Montréal

Jadidi A., Mostafavi M.A., Long B., Where are we in Coastal Erosion Risk Assessment Using GIS?, Géomatique 2009, 22-23 Oct., Montréal.

"Making the simple complicated is commonplace; making the complicated simple, awesomely simple, that's creativity."
Charles Mingus

Chapter 1 Introduction

1.1 Research Context

Recent advances of Geospatial Business Intelligence (GeoBI) technologies, in handling large amounts of data led to intense growth in the adoption of such technologies for decision-making and integrated risk assessment processes. This growth has been powered by the decreasing cost of acquiring and storing large amount of data coming from multiple sources. Reducing the gap between data acquisition and decision-making is the main objective of business intelligence technologies (Chaudhuri et al. 2011). However, spatial uncertainty, which is an inherent characteristic of spatial data limits efficient decision-making process using GeoBI technologies (Kentel & Aral 2007; Bejaoui 2009). Spatial uncertainty can be propagated during integration and aggregation processes and hence in the decision-making process. Ignoring spatial uncertainty may result in unrealistic or misleading conclusions and decisions that yield undesired or even catastrophic consequences. For instance, confronting a disaster situation involves always with analyzing, and interpreting large amounts of spatiotemporal data and information from multiple sources with different types and levels of detail and uncertainty in a minimum time period. Therefore, coping with any undesired situation in our lives due to a natural or human-induced disaster requires an integrated analytical system such as GeoBI technologies. Hence, developing any system for such situations necessitates understanding and characterizing of phenomenon and related parameters and then conceptualizing the system. In this regard, an undesirable natural situation is described with some issues and limits to adapting GeoBI paradigms and handling spatial uncertainty are emphasized.

The earth with its natural climate variability is always coping with hazards caused by natural phenomena. However, not every natural hazard leads to a disaster. Whether natural hazards cause a disaster or not depends on their magnitude and the socio-economic and environmental vulnerabilities of the region being affected (ISDR 2004). Coastal regions are good examples in which natural hazards severely affect human-life and the present and future socio-economic activities of the people living along the coast (see Figure 1.1). Coastal regions are usually densely populated. It is claimed that 21% of the world's population install their accommodation, assets, and socio-economic activities less than 30 Km away from of coastline (IPCC 2007). Coastal communities are therefore increasingly concerned with the risk associated with coastal hazard, especially coastal erosion, given the fact that 70% of coastal regions around the world are subject to severe erosion (IPCC 2007). Canada specifically is a remarkable region with almost 240 000 Km of coastline-more than any other country in the world (Shaw et al. 1998). 7 000 Km of Canadian coasts are highly sensitive to climate change and Sea Level Rise (SLR) (Shaw et al. 1998; Stanton et al. 2010). Moreover, most of Canada's coastal zones are home to a concentrated population, economic centers, and valuable ecosystems which are excessively vulnerable to coastal erosion (Stanton et al. 2010). Furthermore, 38.3% of Canadians (seven millions) live within 20 km of a coast and their living expenses depend directly on the marine resources and tourism industry (Stanton et al. 2010). Considering that, an accurate modeling and presentation of risk associated to coastal erosion in the sensitive regions are crucial from socio-economic and environmental points of view. The misunderstanding of the risk may cause illusory protections or restrictions which may lead to natural, social, and economic disasters or unnecessary constraints on the development of regions. A sustainable development plan is consequently designed for optimal use of available resources while protecting people and socio-economic and environmental values of these regions.

Figure 1.1: Examples of disasters caused by coastal erosion, floods and storms along the Canadian coasts.

A sustainable development plan for coastal regions requires a significant understanding of the phenomena spatiotemporally. To do so, an integrated information system is required to perform necessary computation and representation spatiotemporally to convince human perception about the necessity of this plan (Goodchild & Glennon 2008). Several plans are elaborated on community, municipality, regional, national, and global scales to protect people and their assets from coastal risk. One of the most popular plans is Integrated Coastal Zone Management (ICZM) that emphasizes the sensitivity of coastal zones and their management (Varghese et al. 2008).

ICZM comprises information collection, planning, decision-making, management, and monitoring (Cheng et al. 2009). The schematic framework of ICZM is shown in Figure 1.2. Risk assessment is then a predominant sub-stage of any sustainable development plan and decision-making processes by providing a comprehensible estimation of damage and loss in terms of human, socio-economic, and environmental aspects. Risk assessment and representation provide necessary knowledge and essential guidelines for risk management and policy-making purposes. Coastal Erosion Risk Assessment (CERA) is thus defined as developing methods to determine the nature and extent of risk. To do so, analyzing the probability of hazard (erosion) and evaluating the existing vulnerable conditions of elements at risk, such as people, infrastructures, houses, etc. in given time periods is required within an integrated information system (Cutter et al. 2000; Cutter 2002; Karvetski et al. 2011; Boruff et al. 2005; Alexander 2000).

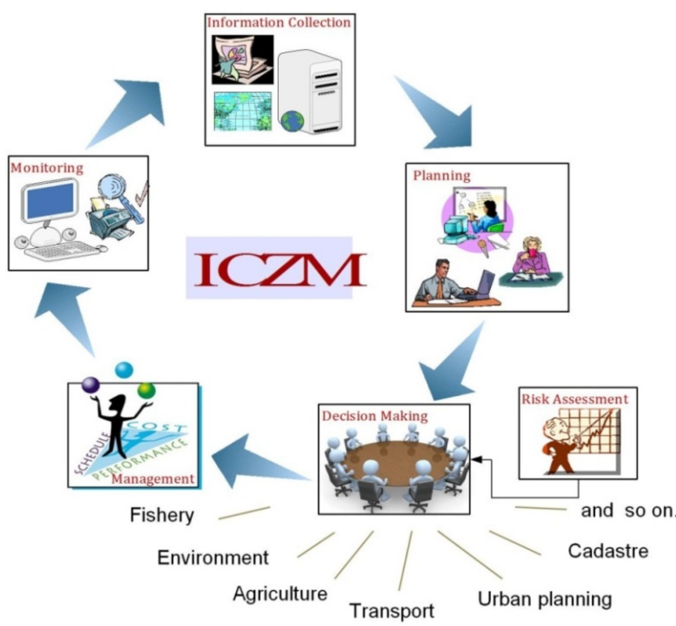


Figure 1.2: Integrated Coastal Zone Management cycle scheme

Coastal regions are managed by multiple stakeholders under the authority of different organizations in local, provincial and federal governments. These organizations are the main actors to make decisions for response option practices (i.e. protection, adaptation, and retreat) (Linham & Nicholls 2012). The main challenge arises when each organization prefers to take its own criteria and its own sources of data to assess the risk. The data and the criteria are often in conflict and it becomes difficult to provide a coherent vision of the overall risk, as well as support an efficient decision-making process. In other words, multiple interests are involved to evaluate risk and to make appropriate decisions. In this context, CERA is characterized as a complex, multi-criteria, spatiotemporal, and multi-scale process (Cutter et al. 2008; Cutter et al. 2003; Boruff et al. 2005; Blaikie et al. 2004). Therefore, assessing risk requires dealing with multidisciplinary studies, with multiple interests, and handling large amounts of data and information from multiple sources with different types, multiple scales and multi-epochs. On the other hand, the main challenge in decision-making processes especially in the case of disaster management and risk assessment is to shorten the time lag between data acquisition and decision-making. This requires a system to perform analytical analysis, fast-synthesis, complex queries, aggregated information, spatial comparisons, cross-tabulated, and interactive knowledge discovery.

From a technological point of view, an information system is required to manage large amounts of heterogeneous data from multiple sources, with different levels of details, various reference systems and different types (i.e. spatial, non-spatial, and temporal). This system has to handle analytically different types of data in multiple levels of details and perform complex queries within an interactive knowledge discovery. The system should also allow on-the-fly aggregation, analysis, and synthesis, then, report the resulting information to the users and decision-makers. Considering the technical specification for such a situation, a revision of the existing technological advances is necessary (see Figure 1.3).

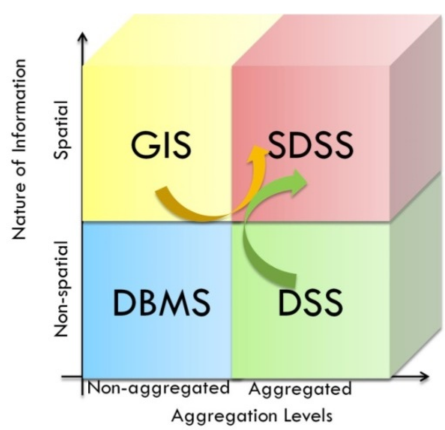


Figure 1.3: Scheme of existing information systems (adapted from (Bédard et al. 2007))

Database Management Systems (DBMS) are primary technologies hosting non-spatial data that does not support aggregation operations (Longley et al. 2005). On the other hand, Decision Support Systems (DSS) are mainly designed to deal with decision-making processes by handling non-spatial data at different levels of details (Van Kouwen et al. 2007; Roca et al. 2008). DSS also play also a major role in anticipating the outcomes of alternate management plans and fulfilling what-if analyses. DSS help the mangers and authorities to visualize long-term consequences of their decisions and evaluate possible solutions for the issues under study (Varghese et al. 2008). As mentioned, neither DBMS nor DSS handle spatial data. Indeed, 80% of data have a spatial component (Franklin et al. 1992). Geographical Information Systems (GIS) provide a wide range of capabilities to store, manipulate, classify, integrate, process, and represent spatiotemporal data (Longley et al. 2005). However, GIS marginally support transactional systems for multi-scales, multi-epochs, and multi-themes analysis. Moreover, on-the-fly multidimensional analysis can hardly be performed by today's GIS. Current advances in Decision Support Systems (DSS) and their integration with GIS provide interesting solutions for efficient integrated processes. Of particular interest are those advances coming from the field of Business Intelligence (BI), where a category of Spatial Decision Support Systems (SDSS) called Spatial Online Analytical Processing SOLAP technology (SOLAP) has been developed (Bédard et al. 1997). SOLAP is designed specifically to overcome the previously listed limitations of GIS and to allow rapid ad hoc multi-scale, multi-epoch, and multi-theme information retrieval and to perform complex querying by simply clicking on the desired level of information detail for given regions, epochs, and themes.

1.2 Problems Statement

Indeed, in adapting a GeoBI system for CERA, some issues and limits are important to be carried out which are characterized in our research projects as follows:

<u>Requirement for an Integrated Framework</u>

Coastal Erosion Risk has an inherently complex, multidimensional, and continuous nature which depends on several criteria, time periods, and scales of information in the presence of uncertainty at different levels. Risk assessment is not therefore as straightforward as one might imagine. Numerous frameworks and methods have already been developed for coastal risk assessment (IPCC 2007; ISDR & Nations 2004; Klein & Nicholls 1999; UNFCCC 1999; NOAA (National Oceanic & Atmospheric Administration) 2003; Mai & Liebermann 2002; Hinkel 2005; McFadden et al. 2007). These approaches are mostly scale dependent for one spatial scale and are unsuitable to be used for multi-scale analysis purposes. The main challenges in the stated methods are:

- Handling and analyzing large amounts of data,

- Performing fast synthesis, fast summarizing, easy comparisons and multi-level querying (on-the-fly analysis) in emergency cases for efficient decision-making processes.

Integration and aggregation of large amounts of data compel many of the stated methods and frameworks in any ICZM plans to focus only on single-oriented or few components involved in risk assessment, such as risk for people or infrastructure- not an integrated approach. Hence, a comprehensive tool for multi-scale representation of risk- not only in various spatial and temporal resolutions, but also in different levels of abstraction in thematic characteristics of risk for any aggregation or detailed information levels- is crucial for sustainable development of coastal regions.

Need to Deal with Large Amounts of Spatiotemporal Data in Multiple Scales

The key challenge in CERA is integrating and analyzing conflicting data in various time periods and spatial scales through dissimilar environmental, social, and economic criteria. Each organization carries out its own data acquisition based on its needs and standards. The resulting data are, therefore, generally heterogeneous and hence difficult to integrate, analyze and share. Available data are generally dissimilar in type, having been acquired based on different standards in different spatial and temporal scales. For instance, socio-economic data are available at the census level of detail whereas environmental data are at coastline-segment scale (Hegde & Reju 2007). In addition, the environmental data include both long-term (such as Sea Level Rise) as well as daily tide average while social data represent a snapshot of one census year (Boruff et al. 2005). Adapting geospatial BI paradigms for risk assessment is not straightforward. This includes characterizing all related risk components (i.e. hazard, vulnerability, and elements at risk) and translating them into multiple hierarchical dimensions.

Need to Deal with Expert-Based Knowledge and Existing Uncertainty

Available data and information for CERA are mainly expert-based and are classified with respect to decision-makers or stakeholders' interests. This can bring the problem of semantics of interested criteria or object for their classification. In other words, risk, by itself, is a definition which is given to demonstrate the impact of a physical or human-induced phenomenon causing disasters or harming human-life. Spatial modeling and representing risk through risk zones (as crisp object) bring up the issue of dealing with the fiat nature of risk which is based on some interested criteria introduced by stakeholders and experts. In this regard, the issue of information vagueness appears in definition of risk zones as well as geometries. Therefore, representing risk zones is a challenge due to the complex nature of risk and the existence of data uncertainty and information vagueness. This complicates the adequate communication of risk values and their consequences to stakeholders and authorities. The proposed information system should also be able to manage the existence of uncertainty and vagueness at different data levels.

Extending the concept of spatial object modeling from crisp to broad boundaries, two types of spatial objects are distinguished that are fiat (poorly-defined) and bona fide or crisp (well-defined) objects (Fisher et al. 2010; Smith & Varzi 2000). A fiat object has a broad boundary that cannot be observed and measured or is not known precisely (Fisher et al. 2010; Smith & Varzi 2000). For example, the spatial extent of a forest stand cannot be determined without any uncertainty. Bona fide object represent an object with a sharp-line boundary (Fisher et al. 2010; Smith & Varzi 2000). This is typically the case for the objects that are built by humans. Examples are buildings and roads. However, since every map and database is a model or representation of reality made for a specific purpose, they all carry some level of uncertainty (Bédard 1988). The difference between fiat and bona fide objects is more a matter of degree than a matter of absolute difference depending on the applications. Nevertheless, for the purpose of this thesis and in accordance with day-to-day practice, it is useful to note this difference.

Requirement for Fuzzy Representation of Risk Zones

Traditionally, coastal risk zones are represented by polygons that result from the aggregation of a series of spatial units defined with respect to stakeholders/authorities' interests or are based on national census parcels (Hegde & Reju 2007; Boruff et al. 2005; McFadden et al. 2007; Vafeidis et al. 2008). The spatial units may have different forms and sizes and may be distributed regularly or irregularly. The area and shape of these units may vary from a specific infrastructure to a very small cadastral parcel, municipality, state, or even a country. For each unit, a vulnerability index and a degree of risk is assigned (ISDR & Nations 2004; Abuodha & Woodroffe 2006; McFadden et al. 2007; Vafeidis et al. 2008; Füssel & Klein 2006). Polygons are separated by well-defined boundaries while the degree of risk is attributed homogenously within each polygon considering multiple criteria (Cheng et al. 2009; Bruce 2004; Burrough & Frank 1996). In fact, risk zones can be obtained from an overlay of vulnerability indicators' maps derived from elements at risk and hazard maps at different time periods. A typical risk representation is illustrated in Figure 1.4.

Therefore, risk zones representation is performed by simply overlaying several data layers and superimposing them into one layer yielding a risk map. This overlay operator is well implemented in existing GIS tools. However, the problem arises when different types of uncertain and vague information with different levels of details in different time periods should be overlaid and aggregated. In this case, any simple query becomes complex and its execution is a time-consuming task; whereas in risk assessment, fast synthesis and rapid analysis are essential.

Figure 1.4: An example of risk zone representation

Furthermore, the way that the geometry of polygons is defined differs among experts depending on their interests (Hegde & Reju 2007; McFadden et al. 2007; Cutter et al. 2003). It should also be noted that risk has multi-scale characteristics, due to inherent needs and interests of different participants on different organizational levels and their interests in one object or another (e.g. ports, buildings, census segment, city, state, or country). In this regard, risk zones are complex objects with uncertain boundaries and vague definitions in multi-scale space (Cheng et al. 2009; Molenaar 2000; Molenaar & Cheng 2000). The transition from one zone to another with a well-defined boundary provides therefore misleading insights of the degree of risk for decision-makers of any strategic actions.

Considering the fact that coastal erosion risk is a function of several continuous parameters (such as erosion rate, land cover change, and climate change), the transition from one spatial unit to another can loosely be explained with a crisp line and risk value changes gradually from a place to another (Cheng et al. 2009). For instance in Figure 1.4, the transition from red to yellow zone is sharp whereas in reality, risk level changes smoothly and continuously. This problem led us to explore and develop a more flexible method to represent continuous risk values.

Uncertainties exist and propagate from the collection, analysis, and representation of spatial data to interpretation and decision-making. This often appears in modeling environmental systems. By “uncertainty” we generally mean the vagueness in boundary zones, ambiguities in linguistic terms, discord in semantics of a class or object, fuzziness in process interpretation or existence degree of a spatial object, or a mix of them (Fisher et al. 2010). Semantics uncertainty comes from the fact that a person wants to describe a given phenomenon with different vocabularies or the same words in different ways (Fisher 2008). Vagueness, fuzziness, and ambiguity (discord, and non-specificity) are the principal factors in semantics uncertainty

(Fisher et al. 2010). Moreover, uncertainty increases from a natural sub-system to a human sub-system i.e. the key components of risk, with the large amounts of uncertainties concerning their interaction (Forth report of (IPCC 2007). This necessitates evaluating any imperfection or uncertainty in any level of risk assessment procedure from data and information to modeling techniques.

Indeed, handling and modeling uncertainties have recently become a major concern in risk assessment studies (Darbra et al. 2008; Kentel & Aral 2007; Skanata & Byrd 2007; Szemesova & Gera 2010; Xie et al. 2011). Probabilistic models based on Probability Theory and possibilistics models based on Possibilitics Theory have commonly been used to accommodate uncertainty associated with risk modeling and representation (Darbra et al. 2008; Bruce 2004; Molenaar & Cheng 2000; Choa et al. 2003; Aerts et al. 2003; Fisher et al. 2010; Cheng et al. 2005). Probabilistic models basically use ellipsoid error to deal with positional and measurement uncertainties (Pfoser & Tryfona 2001; Aerts et al. 2003). Exact, rough and fuzzy models are the example of possibilistics models. Exact models are the extension of the crisp spatial models such as Egg-Yolk (Randell et al. 1992; Cohn et al. 1997; Clementini & DiFelice 1997; Erwig & Schneider 1997; Cohn & Hazarika 2001; Cohn & Gotts 1996). Rough models are based on the Rough Set Theory to represent a spatial object as a pair of its maximum and minimum approximations (Worboys 1998; Fisher et al. 2010; Bejaoui et al. 2008). Fuzzy models are based on Fuzzy Set Theory that deal with inherent fuzziness and uncertainty related to modeling an object or identifying criteria through a membership function (Zadeh 1965; Robinson 2003; Fisher et al. 2010; Dilo et al. 2007; Pauly & Schneider 2010). Exact and rough models have commonly been used by scientists and practitioners due to their simplicity and straightforwardness in implementation (Pauly & Schneider 2010; Kanjilal et al. 2010). However, these models have their inherent limitations to represent continuous phenomenon. An example in this regard is coastal erosion. A detailed literature review reveals that the exact and rough models are not the ultimate solution to overcome the inherent uncertainty due to vague and fuzzy nature of spatial objects. This is more significant when the continuity of an object needs to be taken into account. Contrary to exact and rough models, fuzzy models promote stirring results. However, handling uncertainty using the stated methods needs more investigation to identify a way to adapt such methods.

Lack of a Formal Definition of Fuzzy Spatial Datacube

From the technological point of view, existing SDSS systems do not provide built-in capabilities to deal with semantic uncertainty. The need to handle imperfect data, either uncertain or imprecise, and run flexible queries on a spatial datacube is a challenge in performing CERA though current SDSS. SDSS such as SOLAP use multiple indicators that are translated into datacube dimensions to perform risk assessment. The choice of indicators is typically based on participants' interests while the level of details, classification and semantics of these indicators depend on available data from different sources as well. Semantics of objects, geometry, and temporality are often assumed to be exact or good enough in these spatial datacubes for the decision-makers' requirements. As a result, they depict major weaknesses when it comes to accurate analysis of phenomena (such as coastal risk) in the presence of uncertain and vague data and information. Hence, there is a strong willingness to propose a comprehensive solution for this problem. A promising approach that has been explored for almost two decades is fuzzy logic (Schneider 2010).

Several works have already been initiated to deal with the problem of uncertainty and vagueness issued from the definition of objects in On-line Analytical Processing (OLAP) and Spatial OLAP (SOLAP) systems (González et al. 2009; Laurent 2010; Péres et al. 2007; Molina et al. 2006; English et al. 2004; Pedersen & Jensen 2001; Kaya & Alhajj 2006). Nevertheless, the idea of a fuzzy spatial data model is still a young topic in the geospatial community. In addition, several efforts have already been made to embed fuzzy objects in SOLAP through extended crisp models i.e. exact and rough models (Bejaoui 2009; Siqueira & Ciferri 2012). In recent works, (Gervais et al. 2009; Levesque et al. 2007; Edoh-alove et al. 2013) propose two risk-aware approaches to face spatial fuzziness and uncertainties in which the user is informed about the existence of low-quality data. However, none of the cited works use Fuzzy Set Theory explicitly to characterize spatial uncertainty and to integrate it into SOLAP in a systemic way. Embedding spatial uncertainty in the multidimensional model requires formally redefining the principal elements of the spatial datacube.

Requirement for Multi-Scale Fuzzy Representation of Risk Zones

Risk is analyzed in a hierarchical manner due to the needs and interests of different participants at different levels of government (local, regional, national) for risk assessment and management. However, to our knowledge, there are no methods to aggregate and represent spatial measures resulting from uncertain data and vague information into multiple hierarchical dimensions. Thus, a multi-scale fuzzy representation of risk is necessary based on fuzzy objects.

By "aggregation" in the Geospatial BI community, we mean the grouping of data geometrically, thematically, or semantically to a coarser level of detail (Pedersen et al. 2001; Gomez et al. 2009). In other words, aggregation is a summarization process of fuzzy or crisp values or geometries in a datacube. However, the methods and procedures depend directly on the data model used (Péres et al. 2007; Laurent 2010; Gomez et al. 2009; Pedersen et al. 2001). Using fuzzy concept to define appropriate operators for data aggregation in a datacube has been initiated by Laurent (2010) and Molina et al. (2006). A series of operators (such as roll-up, drill-down, slice, dice, and pivot) have also been defined for fuzzy datacubes in Molina et al. (2006) and

Martin-bautista et al. (2013) using both quantitative and qualitative data. This permits a qualitative representation of results on charts and tables. However, the geometric aggregation of fuzzy objects requires redefining fuzzy operators (such as overlay and fusion) for fuzzy objects. More investigation is required in this regard to design a fuzzy spatial datacube.

1.3 Objective of Thesis

In order to provide an adequate solution for the stated problems, the general objective of this thesis consists of developing an approach combining GeoBI paradigm and fuzzy concept to consider the presence of the spatial uncertainty in the representation of risk zones. Ultimately, we assume this should improve coastal erosion risk assessment. To achieve the overall objective of this research project, three specific objectives are elaborated as follows:

1.3.1 Developing a GeoBI-Based Conceptual Framework Applied to CERA

The first specific objective focuses mainly on developing an analytical approach thanks to GeoBI paradigms to design a fuzzy Spatial Multidimensional Conceptual Model (SMCM) for efficient assessment of coastal erosion risk. As well, dealing with information vagueness coming from data modeling is carried out. In this regard, identifying key elements of a fuzzy spatial datacube are required. Conceptualizing such model is the main goal in this step with a reference to the main components of coastal erosion risk (hazard, vulnerability and elements at risk).

1.3.2 Developing a Novel Approach Based on Fuzzy Set Theory to Improve the Representation of Uncertainty Inherent to Risk Zones Representation

Based on the stated arguments, the main concern in this research project is the indeterminate nature of coastal erosion risk and associated uncertainty, characterized as information vagueness. More precisely, one of the focuses of this research project is on the inherent uncertainty and vagueness mostly arisen from data modeling. In CERA, the first step is the tessellation of the coastal region to determine analysis units. Traditionally, this step is an expert-knowledge based (qualitative) procedure. The tessellation strongly depends on stakeholders/authorities' interests. That introduces the first level of uncertainty. A vulnerability index based on expert-knowledge rules is then assigned in the second step to each unit introducing additional uncertainty. Indeed, the associated uncertainties of a vulnerability index stored as attributes of each unit leads to geometric vagueness. The aggregation of spatial units represents risk zones in the third step. The associated uncertainties are propagated through the first, second, and third steps in CERA and finally appear in the geometry of risk zones. On the other hand as stated above, the continuous characteristics of risk indicate that the risk zones are linked to each other with gradual transition zones (Fisher et al. 2007; Cheng et al. 2009).

The second specific objective focuses on characterizing the nature and level of spatial uncertainty appeared as information vagueness and handling it through coastal risk assessment process using Fuzzy Set Theory. A systematic approach to link key elements of CERA and fuzzy membership function concept needs to be taken into account. Extracting IF-THEN rules from the traditional classification of experts and practitioners to elaborate vulnerability index and elaborating risk formula is a major challenge in this step. Once, the IF-THEN rules and risk formula are elaborated, the multiple layers of fuzzy objects are aggregated to represent risk zones.

1.3.3 Formalizing Fuzzy Spatial Data Aggregation in Fuzzy Spatial Datacubes

The third specific objective focuses formalizing a fuzzy-based approach spatial data aggregation to deal with information vagueness in spatial datacubes. The developed approach is tested within an application of multi-scale risk assessment.

1.4 Research Method

To achieve the stated objectives and to obtain the expected results, the following steps are taken during the course of this thesis. The UML activity diagram describing the details of major phases, sub-phases, and their connection is illustrated in Figure 1.5.

1.4.1 Defining Research Project

A deep investigation of relevant works is begun by scrutinizing the fundamental concepts for a better understanding of the research context and elaborating the main objectives of this research project. The main concepts to be covered are:

1) Geospatial Business Intelligence paradigm;

2) Characteristics of risk and its components and assessment;

3) Risk assessment under uncertainty;

4) Spatial data modeling and representation of an object with vague shape and formal definition of fuzzy object, and finally

5) Fuzzy multi-scale representation of risk and formal definition of required aggregation operators.

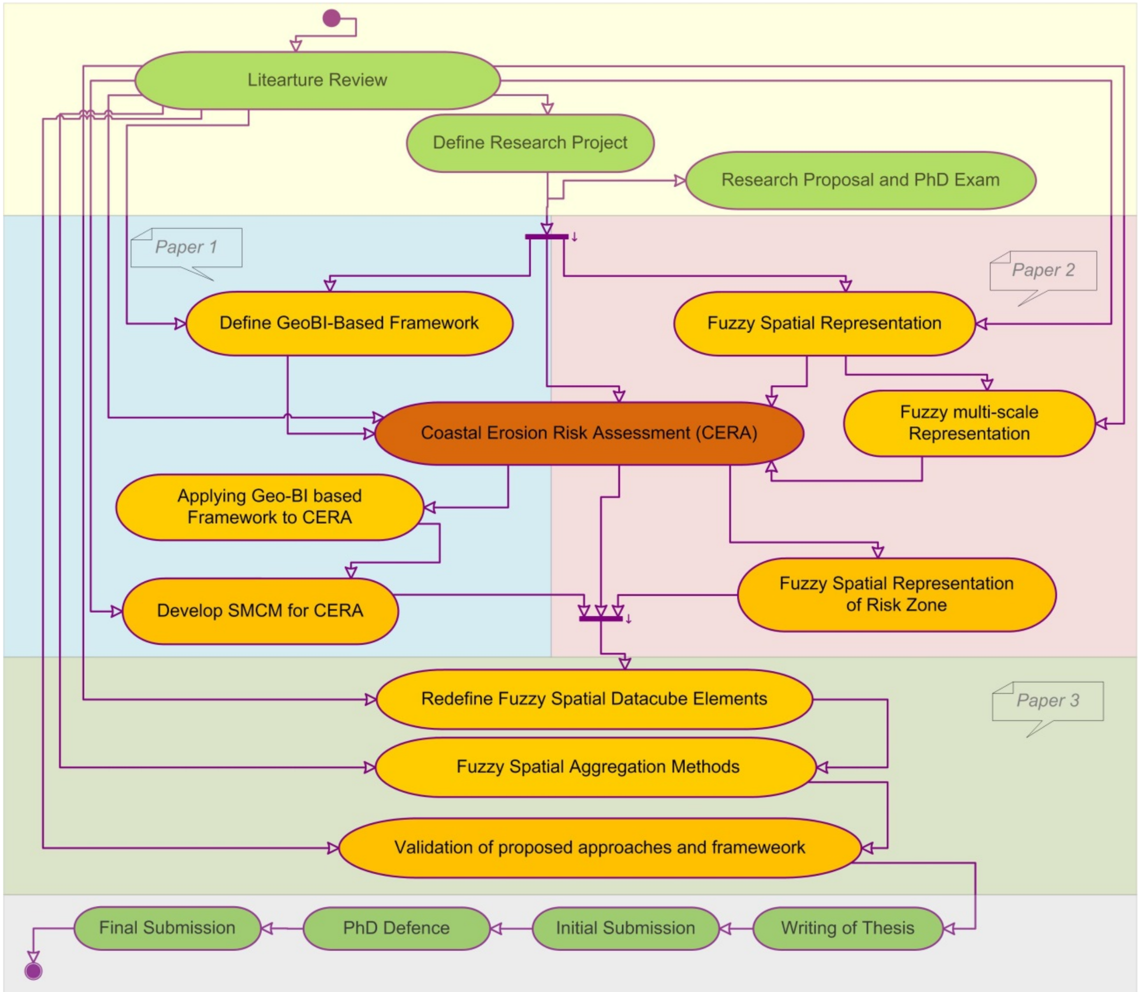


Figure 1.5: UML activity diagram of research methods.

1.4.2 Proposing a Spatial Multidimensional Conceptual Model (SMCM) for Coastal Erosion Risk Assessment

An analytical conceptual framework based on a geospatial BI paradigm is proposed in this phase to elaborate a Spatial Multidimensional Conceptual Model (SMCM) to perform CERA. The model is an example for a comprehensive solution integrating multiple environmental, social, and economic criteria as well as their interactions at various spatial scales and epochs on different themes. It also takes into account multiple elements at risk such as people, infrastructure, and built environments as different dimensions of the analysis. Bear in mind that the main steps followed are listed hereafter and detailed in the upcoming sub-sections:

1) Performing need analysis based on application under study to identify the stakeholders and their needs. This step is done based on literature and consulting the expert of applications such as

(McHugh et al. 2006; Bernatchez, Fraser, Friesnger, et al. 2008; Bernatchez, Fraser & Lefaivre 2008; Xhardé R. 2007);

2) Performing data inventory based on application under study to have an idea about data availability, data types, level of details, metadata, etc. The main sources of data in this step were ETE-INRS, Quebec Province Transport, GeoBase, and erosion projects done by the Industrial Chair of Geospatial Database for Decision Support at Université Laval;

3) Identifying coastal erosion risk key components:

    a. Identifying hazard:

        i. Choosing a method to determine the border of coast and land;

        ii. Calculating erosion rates from coastline changes for different periods.

    b. Identifying elements at risk;

    c. Elaborating an integrated vulnerability index by considering both environmental and socio-economics features;

    d. Elaborating risk equations;

4) Translating the key components of coastal erosion risk into key elements of a spatial datacube;

    a. Determining datacube dimensions (spatial, temporal, and thematic) derived from key components of coastal erosion risk;

    b. Determining datacube measures (spatial and non-spatial) based on risk equations;

    c. Choosing the schema of data structure in spatial datacube for CERA, for instance star schema;

    d. Designing spatial multidimensional conceptual model.

1.4.3 Proposing a Novel Approach to Improve Spatial Risk Representation Based on Fuzzy Set Theory

A fuzzy-based approach is developed in this phase to deal with the uncertainty related to the object and its geometry, or its vagueness. A conceptual framework is also proposed to deal with the problem of spatial uncertainty of risk zones using membership functions. Instead of determining the crisp and sharp boundaries between risk zones, the proposed approach permits a smooth transition from one zone to another. Membership functions are derived from the knowledge of experts. This phase is divided into two main steps, where results from the previous phase can be reused to achieve the determined objectives:

1) Tessellation of the region under study:

    a. Identifying hazard (results reused from the previous phase);

    b. Elaborating vulnerability index (results reused from the previous phase);

    c. Classifying based on erosion rate into five categories (very high, high, medium, low and very low rate of erosion) ;

    d. Loading data related to vulnerability index;

    e. Detecting hotspots of vulnerable indicators and erosion rates;

    f. Defining the size and shape of grid i.e. regular or irregular;

    g. Drawing the grid on the region under study.

2) Fuzzy representation:

    a. Performing fuzzification;

        i. Defining fuzzy membership functions based on the vulnerability index for each layer of indicator;

        ii. Assigning membership value to each cell;

        iii. Calculating the risk value in each cell using respective membership value;

        iv. Representing the risk value for each indicator.

b. Performing fuzzy aggregation operators;

    i. Elaborating risk Formula;

    ii. Extracting IF-THEN rules;

    iii. Aggregating multiple indicators based on elaborated risk formula and rules using the aggregation operators (union, intersection, difference, etc.);

    iv. Representing the risk zones.

1.4.4 Formalizing Fuzzy Spatial Data Aggregation in Fuzzy Spatial Datacubes

In the fourth phase, the fuzzy-based approach developed in the previous phase is embedded into a spatial datacube through redefining the spatial datacube elements (dimensions, members, hierarchy, and facts) and required fuzzy aggregation operators (union, intersection, difference, overlay, and fusion). It is then applied to a spatial datacube with a reference to the CERA multidimensional model. The appropriate methods are proposed to aggregate and represent risks' measures into multiple hierarchical dimensions spatially or numerically. Possible fuzzy spatial aggregation scenarios within a spatial datacube are also presented. This phase is divided into three main steps:

1) Redefining elements of a fuzzy spatial datacube;

2) Proposing fuzzy spatial aggregation methods, and

3) Presenting possible scenarios for spatial aggregation that can be performed in the proposed fuzzy spatial datacube.

1.4.5 Validating the Research Results

Testing the validity of the research results obtained from previous phases is carried out in a case study on the East Coast of Quebec, Canada. This step is partially performed in precedent phases. The validation phase is composed of:

1) Conceptualizing:

    a. Finding a critical coastal zone with respect to erosion rate in recent years;

b. Performing need analysis based on technical reports of different ministers and organizations;

c. Performing data inventory; and

d. Conceptualizing a spatial datacube adapted for risk assessment;

2) Developing and implementing a Matlab code to accommodate the proposed fuzzy spatial objects adapted to risk assessment methods;

3) Comparing the results obtained in this research project to those of existing approaches. This step provides the similarities and differences between the obtained results and those existing before;

4) Proposing the perspective for future research opportunities based on lessons learned and limits of our contributed approaches.

1.5 Thesis Structure

The results of this 5-years research to fulfill the requirements of this doctoral research work are presented in six chapters. However, this thesis is paper-based. Hence, three Chapters **3**, **4**, and **5** are the versions of submitted and published in three peer-reviewed scientific journals. Thus, there are some redundancies between these chapters and **Chapter 2**.

In **Chapter 1**, the research context including the problem statement, the research objectives (general and specific), and the research method to realize the proposed research work have been presented.

In **Chapter 2**, the relevant works identified in the literature review regarding the fundamental key concepts of this research project are discussed.

In **Chapter 3**, the characteristics of coastal erosion risk assessment and in particular, and the multidimensionality aspect of the risk itself are explained. A conceptual framework to model the spatial and non-spatial data required for risk assessment in a SDSS as SOLAP is also proposed. The main contribution of this chapter is a comprehensive multidimensional data model serving as an innovative example to improve the assessment of coastal erosion risk using the SOLAP approach. As a result, a spatial multidimensional conceptual model is presented whilst a deep investigation of the hierarchy model for each analysis axis is discussed. Indeed, the main achievement of this chapter is part of the first specific objective of this research project.

In **Chapter 4**, the spatial fuzzy representation of phenomenon in order to deal with inherent information vagueness due to data modeling as well as semantic definition is presented. Fuzzy Set Theory is then chosen due to its flexibility with human reasoning and perception to deal with this problem. The main contribution of this chapter is to characterize the source and nature of uncertainty appearing as information vagueness in an integrated risk assessment procedure and then handling it in an optimum way. Indeed, this chapter covers the second specific objective of this research project.

In **Chapter 5**, the integration of the fuzzy-based approach within a spatial datacube is explained and the possibility of incorporating fuzzy concept in defining a spatial datacube's elements is discussed. The main contribution of this chapter is therefore to formalize a fuzzy spatial datacube by redefining its key elements and the fuzzy spatial aggregation methods. This allows for dealing with the multidimensionality and hierarchy characteristics of risk assessment. This chapter covers part of first objective and the third specific objective of the current search project.

Finally, **Chapter 6** concludes this thesis and highlights some perspectives for further and future research.

“Somewhere, something incredible is waiting to be known.”
Carl Sagan

Chapter 2 Literature Review

2.1 Introduction

In this chapter, fundamental concepts related to the current research work are reviewed. These include geospatial business intelligence technologies for decision-making, spatial datacubes, spatial data models and spatial data integration methods in a spatial datacube, spatial uncertainty and the methods to deal with it, such as fuzzy set theory. For application purposes, a deep review is done on the assessment and representation of risk with a focus on its spatial characteristics. These concepts are essential to understand the main objectives of this thesis that lead to the fulfillment the expected results. The recent advances in geospatial technologies to handle data modeling and integration are discussed in section 2.2. In Section 2.3, risk assessment and representation with a focus on risk characteristics is introduced. Uncertainty, its nature, and its impact level in the context of decision-making and risk assessment are described in Section 2.4. Also, how data uncertainty and information are handled in a spatial datacube is discussed in this section. The fundamental of spatial data model is discussed in section 2.5 with a focus on fuzzy set theory dealing with uncertainty. This chapter is concluded with a resume of the materials provided in Section 2.6. Since, this thesis manuscript is paper-based; accordingly, there are some redundancies between this chapter and Chapters 3, 4, and 5.

2.2 Advance in Geospatial Technologies for Decision-Making

Geospatial technologies and more specifically GIS provide a wide range of functionality for collection, storage, management, integration, analysis, visualization, and diffusion of spatial and non-spatial data for decision-making. Despite these functionalities, GIS has some limitations for efficient decision-making. For instance, the complex multidimensional characteristics of natural phenomenon and the need to carry out on-the-fly multidimensional analysis are the main challenges of the current GIS.

From another perspective, an essential characteristic of an information system as a decision support system is interconnecting the data from natural processes and human activities into a fast and intuitive solution, by avoiding any prior knowledge about the whole system. In this perspective, the recent evolutions of DSS and the compatibility with GIS leading to SDSS are regarded as a potential solution. In particular, SDSS technology relying on spatial datacubes, which evolved from the field of Business Intelligence (BI), has been rapidly gaining popularity over the last couple of years in several fields.

Most of recent SDSS deal adequately with the majority of conceivable situations by allowing fast synthesis, easy comparisons, and multi-level querying of information for decision-making. It allows decision-makers to focus only on specific criteria and their interactions, intelligently control the overall process, and employ all available data, information, and knowledge into a hierarchical system. It also provides a user-friendly environment for the decision-makers to easily and rapidly navigate within a geospatial database toward the elaboration of an optimal sustainable management and prevention/protection plan (Rivest et al. 2001). Consequently, possible scenarios become rapidly clear and can be presented to decision-makers by creating new relations based on emerged options. Using SDSS technology such as SOLAP and Spatial dashboard, decision-makers are no longer involved in creating complex Structure Query Language (SQL) queries. Non-expert users are able to carry out their analysis with a few mouse clicks and display the results in detail in the form of maps, detailed or aggregated histograms, and detailed or summarized tables (Bédard et al. 2007).

Transactional and analytical databases are two main categories of spatial databases (Salehi et al. 2010). Regardless of its category as either transactional or analytical , the integration of data from different sources is the key functionality of any database (Bejaoui 2009). Transactional databases are most often used due to their facility for storing, integrity checking, manipulation, and visualization of a spatial phenomenon. Analytical spatial databases support fast and efficient decision-making within an interactive spatiotemporal analysis interface. These types of databases are mainly used in application, such as Business Intelligence. Data Warehouses and Data Marts are the typical analytical databases (Kimball & Ross 2002). A data warehouse is a subject-oriented, integrated, time-variant, and non-volatile database that aims to support decision-making processes (Inmon 1992). The multidimensional structures such as a datacube or a hypercube are the main analytical data structures that are employed in analytical databases (Chaudhuri et al. 2011).

The OLAP (On-Line Analytical Processing) tool that has been used for Decision Support Systems (DSS) since the mid-1990s provides a platform to host a datacube or hypercube. However, more than 80% of available data in decision-making processes have spatial components (Franklin et al. 1992). Since the spatial data is involved, the term spatial datacube is proposed (Bédard et al. 1997). This requires the compatibility of a datacube to accommodate geospatial data that results in the first Spatial OLAP (SOLAP) (Bédard et al. 1997; Rivest et al. 2001). SOLAP research was initiated in the mid-1990s in parallel at Laval University (Bédard et al. 1997), Simon Fraser University (Stefanovic 1997; Han et al. 1998), and Minnesota University (Shekhar et al. 2001). Nowadays, research on SOLAP is being carried out around the world. Companies offer commercial products based on SOLAP principals and the Open Geospatial Consortium (OGC) has started to develop standards for Geospatial Business Intelligence.

2.2.1 Spatial On-Line Analytical Processing (SOLAP)

SOLAP, a SDSS, is defined by Rivest et al. (2001) as "a software that allows rapid and easy navigation within spatial databases and that offers many levels of information granularity, many themes, many epochs and many display modes synchronized or not: maps, tables and diagrams". SOLAP provides basic knowledge discovery functions not found in GIS and which make data accessible for non-specialists, including new trials of analysis and queries without any SQL-commands and algorithms. This aspect allows the decision-makers to concentrate on their analysis needs and results interpretation via a friendly interactive graphical interface. This objective can easily be achieved without extensive knowledge of the software being used and without definition of complex queries. Then, analysis results are provided in the form of maps (single maps, multi-maps, and complex thematic maps), statistical diagrams (bar, chart, and pie), and histograms. SOLAP also permits the decision-makers from different organizations to consult and employ the same system (database and server) to navigate through the spatial and descriptive information. The navigation can be performed on different levels of granularity via basic operations such as spatial roll-up or drill-down to go to coarser/finer granularity within a theme, spatial drill-across to show other information at the same level of granularity, and spatial slice and dice. For more details about SOLAP advantages and applications, see (Rivest et al. 2005). The popularity of SOLAP is growing rapidly for applications in different levels of organizations and authorities from public to private sectors (Bédard et al. 2009).

2.2.1.1 Key Elements of SOLAP

SOLAP system architecture consists of spatial databases, a server (relational, multidimensional, or hybrid) and SOLAP clients (see Figure 2.1). In certain cases, there may be a data warehouse between the sources and the SOLAP datacubes. In all cases, there are Extract-Transfer-Load (ETL) processes involved to integrate and aggregate the data. These processes can take place in the server or in dedicated ETL tools, GIS, DBMS, or home-made code. In a SOLAP system, spatial data are structured using dimensions, members, measures, and facts, shown in Figure 2.2, that together constitute a spatial datacube (Bédard et al. 2007).

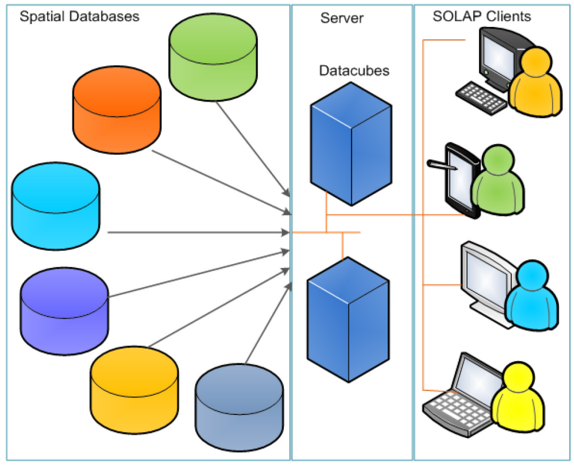


Figure 2.1: SOLAP architecture

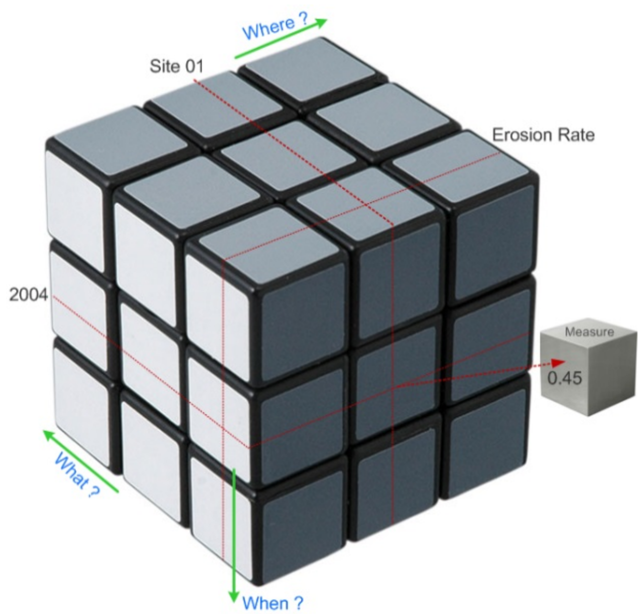


Figure 2.2: Key elements of SOLAP

Dimension is defined as an analysis perspective or theme of interest for a user e.g. products, retailers, regions, and periods (Salehi et al. 2010). A dimension can be spatial to express the spatial component. Spatial dimensions are of three types: non-geometric when only place names are used without maps, geometric when maps are used, and mixed (see Figure 2.3). Dimensions can rely on discrete feature-based or continuous raster-based images (Bédard et al. 2009). A dimension can also be temporal e.g. day, week, month, year or thematic to know "what" is being considered e.g. population or land uses at multiple levels of granularity (Bédard et al. 2009). A dimension includes one or several hierarchies composed of different analysis levels e.g. city, state, country labeled as "administrative region" (Bédard et al. 2009). A member is an instance of hierarchy level that states a position within the hierarchical data structure of a dimension e.g. Canada is a member of country level (Malinowski & Zimànyi 2008). Navigating in the spatial datacube simply translates into selecting the desired members anywhere in the hierarchical structure of the dimensions.

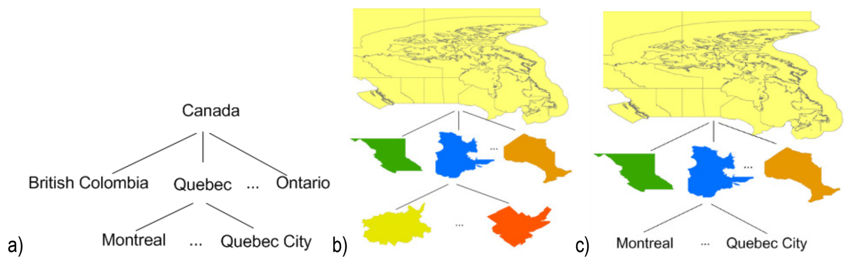
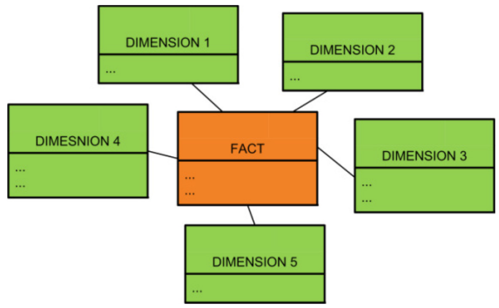


Figure 2.3: Spatial Dimension: a) non-geometric, b) geometric, and c) mixed

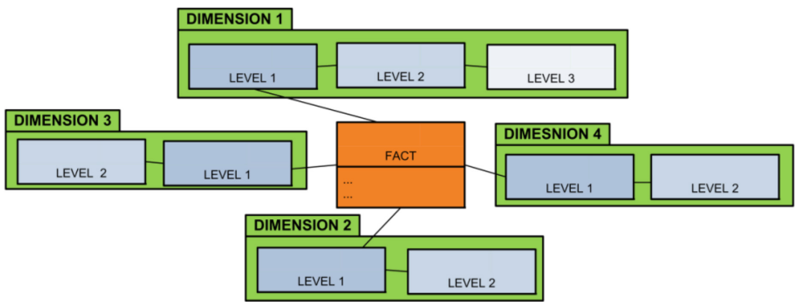
Measures are measurable quantities e.g. the number of victims in an accident with respect to the different levels of "administrative region" and "time" dimensions. These are analyzed against members of different levels of dimensions (Salehi et al. 2010). Values resulting from unique combinations between members of different dimension levels, along with their measures, are called fact e.g. the number of victims in car accidents in Quebec City between May and August of 2010 (Salehi et al. 2010). In other words, measures are dependent variables to dimensions, whilst dimensions are essentially independent variables in nature (Chaudhuri et al. 2011). The latter efficiently supports the multidimensional characteristics of risk analysis where the risk parameters are independent variables in the present context. This makes the whole system more consistent and coherent from the user's perspective.

2.2.1.2 SOLAP Structure

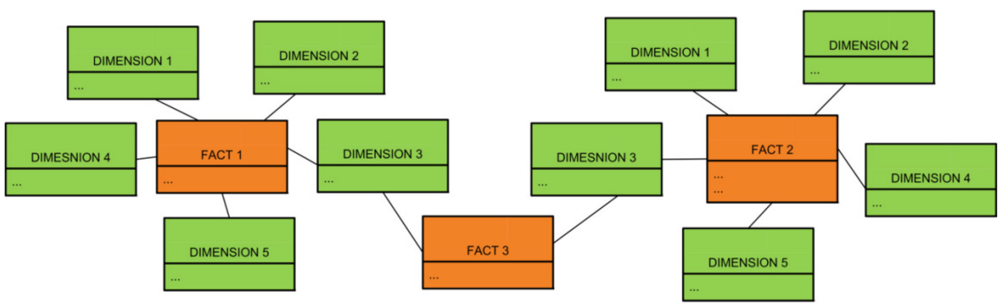
Common methods for structuring the data in SOLAP are star schema, snowflake schema, and mixed or constellation schema (Pedersen & Jensen 2001). A star schema consists of a single table for each dimension and a single fact table containing key attributes of dimensions and measures (Gray et al. 1997). The hierarchy levels are implicitly supported but dimensions are easily navigated by star schema. The snowflake model is obtained through normalizing a single dimension table into a single table for each level of hierarchy of its dimension (Bédard et al. 2009). The hierarchy levels of dimensions are explicitly browsed by normalization of dimensions in snowflake schema. However, this model is less efficient compared to star schema since it requires more joint operations. Fact constellation schema is the most complex structure in which multiple fact tables share dimension tables (Bédard et al. 2009). The possible structures are illustrated in Figure 2.4.



a)



b)



c)

Figure 2.4: SOLAP Structure defined using three different methods: a) star schema, b) snowflake schema, and c) mixed schema

Spatial data integration is the key procedure in establishing any spatial datacubes. The next section focuses on this procedure and the possibility of taking into account data uncertainty accommodation in datacubes and in data integration process.

