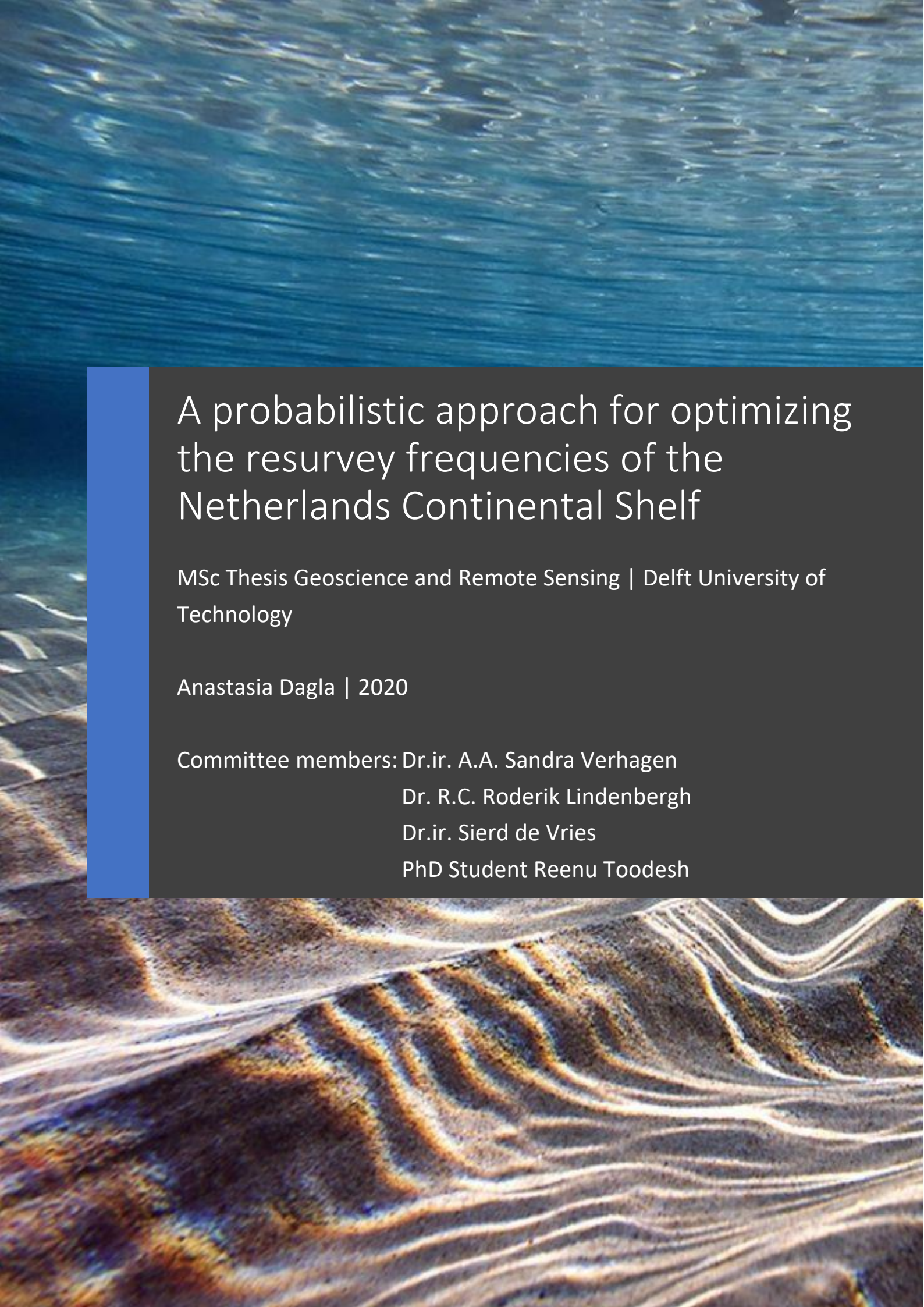


A probabilistic approach for optimizing the resurvey frequencies of the Netherlands Continental Shelf

MSc Thesis Geoscience and Remote Sensing | Delft University of Technology

Anastasia Dagla | 2020

Committee members: Dr.ir. A.A. Sandra Verhagen
Dr. R.C. Roderik Lindenbergh
Dr.ir. Sierd de Vries
PhD Student Reenu Toodesh

A PROBABILISTIC APPROACH FOR OPTIMIZING THE RESURVEY FREQUENCIES OF THE NETHERLANDS CONTINENTAL SHELF

Master of Science in Geoscience & Remote Sensing

by

Anastasia Dagla

January 2020

PREFACE

The present report describes in detail the work that was done during my thesis research as a part of the master's degree of Geoscience and Remote Sensing at Delft University of Technology. This bathymetry project caught my interest as it was related to a real-time problem, which can be summarized to the attempt of surveying the Dutch waters in the most efficient way. I had to develop a result-based, data-driven and automate approach to support the decision-making. For this purpose, working with probabilities was necessary, which has always been a great interest of me as it is this part of mathematics science that can be combined with the engineering set of mind to solve real-world problems. Therefore, it was a delight that I had the time to dig up the literature again and deepen my knowledge in this field of study.

I am grateful to all those people who supported me when doing this research, despite the difficult times of social and physical distancing we are going through. Firstly, I would like to thank my daily supervisor Dr.ir. A.A. Sandra Verhagen for giving me the opportunity to take part in this bathymetry project. I would also like to thank the PhD student Reenu Toodesh for introducing me to the topic and spending several hours for brainstorming together. I thank all the committee members for always being willing to help me with patience, advice and guidance throughout the research process. Finally, t I would like to thank my friends, housemates and parents who supported me with understanding and were always there for me.

I hope you will enjoy reading this thesis.

ABSTRACT

Bathymetric surveying of the Netherlands Continental Shelf (NCS) is taken care of, by the Dutch Hydrographic Office and Rijkswaterstaat. The survey frequencies vary from location to location depending on a set of factors defined by the authorities. These are the minimum depth, draught (difference between seafloor depth and ship's keel), shipping intensity, human interventions and sea floor dynamics that are taking place in an area.