2.2.2 Spatial Data Integration

Spatial data integration is the process of combining data from different sources, a prerequisite of multidimensional databases such as data warehouses (Bakillah 2012). In other words, it consists of establishing the relationships between corresponding instances of an object in multiple data sources representing the same geographic entities (Uitermark et al. 2005). This requires selecting or generating the geometry of an object from multiple sources. An example of the integration of several spatial objects coming from different sources is illustrated in Figure 2.5.

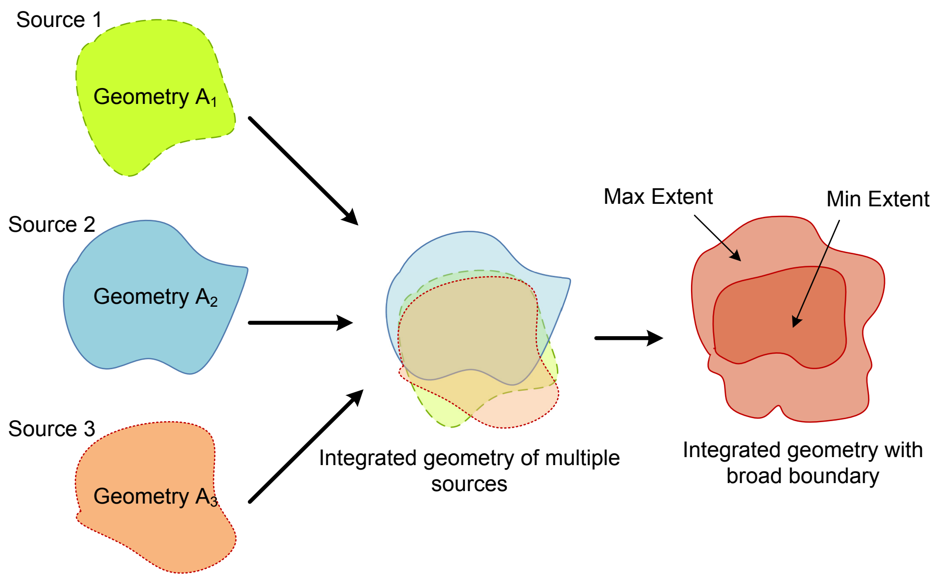


Figure 2.5: Representation of the integrated geometry of an object from different sources (adapted from (Bejaoui 2009)**)**

In a spatial data integration process, data needs to be structurally and semantically adapted in the multidimensional databases (Bejaoui 2009; Bakillah 2012; Sboui 2010; Salehi et al. 2010). That requires characteristics of data sources such as data structure, geometry and semantics to be heterogenic (Bakillah 2012). Principally, spatial data integration may produce a multi-representation of spatial data, change the semantic of data, and improve the completeness and non-redundancy of available data in the database (Bakillah 2012; Sboui 2010; Bejaoui 2009). Data integration can be performed vertically i.e. integrating spatial data describing multiple themes in the same location or horizontally i.e. integrating spatial data describing the same theme but in different locations (Bejaoui 2009). Vertical data integration is used in spatial data warehouses in which the final geometry is represented with a degree of vagueness for the members of hierarchy level.

In geometry integration, the classical method assumes the equal contribution of all crisp sources' geometry to represent final geometry. Therefore, different levels of uncertainties can be introduced and propagated within data integration. The Min-Max extended method based on rough set theory, as illustrated in Figure 2.5, is commonly used to take the uncertainty into account (Bejaoui et al. 2008). The topological relations of the final geometry is, therefore, established between well-defined shapes in which the integrity constraints (IC) control the quality of topological relations through a set of rules (Salehi et al. 2010). (Bejaoui 2009) addresses the problem of topological relationships vagueness within the Qualified Min-Max (QMM) topological relationship model.

2.2.3 Spatial Data Aggregation

Aggregation is the grouping of spatial data at a coarser level of detail or resolution than the level of detail at which the data were collected or represented. Aggregation process is a key operator to any hierarchy system for any ad-hoc solution to support decision-making process. This is of major importance especially where a general rather than detailed analysis is required. This is the case for almost any data warehouse or enterprise data to drill down or roll up into level of details of a specific feature or dimension. The result of aggregation leads to loss of spatial or attribute details through the creation of coarser spatial data. The main drawback is that these details may be requisite by some stakeholder or decision-makers to easily access and rapidly manipulate datasets and report summarized information.

Spatial multidimensional methods facilitate geometrical and thematic aggregation through a variety of functions and techniques. This covers both the geometric extents and associated descriptive attributes. However, the techniques are directly associated to the data model used.

Dissolve and merge are the common aggregation operators performed on the thematic aspect of the same layer of vector data with respect to *Average*, *Sum*, and *Weighted-Average* functions. On the contrary, spatial join is a geometrical aggregation that performs on the spatial components of vector data from different layers with respect to *Sum*, *Min*, *Max*, *Mean*, or *Median* functions. The aggregation of raster data consists of decreasing resolution by increasing the pixel size through multiplying a factor into cell size. SQL Server provides four types of aggregation that are:

- *Union* that combines multiple spatial objects into a single spatial object;
- *Envelope* that covers the multiple objects by a rectangle or circle considering its geometry type or component's location;
- *Collection* that returns a geometry collection with one part for each spatial object in the selection set, and
- Convex Hull that returns a convex hull polygon enclosing one or more spatial objects. This operation can be performed either on vector or raster data.

Overlay is the principal operation to merge the geometries. It intersects different geometry of data sources using a tolerance error value called tolerance match, around the refereed node (Frank 1987). The refereed node is the more qualified data source. If the source geometry is not situated inside the tolerance match zone, the mentioned geometry's source will be excluded from the final geometry. The quality of final geometry depends directly on the quality of data sources, refereed node, and the existence of non-empty differences between union and intersection operation in the multiple sources of geometries.

As stated before, using multidimensional databases for effective decision-making is in constant increase. The main challenge in this regard remains integrating data uncertainty and information vagueness into spatial datacubes and decision-making processes. Some works were the pioneers to use fuzzy set theory in datacubes (Feng & Dillon 1999; Pedersen et al. 1999; Pedersen & Jensen 2001; Laurent 2010). A series of operators such as roll-up, drill-down, slice, dice, and pivot has been defined for fuzzy datacubes in Molina et al. (2006) and Martin-bautista et al. (2013) using both quantitative and qualitative data. This permits a qualitative representation of results on maps, charts, or tables. The main advantage is being able to include human reasoning in aggregation processes. The thematic aggregation can principally be performed based on traditional methods whereas the geometric aggregation requires handling the spatial aspects and semantics of geometry. This requires defining fuzzy operators such as union, intersection, difference, overlay, and fusion developed for crisp objects that have been explained in previous sections.

As mentioned earlier, the achievement of this research project is applied in the case of coastal erosion risk assessment. The reason to choose such application refers to multi-source multidimensional, multi-theme, multi-epoch, and multi-scale characteristics of coastal erosion risk. Also, the fuzzy nature of coastal erosion along coastal regions makes this choice more appropriate to test developed approaches along this research project. Therefore, a detailed description of risk assessment processes is provided.

2.3 Risk, Risk Assessment, and Risk Representation

Risk is the forecast for an accident or loss (Cutter et al. 2003; Alexander 2000). In other words, risk is the probability of harmful consequences or expected losses arising from spatiotemporal interactions between natural or human-induced hazards and vulnerable condition of elements at risk exposed in different time scales and territorial dimensions (Cutter et al. 2003; Daudé et al. 2009; Alexander 2000; Abuodha & Woodroffe 2008). Risk assessment in this regard includes undertaking the impact of any phenomena that causes economic, social, environmental, and human-life losses within a systematic and comprehensive process (Hessami & Karcanias 2009). More precisely, risk assessment is defined as developing methods to determine the nature and extent of risk by analyzing potential hazards and evaluating existing vulnerable condition of elements at risk such as people, infrastructure, houses, etc. (Cutter et al. 2000; Cutter 2002; Karvetski et al. 2011; Boruff et al. 2005). The risk assessment process and the relationship between its components are illustrated in Figure 2.6. Risk is then computed as the mathematical product of the probability of hazard and the element at risk exposed by the range of vulnerability of these elements as stated in Eq.2.1 (Varnes 1984; Blaikie et al. 2004; Uricchio et al. 2004; Alexander 2000).

$$Risk = Hazard \times Element\ at\ Risk\ \times Vulnerability \quad \text{Eq. 2.1}$$

In natural risk assessment, natural hazards are considered as natural phenomena that cannot be avoided (Alexander 2000). The ultimate objective of risk assessment is forecasting and timely prevention against all undesirable situations such as a disaster or perturbation in human-life trends or their assets (Hessami 2004). Understanding risk provides fundamental information for decision-makers and allows them to select appropriate alternatives or strategies in terms of response options for retreating, adapting, and protecting of the regions (Karvetski et al. 2011; Fenton & Wang 2006; Frontiers 2005; Hill et al. 2005; Li et al. 2012; Ali & Rakus-Andersson 2009).

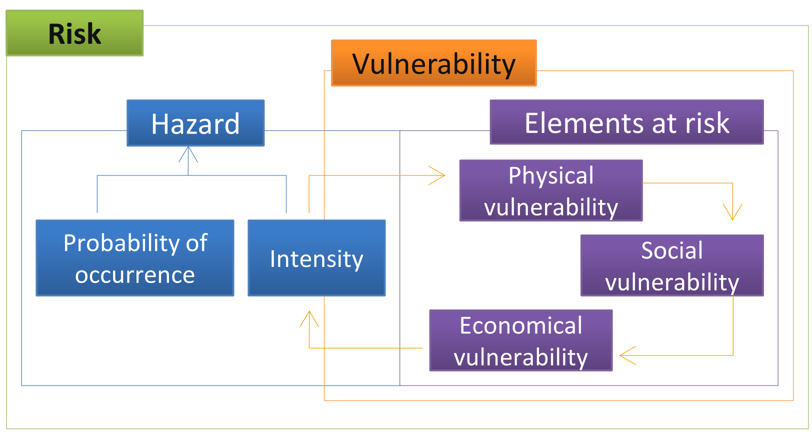


Figure 2.6: Risk scheme and the related components and their relation.

2.3.1 Risk Characteristics: Hazard, Elements at Risk, and Vulnerability

Main components of risk are hazard and vulnerability (see Figure 2.6). These terms are confusingly used interchangeably in the literature. In environmental sciences, we mainly deal with hazard to estimate the degree of risk while in social sciences, the focus is principally on vulnerability studies for risk assessment (Abuodha & Woodroffe 2006; Cutter et al. 2003; Boruff et al. 2005). However, risk assessment should include all physical, social, and economic aspects or an appropriate combination of them considering the context of the study. The physical aspect of risk concerning hazard depends on the location, magnitude, frequency, duration, and process of the hazard under study (Cutter et al. 2003); whereas the socioeconomic aspects of risk are related to vulnerable features. Vulnerable indicators include people at risk, damaged structures, human activities and their assets (Cutter et al. 2003). The following section demonstrates the key characteristics of risk, its components, and the way to assess risk.

Hazard: Hazard is defined as an object, state, or condition with a potential to lead to an accident, event, or process (Cutter et al. 2003; Alexander 2000; Cutter et al. 2008; Adger 2006; ISDR 2004). In other words, hazard is the probability of occurrence of an event or process and its intensity (see Figure 2.6). In particular, natural hazards are processes or phenomena occurring in a biosphere that may constitute a disaster which can be classified according to their geological or biological origins (ISDR 2004). The most common methods to identify natural hazard are statistical or probabilistic approaches. These methods have already been integrated in spatial data analysis (Goodchild & Glennon 2008). Nevertheless, statistical and probabilistic methods require large amounts of data in the past to establish a reliable pattern of occurrence. The sufficient available historic data is the main drawback in many regions. An alternative as a result would be dynamic approaches based on simulation techniques (Uricchio et al. 2004).

Elements at Risk: Elements at risk are the features that are potentially affected by an undesirable event. Examples are people, infrastructure, buildings, tourist industry, and many other socio-economic activities. Elements at risk as shown in Figure 2.6, link the two key components of risk i.e. hazard and vulnerability (Alexander 2000; Daudé et al. 2009). The identification of elements at risk is performed by stakeholders or authorities with recognized interests of society or communities at a particular time. These elements are key features to identify vulnerability indicators which are explained further, and elaborate vulnerability index. This identification is often a challenging task due to the domino effect; one risk may cause subsequent risks which led to multiple risks analysis. The practitioners prefer to perform independent risk assessment for each element at risk to avoid this complexity. However in reality, the risks are dependent and related to each other. A solution is using simulation-based methods (Daudé et al. 2009).

Vulnerability: Vulnerability is defined as susceptibility to injury, fatality or loss (Cutter et al. 2003; Boruff et al. 2005; Alexander 2000; ISDR 2004). Vulnerability functions are developed to link the susceptibility of elements at risk to the magnitude of the respective hazardous processes within risk assessment equation (Füssel & Klein 2006; Abuodha & Woodroffe 2006). This leads to quantitatively assessing individual and collective risks. As illustrated in Figure 2.6, vulnerability is classified into social, economic, and physical indicators. The economic vulnerability represents the risk to production, distribution, and consumption. Social vulnerability emphasizes coping with the capacity of people or society to be resilient from a disaster (Cutter et al. 2008;

Cutter et al. 2003). Physical vulnerability is the inability of an ecosystem to tolerate stressors over time and space (ISDR 2004).

Vulnerability encompasses the idea of response and coping, since it is determined by the potential of a community to react and withstand a disaster. Numerous frameworks, conceptual models, and vulnerability assessment techniques have been developed to advance both the theoretical underpinnings and practical applications of vulnerability (Adger 2006; Füssel & Klein 2006; Füssel 2010b; Füssel 2010a; Cutter et al. 2008). The assessment of vulnerability requires an ability to both identify and understand the susceptibility of elements at risk and in a broader sense, of the society to these hazards (Füssel & Klein 2006). Multi-criteria methods i.e. parametric models, indices-based analyzing, and expert-knowledge-based are among the most popular methods to perform vulnerability analysis (Adger 2006; Füssel & Klein 2006; Füssel 2010b; Füssel 2010a; Cutter et al. 2008; Nicholls et al. 2008; Abuodha & Woodroffe 2006). The principal components of vulnerability are categorized as parameters or indicators and are weighted according to their importance (ω_i). These components vary in space and over time (Adger 2006; Füssel & Klein 2006; Füssel 2010b; Füssel 2010a; Cutter et al. 2008; Nicholls et al. 2008; Abuodha & Woodroffe 2006; Uricchio et al. 2004; Xhardé R. 2007; Vafeidis et al. 2004; Vafeidis et al. 2008; Gornitz et al. 1997). The vulnerability is then calculated from:

$$Vul = \sum_{i=1}^{n} F(x_i) \times \omega_i \qquad \text{Eq.2.2}$$

where $\mathrm{F}(\mathrm{x_i})$ is the impact factor of component $\mathrm{x_i}$ and ω_i is the respective weight. n in this equation is the number of elements.

The identification of relevant components and the respective weight values are the most challenging part of vulnerability analysis. This requires an integrated assessment that brings together experts from a wide range of disciplines and from different organizations. Each expert brings its own conceptual models to study vulnerability which often address similar problems and processes using different semantics (Brooks 2003; Füssel & Klein 2006; Nicholls et al. 2008; Vafeidis et al. 2008; McFadden et al. 2007).

2.3.1.1 Coastal Erosion Risk Assessment

Coastal erosion is a complex dynamic natural phenomenon and as a result, the assessment of the associated risk is a challenging task. In this context, hazard is defined as the potential occurrence of the erosion process and its intensity along the coast. Coastal vulnerability in its turn, describes the measure of damage to elements at risk along the coast such as people or infrastructure (Cutter et al. 2008; Cutter et al. 2003; Boruff et al. 2005; Blaikie et al. 2004).

Coastal Erosion: Coastal erosion is measured through the probability of physical removal of sediment along the coast either in short or long terms (Boruff et al. 2005). This probability, in its turn, is a function of various parameters such as waves, currents, winds, tides, Sea Level Rise (SLR), and storms as well as human-induced activities in different time periods on daily, seasonal, and yearly scales or even over a century (Morang & Szuwalski 2003; Xhardé R. 2007). Human-induced activities may accelerate the sediment transport cycle. These activities include beach mining, removal of protective vegetation, explosion of tourism, infrastructure construction, and unsustainable use of land by private property owners or governmental institutes while applying non-appropriate managing plans to these regions (Xhardé R. 2007). The intensity of hazard is directly associated to the duration of the erosion process and the region under study (Boruff et al. 2005).

The best key indicator to identify coastal erosion is coastline or shoreline changes which are the common target of the most of previous works (Genz et al. 2007). Coastline or shoreline is the first boundary line which forms the intersection between the land and water (Morang & Szuwalski 2003). Studying coastal erosion based on coastline changes requires selecting a reference line which can be described by either of Mean Sea Level (MSL), Low Water Line (LWL), High Water Line (HWL), Mean High Water Line (MHWL), Beach, Cliff, Dune Volume, Vegetation line, or Tide level (TL) as illustrated in Figure 2.7. Selecting one of these references is an expert-based choice that led to fuzziness in the nature of this line and consequently, the coastal erosion rate computation.

In practice, linear regression is the dominant method used to calculate erosion rates although several other statistical approaches have been developed and employed (Genz et al. 2007). Considering the strong tendency toward using GIS tools, several toolkits have been developed to calculate erosion rate. Digital Shoreline Analysis System (DSAS), an extension of ESRI ArcGIS version 9+, is an example (Thieler et al. 2009). It evaluates coastline changes in an object-based model using a statistical analysis through a time-series of multiple coastline positions (Thieler et al. 2009). DSAS then defines a baseline or reference line and produces orthogonal transect at user-defined space along the coastline trend (Thieler et al. 2009). Another example is Soft Cliff and Platform Erosion (SCAPEGIS), a process-based model, which analyzes the coastline retreat and visualizes the result in GIS environment (Koukoulas et al. 2005).

The main challenge identifying erosion rates in coastal risk assessment are:

1) Data integration from different sources and types of data;

2) The fuzzy nature of coastline, as stated above, which consists of selecting an appropriate reference line to evaluate it despite hard-to-measure characteristics, and finally

3) Addressing the continuity of erosion and distinguishing between discrete and continuous events. This applies to soft coasts against cliff coasts. Cliff retreat along the coastline occurs in long-term in time scale (~10 to 10^2 years) and is a discrete event type while erosion on soft coasts happens continuously in daily, monthly, or yearly basis.

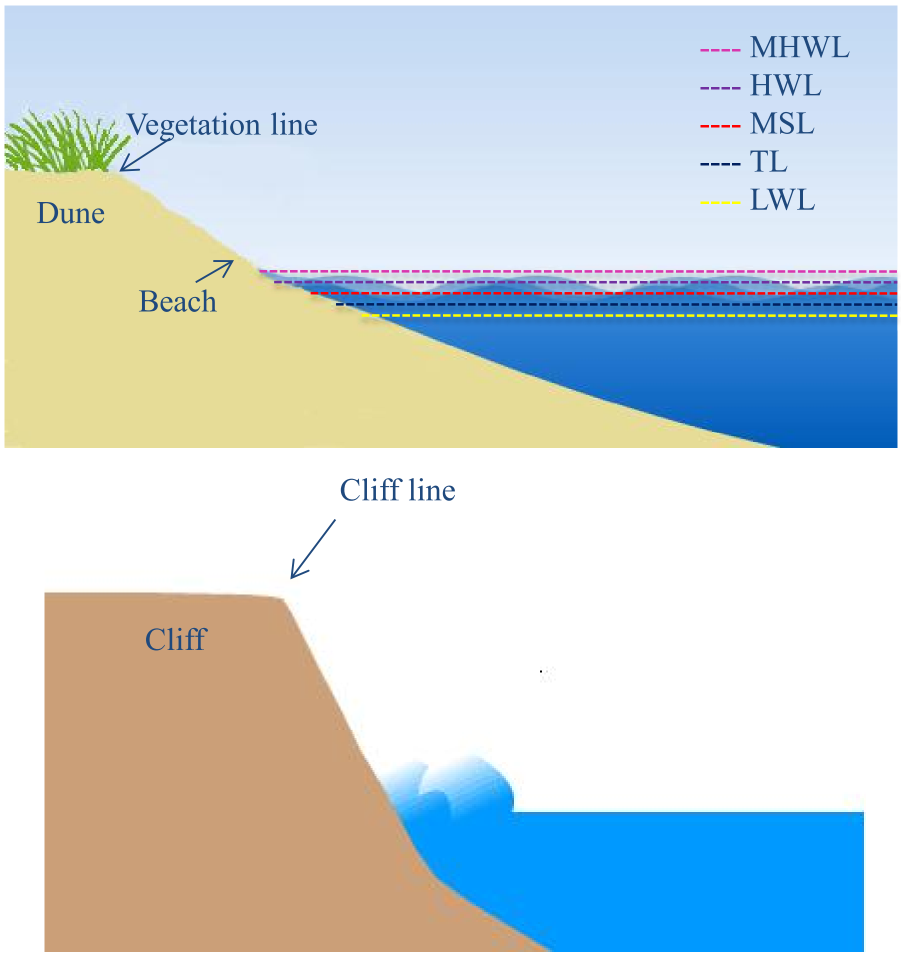


Figure 2.7: Different reference lines for coastal erosion modeling

Coastal Vulnerability: Coastal regions are considered to be extremely vulnerable areas (Klein & Nicholls 1999; Hanson et al. 2007; Füssel & Klein 2006). Coastal vulnerability in this context refers to the elements at risk that magnify or attenuate the effect of the erosion process (Boruff et. al. 2005). Indeed, current coastal erosion risk assessment methods are principally based on defining spatial units and elaborating a vulnerability index for those spatial units in a given time period (Abuodha & Woodroffe 2006; McFadden et al. 2007;

Vafeidis et al. 2008; Vafeidis et al. 2004). The degrees of vulnerability of coastal regions depend especially on human settlement structure and values, response option and adoption, politics and management strategies as well as awareness and information (Füssel 2010b; Füssel 2010a; Füssel & Klein 2006). Hence, coastal vulnerability is a function of time and its evolution through time should be considered in any study. It is important to consider the following aspects in a vulnerability analysis (ISDR 2004):

1) Objective of the study is described in international, regional, and local policies;
2) Scale is recognized as global, national, regional, local and site specific;
3) Time and money resources are characterized as low, medium, and high, and
4) Data availability is characterized as coarse, medium, and detailed.

As stated previously, the most common methods to evaluate vulnerability are index-based (Adger 2006; Füssel & Klein 2006; Füssel 2010b; Füssel 2010a; Cutter et al. 2008; Nicholls et al. 2008; Abuodha & Woodroffe 2006; Uricchio et al. 2004; Xhardé R. 2007; Vafeidis et al. 2004; Vafeidis et al. 2008; Gornitz et al. 1997). A summary of these indices, their variables, and their geographical distributions are provided in Table 2.1.

Coastal Vulnerability Indices (CVI) have been developed as a rapid and easy-to-perform method for characterizing the relevant factors. By evolution of GIS in storing, integrating, and managing large amounts of data, as well as its capability to analyze spatial data, the CVI method is extensively employed in vulnerability analysis using GIS technology (Gornitz et al. 1997). CVI is a static metric with limited predictive capability which is calculated as the square root of the product of the vulnerability degree of features divided by the total number of features:

$$CVI = \frac{\sqrt{(a_1 \times a_2 \times ... \times an}}{n} \quad \text{Eq. 2.3}$$

The index is applicable only to the region for which it is calculated, i.e., it is a local index that cannot be exported to other areas.

To include social aspects and to make a link between social criteria and hazard, the concept of place vulnerability model is introduced in Cutter et al. (2000). This model highlights two important issues i.e. spatial dimension and spatial scale in coastal risk assessment (Abuodha & Woodroffe 2006). Social Vulnerability Index (SoVI) is then developed based on the place vulnerability model (Boruff et al. 2005). CVI and SoVI are finally integrated to create CsoVI which takes into account both physical and socioeconomic aspects of vulnerability (Boruff et al. 2005). The different nature of the spatial and temporal scale of socioeconomic and physical data is an issue in the elaboration of CSoVI. Social data are available at census level, while physical attributes are at shoreline-segment. In addition, the physical data includes both long-term (e.g. SLR), as well as daily tide average, while social data represent a snapshot for one census year (Boruff et al. 2005).

Table 2.1: Summary of coastal vulnerability indices, their geographical application and the variables needed to implement them (adapted from (Abuodha & Woodroffe 2006)**).**

Index Name	Variables	Geographical Application
Coastal vulnerability index (CVI)	Mean elevation variable, Relative sea level rise, Geology, Landform/ geomorphology, Historical shoreline erosion/accretion, Tide range, Wave height, Hurricane probability of occurrence, Tropical storm probability of occurrence, Hurricane strike frequency, Tropical cyclone forward velocity, Mean annual number of extra-tropical cyclones, Mean hurricane surge height variables	USA
Coastal vulnerability index (CVI)	Historic shoreline erosion rates, geomorphology, relative rates of sea-level rise, coastal slope, wave height , tidal range	USA
Social vulnerability index (SoVI)	Principal components analysis of Census-derived social data	USA
Coastal social vulnerability score (CSoVI)	Combination of CVI and SoVI	USA
Sensitivity index (SI)	Relief, sea-level trend, geology, coastal landform, shoreline displacement, wave energy, tidal range	Canada
Erosion hazard index	As SI, exposure, storm surge water level, slope	Canada
Risk matrix	Location, infrastructure (economic value), hazard	South Africa
Sustainable capacity index (SCI)	Vulnerability and resilience of natural, cultural, institutional, infrastructural, economic and human factors	South Pacific
Sensitivity index	Shore face slope, coastal features, coastal structures, access, land use	Ireland
Vulnerability index	Disturbance event frequency, relaxation (recovery) time	UK

As mentioned earlier, vulnerability consists of two factors that are the indicator and the importance or weight of the indicator. The obtained results for any of these factors are declared either quantitatively in an interval scaling [1…5] or translating qualitatively using low, medium, and high levels. Elaborating vulnerability index is an expert-based task. Therefore, these criteria and their importance can easily vary from an expert to another.

The consistency of vulnerability indices for a given region in a specific time period is hence an important issue in risk assessment.

2.3.2 Risk Assessment Methods

In general, risk assessment methods are categorized into quantitative, semi-quantitative, or qualitative approaches (Dziubinski et al. 2006; Abuodha & Woodroffe 2006). The quantitative and semi-quantitative methods are based on probabilistic analysis i.e., likelihood-consequence risk matrix through mathematical technique and engineering evaluation (Muhlbauer 1996). Qualitative methods are mainly based on what/if checklist analysis, event tree analysis, cause-consequence analysis, human-error analysis, and safety review (Thomasoni 2010). Depending on the applied method, the results of risk assessment are represented either by a chart, a table including qualitative expression (low, medium, strong), or a risk map. The level of risk is directly related to the range of vulnerability as well as its physical value. In a multi-criteria risk assessment, risk rating systems including economic, social and environmental damage categories are considered all together.

Numerous frameworks and methods have already been developed for coastal erosion risk assessment (IPCC 2007; ISDR 2004; Klein & Nicholls 1999; UNFCCC 1999; NOAA (National Oceanic & Atmospheric Administration) 2003; Mai & Liebermann 2002; Hinkel 2005; McFadden et al. 2007). Table 2.2 presents the most frequently used methods that were developed by different research groups and scientists. The principal criteria for comparing these methods are 1) their ability to assess risk elements i.e. vulnerability and hazard, 2) input data type, 3) handling both time and space scales, 4) visualization type, 5) result types, and 6) the role of GIS technology. A GIS-based spatiotemporal representation of the assessed risks then allows communicating of the results with expert or non-expert people conveniently to anticipate the range of danger in coastal regions (Fuchs et al. 2011).

The descriptions presented in Table 2.2, demonstrate the considerable impact of GIS on the advancement of coastal risk assessment methods and techniques. Although, the first two methods that are IPCC Common Methodology (IPCC 2007) and UNEP Handbook Methodology (2002), the others are GIS based methods. For instance, the Coastal zone Simulation Model (Cosmo) is a GIS-based Decision Support System which allows coastal managers to evaluate management strategy options under different scenarios including long-term climate change. The method has been developed in the Netherlands to support decision-making subroutine of Integrated Coastal Zone Management (ICZM) plan. The main aspect of this tool is its ability to simulate day-to-day management parameters of coastal zone. This tool can easily be applied to site specific case studies and national scale as an educational tool. However, the specific case study dependency is a drawback of this method to be applied in other cases (UNFCCC 1999).

Table 2.2: Most common used methods and tools for risk assessment, especially in coastal zone management.

Method	Risk Assessment: Hazard	Risk Assessment: Vulnerability	Input	Output	Scale	GIS Role
IPCC	No	Yes	Physical and socio-economic characteristic of study area	Vulnerability profile, list of future policy needs	Sub-national, National, regional, global	No role
UNEP	No	Yes	Bio-physical and socio-economical of data	Provide input for future modeling	Regional, national, specific regions	No role
COSMO	No	Yes	The user's chosen management strategy	range different options	Site specific, national	DSS tool
SimLuicia	No	Yes	Natural, social, economic data	Table, map, text file	Macro, micro	DSS,
DIVA	No	Yes	The user's chosen scenario	Table, map, chart	Regional, global	Database developing, graphical user interface
CVAT	Yes	Yes	Environmental, social, and economical data in GIS format	Static GIS map, relative risk or vulnerability value	Community level	Map viewer, and analysis tool
CV&A	No	Yes	Local experience in relation to climate variability	Vulnerability range and adaptation recommendation	Community level	Map viewer, modeling event or scenario-based method
SmartLine	Yes	Useful	Aerial photograph, cartographic map, geomorphology data	GIS based geomorphic map of coastal sensitivity	Large scale, site specific	Capture geographical data in object-based format, map viewer
RISC	Yes	Yes	Water levels, and geometry of coastal zone	Flood zoned map, loss and risk value	Specific site	DSS tool, using Arcview programming abilities (AVENUE), visualization the result
SimCoast	No	Yes	ngineers, natural and social scientists, law-makers, administrators, community and national leaders experts knowledge	Map, table, graph	Specific site	Integrating data, visualization, process, event, and object based analysis, Zone-based visualization

Dynamic Interactive Vulnerability Assessment (DIVA) is another GIS corporation tool that allows analyzing a wide range of mitigation and adaptation scenarios by providing a predication to the global impact of climate change on coastal zone for the next 100 years. DIVA involves a global database of natural systems and socio-economic factors, relevant scenarios, a set of impact-adaptation algorithms, and a customized graphical-user interface. Considered factors include erosion, flooding, salinization, and wetland loss. DIVA can provide a reasonable result on both regional and global scales by producing table, map, and chart for each scenario. It should be reminded that applications on national scale does not supply particular perspective on vulnerability of the coastal zone (Hinkel 2005; Hinkel & Klein 2007).

Community Vulnerability Assessment Tool (CVAT) has been developed by the coastal center of the National Oceanographic and Atmospheric Administration (NOAA) (NOAA (National Oceanic & Atmospheric Administration) 2003). CVAT is equipped with a static GIS map overlay procedure that allows for analyzing relative risks or vulnerabilities in coastal regions to a series of existing threats. The computer requirements to for operation are Web-Browser (for viewing the text), ESRI suite (version 3+). CVAT requires well-customizing/initialization with respect to the case at hand and relatively depends on data availability. In spite of the ability of GIS in spatial dynamics analysis and visualization, CVAT still uses static modeling and representation methods for a highly dependent time and space changing phenomena. Later, a new version of CVAT known as Community Vulnerability Adaptation Assessment (CV&A) was developed which benefits local experiences in relation to climate changes over time and for extreme events. Modeling and scenario generation play an important role in CV&A that requires skill and expertise in coastal engineering to convey the drawbacks and limitations of this tool (Nakalevu 2006).

The SmartLine tool provides an easy and rapid capturing of geographical information in a segmented line within GIS interface and enables the user to visualize coastal sensitivity within an object-based model. However, adaptation of SmartLine to local and site specific scales is still questionable and requires testing and validation (Sharples 2006). In parallel, numerous other methods and tools (such as SimCoast (SimCost 2013) and SimLucia) have been developed which are equipped with simulation methods (Van Kouwen et al. 2007). Nevertheless, data availability, data uncertainty, prone spatial analysis and visualization, and their sensitivity to region and scale are considered as the limits of these methods.

RISC is a Decision Support System (DSS) tool that provides information on the risk of the probability of failure of dikes and degree of loss in case of failure. This tool is adapted for possible risk of storm surge on the German North Sea coast derived from water levels and geometry of coastal zone. RISC is developed using GIS –ArcView– with expanded capabilities of being programed in AVENUE. The consequence of dike failure is presented as a map of flood zone and the calculation of loss and risk (Mai & Liebermann 2002).

The main limitations of the stated methods and tools in Table 2.2 are:

1) The majority of them consider only the physical characteristics of coastal vulnerability and ignore their socioeconomic aspects for risk assessment,

2) These approaches are mostly developed for one spatial scale and are not suitable to be used for multi-scale analysis purposes,

3) They do not consider non homogenous distribution of elements at risk exposed along the coastal area,

4) They do not consider the variable intensity of the erosion over time and in space,

5) Coastal regions are occupied by diverse communities and managed by different organizations such as fisheries, natural resources, agriculture, transport, and municipalities with local, provincial and federal authorities. These organizations have their own sources of data and criteria that may result in different and conflicting risk values. This may consequently lead to inconsistent and inefficient decisions to prevent potential damages.

A comprehensive tool for multi-scale representation of risk, not only in various spatial and temporal resolutions, but also in different levels of abstraction is still missing. This tool is crucial to sustainable development perspective of coastal regions. In the next section, the methods allowing the representation of the generated results to decision-makers and potential users are discussed.

2.3.3 Spatial Risk Representation

The main goal of risk representation is to allow the potential impact of possible risks to be accurately communicated to decision-makers. It also provides a wide range of information in a user–friendly format easily understood by non–experts. Risk representation is a mathematical combination of vulnerability maps (derived from elements at risk) and hazard maps from different periods of time (Vanneuville et al. 2005; Alexander 2000; Manche 2000). In other words, risk representation is the spatial relation of hazard and vulnerable features at risk (Daudé et al. 2009). In a layer map, risk assessment is generally performed in each spatial unit, a sub-division of the region under study, subject to potential hazards (McFadden et al. 2007; Cutter et al. 2003). The spatial units may have different forms and sizes and may be distributed regularly or irregularly. The area and forms of these units may vary from a specific infrastructure to a very small cadastral parcel, municipality, state, or even a country. For each unit, a vulnerability index and subsequently, a degree of risk are assigned (ISDR 2004; Abuodha & Woodroffe 2006; McFadden et al. 2007; Vafeidis et al. 2008; Füssel & Klein 2006).

The traditional way to assess and represent risk zones is illustrated in Figure 2.8. The first impression is that risk mapping is performed by simply overlaying several data layers and superimposing them into one layer yielding a risk map. This simple overlay is well implemented in existing GIS tools. However, the problem arises when different types of information with different levels of details in different time periods are overlaid and aggregated. In this case, any simple query becomes complex and its execution turn out to be a timely task, while fast synthesis and rapid analysis are essential in risk assessment. Regardless of the discretization form of the region under study, the final results for risk zones are the aggregation of a series of these units with the same degree of risk. The degree of risk, as stated before, is generally expressed by a numerical value in interval [1 5] or by a qualitative expression (very low, low, medium, high and very high level) and represented by a color hue. A typical representation of the zones at risk is illustrated in Figure 2.9.

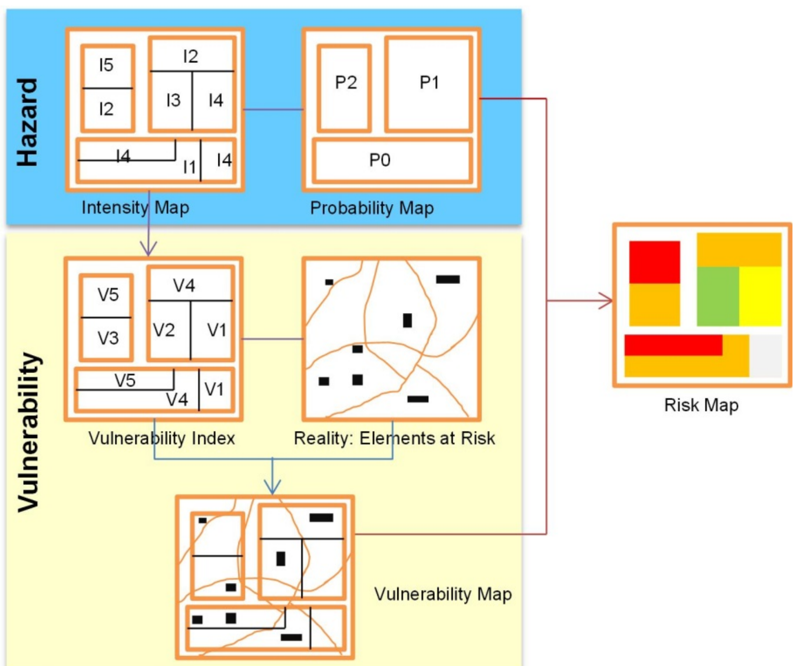


Figure 2.8: Traditional way to represent risk zones through a risk map (adapted from (Manche 2000)**)**

In traditional risk representation, whether the spatial units are regular or irregular, they are defined with well-defined polygons while their boundaries are expressed with a crisp line. Considering the fact that coastal erosion risk is a function of several continuous parameters such as erosion rate, land cover change, and climate change, the transition from a spatial unit to another can loosely be explained with a crisp line. For instance in Figure 2.9, the transition from red zone to yellow zone is sharp, whereas, risk changes smoothly and continuously. This problem, together with other issues of spatial uncertainty, led us to explore and develop more flexible method to represent risk value considering the continuity, heterogeneity, scale-dependency, and fuzzy characteristic of coastal erosion risk. In Chapter 4 and 5, we discussed these aspects in detail.

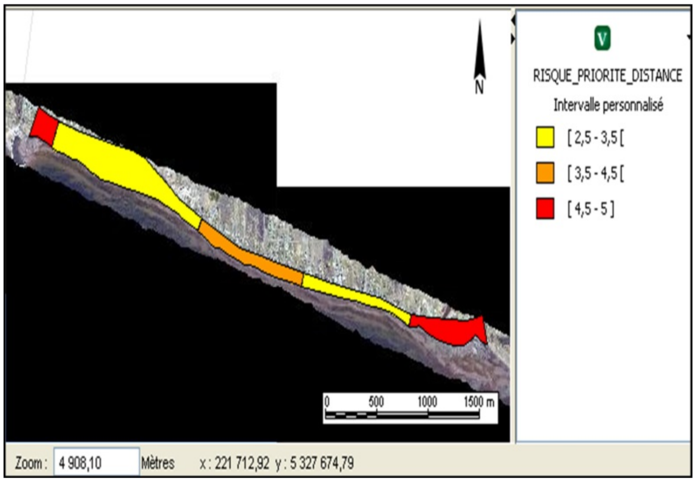


Figure 2.9: Coastal Erosion Risk Representation along the coast (after (McHugh et al. 2006))

2.4 Spatial Uncertainty: Characteristics and Methods to Deal with

This section addresses the fundamental concepts related to spatial uncertainty. Spatial data imperfection is occasionally used interchangeably as spatial uncertainty (Dilo 2006; Bejaoui 2009; Devillers et al. 2010). However, the term "uncertainty" is used in the course of this thesis.

2.4.1 Characteristics

Different kinds of spatial uncertainties in geospatial data quality are distinguished in Worboys (1998) . Spatial uncertainty may be caused by:

- Inaccuracy and error deriving from true values,
- Incompleteness resulting from lack of relevant information,
- Inconsistency arising from conflicts in information,
- Imprecision due to limitation on the granularity or resolution at which the observation is made or the information is represented, and
- Vagueness because of imprecision in concepts used to describe the information.

Before going further, some fundamental terms which are used throughout this thesis are overviewed and their definitions are made clear in the context of this research project. These terms/concepts are error, imprecision, vagueness, ambiguity, incompleteness, inconsistency, and uncertainty. They are considered as the source of uncertainty in modeling and representing a given phenomenon which is erosion in the case of this study.

Error is the difference between the available value and the true value which can often be measured by a calibrated instrument and more-educated-observers (Worboys 1998). Probability Theory is commonly used to handle it.

Imprecision is a limitation in the granularity or resolution related to the observation or representation manners (Worboys 1998). Imprecision is handled using probability analysis that conveys a numerical value for it e.g. standard deviation or an ellipse of error (Aerts et al. 2003).

Incompleteness is the lack of relevant information of a spatial object to describe a phenomenon (Bakillah 2012).

Inconsistency refers to the logical ambiguity in the same database in which the integrity constraint is falsely defined (Worboys & Duckham 2004; Salehi et al. 2010). Inconsistency can arise from wrong descriptions or locations of spatial objects, their geometric primitives, or semantic conflicts (Bejaoui et al. 2008; Bakillah 2012).

Vagueness is an inherent imperfection (Fisher 1993). Vagueness is related to the definition of an object and criteria (Molenaar & Cheng 2000; Cheng et al. 2009; Fisher et al. 2010). It cannot be described through a binary digit, zero or one, based on Boolean logic. Vagueness is the realm of philosophy and logic and has been described as one of the fundamental challenges in every discipline. Some scientists describe vagueness to be linguistics while many researchers suggest that vagueness is ontological (Molenaar & Cheng 2000; Cheng et al. 2009; Fisher et al. 2010). In the context of spatial data and objects, vagueness arises in the presence of ill-defined boundaries to describe an object either resulting from a vague linguistic term or ontological nature. Forest stand is an example of an object with ill-defined boundaries, linguistic or ontological. The definition of a forest stand's boundaries may vary with respect to different views of researchers, partners in a project, or authorities. Moreover, the limits of one class to another on a forest stand cannot be represented as a crisp line since the transition zone from one class to another or one to another object may be gradual as between different types of species in a forest. Soil type, land cover, vegetation, pollution zones, and lakes are other examples of such spatial objects in this regard.

Ambiguity appears when different conceptual schemes are defined for the same object or classes corresponding to different perceptions of it, i.e. different semantics (Fisher et al. 2010). However, non-overlapping classification of diverse users for a phenomenon can also be introduced as a level of uncertainty; a discord that is addressed by semantic analysis (Worboys 1998; Fisher 2008). For instance, coastline definition can vary from one region to another. Deriving different reference lines introduces ambiguities to the coastline change modeling and subsequently, the risk related to this phenomenon.

Spatial uncertainty is the lack of knowledge of the true value of a parameter as an attribute of information (Walker et al. 2003; Zadeh 2005). It can also originate from inherent limitations of the modeling process (Bédard 1988). In other words, a gap between the geographic reality and spatial data modeling can cause the uncertainty. The spatial uncertainty often appears in modeling natural systems originated from vagueness in boundary zones, ambiguities in linguistic terms, discord in semantic of a class or object, fuzziness in process interpretation, existence degree of a spatial object, or a mix of them (Fisher et al. 2010). Inappropriate data representation is also a source of uncertainty on another level in the decision-making process (Dilo 2006).

More precisely and in the context of this research project, the uncertainties associated with coastal risk assessment mainly originate from data and objects or criteria definitions (Cheng et al. 2009; Fisher 2008). The uncertainty from data includes sampling and measurements of phenomenon under study (Cheng et al. 2009; Fisher 2008). This type of uncertainty is essentially random in nature and the probability theory can be applied to handle it appropriately (Fisher et al. 2010; Aerts et al. 2003; Heuvelink et al. 2007). The uncertainty from object definition refers to unsure knowledge of how to define the boundary of the coast and land precisely (Cheng et al. 2009; Fisher 2008). This uncertainty is fuzzy in nature and can be handled by possibility theory (such as Fuzzy Set Theory) using membership functions (Molenaar & Cheng 2000; Fisher et al. 2005; Fisher et al. 2010).

Four levels of uncertainty are proposed by Bédard (1988) for spatial data that consist of conceptual (when identification of entity classification and its existence are fuzzy or imprecise), descriptive (when the definition of an attribute value is fuzzy or imprecise), positioning (spatiotemporal aspect of an observed reality is fuzzy or imprecise), and meta-uncertainty (unknown degree of the preceding uncertainties) (see Figure 2.10). At first glance, the uncertainties associated with coastal risk assessment can be a combination of all those levels with a degree of vagueness and ambiguity resulting from the continuity and heterogeneity of the risk zones and the scale used in risk analysis respectively.

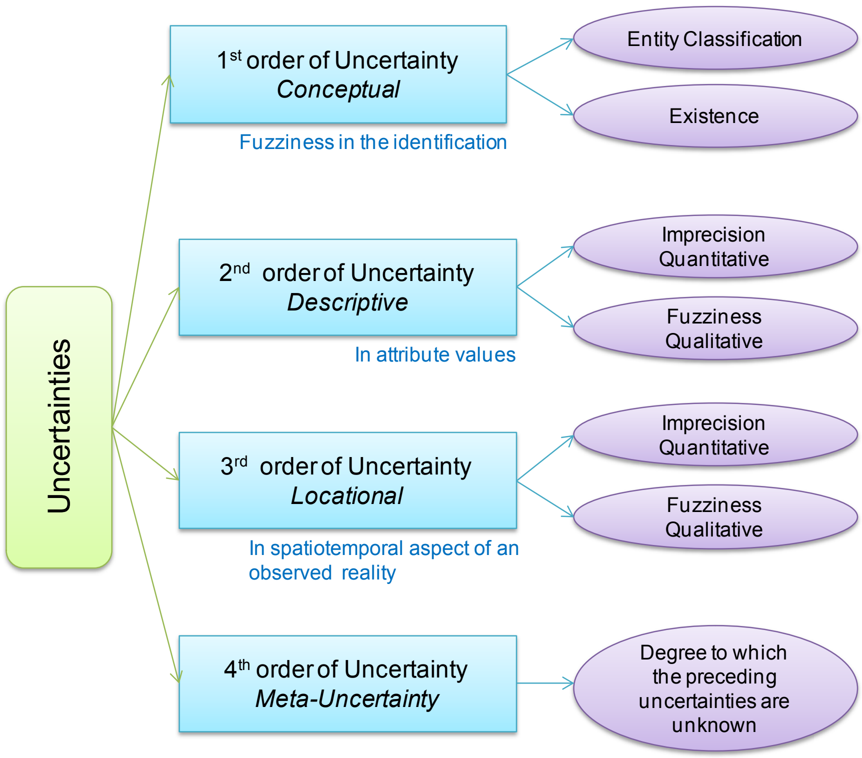


Figure 2.10: Level of uncertainties (Bédard 1988)

According to the standard ISO 19113 (ISO/TC211 2003), spatial data quality is "the totality of a feature and characteristics of a product or service that stand on its ability to satisfy stated or implied needs". In the geospatial community, there are numerous definitions for data spatial quality involving two main categories i.e. internal and external (Devillers et al. 2010). The internal quality refers to the quality of a product or service that refers to the producer of data, contrary to the external quality that includes the satisfaction of users (Devillers et al. 2010). Key elements of internal quality are actuality of data, geometric and thematic accuracy, lineage, logical consistency, and completeness (Mostafavi et al. 2004). On the other hand, external quality refers to the fitness for uses that describes the subjective characteristics (Servigne et al. 2010). The descriptions of the fundamental elements of internal spatial data quality are as follow:

Lineage (genealogy) describes the history of the available geographic dataset. It provides the principal information such as the source of data and acquisition and derivation methods including all information involved in data production (Servigne et al. 2010; Mostafavi et al. 2004).

Completeness refers to the spatial and thematic properties of data, data omission in the case of data missing, and commission in the case of data availability (Servigne et al. 2010; Mostafavi et al. 2004).

Logical Consistency is the degree of consistency of data with respect to its specification that also describes the fidelity of relationships encoded in the data structure (Servigne et al. 2010; Mostafavi et al. 2004). Logical consistency constitutes geometry constraints, topology constraints, and semantics constraints (Salehi et al. 2010; Bakillah 2012).

Positional Accuracy refers to position exactness of an object in space. Positional accuracy can be described as relative and absolute accuracies (Servigne et al. 2010; Mostafavi et al. 2004). The differences between measured distance in the map and on the real world are described as relative accuracy (Mostafavi et al. 2004). Absolute accuracy is the difference of a geographic position on the map and its real position on the earth (Mostafavi et al. 2004).

Attribute Accuracy refers to the accuracy of information assigned to the object as attribute (Mostafavi et al. 2004). Attribute accuracy measures the accuracy of a qualitative or quantitative value assigned to the thematic attribute of a spatial object involved.

Temporal Accuracy describes the accuracy of temporal information assigned to the geographical entity and their temporal relationship (Mostafavi et al. 2004).

All these data quality elements can be involved to create or increase the level of uncertainties i.e. inaccuracy, incompleteness, inconsistency, imprecision, and vagueness as stated above (Bejaoui 2009). Consequently, any enhancement of data quality reduces the uncertainty in CERA.

2.4.2 Methods to Deal with Spatial Uncertainty

As for uncertainty with respect to the nature of spatial object modeling, two approaches are proposed in Fisher et al. (2005). Spatial objects are modeled by either fiat (poorly-defined) and bona fide or crisp (well-defined) objects (Fisher et al. 2010; Smith & Varzi 2000). If a well-defined model is used, the uncertainty is probabilistic and appears as error. It can then be characterized using Probability Distribution Function (PDF). If a poorly-defined model is used, the uncertainty principally originates from the attributes or ambiguity in the definition of the object. The uncertainty from the attribute is described as vagueness and Fuzzy Set Theory is a solution to handle it (Fisher et al. 2005). The uncertainty from ambiguity differs in semantics classification (Fisher 2008). It appears as discord when one object is clearly defined but is shown to be member of two or more different classes under differing schemes or interpretations of evidence. When the process of assigning an object to a class is open to interpretation, the problem is then non-specificity (Fisher 2008). The hierarchy of uncertainty originated from the nature of spatial models is illustrated in Figure 2.11.

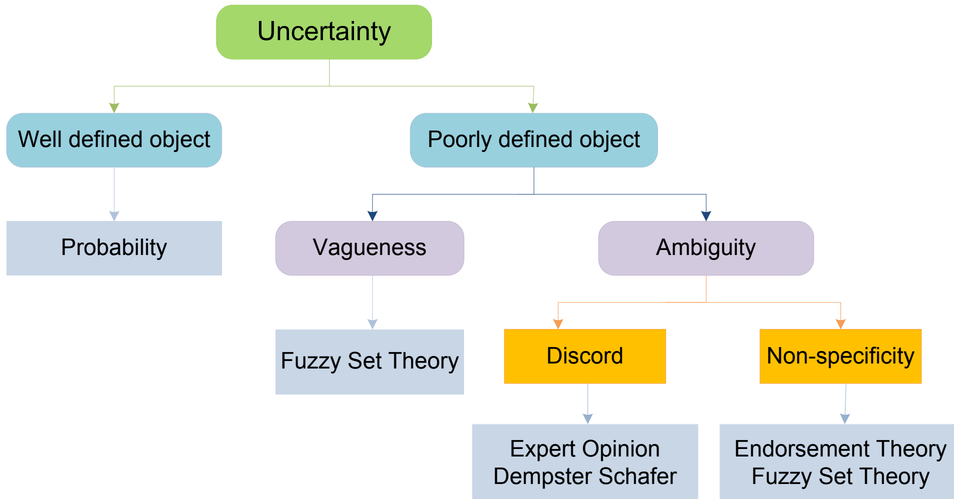


Figure 2.11: Hierarchy of uncertainty in spatial data model (Fisher et al. 2010)

Several works have already been carried out to deal with spatial uncertainty in spatial data modeling and representation. Table 2.3 summarizes a series of these methods. Fuzzy and rough models are used for uncertainties originating from the attributes of information (Robinson 2003; Burrough 1989; Usery 1996; Fisher et al. 2010). Probability, exact, rough, and fuzzy models are investigated to deal with the geometrical aspect of uncertainty (Altman 1994; Brown 1998; Cheng et al. 2001; Fisher et al. 2010). A formal syntax model for conventional crisp objects based on the fuzzy object model is also proposed in Molenaar (2000) to integrate - both descriptive and geometrical aspects of uncertainty. This latter is considered as a revolution in dealing with uncertainty through a concept which absorbs uncertainty inherently. Many researchers propose fuzzy set theory as the ultimate solution for handling different types of uncertainty (Molenaar 2000; Cheng 2002; Cheng et al. 2005; Fisher et al. 2010; Cheng et al. 2009; Dilo et al. 2007; Robinson 2003). A deep investigation on fuzzy object representation of fuzzy set theory is therefore provided in the next section.

Table 2.3: A summary of different types of uncertainties and the respective modeling solutions

Uncertainty Type	Methods	Reference
Related to attributes	Fuzzy Model Rough Model Exact Model	(Robinson 2003; Burrough 1989; Usery 1996; Fisher et al. 2010)
Related to geometry	Exact Model Fuzzy Model Rough Model Probability Model	(Altman 1994; Brown 1998; Cheng et al. 2001; Fisher et al. 2010)
Related to the geometry and attributes (an integrated aspect)	Fuzzy Model Exact Model	(Molenaar 2000; Cheng 2002; Cheng et al. 2005; Fisher et al. 2010; Cheng et al. 2009; Dilo et al. 2007)

Some works have been initiated to deal with the problem of uncertainty and vagueness in OLAP systems (González et al. 2009; Laurent 2010; Péres et al. 2007; Molina et al. 2006; English et al. 2004; Pedersen & Jensen 2001; Kaya & Alhajj 2006; Sboui 2010; Bejaoui 2009; Gervais et al. 2009; Levesque et al. 2007). (English et al. 2004) employs fuzzy logic to execute a spatiotemporal query in OLAP. (Pedersen & Jensen 2001) handles imprecision through dimensions in OLAP. (Laurent 2010) proposes a multidimensional model which deals with fuzzy facts by considering fuzzy relations and partitions in dimensions. (Kaya & Alhajj 2006) uses fuzzy association rules by allowing fuzzy labels to define dimensions. (Molina et al. 2006; Delgado et al. 2004) handle the imprecision by defining fuzzy hierarchies and facts. Some efforts have been made on coupling a possibilistics approach such as exact and rough models in SOLAP (Bejaoui 2009; Siqueira & Ciferri 2012; Edoh-alove et al. 2013). They redefine spatial attributes, measures, dimensions, and hierarchy. In recent works, (Edoh-alove et al. 2013; Gervais et al. 2009; Levesque et al. 2007) propose two risk-aware approaches for spatial datacube design to face spatial fuzziness and/or uncertainties. The proposed approaches inform the users about the existence of low-quality data and using it in spatial datacubes. (Levesque et al. 2007; Gervais et al. 2009) also propose the enriching of metadata through additional information about uncertainty and data quality via contextual warning while using the system. However, none of the stated works explicitly characterize and handle information vagueness in SOLAP by means of fuzzy model. There is, therefore, a great need to couple uncertainty and vagueness in a spatial data model and embed them into spatial datacubes. This includes the methods to aggregate and represent spatial measures resulting from uncertain data into multiple hierarchical dimensions.

2.5 Spatial Data Model

Modeling is an abstraction of the real world used to represent the phenomena occurring on or near a surface (Bédard et al. 2007; Goodchild & Glennon 2008). This necessitates facing and dealing with complex and continuous systems by simplifying them in an efficient manner (Goodchild & Glennon 2008). The spatial data management complexity comes from the fact that the world is in continuous change both over space and time.

In practice, the modeling of reality constitutes three levels of abstraction that are conceptual, data model, and data structure which refer respectively to the conceptual representation, functionally-oriented representation, and implementation (Burrough & Frank 1996). The conceptual level is the perception of a phenomenon by addressing its key characteristics (Burrough & Frank 1996). Data model is the formal definition of a conceptual model, regardless of the implementation level (Burrough & Frank 1996). In other words, a data model is a set of entities and their relations used to express the complexity of reality (Goodchild & Glennon 2008). Data structure is the digital format of a data model, in which the data are required to be stored in a computer (Goodchild & Glennon 2008). This is illustrated in detail in Figure 2.12.

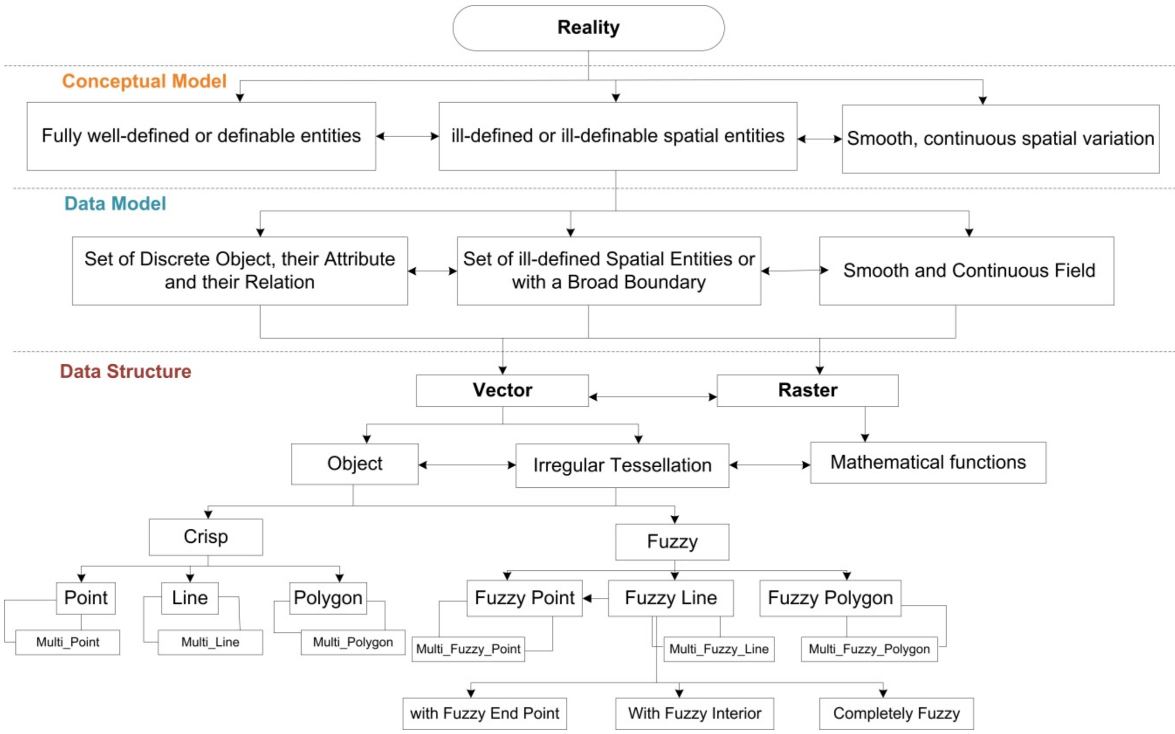


Figure 2.12: Three levels of the real world modeling (adapted from (Burrough & Frank 1996)**)**

Extending the concept from crisp to broad boundaries, two types of spatial objects are distinguished that are fiat or fuzzy (poorly-defined) and bona fide or crisp (well-defined) (Fisher et al. 2010; Smith & Varzi 2000). A fiat object has a broad boundary that cannot be observed and measured or is not known precisely (Fisher et al. 2010; Smith & Varzi 2000). An example as stated before is forest stands whose extents cannot be determined without any uncertainty. A Bona fide object represents an object discontinuous from the space with a sharp-line boundary (Fisher et al. 2010; Smith & Varzi 2000). This is typically the case for artificial objects that are built by humans. Examples are buildings and roads. Traditional geometrical models have certain limits to representing of spatial objects with vague shapes (Robinson 2003); they principally model only the certain part of a vague shape object and simply ignore the vague part (Bejaoui 2009). This problem arises mainly from the complexity of vague parts and technological limits in storing and managing the uncertain part. Subsequently, the simplified reliability is modeled. For instance, risk zones in a traditional risk assessment procedure are represented as crisp objects with sharp boundaries whilst they generally are fuzzy. Consequently, some of the inherent characteristics of these objects are missing through the representation process.

To model spatial vagueness, four main approaches are used such as exact, rough, probabilistic, and fuzzy models. Exact models are the extension of the crisp spatial models such as Egg-Yolk (Randell et al. 1992; Cohn et al. 1997; Clementini & DiFelice 1997; Erwig & Schneider 1997; Cohn & Hazarika 2001; Cohn & Gotts 1996). Rough models are based on the *Rough Set Theory* to represent a spatial object as a pair of its maximum and minimum approximations (Worboys 1998; Fisher et al. 2010; Bejaoui et al. 2008). Rough models are apt to handle the ambiguity of a spatial object. Probabilistic models based on *Probability Theory* use ellipsoid error to deal with positional and measurement uncertainties (Pfoser & Tryfona 2001; Aerts et al. 2003). Fuzzy models are based on Fuzzy Set Theory that deal with inherent fuzziness and uncertainty related to modeling an object or identifying criteria through a membership function (Zadeh 1965; Robinson 2003; Fisher et al. 2010; Dilo et al. 2007; Pauly & Schneider 2010).

A detailed literature review reveals that the exact and rough models are not the ultimate solution to overcome the inherent uncertainty due to the vague and fuzzy nature of spatial objects. This is more significant when the continuity of an object needs to be taken into account. An example is pollution modeling. It is also the case when an object has partial vagueness e.g. a lake with a sharp rocky boundary from one side and swamp boundary on the other side (Bejaoui 2009). Contrary to exact and rough models, fuzzy models promote stirring results. The principal concept of spatial fuzzy object modeling and representation is, therefore, synthetized in the next section.

2.5.1 Fuzzy Set Theory: An Approach to Deal with Spatial Uncertainty

The fundamental concept of *Fuzzy Set Theory* was originally proposed by Zadeh (1965) to deal with the inherent uncertainty and complexity of a complex object or phenomenon (Robinson 2003). Fuzzy context was introduced to the geospatial community in 1972 by Gale (1972), who employed *Fuzzy Set Theory* to describe the natural behavior of geographic domain. In the 80s and 90s, hundreds of papers, books, and reports were published to append fuzzy concepts into GIS to model and represent a dynamic phenomenon with ill-defined boundaries and in the presence of uncertainty (Cheng et al. 2001; Wang et al. 1996; MacEachren et al. 2005; Morris & Jankowski 2005; Burrough 1989; Burrough & Frank 1996; Pauly & Schneider 2010; Schneider 1999; Schneider 2003a; Schneider 2003b; Schneider et al. 2011; Zhan & Lin 2003; Zhan 1997; Dilo et al. 2007; Dilo 2006; Dilo et al. 2006).

Principally, a fuzzy set is a set of objects whose membership to this set takes a value between zero and one (Robinson 2003). A fuzzy set X is defined as:

$$A = \{(x, \mu_A(x) | x \in X \wedge \mu_A : X \rightarrow [0,1])\} \qquad \text{Eq. 2.4}$$

In Eq. 2.4, $\mu_A(x)$ is called the membership function that quantifies to what degree member *x* belongs to set *X*. This definition can easily be applied in 2D to represent a spatial object. The core, the transition zones or the indeterminate boundaries, α-cut boundary, and the conditional boundary of a simple fuzzy object as well as its membership function are illustrated in Figure 2.13 (Zhan & Lin 2003). In this thesis, spatial objects with vague shapes are called *spatial fuzzy objects*. Note that the term *fuzzy* refers to the mathematical approach to deal with uncertainty and vagueness, regardless of the type and source of spatial uncertainty.

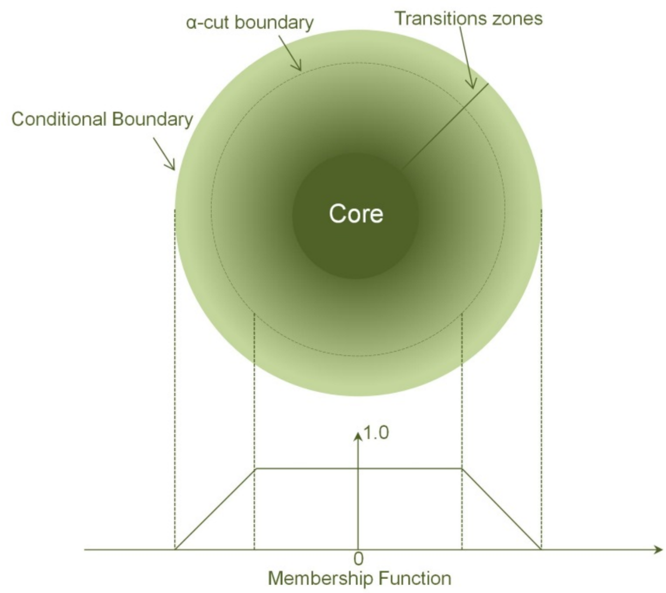


Figure 2.13: A simple fuzzy object model (Zhan & Lin 2003)

Spatial fuzzy objects are categorized into four models that are Fuzzy-Fuzzy (FF) objects, α-cut boundaries (αF), Fuzzy-Crisp (FC) objects and Crisp-Fuzzy (CF) objects (Cheng 2002). See Figure 2.14 for a graphical representation. A FF-object is the result of fuzzy classification adapted to model an object with continuous characteristics (Foody 1999; Schneider 2003a; Molenaar & Cheng 2000; Cheng et al. 2009). α-cut boundaries is another way to represent fuzzy objects in which the membership function is described by n α-cut (Zhan 1997) (Wang et al. 1996; Zhan & Lin 2003) . In other words, α-cut is a crisp set in which the element's membership degrees are equal or higher than $\alpha \in [0,1]$ i.e. $\mu_\alpha = \{x \in \Re^n \mid \mu(x) \geq \alpha\}$. The broad boundary can be decomposed into n α-cuts. Likewise, a strict α-cut of μ is defined as $\mu_{\bar{\alpha}} = \{x \in \Re^n \mid \mu(x) > \alpha\}$. The strict 0-cut of a fuzzy set $\mu_{\bar{0}} = \{x \in \Re^n \mid \mu(x) > 0\}$ is called its support set and presented as *supp(μ)*. Similarly, 1-cut of a fuzzy set $\mu_1 = \{x \in \Re^n \mid \mu(x) \geq 1\}$ is denoted as the *core of μ*.

Objects with partially broad boundaries, such as a lake enclosed in rocky banks cannot be modeled using this approach (Bejaoui 2009).

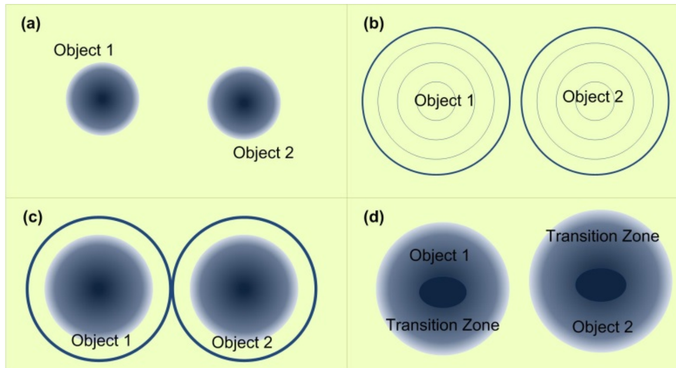


Figure 2.14: Four ways to represent fuzzy objects: (a) Fuzzy-Fuzzy areas; (b) α-cut boundaries; (c) Fuzzy-Crisp object; and (d) Crisp-Fuzzy object (Cheng 2002)

A FC-object is defined as a spatially disjoint encountering a conditional boundary arising from discretization upon conditions or criteria. This is similar to the "Egg-Yolk" model presented in several works (Randell et al. 1992; Cohn et al. 1997; Clementini & DiFelice 1997; Erwig & Schneider 1997; Cohn & Hazarika 2001; Cohn & Gotts 1996). The egg-yolk model of object A is defined as the composition of inner and outer parts with crisp boundaries A_1 and A_2, with $A_1 \subseteq A_2$. The inner region or "yolk" is the certain part of the object. The outer region or "white" is the indeterminate boundary which delineates limits on the range of vagueness. The white and yolk together form the egg that is the full extent of the vague object. The broad boundary A refers to the difference between A_1 and A_2, $\Delta A = A_1 - A_2$ (Clementini & DiFelice 1997). This representation also makes sense with the Region Connection Calculus (RCC) theory (Randell et al. 1992). The drawback of the RCC method is that it cannot model a point or a region having broad boundaries with an empty yolk or empty egg.

The fourth model is the CF-model which represents an object with a certain core (crisp region) and a fuzzy spatial extent (transition zones). The crisp region is a special case of a fuzzy region (Erwig & Schneider 1997). The crisp and transition zones are two disjointed sets, while the boundary between them can be a line or a region. The CF-model is not compatible with Egg-Yolk representation since it may not be possible to determine that an object belongs to which grid within the transition zone.

The four cited models are the ancestors of a novel data model commonly called vague spatial data model or vague spatial algebra (VASA) (Pauly & Schneider 2010). It permits the representation of vague points, vague lines, and vague regions in fuzzy sets $\Re^2$ (Schneider 1999; Schneider 2003b).

Dilo (2006) has characterized fuzzy objects into simple and general objects. Simple fuzzy object represents an identifiable object with simple structure, such as a point, a line or a region. General fuzzy types are a collection or class of simple objects such as multipoints, multilines, and multiregions. A fuzzy partition is also introduced in Dilo (2006) to represent the soft classification of space (continuous with a broad boundaries between zones) that is considered a special case of a fuzzy region. Schneider (2003a) has defined fuzzy objects as finite collections of elements from a regular grid. This approach establishes a regular tessellation of a bounded subspace of $\Re^2$ as shown in Figure 2.15. Membership values are attributed to each cell of the grid. The color hue illustrates the membership value; dark color for high membership degree close to 1 and light color for low membership value close to 0. Raster data structure can easily be implemented using this model. Moreover, any other tessellation methods such as TIN-based that use curves and surfaces can also be employed based on Schneider (2003a)'s model.

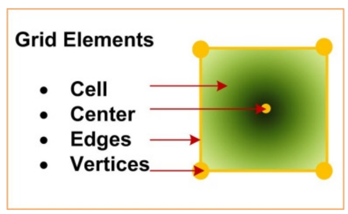


Figure 2.15: Fuzzy Object definition based on Schneider's method (Schneider 2003a)

This section is dedicated to different ways of modeling a fuzzy object. The uncertainty may come from the location, geometry, or attribute of an object. We focus on the fuzziness due to geometry and attribute rather than location of a spatial object. Indeed, the fuzziness in location may appear as fuzziness in attributes or my increase it.

Fuzzy Point

A point is a 0-dimensional object which is defined by its location and attributes. Two categories of fuzzy points are introduced in Dilo (2006) that are fuzzy point and fuzzy multipoint that is a class of fuzzy points. Fuzzy point and fuzzy multipoint are illustrated in Figure 2.16. The formal definition of a fuzzy point is provided in Appendix B.

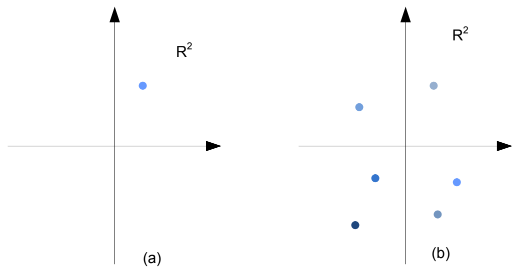


Figure 2.16: Fuzzy point object: (a) a fuzzy point, and (b) a fuzzy multipoint (Dilo 2006)

Fuzzy Line

A crisp line is defined as one-dimensional object where an interior trajectory (set of points) is ended by two endpoints (disconnected boundary). The two endpoints are the boundary of a crisp line. A fuzzy line can be the result of the vagueness on the interior or on the boundary (endpoints) of a crisp line. Hence, a fuzzy line can be partially or completely fuzzy with a broad boundary or a broad interior. The classification proposed by Dilo (2006) is also valid for fuzzy lines; a fuzzy simple line or fuzzy multiline (a set of simple fuzzy lines) as shown in Figure 2.17.

Figure 2.17: Fuzzy Line Objects: (a) Fuzzy line and (b) Fuzzy Multiline (Dilo 2006)

Another classification is proposed in Bejaoui (2009). Four categories for a fuzzy line are distinguished i.e. a line with broad boundary, a line with broad interior, a completely broad line, or a combination of the first two. A fuzzy line is, hence, represented as a union of one-dimension (minimal extents of line) and two-dimension extents (maximal extents of line). The vagueness of a line is also clustered into five levels that are none (crisp line), weakly, fairly, strongly, and completely.

Fuzzy Region

Many works have already been done on fuzzy regions and their topological relation. A fuzzy region is a region with an indeterminate boundary i.e. a transition zone. A fuzzy region can be simple or multiple as illustrated in

Figure 2.18. The indeterminate boundary can be a homogenous unit as suggested by the Egg-Yolk model (Randell et al. 1992; Cohn et al. 1997; Clementini & DiFelice 1997; Erwig & Schneider 1997; Cohn & Hazarika 2001; Cohn & Gotts 1996) or a gradual transition zone with different degrees of membership to the region (Schneider 1999; Pauly & Schneider 2010; Schneider 2003a; Schneider 2003b; Tang 2004; Zhan & Lin 2003; Zhan 1997; Dilo et al. 2006).

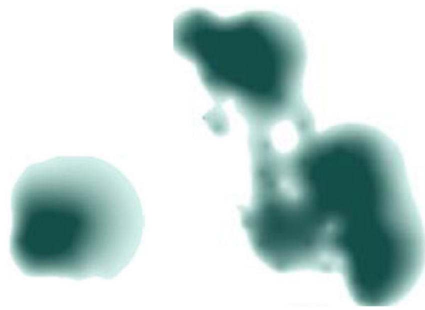


Figure 2.18: Fuzzy Region Objects: Left: Fuzzy Region, Right: Fuzzy Multi-region (Dilo 2006)

Fuzzy tessellation of space into a regular or irregular grid e.g. raster or Voronoi diagram, is practically a fuzzy multi-regions type representation (see Figure 2.19). A fuzzy partition is a soft classification. In a fuzzy multi-regions representation, different classes may not be disjointed and may intersect within the transition zones.

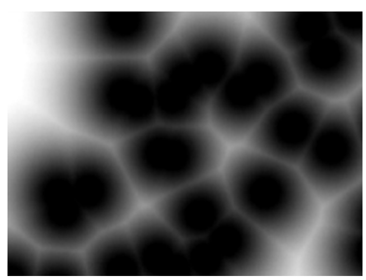


Figure 2.19: A fuzzy tessellation of space based on Voronoi diagram

2.5.2 Fuzzy Membership Function

In fuzzy set theory, a membership function (MF) determines the degree of belonging of an element to a set. Two methods are introduced in Tang (2004) to define MFs that are active or passive. In active methods, the MFs are defined based on expert knowledge. Active methods are principally Semantic Import Model (SIM) (Wang et al. 1996; Usery 1996) Fuzzy C-Mean approach (Bezdek et al. 1984; Chi et al. 1995), Self-Organized Map (Chi et al. 1995), fuzzy supervised classification (Mannan et al. 1998) and neural network methods (Robinson 2003). The shape of the membership function e.g. a simple vector, S-function, triangular, trapezoid is optimized in a passive method through successive observations depending on the application and the need to capture different levels of uncertainty.

Fuzzy object models describe the internal structure of uncertain parts of spatial objects with vague shapes. Hence, an initial quantitative hypothesis is required to characterize the vague, uncertain, or transition part of a fuzzy object and define the appropriate MF to complete modeling and representation of the respective object. MF definition remains a major challenge in fuzzy object modeling and representation.

2.5.3 Fuzzy Operators

Fuzzy objects are manipulated and represented through fuzzy operators. This section provides some basic definition of fuzzy operators for fuzzy objects. In general, there are three types of fuzzy operators with respect to their return types that are spatial, numerical, or a collection of spatial objects (Dilo 2006). Operators returning special types produce a new geometry of a new spatial object. This results from the integration of two objects' geometry. Union, intersection, difference, common border, vertices, contour, and interior operators are some examples. The formal definition of the stated operators is provided in Table 2.4.

Table 2.4: Mathematical definition of spatial type fuzzy operators (Dilo 2006)

Operators	Definition
Union	$\mu \bigcup \nu = \{(x, \max\{\mu(x), \nu(x)\}) \mid x \in \Re^n\}$
Intersection	$\mu \bigcap \nu = \{(x, \min\{\mu(x), \nu(x)\}) \mid x \in \Re^n\}$
Difference	$\mu - \nu = \mu \bigcap (1^{\Re^n} - \nu)$
Symmetric Difference	$\mu \Delta \nu = (\mu - \nu) \bigcup (\nu - \mu)$
Bounded Difference	$\forall x, \mu \nabla \nu(x) = \max\{0, \mu(x) - \nu(x)\}$
Absolute Difference	$\forall x, \mu \lvert - \rvert \nu(x) = \lvert \mu(x) - \nu(x) \rvert$

The *Union* of two fuzzy sets is the smallest fuzzy set containing both of them. The *Intersection* of two fuzzy sets is the biggest fuzzy set contained in both of them. The *Difference* of two fuzzy sets is the intersection of the first with the complement of the second. There are three other types of differences for two fuzzy sets that are symmetric, bounded, and absolute difference. It is worth stating that *fuzzy difference* is the extension of *bounded difference* and *fuzzy symmetric difference* is the extension of *absolute difference* on crisp sets. *Common border* acts on two lines or regions and returns the common boundary multiline. *Vertices* return the end points of a multiline and vertex points of a multi-polygon. *Contour* is applied on a multipolygon and returns its boundary as a multiline whereas *Interior* is applied on a multiline and returns the multipolygon enclosed by the multiline (Dilo 2006).

There are also some operators that return numbers such as *No_of_components, Dist, Diameter, Length, Area,* and *Perimeter* (Cheng 2002). *No_of_components* returns the number of components of an object by an integer value. *Dist* returns the minimum distance between two spatial objects by a real value. *Diameter* calculates the largest distance between a spatial object and its location. *Length* indicates the total length of the segments of a multiline. *Area* computes the sum area of the components of a multipolygon. *Perimeter* returns the sum of the length of the boundary line of a multipolygon.

As stated above, the third type of operator is the one whose return type is a collection of spatial objects. This kind of operator is mostly applied on fuzzy multi-regions or fuzzy partitions. Examples are *Sum, Closest, Decompose, Overlay,* and *Fusion* (Cheng et al. 2009). *Sum* aggregates the value of some spatial attributes of an object set and computes the geometric union of all these values such as reporting density of people in country by aggregating municipality. *Closest* returns the nearest object to a reference object in set of objects. *Decompose* provides the set of connected components of an object. *Overlay* returns the superimposition of two partitions; the resulting partition is the intersection of the first partition with the second partition. *Fusion* merges regions based on the equality of some attribute value.

2.6 Conclusion

Coastal erosion risk assessment is a complex process that requires handling large amounts of data and information from different sources and types. Modeling and representing such amounts of multidimensional data necessitates employing the most recent technologies such as multidimensional databases. In addition, inherent uncertainties exist arising from data and modeling. These inherent uncertainties add more complexities in modeling and representation of risk zones. To handle inherent spatial uncertainty, a new spatial data model called fuzzy spatial data model is proposed in the literature. The implementation of this data model in multidimensional databases requires a good understanding of its structure and its related operators. A detailed literature review was performed and the synthesis was provided in this chapter to highlight all fundamental concepts. The following chapters cover the design of a multidimensional datacube and how to deal with spatial uncertainty by a fuzzy logic approach on behalf of the datacube's elements.

"I can't change the direction of the wind, but I can adjust my sails to always reach my destination."
Jimmy Dean

Chapter 3 Using Geospatial Business Intelligence Paradigm to Design a Multidimensional Conceptual Model for Efficient Coastal Erosion Risk Assessment

Published in Journal of Coastal Conservation, DOI 10.1007/s11852-013-0252-5
Jadidi A., Mostafavi M.A., Bédard Y., Long B., Grenier E

Keywords: *Coastal Erosion Risk Assessment, SOLAP, Decision-making, GIS, Spatial Datacube, Geospatial Business Intelligence.*

3.1 Preface

This chapter is composed of the first paper published within the framework of the present thesis in the Journal of Coastal Conservation: planning and management. The paper is focused on the first objective of the thesis which proposes to develop a GeoBI-Based Conceptual Framework Applied to CERA. The main focus of the paper is to develop a Spatial Multidimensional Conceptual Model (SMCM) to perform coastal erosion risk assessment. Through the chapter, the methodology proposed for the definition of risk and its components and their translation to a spatial multidimensional database is described in details.

3.2 Abstract

One of the main challenges in Coastal Erosion Risk Assessment (CERA) is integrating and analysis of conflicting data in various time periods and spatial scales through dissimilar environmental, social, and economic criteria. Currently, Geographical Information Systems (GIS) are widely used in risk assessment despite their drawbacks and limitations as transactional systems for multi-scales, multi-epochs, and multi-themes analysis. Hence, an analytical conceptual framework is proposed in this paper based on geospatial business intelligence paradigm to develop a Spatial Multidimensional Conceptual Model (SMCM) to assess coastal erosion risk. The model is designed based on Spatial On-Line Analytical Processing (SOLAP) platform, on the top of both analytical and transactional paradigms, to allow fast synthesis of cross-tabulated data and easy comparisons over space, scales, epochs, and themes. This objective is achieved through a comprehensive integration of multiple environmental, social, and economic criteria as well as their interactions at various scales. It also takes into account multiple elements at risk such as people, infrastructure, and built environment as different dimensions of analysis. Using this solution allows decision-makers to benefit from on-demand, interactive, and comprehensive information in a way that is not possible using GIS alone. The developed model can easily be adapted for any other coastal region through the proposed framework to perform risk assessment. The advantages and drawbacks of the proposed framework are also discussed and new research perspectives are presented.

3.3 Introduction

Coastal communities are increasingly concerned with the risk associated to erosion given the fact that 70% of coastal regions around the world are subject to severe erosion (IPCC 2007). In this context, many scientists and experts have raised, elaborated, and studied the problem of efficient Coastal Erosion Risk Assessment (CERA) since many years (Li et al. 2012; Linham & Nicholls 2012; Karvetski et al. 2011; Santini et al. 2010).

Coastal Erosion Risk (CER) is a complex spatial dynamic phenomenon that results from spatiotemporal interactions between hazard and vulnerability on involved elements at risk in different time periods and spatial scales (Blaikie et al. 2004; Daudé et al. 2009; Varnes 1984). The three stated components are related to multiple environmental, social and economic criteria that depend in their turn on the specific needs of different actors and coastal organizations on local, regional and national levels (Cutter et al. 2003; Boruff et al. 2005). In addition, each organization carries out its own data acquisition based on its needs and standards. The resulting data are, therefore, generally heterogeneous and hence difficult to integrate, analyze and share.

To perform CERA efficiently, an information system is required to accommodate and integrate available heterogeneous data from different sources. This system should also allow on-the-fly aggregation, analysis, synthesis, and reporting of resulting information for users and decision-makers. During the past two decades, Geographical Information Systems (GIS) are widely used for the assessment, analysis, and visualization of coastal risk (Vafeidis et al. 2008; Li et al. 2012; Linham & Nicholls 2012; Karvetski et al. 2011; Santini et al. 2010). However, GIS are limited when it comes to performing complex multi-scales, multi-epochs, and multi-themes queries (Salehi et al. 2010). The operations are executed slowly and the complex queries are out-of-reach for non-GIS experts.

Recent advances in Decision Support Systems (DSS) incorporated with traditional GIS provide interesting solutions for efficient risk assessment processes. Of particular interest are the advances coming from the field of Business Intelligence (BI), where a category of Spatial Decision Support Systems (SDSS) called Spatial On-Line Analytical Processing (SOLAP) has been developed (Bédard et al. 1997). SOLAP has been designed specifically to overcome the previously stated limitations of GIS through a nested hierarchy system with several levels of abstraction (Rivest et al. 2005). Given these advantages, an analytical conceptual framework based on geospatial BI paradigm is proposed in this paper to elaborate a Spatial Multidimensional Conceptual Model (SMCM) to perform CERA. The model is an example of a comprehensive solution integrating multiple environmental, social, and economic criteria as well as their interactions at various spatial scales and epochs on different themes. It also takes into account multiple elements at risk such as people, infrastructure, and built environment as different dimensions of analysis. Through the proposed framework, the model can eventually be adapted to the context of any coastal region to perform risk assessment.

In the following section, the state of the art of the existing GIS-based CERA methods is presented. Next, fundamental concepts of geospatial BI paradigm and its adaption for CERA are described. Then, an analytical conceptual framework based on this paradigm is proposed and a SMCM is developed for CERA. Finally, the advantages and drawbacks of the proposed framework are discussed and new research perspectives are presented.

3.4 Related Works

Conceptually, risk can be assessed through quantitative, semi-quantitative, or qualitative methods (Dziubinski et al. 2006; Abuodha & Woodroffe 2006). The quantitative and semi-quantitative methods are based on probabilistic analysis such as likelihood-consequence risk matrix through mathematical technique and engineering evaluation (Totschnig et al. 2010; Muhlbauer 1996). Qualitative methods are mainly based on what-if/check-list, event tree, cause-consequence, and human-error analysis as well as safety reviews (Thomasoni 2010). The results of risk assessment are, therefore, represented either statistically by a chart, a table including qualitative expression (low, medium, strong), or as a risk map.

Traditionally, CERA is performed on spatial units with a pre-defined form (McFadden et al. 2007). A vulnerability index with respect to potential hazard i.e. coastal erosion is then assigned to those spatial units (Abuodha & Woodroffe 2006; Hinkel & Klein 2007; Füssel & Klein 2006; Gornitz et al. 1997; Cutter et al. 2003; Boruff et al. 2005; Füssel 2010a). While most of indices are designed only for environmental or socio-economic aspects of vulnerability separately, the integrated indices are relatively poorly considered (Abuodha & Woodroffe 2006; Füssel 2010a; Cutter et al. 2003). On the other hand, the common indicator to measure coastal erosion is the coastline change rate which is determined using a probabilistic approach, a simulation-based technique, or is derived from Digital Terrain Models (DTM) (Genz et al. 2007; Uricchio et al. 2004; Limber et al. 2007).

As stated before, one of the main issues in CERA is the efficient integration of large volumes of heterogeneous spatiotemporal data from different sources to determine hazard and vulnerability index (Vafeidis et al. 2004; Vafeidis et al. 2008). Available data are generally dissimilar in types, acquired based on different standards in different spatial and temporal scales. For instance, socio-economic data are available at the census level of detail whereas environmental data are at coastline-segment scale (Hegde & Reju 2007). In addition, the environmental data include both long term e.g. Sea Level Rise as well as daily tide average while social data represent a snapshot of one census year (Boruff et al. 2005).

Geospatial information technologies and more specifically GIS, provide a wide range of functionality from collection, storage, management, integration, aggregation, analysis, visualization, and diffusion of spatial and non-spatial data to perform CERA (Li et al. 2012; Linham & Nicholls 2012; Karvetski et al. 2011; Santini et al. 2010; Vafeidis et al. 2008). These capabilities can be used to identify hazard, place vulnerability indicators and spot elements at risk, and represent and communicate associated risk to users and decision-makers (Van Kouwen et al. 2007; McFadden et al. 2007; Nakalevu 2006; Hinkel 2005; Zuzekt et al. 2003; Mai & Liebermann 2002). During recent years, several GIS-based methods and frameworks have been developed by experts and scientists to deal with problems related to coastal zones management. Dynamic Interactive Vulnerability Assessment (DIVA) (Hinkel & Klein 2007), CoastBase (Kazakos et al. 2000), SimCost (SimCost 2012), Community Vulnerability, Adaptation Assessment (CV&A) (Nakalevu 2006), Community Vulnerability Assessment Tool (CVAT) (NOAA (National Oceanic & Atmospheric Administration) 2003), SmartLine (Sharples 2006), RISC (Mai & Liebermann 2002), and Coastal zone Simulation Model (Cosmo) (UNFCCC 1999) are the examples that potentially can also be applied for CERA. The characteristics of the methods and frameworks are summarized in Table 3.1.

Table 3.1: A summary of GIS-based methods developed for coastal issues.

Methods	Type of Database	Scale	Limits
DIVA	Transactional Database	Regional and global	For local scale use is inappropriate due to low quality of data in mentioned level of details.
SimCost	A simulation toolkit, visualize integrated data in a map	Specific site	Some computer and programing skills are necessary and extensive scientific training is required.
CV&A	Risk map and adaptation scenario	Local	Expertise on coastal management and query performing skill are required. It is scale dependent.
CVAT	Static risk map	Local	Query performing skill is required. It uses static modeling methods. It is scale dependent.
SmartLine	Coastal sensitivity map	Global	Adaptation for a specific site needs still to be tested and validated.
RISC	DSS tool	Local	Programing and query performing skills are necessary. It is scale dependent.
Cosmo	DSS tool	Specific site	It is appropriate for academic use and valid for specific case study.

According to Table 3.1:

1) The majority of these methods and frameworks are based on transactional GIS databases or use simple digital maps. Transactional databases are less efficient for intensive and effective analysis of multi criteria and multi- scale spatial information that is a requirement for CERA;

2) These methods are mostly developed for a single spatial scale (local, regional or global). They are not suitable to be used for multi-scale analysis purposes while navigation from a fine resolution to a coarser one and vice versa is poorly supported;

3) Data integration and aggregation at different time scales (yearly, seasonally, monthly, daily, etc.) are poorly considered; and

4) CERA requires complex analysis on large volumes of data. It also requires complex queries and syntaxes that are difficult to be efficiently expressed and supported by transactional approaches.

Merging GIS with geospatial solutions evolved in BI addresses major and principal requirements in CERA. This turns our attention toward geospatial BI paradigm as an alternative to efficiently accommodate and perform CERA analysis which is scrutinized in the next section.

3.5 Geospatial Business Intelligence Paradigm

Goespatial intellegence paradigm , evolved from BI, provides promoting solutions to overcome limitations of traditional GIS based on transactional databases. In practice, decision-makers need to analyze multiple criteria that are related to the issue under study, to summarize and aggregate data along these criteria for a global view, and sometimes to go through the details of each criterion and view and visualize the results. It is more efficient in this regard to use aggregated data with a certain time period rather than individual records from trasactional databases (Salehi et al. 2010). Using this approach, decision-makers can then focus only on specific criteria and their interactions while intelligently control the overall process, and employ all available data into a nested hierarchy system. Consequently, possible scenarios become rapidly clear and can be presented by creating new relations based on emerged options without being involved in complex queries execution.

To enable such complex analysis and visualization operations, data warehouses are employed that are commonly modeled using a datacube or multidimensional paradigm (Abelló et al. 2006). The key concepts in a typical datacube, as illustrated in Figure 3.1, are dimensions, member, measures, and facts (Kimball & Ross 2002; Torlone 2003).

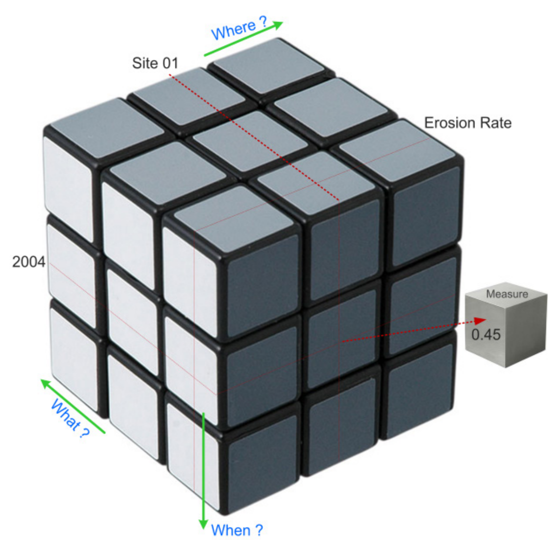


Figure 3.1: Datacube and its key elements i.e. Dimension, Member (e.g. Site 01, 2004, and Erosion Rate), Measures (e.g. numerical value 0.45), and Fact (e.g. the erosion rate of Site 01 in 2004 is 0.45 m/yr.).

Dimension is defined as an analysis perspective or theme of interest for a user (Salehi et al. 2010). A dimension can be spatial e.g. regions, temporal e.g. period and thematic e.g. products or retailers (Bédard et al. 2007). Spatial dimensions are considered as non-geometric, geometric, or mixed (see Figure 3.2); they can rely on discrete object-based or continuous raster-based structures (Bédard et al. 2009). A dimension includes one or several hierarchies composed of different analysis levels e.g. city, state, country labeled as "administrative region". A *member* is an instance of hierarchy level that states a position within the hierarchical data structure of a dimension. For instance, Canada is a member of country level (Malinowski & Zimànyi 2008).

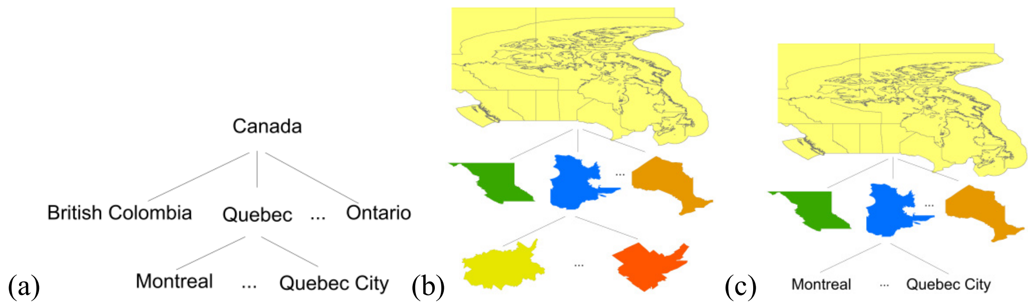


Figure 3.2: Three Types of Spatial Dimensions: (a) Non-Geometric, (b) Geometric, and (c) Mixed.

Measures are measurable quantities e.g. the number of victims in an accident with respect to the different levels of "administrative region" and "time" dimensions; these are analyzed against members of different levels of dimensions (Bédard et al. 2007). The values resulting from unique combinations between members of different dimension levels, along with their measures, are called *facts* (Rivest et al. 2005). For instance, the number of victims in car accidents in Quebec City between May and August 2010 is a fact (Salehi et al. 2010). Indeed, measures are dependent variables to dimensions, whilst dimensions are essentially independent variables in nature (Chaudhuri et al. 2011). The datacube concept efficiently supports the multidimensional characteristics of coastal erosion risk assessment where the risk components are assumed independent. This makes the whole system more consistent and coherent with the user's perception. SOLAP is an example of a geospatial BI system based on the datacube concept (Bédard et al. 1997; Stefanovic 1997; Caron 1998; Han et al. 1998; Rivest et al. 2005; Rivest et al. 2001; Bimonte et al. 2012).

To accommodate data in SOLAP, three data structures are commonly used that are star schema (see Figure 3.3.a), snowflake schema (see Figure 3.3.b), and a mixed structure or fact constellation schema (see Figure 3.3.c) (Pedersen & Jensen 2001). Star schema is very simple and intuitive compared to the other models and is supported by a large number of BI tools (Bédard et al. 2009; Chaudhuri et al. 2011).

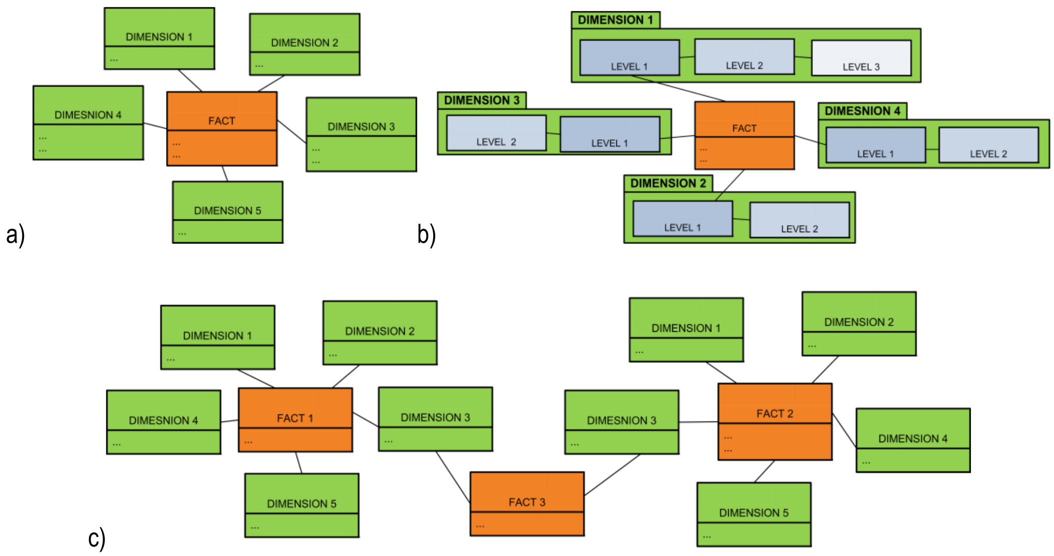


Figure 3.3: SOLAP data structure: a) Star schema, b) Snowflake schema, and c) Mixed schema.

The first works that use SOLAP for risk assessment are (McHugh et al. 2006; Iris 2009) in which the potential of SOLAP technology in risk assessment is explored focusing only on a single-element at risk-based analysis that is in this case, road network and residential buildings. However, performing multiple elements at risk analysis remains a challenge (Desprats et al. 2010; Totschnig et al. 2010). Further to the discussion elaborated in the section, an analytical conceptual framework is proposed in the next section to design an adaptive SMCM to perform a more efficient CERA.

3.6 An Analytical Conceptual Framework for Coastal Erosion Risk Assessment

An analytical conceptual framework is proposed in this section to develop a SMCM for CERA. This framework includes four main steps that consist of performing needs analysis; accomplishing data inventory; defining risk components i.e. hazard, elements at risk, and associated vulnerability index; and finally designing a SMCM that includes identifying dimensions and measures to calculate associated risk. The scheme of the proposed framework is illustrated in Figure 3.4.

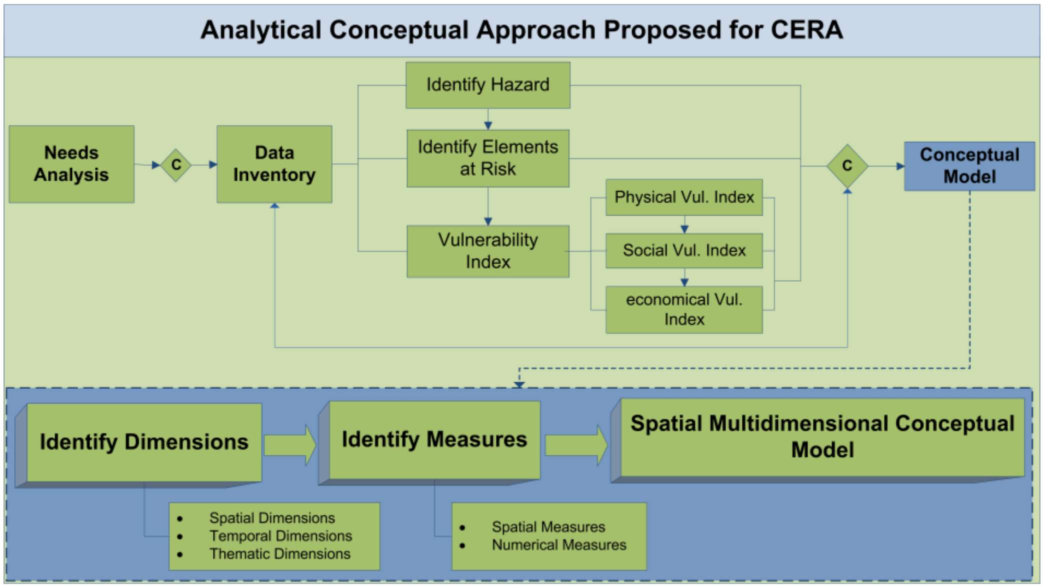


Figure 3.4: Analytical conceptual framework proposed for coastal erosion risk assessment.

Needs analysis is the analysis of a client's requirements to solve or facilitate problems or difficulties faced by an organization or a society. This calls for meetings or discussions with potential clients. As mentioned before, coastal regions are managed and used by diverse organizations and stakeholders under local, provincial, and national authorities. Transport ministry, urban planners, municipalities, fishery and ocean organizations, public security, natural resources, environment, and tourist industry are the potential users of a SDSS tool for CERA.

Data inventory permits identification of available datasets and defines how such data are currently being used by different stakeholders. The first step in data inventory is the selection of the regions that are undergoing severe erosion and are of interest to authorities and stakeholders. In the next step, available documents related to erosion sites such as maps, plans, aerial photographs, LiDAR data, technical reports, census data, economic values, or any other pieces of information are collected. However, it is vital to take into account the semantic of geographical entity (what), spatial aspect (where), temporal aspect (when), data format (raster, vector), data quality, spatial and temporal scale and resolution, reference system, and data accessibility i.e. monetary means (Larrivée 2011). Moreover, certain metadata standards such as ISO/TC211 international standards (ISO/TC211 2003) have to be respected to define characteristics of the spatial data and geographical features. Providing a deep insight of available data and distribution of features of interest in erosion sites plays an important role in facilitating the identification of the coastal erosion risk parameters. These factors directly influence the structure of a datacube for CERA.

Once the data inventory is completed, the next step is to identify coastal erosion risk elements. CER, or $R(T,t)$ is defined as a cross measuring of coastal vulnerability, or $V(T,t)$, onto intensity of hazard, or $H(T,t)$, which occurs on a time scale t (Blaikie et al. 2004). Elements at risk, T, are every element in the exposed regions with a recognized interest of society or organizations. Elements at risk link together two components of risk i.e. hazard and vulnerability (Daudé et al. 2009). The concepts of element at risk and time scale have been adapted from the work of Daudé et al. (2009) and Alexander (2000) and have been integrated into the definition provided by Blaikie et al. (2004). As a result, the following definition of risk is adapted in this work:

$$R(T,t) = H(T,t) \times V(T,t) \qquad \text{Eq.3.1}$$

Since each sub-step is subsequent to the previous steps, the more precise hazard, target, and vulnerability index are characterized; the more risk assessment results are realistic. The interaction between hazard, elements at risk and vulnerability in risk concept is illustrated in Figure 3.5.

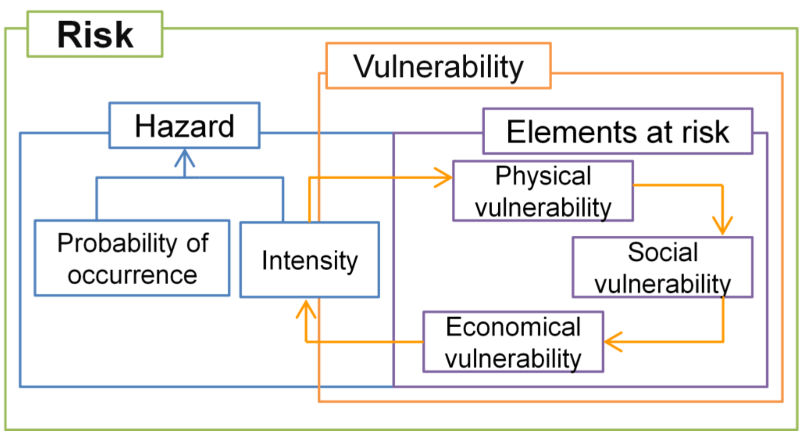


Figure 3.5: Risk elements i.e. hazard, vulnerability, targets and their interactions.