The project is about developing a methodology using the available time-series of bathymetric data to automatically make a well-informed decision on the survey frequency of critical areas of the NCS. This study aims to contribute to the decision-making by performing a risk-based analysis using a probabilistic approach. The main objectives are to estimate and map the risk that the seafloor is changing significantly. Risk severity should be measured and the probability of risk's consequence to occur should be estimated and mapped in temporal and spatial scale indicating if possible, the optimal surveying frequency.

Firstly, the accuracy, quality and reliability of the prediction values was defined and computed. This was a necessary step to realize how powerful are the predictions used as an input dataset for the risk assessment of the seafloor to change significantly. As a step further, the risk's impact and probability to occur should be determined. Hence, it was necessary to define quantities related to seabed dynamics that express the evolution of the seafloor. These are mentioned as risk indicators and are the Relative Depth Change (RDC), the Rate of change (RoC) and the direction of change. The RoC was used for probability estimation and the rest of it as informative variables for the analysis. Finally, an adjustment of the method used here with the Netherlands Hydrographic Office (NLHO) standards was done and conclusions were made regarding the survey frequency of the chosen study areas. The analysis was applied for two study areas: West IJmuiden and West Rotterdam, as these areas according to NLHO are highly prioritised (should be resurveyed every 2 years) and thus, they are critical areas of the NCS. The method applied and the results can be used as an informative tool in designing a successful, reliable and cost-effective bathymetric survey plan of the NCS.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

1 INTRODUCTION

In this chapter an introduction to the problem of optimizing the resurvey frequencies of the NCS is being made. Firstly, the necessary background knowledge to introduce the topic is pointed out. This includes the special characteristics of the seafloor of the NCS, as well as, the chosen bathymetric measurement techniques and an overview of the survey policy and the factors that it is based on. Considering the current situation, the research question, as well as, the sub-questions of this study are formulated. Finally, the structure of the thesis is provided in the last section, to give an overview of each chapter's content.

1.1 BACKGROUND

Oceans are a vital component of our ecosystem. They cover more than 70% of Earth's surface and until now 95% of the underwater is completely unexplored. In recent years there is a growing interest of the world's economy in maritime activities and sustainable use of ocean's wealth. The concept of "Blue Economy" was introduced to sustainably manage ocean resources. It is defined as *"All economic activities related to oceans, seas and coasts. Blue Economy covers a wide range of interlinked established and emerging sectors."* [7]. These sectors are associated with the renewable energy, maritime transport, fisheries, tourism, climate change and waste management. European Union (EU) statistics of 2015 show that around 5% of the EU's Gross Domestic Product (GDP) comes from maritime activities with more than 5 million people working on this field, which yields EUR 495 billion to the European economy. 90% of the world trade and around 43% of EU trade takes place over water [8].

The North Sea is a typical example of Blue Economy's high potential. It borders with strong economies and is one of the busiest seas in the world. Activities like extensive shipping, submarine pipeline networking and renewable energy use (growing infrastructure of offshore wind platforms) are taking place simultaneously. Additionally, one of the world's largest ports, the Dutch port of Rotterdam ends up in the North Sea [9]. When it comes to the Netherlands, it is obvious that "Blue Economy" plays a significant role to its economic success. According to the Netherlands Maritime Technology (NMT) trade association, about 2.5 million people are living close to maritime regions and are employed in this sector. Offshore infrastructure, modern hydrographic products and services are a keystone for sustaining Dutch wealth [10].

At the same time, the NCS as part of the North Sea presents a very dynamic seafloor (Figure 1.1). The seabed of the Netherlands is very impressive, especially when compared to its rather smooth topography. There is a combination of very shallow and sandy seafloor, which is characterized by particularly intense and relatively fast natural processes that cause significant changes to the submarine terrain. Various structures and spatial patterns of sea bed dynamics such as ripples, megaripples, sand banks and tidal sand waves with different characteristics migrate in the NCS. The latter, sand waves, is known as the most relevant seabed form. Large sand waves can change seafloor depths very fast [11]. They typically have heights of tens of meters, wavelengths of hundreds of meters and their migration rate is up to 10m/year [5], [12], [11]. Sand waves have been subject of research repeatedly over

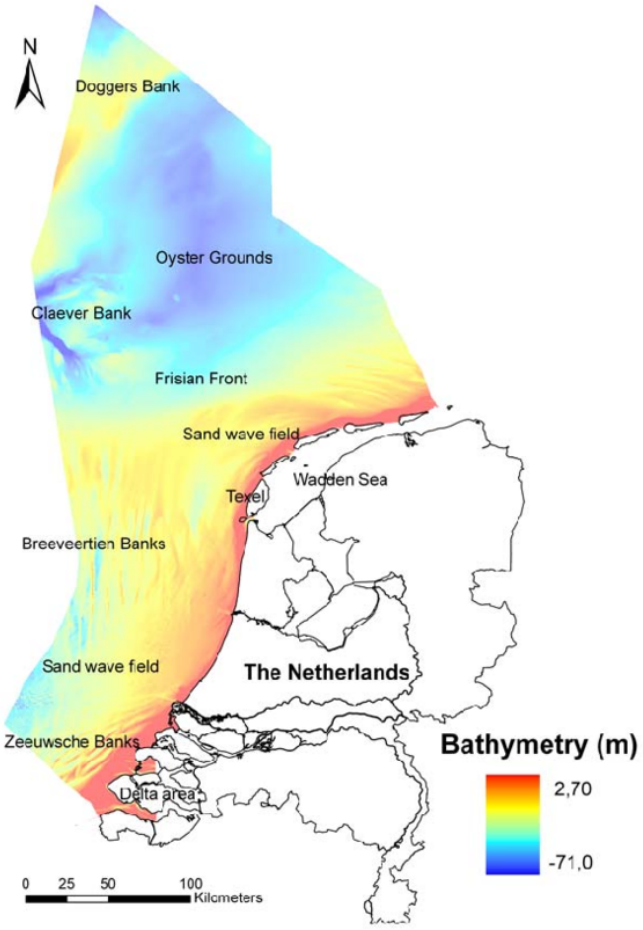


Figure 1.1: Bathymetric map of the NCS, which also presents an overview of the borders of the NCS, [1]. Red (close to the coast) and yellow areas are shallow areas with depth values below 40m, while blue areas are deeper with depth values that reach 70m.

the last years as their presence is recorded in many parts of the North Sea and could cause danger to passing ships and to submarine pipelines and cables. Additionally, it is assumed to be the most complicated factor to model among the factors that drive the current resurvey policy.