Identify Hazard (Coastal Erosion): Hazard, is defined as the probability of coastal erosion process occurrence and its intensity along the coast (Boruff et al. 2005). Coastal erosion is measured through the probability of physical removal of sediment along the coast either in short or long terms (Boruff et al. 2005). This probability and its intensity are functions of various factors including waves, currents, winds, tides, and storms as well as human-induced activities in different time periods on daily, seasonal, and yearly scales or even over a century (Morang & Szuwalski 2003). As stated before, various numerical and statistical methods exist to analyze coastal erosion. The coastline change rate is used in this study to accomplish CERA, while coastline change is extracted from DTM in multiple epochs. There are many tools or models to calculate the erosion rate such as Digital Shoreline Analysis System (Thieler et al. 2009) that is employed to obtain the erosion rate in this study.

Identify Elements at Risk: The identification of elements at risk is fundamental to elaborate the vulnerability index. There are a few generic methodologies to identify elements at risk. However, the need analysis and data inventory play a critical role in this context. Most of the existing studies in target identification have been performed for specific case studies that limit their application elsewhere. Examples of specific target-oriented studies are road networks (McHugh et al. 2006), and buildings close to the coast (Desprats et al. 2010; Totschnig et al. 2010). There are few studies on multiple interacting elements at risk. In this study, we consider multiple potentially elements at risk on or close to the coast. The elements at risk are classified with respect to different typologies regarding physical, social, and economical criteria. Examples are structures e.g. infrastructure and buildings along the coast; people at risk e.g. residents, employees, tourists; and land use classification e.g. urban, rural, agricultural, and industrial.

Vulnerability Index: Vulnerability describes both the measure of damage to elements at risk exposed e.g. people or infrastructures and the ability of a society or elements at risk to resist or recover from a disaster (Cutter et al. 2003). Vulnerability indicators (I_i) within a vulnerability index are categorized in environmental, social, and economic groups. The degree of damage or resistance of such indicators is described by $Rank(I_i)$ (Abuodha & Woodroffe 2006; Karvetski et al. 2011). The importance of these indicators to a society or organizations is considered as $\omega_i(T,t)$. This definition can mathematically be stated as follows:

$$V(T,t) = \sum_i^n Rank(I_i) \times \omega_i(T,t) \quad \text{Eq.3.2}$$

The quality of a vulnerability index is related extremely to experts' knowledge to characterize the susceptibility of exposed elements at risk with ranking score 1 to 5 and their importance. The reason to define the scores from 1 to 5 is associated with human feeling perception from a low sensible situation to a high sensible situation i.e. very low, low, medium, high, and very high. This standard is also employed by researchers to consider the risk degree (Boruff et al. 2005). On one hand, the ranking scores and weighting values are mainly derived from experimental studies on environmental criteria (Xhardé R. 2007) or statistical analysis such as Principal Component Analysis (PCA) for socio-economic indicators (Boruff et al. 2005). On the other hand, the choice of vulnerability indicators depends strongly on the data availability, the objectives of study, and the importance of the region under study from environmental and socio-economic aspects (Füssel 2010a). Since the scope of this paper is confined to developing a SMCM for CERA, the vulnerability index is adopted directly from (Xhardé R. 2007) and (Boruff et al. 2005). This index is also presented in Table 3.2. While the difference between social and economic indicators is the dollar-value associated to each of them, the dollar-value of vulnerable indicators is added up to estimate the total loss value. Loss is damage or harm to a natural habitat or built environment and structures, physical harm to people or a combination of them (Hessami 2004).

Table 3.2: Adapted vulnerability index for coastal erosion, their categories, and the ranking score of vulnerability indicators.

Category	**Indicators**	**Ranking of each indicator with respect to coastal erosion**				
		Rank 1*	**Rank 2***	**Rank 3***	**Rank 4***	**Rank 5***
Environmental indicators	Geology & geomorphology (type of coast)	Cliff, fjords beaches	Talus, stable beach (with vegetation)	Talus, and instable beach (without vegetation)	Beach	Delta, marsh, dune
	Coast Elevation (DEM)	> 25m	17-24m	11-17m	4-10m	0-3m
	Slop average	1-13 %	14-20 %	21-28 %	29-35 %	> 36 %
	Tide variation (m/year)	<-1	-1-0.99	1-2	2.1- 4	> 4
	Tide range (m)	<1	1-1.9	2-4	4.1-6	> 6
	Wave height	0- 2.9	3 - 4.9	5- 5.9	6- 6.9	> 6.9
	Hydrology and drainage network	Non-presence	----	----	----	Presence
	Distance between shore and depth of 5m	1001-1200m	701-1000m	700-401m	301-400m	< 300m
	Distance between coastline and the vulnerable object	> 61m	31-60m	21-30m	11-20m	0-10m
	Nature of geological structure	Absence of faults, fractures or subsidence	----	----	----	Presence of faults, fractures or subsidence
Socio-economic indicators	Land occupation	Park, forest, marsh, vegetation land	Rural zone	Mixed rural zone (rural and other)	Urban zone (residential)	Mixed urban zones (commercial, residential, industrial, school,

						church, and etc.)
	Protection structure (type, status, date of construction)	Yes, good state	----	Yes, but destroyed or need to repair	----	Non
	People at risk (age, sex, income, occupation, education)	Principal component analysis of census data				
	Tourism	Principal component analysis of census data				
	Structure (house, manufacture, built environment) and Infrastructure (road, railway, port, bridge, power transfer)	Principal component analysis of census data				

* Rank 1= very low, Rank 2= low, Rank 3=Medium, Rank 4=High, and Rank 5= Very high

3.7 Results: Development of Spatial Multidimensional Conceptual Model

As illustrated in Fig. 4, the design of a SMCM requires in the first step identifying key elements of the model, which are *dimensions* and *measures*. Here, dimensions constitute hazard and vulnerability, interested elements at risk, their interactions, and the location under study at a given time period.

3.7.1 Identify Dimension

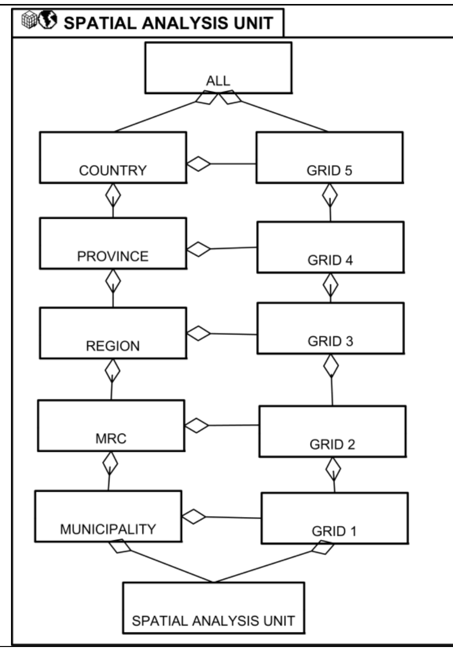
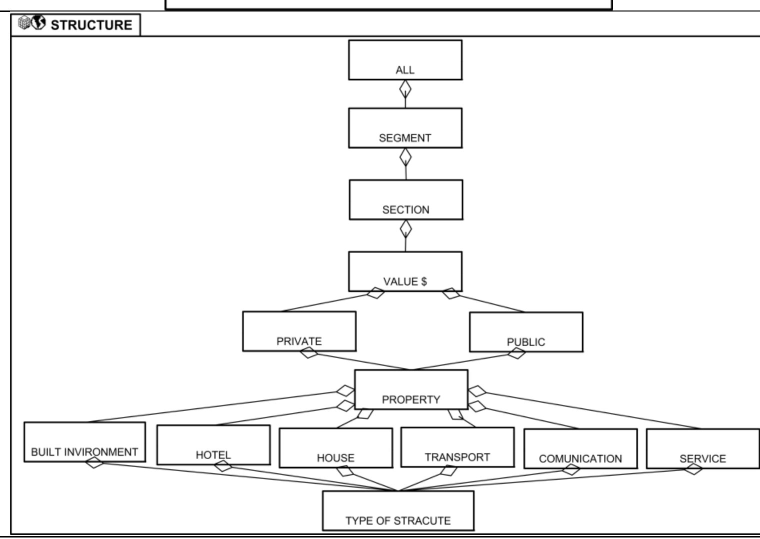
Spatial dimensions include two geographical features, “spatial analysis unit” and “structures” that are located along the coast regardless of their representation method i.e. by geometries or by text. The “spatial analysis unit” is stored as a grid cell and then connected to administrative boundaries while the “structures” are stored by different geometry primitives such as points, lines, and polygons depending on the scale of the representation and the level of granularity. For instance, a building may be considered a polygon on finer hierarchy or a point on a coarser level of hierarchy.

Primarily, “spatial analysis unit” is defined by employing a segmentation technique. A regular grid-cell is proposed in this work considering the typology of the elements at risk at any level of granularity which can capture the variability of risk components with an alternative hierarchy regarding with administrative boundaries (municipality, MRC, region, province, and country). The advantage of grid-based spatial unit is dividing space into continuous grid pixel along the coast that can facilitate the aggregation of homogenous cells with independent characteristics. This permits to aggregate and then represent the similar pixels independently from their initial discretization and provides a homogeneous aggregation from fine resolution to coarser level and vice versa. Members and hierarchy levels of “spatial analysis unit” dimension, as well as the UML (Unified Modeling Language) formal representation of hierarchy model are presented in Table 3.3.

Primarily, “spatial analysis unit” is defined by employing a segmentation technique. A regular grid-cell is proposed in this work considering the typology of the elements at risk at any level of granularity which can capture the variability of risk components with an alternative hierarchy regarding with administrative boundaries (municipality, MRC, region, province, and country). The advantage of grid-based spatial unit is dividing space into continuous grid pixel along the coast that can facilitate the aggregation of homogenous cells with independent characteristics. This permits to aggregate and then represent the similar pixels independently from their initial discretization and provides a homogeneous aggregation from fine resolution to coarser level and vice versa. Members and hierarchy levels of “spatial analysis unit” dimension, as well as the UML (Unified Modeling Language) formal representation of hierarchy model are presented in Table 3.3

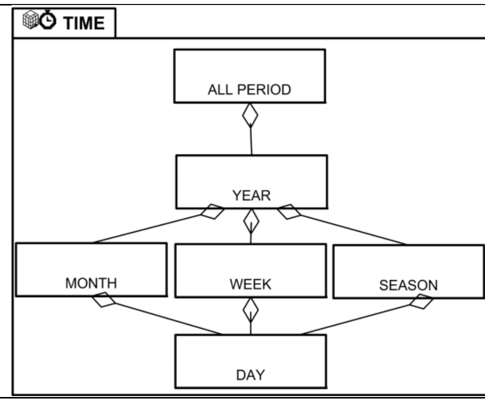
The “structure” dimension constitutes all infrastructures (transport networks, water supply networks, communication networks, etc.), buildings (houses, hotels, schools, hospitals, etc.), and built-environments (parks, zoos, aquariums, tourism sites, etc.) in the region at multiple levels of granularity. Targeting a specific structure in the database depends on the needs of stakeholders or users. For instance, the “structure” dimension permits looking into information cartographically or statistically at different levels of details. This information may be associated with possible socio-economic vulnerable structures and infrastructures as a vulnerable object e.g. a bridge (finest level of hierarchy), a section e.g. a part of road network consisting of bridges and roads (higher level of hierarchy) or a segment e.g. road segment consistent with National Road Network definition in “Geobase" (coarser level of hierarchy) regarding either the type of property or their dollar values.

Table 3.3: "Spatial" dimensions, their members, their hierarchies, and their formal representation in UML.

Dimensions	Members and Hierarchy of Dimension	UML Formal Representation of Hierarchy Model
Spatial analysis unit	Spatial analysis unit → Municipality→ MRC or County→ Region→ Province or State→ Country→all Spatial analysis units → grid1→grid2→grid 3→grid4→ grid5→ all	
Structure	Type (transport (road, railway, airport, port, etc.), communication network (water, electricity, etc.), service (school, library, etc.), house, hotel, built environment (national park, beach, etc.) → Property (public, private)→value$ →Section → Segment →all	

"Time", known as temporal dimension, should also be considered in the model in order to allow the decision-makers to look at the level of risk at a given time period e.g. day, week, month and year, through a hierarchy system. An innovative aspect of the proposed model in this work is related to an alternative hierarchy of "Time" dimension regarding the weather or tide variation. For instance, in addition to a calendar hierarchy e.g. day, week, month and year, a seasonal hierarchy of "Time" that depends on the weather variation of the region is also included i.e. early spring (snow melting period), late spring (without snow), summer, early fall, late fall, and winter. The seasonal hierarchy is not an aggregation of months, but date to date. This aspect enriches the proposed model to investigate the risk degree and its consequence at high and low rates of erosion seasonally throughout the years. This can play an important role within a strategic period for decision-makers. An illustrative schema for “Time” dimension is provided in Table 3.4.

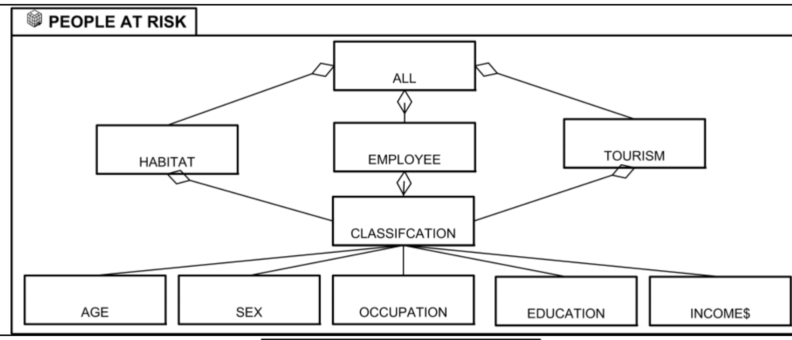
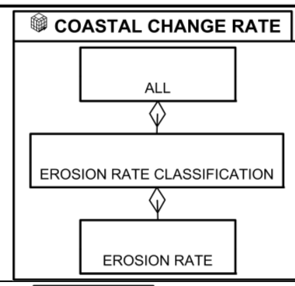
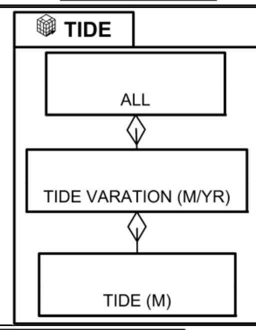
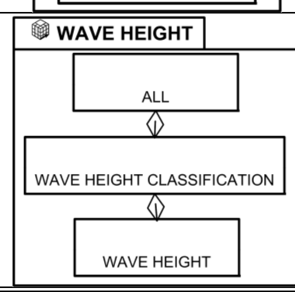
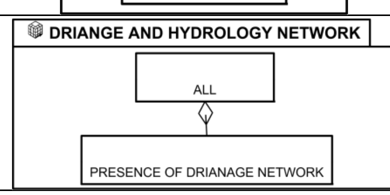
Table 3.4: “Time” dimension, its members, its hierarchies and its formal representation in UML.

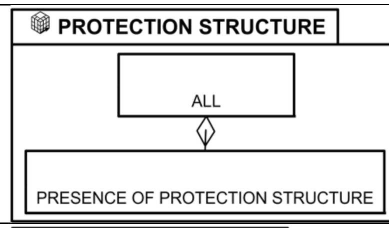
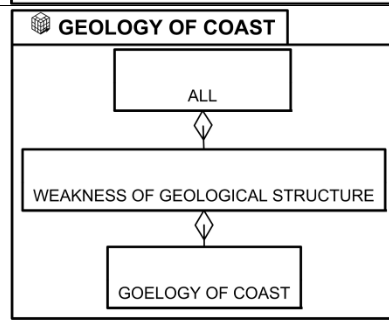
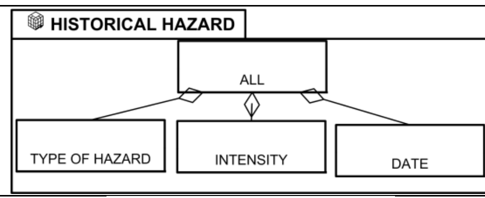
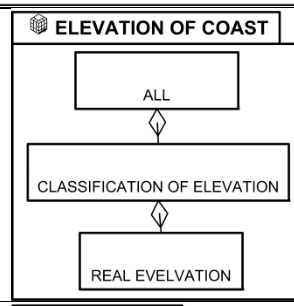
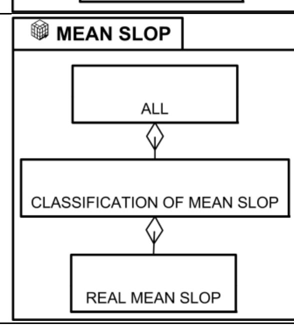
Dimension	**Members and Hierarchy of Dimension**	**UML Formal Representation of Hierarchy Model**
Time	Day→ Week→ Month→ Year→all period Day→Season (Early Spring, Late Spring, Summer, Early Fall, Late Fall, Winter) →Year→all period	

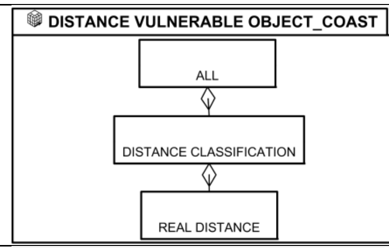
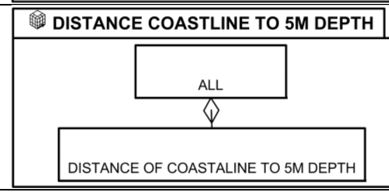
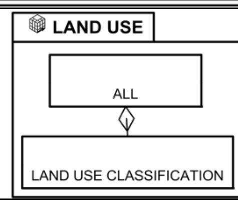
Thematic dimensions consist of sets of criteria and their attributes corresponding to the vulnerability index with a possible multiple levels of granularity for each indicator that are derived from Table 3.2. Examples are “people at risk”, “coastline change rate”, “tide”, “wave height”, “hydrology network”, “protection structure”, “geology of coast”, “historical hazard”, “elevation of coast”, “mean slope”, “land use”, “distance between coastline and the depth of 5-meter”, and “the distance of vulnerable object from the coast”. Members and hierarchy levels of each dimension as well as the UML formal representation of hierarchy model are presented in Table 3.5.

In thematic dimensions, “People at risk” is of particular interest defined based on census unit division. Non-consistency of the census and spatial analysis units is an issue in the implementation stage. An appropriate solution to extract the number of people at risk in a spatial analysis unit is using an ecumene map of the inhabited regions under study. The ecumene regions are lands where people have made their permanent home, and to all work areas that are considered occupied and used for agricultural or any other economic purposes (StatisticCanada 2012). Otherwise, counting the number of houses, hotels, schools, hospitals, etc. situated on each spatial analysis unit may provide the approximate number of people in the region.

Table 3.5: "Thematic" dimensions, their members, their hierarchies, and their formal representation in UML.

Dimensions	Members and Hierarchy of Dimension	UML Formal Representation of Hierarchy Model
People at risk	Age Sex Education Occupation Income \$ Classification (habitat, employee, tourism) →all	
Coastline change rate (m/yr)	> +0.1 accretion → 0 stable → -0.1 to -0.5 erosion (average) → -0.6 to -1.0 erosion (high) → > -1.0 erosion (very high)→ all	
Tide	Tide range (m) <1 →1-1.9 →2-4→4.1-6→ > 6 Tide variation (m/year) <-1 →-1-0.99 →1-2 →2.1-4 →>4→ all	
Wave height	Real value of wave height → Wave height classification→ all	
Hydrology network	Presence of hydrology network (Yes or No)	

Protection structure	Type (good state, destroyed and need to repair, not available)→all Date of construction	
Geology of Coast	Type of geology (Cliff, fjords beaches; Talus, stable beach; Talus, and instable beach; Beach; Delta, marsh, dune)→ Weakness of geological structure (Absence of faults, fractures or subsidence; Presence of faults, fractures or subsidence) → all	
Historical hazard	Intensity of hazard →all Date → all Type of Hazard → all	
Elevation of coast (DEM)	Real value of elevation → Classification→all	
Mean slop of coast	Real value of mean slop → Classification→all	

Distance of vulnerable object to coast	Real distance of object to coastline → Classification→all	
Distance between shore and depth of 5m	Distance of coastline to 5m depth→all	
Land use	Type of land use→all	

The hierarchy model of each dimension as a formal representation in UML allows the user to define the possible measures in any required combination of the dimensions and levels of aggregation by drilling down or rolling up in the hierarchies. In fact, the hierarchy model provides a significant insight into the whole system and lets the user navigate between dimensions and measure levels. Different types of hierarchy models employed in this work involved strict hierarchy, such as the most of thematic dimensions and non-strict hierarchy, the “spatial analysis unit”, “structures”, “time”, and “people at risk” dimensions. More details on the types of hierarchies and their implementation methods are provided in Malinowski and Zimanyi (2006).

3.7.2 Identifty Measures

Since the proposed SMCM is developed to perform CERA, the CER is considered as a measure variable. The value of CER is computed using Eq. 1 through the intersection of the spatial, temporal, and thematic dimensions. The risk level is expressed by values between 1 and 5 where 1 corresponds to very low and 5 to very high level of risk. These measures can be expressed either spatially or numerically is presented in Table 3.6.

Table 3.6: The list of potential measures based on developed SMCM

Measures	
Spatial Measures	Level of risk for all vulnerable elements at risk in any level of detail with respect to one or several dimensions in a particular region and time period with a priority of distance from the coastline.
	Level of risk for all vulnerable elements at risk in any level of detail with respect to one or several dimensions in a particular region and time period with a priority of mean slope of the region.
	Level of risk for all vulnerable elements at risk in any level of detail with respect to one or several dimensions in a particular region and time period with a priority of geology type of the coast.
	Level of risk for any physical feature in any level of detail in a particular region and time period.
	Level of risk for any socio-economic feature in any level of detail in a particular region and time period.
	Overall risk degree for any physical feature in any level of detail with respect to one or several dimensions in a particular region and time period.
Numerical Measures	Number of spatial analysis units in any level of detail in a particular region and time period.
	Number of people at risk in any level of detail in a particular region and time period.
	Number of structures in any level of detail in a particular region and time period.
	Size of structures under severe erosion risk, e.g. road or railway at risk.
	Loss of value in any level of detail in a particular region and time period.

3.7.3 Formal Presentation of Spatial Multidimensional Conceptual Model

The proposed SMCM is developed based on a star schema model illustrated in Figure 3.6. The star schema graphically represents the end-user's perception of how the information can be accessed. The main advantages of this schema are direct and intuitive mapping between the features of interest being analyzed by end-users and the schema design, highly optimized performance for typical star queries, and widely supported by a large number of BI tools (Chaudhuri et al. 2011). The star schema is perhaps the simplest multidimensional data modeling technique. Principally, it is composed of a single fact table in the center linked with a line to a set of dimension tables as a star with several granularity levels (Bédard and Han 2009). The fact table contains the measures and one foreign key per dimension to link the fact with the dimension's member (Bédard and Han 2009). In a star implementation, each dimension is stored in one table, independent of the member's hierarchical level and identified by a primary key or ID that serves as a foreign key to join to the fact table (Kimball and Ross2002).

There are many possibilities for building datacubes based on the developed conceptual models such as Relational OLAP (ROLAP), Multidimensional OLAP (MOLAP), or a combination of both, Hybrid OLAP (HOLAP) within the SOLAP system (Imhoff et al. 2003). The implementation of SMCM in the form of a physical spatial datacube is out of the scope of this paper.

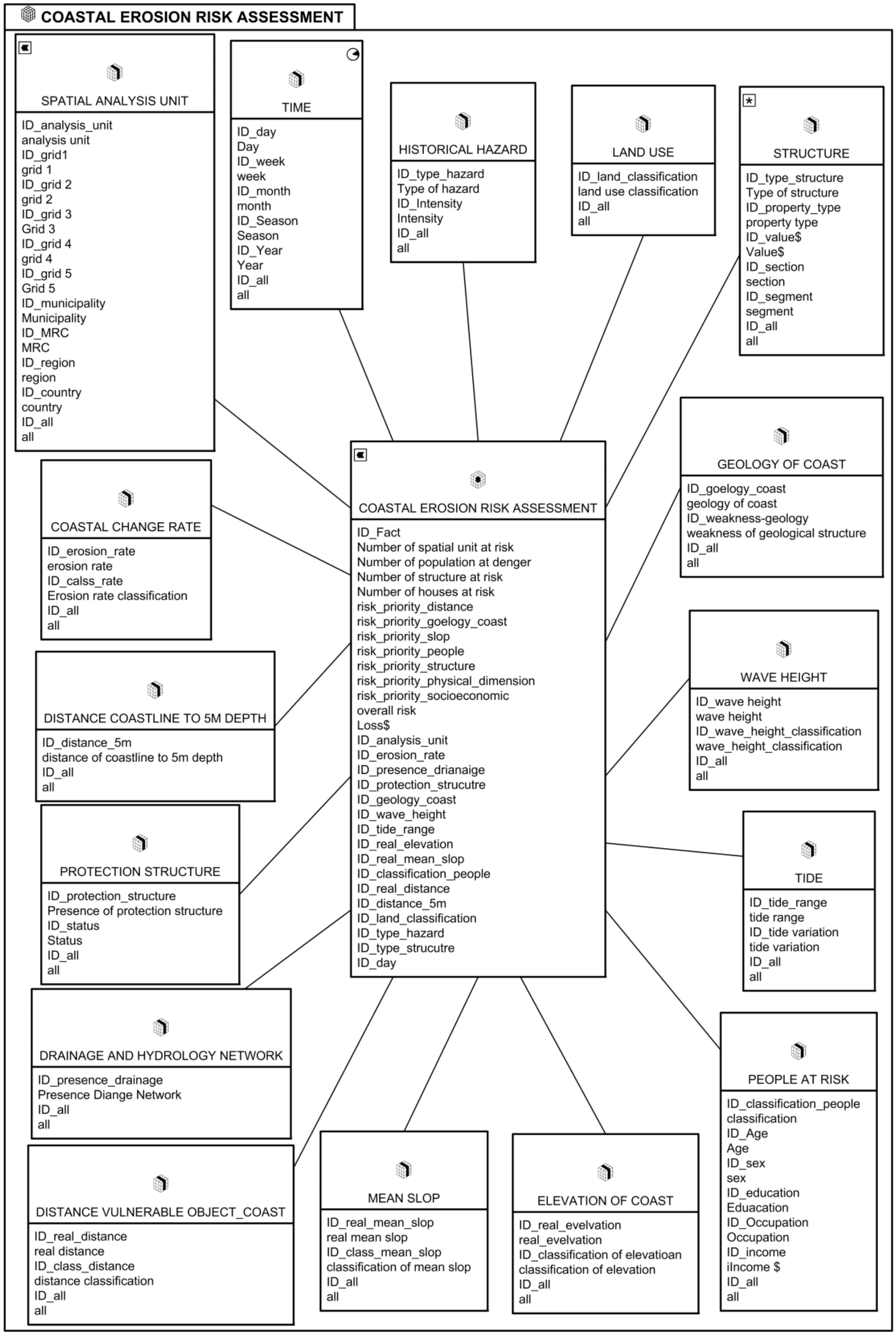

COASTAL EROSION RISK ASSESSMENT
SPATIAL ANALYSIS UNIT
ID_analysis_unit
analysis unit
ID_grid1
grid 1
ID_grid 2
grid 2
ID_grid 3
Grid 3
ID_grid 4
grid 4
ID_grid 5
Grid 5
ID_municipality
Municipality
ID_MRC
MRC
ID_region
region
ID_country
country
ID_all
all
TIME
ID_day
Day
ID_week
week
ID_month
month
ID_Season
Season
ID_Year
Year
ID_all
all
HISTORICAL HAZARD
ID_type_hazard
Type of hazard
ID_Intensity
Intensity
ID_all
all
LAND USE
ID_land_classification
land use classification
ID_all
all
STRUCTURE
ID_type_structure
Type of structure
ID_property_type
property type
ID_value$
Value$
ID_section
section
ID_segment
segment
ID_all
all
GEOLOGY OF COAST
ID_goelogy_coast
geology of coast
ID_weakness-geology
weakness of geological structure
ID_all
all
COASTAL CHANGE RATE
ID_erosion_rate
erosion rate
ID_calss_rate
Erosion rate classification
ID_all
all
COASTAL EROSION RISK ASSESSMENT
ID_Fact
Number of spatial unit at risk
Number of population at denger
Number of structure at risk
Number of houses at risk
risk_priority_distance
risk_priority_goelogy_coast
risk_priority_slop
risk_priority_people
risk_priority_structure
risk_priority_physical_dimension
risk_priority_socioeconomic
overall risk
Loss$
ID_analysis_unit
ID_erosion_rate
ID_presence_drianaige
ID_protection_strucutre
ID_geology_coast
ID_wave_height
ID_tide_range
ID_real_elevation
ID_real_mean_slop
ID_classification_people
ID_real_distance
ID_distance_5m
ID_land_classification
ID_type_hazard
ID_type_strucutre
ID_day
DISTANCE COASTLINE TO 5M DEPTH
ID_distance_5m
distance of coastline to 5m depth
ID_all
all
WAVE HEIGHT
ID_wave height
wave height
ID_wave_height_classification
wave_height_classification
ID_all
all
PROTECTION STRUCTURE
ID_protection_structure
Presence of protection structure
ID_status
Status
ID_all
all
TIDE
ID_tide_range
tide range
ID_tide variation
tide variation
ID_all
all
DRAINAGE AND HYDROLOGY NETWORK
ID_presence_drainage
Presence Diange Network
ID_all
all
PEOPLE AT RISK
ID_classification_people
classification
ID_Age
Age
ID_sex
sex
ID_education
Eduacation
ID_Occupation
Occupation
ID_income
iIncome $
ID_all
all
DISTANCE VULNERABLE OBJECT_COAST
ID_real_distance
real distance
ID_class_distance
distance classification
ID_all
all
MEAN SLOP
ID_real_mean_slop
real mean slop
ID_class_mean_slop
classification of mean slop
ID_all
all
ELEVATION OF COAST
ID_real_evelvation
real_evelvation
ID_classification of elevatioan
classification of elevation
ID_all
all

Figure 3.6: Formal presentation of spatial multidimensional conceptual model for CERA

For the estimation of the measures, a star-query model that is a common technique in star schema modeling is proposed in this work. The star query model is also known as a junction between a fact table and a number of dimension tables. A simplified version of the SMCM developed in this paper is illustrated in Figure 3.7. Each point on the axis presents a level of hierarchy for each dimension. It is clearly demonstrated in this figure how the levels of dimensions are interconnected and the measures are computed. For instance, the degree of risk is calculated using Eq. 3.1, as elaborated in Eq. 3.3, on a specific spatial unit *a* regarding a road section *b* (target) located on high erosion rate *c* with the elevation classification *d,* geology type *e*, *f* meter from the coastline, in an urban area *g*, and on a given year *h*.

$$R(a,h) = \frac{1}{6}\left[c(a,h)\times((b\times\omega_b)+(d\times\omega_d)+(e\times\omega_e)+(f\times\omega_f)+(g\times\omega_g))\right] \quad \text{Eq. 3.3}$$

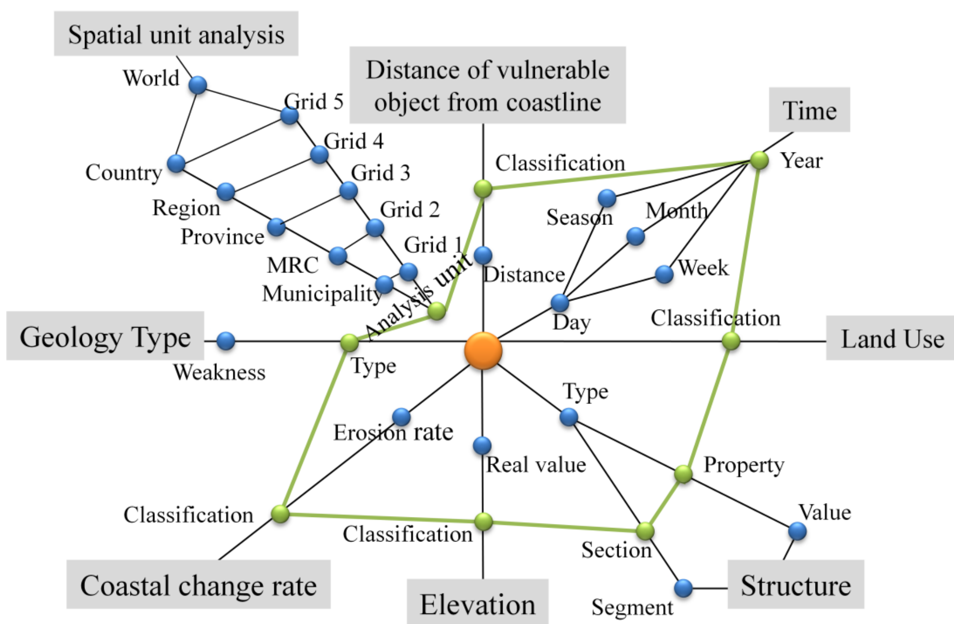


Figure 3.7: Star query model of simplified SMCM

3.8 Discussion

This paper aims to present a generic analytical framework to elaborate a spatial multidimensional conceptual model for coastal erosion risk assessment. This model takes into account multiple-categories of potentially vulnerable features and elements at risk as well as their interactions to compute the coastal erosion risk in hazardous regions at a given time.

The identification of dimensions and their hierarchies in the proposed conceptual multidimensional model is based on the vulnerability index presented in Table 3.2 resulted from experts' knowledge and empirical practices. However, the proposed model can easily be elaborated based on other vulnerability indices through the proposed framework and adapted to other coastal regions.

In the proposed SMCM, Eq.3.1 is evaluated with respect to seven different priorities within the fact table to provide some pre-calculated measures for users. The evaluated priorities are the distance of vulnerable features from the coastline, geology type, mean slope, people density, structures, physical and socio-economic vulnerable features (see Figure 3.6). This provides the users the possibility to select one or several elements at risk and represents the degree of risk at different levels of details and time periods.

As stated before, the star query model is used to execute pre-defined or user-defined measures; it is consistent with the users' perception while a vast majority of BI tools support star schema and star query models. Some typical examples of the queries that can be carried out by a client using the proposed model for CERA are presented in Table 3.7.

Table 3.7: Some typical examples of complex queries that can be executed within developed SMCM.

Examples of Complex Queries
Representing the degree of risk via a map or diagram with respect to geology of the region, the density of people at risk, and the structures that are located less than *x* meter from the coastline, for a high erosion rate at municipality *y* where the elevation of the coast is more than *z* meter.
Representing the same results, but this time on a finer scale of spatial analysis unit to observe vulnerable objects whose distance from the coastline is *x* to y where the protection structures exist but are destroyed.
Estimating the number of public structures which are affected by the high erosion rate at MRC *x*.
Computing the length of roads which are located within high risk zones.
Representing the risk zones with a priority of *i* indicator which is mentioned as a measure.

The flexibility of the proposed model to select the desired time period as a calendar or defined season, with respect to the weather variation and environmental factors, is also an enrichment of the proposed model. In addition, the grid-based "spatial analysis unit" and interconnecting with administrative boundaries is another richness of the SMCM. This allows the stakeholders and decision-makers to make strategic decisions or take actions at the right period and in the right place in order to protect the region at risk.

Despite the advantages of using a multidimensional paradigm to assess risk in coastal areas, SMCM development has its own challenges. The proposed model in this paper is based on the data inventory in the

Gaspe region, Quebec, Canada. The main data sources in this study are LiDAR data (INRS-ETE Quebec Canada), Geobase databases (Natural Resources Canada), Census data (Statistics Canada), and BGR database (Ministry of Transport Quebec). One issue is data availability, which has an impact on the structure of the resulting model. By data unavailability we mainly mean 1) the lack of data in the required details in spatial, temporal or thematic aspects, 2) the lack of an integrated vulnerable index in the region under study, and 3) the non-consistency of census parcel with spatial analysis units to calculate the precise number of people at risk. Since the census units are not identical to the spatial analysis units, the estimated number of people in each analysis unit is affected by an uncertainty degree. Possible solutions are performing ecumene maps (if available) or counting the number of habitats, houses, or buildings. Nevertheless, it does not mean that the proposed approach cannot be employed in the regions with considerable limited data. Though, the proposed analytical conceptual framework in this paper is still effective in any region to develop an appropriate spatial multidimensional model.

Available information technology also imposes some limitations regarding the efficient computation in datacubes. Since a multidimensional database consists of numerous dimensions with multiple levels of granularity, there exist a large number of possible combinations of dimensions and levels each of which forms an aggregated multidimensional cube called a cuboid (Bédard et al. 2009). Managing high numbers of cuboids in a user-friendly interface is a challenge. In addition, technological constraints may limit the number of dimensions in a datacube. An intermediate solution is to integrate dimensions with the same typology. Examples are integrating “geology type” with “the presence of weaknesses in geological zones” or “mean tide variation (mm/yr.)” with “tide value (mm/day)”. Moreover, all infrastructures, buildings, and built-environments can be integrated into the same spatial dimension “structure”.

Determining the form and the size of the grid cell in spatial analysis units remain also a challenge in CERA. The size of the grid is generally defined considering empirical study and the resolution of the available data. The degree of hazard and the distribution of vulnerable features also play an important role in selecting the unit form and size.

Managing uncertainties in CERA is also the issues to be taken into account in future works. Uncertainties stem from data itself, the way spatial analysis units are defined, and identified risk elements as well as their interactions.

3.9 Conclusion

This paper outlined the development of a comprehensive integrated system for coastal erosion risk assessment. First, the state of the art in prevailing GIS-based CERA methods was presented. The arguments were provided to indicate the limits of the existing methods including the lack of an integrated system. An analytical conceptual framework was proposed to overcome aforementioned limits by accomplishing a comprehensive CERA system through the geospatial BI paradigm. The integration of information from different stakeholders and extracting different dimensions of risk assessment through multiple criteria with different levels of details can be extremely facilitated through geospatial BI. This also allows aggregation of qualitative and quantitative information and knowledge to estimate the risk in the several levels of hierarchies. A SMCM was developed for CERA through the proposed framework. The proposed model provides a complete and coherent vision of the CER phenomenon by integrating different spatial, temporal, and thematic dimensions. It performs the cross-measuring of the information to estimate the CER value at a given time period onto a grid-based spatial unit interconnecting with administrative boundaries. Several aspects of the proposed SMCM can be enhanced. Two immediate improvements are through applying an appropriate segmenting method to analysis units and assigning uncertainty propagation impact.

"There is nothing worse than a sharp image of a fuzzy concept."
Ansel Adams

Chapter 4 Spatial Representation of Coastal Risk: A Fuzzy Approach to Deal with Uncertainty

In process at ISPRS International Journal of Geo-Information, 2013,
Jadidi A., Mostafavi M.A., Bedard Y., Shahriari K.,

Keywords: *Uncertainty; Fuzzy Set Theory; Coastal Erosion Risk Assessment; Spatial Representation; Fuzzy Object*

4.1 Preface

This chapter is composed of the second paper published within the framework of the present thesis which is submitted to the ISPRS International Journal of Geo-Information. The paper is focused on the second objective of the thesis which proposes to develop a novel approach based on Fuzzy Set Theory to improve the representation of uncertainty inherent to risk zones representation. First, the spatial uncertainty related the risk assessment and representation are defined. Then, a fuzzy based approach is proposed to deal with inherent spatial uncertainty in the risk zones representation. Here, the conceptual model presented in chapter 3 on risk components (i.e. hazard, elements at risk and vulnerability index) is used to apply the proposed fuzzy logic based approach for risk representation.

4.2 Abstract

Spatial information for coastal risk assessment is inherently uncertain. This uncertainty may be at different spatial, semantic, and temporal components of spatial data and can be expressed either quantitatively or qualitatively. Uncertainty in coastal risk assessment itself arises from poor spatial representation of risk zones.

Indeed, coastal risk is inherently a dynamic, complex, scale-dependent, and vague phenomenon in concept. In addition, representing the associated zones with polygons having well-defined boundaries does not provide a realistic manner for efficient and accurate representing of the risk. This paper proposes a conceptual framework based on fuzzy set theory to deal with the problems of ill-defined boundaries of risk zones and of its inherent uncertainty issues. To do so, the nature and level of uncertainty as well as the way to model it is characterized. Then, a fuzzy representation method is developed where the membership functions are derived based on expert-knowledge. The proposed approach is then applied in the Perce region (Eastern Quebec, Canada) and results are presented and discussed.

4.3 Introduction

Characterizing the uncertainty associated with assessed risk and representing the results together with risk value has a direct impact on decision-making processes (Darbra et al. 2008; Kentel & Aral 2007; Xie et al. 2011; Dolan & Walker 2006). In fact, a large amounts of spatiotemporal data, either certain or uncertain, from multiple sources with different levels of detail are used in risk assessment (Jadidi et al. 2013). In this regard, realistic decisions can't be made without knowing both the value of the risk and the respective modulated uncertainty.

In risk assessment processes, spatiotemporal data is classified based on expert knowledge, through a vulnerability index (Boruff et al. 2005; Jadidi et al. 2013). Spatial uncertainty is correspondingly defined as the lack of knowledge regarding the true value of a parameter or an attribute of information (Walker et al. 2003; Zadeh 2005). This uncertainty has both qualitative and quantitative natures, which may consist of multiple dimensions. In modeling of natural phenomena such as coastal erosion risk assessment, uncertainty often appears as imperfect knowledge. This imperfect knowledge includes vagueness in boundary zones, ambiguities in linguistic terms, fuzziness in semantics of spatial objects, and a mix of these that pertain ontologically only to the fiat world (Fisher et al. 2010; Smith & Varzi 2000). This brings us to the realm of spatial modeling by the concept of bona fide (spatial object models with well-defined crisp boundaries) and fiat object models (spatial object models with broad boundaries) (Smith & Varzi 2000; Smith & Mark 2003). The fiat objects refer to boundaries induced through human demarcation, qualitative differentiation, or spatial discontinuity, for example river, mountain, coastline, and risk zone (Fisher et al. 2010; Smith & Varzi 2000). Indeed, fiat objects are conventionally approximated in the databases to be represented like bona fide objects. Representing risk zones by polygons with well-defined boundaries is an example of such approximation. These polygons are created using aggregations of a set of spatial units defined based on either the stakeholders' interests or national census divisions (McFadden et al. 2007; Jadidi et al. 2013). Despite spatiotemporal variation of the multiple criteria involved at risk extent, each polygon has a unique value of risk attributed homogeneously over its spatial extent (Cheng et al. 2009). In reality, risk value changes gradually from one polygon to another (Cheng et al. 2009). The transition from one zone to another is not therefore properly represented with bona fide (crisp) object models. Therefore, the main dimension of uncertainty in Coastal Erosion Risk Assessment (CERA) arises from this poor spatial representation of risk that is related to spatial object modeling of risk zones.

Two main approaches are widely employed to characterize spatial uncertainty associated with risk modeling and representation. These are probabilistic and possibilistics approaches (Aerts et al. 2003; Choa et al. 2003; Cowell & Zeng 2003; Darbra et al. 2008; Fisher et al. 2007; Kentel & Aral 2007). Considering the nature of uncertainty, whether probabilistic or possibilistics, either approach can be employed (Walker et al. 2003). However, the flexibility of the possibilistics approach in dealing with uncertainties related to spatial object modeling suggests that it can be an efficient solution for spatial representation of risk (Kentel & Aral 2007). The possibilistics approach consists of exact models, rough models, and fuzzy models (Pauly & Schneider 2010; Schneider 2003a; Schneider 2003b; Dilo et al. 2007; Cheng et al. 2009; Kanjilal et al. 2010; Bejaoui et al. 2008; Cohn & Hazarika 2001). Extensive studies have been done on implementation of exact and rough models, i.e. an extension of crisp spatial data model, with a few references on fuzzy model by means of fuzzy set theory (Pauly & Schneider 2010; Cheng et al. 2009; Kanjilal et al. 2010). For instance, a spatial data model based on the possibilistics approach has been developed by Schneider (Schneider 2003a), called vague spatial data model based on exact approach and is extended by Kanjilal et al. (2010) and Pauly & Schneider (2010), called Vague Spatial Algebra (VASA). These models are simple to use and consistent with crisp spatial data model which is well known by geospatial communities. However, these models have their inherent limitations to represent the phenomena when it is impossible to determine their sharp boundaries (Molenaar & Cheng 2000). Moreover, if a fine-grained modeling of objects is demanded, the exact and rough models are not well supported (Pauly & Schneider 2010). Also, continuous change of features is not completely modeled by three-value logic of exact and rough methods such as modeling of pollution zones, flood zones, erosion zones, etc. In other words, the extents of these spatial objects cannot be bounded by precise boarders (Fisher et al. 2005; Fisher et al. 2007; Pauly & Schneider 2010). Nevertheless in fuzzy models, the degree of broadness of spatial object boundaries is expressed by an assigned membership function. Fuzzy models allow a continuous change of the degree of spatial uncertainty in the interior of such objects based on multi-value logic of fuzzy set theory. This is the main motivation behind exploring fuzzy models to represent a continuous phenomenon such as coastal erosion risk.

The main objective of this paper is to develop a generic fuzzy approach for spatial representation of risk zones. First, the nature and level of spatial uncertainty in CERA is characterized. Fuzzy models based on fuzzy set theory are then explored to handle such uncertainty. To do so, a conceptual framework is proposed to deal with the problem of spatial uncertainty of risk zones using membership functions. Instead of determining the bona fide boundaries between the risk zones, the proposed approach permits a smooth transition from one zone to another. Membership functions are derived from the knowledge of experts (e.g. from vulnerability index). The membership values of multiple indicators are then aggregated based on risk formula and Fuzzy IF-THEN rules to represent risk zones. The proposed approach is applied to Perce region (Eastern Quebec, Canada) as a case study.

The remainder of this paper is organized as follows: in Section 4.4, the background on the spatial representation of the risk and the nature of spatial uncertainty in CERA is provided. In Section 4.5, the fundamental concepts related to spatial fuzzy representation are explained. A novel conceptual framework to achieve a fuzzy spatial representation of risk is described in Section 4.6 while results and key lessons of this study are explained in Sections 4.7 and 4.8 respectively. Finally, the conclusion and perspectives for future work are provided in Section 4.9.

4.4 Background

4.4.1 Spatial Representation of Coastal Erosion Risk

CERA includes identifying hazard, detecting elements at risk (roads, houses, and people), elaborating vulnerability index, and characterizing the associated uncertainty (Jadidi et al. 2013; Darbra et al. 2008). However, this process is not as straightforward as one might imagine. The difficulties in CERA are especially due to the dynamic, complex, scale-dependent, and vague nature of erosion in coastal zones (Cheng et al. 2009).

Principally, spatial phenomena are modeled and represented either as discrete-object data models (vector data structure: point, line, and polygon) or continuous-field data models (raster data structure) (Goodchild & Glennon 2008). The extent of a spatial object is defined by its attribute and its geometry (shape and position) (Goodchild & Glennon 2008) as illustrated in Figure 4.1. Spatial representation of risk requires mathematical aggregation of vulnerability indicators (derived from elements at risk) and hazard maps in different time periods (Vanneuville et al. 2005). Risk is hence a spatial relation between hazard, elements at risk and vulnerability (Jadidi et al. 2013).

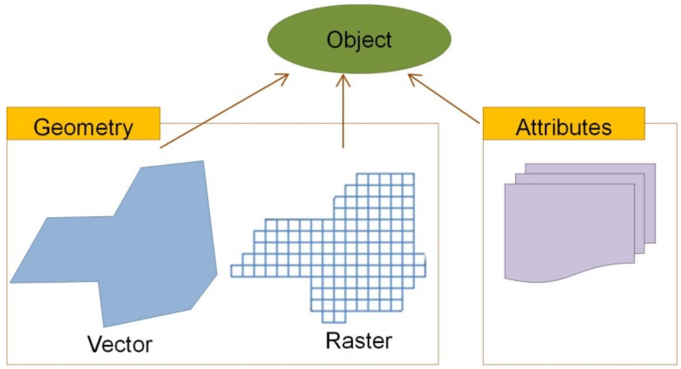


Figure 4.1: Representation of a spatial object, its geometry and attributes through vector and raster data structures

Current approaches for the representation of coastal risk are mostly based on using polygons with well-defined boundary. The boundaries are defined according to the stakeholders' interests or census division along the coast (Figure 4.2). The risk level is then attributed homogeneously within these units. A set of these units is aggregated to form risk zones. However, in reality, the degree of risk changes gradually from one point to the other e.g., it decreases when we get farther from the coastline. If bona fide (crisp) boundaries are applied to define the risk zones, the transition from one zone to another one is sudden and sharp which is not an appropriate representation of the reality in most cases. Though, the representation of natural phenomena (such as erosion risk) is often deliberated since their definitions are fiat and vague (Fisher et al. 2007; Smith & Varzi 2000; Smith & Mark 2003).

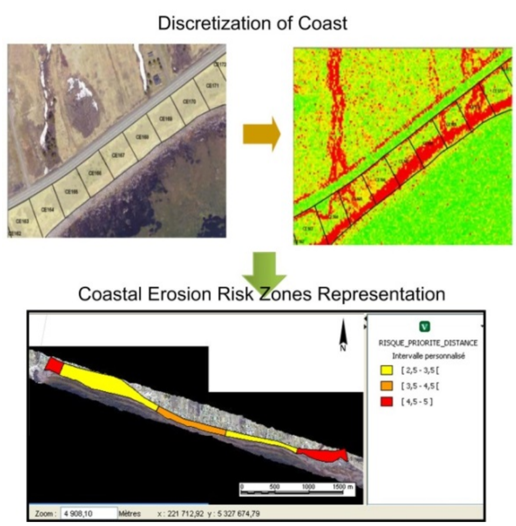


Figure 4.2: An example of coastal erosion risk representation (McHugh et al. 2006)

On the other hand, the technological capability of Geographical Information Systems (GIS) in integrating uncertainty in spatiotemporal modeling and representation is still a great challenge. The main limit of current

GIS tools is representation of complex fiat objects with uncertain boundaries (Cheng et al. 2009; Darbra et al. 2008; Kentel & Aral 2007; Pauly & Schneider 2010). Indeed, conventional GIS models geographical entities using a bona-fide concept with well-defined geometry such as points, lines, polygons in Euclidean space whereas the spatial representation of risk zones needs to deal with fiat objects with broad boundaries. The following sections seek out the different aspects of uncertainty in spatial representation of a coastal risk zone and then the approaches to handle it properly.

4.4.2 Uncertainty Characterization

The uncertainties associated with coastal risk assessment mainly originate from data (here, to estimate risk) and object definition (here, risk zones) (Cheng et al. 2009; Walker et al. 2003). A comprehensive scheme of uncertainty in spatial data modeling and the methods to handle it is illustrated in Figure 4.3. Uncertainty from data includes sampling and measurement errors that are categorized as epistemic uncertainty (Fisher et al. 2010). This type of uncertainty is random in nature and the probability theory can be applied to handle it appropriately (Fisher 2008). The uncertainty from object definition is characterized as ontological uncertainty that is related to the semantics of the object and its geometry (Cheng et al. 2009; Fisher et al. 2010). This refers to not only imperfect knowledge, but also lack of knowledge about an object or a phenomenon (Cheng et al. 2009; Walker et al. 2003; Dilo 2006).

Bédard (1988) introduced uncertainty into four levels of spatial data modeling: conceptual (when identification of entity classification and its existence are fuzzy or imprecise), descriptive (when the definition of an attribute value is fuzzy or imprecise), positioning (spatiotemporal aspect of an observed reality is fuzzy or imprecise), and meta-uncertainty (unknown degree of the preceding uncertainties). The spatial uncertainty associated with CERA is the combination of all these levels with a degree of vagueness, fuzziness, and ambiguity during the definition of risk zone objects. The intrinsic nature of coastal risk, i.e. the continuity, heterogeneity, dynamics and scale-dependence (Cheng et al. 2009), includes both epistemic and ontological aspects of uncertainty in CERA. The vagueness, fuzziness and ambiguity that are the principal factors in semantics uncertainty (Fisher et al. 2010), result from the continuity and heterogeneity of the risk zones and the scale issues in risk analysis respectively (Fisher 2008; Robinson 2003). Vagueness and fuzziness refer to the gradual transition boundaries of the risk zones in space and the impossibility of determining these boundaries (Dilo et al. 2007; Fisher 2008; Fisher et al. 2010). Ambiguity is related to discord and non-specificity description of classified vulnerability index to calculate risk degree in a risk zones (Fisher et al. 2010).

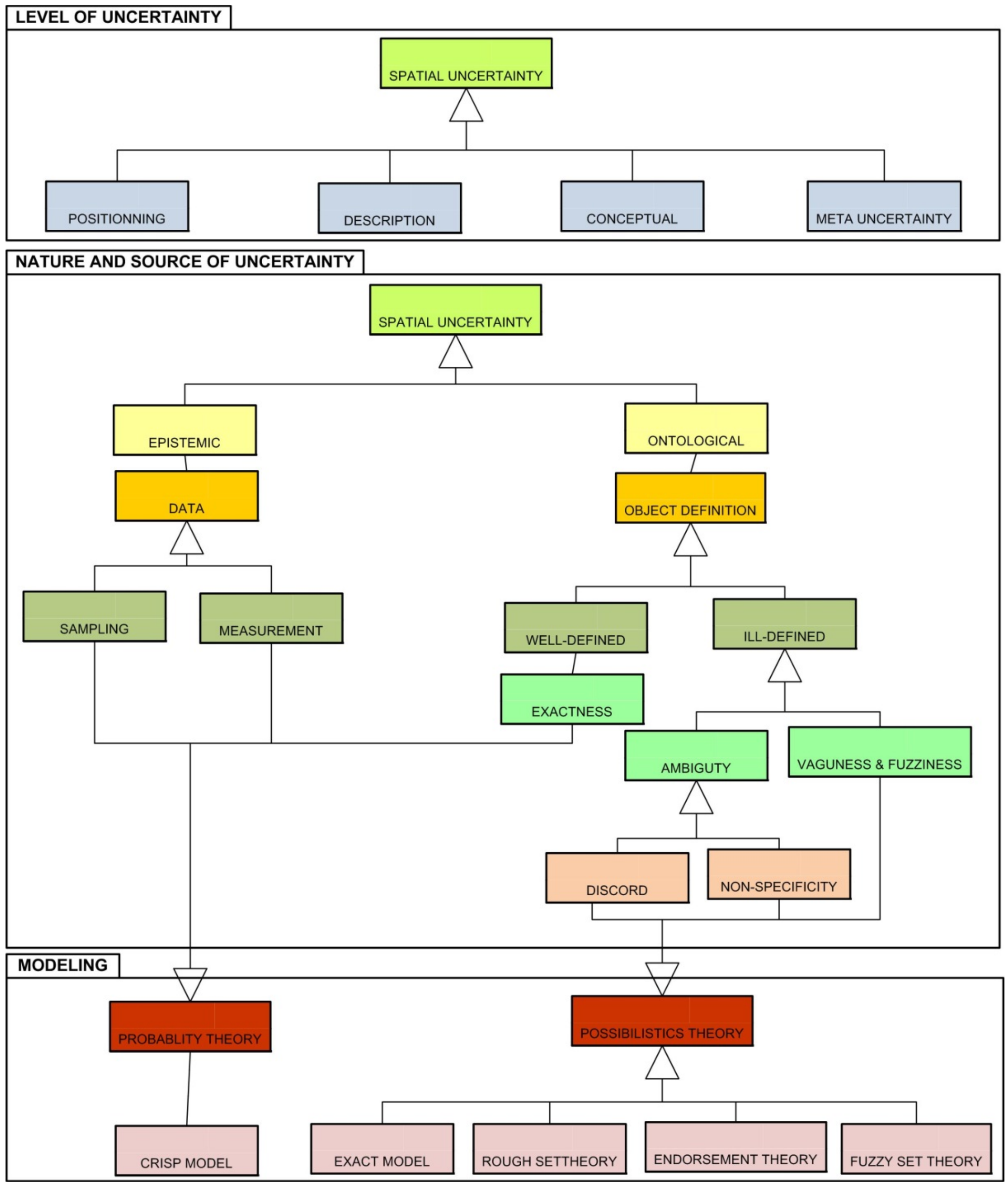


Figure 4.3: A comprehensive UML class diagrams of spatial uncertainty in spatial data modeling and the methods to hand it

Handling spatial uncertainty related to the ontological aspect demands an integrated method which take into account both semantic and geometrical imperfections. Exact, endorsement, rough, and fuzzy models are some examples in this regard (Fisher et al. 2010; Pauly & Schneider 2010; Fisher et al. 2007). Robinson (2003), Burrough (1989), and Usery (1996) worked on handling the descriptive; while Altman (1994) and Brown (1998) dealt with geometrical aspect of uncertainty. Molenaar (2000) integrated both descriptive and geometrical aspects within a formal syntax model for conventional crisp objects based on the fuzzy object model. This was a revolutionary approach through a concept, which takes into account uncertainty inherently. Many papers and research results demonstrate the flexibility of fuzzy set theory to handle the ontological spatial uncertainty (Vassur et al. 2004; Dilo et al. 2007; Molenaar & Cheng 2000; Cheng et al. 2005; Cheng 2002; Fisher et al. 2007; Fisher et al. 2005; Cheng et al. 2009; Chen 2009; Schneider 2003a; Roy & Mandal 2011; Dragicevic & Marceau 2001; Kanjilal et al. 2010; Chowdhury et al. 2009; Morris & Jankowski 2005). In this regard, the following section describes how to use fuzzy set theory to model a spatial object.

4.5 Fuzzy Representation

The concept of fuzzy set theory is originally proposed by Zadeh (Zadeh 1965) to model ill-defined concepts such as the distinction between “tall person” or “short person”. A fuzzy set is a set of objects whose membership to the set takes a value between zero and one. Each fuzzy object can have partial or multiple memberships (Robinson 2003). A fuzzy set A in X is mathematically characterized by a membership function $\mu_A(x)$ which associates with each ***x*** in *X* a real number in the interval [0,1], with the values of $\mu_A(x)$ at ***x*** representing the "degree of membership" of ***x*** in **A** (Zadeh 1965):

$$A = \{(x, \mu_A(x) | x \in X \wedge \mu_A : X \to [0,1])\} \quad \text{Eq. 4.1}$$

Cheng (2002) distinguished four approaches to model fuzzy objects. These include Fuzzy-Fuzzy (FF) objects, objects with α-cut boundaries (αF), Fuzzy-Crisp (FC) objects, and Crisp-Fuzzy (CF) objects. FF-object model or smooth fuzzy object results from fuzzy classification where the spatial extent of an object and its attributes are uncertain (Schneider 2003a). This uncertain part is then described by a membership function. α-cut boundaries is another way to represent a fuzzy object by assuming a threshold value α for each cell of each layer (Zhan & Lin 2003). The main advantage of this method is that it can be applied to known geometric data structures such as Triangular Irregular Network (TIN). FC objects are similar to “Egg-Yolk” model upon conditional boundaries (Cohn & Hazarika 2001). The inner region, the “yolk”, gives the certain part of the object. The outer region, the “white”, is the indeterminate boundary that delineates limits on the range of vagueness. The white and yolk together form the egg that is the full extent of the fuzzy object (Randell et al. 1992). An object with fuzzy spatially extent (transition zones) and a certain core is called a CF-object (Schneider 2003a). The main advantage of this model lies in the implementation phase since efficient representation algorithms from crisp models can be adopted and reused (Kanjilal et al. 2010; Pauly & Schneider 2010). However, the insufficient knowledge to characterize the uncertain part of the object is the main drawback.

All these methods suggest another type of data model that is often called vague or fuzzy spatial data model. The difference between the terms vague and fuzzy refers to the approach used to define the data model as an exact model and approach based on mathematical theories (rough set theory and fuzzy set theory) as illustrated in Figure 3. (Pauly & Schneider 2010; Erwig & Schneider 1997; Schneider 2003b; Kanjilal et al. 2010; Dilo et al. 2007; Bejaoui et al. 2008) used exact and rough approaches to define the vague spatial data model. In contrast, (Schneider 1999; Schneider 2003a; Cheng et al. 2001; Fisher et al. 2005; Molenaar & Cheng 2000) used fuzzy set theory to define the fuzzy spatial data model. Regardless the term vague or fuzzy, this type of spatial data model consists of vague or fuzzy points, vague or fuzzy lines, vague or fuzzy polygons, and vague or fuzzy partitions/grids (Dilo et al. 2007; Schneider 1999; Pauly & Schneider 2010; Schneider 2003a; Cheng et al. 2001; Cheng et al. 2005; Fisher et al. 2005; Molenaar & Cheng 2000; Kanjilal et al. 2010). Our focus in this study is fuzzy set theory due to its flexibility in model the continuous and heterogeneous phenomena such as coastal erosion risk. Hereafter, we will use the term fuzzy spatial data model. Extending the idea of (Dilo 2006; Schneider 2003a), a fuzzy object is defined by its position, geometry, and attributes where each of these components can be uncertain and considered by a degree of membership with respect to their vulnerability index classification.

Fuzzy Membership Function (MF) is a subjective study in nature that is defined by adapting crisp classification of the region by extending these boundaries into a transition zone. However, defining the shape of a MF is a challenge. Two active and passive approaches are used to find the shape of a MF (Tang 2004). In the active approach, MF is defined based on expert knowledge, which is the case of CERA in this paper. Examples are Semantic Import Model (SIM) or Fuzzy IF-THEN rules (Usery 1996; Wang et al. 1996). In the passive approach, numerical taxonomy methods such as Fuzzy C-Mean approach (Bezdek et al. 1984), Self-Organized Map (Chi et al. 1995), fuzzy supervised classification [59], or neural network methods (Nauck & Kruse 1997) are employed. In fact, all input variables (qualitative or quantitative) are initially converted into fuzzy variables using membership functions, a process known as fuzzification. The shape of the membership function is then optimized through successive observations (Robinson 2003). The most difficult step in fuzzy modeling is defining MF. Here, we use the active approach based on expert classification of vulnerability index. The details are provided later in this manuscript.

4.6 Fuzzy Representation of Coastal Erosion Risk: A Conceptual Framework

The proposed conceptual framework for fuzzy representation of risk zones is illustrated in Figure 4.4. This framework consists of two main steps: (1) Tessellation and (2) Fuzzy representation including (a) Fuzzification and (b) Fuzzy aggregation sub-steps. The applied algorithm is presented in Table 4.1.

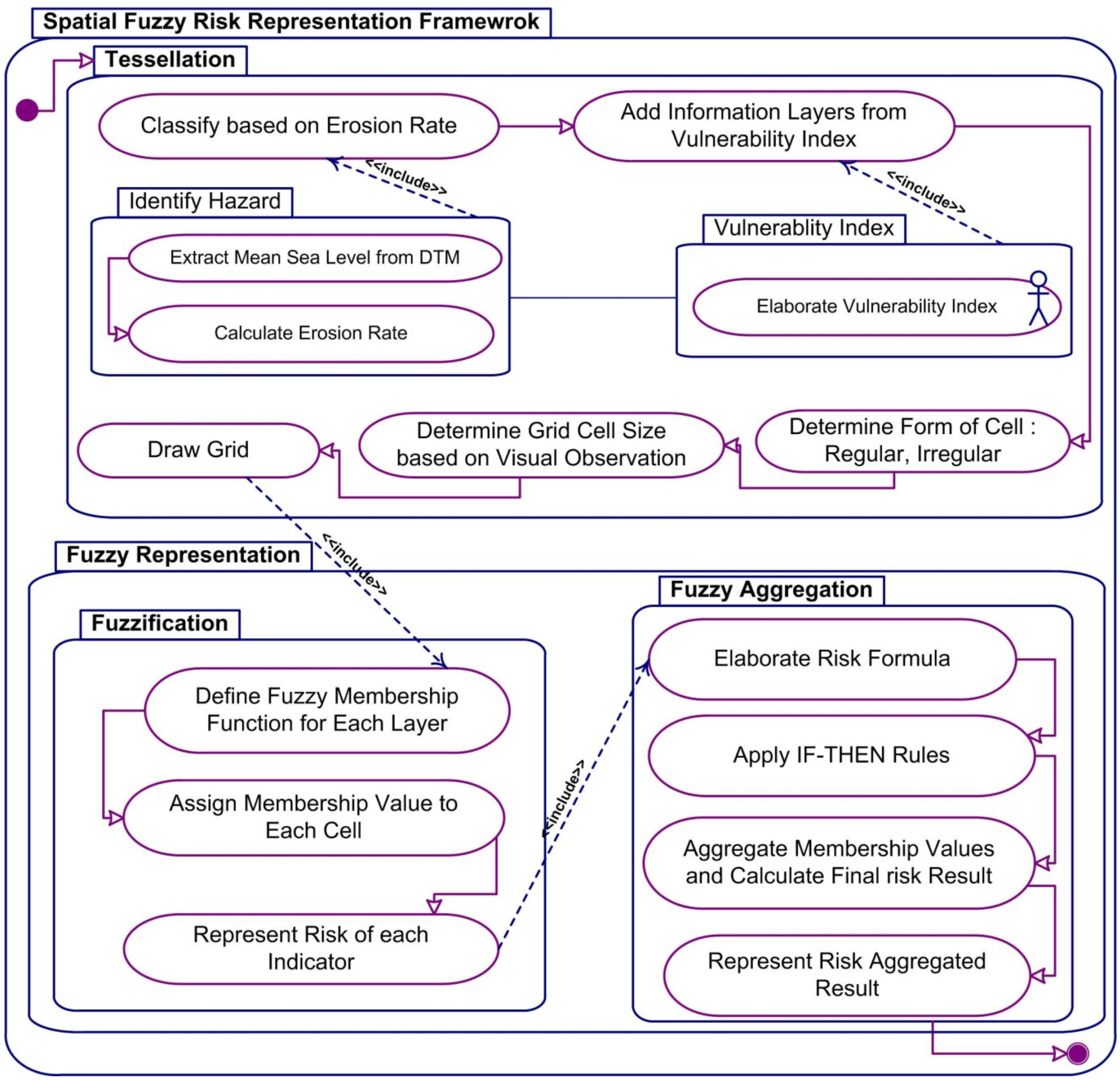


Figure 4.4: UML activity diagram of conceptual framework for spatial fuzzy representation of coastal risk zones

Table 4.1: Applied algorithm for spatial fuzzy representation of coastal risk zones.

1. Draw the grid on the region regarding with the identified hazard and elaborated vulnerability index, *2. For i=1:number of vulnerable indicators,* *a. For j=1:length of the grid,* *i. Determine the fuzzy membership value for each cell of the grid using defined fuzzy membership function of each indicator,* *ii. Assign membership value to center of each cell for each indicator,* *iii. Represent the risk value for each indicator,* *b. End* *3. End* *4. Aggregate the risk value of different indicators based on assign operator,* *i. Elaborate risk formula,* *ii. Apply IF-THEN rules,* *iii. Calculate the risk value,* *5. Represent the aggregated result,* *6. End.*

4.6.1 Tessellation

The tessellation step is performed with respect to identified hazard and elaborated vulnerability index. Hence, identification of hazard and elaboration of vulnerability index are indispensable actions in this step. Indeed, the proposed method is a generic framework to represent risk zones based on fuzzy approach (Figure 4.4.). This framework can be adapted to any other applications and phenomena for risk assessment. Here in this paper coastal erosion risk representation is chosen as a special case study and hence erosion rate and coastal vulnerability index are included in the diagram.

The common indicator for coastal erosion identification is coastline change by calculating erosion rate in different epochs (Genz et al. 2007). Mean Sea Level (MSL) extracted from Digital Terrain Model (DTM) is commonly used to estimate erosion rates. The coastal erosion rates are then calculated by transecting a perpendicular line as profile along these MSL lines with respect to different time periods.

The elaboration of vulnerability index is commonly based on expert-knowledge and the interests of stakeholders and decision-makers (Jadidi et al. 2013; Füssel & Klein 2006). A schema of such vulnerability index is presented in Figure 4.5. Since the scope of this paper is confined to investigating fuzzy approach for spatial representation of risk zones, the vulnerability index is used from Jadidi et al. (Jadidi et al. 2013). In fact, vulnerability indicators are classified based on experts' knowledge to characterize the susceptibility of exposed elements at risk by ranking scores from 1 to 5 and their importance. These classifications are always done based on experimental studies and the specific needs of stakeholders. The reason to define the scores from 1 to 5 is associated with human feeling perception from low sensitive to high sensitive situations, i.e. very low, low, average, high, and very high. This standard is also employed by researchers to consider the risk degree (Boruff et al. 2005).

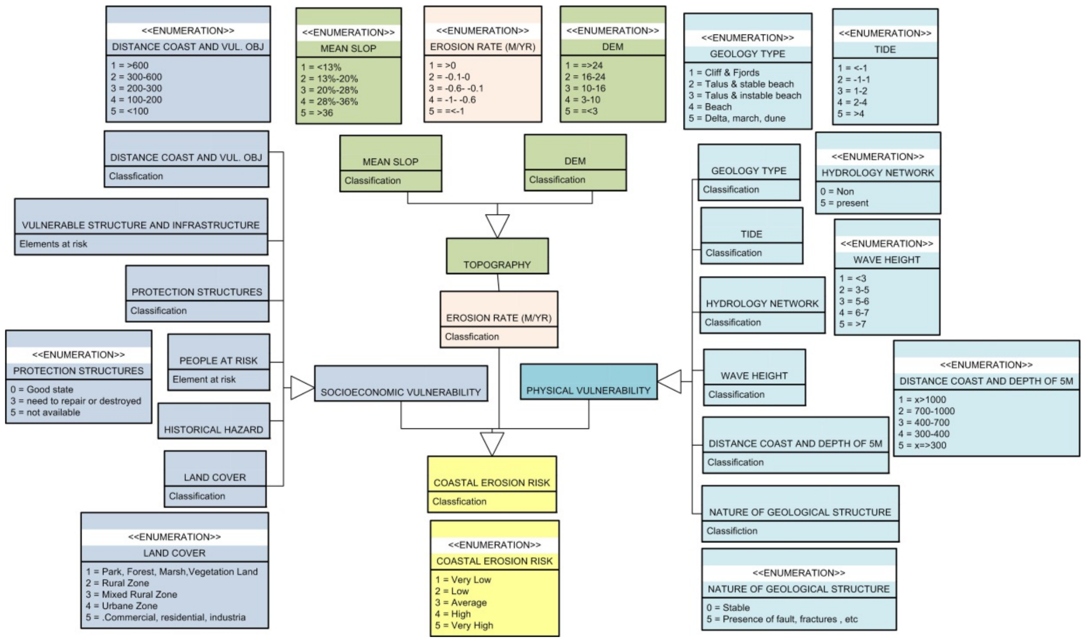


Figure 4.5: UML class diagram of vulnerability index for coastal erosion risk assessment adapted from (Jadidi et al. 2013)

Henceforth, with respect to the estimated erosion rate, the region is classified into 5 categories. A risk level is assigned to each class of erosion rate from very low, low, average, high and very high as presented in Figure 4.5, "Erosion Rate" class. The information about vulnerable indicators, e.g. road network, houses, people density, etc. are then integrated to provide a clear perception of the distribution of elements at risk and the variation of erosion. This step can be done using any available GIS tools such as ArcGIS. However, choosing the size and shape of the grid cells is always a challenging issue in this step and it is totally experimental. A regular tessellation is chosen in this study because of its simplicity to perform in many GIS tools and in the fuzzification step. The size of the cell depends on required scales and available information. Indeed, if a fine-grain risk representation is needed and high dense data (e.g. LiDAR) is available, the cell size can vary from the same resolution of derived DTM from LiDAR data to census units (hundreds square meters).

4.6.2 Fuzzy Representation

Fuzzy representation step consists of two sub-sections (a) Fuzzification and (b) Fuzzy Representation explained hereafter.

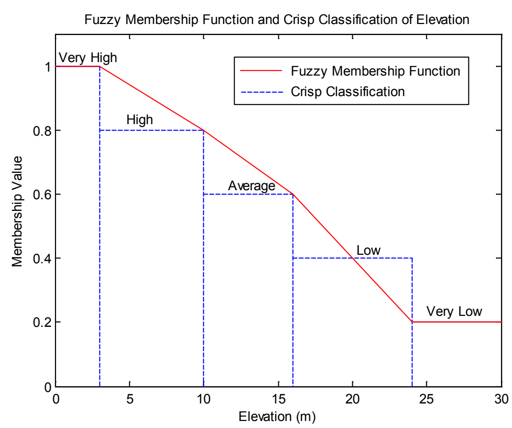
4.6.2.1 Fuzzification

The fuzzification consists of determining the membership functions and the respective membership value for each cell of a grid of each vulnerable feature. As stated previously, the fuzzy set theory is employed in this study to deal with uncertainty related to spatial fiat objects. Regarding the intrinsic nature of coastal erosion risk, the FF-objects model is used in this paper to represent risk zones. This kind of representation comes intuitively from fuzzy classification results (here from vulnerability index classification). Risk zones are extracted from these classifications consisting of continuous sets of grid cells belonging to one class. The objects of one class are then represented as a layer of the grid, so that N layers of objects will be formed, each consisting of fuzzy regions. A membership value is assigned to each element of the grid called a cell. Worth stating that based upon Schneider (Schneider 2003a)'s definition, fuzzy objects are finite collection of elements from a regular tessellation, forming a partition of bounded subspace of $\Re^2$.

In the present paper, the membership functions are derived from vulnerability index classification, where a degree of risk is attributed to each indicator based on experimental studies (Jadidi et al. 2013). The membership values are determined by associating the vulnerability index with the independent variable of each indicator (horizontal axis, Figure 6) and dependent variable of the membership value (vertical axis, Figure 4.6). In Figure 6(a) and 6(b), the membership functions of elevation and erosion rate are presented as graphical examples to compare with their respective crisp classifications.

(a) Elevation

$$\mu = \begin{cases} 1 & x \le 3 \\ \dfrac{0.8-1}{10-3}x + 1 & 3 < x \le 10 \\ \dfrac{0.6-0.8}{16-10}x + 0.8 & 10 < x \le 16 \\ \dfrac{0.4-0.6}{24-16}x + 0.6 & 16 < x \le 24 \\ 0.2 & x > 24 \end{cases}$$



(b) Erosion Rate

$$\mu = \begin{cases} 1 & x \le -1 \\ \dfrac{0.8 - 1}{-0.6 - (-1)} x + 1 & -1 < x \le -0.6 \\ \dfrac{0.6 - 0.8}{-0.1 - (-0.6)} x + 0.8 & -0.6 < x \le -0.1 \\ \dfrac{0.4 - 0.6}{0 - (-0.1)} x + 0.6 & -0.1 < x \le 0 \\ 0.2 & x > 0 \end{cases}$$

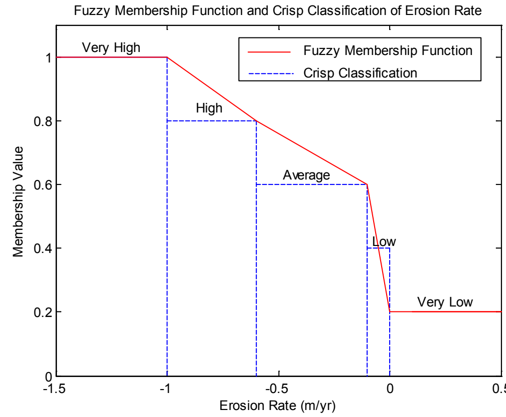


Figure 4.6: A graphical Example of membership functions of some indicators and their crisp classifications: (a) "Elevation" and (b) "Erosion Rate".

The risk zones are represented by a grid based data structure. Each grid cell is identified by its center, vertices, and edges (see Figure 4.7.(a)). The membership value is assigned to the center of each cell and scattered as Gaussian function toward the outside (Kentel & Aral 2007). For any other point (X,Y), the membership value is calculated from Eq.4.2:

$$F(X,Y) = mv \times e^{(-\sqrt{(X-X_c)^2+(Y-Y_c)^2})} \quad \text{Eq. 4.2}$$

where *mv* is the amplitude at the cell's center (X_c,Y_c). Indeed, the Gaussian function is used to feed the cell neighbors with respect to the inverse of weighted distance from the cell's center (see Figure 4.7.(a)). A risk zone in this case is generated by aggregating a set of cells with the same values (see Figure 4.7(b)). In Figure 4.7(b), the color hue represents the risk value. Dark red represents higher risk with membership value close to 1 and light blue represents lower risk with membership values close to 0. It is worth to mention that the hypothesis is that the data are already cleaned regarding positional and measurement uncertainty using probability theory by an accepted confidence level.

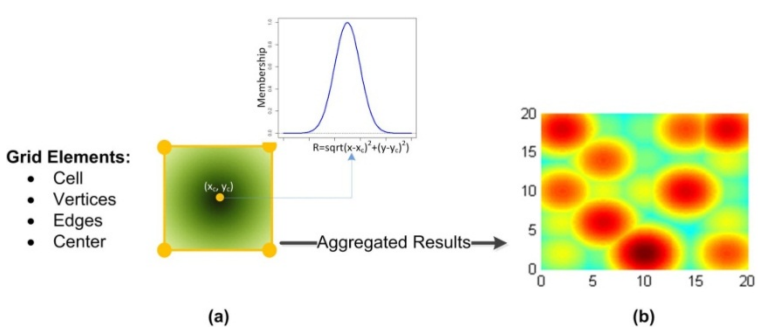


Figure 4.7: (a) Proposed approach based on fuzzy model. (b) Fuzzy representation.

4.6.2.2 Fuzzy Aggregation

To calculate the overall risk value for a given region, the aggregation of multiple layers is required. A risk formula is thus elaborated in this step consisting of hazard, element at risk, and vulnerability index's parameters (Jadidi et al. 2013) to perform a vertical integration of information (Bejaoui 2009). IF-THEN rules are defined based on the risk formula and the priority of stakeholders and authorities of the region under study. An example of a fuzzy rule is shown in Table 4.2. Fuzzy operators translate IF-THEN rules and combine the individual output of membership values. In fact, IF-THEN rules link computationally multiple outputs with multiple inputs. To aggregate multiple layers, The "Overlay" operation is performed using "Union", "Intersection", "Mean", or "Mean Weighted" (see Figure 4.8). The "Union" operator returns the maximum membership value of compared layers. That is also called the "MAX" or "OR" operator. The "Intersection" operator returns the minimum membership value of compared layers that is also called "MIN" or "AND" operator. "Mean" and "Mean Weighted" operators calculate the average and weighted average membership value of compared layers respectively.

Table 4.2: An example of fuzzy IF-THEN rules

IF *(HydroNetwork is VH) and (ProtectStructure is VH) and (DistObjVul is VH) and (ErosionRate is H)* ***THEN*** *(Use "MAX" operation for "Erosion Risk" calculation)*
VH: very high, H: high, A: Average, L: low, VL: very Low

The choice of fuzzy operators reflects the specific needs laid out by CERA that requires the knowledge of the maximum, minimum, average, or weighted average of the erosion risk associated in the given region. The results from the risk assessment can then be represented either using risk map (see Figure 4.9.) or using tables and charts after defuzzification of the fuzzy risk values. Figure 4.9 illustrates the fuzzy representation of a region with respect to five arbitrary varied indicators (Figure 4.9.(a)) and their overlaid results using fuzzy operators (Figure 4.9.(b)). If there are some cells with no data, the membership values for these cells are calculated using their neighbours' membership values. This operation is likewise based on one of fuzzy operators that are "Union", "Intersection", "Mean", or "Mean Weighted".

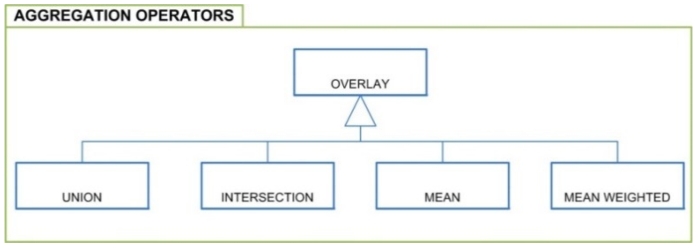


Figure 4.8: UML class diagram of Fuzzy Aggregation Operators

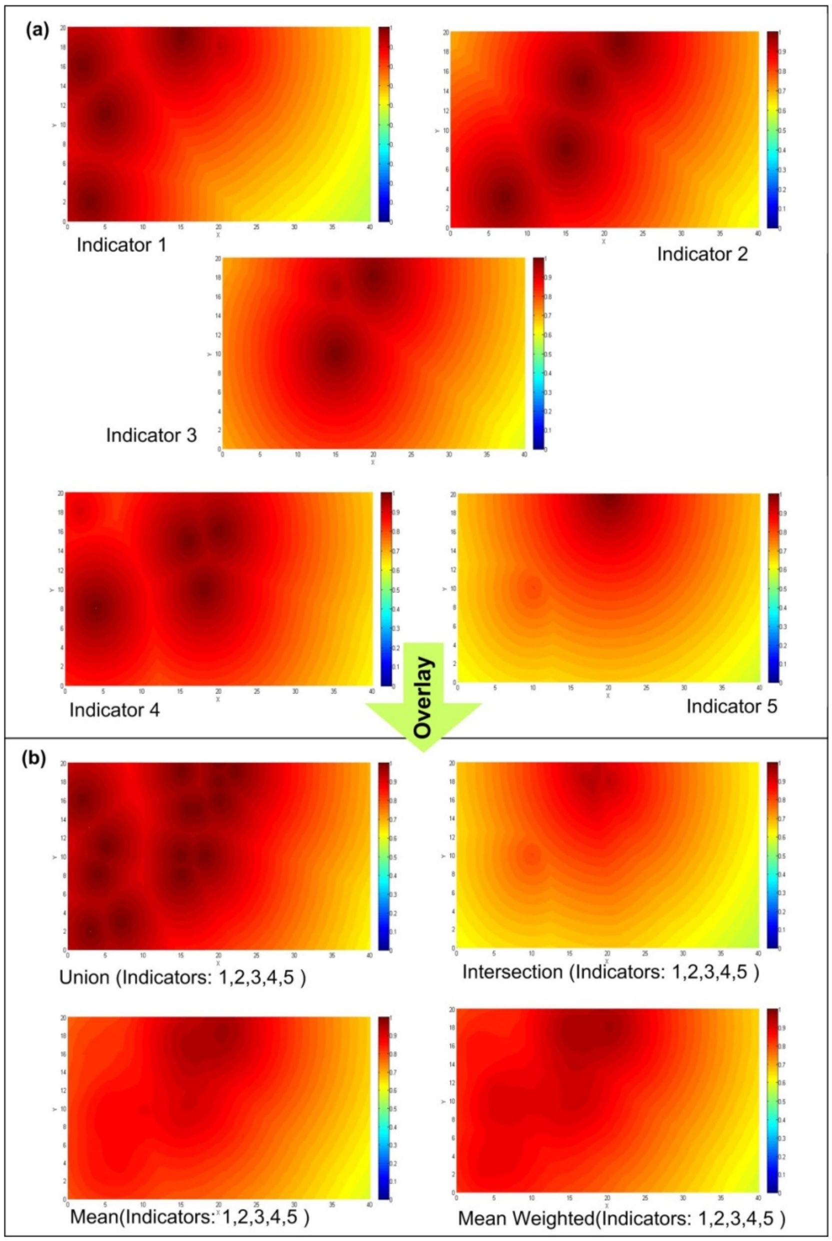


Figure 4.9: The representation of five different indicators in each layer. (b) Fuzzy aggregation of these indicators: an overlay operation (union, intersection, mean, and mean weighted).

Defuzzification is the process of translating fuzzy result values into crisp values or linguistic expressions (Kentel & Aral 2007). In fact, defuzzification is the reverse process of fuzzification. This step is essential in CERA because some decision-makers are more comfortable and prefer to work with the traditional method with crisp values of linguistic expressions. The common defuzzification methods are centroid, maximum, and mean methods (Kentel & Aral 2007). The centroid method is employed in this study for its reputed performance on extremely large amounts of uncertain data. The centroid method determines the centre of the area of the combined membership values. Accordingly, the relation between risk degree for each fuzzy value, crisp value and linguistic expression is established. The defuzzification relationships used and proposed in this paper are provided in Table 4.3. This classification matches also with linguistic variables in the fuzzy membership values of vulnerability index. It should be noted that defuzzification generally causes information loss embedded in fuzzy values.

Table 4.3: Defuzzification result for final risk classification

Linguistic Expression	Crisp Value	Fuzzy Risk Value
Very Low Risk	Risk(Erosion)=1	0 ≤ Risk (Erosion) ≤0.175
Low Risk	Risk(Erosion)=2	0.175< Risk (Erosion) ≤0.375
Moderate Risk	Risk(Erosion)=3	0.375< Risk (Erosion) ≤0.575
High Risk	Risk(Erosion)=4	0.575< Risk (Erosion) ≤0.775
Very High Risk	Risk(Erosion)=5	0.775< Risk(erosion) ≤1

4.7 Results: A Case Study

To validate the proposed framework, a case study was carried out by executing multiple steps mentioned in section 4.5.

4.7.1 Study Site

The region along the coast of the St-Laurence River in Perce, near the tip of the Gaspe Peninsula in Eastern Quebec, Canada (see Figure 4.10) is identified as a potential study site to implement and validate the proposed framework. This region is one of the most attractive touristic locations in Quebec because of the well-known Perce Rock and Bonaventure Island. It has 52km of coastline, 432.39km^2 surface, a population density equals to 7.7 people per km^2 , and a total population of 3312 (StatisticCanada 2012).

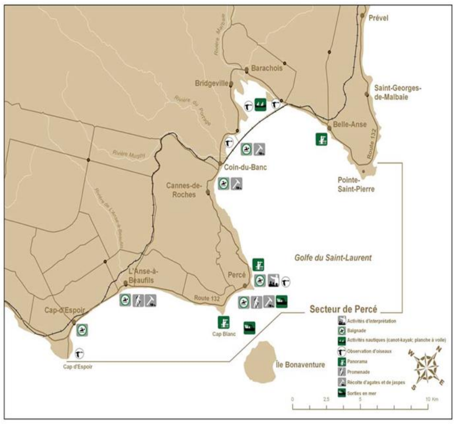


Figure 4.10: Geographical view of Perce, Eastern Quebec, Canada

Perce is principally characterized by rocky cliffs consisting of sedimentary rock, sandstone, and limestone of Gaspe covering 76% of surface area of the region. These are then divided into 17 smaller segments, rocky inlets, and beach terraces for 12% and a spit of beach and tidal marsh for 11%. 80% of Perce coasts are under severe coastal erosion with an average rate of -0.20 m/year (1994-2001) (Bernatchez, Fraser, Friesnger, et al. 2008). Sea Level Rise (SLR) also has a particular impact on coastal erosion by reducing the width of beaches according to study results from 1934 to 2001 (Bernatchez, Fraser & Lefaivre 2008). Heavy rain is another factor that accelerates the erosion process at cliff bottoms which may even provoke sudden and discrete damaging events such as rock falls and landslides along or far from the coasts. An average erosion rate of -0.49m/yr is predicted until 2050 in the region (Bernatchez, Fraser & Lefaivre 2008).

Infrastructures such as the road network (132 national road) and the railway along the spit of Barachois include 34.6% of the socio-economic features at risk in this region; 79% of these features have already been affected by erosion with a total cost of $15.5 million (Bernatchez, Fraser, Friesnger, et al. 2008). Businesses related to the tourist industry, together with the majority of jobs associated with this industry, are at a high risk. In addition, the residential areas (23.4%) and villages (5.8%) located around the ports and along the coast are also affected extensively by erosion. It is reported that a total of 13 residential houses, three rural houses, seven commercial buildings, and two industry buildings are in danger with regard to coastal erosion (Bernatchez, Fraser & Lefaivre 2008). 30.7% of the coastlines are natural and wild areas. Any change along the beach caused by erosion may significantly reduce the tourist attraction of the region. The landscapes and natural areas such as lagoons and beaches can then be considered as vulnerable elements that should be taken into account for any sustainable planning for the region. Fishing ports, harbors, and docks are also vulnerable features in this regard.

4.7.2 Implementing Proposed Framework on Study Site

Multiple sources of data are used to accomplish coastal risk assessment of the study site. Table 4.4 presents a list of datasets and the related parameters that are used for CERA in this case study. Most of the information is extracted from technical reports and is then projected to their respective geographical position. ArcGIS 10 is used to produce DTM from LiDAR data and Digital Shoreline Analysis System (Thieler et al. 2009) is employed to obtain the erosion rate in this study. The information about vulnerability index is extracted from technical and research reports and formally investigated in ArcGIS 10 yielding a regular tessellation. A grid of 40mx40m is produced along the coast with a width of 1.4 km to 2.760 km (dependent upon data availability). The 40mx40m grid was selected based on the resolution of produced DTM from available LiDAR data on the region under study. This size can change from one to another depending upon needed specifications and the objective of performing CERA.

Table 4.4: The list of data sets used for erosion risk assessment

Source	Extracted Information
LiDAR Data	Slop , DEM, erosion rate
Technical and Research Reports	Protection structure, Infrastructure situation, type of coastline, state of coastline, land use information, distance coast and 5m depth, distance coast element at risk
Geobase	Hydrology network and drainage, land use
Quebec Prov. Transport Dept.	Road network

Matlab code is developed to perform fuzzy representation of risk zones. A fuzzy membership value is assigned to each cell based on the defined membership function of each indicator. The risk with respect to a specific priority i.e. element at risk is then calculated using the following formula (Jadidi et al. 2013):

$$CoastalErosionRisk(ElementAtRisk,time) = ErosionRate(ElementAtRisk,time) \times \sum_{i=1}^{n} v_i \times w_i\,(ElementAtRisk,time)$$

Eq.4.3

The element at risk in this case study is the road network. v_i is the vulnerability indicator and w_i is the weight value. Table 4.5 presents the list of eight indicators and associated weights derived from technical reports from the Quebec Provincial Transport Department that are used in our case study (Xhardé R. 2007). The calculated fuzzy risk values are then aggregated using the “Mean Weighted” operator that is best fitted to risk formula (Eq.4.3). Fuzzy representation of risk zones is shown in Figure 4.11. It is worth stating that the list of vulnerability indicators provided in Figure 4.5 is a complete list of coastal vulnerability indices. Nevertheless, all of them are not available in all case studies.

Table 4.5: Associated weights of vulnerability indicators used in case study (adapted from (Xhardé R. 2007))

Vulnerability Indicators (v_i)	Weight (w_i)
Protection Structure	34%
Distance shore and element at risk	17%
Mean Slop	13%
DEM	10%
Geology Type	8%
Land Cover	7%
Hydrology Network	6%
Distance shore to 5m depth	5%

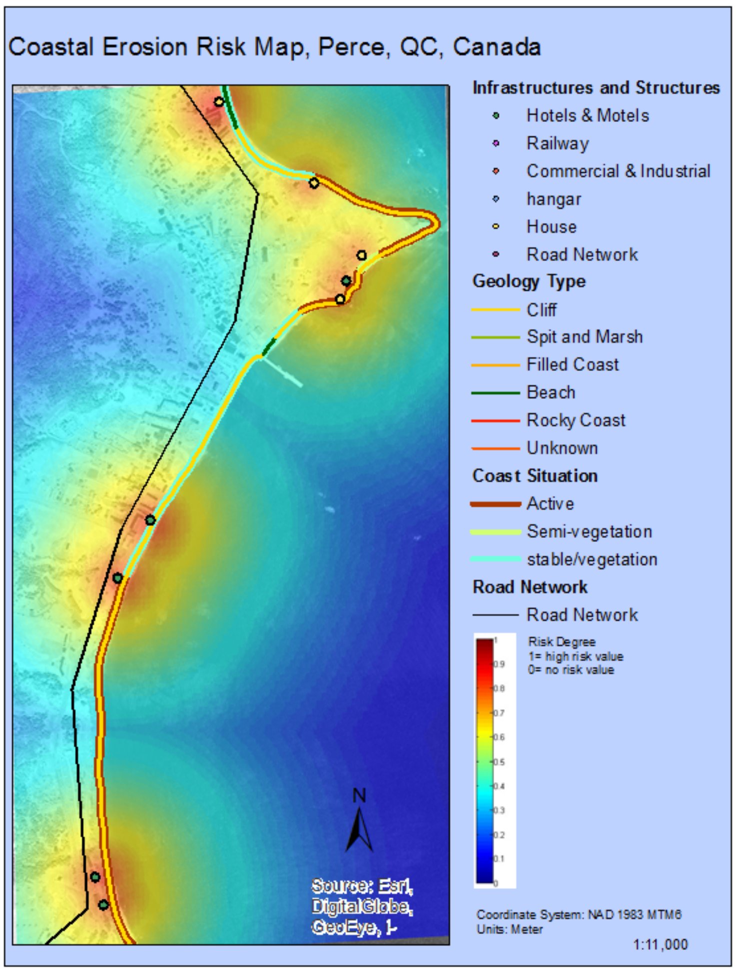


Figure 4.11: Fuzzy representation of coastal erosion risk zones on the study site.

4.7.3 Results Interpretation

As illustrated in Figure 4.11, regions at high risk are along the coast. About 800m of 132 national road is severely at risk (north, center and southwest). The residential area (total of 4 houses in yellow circles) and two motels (green circles) on the nose of Perce rocks are also at very high risk. From the nose of Perce rocks toward southwest, 4 motels are also at high risk. Additionally, having the high risk region (red zone) along the road 132 confirms the results of previous studies in the region (McHugh et al. 2006; Xhardé R. 2007). This region is reported as an active cliff coastline with 0 m/year of erosion rate. However due to its geology type, the erosion may happen all of sudden as a land-slide. The main difference between our method and other studies is how spatial uncertainty related to risk modeling is handled through fuzzy approach.

The high risk areas are well recognized with respect to existing infrastructures, buildings, people, and their properties. Indeed, the obtained results in this study are more consistent with human reasoning and perception, conveying the level of risk in a continuous and smooth manner. The continuity is not only handled by raster format but is also conducted by the fuzzy representation.

4.8 Discussion and Remarks

In addition to spatial uncertainty, defined as the lack of knowledge (Walker et al. 2003), the imperfection of very large amounts of data and information should also be considered in CERA. CERA is traditionally an expert-based process. Giving the flexibility and capability to integrate expert knowledge (structured as linguistic expression) and characterizing and handling data uncertainty have always been stated as important issues to improve the quality of results in CERA. The proposed approach in this study was to employ knowledge-based solutions such as fuzzy set theory with the main following advantages:

- Spatial uncertainty associated with object definition is explicitly dealt through fuzzy approach. It is also possible to attach a probability density to the values of position and measurement uncertainty. If this is the case, it is recommended to clean the data using probability approaches with an accepted confidence level before using this data in CERA.

- Membership function definition issues are resolved by converting the crisp classification of vulnerability index to a fuzzy classification. Accordingly, the integration of multiple criteria is performed by aggregating their respective membership values using fuzzy aggregation operators. If the vulnerability index classification is not available, the methods such as Fuzzy C-Mean and fuzzy K-Mean are recommended to define the required membership functions based on available data.

- The grid-based structure of the proposed approach avoids the difficulties of combining different membership values to compute the fuzziness inside of objects where various criteria lead to fuzziness. Besides, several studies confirm that the fuzzy approach works well with grid-based format (Pauly & Schneider 2010; Bejaoui 2009).

- Elaborating risk formula and then constructing IF-THEN rules of associated indicators allow a direct control on the entire CERA process. In addition, this provides more flexibility if one or some indicators or their classifications are changed. In this case, updating the desired information by re-running fuzzification step or modifying the IF-THEN rules by re-executing fuzzy aggregation step will be sufficient.

- The proposed approach allows performing multiple fuzzy aggregation operators (union, intersection, mean, mean weighted) that is required in any CERA process. The result in Figure 4.9.(b) shows how significantly, the choice of fuzzy operators can affect the end result. Therefore, regarding the needs of decision-makers and the emergence of protection actions, the choice of these operators is also varied.

- The flexibility of fuzzy set theory to characterize and handle inherent spatial uncertainty through the entire assessment process widely increases the confidence levels of adapted strategies for protection regions under study. It also accelerates the implementation of response plans in case of a disaster. This happens through the interpretation of the result and prioritization of the planning actions based on expert perception in an accurate manner. From another point of view, traditional risk assessment methods lead to crisp decisions i.e. “Yes” or ”No” while the fuzzy approach leads to smooth transitions between these two extremes.

- Fuzzy risk representation is a relatively new concept for decision-makers. In this new context, decision-making processes need to be adapted and meaningful criteria need to be established to accept and manipulate fuzzy risk values. Changing the decision-making culture to use fuzzy results requires finding evidences to convince the decision-makers of the benefits of this new approach. The defuzzification step explained briefly in this paper is an alternative in this regard to translate fuzzy values to measurable values, making them understandable for decision-makers. Kentel and Aral (Kentel & Aral 2007) propose a risk tolerance measure method based on a crisp compliance guideline which is already available in some domains such as in the health system.

- On the other hand, the main limitations of the proposed approach in this paper are as follows:

- The proposed fuzzy representation is tested only on regular tessellation. The neighborhood relation is implicit, based on the ID of a cell. If an irregular tessellation is needed, more effort in neighborhood concepts and topological predicates are required.

- The temporal aspect of the fuzzy object is not taken into account in this approach. This paper only discusses the spatial extent of fuzzy objects and the situations that the fuzzy classification is due to multi-criteria nature of CERA and spatial uncertainty associated with object definition. This means that the risk zones are represented spatially as a snapshot of a given time period. How to handle fuzzy objects that change in different time periods needs more investigation.

- The proposed approach is employed only on a small region with a given level of details (scale). When analyzing extremely large amounts of data within a hierarchical system is required, the proposed approach needs to be adjusted with respect to selected technology. In this regard, efforts are mainly needed on fuzzy aggregation operators such as "Fusion" where the multi-scale representation is required.

4.9 Conclusions

Characterizing spatial uncertainty is important in coastal risk assessment for effective decision-making. Also, accurate spatial representation of coastal risk is essential to provide the necessary knowledge of the potential impact of the risk and to help decision-makers to take the necessary actions to better protect people, infrastructures, and other installations along the coast. This paper has focused on the improvement of spatial representation of coastal erosion risk by taking into account the inherent uncertainty related to spatial objects and risk zone representation. The associated uncertainties were characterized as vagueness and fuzziness of object and then handled by fuzzy set theory. A conceptual framework was proposed to represent risk zones based on fuzzy model. Vulnerability index classifications were used to determine membership functions for each indicator as a separate layer. A regular tessellation of the region was generated for each indicator by assigning an appropriate membership value to each cell indicating the degree of risk. IF-THEN rules were defined to aggregate multiple layers of indicators by using aggregation operators such as Union, Intersection, Mean, and Mean Weighted. Finally spatial fuzzy representation of CERA was presented. The proposed approach was applied to a study site in Perce, Quebec, Canada to demonstrate the validity and advantages of the proposed method. This method provides a better tool for decision-making as the risk values were better adapted to the reality of the region and better help the decision-makers in their work. Since the ultimate objective in any risk assessment is assisting efficient decision-making, so confronting very large amounts of data is unavoidable. Nowadays, geospatial intelligence systems are extensively used in this regard.

Implementing the proposed method can open new insights on how to deal with spatial uncertainty in such databases. Manipulating fuzzy models in a hierarchy system requires fuzzy aggregation operators. The expansion and redefinition of these concepts in a spatial multidimensional systems context is one of the concerns for future work.

"If the facts don't fit the theory, change the facts."
Albert Einstein

Chapter 5 Fuzzy Spatial Datacube for Multi-Scale Coastal Risk Assessment: Towards Fuzzy Spatial Aggregation to Support Geospatial Decision Models

Submitted on International Journal of Geographical Information Science, 2014
Jadidi A., Bedard Y., and Mostafavi M.A.,

Keywords: *Fuzzy Set Theory, Spatial Datacube, Data Aggregation, Information Vagueness.*

5.1 Preface of Chapter

This chapter is composed of the third paper within the framework of the present thesis which is submitted to the International Journal of Geographical Information Science. The paper is focused mainly on the third objective of the thesis which proposes to formally define the components of a fuzzy spatial datacube and its fuzzy aggregation operators based on Fuzzy Set Theory to improve the representation of risk zones. Hence, the fundamental concepts to design a fuzzy spatial datacube are discussed with scrutinizing on potential aggregation procedure. The results obtained from Chapters 3 and 4; risk components (i.e. hazard, elements at risk and vulnerability index), elaborated risk formula, fuzzy representation approach to deal with spatial uncertainty due to object and criteria definition are used as example in this chapter to describe the main elements of fuzzy spatial datacubes and to present some aggregation operations results.

5.2 Abstract

Dealing with spatial information vagueness from multiple sources is a major challenge in any decision-making process. It is particularly the case in assessing the risks caused by natural phenomena when the decision models rely on a number of natural, economic and social indicators. In this context, many of these indicators are semantically vague, others are spatially or temporally uncertain. Most of the indicators can be assessed at different scales depending on the needs of different organizations. Current decision support systems typically do not take into account information vagueness, they assume that the indicators have well-defined and exact semantics, geometry, and temporality. Hence, this paper presents a fuzzy-based approach to deal with information vagueness in the decision-making process. Fuzzy Set Theory is employed to characterize semantic, spatial or temporal vagueness. The proposed use of fuzzy dimensions leads to build a fuzzy spatial datacube within Spatial OLAP (SOLAP) systems. In particular, we propose a fuzzy spatial aggregation method dealing with multiple hierarchical dimensions to produce and represent fuzzy datacube measures, either spatially or numerically. To demonstrate the potentials of a fuzzy spatial datacube, the proposed approach is applied for the assessment of the coastal erosion risks. Results are presented, learned lessons are discussed, and new research directives are established for future works.

5.3 Introduction

Information vagueness is an inherent characteristic of data that should be considered in any decision-making process (Kentel & Aral 2007; Bejaoui 2009). Information vagueness can be propagated during integration and aggregation processes and hence in the decision-making process. Ignoring information vagueness may result in unrealistic or misleading conclusions and decisions that yield into undesired or even catastrophic consequences. For instance, decision-making based on vague environmental, economic, and social indicators may result in incomplete spatial coverage for flood insurances, badly positioned erosion protection infrastructures, changes in emergency plans, more casualties, etc.

The information vagueness in assessing the risks caused by natural phenomena can be characterized as semantic, spatial, and temporal. Most of these phenomena have also multi-scale characteristics (Cheng et al. 2009). Several works have already been initiated to deal with information vagueness (Pauly & Schneider 2010; Bejaoui 2009; Edoh-alove et al. 2013; Schneider et al. 2011; Schneider 2010; Schneider 2003a; Fisher 2008). From the technological point of view, existing tools do not provide built-in capabilities to deal with information vagueness either. For instance, in the case of risk assessment, multiple decision indicators are required to define and then translate into datacube dimensions to calculate potential risk. The choice of indicators is typically based on participants' interests while the semantics of these indicators and their classification and their levels of details depend on stakeholders 'requirements. The information about these indicators is often assumed to be exact or good enough in such spatial data cube for decision-making process. As a result, these datacubes may depict major flaws for the accurate analysis of phenomenon, predictive modeling and comparison of phenomena from different regions, periods and scales. There is hence a strong willingness to propose a comprehensive solution for the issues from information vagueness. A promising approach that is being explored for almost two decades is fuzzy logic (Schneider 2010). This approach is generic and compatible with human reasoning in expert-based processes and procedures; largely employed in current practices in risk assessment (Cheng et al. 2009; Kentel & Aral 2007; Darbra et al. 2008; Skanata & Byrd 2007).