To support the existing maritime activities and to encourage future investments in the territory, it is necessary to detect the seafloor changes and map them frequently to ensure security of navigation. Hence, acquiring full and accurate geospatial data of the seafloor and coastal areas, the so-called *bathymetric data*, and use them for generating up-to-date nautical charts is of great importance. The NLHO and Rijkswaterstaat (RWS) are both responsible for mapping and charting the NCS as well as for ensuring safe navigation in the dynamic and shallow Dutch waters [13]. RWS is in charge of the nearshore coastal areas, while the NLHO is accountable for the deep waters of the NCS.

NLHO and in particular, the Hydrographic Service of the Royal Netherlands Navy is responsible to provide information to the mariners related to the shipping routes, the seabed and underwater hazards such as shipwrecks. This means, that bathymetric surveys are carried out frequently in order to update the nautical charts. The NLHO, thus, is in charge of protecting the Dutch interests at sea and should ensure safe shipping navigation inside the borders of the NCS [14].

Bathymetric surveying of the NCS is carried out by two hydrographic survey vessels of NLHO that use echo sounding; HNLMS Luymes and HNLMS Snellius. Sounding is the most suitable technique for measuring underwater depths. Acoustic waves have the ability to travel deep down to the sea bottom, compared to electromagnetic radiation which fades relatively fast when it penetrates into the water. Single-Beam Echo Sounder (SBES) and the evolution to Multi-Beam Echo Sounder (MBES) are used to determine the depth values at specific locations. The MBES introduced a new era in bathymetric surveying, as it uses a swath system that produces multiple acoustic beams and provides increased bottom coverage, compared to the SBES which measures the depth at nadir by using a single beam width. Despite the evolution in the surveying equipment, dealing with the measurement uncertainties is still a challenging topic [6].

Specific accuracy requirements for bathymetric surveys are defined by the International Hydrographic Organisation (IHO) and, in particular, by the latest version of "IHO Standards for Hydrographic Surveys, special publication N. 44, February 2020" (S-44) [15]. However, in the international policy of IHO there are no recommendations regarding how often the seabed should be measured, especially in dynamic seas. NLHO designs the survey policy and planning of the NCS, according to S-44 regulations, but at the same time it should develop its own policy regarding issues not covered by the framework of IHO. According tho this policy the NLHO assigns resurvey frequencies to all areas depending on five factors; the minimum depth, the shipping draught , the shipping intensity, the seabed dynamics and the undertaken human intervention in each area of the NCS [5], [13]. This process results into prioritizing all the areas of the NCS (categorisation) based on their assigned resurvey frequency. Areas with high resurvey frequencies are the so-called *critical areas* of the NCS. For instance, very shallow and dynamic areas, which are principal shipping routes of large and deep draught ships will be highly prioritized and hence most frequently resurveyed.

The operation of hydrographic vessels is very expensive and limited resources, in terms of survey ships, are available. At the same time, the nautical charts of the shallow and dynamic waters of NCS become outdated very fast. Hence, it is necessary to carefully design a reliable and efficient resurvey plan of the NCS, without areas being over- or under-surveyed. Compared to the other neighboring countries, the assigned resurvey frequencies of NLHO are rather ambitious and it remains an open question; if an optimization of the current survey policy is possible, by reducing the resurvey frequencies or the size of the areas that should be resurveyed [13]. Further research is still needed in order to include fundamental knowledge regarding seafloor dynamics for optimising the assigned resurvey frequencies and this is the aim of this study.

Firstly, it is necessary to develop a scientific and data-driven approach in order to understand the behaviour of seabed dynamics, model them, deal with their uncertainties and make reliable predictions. The amount of data produced by modern survey equipment is much higher and requires efficient, automatic methods [12]. Lindenbergh [12] and Dorst [5] used the method of deformation analysis to study the behaviour of sand waves. This method is based on the adjustment theory, with estimation of parameters of interest, and on the testing theory, with different hypothesis testings [5], [16], [17]. After [12] and [12], other studies like [13] and [6] focused on modeling sand waves by applying deformation analysis and also introduced the concept of risks related to seafloor dynamics. However, there is still no applicable approach of defining and quantifying the risk of the seafloor to change significantly using a probabilistic approach.

Prioritization of areas is also based on how busy they are. Densely used areas are of greater risk when significant seafloor changes remain undetected. The risk of the seafloor to change significantly entails additional risks such as; navigational risks related to the grounding probability of ships or other maritime application risks related to the probability that pipelines or cables get exposed or damaged. Conse-

quently, a risk-based analysis, regarding the evolution of the seafloor, is required for optimizing and validating the current resurveying frequencies as well as for avoiding additional risk related to maritime applications. According to Dijk, and the study [11], a validation of the current survey policy was implemented by using the existing knowledge about sand waves and assessing the navigational risks by estimating the grounding probability and grounding danger.

It becomes clear that there is still more research needed on getting reliable predictions of the seafloor as well as on quantifying the risks associated with seabed dynamics on the shallow and densely used locations of the NCS. Dealing with these issues in an automate and interactive way, is a prerequisite to define optimal resurvey frequencies for the NCS.

1.2 RESEARCH OBJECTIVES

The project goal is to optimize and validate the resurvey frequencies of NCS by perform a risk assessment regarding the relevance of seafloor changes. A risk assessment can contribute to the design of a successful bathymetric survey plan. Successful means cost-effective and with feasible resurvey frequencies that could be adapted automatically in the areas of NCS. Studies like [5], [13], [12] and [11], are focusing on the morphology of the sand waves and hence, the evolution of the seafloor of the NCS. The epochs of bathymetric surveys create suitable time series for applying deformation analysis in order to make reliable predictions. These predictions can be used as an input to assess the risk associated to the seafloor changes. Obtaining accurate and high quality predictions gives the potential for defining and quantifying risk-alarm levels related to seabed dynamics that influence the shipping sector as well as the pipeline and cable industry. Hence the following research question is formulated:

"How to determine the risk of the seafloor to change significantly, and what is the impact on the survey frequencies in critical areas of the NCS?"