The main objective of this paper is to present how to design a fuzzy spatial datacube based on fuzzy logic approach to deal with information vagueness. To do so, a background of the existing solutions to the problem of information vagueness in coastal erosion risk assessment (CERA) is provided. The proposed fuzzy approach is then explained to handle the spatial vagueness. It is then applied to a spatial datacube with a reference to a CERA multidimensional model. In this regard, fuzzy spatial datacube elements like dimension, hierarchy, measure and fact are redefined. Appropriate methods are then proposed to aggregate and represent risk measures at different hierarchical levels, either spatially or non-spatially. Finally, a discussion on the proposed approach, obtained results, and some guidelines for future works are provided at the end of this paper.

5.4 Background

CERA requires a large amounts of data from multiple sources (Jadidi et al. 2013). These data are often very heterogeneous in their semantics, i.e. definitions of spatial entities and their attribute, domain values, temporalities and geometries (Sboui 2010; Bakillah 2012). This introduces spatial uncertainty as information vagueness in the risk assessment process. In addition, the assessment of erosion risk should be done different scales based on the needs and interests of users and decision-makers (Jadidi et al. 2012). This latter emphasizes the need for hierarchical risk data aggregation method dealing with vagueness in any CERA process (Cheng et al. 2009; Jadidi et al. 2013). Furthermore, definitions of the coastal risk zones are often vague and their spatial and temporal delineation are uncertain. During recent years, spatiotemporal representation of these zones has been under intense investigations (Cheng et al. 2009; Dilo 2006). Although some preliminary solutions has been proposed for the stated problems in the current literature. However, a comprehensive solution for spatiotemporal, multidimensional representation and assessment of coastal erosion risk in the presence information vagueness is necessary for efficient CERA.

Spatial information vagueness is a kind of imperfection arising from the presence of gradual transition between objects or classes (Dilo 2006; Fisher 2008; Smith & Varzi 2000). The vagueness is sometime divided to ontological or linguistic components (Dilo 2006; Jeansoulin et al. 2010; Smith & Varzi 2000). Indeed, an explicit distinction on whether information vagueness is predominantly ontological or linguistic, is difficult and often impossible (Jeansoulin et al. 2010). Therefore, information vagueness refers to the gradual transition boundaries of an object or phenomena in space, for example the boundary between grassland and forest (Dilo 2006; Jeansoulin et al. 2010). Moreover, the impossibility of determining where sharp boundaries of an object are situated is also related to information vagueness (Dilo 2006; Jeansoulin et al. 2010). For instance, on one hand, one defines shoreline as the border of water and land with respect to high tide range while another one prefers low tide range. In all situations, the ontological vagueness is present where the boundary between water and land is a transition zone. On other hand, when we talk about the shoreline without mentioning its type and origin, linguistic vagueness is present. In addition, the definition of shoreline can be even varied during the day and seasons with respect to the tide changes. As a result, the confusion in knowing the sharp boundaries of land and water in a specific time period is perceived as the vagueness in geometry of the shoreline object and time.

Different approaches exist in the literature to handle with information vagueness such as exact, rough, probabilistic, and fuzzy models (Pfoser & Tryfona 2001; Cohn & Hazarika 2001; Bejaoui 2009; Pauly & Schneider 2010; Kanjilal et al. 2010; Haunert et al. 2009; Dilo et al. 2007). Probabilistic models are out of the scope of this paper because they deal only with quantitative positional and measurement uncertainty predominantly and are well documented in the literature. Exact and rough models have commonly been used by scientists and practitioners due to their simplicity and straightforwardness in implementation (Pauly & Schneider 2010; Kanjilal et al. 2010). However, these models have their inherent limitations to represent of an object when it is impossible to determine its sharp boundaries (Molenaar & Cheng 2000). Moreover, if a fine-grained modeling of object is demanded the exact and rough models are not supported well (Pauly & Schneider 2010). In addition, evolutional objects such as a pollution zone, flood zone, or an erosion zone cannot be precisely modeled by three-value logic-based of exact and rough methods., Because, these objects have an extent but cannot be bounded by a precise border (Pauly & Schneider 2010; Fisher et al. 2007; Fisher et al. 2005) . Fuzzy models allow a continuous change of the degree of spatial vagueness in the interior of such object based on multi-value logic of Fuzzy Set Theory. In this regard, the degree of vagueness of an object is expressed by an assigned membership function in fuzzy approach. Therefore, fuzzy models based on *Fuzzy Set Theory* are used in this paper to overcome the stated issues.

Ever-increasing demand to handle and process spatial data in multidimensional databases plays an important role in modern decision-making (Salehi et al. 2010; Laurent 2010; Santini et al. 2010; Bédard et al. 2007). Nevertheless, the idea of fuzzy spatial data model is still a new topic in the geospatial business intelligence communities. Few efforts have been made to consider the vagueness of spatiotemporal information in SOLAP (Bejaoui 2009; Siqueira & Ciferri 2012). Most of the proposed solutions are mainly based on the extended spatial crisp models (i.e. exact and rough models). In a recent work, Edoh-alove *et al.* (2013) proposes a risk-aware approach to face spatial fuzziness in which the user is informed about the existence of low-quality data. However, none of the cited works use Fuzzy Set Theory explicitly to characterize semantics uncertainty due to vagueness in SOLAP. Beside, some works have been initiated to deal with the problem of information vagueness only in On-line Analytical Processing (OLAP) (González et al. 2009; Laurent 2010; Péres et al. 2007; Molina et al. 2006; English et al. 2004; Pedersen & Jensen 2001; Kaya & Alhajj 2006). For instance, Pedersen & Jensen (2001) handles imprecision through non-spatial dimensions in OLAP. Laurent (2010) proposes a multidimensional model within OLAP which deals with fuzzy facts by considering fuzzy relation and partition in dimensions. Kaya & Alhajj (2006) use fuzzy association rules by allowing fuzzy labels derived from membership degrees to define dimensions. Delgado *et al.* (2004) and Molina *et al.* (2006) handle the imprecision by defining fuzzy hierarchies and facts in OLAP, similar to the work presented in Laurent (2010). English *et al.* (2004) employ fuzzy logic to execute a spatiotemporal query by establishing the degree of satisfaction of a space to corresponding selected criteria. Hence, these methods need to be adapted explicitly for spatial aspect of analyzed phenomena in SOLAP systems. Moreover, methods to aggregate and represent spatial measures resulting from the uncertain data and information vagueness found into multiple hierarchical dimensions are still missing.

5.5 Handling Information Vagueness: A Fuzzy Approach

Typically, spatial objects are represented either by crisp or fuzzy models (Fisher et al. 2007). Crisp models are used to represent objects with sharp boundaries such as a building or a road (Schneider 2003a). Contrary to crisp objects, fuzzy objects have broad boundaries (i.e. transition zones) that cannot be observed or determined precisely, or have unclear position which cannot be measured exactly (Cheng et al. 2009). Examples are the limits of a lake or of forest stands and the limits between a mountain and the neighboring objects such as plateau. Fuzzy models are the best fits to represent the broad boundaries of these types of objects (Smith & Mark 2003).

The idea of using Fuzzy Set theory to handle information vagueness led us to a specific type of data model called Fuzzy Spatial Data Model. This model constitutes fuzzy points, fuzzy lines, fuzzy regions, and fuzzy partitions in $\Re^2$ (Schneider 2003b; Pauly & Schneider 2010; Dilo 2006; Molenaar & Cheng 2000; Robinson 2003; Kanjilal et al. 2010; Schneider 2003a). The proposed approach for CERA in this paper is inspired from (Schneider 2003a) which is briefly illustrated in Figure 5.1. Schneider (2003a) defines fuzzy objects based on a finite collection of elements from a regular grid, establishing a discretization of a bounded subspace of $\Re^2$ called fuzzy partition. A membership value is then assigned to each elements of the grid called cell. As we can see from the figure, a regular tessellation is performed on the region under study. Each cell is identified by its center, vertices, and edges. A membership value is assigned to each cell based on assigned membership function to each grid. The membership functions are generally defined by an expert-based approach. A risk zone in this case is generated by aggregating a set of cells with the same values. The overall risk in the region is calculated by aggregating multiple grids of indicators using an appropriate fuzzy operator. In Figure 1, the color hue represents risk value; the dark red colors for higher risk value close to 1 and light blue colors for lower risk values close to 0.

Extending the proposed approach to model information vagueness in spatial multidimensional databases necessitates identifying where the fuzziness happens, how this can be embedded into the database and managed while performing the queries and representing the results. The following sections explain these situations step by step using a CERA conceptual multidimensional model presented in Jadidi et al. (2013).

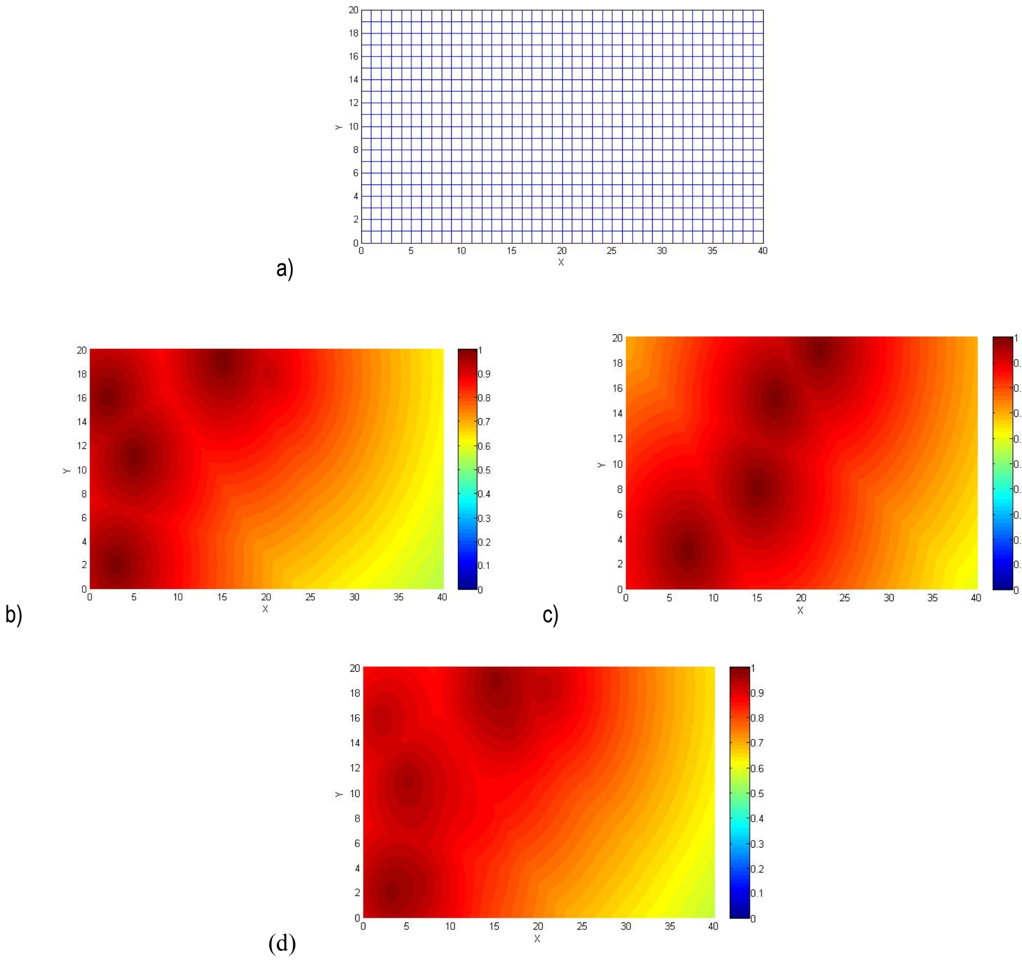


Figure 5.1: Spatial fuzzy representation of risk: (a) Regular tessellation, (b) and (c) Fuzzy spatial risk zones representations of indicator 1 and 2, and (d) the result of fuzzy Spatial Aggregation.

5.6 Fuzzy Spatial Datacube

Spatial datacubes are designed to deal with the cross-tab of hierarchical semantic systems (Bédard et al. 2009). The main elements of a spatial datacube (i.e. level's attributes, spatial levels, spatial members, spatial dimension, spatial hierarchy, spatial measures, and spatial facts) are formally defined in Salehi (2009). Data in a spatial datacube can also be uncertain or vague. Moreover, the information vagueness may also arise in the definition of dimensions, hierarchies, aggregation relationships, aggregation functions and spatial measures (Sboui 2010). Thus, information vagueness should appropriately be handled in a spatial datacube. Embedding information vagueness in the multidimensional model requires redefining the principal elements of the spatial datacube which is explained hereafter. In this regard, a fuzzy approach based on Fuzzy Set Theory is proposed and is applied to a spatial multidimensional model proposed for CERA in Jadidi *et al.* (2013). This model is illustrated in Figure 5.2. It is originally designed to assess coastal erosion risk associated to structures, infrastructures and people at danger in different administrative regions in a coastal area at multiple time periods. The Proposed CERA model consists of 15 dimensions (two spatial dimensions, one temporal and 12 thematic dimensions) and 13 spatial measures (eight measures with geometry, five numeric measures). These measures represent geometrically the risk zones and indicate the number of affected structures or people with regards to the desired indicators or dimensions. These elements are derived from vulnerability index studied by experts based on decision-makers and practitioners' interests (Jadidi et al. 2013).

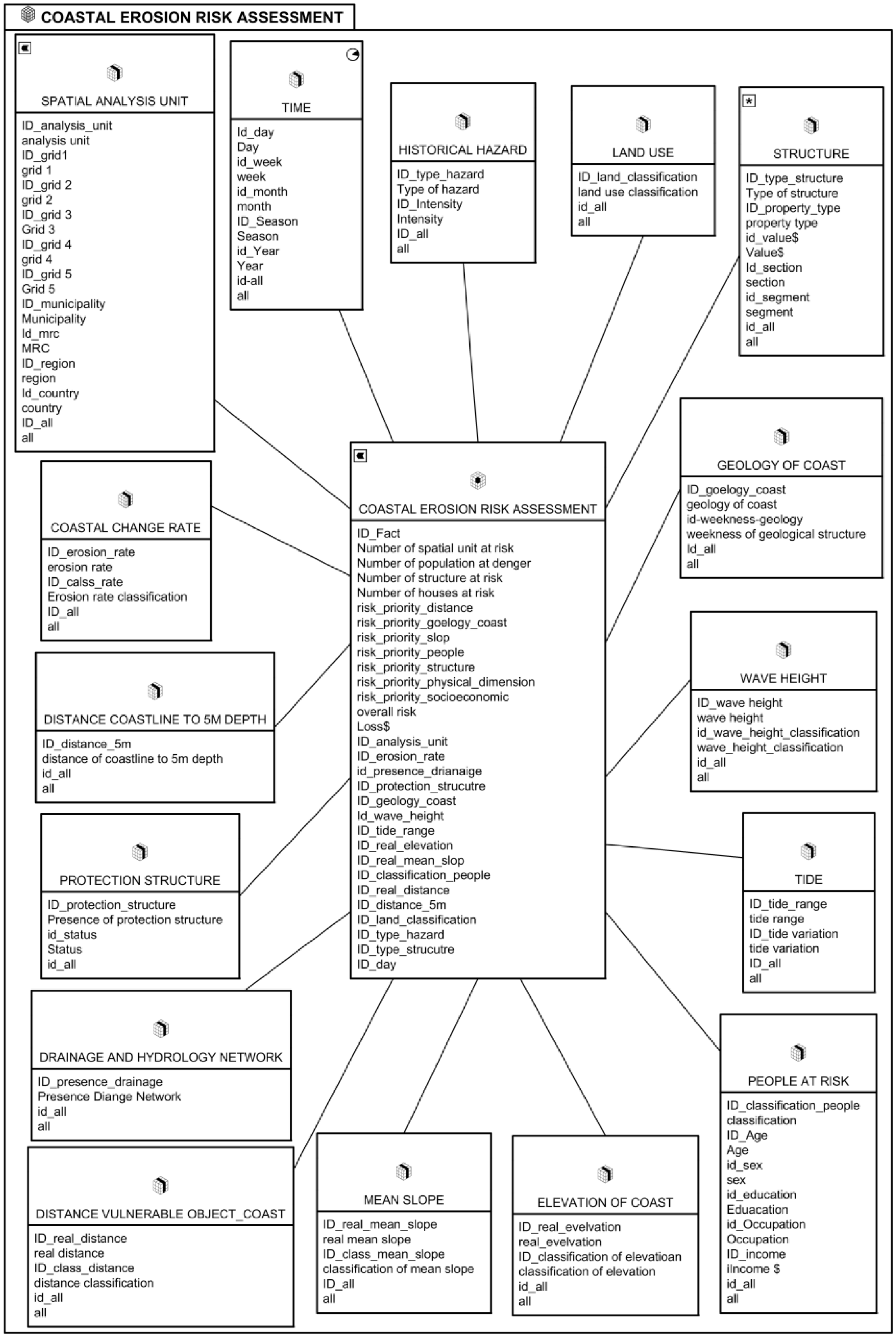

COASTAL EROSION RISK ASSESSMENT
SPATIAL ANALYSIS UNIT
ID_analysis_unit
analysis unit
ID_grid1
grid 1
ID_grid 2
grid 2
ID_grid 3
Grid 3
ID_grid 4
grid 4
ID_grid 5
Grid 5
ID_municipality
Municipality
Id_mrc
MRC
ID_region
region
Id_country
country
ID_all
all
TIME
Id_day
Day
id_week
week
id_month
month
ID_Season
Season
id_Year
Year
id-all
all
HISTORICAL HAZARD
ID_type_hazard
Type of hazard
ID_Intensity
Intensity
ID_all
all
LAND USE
ID_land_classification
land use classification
id_all
all
STRUCTURE
ID_type_structure
Type of structure
ID_property_type
property type
id_value$
Value$
Id_section
section
id_segment
segment
id_all
all
GEOLOGY OF COAST
ID_goelogy_coast
geology of coast
id-weekness-geology
weekness of geological structure
Id_all
all
COASTAL CHANGE RATE
ID_erosion_rate
erosion rate
ID_calss_rate
Erosion rate classification
ID_all
all
COASTAL EROSION RISK ASSESSMENT
ID_Fact
Number of spatial unit at risk
Number of population at denger
Number of structure at risk
Number of houses at risk
risk_priority_distance
risk_priority_goelogy_coast
risk_priority_slop
risk_priority_people
risk_priority_structure
risk_priority_physical_dimension
risk_priority_socioeconomic
overall risk
Loss$
ID_analysis_unit
ID_erosion_rate
id_presence_drianaige
ID_protection_strucutre
ID_geology_coast
Id_wave_height
ID_tide_range
ID_real_elevation
ID_real_mean_slop
ID_classification_people
ID_real_distance
ID_distance_5m
ID_land_classification
ID_type_hazard
ID_type_strucutre
ID_day
DISTANCE COASTLINE TO 5M DEPTH
ID_distance_5m
distance of coastline to 5m depth
id_all
all
WAVE HEIGHT
ID_wave height
wave height
id_wave_height_classification
wave_height_classification
id_all
all
PROTECTION STRUCTURE
ID_protection_structure
Presence of protection structure
id_status
Status
id_all
TIDE
ID_tide_range
tide range
ID_tide variation
tide variation
ID_all
all
DRAINAGE AND HYDROLOGY NETWORK
ID_presence_drainage
Presence Diange Network
id_all
all
PEOPLE AT RISK
ID_classification_people
classification
ID_Age
Age
id_sex
sex
id_education
Eduacation
id_Occupation
Occupation
ID_income
iIncome $
id_all
all
DISTANCE VULNERABLE OBJECT_COAST
ID_real_distance
real distance
ID_class_distance
distance classification
id_all
all
MEAN SLOPE
ID_real_mean_slope
real mean slope
ID_class_mean_slope
classification of mean slope
ID_all
all
ELEVATION OF COAST
ID_real_evelvation
real_evelvation
ID_classification of elevatioan
classification of elevation
id_all
all

Figure 5.2: A Star schema of spatial multidimensional model for CERA (Jadidi et al. 2013)

Level Attribute a_i is defined by the triple of a_i = *(type, nature, domain)* for spatial datacube. The type of attribute a_i is the data type such as numeric (real and integer), textual, date (instant and interval), or geometric (vector-based i.e. point, line, and polygon and raster-based i.e. cell). The nature of the attribute refers to spatial, temporal, or thematic characteristics of the attribute and indicates if it is crisp or fuzzy. Crispness and fuzziness of an attribute may be addressed to its spatial, temporal or thematic nature. The domain of the attribute refers to its value domain. For instance, *location (geometric: cell, spatial-fuzzy, cell in a grid)"* and *name(textual, spatial-fuzzy, {"Site01", "Site02", "Site03", "Site04",…})* are the attributes of a "Spatial Analysis Unit" level in "Spatial Analysis Unit" dimension (see Figure 5.2).

Hierarchy Levels describe the granularity of analysis along a dimension, and are conveyed by $l=\{a_1 \ldots, a_n\}$, where l is the name of the level and $\{a_1 \ldots, a_n\}$ are the set of its attributes. A level of a dimension can be spatial or non-spatial. For instance, the "Municipality" level in "Spatial Analysis Unit" dimension in the proposed CERA model (Figure 2) is defined as *Analysis Unit={name,location}*. In a fuzzy datacube, an instance of a level is a **Fuzzy Member** of that level. The membership value $\mu \in [0,1]$ for each member is calculated through expert-based IF-THEN rules or an assigned membership function. A member m of level l is therefore defined by quadruple $m = \langle AT_m, V, \mu, f(a_i, v_i) \rangle$. In this definition, $AT_m = \{a_1 \ldots a_n\}$ is the set of members' attributes that includes a subset of levels' attributes, $V = \{v_1 \ldots v_n\}$ is the set of domain values of attributes, $\mu \in [0,1]$ is the membership degree of m in l, and $f(a_i, v_i)$ is the function from elements in AT_m to elements in V. The members are either geometrical or non-geometrical. For instance, "Spatial Analysis Unit" is a geometric member with attributes *(name: Site05, location: grid (i,j), μ=0.56, …)*, "Erosion Rate Classification" is a non-geometric member with attributes *(calss1, location: grid (i,j), μ=0.87,…)*, and "Year" a non-geometric member with attributes *(date: 2006,…)*.

Dimension includes a number of related hierarchy levels. The levels are ordered from detailed to general and build a hierarchy of abstraction levels. Bimonte *et al.* (2012) define an instance of a dimension as pair $d = (L, \leq)$. $L = \{m_1 \ldots m_n\}$ is a set of members with an order (roll-up) relation (≤)between these members for a level L. A dimension can be spatial (geometric, non-geometric (i.e. using place names), and mixed), temporal or thematic (Salehi et al. 2010) and may have one or more alternate hierarchies $h_1 = (L_{h_1}, \leq)$, $h_2 = (L_{h_2}, \leq)$, For instance, "Spatial Analysis Unit" in the presented CERA model is a spatial dimension (see Figure 3) that has two hierarchies (h_1, h_2):

h_1: a vector-based data: “Spatial Analysis Unit” < “Municipality” < “MRC” < “Region” < “Province” < “Country” < All and

h_2: a raster-based data: “Spatial Analysis Unit” < “Grid 1 (with very small cell)” < “Grid 2 (with small cell)” < “Grid 3 (with medium cell)” < “Grid 4 (with large cell)” < “Grid 5 (with very large cell)” <All.

Each cell of a raster-based hierarchy level of data is related to a unit of the vector-based administrative level which in its turn has several attributes such as location, name, and geometry called levels’ attributes.

The fuzziness in a dimension comes from relation between hierarchy levels or the belongingness of a member to a hierarchy level (Laurent 2010). To define fuzzy relation between hierarchy levels, the spatial fuzzy partitions are presented as $grid_i$ with a specific level of hierarchy that links them to their respective vector-based administrative data. A grid is superimposed to each level of administrative data (vector-based) while each cell of grid can be overlapped by many of them in an information system. However, it is only attributed to one of them in practice.

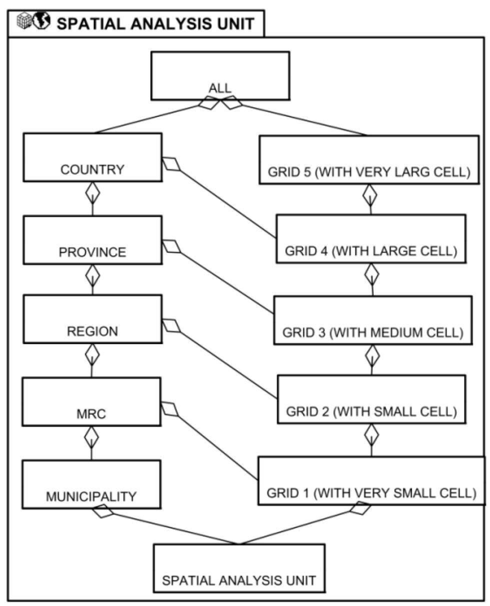


Figure 5.3: Spatial Analysis Unit in presented model for CERA (Jadidi et al. 2013).

Belongingness of a member to each level of hierarchy can also be fuzzy (Laurent 2010) Dimension “Time” is as an example of a fuzzy hierarchy dimension based on a fuzzy relation. Two hierarchy levels are defined; one with respect to the calendar called “Day” and the other with respect to season change called “Season”. In the proposed CERA model, “Season” is semantically defined with regards to “Day” based on a set of rules as

Early Spring (snow melting period), late spring (without snow), summer, early fall, late fall, and Winter. The belongingness of “Day” to Season is then determined by a membership degree ν . For instance, March 15 may belong 70% to early spring and 30% to winter in a specific year. This belongingness changes also from year to year that is typically the case in Quebec, Canada.

Therefore **Fuzzy Spatial Dimension** is a tuple $D = \langle L, (\leq, \nu) \rangle$ where $L = \{m_1, m_2, ..., m_n\}$ is the set of members, ≤ the relationship between the members for a level in *L*, and membership degree ν , that represents the belongingness of a level to its successor.

Fuzzy Spatial Measures are fuzzy spatial attributes that are analyzed based on different levels of dimensions and their respective membership degrees. To calculate measures, one or a series of fuzzy operators (Union, Intersection, Difference, Mean, etc.) are applied to select the necessary fuzzified dimensions. According the CERA model in Figure 2, the measures are fuzzy for each cell of the grid that is associated to the respective administrative **vector-**based regions. A series of these cells with the same membership value form the risk zones. Spatial measures are generally represented either numerically or geometrically. In the presented CERA model for instance, “number of structures at risk” is a numerical measure described as *number of structures at risk*= *(numeric: integer, spatial, natural number)*. Another example is “overall risk” that is a geometrical measure described by a geometric spatial attribute (grids in this study) as: *overall risk*= *geometric: cell, spatial, set of cell representing the risk zones regarding to calculated membership degree).* Reminding that, a fuzzy measure can be performed onto the combination fuzzy dimensions and conventional dimensions (where the values are exact with an acceptable degree of precision).

A Fact describes an event or interest of decision-makers to process. For instance “Coastal Erosion Risk” is considered as a fact in our CERA model. In this regard, a **Fuzzy Spatial Fact** is defined as pair $F = \langle M, Ms \rangle$ of finite set of members of dimensions $M = \{m_1, ..., m_n\}$ and finite set of fuzzy measure values $Ms = \{Ms_1, ..., Ms_m\}$ calculated with respect to members of *M* and their respective membership degrees $\mu \in [o,1]$. A fuzzy spatial fact can be represented geometrically or numerically.

Fuzzy Spatial Datacube is defined as pair $FSC = \langle D, F \rangle$ of finite set of dimension instances $D = \{d_1, ...d_n\}$ and finite set of facts $F = \langle M, Ms \rangle$ over dimension instances, while at least there are one fuzzy spatial dimension and fuzzy spatial facts. Each cell of the datacube contains a measure or a set of measures that are semantically aggregated using fuzzy operators (Bédard et al. 2009). Dimensions are structured in hierarchies that support aggregation process.

5.7 Fuzzy Aggregation in Spatial Datacube: A Multi-scale Representation

Aggregation in Geospatial Business Intelligence (GeoBI) community is the grouping of data geometrically, thematically, or semantically to a coarser level of detail (Pedersen et al. 2001; Gomez et al. 2009). This concept of aggregation is very different from the map generalization process (Bédard et al. 2007) and is not the same as in object-oriented modeling (Laurent 2010). In fact, aggregation in GeoBI is a summarization process of values or geometries in a datacube that directly depends on the data model used (Péres et al. 2007; Laurent 2010; Gomez et al. 2009; Pedersen et al. 2001). It typically uses SUM, AVG, MIN, MAX, COUNT and similar operators, but also more complex ones such as spatial operators (e.g. overlay, intersect, include, and fusion), and advanced statistical formula or simulation algorithms. In addition, the geometry used at the different levels of abstraction often comes from different datasets.

Using fuzzy concepts to define appropriate operators for data aggregation in a datacube has been initiated by Laurent 2010 and Molina et al. (2006). A series of operators such as roll-up, drill-down, slice, dice, and pivot have been defined for fuzzy datacubes in Molina et al. 2006 and Martin-bautista et al. (2013) using both quantitative and qualitative data. This permits a qualitative representation of results on charts and tables. The thematic aggregation can principally be performed based on Laurent 2010 and Molina et al. (2006). However, the geometric aggregation involving spatially fuzzy or crisp members requires redefining fuzzy operators such as overlay and fusion for fuzzy spatial objects. Spatial relations in fuzzy aggregation follows the ISO standard model (Dilo 2006) and the true/false values of the ISO relations have been extended to a fuzzy degree between 0 and 1 that represents its truthiness. The fuzzy operators are equivalent to their crisp counterparts when applied to crisp objects, since crisp objects have maximum degree of membership of 1.

5.7.1 Fuzzy Overlay

Fuzzy overlay superimposes multiple dimensions' information, here fuzzy partitions, to combine, modify, or update attributes and their membership values, resulting a new fuzzy partition obtained from the union, intersection, difference, any arithmetic operations, or mixed of them of a cell of first partition with cell of second partition and so on (Dilo et al. 2007; Schneider 2003a). Overlay is the key operator to compute measures in a spatial datacube that defines fuzzy partitions by combining their fuzzy membership functions (Laurent 2010; Gomez et al. 2009; Molina et al. 2006). It explores the likelihood of a cell being a member of each set defined by multiple dimensions. There are typically two methods for overlay operation i.e. vector-based or raster-based. Since, the proposed fuzzy approach in this paper is grid-based, so fuzzy raster based overlay operators are discussed hereafter.

Fuzzy raster-based overlay quantifies each location's possibility of belonging to specified cell sets from multiple fuzzy partitions. Fuzzy union, intersection, difference, operators are commonly used with the combination of some arithmetic's operators such as SUM, Mean, and Weighted Mean to perform overlay operation. For instance, representing the spatial measure *"risk_priority_structure"* (see Figure 2) is an overlay operator by performing a combination of fuzzy union and Weighted Mean operators. The result of *"risk_priority_structure"* measure is map that visualizes fuzzy risk zones by emphasizing the vulnerable infrastructures and structures such as road, railway and houses. In addition, to obtain the measure of "Number of Population in danger" (see Figure 2) we need to perform overlay operator to identify risk zones and the count the number of people in danger.

5.7.1.1 Fuzzy Union

Fuzzy union operator results a new fuzzy object from combining multiple fuzzy objects, fuzzy partitions in this paper, with the membership values of whom with higher membership degree (Dilo et al. 2007; Schneider 2003a). In other words, fuzzy union builds a new fuzzy region by taking the cells with higher membership value inside any overlapping area. Table 5.1 illustrates the formal definition of fuzzy union and an example of numerical and geometrical representation. Two fuzzy regions, here fuzzy partitions are presented in Table 5.1. The result of union operator in Table 5.1 is a new fuzzy region including both fuzzy regions of fuzzy regions of 1 and 2.

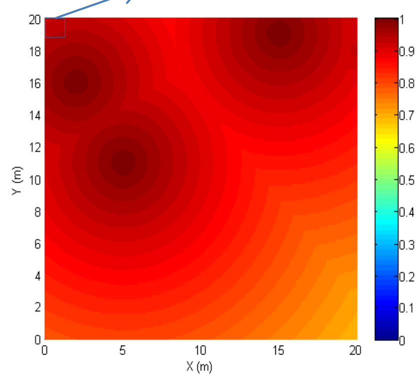
Table 5.1: The formal definitions of fuzzy union, example of numerical values and their geometrical representations

Union

$$
\begin{aligned}
&FuzzyOverlay : FuzzyPartition \times FuzzyPartition \rightarrow FuzzyPartition \\
&\forall P_1 = \{\mu_i\}_{i=1}^{n}, P_2 = \{\nu_j\}_{j=1}^{n} \in FuzzyPartition, \\
&FuzzyOverlay(P_1, P_2) = \left\{ \begin{array}{l} \zeta_{i,j} \mid i \in \{1 \ldots n\}, j \in \{1 \ldots m\}, \\ \zeta_{i,j} = Union(\mu_i, \nu_j) = Max(\mu_i, \nu_j) = \mu \cup \nu \end{array} \right\}
\end{aligned}
$$

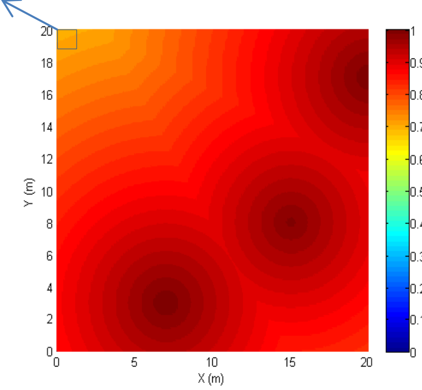
Fuzzy Region 1

1.0000	0.7856	0.7858	0.7861
0.7859	0.7862	0.7864	0.7867
0.7865	0.7867	0.7870	0.7872
0.7870	0.7873	0.7876	0.7878

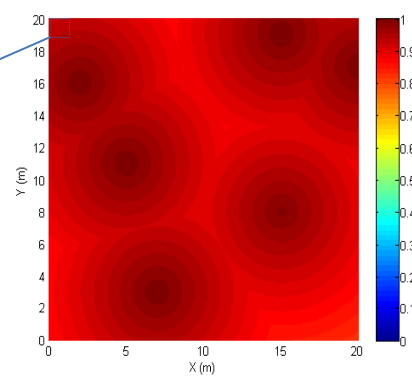


Fuzzy Region 2

1.0000	0.8593	0.8600	0.8606
0.8590	0.8596	0.8603	0.8609
0.8593	0.8599	0.8605	0.8612
0.8595	0.8602	0.8608	0.8614



1.0000	0.8593	0.8600	0.8606
0.8590	0.8596	0.8603	0.8609
0.8593	0.8599	0.8605	0.8612
0.8595	0.8602	0.8608	0.8614



5.7.1.2 Fuzzy Intersection

Fuzzy intersection operator results a new fuzzy object from combining multiple fuzzy objects, here fuzzy partitions, with the membership values of whom with lower membership degree (Dilo et al. 2007; Schneider 2003a). That means fuzzy intersection produces a new fuzzy region by taking the cells with lower membership value inside any overlapping area. Table 5.2 illustrates the formal definition of fuzzy intersection and an example of numerical and geometrical representation. Two fuzzy regions, here fuzzy partitions are presented in Table 5.2. The result of intersection operator in Table 5.2 is a new fuzzy region including both fuzzy regions of fuzzy regions of 1 and 2.

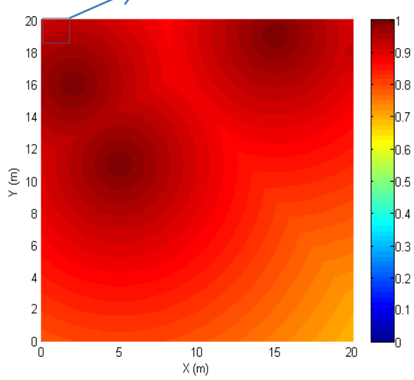
Table 5.2: The formal definitions of fuzzy intersection, example of numerical values and their geometrical representations

Intersection

$$FuzzyOverlay : FuzzyPartition \times FuzzyPartition \rightarrow FuzzyPartition$$
$$\forall P_1 = \{\mu_i\}_{i=1}^{n}, P_2 = \{\nu_j\}_{j=1}^{n} \in FuzzyPartition,$$
$$FuzzyOverlay(P_1, P_2) = \begin{Bmatrix} \zeta_{i,j} \mid i \in \{1 \ldots n\}, j \in \{1 \ldots m\}, \\ \zeta_{i,j} = Intersection(\mu_i, \nu_j) = Min(\mu_i, \nu_j) = \mu \cap \nu \end{Bmatrix}$$

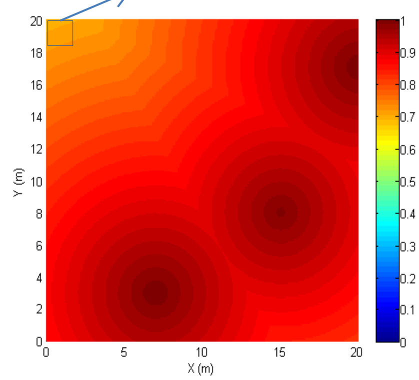
Fuzzy Region 1

1.0000	0.7856	0.7858	0.7861
0.7859	0.7862	0.7864	0.7867
0.7865	0.7867	0.7870	0.7872
0.7870	0.7873	0.7876	0.7878

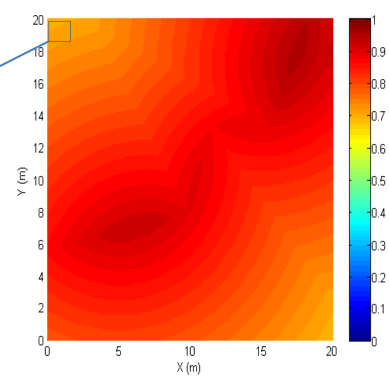


Fuzzy Region 2

1.0000	0.8593	0.8600	0.8606
0.8590	0.8596	0.8603	0.8609
0.8593	0.8599	0.8605	0.8612
0.8595	0.8602	0.8608	0.8614



1.0000	0.7856	0.7858	0.7861
0.7859	0.7862	0.7864	0.7867
0.7865	0.7867	0.7870	0.7872
0.7870	0.7873	0.7876	0.7878



5.7.1.3 Fuzzy Difference

Fuzzy difference results a new fuzzy object by intersecting of the first fuzzy partition with the complement of the second. In fact, fuzzy difference calculates a new membership value as their difference and considers it as the result value of the partition for the common location (Dilo et al. 2007; Schneider 2003a). SUM, Mean, and Weighted Mean operators follow the arithmetic theory to calculate the final result. Table 5.3 illustrates the formal definition of fuzzy difference and fuzzy arithmetic operators such as Mean, and Weighted Mean as well as their numerical and geometrical representations. Two fuzzy regions, here fuzzy partitions are presented in Table 5.3. The result of indicated operators in Table 5.3 is a new fuzzy region including both fuzzy regions of fuzzy regions of 1 and 2.

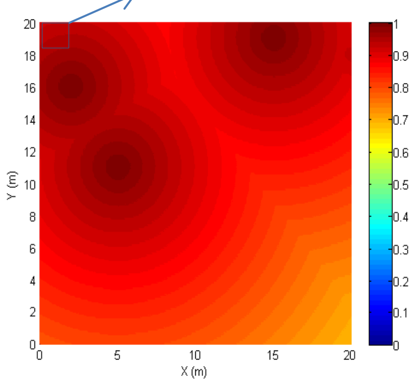
Table 5.3: Fuzzy difference, Mean and Weighted Mean, their formal definitions and examples of numerical values and their geometrical representations

Diffrence

$$FuzzyOverlay : FuzzyPartition \times FuzzyPartition \to FuzzyPartition$$
$$\forall P_1 = \{\mu_i\}_{i=1}^{n}, P_2 = \{\nu_j\}_{j=1}^{n} \in FuzzyPartition,$$
$$FuzzyOverlay(P_1, P_2) = \begin{Bmatrix} \zeta_{i,j} \mid i \in \{1...n\}, j \in \{1...m\}, \\ \zeta_{i,j} = Differnce(\mu_i, \nu_j) = Intersection(\mu_i, 1^{\Re^2} - \nu_j) = \mu - \nu \end{Bmatrix}$$

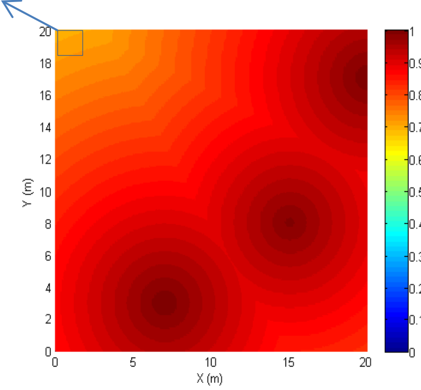
Fuzzy Region 1

1.0000	0.7856	0.7858	0.7861
0.7859	0.7862	0.7864	0.7867
0.7865	0.7867	0.7870	0.7872
0.7870	0.7873	0.7876	0.7878

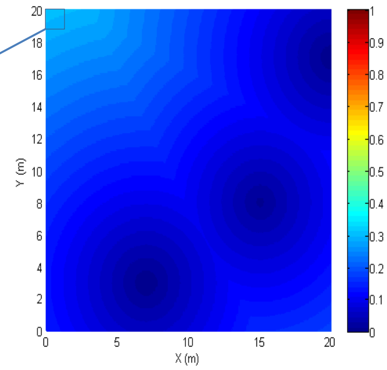


Fuzzy Region 2

1.0000	0.8593	0.8600	0.8606
0.8590	0.8596	0.8603	0.8609
0.8593	0.8599	0.8605	0.8612
0.8595	0.8602	0.8608	0.8614



1.0000	0.1407	0.1400	0.1394
0.1410	0.1404	0.1397	0.1391
0.1407	0.1401	0.1395	0.1388
0.1405	0.1398	0.1392	0.1386



Artithmetic Operators: Mean and Weighted Mean

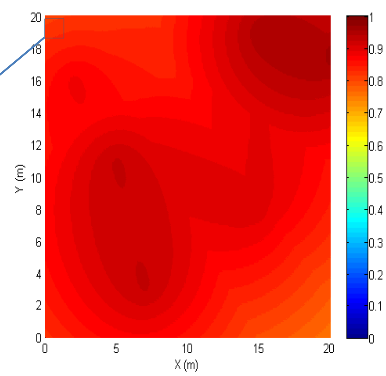
$$FuzzyOverlay : FuzzyPartition \times FuzzyPartition \rightarrow FuzzyPartition$$

$$\forall P_1 = \{\mu_i\}_{i=1}^{n}, P_2 = \{\nu_j\}_{j=1}^{n} \in FuzzyPartition,$$

$$FuzzyOverlay(P_1, P_2) = \left\{ \begin{array}{l} \zeta_{i,j} \mid i \in \{1 \ldots n\}, j \in \{1 \ldots m\}, \\ \zeta_{i,j} = Mean(\mu_i, \nu_j) = \dfrac{\mu_i + \nu_j}{2} \\ \text{OR} \\ \zeta_{i,j} = WeightedMean(\mu_i, \nu_j) = (\omega_i \times \mu_i) + (\omega_j \times \nu_j) \end{array} \right\}$$

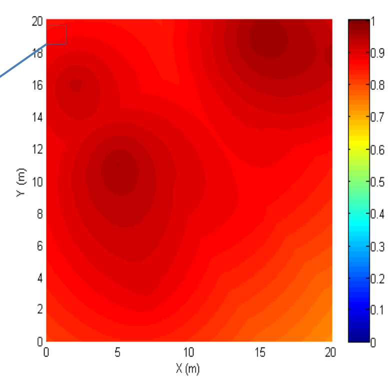
Mean

1.0000	0.8225	0.8229	0.8234
0.8224	0.8229	0.8233	0.8238
0.8229	0.8233	0.8238	0.8242
0.8233	0.8237	0.8242	0.8246



Weighted Mean

1.0000	0.8077	0.8081	0.8084
0.8078	0.8082	0.8086	0.8089
0.8083	0.8087	0.8090	0.8094
0.8088	0.8092	0.8095	0.8099



5.7.2 Fuzzy Fusion

Fuzzy fusion operator dissolves a fuzzy partition by merging the cells based on grouping, likeness, or equality of some attributes value of the regions (Dilo et al. 2007; Schneider 2003a). In other words, fusion operator reduces the resolution of grid by increasing the grid cell size multiplying a factor to obtain desired level of hierarchies and scale. Indeed, the value of this factor should be integer and greater than 1. For example, the value of three as factor would result in an output cell size three times larger than that of the input grid. To perform fuzzy fusion operator, two situations can be happened:

1) Performing overlay operation on multiple layers of information on finer level, and then creating merged cells in coarser scale,

2) Performing merged-cell operation and then assign the result of overlay operation to obtain the final result in coarser scale.

The result of a fusion operator is a new fuzzy region. The Syntax of Fusion operator is as follows:

FuzzyFusion (Input grid, cell_factor, {Fuzzyoverlay_type}, output grid)

For example, to represent the spatial measure "overall risk" in multiple level of hierarchies (rolling from grid 1 to grid 2), fusion operator is required allowing a multi-scale representation of overall risk zones. The numerical and geometrical examples of such operation are provided in Table 5.4. The resolution of resulted cells (cell factor) is 10 times larger than the original one.

Table 5.4: The formal definitions of fuzzy fusion operators, example of numerical values and their geometrical representations

Fuzzy Fusion

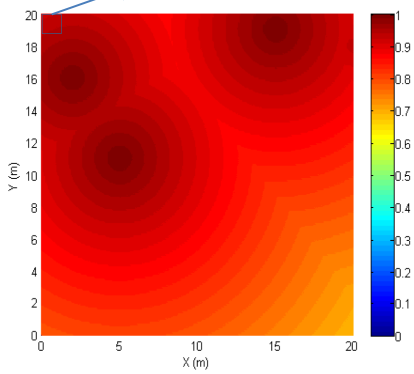
$$FuzzyFusion : FuzzyPartition \times FuzzyPartition \rightarrow FuzzyPartition$$
$$\forall P_1 = \{\mu_i\}_{i=1}^{n}, P_2 = \{\nu_j\}_{j=1}^{n} \in FuzzyPartition, and\ Cell\_factor = C$$
$$FuzzyFusion(P_1, P_2) = \left\{ \begin{array}{l} \zeta_{i,j} \mid i \in \{1...n\}, j \in \{1...m\}, andC \\ \zeta_{i,j} = FuzzyOverlay(\mu_i, \nu_j) \\ \end{array} \right\}$$

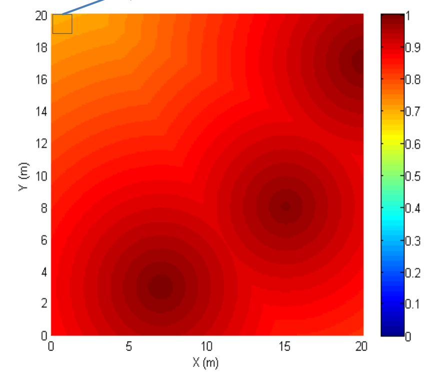
Fuzzy Region 1

1.0000	0.7856	0.7858	0.7861
0.7859	0.7862	0.7864	0.7867
0.7865	0.7867	0.7870	0.7872
0.7870	0.7873	0.7876	0.7878

Fuzzy Region 2

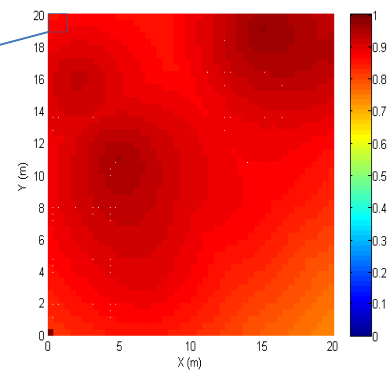
1.0000	0.8593	0.8600	0.8606
0.8590	0.8596	0.8603	0.8609
0.8593	0.8599	0.8605	0.8612
0.8595	0.8602	0.8608	0.8614





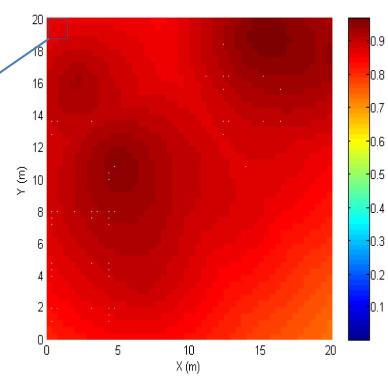
Based on Union Operation, Cell_Factor= 10

1.0000	0.8187
0.8198	0.8236



Based on Mean Operation, Cell_Factor= 10

0.8131	0.8148
0.8160	0.8197



5.7.3 Five Possible Scenarios of Aggregation in a Fuzzy Spatial Datacube

Based on the presented fuzzy operators, five aggregation scenarios are proposed to support a fuzzy model in such spatial multidimensional datacube:

Scenario 1 (Crisp Aggregation onto Crisp Data): This is what is typically done with any traditional datacube (Salehi et al. 2010; Bédard et al. 2009).

Scenario 2 (Fuzzy Aggregation onto Crisp Data): It is the case when a dimension includes precise data such as “Elevation of Coast” dimension or road network in “Structure” Dimension (See Figure 2) while the rules to calculate the measure "degree of risk" from all dimensions. In this case, members are considered as crisp and fuzzy aggregation is performed to calculate measure. The result of this scenario is a fuzzy risk zones where fuzziness happens in elaborating risk formula to calculate measure.

Scenario 3 (Crisp Aggregation onto Fuzzy Data): This scenario is only feasible if fuzzy data are translated into crisp data that results in the loss of information on vagueness through the aggregation process. In this case, we find ourselves again in the first scenario having traditional datacubes with crisp data onto which crisp operators are applied.

Scenario 4 (Fuzzy Aggregation onto Fuzzy Data): This scenario is applicable when information vagueness is characterized as degree of membership in a spatial fuzzy datacube. For instance in the CERA (see Figure 2), the available data and information to calculate vulnerability index and to assign a degree of risk to each indicator and cell in spatial grid are expert-based with degree of vagueness. The semantics uncertainty therefore appears in each member of dimensions and its hierarchy level. The defined fuzzy operators can then be employed to calculate required measures and represent risk zones spatially.

Scenario 5 (Fuzzy Aggregation onto mixed of Crisp and Fuzzy Data): This scenario happens when a mixed of fuzzy and crisp dimensions should be analyzed that is the case of CERA model (see Figure 2). This situation brings us to a mixed of scenarios of 2 and 4, where a membership degree of 1 assigned to crisp dimensions, then the fuzzy aggregation operators are applied to calculate desired measures. For instance, the measure of “risk_prioriry_socioeconomic” requires dealing with such situation where the precise census data should be integrated with fuzzy spatial dimension of “Spatial Analysis Units” and so many other fuzzy dimensions which are classified based on expert-knowledge. The scenario 5 is a general case of the other scenarios that demonstrates the generality of the proposed solution to handle information vagueness based on fuzzy model in the context of spatial datacubes.

5.8 Discussion

In a fuzzy spatial datacube, spatially referenced facts and their respective membership values are stored. Assigning the label of “fuzzy” to a spatial datacube requires at least one dimension of the datacube to be defined based on fuzzy partition or fuzzy hierarchy relation.

The proposed fuzzy spatial datacube for CERA (see Figure 2) is well adapted for both vector and grid-based information. Spatial data is stored in a grid-based structure that is associated to a vector-based data structure e.g. census division in a spatial dimension. Performing the aggregation on this dimension permits navigating from the grid-based to the vector-based structure and vice versa. This solves systemically the problem of spatial aggregation-generalization mismatch (Bédard et al. 2007). Furthermore, having a grid for each hierarchy level in a spatial dimension is an efficient way to design raster spatial datacube with increased on-the-fly spatial analysis capabilities (Plante 2013).