A probabilistic approach regarding risk assessment, can aid the decision-making of the NLHO regarding the optimization of the resurvey frequencies of NCS, which is the general objective of this project. For answering the research question, a list of sub-questions should be answered :

1. *What is the accuracy and quality of the depth prediction values, that resulted from deformation analysis?*
2. *How can risk be defined using a probabilistic approach?*
3. *Which is an appropriate measure of risk severity?*
4. *How to go from nodal to regional analysis?*
5. *How to validate the current survey frequencies with the results of the risk assessment?*

This study will, firstly, focus on evaluating the accuracy and quality of the depth predictions, derived from the deformation analysis and, secondly, on suggesting a quantified risk assessment regarding seabed dynamics. This is a data-driven approach, based on the vertical dynamics of the NCS, rather than, a physical modelling approach, which is useful to study the morphology of the sand waves and how they migrate through time. Predictions of depth values are more interesting regarding their order of magnitude, rather than being used to physically interpret the characteristics of the sand waves. The principle followed is that significant depth changes, which are possible to occur in short time scale should be detected and raise a risk-alarm level, which differentiates for each maritime application. Seafloor changes should be expressed and estimated in a probabilistic way to assess the risks introduced and, finally, validate the assigned resurvey frequencies.

1.3 THESIS OUTLINE

Before dealing with the research question some necessary background knowledge related to bathymetry and map monitoring techniques, seabed dynamics and the current survey policy of NCS, as well as, a probabilistic definition of risk is given in Chapter 2. The selection of study areas and the formulation of the time series is explained in detail in Chapter 3. Chapter 4 includes the methodology that was used to answer the sub-objectives and the main research objective of this study. In Chapter 5 the results are presented and discussed, in order to make final conclusions and recommendations for future related research in Chapter 6. A flow chart is designed to depict the structure of the thesis and is presented in Figure 1.2.

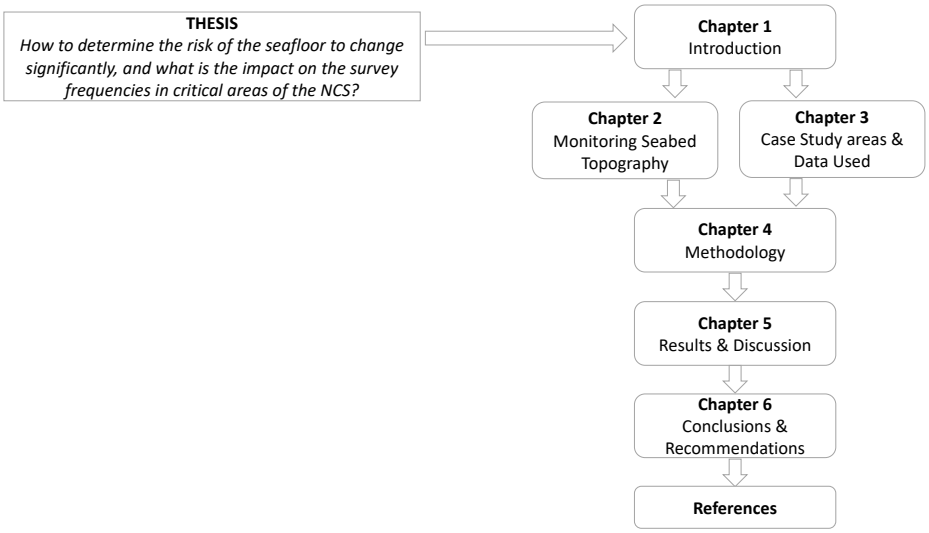


Figure 1.2: Flow chart representing the structure of the thesis and a detailed division of the chapters.

2 MONITORING SEABED TOPOGRAPHY

After introducing the topic of this research, a review of the available related literature is necessary to be presented and is done in this chapter. The concept of bathymetry is defined and the measurement process is described, in the first section. The current survey policy of the NCS is described in detail in the next section. Especially, the reason that Dutch seafloor is very dynamic is discussed; the different forms of seabed dynamics are presented in a separate section, as well as their special characteristics according to past research. Finally, in the last section, an introduction to the concept of risk is presented and the different approaches to determine it.

2.1 BATHYMETRY: CONCEPT & TECHNIQUES

According to National Oceanic and Atmospheric Administration (NOAA) bathymetry is *"the study of the 'beds' or 'floors' of water bodies, including the ocean, rivers, streams, and lakes"* [18] . The term is actually used to describe the submarine terrain or seabed topography. It is the measurement of the vertical distance, h, from the water bed to the free water surface (Figure 2.1). The reference datum used for this project is the MSL and that is why the depths get negative values. The products derived from acquiring and processing bathymetric data are bathymetric maps and nautical charts. These maps follow the same concept with the Three-Dimensional (3D) topographic maps that represent the land terrain. The submarine terrain is characterized by mountains, valleys, plains, and other sea floor features. Measured underwater depth values are usually visualized as contour lines, known as *isobaths*. Over the years various bathymetric techniques have been developed for measuring the seafloor depths as accurately as possible. Until now, this proved to be a very challenging process ([19], [18]).

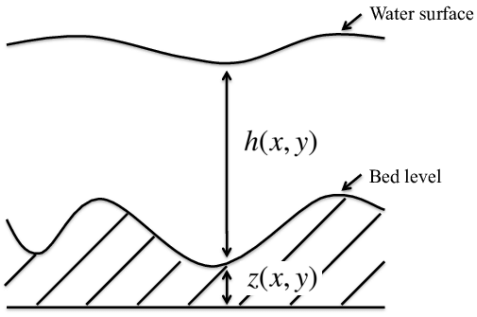


Figure 2.1: Bathymetry overview: Height, h, as the vertical distance of the bed level to the water surface, [2].

Sounding is considered to be the most accurate way of measuring depths in the deep ocean. For this purpose, sonar mapping systems are used and they can be divided into three categories:

1. SBES
2. MBES
3. Side Scan Systems (SSS)

The basic principle of sonar mapping systems is that the depth value is derived from the travel time of the transmitted acoustic waves. Depth measurement is done using the following equipment:

1. Transmitter
2. Transducer
3. Receiver
4. Recorder

The transmitter produces the acoustic waves. The transducer transforms this to the acoustic signal and receives the echo and transforms it to electrical pulse. Receiver processes the signal and sends it to the recorder which calculates the travel of the signal as the time difference between signal's transmission and reflection. This information is then stored and is actually the depth value [20], [13] (Figure 2.2).

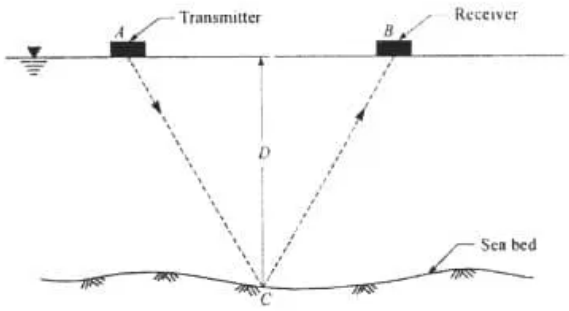


Figure 2.2: Sonar mapping system's main principle: A transmitter sends acoustic waves that reach the sea bed and they are reflected back to the receiver. This process results to the observed depth values, using as reference the water surface, which are finally adjusted to the final chart depth values, [3].

For high quality depth measurements, it is a prerequisite to specify accurately the horizontal location of the point/area measured. For this purpose, special instruments such as GPS/DGPS or triangulation technique in the previous years are used [21]. The acoustic waves are transmitted from the water surface to the bottom of the ocean. The travel time is recorded as the sound wave travels from free water surface to the submarine terrain and back. By determining the sound velocity, the depth value can be estimated [3].

To derive the final depth values, used for nautical charting, a pre-process is necessary. The observed depth should be, firstly, adjusted depending on a reference datum, the so-called *MSL*. Additionally, depth measurement should be corrected from specific systematic biases that affect the measurement, such as; the effect of tides and the errors propagating from the sound velocity profile as the wave travels underwater.(Figure 2.3)

Sounding techniques have some major drawbacks that should be addressed. Although MBES can achieve an accuracy of 0.05m, their spatial coverage is very limited and they suffer from logistical constraints due to the high costs needed for their operation. Another issue that arises, is that they are designed for measuring deep water depths and hence, close to the coast the measurement process gets more complicated. However, the need for monitoring and mapping coastal areas is increasing rapidly, as they are very dynamic (experience intense short- and long-term changes) and they are of great economic value. Remote sensing techniques and the use of satellites, for bathymetric purposes to the coastal areas, show great results but they are still only used as a complementary tool [22].

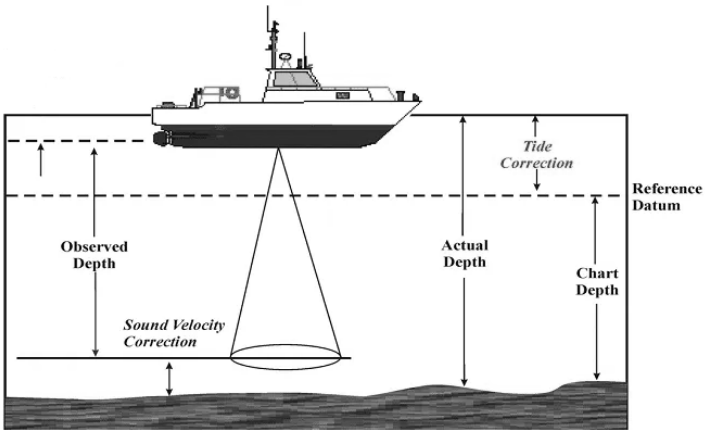


Figure 2.3: Overview of the corrections that need to be applied to the measurement depth values, as a part of the pre-process to derive the final chart depth values, [3]

.

Hydrographic surveys, worldwide, are held according to the framework of the IHO. IHO is an intergovernmental organisation responsible for maintaining the greatest possible uniformity of nautical charts and for publishing and updating international standards for hydrographic data [23]. The countries members should follow these standards when performing the regional hydrogrpahic survey planning. However, as every country has special characteristics and problems to deal with, regarding bathymetry, it is important to develop its own framework and use IHO's standards complementary. For example, North Sea is a typical example of how further research needs to be done on designing a cost-effective and reliable survey plan that could be applied automatically. Its special characteristics; it is a very shallow, dynamic and busy sea with many offshore activities taking place, could cause various risks.

2.2 SURVEY POLICY OF THE NETHERLANDS CONTINENTAL SHELF

In the Netherlands, the hydrographic surveys of the deep water depths of NCS are scheduled and implemented by the NLHO and for surveying the nearshore coastal areas, RWS is responsible. Two survey vessels of NLHO are used for bathymetric surveys under specific policy design.

The operation of the vessels is very expensive and mapping the NCS needs a lot of time as its area is very large. A plan that divides the areas of NCS to different categories is scheduled and published frequently, in order to avoid over-measuring areas or missing other critical regions. Hence, decision-making should be done on when or how frequently to resurvey which areas. This problem is dealt in a regional level, as IHO standards do not cover issues related to the specification of the resurvey frequencies or other additional criteria that could be used for this purpose.

The resurvey frequencies of all the areas in NCS are decided by NLHO. The current policy divides the NCS into (critical) zones of different frequencies, based on the following five factors:[5]

1. Minimum depth
2. Draught (difference between seafloor depth and ship's keel
3. Shipping intensity

4. Human interventions
5. Sea floor dynamics

As it is expected, shallow areas (depths $< 40m$) are automatically associated with the highest resurvey frequencies in the planning. Areas with very small depths are the most interesting as they could cause shipping risks but at the same time they usually are suitable candidate sites for various submarine applications such as the establishment of cable/pipeline networks or offshore wind turbines. Hence, the minimum depth is used as the first factor for prioritizing the areas.