Some fuzzy spatial aggregation operators are presented in this paper. The overlay and fusion of fuzzy partitions use respectively the concept of the intersection and the union of fuzzy objects in fuzzy set theory with a combination of arithmetic's operators such as SUM, Mean, and Weighted Mean. The overlay operator combines two fuzzy regions to form a new fuzzy region whilst the fusion operator allows the generalization of a fuzzy region.

Five feasible scenarios are presented which are supported by the proposed fuzzy spatial datacube model in this paper using proposed aggregation operators. These scenarios are as follows:

1) Crisp aggregation onto crisp data;

2) Fuzzy aggregation onto crisp data;

3) Crisp aggregation onto fuzzy data;

4) Fuzzy aggregation onto fuzzy data;

5) Fuzzy aggregation onto mixed of crisp and fuzzy data.

The implementation of the proposed spatial datacube for CERA (see Figure 5.2) is identical to the implementation of typical crisp spatial datacubes. The only difference is in measures calculation in which a fuzzy operator instead of a crisp one should be applied to respective dimensions, members, and hierarchy relations. The fuzzy measures remain values like crisp measures although they have a membership value which is itself a value like crisp values (although between 0 and 1) and is used to properly interpret a fuzzy measure. Such an interpretation can be achieved by reading the membership value of a measure or by using one additional visual variable (ex. transparency, pattern, chroma) in thematic maps, tables, or charts.

5.9 Conclusion

Fuzzy spatial datacubes are essential to perform more comprehensible knowledge discovery for effective decision-making. An example in this regard is assessing the risks caused by natural phenomena like erosion in coastal regions. A fuzzy-logic-based approach was proposed in this paper to deal with information vagueness originated from the uncertainty of an object and its geometry definition. This concept was then embedded into a spatial datacube through redefining the spatial datacube elements (dimensions, members, hierarchies, measure and facts) as fuzzy dimensions, fuzzy members, fuzzy hierarchies, fuzzy measures, fuzzy facts and required fuzzy aggregation operators (union, intersection, difference, overlay, and fusion). A multidimensional model to perform CERA (see Figure 5.2) was used throughout this paper to demonstrate the design of a fuzzy spatial datacube. One of the main advantages of using a fuzzy spatial datacube is the capability to present and report the results to end-users using linguistic expressions. Another advantage is representing separately the level of uncertainty and vagueness of the calculated measures for a more realistic decision-making. These two aspects were not the subject of the current paper. More studies are required to establish them in a more explicit way which can be potential opportunities for future works.

"What you get by achieving your goals is not as important as
what you become by achieving your goals."
Henry David Thoreau

Chapter 6 Conclusion and Future Work

This thesis aimed to present the fundamental concepts towards the development of fuzzy spatial datacubes. The main contribution of this thesis is proposing an approach on the combining Geospatial Business Intelligence (GeoBI) paradigm and fuzzy concept considering the presence of the spatial uncertainty. The proposed approach is applied for a coastal erosion risk assessment process. This addresses the problems of integrating spatiotemporal components of coastal erosion risk and the spatial uncertainty that appears during modeling geospatial data as information vagueness. As stated before, one of the main challenges of geospatial technology is to reduce the time lag between data acquisition and decision-making as well as taking into account the inherent spatial uncertainty within user-friendly and efficient systems. In this thesis, we progressed in this direction by developing the fundamental concepts as our contributions.

Chapter 1 of this thesis describes the research project including the problem statements, elaboration of the main and specific objectives, and proposed methodology. As the main objective of this research project, we needed to scrutinize some fundamental concepts related to geospatial business intelligence systems, spatial data models and integration, risk characteristics, risk assessment methods, risk representation, uncertainty, Fuzzy Set Theory, and fuzzy aggregation operators. Therefore, **Chapter 2** of this thesis presents an exhaustive literature review of the state-of-the-art methods concerning these topics. Indeed, existing spatial multidimensional databases do not deal with the inherent uncertainty which comes with the integration of multiple sources of data, and appears as information vagueness. In addition, the application under study in this thesis, coastal erosion risk assessment, is a complex process which requires handling large amounts of data and information from several sources with different types and levels of detail. Dealing with such situations was the main concern along this research project. In this regard, a generic framework was proposed to develop a spatial multidimensional conceptual model with an application for coastal erosion risk assessment (**Chapter 3**). Handling inherent uncertainty based on fuzzy set theory was described in **Chapter 4**. Then, the proposed fuzzy approach was combined with the GeoBI paradigm to develop a fuzzy spatial datacube (**Chapter 5**). The following section describes in detail the main contributions of our research project.

6.1 Contributions and Discussion

The main contributions of this thesis are detailed in chapters of 3, 4 and 5. These chapters demonstrate the contribution, proposition, and perspective of this research project.

6.1.1 Spatial Multidimensional Conceptual Model (SMCM) for CERA

The main contribution of Chapter 3 was adapting the GeoBI paradigm for CERA. In fact, the key components of coastal erosion risk were identified and then translated as key elements of a spatial datacube. Back to Objective 1, developing a GeoBI-Based conceptual framework applied to CERA, Chapter 3 presented the development of a comprehensive integrated system for coastal erosion risk assessment. First, the state of the art in prevailing GIS-based CERA methods was presented. The limitations of existing methods were highlighted by concluding that an integrated multi-scale system is needed. An analytical conceptual framework was proposed to perform CERA through the GeoBI paradigm. A Spatial Multidimensional Conceptual Model (SMCM) was developed for CERA through the proposed framework. The model provides a complete and coherent vision of the coastal erosion risk by integrating different spatial, temporal, and thematic dimensions. This model takes into account multiple-categories of potentially vulnerable features and elements at risk as well as their interactions to compute the coastal erosion risk in hazardous regions at a given time. Although the resulting SMCM model aims only to support the main research of this thesis, it also provides a good indication to CERA on a novel way to go beyond existing CERA models.

The key elements of SMCM (dimensions, measures, fact) are derived based on risk components (i.e. hazard, elements at risk, and vulnerability). The SMCM consists of 15 dimensions (two spatial, one temporal and 12 thematic dimensions) and thirteen spatial measures (eight measures with geometry, five numeric measures). The identification of dimensions and their hierarchies in the SMCM is based on the vulnerability index's classifications resulting from experts-based-knowledge. However, the proposed model can easily be adapted based on other vulnerability indices through the proposed framework for any other phenomena and other CERA projects. Thirteen measures are determined with respect to the seven different priorities elaborated of risk formula in SMCM. These measures geometrically represent the risk zones and indicate the number of affected structures or people with regard to the desired indicators or dimensions. A star query model is used to execute pre-defined or user-defined measures. This is an advantage of proposed SMCM, because a vast majority of BI tools support star query models. Also, the star query model is easily understood by users. The integration of multiple criteria with different levels of detail from different stakeholders was facilitated through geospatial BI. This also allows aggregation of qualitative and quantitative information and knowledge to estimate the risk at different levels of hierarchy. It performs the cross-measuring of the information to estimate the risk value at a given time period onto a grid-based spatial unit interconnecting with vector-based administrative boundaries' information.

Two main advantages of proposed SMCM refer to "Spatial Analysis Unit" and "Time" dimensions. The flexibility of "Time" dimension to select the desired time period as a calendar or defined season, with respect to the temperature variation and environmental factors, is an advantage of the proposed model in CERA. Additionally, a hybrid data structure of "Spatial Analysis Unit" dimension permits vector-based administrative division data being linked to its successor defined grid-based risk analysis unit. This aims at allowing the stakeholders and decision-makers to make strategic decisions or take action at the right time and in the right place in order to protect the population, infrastructure and region at risk.

Despite the advantages and achievement of proposed SMCM as well as its flexibility to adapt to other applications, there are some issues and limits that should be taken into account for any CERA project. The data availability is a challenge, which has an important impact to the structure of the resulting datacube and its dimensions, measure, and fact. Some other issues that have been considered in this research project were:

- Lack of data in the required details of spatial, temporal or thematic aspects that impact the quality of final results or may add some inherent uncertainty;
- Lack of an integrated vulnerability index in the region under study: this issue is handled by integrating commonly used vulnerability indices around the world; and
- Inconsistency of census parcels with spatial analysis units to calculate the precise number of people in risk zones: possible solutions are performing ecumene maps (if available) or counting the number of people, houses, or buildings.

All stated issues add some uncertainty to risk assessment process. To overcome these issues, in Chapter 4, a novel approach based on fuzzy set theory was introduced and, in Chapter 5, is adapted within spatial datacubes. The details are provided in next sections.

Additionally, technological constraints may limit the number of dimensions in a datacube. An intermediate solution used in this thesis was integrating dimensions with the same typology. Examples are integrating “geology type” with “the presence of weaknesses in geological zones” and integrating “mean tide variation (mm/yr)” with “tide value (mm/day)”.

6.1.2 Fuzzy Spatial Representation of Risk Zones

The main contribution of Chapter 4 was developing a generic approach for dealing with spatial uncertainty within risk zone representation based on fuzzy set theory. In this regard, it was important to first characterize spatial uncertainty during coastal risk assessment process. This spatial uncertainty comes from lack of knowledge (Walker et al. 2003) and the imperfection of large amounts of data and information. Indeed, Chapter 4 is concerned with issues related to Objective 2 of our research project: “developing a novel approach based on Fuzzy Set Theory to improve the representation of uncertainty inherent to risk zones representation”.

In fact, Chapter 4 has focused on the improvement of spatial representation of coastal erosion risk by taking into account the inherent uncertainty related to spatial objects and risk zone representation. This inherent uncertainty was identified as information vagueness and explicitly dealt with through Fuzzy Set Theory. A regular tessellation of the region was generated for each vulnerability indicator by assigning an appropriate membership value to each cell indicating the degree of risk. Membership function of each vulnerability indicator was derived from the classification of vulnerability index, elaborated in Chapter 3. The definition of membership functions is resolved by converting the crisp classification of vulnerability index to a fuzzy classification. Accordingly, the integration of the multiple criteria is performed by aggregating their respective membership values using fuzzy aggregation operators. IF-THEN rules were defined to aggregate multiple layers of indicators by using aggregation operators such as Union, Intersection, Mean, and Mean Weighted. Therefore, with respect to the needs of decision-makers and the emergence of response actions, the choice of these operators can also be varied.

Using regular cells for the tessellation avoids the difficulties of combining different membership values to compute fuzziness where various criteria lead to fuzziness inside of the objects. Indeed, several studies confirm that the fuzzy approach works well with raster format. Elaborating the risk formula and then constructing IF-THEN rules of the associated indicators allow a direct control on the entire CERA process. In addition, this provides more flexibility if one or some indicators or their classifications are changed. In this case, updating the desired information by re-running the fuzzification step or modifying the IF-THEN rules by re-executing the fuzzy aggregation step will be sufficient.

Fuzzy risk representation is a relatively new concept for decision-makers. Traditional risk assessment methods lead to crisp representations of the risk and hence to crisp decisions (i.e. “Yes” or “No”) while the fuzzy approach leads to smooth transitions between these two extremes. In this new context, decision-making processes can adapt and add more flexibility to how to handle an issue. Adapting the decision-making culture towards fuzzy results necessitates finding evidence to convince the decision-makers of the benefits of this new approach. The defuzzification step, explained briefly in Chapter 4, was an alternative in this regard, rendering fuzzy results to be understandable to decision-makers. A risk tolerance measure method based on a crisp compliance guideline is another option already available in some domains such as health system (Kentel & Aral 2007).

The proposed fuzzy approach was applied to a study site in Perce, Quebec, Canada. . As illustrated in Figure 4.11 in chapter 4 about 800m of 132 national roads are severely at risk (north, center, and southwest). The residential area (total of 4 houses in yellow circles) and two motels (green circles) on the nose of Perce Rocks are also at very high risk. From the nose of Perce Rock towards the southwest, four other motels are at high risk as well. This region is reported as an active cliff coastline with 0 m/year of erosion rate. However, due to its geology type, the erosion may happen suddenly in the form of a landslide. The main difference between our method and other studies is how the spatial uncertainty related to risk modeling is handled through a fuzzy approach. Replacing the crisp classification of the vulnerability index with continuous membership functions brings the result closer to the existing reality. Indeed, the obtained results of this study are more consistent with human reasoning and perception by conveying the level of risk in a continuous and smooth manner. The continuity is not only handled by raster format but is also carried out by the fuzzy representation.

A Matlab code was developed to perform tessellation and fuzzy representation steps. In addition, the membership functions of multiple indicators are defined within Matlab. In this thesis, recalling the appropriate fuzzy membership values for each indicator in the Matlab code is done manually. There is potential to automate this step that requires an intermediate code. ArcGIS 10 is used to prepare the data and to visualize the output of fuzzy representation (Matlab code) on top of a Basemap and some useful information using geo-referencing operation. Migrating to Python programming language which is supported by many GIS tools will contribute to the automation of tessellation, membership function definition, fuzzy representation, and, visualization.

However, the proposed method in Chapter 4 was tested only on regular tessellation. The neighborhood relation is implicit, based on the ID of a cell. Additionally, the temporal aspect of the fuzzy object is not taken into account in this approach. This means that the fuzzy risk zones are represented spatially as a snap-shot of a given time period. How to handle fuzzy objects which change in different time periods needs more investigation. Also, the proposed approach is employed only on a small region with a given level of details (scale). When analyzing extremely large amounts of data within a hierarchical system, the proposed approach needs to be adjusted with respect to the selected technology. In this regard, efforts are mainly needed for fuzzy aggregation operators such as "Fusion", if the multi-scale representation is required. These latter were discussed in Chapter 5.

6.1.3 Development of a Fuzzy Spatial Datacube with an Application for CERA

Chapter 5 of this thesis responds to Objectives 1 and 3 of our research project. The concept of fuzzy set theory is integrated with GeoBI paradigm. In this regard, the term of "fuzzy spatial datacube" was coined in this chapter. The issue of information vagueness coming from integrating multiple sources of data in datacube systems was addressed. Also, ignoring such vagueness by depicting the data with exact semantics, geometry, and temporality in datacubes was considered. As presented in Chapter 4, fuzzy logic was a promising approach that has been explored for almost two decades (Schneider 2010). This approach has started being used in current practices of risk assessment (Cheng et al. 2009; Kentel & Aral 2007; Darbra et al. 2008; Skanata & Byrd 2007). Extending the proposed fuzzy approach to spatial databases requires identifying where the fuzziness happens, how it can be embedded into the database while performing queries and representing the results.

To do so, a fuzzy approach developed in Chapter 4 was embedded into a spatial datacube through redefining the spatial datacube elements as fuzzy dimensions, fuzzy members, fuzzy hierarchies, fuzzy measures, and fuzzy facts by adding membership degrees for these elements. Table 1 presents a summary of fuzzy spatial datacubes' elements.

Table 6.1: Formal definition of fuzzy spatial datacubes' elements

Fuzzy Spatial Datacube:	$FSC = \langle D, F \rangle$ Set of dimension instances: $D = \{d_1, ...d_n\}$ Finite set of facts $F = \langle M, Ms \rangle$
Fuzzy Spatial Fact:	$F = \langle M, Ms \rangle$ Finite set of members of dimensions: $M = \{m_1, ..., m_n\}$ Finite set of fuzzy measure values: $Ms = \{Ms_1, ..., Ms_m\}$
Fuzzy Spatial Dimension:	$D = \langle L, (\leq, \nu) \rangle$ Set of members: $L = \{m_1, m_2, ..., m_n\}$ Relationship between the members: ≤

	Membership degree of a hierarchy level to its successor Instance of a dimension: $d = (L, \leq, \nu)$
Fuzzy Member:	$m = \langle AT_m, V, \mu, f(a_i, v_i) \rangle$ Set of members' attributes : $AT_m = \{a_1 ... a_n\}$ Set of domain values of attributes: $V = \{v_1 ... v_n\}$ Membership degree of *m* in *l* $\mu \in [0,1]$ Function from elements in AT_m to elements in *V:* $f(a_i, v_i)$
Hierarchy Levels	*l={ a1 ...,an}* Set of attributes: *ai*
Level Attribute	*ai= (type, nature, domain)* type={numeric, textual, date, geometric} nature={spatial, temporal, thematic, crisp or fuzzy} domain={value}

A fuzzy spatial datacube is a pair $FSC = \langle D, F \rangle$ of finite set of dimension instances $D = \{d_1, ... d_n\}$ and finite set of facts $F = \langle M, Ms \rangle$ over dimension instances, while there are at least one fuzzy spatial dimension and fuzzy spatial facts. In this case, fuzzy spatial fact is a pair $F = \langle M, Ms \rangle$ of finite set of members of dimensions $M = \{m_1, ..., m_n\}$ and finite set of fuzzy measure values $Ms = \{Ms_1, ..., Ms_m\}$ calculated with respect to members of M and their respective membership degrees $\mu \in [o,1]$. Also, fuzzy spatial dimension is a tuple $D = \langle L, (\leq, \nu) \rangle$ where $L = \{m_1, m_2, ..., m_n\}$ is the set of members, ≤ the relationship between the members for a level in L, and membership degree ν, that represents the degree of belonging of a level to its successor. The fuzziness in a dimension comes from the relation between hierarchy levels or the degree of belonging of a member to a hierarchy level (Laurent 2010). The nature of the attribute in the level attributes indicates whether it is fuzzy or crisp.

As discussed in Chapter 5, aggregation in GeoBI is a summarization process of values or geometries in a datacube that directly depends on the data model used (Péres et al. 2007; Laurent 2010; Gomez et al. 2009; Pedersen et al. 2001). Using fuzzy concepts to define appropriate operators for data aggregation in a datacube has been initiated by Laurent 2010 and Molina et al. (2006). A series of operators such as roll-up, drill-down, slice, dice, and pivot have been defined for fuzzy datacubes in Molina et al. (2006) and Martin-bautista et al. (2013) using both quantitative and qualitative data. The thematic aggregation can principally be performed based on Laurent (2010) and Molina et al. (2006). However, the geometric aggregation involving spatially fuzzy or crisp members required redefining fuzzy operators such as overlay and fusion for fuzzy spatial objects. In this regard, fuzzy overlay and fusion operators were developed in Chapter 5. The result of such operators is a new fuzzy object, in this research project a fuzzy partition or grid, where the membership degree of each cell or part of an object is calculated using arithmetic or logic operators.

Fuzzy overlay operators consist of union, intersection, difference, mean, and weighed mean to regroup information at the same level of hierarchy or scale. Fuzzy fusion was also introduced to deal with regrouping of information from fine level to coarser level of spatial datacube. The syntax of fuzzy overlay and fusion operators is presented in Table 6.2. Keep in mind that, the cell size of multiple fuzzy partitions should be the same (if that is not the case, all grids should be equalized with respect to a reference layer's cell size).

Table 6.2: The formal syntax of fuzzy aggregation operators

Fuzzy Overlay:	*FuzzyOverlay (Input grid, {FuzzyOverlay_Type}, output grid)* {FuzzyOverlay_Type}= {Union, Intersection, Difference, Mean, and Weighted Mean}
Fuzzy Fusion:	*FuzzyFusion (Input grid, cell_factor, {FuzzyOverlay_Type}, output grid)*

Five feasible scenarios were presented in Chapter 5, which are supported by the proposed fuzzy spatial datacube model using fuzzy aggregations operators. These scenarios are as follows:

1) Crisp aggregation onto crisp data;
2) Fuzzy aggregation onto crisp data;
3) Crisp aggregation onto fuzzy data;
4) Fuzzy aggregation onto fuzzy data;
5) Fuzzy aggregation onto mixed crisp and fuzzy data.

The SMCM model, developed in Chapter 3 was used to demonstrate the design of a fuzzy spatial datacube. The proposed fuzzy spatial datacube for CERA is well adapted for both vector and grid-based information. Spatial data is stored in a grid-based structure that is associated with a vector-based data structure e.g. census division in a spatial dimension. Performing the aggregation on this dimension permits navigating from the grid-based to the vector-based structure and vice versa. In addition, (Plante 2013) shows in his master thesis, having a grid for each hierarchy level in a spatial dimension is an efficient way to design raster spatial datacubes with increased on-the-fly spatial analysis capabilities. The implementation of such datacubes is identical to the implementation of typical crisp spatial datacubes. The only difference is in measures calculation in which fuzzy operators instead of a crisp one should be applied to respective dimensions, members, and hierarchy relations.

6.2 Research Perspectives

The fundamental concepts for developing a fuzzy spatial datacube are provided in this thesis. In this context, several new research perspectives to enhance decision-making process within GeoBI systems can be identified. It is worth remembering that this Ph.D. thesis was a multidisciplinary research project. In this case, the obtained results drive us to different themes of research as follows:

- **Enhancement in developing fuzzy spatial datacube**

  The proposed approaches to design fuzzy spatial datacubes in this research project do not support the capability of reporting the end-results using linguistic expressions (very high, high…low) or representing the level of uncertainty and vagueness separately.

- **Develop PictograF for fuzzy object**

  On a conceptual level of multidimensional database design, using pictograms would be very useful to show schematically the fuzzy elements of models and for communication purposes. Extending a visual language such as PictograF (http://pictograf.scg.ulaval.ca/) seems promising. Adding symbols for fuzzy objects (fuzzy point, fuzzy line, fuzzy polygon) and fuzzy grids can be a new insight for future work. The advantage of such symbols is that fuzzy and non-fuzzy elements in a spatial datacube can be easily analyzed and understood.

- **Modeling complex fuzzy spatial objects in spatial datacubes**

  This thesis focused only on grid-based fuzzy representation. In practice, this is a simplified type of fuzzy model. If complex objects (vector format) with broad boundaries, holes and multiple certain parts are the case, the research is open to finding fuzzy aggregation operators and topological relationships and neighborhood concepts in the vector world.

- **Evolution of Fuzziness in Fuzzy Objects**

  Fuzzy objects change within time. This change can happen in geometry, attributes, or both. Moreover, the fuzziness (membership degree) associated with these two aspects also changes. By comparing the spatial extents of an object in two successive periods, we can derive the change of shape, size, and the degree of membership. It is important to study these influences in order to provide accurate information to decision-makers. Efficient handling of the dynamic fuzzy objects in spatial datacubes is a topic for future research. This provides the advantage of calculating difference between two successive levels of hierarchy of given fuzzy dimensions and representing it geometrically.

- **Fuzzy TIN and Voronoi diagram**

As stated, a regular tessellation (square) was performed in this research project. Testing the proposed approach within an irregular tessellation such as TIN and Voronoi diagrams can be another research perspective to achieve modeling and representation of the a dynamic phenomenon. To do so, a self-standing code should be developed to store and manage a dynamic data structure. An example of such an approach can be performed in surveillance of infrastructure or traffic control system from a sensor network. Extending this idea within GeoBI systems allows the richness of managing large amounts of data, enriching with simulation methods and handling inherent spatial uncertainty through a fuzzy spatial datacube.

- **Testing the fuzzy based approach in the other domain of applications**

As mentioned previously, there are many possibilities for the fuzzy based approach. This thesis was focused on CERA. In each Chapter (3, 4, and 5), a generic framework is provided, so adapting the proposed methods should be easily done for other CERA models. Examples of such application in other domains can also be foreseen, such as in health, forestry, mining, etc.

- **Adapting Fuzzy Spatial Datacube with moving objects: a simulation situation**

Manipulating moving objects in spatial datacubes is an interesting topic when the fuzzy concepts are included in the picture. The notion of semantic trajectories (i.e., trajectories expressed in terms of places of interest instead of x, y coordinates thorough fuzzy IF-THEN rules) allows the inferring of interesting patterns of movement. An example of such application is an urgency system or traffic control.

- **A semantic analysis in elaboration of vulnerability index**

As stated previously, the vulnerability index is elaborated with respect to the interests of stakeholders and experts. Integrating multiple vulnerability indices brings the issue of semantics when defining indicator and related parameters. A semantics analysis to elaborate such index is missing. Also, there is the issue of the supplementary information (metadata) of involved indicators. This complicates the situation identifying risk components properly, translating them as datacube's elements, and handling the inherent spatial uncertainty.

- **Testing proposed approaches on a large territory and validating the results with experts and decision-makers**

  In addition, testing the proposed approaches on very large territory can be another future work. This helps to identify the possible problem caused by large amounts of real data coming from semantics issues, as well as dealing with large process of aggregation (e.g. level of province to country, continent, and world). Then, validating the results based on proposed fuzzy solutions by stakeholders, experts and decision-makers is recommended.

In conclusion, based on the achievement of this research project, we found that fuzzy spatial datacubes are essential to perform more coherent knowledge discovery for effective decision-making. These datacubes enable the extraction of the relevant knowledge in a more natural way and provide the results to queries with a certain precision about the reliability of the knowledge. The decision model can be performed based on fuzzy spatial datacubes to deal with inherent spatial uncertainty. Fuzzy set theory and the SOLAP system, together, render the discovered knowledge more easily understood, since fuzzy set theory treats numerical values in a more natural way.

Reference

Abelló, A., Saltor, F. & Samos, J., 2006. YAM2: A Multidimensional Conceptual Model Extending UML. *Information Systems*, 31(6), pp.541–567.

Abuodha, P.A. & Woodroffe, C.D., 2006. *International Assessments of the Vulnerability of the Coastal Zone to Climate Change, Including an Australian Perspective* D. of the E. and H. Australian Greenhouse Office, ed., Australian Greenhouse Office.

Adger, W., 2006. Vulnerability. *Global Environmental Change*, 16(3), pp.268–281.

Aerts, J., Goodchild, M.F. & Heuvelink, G., 2003. Accounting for Spatial Uncertainty in Optimization with Spatial Decision Support Systems. *Transactions in GIS*, 7(2), pp.211–230.

Alexander, D., 2000. *Confronting Catastrophe: New Perspectives on Natural Disasters*, Terra Publishing.

Ali, A. & Rakus-Andersson, E., 2009. Fuzzy Decision Making in Business Intelligence Application of fuzzy models in retrieval of optimal decision. *School of Engineering, Mathematics Dept*, Msc, p.31.

Altman, D., 1994. Fuzzy set theoretic approaches for handling imprecision in spatial analysis. *International Journal of Geographical Information Science*, 8(3), pp.271–289.

Bakillah, M., 2012. *Real Time Semantic Interoperability in Adhoc Networks of Geospatial Databases : Disaster managment*. Université Laval, Qubec City, Canada.

Bédard, Y. et al., 1997. *Étude de l'état actuel et des besoins de R&D relativement aux architectures et technologies des data warehouses appliquées aux données spatialesResearch*, Quebec City: Laval university.

Bédard, Y., 1988. Uncertainties in Land Infromarion Systems Databases. In N. Chrisman, ed. *Auto-Carto*. Baltimore, United States, pp. 175–184.

Bédard, Y., Merrett, T. & Han, J., 2009. Fundamentals of spatial data warehousing for geographic knowledge discovery. In H. J. Miller J.H., ed. *Geographic Data Mining and Knowledge Discovery*. Taylor & Francis, pp. 53–73.

Bédard, Y., Rivest, S. & Proulx, M.-J., 2007. Spatial Online Analytical Processing (SOLAP): Concepts, Architectures, and Solutions from a Geomatics Engineering Perspective. In R. Wrembel & C. Koncilia, eds. *Data Warehouses and OLAP: Concepts, Architectures, and Solutions*. London, UK: IRM Press, EDA Group Inc., pp. 298–319.

Bejaoui, L. et al., 2008. Logical consistency for vague spatiotemporal objects and relations. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 34.

Bejaoui, L., 2009. *Qualitative topological relationships for objects with possibly vague shapes: imolication on the specification of topological integrity constraint in transactional spatial databases and in spatial data warehouses*. Quebec City: Université Laval.

Bernatchez, P., Fraser, C., Friesnger, S., et al., 2008. *Sensibilité des côtes et vulnérabilité des communautés du golfe du Saint-Laurent aux impacts des changements climatiques*,

Bernatchez, P., Fraser, C. & Lefaivre, D., 2008. Effets des structures rigides de protection sur la dynamique des risques naturels côtiers: érosion et submersion. In J. Locat et al., eds. *4th Canadian Conference on Geohazard: from Causes to Managment*. Quebec City: Universite Laval, pp. 594–604.

Bezdek, J., Ehrlich, R. & Full, W., 1984. FCM: The fuzzy c-means clustering algorithm. *Computers & Geosciences*, 10(2), pp.191–203.

Bimonte, S. et al., 2012. Definition and analysis of new agricultural farm energetic indicators using spatial OLAP. In B. Murgante et al., eds. *Computational Science and Its Applications – ICCSA 2012*. Salvador de Bahia, Brazil: Springer-Verlag Berlin Heidelberg, pp. 373–385.

Blaikie, P. et al., 2004. *At Risk Natural Hazards, People's Vulnerability and Disaster* second., New York: Routledge.

Boruff, B., Cutter, S. & Emrich, C., 2005. Erosion Hazard Vulnerability of US Coastal Counties. *Journal of Coastal Research*, 21(5), pp.932–942.

Brooks, N., 2003. *A conceptual framework Vulnerability , risk and adaptation : A conceptual framework*,

Brown, D., 1998. Classification and boundary vaguen ess in mapping presettlement forest types. *International Journal of Geographical Information Science*, 12(2), pp.105–129.

Bruce, E., 2004. Spatial Uncertainty in Marine and Coastal GIS. In J. Smith & D. Bartlett, eds. *GIS fo Coastal Zone Managment*. CRC Press, p. 12.

Burrough, P., 1989. Fuzzy mathematical methods for soil survey and land evaluation. *Journal of Soil Science*, 40(3), pp.477–492.

Burrough, P. & Frank, A.U., 1996. *Geographic Objects with Indeterminate Boundaries (GISDATA)* first., CRC Press.

Caron, P., 1998. *Étude du potentiel OLAP pour supporter l'analyse spatio-temporelle*. Quebec City: Laval.

Chaudhuri, S., Dayal, U. & Narasayya, V., 2011. An Overview of Business Intelligence Technology A. D'Atri et al., eds. *Communications of the ACM*, 54(8), pp.88–98.

Chen, K., 2009. Quantifying environmental attributes from Earth Observation data products by spatial upscaling: Three case studies. In *2nd International Conference on Earth Observation for Global Changes (CD-ROM),*. Chengdu, Sichuan Province, China. 1, pp. 1600–1610.

Cheng, T., 2002. Fuzzy Objects : Their Changes and Uncertainties. *Photogrammetric Engineering & Remote Sensing*, 68(1), pp.41–49.

Cheng, T., Fisher, P. & Li, Z., 2005. Double Vagueness: Effect of Scale on the Modelling of Fuzzy Spatial Objects. In F. P., ed. *11th International Symposium on Spatial Data Handling,*. Springer, pp. 299–313.

Cheng, T., Molenaar, M. & Lin, H., 2001. Formalizing Fuzzy Objects from Uncertain Classification Results. *International Journal of Geographical Information Science*, 15(1), pp.27–42.

Cheng, T., Molenaar, M. & Stein, A., 2009. Fuzzy Approach for Integrated Coastal Zone Management. In Y. X., ed. *Remote Sensing and Geospatial Technologies for Coastal Ecosystem Assessment and Management*. Springer Berlin Heidelberg, pp. 67–90.

Chi, Z., Wu, J. & Yan, H., 1995. Handwritten numeral recognition using self-organizing maps and fuzzy rules. *Pattern Recognition*, 28(1), pp.59–66.

Choa, H.-N., Choi, H.-H. & Kim, Y.-B., 2003. A risk assessment methodology for incorporating uncertainties using fuzzy concepts. *Reliability Engineering and System Safety*, 78, pp.173–183.

Chowdhury, S., Champagne, P. & McLellan, P.J., 2009. Uncertainty characterization approaches for risk assessment of DBPs in drinking water: a review. *Journal of environmental management*, 90(5), pp.1680–91.

Clementini, E. & DiFelice, P., 1997. Approximate topological relations. *International Journal of Approximate Reasoning*, 16, pp.173–204.

Cohn, A. et al., 1997. Qualitative Spatial Representation and Reasoning with the Region Connection Calculus. *GeoInformatica*, 1, pp.275–316.

Cohn, A. & Gotts, N., 1996. The “Egg-Yolk” Representation of Regions with Indeterminate Boundaries. In P. Burrough & A. Frank, eds. *Geographic Objects with Indeterminate Boundaries*. Taylor & Francis, pp. 171–189.

Cohn, A. & Hazarika, S., 2001. Qualitative Spatial Representation and Reasoning: An Overview. *Fundam. Inf.*, 46(1-2), pp.1–29.

Cowell, P. & Zeng, T., 2003. Integrating uncertainty theories with GIS for modeling coastal hazards of climate change. *Marine Geodesy*, 26(1-2), pp.5–18.

Cutter, S. et al., 2008. A place-based model for understanding community resilience to natural disasters. *Global Environmental Change*, 18(4), pp.598–606.

Cutter, S., 2002. Historic assessment of the socio-economic vulnerability of United States coastal counties. *Vulnerability Assessment Techniques (VAT) III*

Cutter, S., Boruff, B. & Shirley, W., 2003. Indicators of social vulnerability to environmental hazards. *Social Science Quarterly*, 84(1), pp.242–261.

Cutter, S., Mitchell, J. & Scott, M., 2000. Revealing the Vulnerability of People and Place : a Case Study of Georgetown County, South Carolina. *Annual of the Association of American Geograpers*, 90(4), pp.713–737.

Darbra, R., Eljarrat, E. & Barceló, D., 2008. How to measure uncertainties in environmental risk assessment. *TrAC Trends in Analytical Chemistry*, 27(4), pp.377–385.

Daudé, E. et al., 2009. Spatial risks and complex systems : methodological perspectives. In Aziz-Alaoui M. A. and Bertelle C., ed. *From System Complexity to Emergent Properties Understanding Complex Systems*. Springer.

Delgado, M., Molina, C. & Sánchez, D., 2004. A fuzzy multidimensional model for supporting imprecision in OLAP. In *Fuzzy Systems*. Budapest, Hungary, pp. 1331–1336.

Desprats, J.-F. et al., 2010. A "coastal-hazard GIS" for Sri Lanka. *Journal of coastal Conservation*, 14(1), pp.21–31.

Devillers, R. et al., 2010. Thirty Years of Research on Spatial Data Quality: Achievements, Failures, and Opportunities. *Transactions in GIS*, 14(4), pp.387–400.

Dilo, A., *Representation of and reasoning with vagueness in spatial information*,

Dilo, A., 2006. *Representation of and reasoning with vagueness in spatial information: A system for handling vague objects*. Enschede: ITC.

Dilo, A., By, R.A. de & Stein, A., 2007. A system of types and operators for handling vague spatial objects. *International Journal of Geographical Information Science*, 21(14), pp.397–426.

Dilo, A., Kraipeerapun, P. & By, R.A. de, 2006. Storage and Manipulation of Vague Spatial Objects Using Existing GIS Functionality. *Studfuzz*, 203, pp.293–321.

Dolan, A. & Walker, I., 2006. Understanding vulnerability of coastal communities to climate change related risks. *Journal of Coastal Research*, 2004(39).

Dragicevic, S. & Marceau, D.J., 2001. Space, time, and dynamics modeling in historical GIS databases: a fuzzy logic approach. *Environment and Planning B: Planning and Design*, 28(545-562).

Dziubinski, M., Fratczak, M. & Markowski, A., 2006. Aspects of risk analysis associated with major failures of fuel pipelines. *Journal of Loss Prevention in the Process Industries*, 19, pp.399–408.

Edoh-alove, E. et al., 2013. Exploiting Spatial Vagueness in Spatial OLAP : Towards a New Hybrid Risk-Aware Design Approach. In *AGILE*. Leuven, pp. 1–4.

English, J., George, R. & Yazici, A., 2004. A Fuzzy Spatiotemporal query model for facilating decision support in operational planning. In *In current issue in data and knowledge Engineering*. Varsovia, pp. 225–232.

Erwig, M. & Schneider, M., 1997. Vague regions. In M. Scholl & A. Voisard, eds. *Advances in Spatial Databases*. Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp. 298–320.

Feng, L. & Dillon, T., 1999. Enhancing data warehousing with fuzzy technology. In T. Bench-Capon, G. Soda, & A. Tjoa, eds. *Database and Expert Systems Applications*. Springer-Verlag Berlin Heidelberg 1999, pp. 872–881.

Fenton, N. & Wang, W., 2006. Risk and confidence analysis for fuzzy multicriteria decision making. *Knowledge-Based Systems*, 19(6), pp.430–437.

Fisher, P., 1993. Algorithm and implementation uncertainty in viewshed analysis. *International Journal of Geographical Information Science*, 7(4), pp.331–3347.

Fisher, P., 2008. Uncertainty, Semantic. *Encyclopedia of GIS*, pp.1194–1196.

Fisher, P., Cheng, T. & Wood, J., 2005. Fuzziness and Ambiguity in Multi-Scale Analysis of Landscape Morphometry. In R. V. B. Petry F. abd Cobb M.A., ed. *Fuzzy Modeling with Spatial Information for Geographic Problems*. Springer Berlin Heidelberg, pp. 209–232.

Fisher, P., Cheng, T. & Wood, J., 2007. Higher Order Vagueness in Geographical Information: Empirical Geographical Population of Type n Fuzzy Sets. *GeoInformatica*, 11(3), pp.311–330. Available at: http://dx.doi.org/10.1007/s10707-006-0009-5.

Fisher, P., Comber, A. & Wadsworth, R., 2010. Approaches to uncertainty in spatial data. In R. Devillers & R. Jeansoulin, eds. *Fundamentals of Spatial Data Quality*. ISTE Ltd, pp. 43–59.

Foody, G., 1999. The continuum of classification fuzziness in thematic mapping. *Photogrammetric Engineering & Remote Sensing*, 65(4), pp.443–451.

Frank, A.U., 1987. *Overlay Processing in Spatial Information System*. University of Maine.

Franklin, N., Tversky, B. & Coon, V., 1992. Switching points of view in spatial mental models. *Memory & cognition*, 20(5), pp.507–18.

Frontiers, R., 2005. Decision support tools for managing rising disaster risk : A survey.

Fuchs, S., Kuhlicke, C. & Meyer, V., 2011. Editorial for special issue: Vulnerability to natural hazards—the challenge of integration. *Natural Hazards*, 58(2), pp.609–619.

Füssel, H.-M., 2010a. How inequitable is the global distribution of responsibility, capability, and vulnerability to climate change: A comprehensive indicator-based assessment World Bank, ed. *Global Environmental Change*, 34(4), pp.597–611.

Füssel, H.-M., 2010b. Review and quantitative analysis of indices of climate change exposure, adaptive capacity, sensitivity, and impacts World Bank, ed. *World Development Report*, 34(4), pp.597–611.

Füssel, H.-M. & Klein, R.J., 2006. Climate Change Vulnerability Assessments: An Evolution of Conceptual Thinking. *Climatic Change*, 75(3), pp.301–329. Available at:

Gale, S., 1972. Inexactness, Fuzzy Sets, and the Foundations. *Geographical Analysis*, 4, pp.337–349.

Genz, A. et al., 2007. The Predictive Accuracy of Shoreline Change Rate Methods and Alongshore Beach Variation on Maui, Hawaii. *Journal of Coastal Research*, 23(1), pp.87–105.

Gervais, M. et al., 2009. Data quality issues and geographic knowledge discovery. In H. J. Miller & J. Han, eds. *Geographic Data Minning and Knowldege Discovery*. CRC Press, pp. 99–116.

Gomez, L. et al., 2009. Spatial Aggregation: Data Model and Implementation. *information systems*, 34(6), pp.551–576.

González, C., Tineo, L. & Urrutia, A., 2009. Fuzzy OLAP : A Formal Definition. In W. Yu & E. N. Sanchez, eds. *Advances in Computational Intelligence*. Springer-Verlag Berlin Heidelberg, pp. 189–198.

Goodchild, M.F. & Glennon, A., 2008. Representation and Computation of Geographic Dynamics. In K. Hornsby & M. Yuan, eds. *Understanding Dynamics of Geographic Domains*. Taylor & Francis Group, LLC, pp. 13–30.

Gornitz, V.M., Beaty, T.W. & Daniels, R.C., 1997. *A Coastal Hazards Database for the US West Coast* T. O. R. N. L. Oak Ridge, ed.,

Gray, J. et al., 1997. Data Cube : A Relational Aggregation Operator U. Fayyad, H. Mannila, & G. Piatetsky-Shapiro, eds. *Data Mining and Knowledge Discovery*, 53(1), pp.29–53.

Han, J., Stefanovic, N. & Koperski, K., 1998. Selective materialization: an efficient method for spatial data cube construction. In X. Wu, R. Kotagiri, & K. Korb, eds. *Second Pacific-Asia Conference on Knowledge Discovery and Data Mining,*. Lecture Notes in Computer Science. Melbourne, Australia: Lecture Notes in Computer Science, pp. 144–158.

Hanson, S. et al., 2007. *Capturing coastal morphological change within regional integrated assessment : an outcome-driven fuzzy logic approach*,

Haunert, J.-H., Dilo, A. & van Oosterom, P., 2009. Constrained set-up of the tGAP structure for progressive vector data transfer. *Computers & Geosciences*, 35(11), pp.2191–2203.

Hegde, A. & Reju, V., 2007. Development of Coastal Vulnerability Index for Mangalore Coast, India. *Journal of Coastal Research*, 23(5), pp.1106–1111.

Hessami, A., 2004. A Systems Framework for Safety and Security: The Holistic Paradigm. *Systems Engineering*, 7(2), pp.99–112.

Hessami, A. & Karcanias, N., 2009. Complexity , Emergence and the Challenges of Assurance The need for a Systems Paradigm II . Complexity and Emergent. In *IEEE Systems Conference*. Vancouver, Canada: IEEE.

Heuvelink, G., Brown, J. & VanLoon, E., 2007. A probabilistic framework for representing and simulating uncertain environmental variables. *International Journal of Geographical Information Science*, 21(5), pp.459–513.

Hill, M. et al., 2005. Multi-criteria decision analysis in spatial decision support: the ASSESS analytic hierarchy process and the role of quantitative methods and spatially explicit analysis. *Environmental Modelling & Software*, 20, pp.955–976.

Hinkel, J., 2005. Advances in Geosciences DIVA : an iterative method for building modular integrated models. *Advances In Geosciences*, 4, pp.45–50. Available at: http://www.adv-geosci.net/4/45/2005/adgeo-4-45-2005.pdf.

Hinkel, J. & Klein, R.J., 2007. Integrating knowledge for assessing coastal vulnerability to climate change. In N. R. J. and P.-R. E. C. McFadden L, ed. *Managing Coastal Vulnerability: An Integrated Approach. Amsterdam*. Elsevier Science, pp. 61–78.

Imhoff, C., Gallemo, N. & Geiger, J., 2003. *Mastering Data Warehouse Design: Relational and Dimensional Techniques* J. Wiley, ed.,

Inmon, W., 1992. *Building the Data Warehouse*, New York: John Wiley & Sons, Ltd.

IPCC, 2007. *IPCC Fourth Assessment Report: Climate Change 2007* W. G. I. R. T. P. S. Basis, ed.,

Iris, J., 2009. *Contribution de la méthodologie et de la technologie géodécisionnelle pour l'aide à l'évaluation des risques naturels dans le secteur de l'assurance en France*. Paris: l'Ecole des Mines de Paris.

ISDR, 2004. *Living with Risk*, Geneva, Switzerland.

ISDR & Nations, U., 2004. *Living with Risk*, Geneva, Switzerland.

ISO/TC211, 2003. *Geographie Information - Spatial Referencing by Geographic Identifiers* R. 19112, ed.,

Jadidi, A. et al., 2012. Toward an Integrated Spatial Decision Support System to Improve Coastal Erosion Risk Assessment: Modeling and Representation of Risk Zones. In *FIG working week 2012*. Rome, Italy: FIG, p. 10.

Jadidi, A. et al., 2013. Using Geospatial Business Intelligence Paradigm to Design a Multidimensional Conceptual Model for Efficient Coastal Erosion Risk Assessment. *Journal of Coastal Conservation*, 17(3), pp.527–543.

Jeansoulin, R. et al., 2010. *Methods for handling imperfect spatial information*, Springer Berlin / Heidelberg.

Kanjilal, V., Liu, H. & Schneider, M., 2010. Plateau Regions : An Implementation Concept for Fuzzy Regions in Spatial Databases and GIS. In *Proceedings of the 13th International Conference on Information Processing and Management of Uncertainty in KnowledgeBased Systems*. pp. 624–633.

Karvetski, C.W. et al., 2011. Integration of Decision Analysis and Scenario Planning for Coastal Engineering and Climate Change. *IEEE Transactions on Systems Man and Cybernetics Part A Systems and Humans*, 41(1), pp.63–73.

Kaya, M. & Alhajj, R., 2006. Extending OLAP with fuzziness for effective mining of fuzzy multidimensional weighted association rules. In X. Li, O. R. Zaiane, & Z. Li, eds. *Advanced Data Mining and Applications*. Springer-Verlag Berlin Heidelberg, pp. 64–71.

Kentel, E. & Aral, M.M., 2007. Risk tolerance measure for decision-making in fuzzy analysis: a health risk assessment perspective. *Stochastic Environmental Research and Risk Assessment*, 21(4), pp.405–417.

Kimball, R. & Ross, M., 2002. *The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling* 2nd ed., Wiley.

Klein, R.J. & Nicholls, R.J., 1999. Assessment of coastal vulnerability to sea-level rise. *Ambio*, 28(2), pp.182–187.

Koukoulas, S. et al., 2005. A GIS Tool For Analyis and Interpretation of Coastal Erosion Model Outputs (SCAPEGIS). In *Coastal Dynamics*. American Society of Civil Engineers, p. 15.

Van Kouwen, F. et al., 2007. Applicability of Decision Support Systems for Integrated Coastal Zone Management. *Coastal Management*, 36(1), pp.19–34.

Larrivée, S., 2011. Comment choisir les meilleures données géospatiales? , 2011. Available at: http://www.intelli3.com/blog/?p=165.

Laurent, A., 2010. Fuzzy Multidimensional Databases. In J. Kacprzyk, ed. *Uncertainty Approaches for Spatial Data Modeling and Process*. Springer-Verlag Berlin Heidelberg, pp. 43–59.

Leatherman, S.P. & Clow, J.B., 1983. UMD shoreline mapping project. *IEE Geoscience and Remote Sensing Society Newsletter*, 22, pp.5–8.

Levesque, M.-A. et al., 2007. Towards Managing the Risks of Data Misues for Spatial Datacube. In *5th International Symposium on Spatial Data Quality*. Enschede, Netherlands.

Li, X. et al., 2012. A Decision Support Framework for the Risk Assessment of Coastal Erosion in the Yangtze Delta. In A. G. O. Yeh et al., eds. *Advances in Spatial Data Handling and GIS*. Springer Berlin Heidelberg, pp. 213–226.

Limber, P.W. et al., 2007. Using topographic lidar data to delineate the North Carolina shoreline. In *Coastal Sediments '07*. American Society of Civil Engineers, p. 14.

Linham, M. & Nicholls, R.J., 2012. Adaptation technologies for coastal erosion and flooding: a review. *Proceedings of the ICE-Maritime Engineering*, 165(3), pp.95–112.

Longley, P. et al., 2005. *Geographical Information System and Science* 2end ed., John Wiley & Sons, Ltd.

MacEachren, A.M. et al., 2005. Visualizing Geospatial Information Uncertainty: What We Know and What We Need to Know. *Cartography and Geographic Information Science*, 32(3), pp.139–160.

Mai, S. & Liebermann, N., 2002. A decision support system for an optimal design of sea dikes with respect to risk. In S. V. L. V. Babovic and K. Phoon., ed. *5th International Conference on Hydroinformatics*. Cardiff, United Kingdom.

Malinowski, E. & Zimànyi, E., 2008. *Advanced Data Warehouse Design: From Conventional to Spatial and Temporal Applications*, Springer.

Manche, Y., 2000. *Analyse spatiale et mise en place de système d'information pour l'évaluation de la vulnérabilité des territoires de montagnes face aux risques naturels*. Joseph Fourier- Grenoble I.

Mannan, B., Roy, J. & Ray, A., 1998. Fuzzy ARTM AP supervised classification of multi-spectral remotely-sensed images. *International Journal of Remote Sensing*, 19(4), pp.767–775.

Martin-bautista, M.J. et al., 2013. A new multidimensional model with text dimensions : definition and. *International Journal of Computational Intelligence Systems*, 6(1), pp.137–155.

McFadden, L. et al., 2007. A Methodology for Modeling Coastal Space for Global Assessment. *Journal of Coastal Research*, 23(4), pp.911–920.

McHugh, R.-M. et al., 2006. Analyse du potentiel d'une application SOLAP pour une gestion efficace de l'érosion des berges en Gaspésie Iles-de-la-Madeleine. *Géomatique 2006*.

Molenaar, M., 2000. Tree Conceptual Uncertainty Levels for Spatial Objects. *International Archives of Photogrammetry and Remote Sensing*, XXXIII, pp.670–677.

Molenaar, M. & Cheng, T., 2000. Fuzzy spatial objects and their dynamics. *ISPRS Journal of Photogrammetry & Remote Sensing*, 55, pp.164–175.

Molina, C. et al., 2006. A New Fuzzy Multidimensional Model. *IEEE Transactions on Fuzzy Systems*, 14(6), pp.897–912.

Morang, A. & Szuwalski, A., 2003. *Coastal Engineering Manual-Glossory*, Washington, D.C.: U.S. Army Corps of Engineers.

Morris, A. & Jankowski, P., 2005. Spatial Decision Making Using Fuzzy GIS. In R. V. B. Petry F. abd Cobb M.A., ed. *Fuzzy Modeling with Spatial Information for Geographic Problems*. Springer Berlin Heidelberg, pp. 275–298.

Mostafavi, M.A., Edwards, G. & Jeansoulin, R., 2004. An ontology-based method for quality assessment of spatial data bases. In *ISSDQ'04, GeoInfo Series*. Austria.

Muhlbauer, W., 1996. *Pipeline Risk Management Manual*, Gulf Publishing Company, Houston, Texas.

Nakalevu, T., 2006. *CV&A : a guide to community vulnerability and adaptation assessment and action* S. : S. Apia, ed., Apia, Samoa : SPREP.

Nauck, D. & Kruse, R., 1997. A neuro-fuzzy method to learn fuzzy classification rules from data. *Fuzzy Sets and Systems*, 89(3), pp.277–288.

Nicholls, R.J. et al., 2008. Climate change and coastal vulnerability assessment: scenarios for integrated assessment. *Sustainability Science*, 3(1), pp.89–102. Available at: http://dx.doi.org/10.1007/s11625-008-0050-4.

NOAA (National Oceanic & Atmospheric Administration), 2003. *Science of Tsunami Hazards 21(1).*,

Pauly, A. & Schneider, M., 2010. VASA: An algebra for vague spatial data in databases. *Information Systems*, 35(1), pp.111–138.

Pedersen, T. et al., 2001. Pre-aggregation in Spatial Data Warehouses Advances in Spatial and Temporal Databases. In Springer Berlin / Heidelberg, pp. 460–478.

Pedersen, T. & Jensen, C., 2001. Multidimensional database technology. *IEEE Computer Society*, 34(12), pp.40–46.

Pedersen, T., Jensen, C. & Dyreson, C., 1999. Supporting imprecision in multidimensional databases using granularities. In *Proceedings. Eleventh International Conference on Scientific and Statistical Database Management*. IEEE Comput. Soc, pp. 90–101.

Péres, D. et al., 2007. Fuzzy Spatial Data Warehouse: A Multidimensional Model. In *Current Trends in Computer Science,ENC 2007. Eighth Mexican International Conference*. IEEE, pp. 1–3.