Draught is known as the difference between the depth and the Under Keel Clearance (UKC). Draught values should be frequently updated and maintained (mostly done with dredging techniques) for critical areas of NCS. These are areas with high shipping intensity and concurrent shallow depths, which are mostly located close to the coast. The requirement of deep draught values could cause major problems associated with navigational risks and the probability of a ship going aground. For studying this factor knowledge regarding shipping intensity and the technical characteristics of the ships are necessary.

Shipping intensity is a very important factor that strongly influences the decision on when to resurvey an area. This is because, areas with high shipping intensity are widely used, which as a fact increases the risks caused by unexpected-undetected seafloor depth changes. Therefore, in these areas more unpleasant events could happen if not surveyed frequently and hence they are highly prioritized.

Human intervention in the dynamic evolution of the submarine terrain has an indirect effect that could cause long-term depth changes in the areas close to these inventions and that is why it should be taken into account when deciding upon resurvey frequencies [5]. For example, when maintaining the depth of an area by dredging it is logical to expect that the seafloor will 'try' to go back to its initial state and probably with faster rhythm. Hence, the natural processes are interrupted and especially with intense maritime activities and large offshore projects. These include dredging, sand mining, land reclamation or placing objects in the seafloor such as cables, pipelines or wind farms[5]. At the same time, large offshore projects like Maasvlakte II and Zandmotor project disturb the sediment processes[24]. It is clear, that this factor becomes more significant as the recent years seas are densely used and oceans resources highly exploited.

Finally, seafloor dynamics formulate the main factor that causes direct changes in depth values through the years. Therefore, modeling and predicting the evolution of the seabed dynamics is a key process for optimising the resurvey plan. However, this is a very challenging process and that is why various studies have been done focusing on this factor. These studies made a contribution in modelling the seabed dynamics, as natural processes and especially, tidal sand waves; their morphology [5], [11], their migration rate , [25] and, in general, their vertical and nodal dynamics [12], [6] and [26]. Despite the research done, it still remains a challenging process to assign resurvey frequencies, that could be adapted automatically, in areas with high levels of seabed dynamics being widely present [13].

The current prioritization of the areas in the NCS is presented in Figure 2.4, which was provided from NLHO and was published on 2017. For example, the most critical areas should be resurveyed once every 1 to 2 years. However, the survey plan and hence the prioritization, is based more on empirical and intuitive manners. A further scientific justification, that should include knowledge of seafloor dynamics, is needed for resurveying successfully the North Sea, which means that it should not be under-/over-surveyed as it was explained before.

Apart from this, a more reliable survey plan is needed, as the current resurvey frequencies are not always satisfied over the past years. Compared to the neighbouring countries which we can assume that they have the same bathymetry, the assigned resurvey frequencies for NCS are relatively high and difficult to achieve. According to Huizenga and [13], this causes delays in terms of total Hydrographic

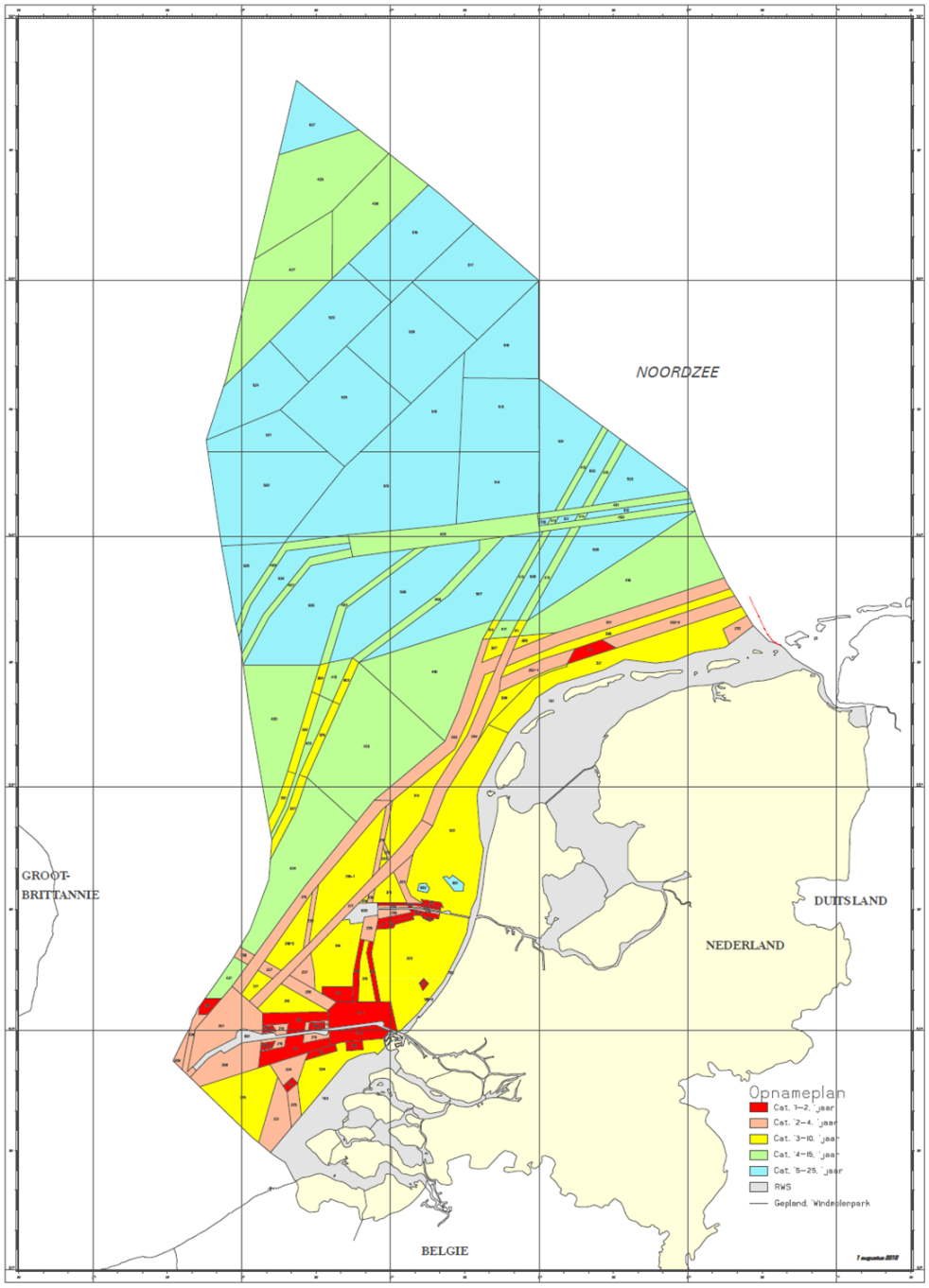


Figure 2.4: Map of 2017 with the current resurvey policy: Different colored areas of the NCS, depending on their resurvey frequency/category (provided after contact with NLHO). Areas with red and pink color are assumed to be the most critical areas of the NCS.

Days per year, which is translated to a backlog of around 800 days on the initial plan and it increases each year. Hence, there are many gaps in the data collection, for which estimations of depth values are done. Unsuccessful survey plan and gaps in the time series entail not only additional risks, but also more working hours.

An example of different survey plans, can be seen in Figures 2.5 and 2.6, which show the survey plans of 2003 and 2007 respectively. When comparing the plans with each other, and especially when comparing with the plan of 2017 (Figure 2.4), there is an obvious change in the prioritization of areas with more and larger areas planned to be resurveyed more frequently.

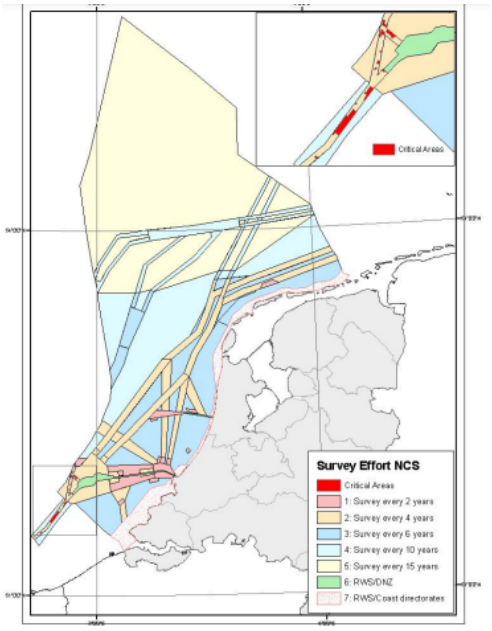


Figure 2.5: Resurvey frequencies of the NCS, 2003

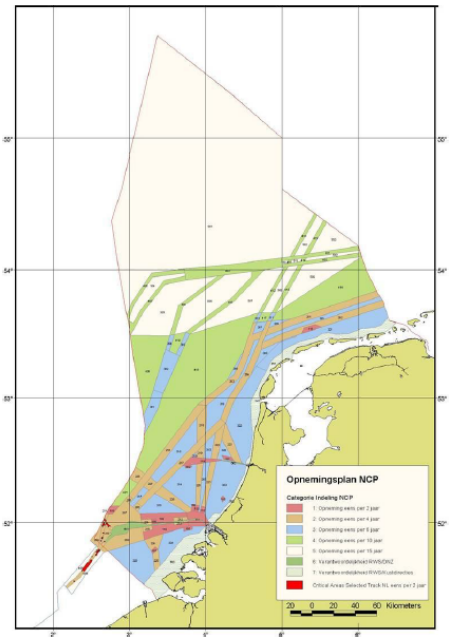


Figure 2.6: Resurvey frequencies of the NCS, 2007

There is an increasing interest on the seafloor dynamics, which resulted in the *"Seafloor Monitoring"* project. This project is based on applying statistical methods to perform a complete time series analysis using as input bathymetric data. This study can be considered as a part of this project and its scope is to schedule more efficiently which parts of the North Sea should be resurveyed and when [13]. For a better understanding of North Sea's bathymetry, the sea bed dynamics taking place will be further analysed in the following section.

2.3 SEABED DYNAMICS

North Sea is known for its highly dense used territory, as it hosts some of the largest ports in the world, such as the ports of Rotterdam, Antwerp, Hamburg, and London. Its seafloor is shallow and sandy and various dynamic processes are taking place. The average depth is estimated to be around 90 meters [27] and close to the south, in the NCS, it is, even more, shallower [13], Figure 2.7.

The seafloor of North Sea is not only influenced by physical processes, but also by large-scale human interventions and other maritime activities. However, the focus here is given on the physical processes that can be described by various seabed forms of different shapes and sizes. The main characteristics of the observed seabed dynamics are presented in Table 2.1. Five different categories of seabed dynamics have been detected in the North Sea: ripples, mega-ripples, tidal sand waves, long bed waves and tidal sand banks. It is possible that they occur at the same time in the same area. This creates a complicated and dynamic seafloor topography [13].

The most interesting bed form are the tidal sand waves as their wavelength, height and migration rate are relatively large and hence, they influence significantly the depth values and the changes they cause on the seafloor, happen rather quickly. Therefore, areas that sand waves are present are assumed to be very dynamic and it is important to monitor them [5]. An example of the seabed topography of an area that sand waves are present, can be seen in figure 2.8.

Especially, in the NCS, it has been observed, that tidal sand waves are the most common and widely present feature. They can transform shallow areas of the NCS

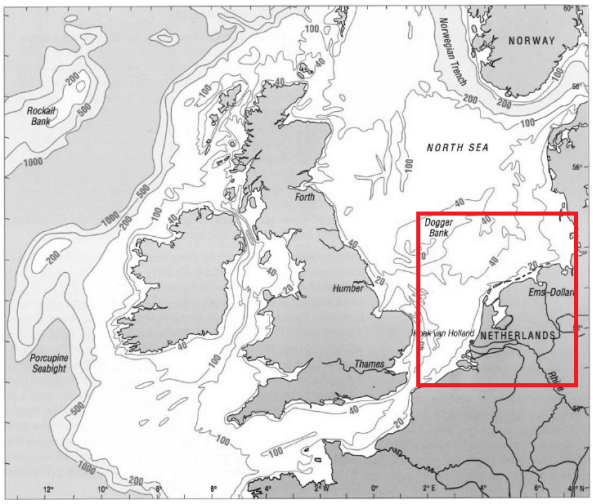


Figure 2.7: Overview of the North Sea and the neighboring countries. Isobaths are drawn that present the depth values below MSL. Most parts of the NCS (area indicated with the red rectangle) are very shallow (below 40 meters), [4].