Pfoser, D. & Tryfona, N., 2001. Capturing Fuzziness and Uncertainty of Spatiotemporal Objects. In and E. J. Caplinkskas A., ed. *Advances in Databases and Information Systems 5th East European Conference, ADBIS 2001 Vilnius, Lithuania, September 25–28, 2001 Proceedings*. Springer Berlin / Heidelberg.

Plante, M., 2013. *Intégration du matriciel dans les cubes matriciels*. Université Laval.

Randell, D., Cui, Z. & Cohn, A., 1992. A spatial logic based on regions and connection. In B. Nebel, C. Rich, & W. Swartout, eds. *international conference on knowledege Representation and Reasoning (KR92)*. San Mateo, CA: Morgan Kaufmann, pp. 165–176.

Rivest, S. et al., 2005. SOLAP: Merging Business Intelligence with Geospatial Technology for Interactive Spatio-Temporal Exploration and Analysis of Data. *Journal of ISPRS: Advances in spatio-temporal analysis and representation*, 60(1), pp.17–33.

Rivest, S., Bédard, Y. & Marchand, P., 2001. Toward Better Support For Spatial Decision Making: Defining the Characteristics of Spatial On-Line Analytical Processing (SOLAP). *Geomatica*, 55(4), pp.539–555.

Robinson, V.B., 2003. A Perspective on the Fundamentals of Fuzzy Sets and their Use in Geographic Information Systems. *Transactions in GIS*, 7(1), pp.3–30.

Roca, E., Gamboa, G. & Tàbara, J.D., 2008. Assessing the multidimensionality of coastal erosion risks: public participation and multicriteria analysis in a Mediterranean coastal system. *Risk analysis : an official publication of the Society for Risk Analysis*, 28(2), pp.399–412.

Roy, P. & Mandal, J., 2011. A Novel Fuzzy-GIS Model based on Delaunay Triangulation to Forecast Facility Locations(FGISFFL). *International Symposium on Electronic System Design*.

Salehi, M., 2009. DEVELOPING A MODEL AND A LANGUAGE TO IDENTIFY AND SPECIFY THE INTEGRITY.

Salehi, M., Rivest, S. & Bédard, Y., 2010. A Formal Conceptual Model and Definition Framework for Spatial Datacubes. *Geomatica*, 64(3), pp.313–326.

Santini, M. et al., 2010. A multi-component GIS framework for desertification risk assessment by an integrated index. *Applied Geography*, 30(3), pp.394–415.

Sboui, T., 2010. *A Conceptual Framework And A Risk Management Approach for the Interoperability between Geospatial Datacubes*. Laval University.

Schneider, M., 2003a. Design and implementation of finite resolution crisp and fuzzy spatial objects. *Data & Knowledge Engineering*, 44(1), pp.81–108.

Schneider, M., 2010. Soft Computing Techniques in Spatial Databases. *Soft Computing Applications for Database : Techniques and Issues*, p.49.

Schneider, M., 1999. Uncertainty Management for Spatial Datain Databases: Fuzzy Spatial Data Types Advances in Spatial Databases. In R. Güting, D. Papadias, & F. Lochovsky, eds. *Advances in Spatial Databases, 6th International Symposium, SSD'99*. Hong Kong, China: Springer Berlin / Heidelberg, pp. 330–351.

Schneider, M., 2003b. Vague Spatial Data Types R. Güting, D. Papadias, & F. Lochovsky, eds. *Advances in Spatial Databases, Lecture Notes in Computer Science*, 44(1), pp.81–108.

Schneider, M., Vossen, G. & Zimànyi, E., 2011. Data Warehousing: From Occaional OLAP to real-time Business Intelligence. *Dagstuhl Reports*, 1(9), pp.1–25.

Servigne, S., Lesage, N. & Libourel, T., 2010. Quality components, standards, and metadata. In R. Devillers & R. Jeansoulin, eds. *Fundamentals of Spatial Data Qualityof spatial data quality*. ISTE Ltd, pp. 179–210.

Sharples, C., 2006. *Indicative mapping of Tasmanian coastal vulnerability to climate change and sea level rise: explanatory report (second edition)* W. and E. Tasmanian Department of Primary Industries, ed.,

Shaw, J. et al., 1998. Potential Impacts of Global Sea-Level Rise on Canadian Coasts. *The canadian Geographer*, 42(4), pp.365–379.

Shekhar, S. et al., 2001. Map Cube: A Visualization Tool for Spatial Data Warehouse. 2000. In H. Miller & J. Han, eds. *Geographic Data Mining and Knowledge Discovery*. London, UK: Taylor & Francis.

Siqueira, T. & Ciferri, C. de A., 2012. Towards Vague Geographic Data Warehouses. In N. Xiao, ed. *GIScience 2012, Lecture Note in Computer Science 7478*. Springer, pp. 173–186.

Skanata, D. & Byrd, D., 2007. Fuzzy Modelling for Uncertainty Propagation and Risk Quantification in Environmental Water Systems. *Computational Models of Risks to …*, 13, pp.260–278.

Smith, B. & Mark, D.D.M., 2003. Do mountains exist? Towards an ontology of landforms. *Environment and Planning B: Planning and Design*, 30(3), pp.411–427.

Smith, B. & Varzi, A.C., 2000. Fiat and Bona Fide Boundaries. *Philosophy and Phenomenological Research*, 60(2), p.401.

Stanton, E., Davis, M. & Fencl, A., 2010. *Costing Climate Impacts and Adaptation: A Canadian Study on Coastal Zones*, Somerville, USA.

StatisticCanada, 2012. 2011 Census. Statistics Canada Catalogue no. 98-316-XWE, Census profile: Percé, Quebec (Code 2402005) and Quebec (Code 24). *Statistic Canada*.

Stefanovic, N., 1997. *Design and Implementation of On-Line Analytical Processing (OLAP) of Spatial Data*. Vancouver: Simon Fraser University.

Szemesova, J. & Gera, M., 2010. Uncertainty analysis for estimation of landfill emissions and data sensitivity for the input variation. *Climatic Change*, 103(1-2), pp.37–54.

Tang, X., 2004. *Spatial Object Modeling in Fuzzy Topological Spaces with Applications to Land Cover Change*. ITC, The Netherlands.

Thieler, E. et al., 2009. *Digital Shoreline Analysis System (DSAS) version 4.0—An ArcGIS extension for calculating shoreline change* U. S. G. S. O.-F. R. 2008-1278, ed., U.S. Geological Survey.

Thomasoni, A., 2010. *Modèles et méthodes d'évaluation et de gestion des risques appliqués aux systèmes de transport de marchandises dangereuses(TMD) reposant sur les nouvelles technologies de l'information et de la communication (NTIC)*. l'École nationale supérieure des mines de Paris.

Torlone, R., 2003. Conceptual Multidimensianl Models. In M. Rafanelli, ed. *Multidimensional datacubes: Problems and Solutions*. Idea Group Inc., pp. 69–90.

Totschnig, R., Sedlacek, W. & Fuchs, S., 2010. A quantitative vulnerability function for fluvial sediment transport. *Natural Hazards*, 58(2), pp.681–703.

Uitermark, H.T. et al., 2005. Ontology-based integration of topographic data sets. *International Journal of Applied Earth Observation and Geoinformation*, 7(2), pp.97–106.

UNFCCC, 1999. *Compendium of Decision Tools to Evaluate Strategies for Adaptation to Climate Change*, Bonn, Germany.

Uricchio, V.F., Lopez, N. & Giordano, R., 2004. A fuzzy knowledge-based decision support system for groundwater pollution risk evaluation. *Journal of Environmental Management*, 73, pp.189–197.

Usery, E.L., 1996. A Conceptual Framework and Fuzzy Set Implementation For Geographic Features. In P. Burrough & A. Frank, eds. *Geographic Objects with Indeterminate Boundaries*. Taylor & Francis, pp. 71–87.

Vafeidis, A.T. et al., 2008. A New Global Coastal Database for Impact and Vulnerability Analysis to Sea- Level Rise. *Journal of Coastal Research*, 24(4), pp.917–924.

Vafeidis, A.T. et al., 2004. Developing a global database for coastal vulnerability analysis: design issues and challenges. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 35(B), pp.801–805.

Vanneuville, W. et al., 2005. Spatial calculation of flood damage and risk ranking. In Agile, ed. *8th conference on Geographic Information Science*. pp. 549–556.

Varghese, K. et al., 2008. Identifying critical variables for coastal profiling in ICZM planning—A systems approach. *Ocean & Coastal Management*, 51(1), pp.73–94.

Varnes, D.J., 1984. *Commission on Landslides and Other Mass-Movements-IAEG Landslide Hazard Zonation: A Review of Principles and Practices*, Paris: UNESCO Press.

Vassur, B. et al., 2004. Spatio-temporal Ontology for defining the quality of an application. In *ISSDQ*. p. 13.

Walker, W. et al., 2003. Defining Uncertainty A Conceptual Basis for Uncertainty Management in Model-Based Decision Support. *Integrated Assessment*, 4(1), pp.5–17.

Wang, F., Hall, G. & Brent, H., 1996. Fuzzy representation of geographical boundaries in GIS. *International Journal of Geographical Information Science*, 10(5), pp.537–590.

Worboys, M.F., 1998. Imprecision in finite resolution spatial data. *GeoInformatica*, 2(3), pp.257–279.

Worboys, M.F. & Duckham, M., 2004. *GIS: A computing perspective* Second., CRC Press.

Xhardé R., 2007. *Application des techniques aéroportées vidéographiques et lidar à l'étude des risques naturels en milieu côtier*. Quebec City: INRS.

Xie, Y. et al., 2011. Uncertainty information fusion for flood risk assessment based on DS-AHP method. In *19 th International Confrence on Geoinformatics*. IEEE, pp. 1–6.

Zadeh, L.A., 1965. Fuzzy Set. *Information and control*, 8, pp.338–353.

Zadeh, L.A., 2005. Toward a generalized theory of uncertainty (GTU)—an outline. *Information Sciences*, 172(1-2), pp.1–40.

Zhan, F.B., 1997. *Toplogical relations between fuzzy regions*, Zhan, F.B. & Lin, H., 2003. Overlay of Two Simple Polygons with Indeterminate Boundaries. *Transactions in GIS*, 7(1), pp.67–81.

Zuzekt, P.J., Nairnt, R.B. & Thiemet, S.J., 2003. Spatial and temporal considerations for in calculating shoreline change rates the Great Lakes basin. *Journal of Coastal Research*, 38(SPEC. ISS. 38), pp.125–146.

Appendix A: Fuzzy Membership Function

Membership Function(μ)	
Elevation(m) → x $$\mu=\begin{cases}1 & x\le 3\\ \frac{0.8-1}{10-3}x+1 & 3<x\le 10\\ \frac{0.6-0.8}{16-10}x+0.8 & 10<x\le 16\\ \frac{0.4-0.6}{24-16}x+0.6 & 16<x\le 24\\ 0.2 & x>24\end{cases}$$	**Mean Slop(%)→ x** $$\mu=\begin{cases}1 & x\ge 35\\ \frac{1-0.8}{35-28}x+0.8 & 28\le x<35\\ \frac{0.8-0.6}{28-20}x+0.6 & 20\le x<28\\ \frac{0.6-0.4}{20-13}x+0.4 & 13\le x<20\\ 0.2 & x<13\end{cases}$$
Distance vulnerable object from coastline (m) → x $$\mu=\begin{cases}1 & x\le 100\\ \frac{0.8-1}{200-100}x+1 & 100<x\le 200\\ \frac{0.6-0.8}{300-200}x+0.8 & 200<x\le 300\\ \frac{0.4-0.6}{600-300}x+0.6 & 300<x\le 600\\ 0.2 & x>600\end{cases}$$	**Distance of coastline from depth of 5m (m) → x** $$\mu=\begin{cases}1 & x\le 300\\ \frac{0.8-1}{400-300}x+1 & 300<x\le 400\\ \frac{0.6-0.8}{700-400}x+0.8 & 400<x\le 700\\ \frac{0.4-0.6}{1000-700}x+0.6 & 700<x\le 1000\\ 0.2 & x>1000\end{cases}$$
Erosion Rate (m/yr) → x $$\mu=\begin{cases}1 & x\le -1\\ \frac{0.8-1}{-0.6-(-1)}x+1 & -1<x\le -0.6\\ \frac{0.6-0.8}{-0.1-(-0.6)}x+0.8 & -0.6<x\le -0.1\\ \frac{0.4-0.6}{0-(-0.1)}x+0.6 & -0.1<x\le 0\\ 0.2 & x>0\end{cases}$$	**Geology Type → x** $$\mu=\begin{cases}1 & x=\text{"Delta" or "Dune"}\\ 0.8 & x=\text{"Beach"}\\ 0.6 & x=\text{"Talus with no vegetation"}\\ 0.4 & x=\text{"Talus with vegetation"}\\ 0.2 & x=\text{"Cliff" or "Fjords"}\end{cases}$$
Hydrology Network → x $$\mu=\begin{cases}1 & x=Yes\\ 0.6 & 28=\text{small}\\ 0 & x=\text{non}\end{cases}$$	**Protected Structure → x** $$\mu=\begin{cases}1 & x=\text{"Fault" OR "Fracture" OR "Subsidence"}\\ 0 & x=\text{non}\end{cases}$$

Appendix B: Mathematical Definition of Fuzzy Data Model

Some Fundamental definition in Fuzzy Set Theory:

Definition 1: Let Set A, the interior A is A° and the closure of Set A is $\overline{A}$ and the boundary of set A is ∂A where $\partial A = \overline{A} - A^{\circ}$.
Definition 2. *Complement* of the set A is A^{c} where $\mu_{A^c} = 1^{\Re^n} - \mu_A = \{(x, 1 - \mu_A(x)) \mid x \in \Re^n\}$.
Definition 3. The set A is regularly closed if only if $A = \overline{A^{\circ}}$.
Definition 4. *Subset* $(\mu \subseteq \nu)$: a fuzzy set μ is the subset of υ $iff \forall x \in \Re^n, \mu(x) \leq \nu(x)$.

Fuzzy Data Type: Point, Line, Region, and Partition

Fuzzy Point: A fuzzy point is a 0-D object with a known location *(x,y)* that is associated with a membership function F to a phenomenon of interest:

$$FuzzyPoint = \{\mu \in F(\Re^2) \mid \exists!(x,y) \in \Re^2, \mu(x,y) > 0\} \quad (1)$$

A Fuzzy multipoint in its turn is defined as a finite collection of disjoint fuzzy points:

$$FuzzyMultiPoint = \left\{ \mu \in F(\Re^2) \mid \exists \{\mu_i\}_{i=1}^{n} \subset FuzzyPoint, \mu = \bigcup_{i=1}^{n} \mu_i \right\} \quad (2)$$

Fuzzy Line: A fuzzy line is a 1-Dimensional object that can partially or completely be fuzzy with a broad boundary (fuzzy endpoints) or a broad interior (fuzzy trajectory) (Bejaoui 2009). The fuzzy line μ is yielded from the image of homeomorphism $h \in [0,1]$ from which the extension of a fuzzy line is built within interval (0, 1). The formal definition of a fuzzy line is:

$$\begin{aligned} &FuzzyLine \equiv \{\ \mu \in F(\Re^2) \mid \exists\, \eta \in F([0,1]), \eta = \overline{\eta^{\circ}}, \text{and connected}, \\ &\exists h : [0,1] \to \Re^2 \text{homeomorphism in } (0,1), \text{continous in } \{0,1\}, \qquad (3) \\ &\mu = \tilde{\mathrm{h}}(\eta) \text{ and } (\mathrm{h}(0) = \mathrm{h}(1) \Rightarrow \eta(0) = \eta(1)\ \} \end{aligned}$$

where η is a regular closed fuzzy set in [0, 1]. Respectively, a fuzzy multiline is a finite collection of fuzzy lines whose extensions intersect only at their end node and the lines have the same membership function value at the common end nodes (Dilo 2006). The formal definition of a fuzzy multiline is:

$$\begin{aligned}&FuzzyMultiLine \equiv \{\mu \in F(\Re^2) \mid \exists \{\mu_i = \widetilde{h}_i(\eta_i) \in FuzzyLine, i \in \{1...n\}\}, \mu = \bigcup_{i=1}^{n} \mu_i,\\&\forall i, j \in \{1...n\}, i \neq j \Rightarrow \mu_i \cap \mu_j \subseteq \mu_i^{|\{h_i(0), h_i(1)\}} \cup \mu_j^{|\{h_j(0), h_j(1)\}}\ \}\end{aligned} \tag{4}$$

Note that a fuzzy line is a special case of a fuzzy multiline when *n=1*.

Fuzzy Region: A fuzzy region is a 2-Dimensional object with an indeterminate boundary i.e. a transition zone. This definition is compatible with the model Fuzzy-Fuzzy object of Cheng in Cheng et al. (2009). A fuzzy point with an indeterminate boundary is a specific case of fuzzy region (Dilo et al. 2007). The formal definition of fuzzy region is given by *FuzzyRegion* as:

$$FuzzyRegion \equiv \{\mu \in F(\Re^2) \mid \mu \text{ is bounded}, \mu = \overline{\mu^{\circ}}, \mu^{\circ} \text{ is connected}\} \tag{5}$$

The regular closure property for fuzzy region topology excludes the possibility for a fuzzy line or fuzzy multiline to be a specific case of fuzzy region (Dilo et al. 2007). Respectively, the formal definition of fuzzy multi-region is as follow:

$$FuzzyMultiRegion \equiv \{\mu \in F(\Re^2) \mid \mu \text{ is bounded}, \mu = \overline{\mu^{\circ}}\} \tag{6}$$

Fuzzy Partition: Fuzzy partition is a set of fuzzy multi-regions that may intersect on their uncertain parts. However, the core of one region can intersect with other regions only at their boundaries. Fuzzy partition is defined as:

$$\begin{aligned}&FuzzyPartiotion \equiv\\&\left\{\begin{aligned}&\{\mu_i\}_{i=1}^{n} \subset FuzzyRegion \mid \forall \text{i}, \text{j} \in \{1...\text{n}\}, i \neq j \Rightarrow\\&\mu_i \cap \mu_j \subseteq RegionBoundary(\mu_i) \cap RegionBoundary(\mu_j)\end{aligned}\right\}\end{aligned} \tag{7}$$

where *RegionBoundary* is the uncertain part of the partition. This uncertain zone is yielded from the fuzzy boundary μ^b , or the fuzzy extension of the partition that consists of fuzzy regions and fuzzy lines. The formal definition of the fuzzy boundary of a region is:

$$\begin{aligned}&RegionBoundary : FuzzyMultiRegion \rightarrow FuzzyExtention\\&\forall \mu \in FuzzyMultiRegion, RegionBoundary(\mu) = \mu^b\end{aligned} \tag{8}$$

Fuzzy discretization of space into regular or irregular grids e.g. raster format or TIN-based, is practically a specific case of a fuzzy multiregions type representation. A fuzzy partition is a soft classification of space.

Fuzzy Operators: Union, Intersection, Difference, Overlay, Fusion

Union of Fuzzy Multipoints: The union between two fuzzy multipoints is a fuzzy multipoint yield from the union of input point locations. The membership value is then the maximum membership of the input points. The fuzzy point union (*PUnion*) is defined as follows:

$$\begin{aligned}&PUnion : FuzzyMultiPoint \times FuzzyMultiPoint \rightarrow FuzzyMultiPoint\\&\forall \mu, \nu \in FuzzyMultiPoint, PUnion(\mu,\nu) = \mu \bigcup \nu\end{aligned} \qquad (9)$$

Intersection of Fuzzy Multipoints: The intersection of two fuzzy multipoints is a new fuzzy multipoint whose locations are the common locations of input points. The membership of each location is then the minimum of memberships of input points. The mathematical definition of a fuzzy point intersection (*PIntersection*) is:

$$\begin{aligned}&PIntersection : FuzzyMultiPoint \times FuzzyMultiPoint \rightarrow FuzzyMultiPoint\\&\forall \mu, \nu \in FuzzyMultiPoint, PIntersection(\mu,\nu) = \mu \bigcap \nu\end{aligned} \qquad (10)$$

Difference of Fuzzy Multipoints : The difference between two fuzzy multipoints is a fuzzy multipoint whose locations are derived from those of the first input object. The membership is then the difference between the memberships of two input objects. The formal definition is as follow:

$$\begin{aligned}&PDifference : FuzzyMultiPoint \times FuzzyMultiPoint \rightarrow FuzzyMultiPoint\\&\forall \mu, \nu \in FuzzyMultiPoint, PDifference(\mu,\nu) = \mu - \nu\end{aligned} \qquad (11)$$

The union, intersection, and difference operators for fuzzy line are yielded from corresponding fuzzy operators that are explained hereafter.

Union of Fuzzy Multilines: The union of two fuzzy multilines is a fuzzy multiline. The formal definition is given by *LUnion*:

$$\begin{aligned}&LUnion : FuzzyMultiLine \times FuzzyMultiLine \rightarrow FuzzyMultiLine\\&\forall \mu, \nu \in FuzzyMultiLine, LUnion(\mu,\nu) = \mu \bigcup \nu\end{aligned} \qquad (12)$$

Intersection of Fuzzy Multilines: The intersection of two fuzzy multilines is the intersection of two lines interiors or end points. The formal definition of the fuzzy intersection of two fuzzy multilines is given by *LIntersection* as (Dilo 2006):

$$LIntersection : FuzzyMultiLine \times FuzzyMultiLine \rightarrow FuzzyMultipoint$$
$$\forall \mu, \nu \in FuzzyMultiLine, LIntersection(\mu, \nu) = (\mu \bigcap \nu)^{|crsipIntersection(supp(\mu), supp(\nu))} \quad (13)$$

Difference of Fuzzy Multilines: The difference operator of two multilines produces a fuzzy multiline as regular closure of fuzzy difference. Reminding that, a regular closed set is if only if a set A is equal to its closure interior i.e. $A = \overline{A^{\circ}}$. *LDifference* denotes as formal definition of difference operator on two fuzzy multiline following as:

$$LDifference : FuzzyMultiLine \times FuzzyMultiLine \rightarrow FuzzyMultiLine$$
$$\forall \mu, \nu \in FuzzyMultiLine, LDifference(\mu, \nu) = \overline{(\mu - \nu)^{\circ}} \quad (14)$$

The union, intersection and difference operators for fuzzy region and fuzzy multi-region produce a fuzzy multi-region.

Union of Fuzzy Multiregions: The union operator *RUnion* is defined as:

$$RUnion : FuzzyMultiRegion \times FuzzyMultiRegion \rightarrow FuzzyMultiRegion$$
$$\forall \mu, \nu \in FuzzyMultiRegion, RUnion(\mu, \nu) = \mu \bigcup \nu \quad (15)$$

Intersection of Fuzzy Multiregions: Fuzzy intersection of two fuzzy multiregion is the regular closure of their fuzzy set intersection. The formal definition of fuzzy intersection *RIntersection* is indicated as:

$$RIntersection : FuzzyMultiRegion \times FuzzyMultiRegion \rightarrow FuzzyMultiRegion$$
$$\forall \mu, \nu \in FuzzyMultiRegion, RIntersection(\mu, \nu) = \overline{\mu^{\circ} \bigcap \nu^{\circ}} \quad (16)$$

Difference of Fuzzy Multiregions: The fuzzy difference between two fuzzy multi-regions produces a bounded fuzzy set but not always a regular closed fuzzy set. Respectively the formal definition of fuzzy difference for two fuzzy multi-regions is described as:

$$RDifference : FuzzyMultiRegion \times FuzzyMultiRegion \rightarrow FuzzyMultiRegion$$
$$\forall \mu, \nu \in FuzzyMultiRegion, RDiffernece(\mu, \nu) = \overline{\mu^{\circ} \bigcap (1^{\Re^2} - \nu)} \quad (17)$$

Fuzzy Overlay of Fuzzy Partitions: The most common operators over fuzzy partitions are overlay and fusion. Fuzzy overlay superimposes two fuzzy partitions and build a new fuzzy partition with a fuzzy multi-region obtained from the intersection of fuzzy multi-region of the first and the second partition. In fact, the overlay operator is a kind of refinement of two fuzzy partitions by combining two fuzzy classifications. The formal definition of fuzzy overlay *FuzzyOverlay* is defined as:

$$
\begin{aligned}
&FuzzyOverlay : FuzzyPartition \times FuzzyPartition \rightarrow FuzzyPartition \\
&\forall P_1 = \{\mu_i\}_{i=1}^{n}, P_2 = \{\nu_j\}_{j=1}^{n} \in FuzzyPartition, \\
&FuzzyOverlay(P_1, P_2) = \left\{ \begin{array}{l} \zeta_{i,j} \mid i \in \{1...n\}, j \in \{1...m\}, \\ \zeta_{i,j} = RegionIntersection(\mu_i, \nu_j) \end{array} \right\}
\end{aligned} \tag{18}
$$

where $\{\zeta_{i,j} \mid i \in \{1...n\}, j \in \{1...m\}\}$ is the new fuzzy partition.

Fuzzy Fusion of Fuzzy Partitions: The fuzzy fusion operator dissolves a fuzzy partition by merging fuzzy multi-regions based on grouping or likeness of some attribute values of the regions. This operator also allows generalizing a fuzzy partition. Associating fuzzy multi-regions to a fuzzy partition based on an attribute require another operation called attribute extended fuzzy partition (*AFPartition*) (Dilo et al. 2007). Let's present the domain of an attribute by *ADomain*. Then, the relation between *AFPartition*, *ADomain*, and *FuzzyMultiRegion* for a fuzzy partition is as below:

$$
AFPartition \equiv \left\{ \begin{array}{l} \{(\mu_i, \nu_i)\}_{i=1}^{n} \subset FuzzyRegion \times ADomain \\ \mid \{\mu_i\}_{i=1}^{n} \in FuzzyPartition \end{array} \right\} \tag{19}
$$

Let's group the attribute values in function $g : ADomain \rightarrow ADomain$. *g* is an element of Power set $\mathrm{P}(ADomain \times ADomain)$ which is the collection of subsets of the Cartesian product of *ADomain* with itself. The fusion operator is set up follows:

$$
\begin{aligned}
&FuzzyFusion : AFPartition \times P(ADomain \times ADomain) \rightarrow AFPartition \\
&\forall A = \{(\mu_i, \nu_i)\}_{i=1}^{n} \in AFPartition, \forall g \in P(ADomain \times ADomain), \\
&FuzzyFusion(A) = \left\{ \begin{array}{l} \{(\zeta_j, \omega_j)\}_{j=1}^{m} \mid \{\omega_j\}_{j=1}^{m} = ran(g), \forall j \in \{1...m\}, \\ \zeta_j = \bigcup \{\mu_i \mid g(\nu_i) = \omega_j\} \end{array} \right\}
\end{aligned}
$$

Appendix C: Implementation: Matlab Code

A **Matlab** code is developed to implement the framework presented in Figure 4.4. The code is divided in 3 main part: 1) draw grid (Table AppC.1), 2) Prepare information for each indicator (cell center, respective membership value) and represent risk level as fuzzy grid (Table AppC.2), 3) Aggregate multiple indicators using union, intersection, difference, mean, mean weighted, fusion (see Table AppC.3-8).

The input of Draw grid in Matlab code is the minimum and maximum position values (X_min,Y_min; X_max, Y_max) to create a box. Then the cell size should be introduced. The output is the square grid with given cell size.

Table AppC.1: Draw grid

```
clear all

x_max=40;
y_max=20;
step=0.05;

grid_node_x=0:step:x_max;
grid_node_y=0:step:y_max;

for i=1:x_max
    figure(1)
    line([i i],[0 y_max])
end

for i=1:y_max
    figure(1)
    line([0 x_max],[i i])
end
```

To prepare the information about each indicator, the position and respective the membership values should be entered. A Gaussian function is assigned to assign the membership value to center of each cell. A radius of r is used to determine the extension of Gaussian bell. This value is constant along the process. Then, the risk level with respect to given indicator is displayed.

Table AppC.2: Prepare the information for each indicator (center of cells, membership value) and represent risk degree for a specific indicator

```
% -----------------------------------------------------------------
% Indicator 1
% -----------------------------------------------------------------

centers1=[3 6 12 15 18 5 25 20 15 18 5 3 8 2;
          2 8 15 17 12 11 14 18 19 16 8 17 15 16];
```

```
r1=50*[1 1 1 1 1 1 1 1 1 1 1 1 1 1];
Amplitude1=[1 0.65 0.7 0.7 0.79 1 0.0001 0.93 1 0.8 0.7 0.89
0.0001 1];

[X Y]=meshgrid(grid_node_x,grid_node_y);
Z1=nan([length(Amplitude1),size(X)]);

for i=1:length(Amplitude1)
    Z1(i,:,:) = Amplitude1(i)*...
        exp(-1/r1(i)*sqrt((centers1(1,i)-X).^2+(centers1(2,i)-
Y).^2));
end
z1_max=reshape(max(Z1,[],1),y_max/step+1,x_max/step+1);
z1_max(1)=1;
z1_max(end)=0;
figure(2)
hold on
mesh(X,Y,z1_max)
xlabel('X')
ylabel('Y')
```

Once risk is calculated and represented by running Matlab code, transferring the result to a GIS tool is required. For an appropriate visualisation, it is recommended the result is rectified on a BaseMap. This step is done in ArcMap. This permits to realize how the resulted risk can affect the real world.

Fuzzy Aggregation

To calculate the overall risk some aggregation operators are required. In this regard, the Matlab code is developed to perform overlay (union, intersection, difference and mean, and mean weighted) and fusion of two or more fuzzy grids. The details are provided respectively from Table 3 to 8. The theoretical aspects of these concepts are provided in Chapter 2 and 5. As mentioned in Chapter 5, these operators can be easily adapted in a spatial multidimensional database to handle with fuzzy data including in fuzzy dimension. These operators can handle as well the crisp data too. The illustration of each operator is provided in Chapter 5.

Table AppC.3: Union of multiple layers of fuzzy grids

```
% ----------------------------------------------------------------
% Aggregate(MAX)
% ----------------------------------------------------------------
z_final_u_1=max(z1_max,z2_max);
z_final_u_2=max(z_final_u_1,z3_max);
z_final_u_3=max(z_final_u_2,z4_max);
z_final_u_4=max(z_final_u_3,z5_max);
z_final_u=max(z_final_u_4,z6_max);
z_final_u(1)=1;
z_final_u(end)=0;
figure(7)
```

```
hold on
mesh(X,Y,z_final_u)
xlabel('X')
ylabel('Y')
% % legend ('Risk')
xlim([x_min,x_max])
ylim([y_min,y_max])
```

Table AppC.4: Intersection of multiple layers of fuzzy values

```
% -----------------------------------------------------------------
% Aggregate(MIN)
% -----------------------------------------------------------------
z_final_i_1=min(z1_max,z2_max);
z_final_i_2=min(z_final_i_1,z3_max);
z_final_i_3=min(z_final_i_2,z4_max);
z_final_i_4=min(z_final_i_3,z5_max);
z_final_i=min(z_final_i_4,z6_max);
z_final_i(1)=1;
z_final_i(end)=0;
figure(8)
hold on
mesh(X,Y,z_final_i)
xlabel('X')
ylabel('Y')
xlim([x_min,x_max])
ylim([y_min,y_max])
```

Table AppC.5: Calculate Mean weighted of multiple layers of fuzzy values

```
% -----------------------------------------------------------------
% Aggregate(Mean weighted)
% -----------------------------------------------------------------
z_final_mw=((0.1*z1_max)+(0.08*z2_max)+(0.08*z3_max)+(0.23*z4_max)
+(0.17*z5_max)+(0.34*z6_max));
z_final_mw(1)=1;
z_final_mw(end)=0;
figure(9)
hold on
mesh(X,Y,z_final_mw)
xlabel('X')
ylabel('Y')
xlim([x_min,x_max])
ylim([y_min,y_max])
```

Table AppC.6: Mean of multiple fuzzy grids

```
% -----------------------------------------------------------------
% Aggregate(Mean)
% -----------------------------------------------------------------
z_final_m=(z1_max+z2_max+z3_max+z4_max+z5_max+z6_max)/6;
z_final_m(1)=1;
z_final_m(end)=0;
figure(10)
```

```
hold on
mesh(X,Y,z_final_m)
xlabel('X')
ylabel('Y')
xlim([x_min,x_max])
ylim([y_min,y_max])
```

Table AppC.7: Difference of two fuzzy grids

```
% ----------------------------------------------------------------
% Aggregate(Difference)
% ----------------------------------------------------------------
z_final_D=min(z1_max, 1-z2_max);
z_final_D(1)=1;
z_final_D(end)=0;
figure(8)
hold on
mesh(X,Y,z_final_D)
xlabel('X')
ylabel('Y')
```

As mentioned in chapter 5, fusion operator can be performed in two ways:

1) Performing overlay operation on multiple layers of information on finer level, and then creating merged cells in coarser scale,

2) Performing merged-cell operation and then assign the result of overlay operation to obtain the final result in coarser scale.

The result of a fusion operator is a new fuzzy partition. The Syntax of Fusion operator is as follows:

FuzzyFusion (Input grid, cell_factor, {Fuzzyoverlay_type}, output grid)

Regarding to demand of the user, the fuzzy overlay type (union, intersection, difference, mean, mean weighted) can be selected. In Table 6.14, Mean and union operators are illustrated.

Table AppC.8: Fusion of two fuzzy grids toward a multi-scale representation

```
% ----------------------------------------------------------------
% Aggregate(Fusion -Mean weighted)
% ----------------------------------------------------------------
fusion_cell_dim=10;
[y_max x_max]=size(z_final_mw);
x_lower=1:fusion_cell_dim:x_max-fusion_cell_dim;
x_upper=fusion_cell_dim:fusion_cell_dim:x_max;
y_lower=1:fusion_cell_dim:y_max-fusion_cell_dim;
y_upper=fusion_cell_dim:fusion_cell_dim:y_max;
```

```
z_fusion_mean=z_final_mw;

for i=1:length(x_upper)
    for j=1:length(y_upper)

temp=z_final_mw(y_lower(j):y_upper(j),x_lower(i):x_upper(i));
       fusion_result=mean(mean(temp));

z_fusion_mean(y_lower(j):y_upper(j),x_lower(i):x_upper(i))=fusion_
result;
    end

end
figure(9)
hold on
mesh(X,Y,z_fusion_mean)
xlabel('X')
ylabel('Y')
% -----------------------------------------------------------------
% Aggregate(Fusion-Max)
% -----------------------------------------------------------------
fusion_cell_dim=10;
[y_max x_max]=size(z_final_max);
x_lower=1:fusion_cell_dim:x_max-fusion_cell_dim;
x_upper=fusion_cell_dim:fusion_cell_dim:x_max;
y_lower=1:fusion_cell_dim:y_max-fusion_cell_dim;
y_upper=fusion_cell_dim:fusion_cell_dim:y_max;

z_fusion_max=z_final_mw;

for i=1:length(x_upper)
    for j=1:length(y_upper)

temp=z_final_mw(y_lower(j):y_upper(j),x_lower(i):x_upper(i));
       fusion_result=max(max(temp));

z_fusion_max(y_lower(j):y_upper(j),x_lower(i):x_upper(i))=fusion_r
esult;
    end

end
figure(10)
hold on
mesh(X,Y,z_fusion_max)
xlabel('X')
ylabel('Y')
```

Appendix D: Erosion Rate Calculation

Based on the developed framework in chapter 3 (Figure 3.4), coastal erosion risk consists of coastal hazard (here erosion), element at risk (such as road network) and coastal vulnerability index. Each of these parameters can be spatially defined in a GIS environment.

Identity Coastal Erosion

As stated before, various numerical and statistical methods exist to analyze coastal erosion. The coastline change rate is used in this study to accomplish CERA, while coastline change is extracted from Digital Terrain Model in multiple epochs. There are many tools or models to calculate the erosion rate such as Digital Coastline Analysis System (DSAS) (Thieler et al. 2009) that is employed to obtain the erosion rate in this study.

Three epochs of LiDAR data are used in this study to extract coastline change rates based on Mean Sea Level (MSL) (altitude=0). LP360, a toolkit on ArcGIS is used to analyse LiDAR data. LP360 is an extension to ArcMap that allows visualizing and processing of very large point clouds (LIDAR and dense image matching) in a familiar GIS desktop environment. LP360 provides tools from rapid visualization and derived product generation through advanced features such as automatic ground classification and building footprint extraction. The workflow of LP360 is illustrated in Figure AppD.1.

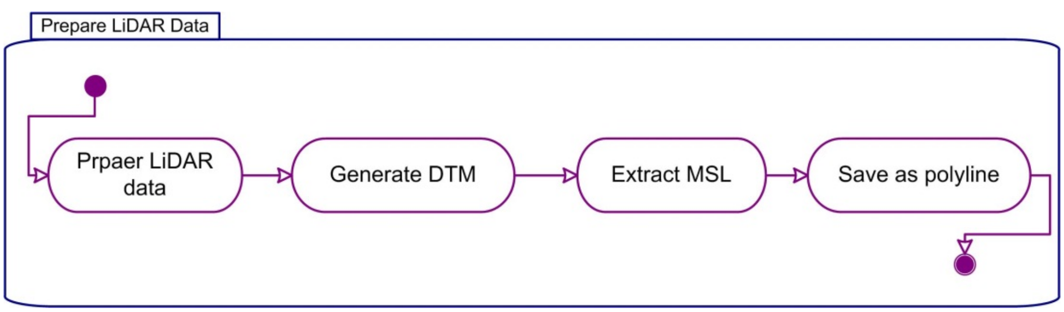


Figure AppD.1: The workflow of using LP360 to prepare LiDAR data for erosion analysis

DTM is generated for each epoch and then the MSL is derived by drawing the contour line of zero (altitude=0). The contour line zero indicates the coastline on a given time period. By calculating the displacement of coastline on different epochs, coastal change rate is estimated. To do so, DSAS is used to calculate the change rate along the coastline.

DSAS is a freely available software application that works within the Environmental Systems Research Institute (ESRI) Geographic Information System (ArcGIS) software. DSAS computes rate-of-change statistics for a time series of coastline vector data. The input of DSAS is the multiple coastlines in different epochs, a baseline as reference, and a distance between two transact cast. DSAS generates an XML-based output table that contains the distance measurements used to compute rate-of-change statistics along each transect (See Figure AppD.2).

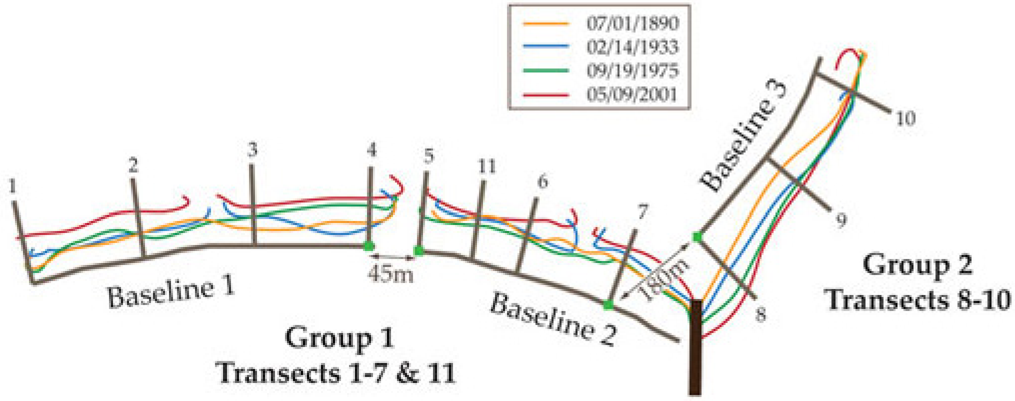
Coastline positions can reference several different features such as the vegetation line, the high water line, the low water line, or the wet/dry line. They can be digitized from a variety of sources (for example, satellite imagery, digital orthophotos, historical coastal-survey maps), collected by global-positioning-system field surveys, or be extracted from LiDAR surveys. It is strongly recommended that initial data-preparation steps be taken to reference all coastline vectors to the same feature (for example, MSL in this study) before using DSAS to compute change statistics. Each coastline vector represents a specific position in time and must be assigned a date in the coastline feature-class attribute table. The measurement transects that are cast by DSAS from the baseline will intersect the coastline vectors. The points of intersection provide location and time information used to calculate rates of change. The distances from the baseline to each intersection point along transect are used to compute the selected statistics. Users have the option of specifying for each coastline an overall uncertainty value, which should account for both positional and measurement uncertainties. DSAS uses a measurement baseline method (Leatherman & Clow 1983) to calculate rate-of-change statistics for a time series of coastlines. The baseline is constructed by the user and serves as the starting point for all transects cast by the DSAS application. Transect intersects each coastline at the measurement points used to calculate coastline-change rates. There are three ways to create a baseline:

1) Start with a new feature class,

2) Buffer an existing coastline,

3) Use a pre-existing baseline.

In this study, we used the option of buffering an existing coastline. There are several statistical methods to calculate the rate of erosion. Some of most commonly used methods are provided in Table AppD.1 and supported by DSAS. (Genz et al. 2007) have overviewed all these methods with concern on their advantages and disadvantages.



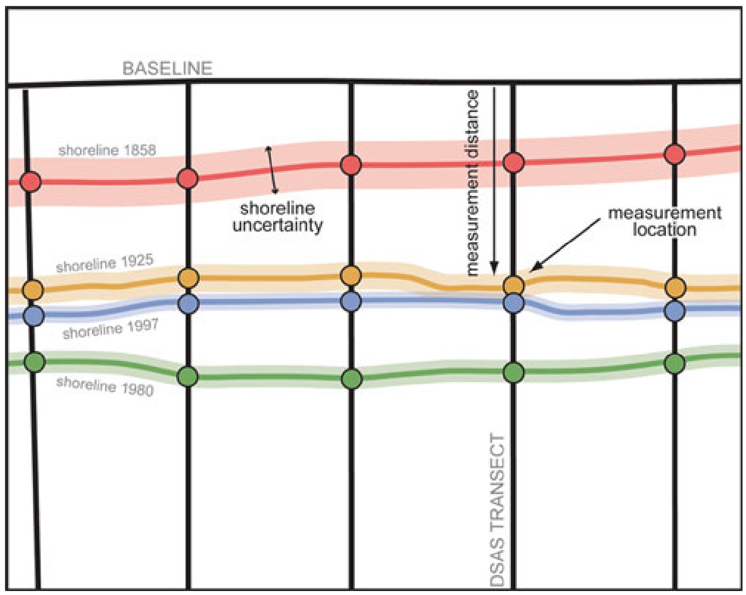


Figure AppD.2: A schematic view how DSAS calculate coastline change rate

Table AppD.1: Statistical methods to calculate erosion rate (derived from (Genz et al. 2007)

Method	Description
Net Coastline Movement	It reports a distance, not a rate. The NSM is associated with the dates of only two coastlines. It reports the distance between the oldest and youngest coastlines for each transect. This represents the total distance between the oldest and youngest coastlines.
Coastline Change Envelope	It is reports a distance, not a rate that is the distance between the coastlines farthest and closest to the baseline at each transect. This represents the total change in coastline movement for all available coastline positions and is not related to their dates.
End Point Rate	It is calculated by dividing the distance of coastline movement by the time elapsed between the oldest and the most recent coastline.
Linear Regression Rate	A linear regression rate-of-change statistic can be determined by fitting a least-squares regression line to all coastline points for a particular transects. The regression line is placed so that the sum of the squared residuals (determined by squaring the offset distance of each data point from the regression line and adding the squared residuals together) is minimized. The linear regression rate is the slope of the line. The method of linear regression includes these features: 1) All the data are used, regardless of changes in

	trend or accuracy, 2) The method is purely computational, 3) The calculation is based on accepted statistical concepts, and 4) The method is easy to employ.
Weighted Linear Regression Rate	In a weighted linear regression, more reliable data are given greater emphasis, or weight towards determining a best-fit line. In the computation of rate-of-change statistics for coastlines, greater emphasis is placed on data points for which the position uncertainty is smaller. The weight (w) is defined as a function of the variance in the uncertainty of the measurement (e): $w = \frac{1}{e^2}$
Least Median of Squares	In ordinary and weighted least-squares regression, the best-fit line is placed through the points in such a way as to minimize the sum of the squared residuals. In the linear regression method, the sample data is used to calculate a mean offset, and the equation for the line is determined by minimizing this value so that the input points are positioned as close to the regression line as possible. In the least median of squares method the median value of the squared residuals is used instead of the mean to determine the best-fit equation for the line.

Once the required geodatabase and input-feature classes have been created or imported from shape files and all necessary feature classes have been added and properly attributed, the DSAS Application can be used within ArcMap to establish transect locations and calculate change statistics. Once the change rate is calculated the end-result can be opened in ArcMap display for future analysis. The workflow of DSAS is illustrated in Figure AppD.3. The final result can be joint or integrated with other information in any GIS environment. An example of final result in ArcMap is presented in Figure AppD.4.

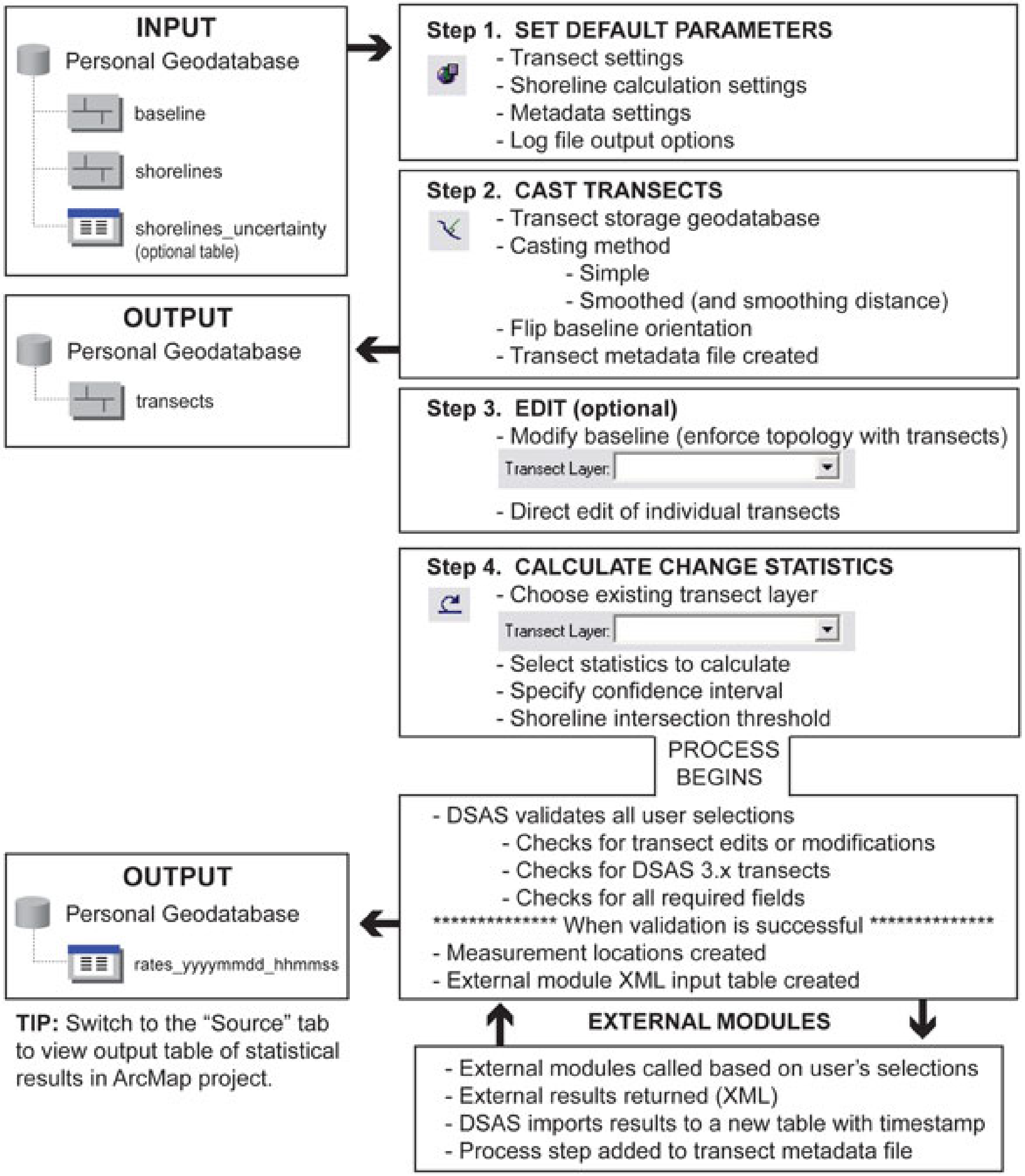


Figure AppD.2: The workflow of DSAS

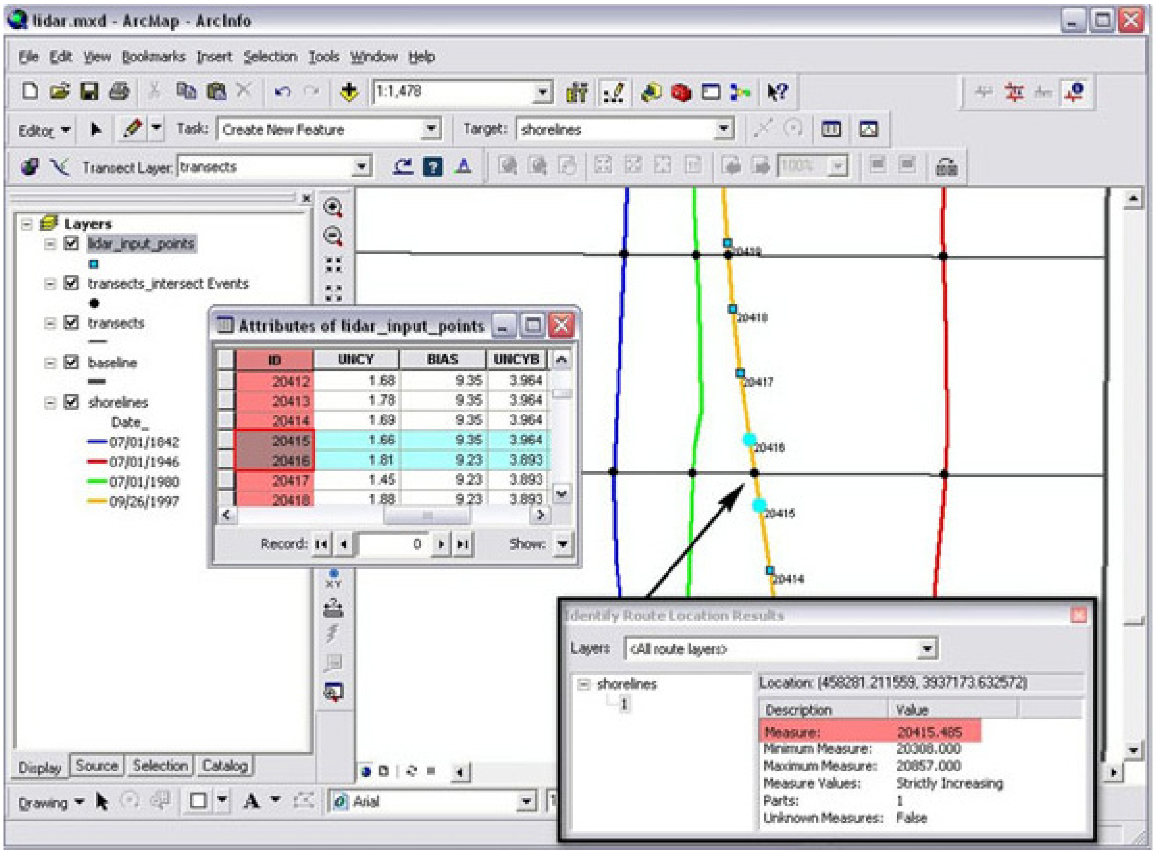


Figure AppD.3: An example of transact profile along the coast derived by DSAS