Table 2.1: Rhythmic features and their main characteristics [5].

Sea bedforms and their main characteristics			
Type	Wavelength	Max. Height	Migration Rate
Ripples	1m	0.01m	1m/hour
Megaripples	10m	0.1m	1m/day
Tidal sand waves	500m	5m	10m/year
Long bed waves	1.5km	5m	unknown
Tidal sand banks	6km	10m	1m/year

to be even more shallow. Due to their large wavelength values, they cover huge parts of the sandy oceans. At the same time, they migrate fast over time and could cause significant depth changes [28]. Their migration is mostly caused by residual currents and due to tides effect [29]. Various studies focused on estimating sand wave's migration rate and different values have been derived. For example, Dijk [11] estimated a migration rate in the magnitude of several decimeters per year. In the study [25], the migration rate of sand waves has been estimated separately for the crests and troughs and is around 3 m/yr for the crests and 1.5 m/year for the troughs. Finally, Dorst [5] argues that the estimated migration rate of sand waves for the areas of West of Rotterdam and West of IJmuiden is up to 7.5 m/yr with a 95% confidence interval that depends on the regularity of the pattern, which is close to the one presented in Table 2.1 which was derived by Knaapen [28].

Due to the proven significance of this phenomenon, there has been an attempt to model the morphological characteristics of the sand waves that are present in the NCS and to detect and quantify the vertical changes that they cause in the seafloor. Various models (linear and non-linear) have been used in order to best describe either their morphological characteristics [5] or the vertical dynamics that they are responsible for, in areas that are widely present [6], [26]. By processing bathymetric data, it is easy to distinguish them as they are characterized by crests and troughs, like the surface water waves. According to Lindenbergh [12], average changes in depth are in the order of 1 dm/yr, while differences up to a few meters may occur in a time interval of ten years.

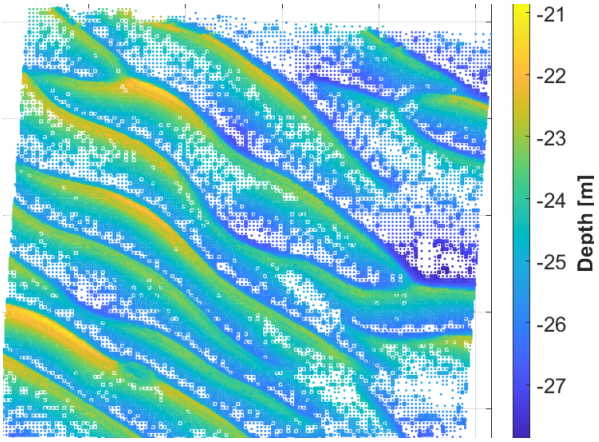


Figure 2.8: General look of the submarine terrain of NF subarea (West IJmuiden study area). The presence of sand waves in the area is obvious; the crests and the troughs can be distinguished.

2.4 RISK DEFINITION & PROBABILISTIC APPROACH

The concept of risk is not only familiar to the engineering world but almost to all important decisions that someone should take in its life. According to [30], two definitions are given to risk and these are:

1. A situation involving exposure to danger
2. The possibility that something unpleasant or unwelcome will happen

The first definition points out the concept of the consequences that are associated with a certain risk level, while the second definition is mostly related to the possibility or probability that an event will happen. According to [30], a common definition for risk is the following:

"Risk is the probability of an undesired event multiplied by the consequences"

This definition considers risk as an expected value and its unit depends on the units of probability and consequences. In general, the probability of an event is expressed as the probability that this event will happen depending on time. One of the often-used equations to describe risk is the following 2.4, which will be also used for the purpose of this project:

Risk = Consequences * Probabilities

For quantifying risk it is necessary to relate and estimate the probabilities and the consequences (Figure 2.9). This process is mostly known as risk analysis or risk assessment or quantitative risk analysis (QRA) [30].

Successful decision-making is based on a comprehensive and precise risk-analysis. Risk analysis can serve various aspects of a problem. Apart from its main scope, to quantify and calculate risk-alarm levels, it can also be used as a tool to examine the capabilities of a system or a plan. The concept of risk analysis can also help users not familiar with a topic to further understand it. It is very powerful for the communication between different disciplines of engineers and scientists. Finally, risk analysis can be used for optimizing the design of a system or a plan. The latter aspect is often-used, as if it gets a step further it can ensure high quality of a system designed.

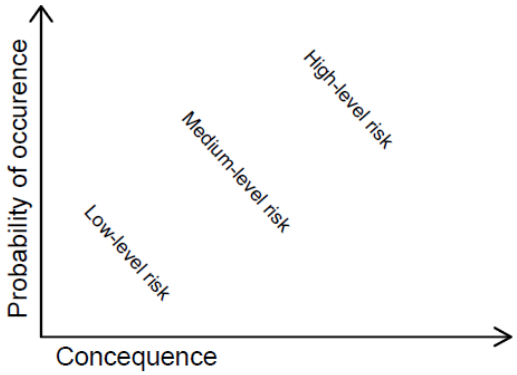


Figure 2.9: Graph that shows the relationship between risk's impact (consequences) and their probability to occur. As both consequence and probability get high values the risk alarm levels become higher.

The concept of Quantitative Risk Assessment (QRA) or Probabilistic Risk Assessment Probabilistic Risk Assessment (PRA) has been used to supply decision-making in many engineering problems and one of its main characteristic is that, it is based on the second risk definition that was addressed before, which points out the probability that an event will take place. According to [31], when performing PRA two main factors should be taken into account:

1. Severity of the consequences
2. The likelihood that these consequences will happen

Estimating the above variables and giving them numerical values makes the risk assessment more powerful and therefore, it is called quantitative. This means that the consequences take a specific value as well as the likelihood that these consequences will happen. On the other hand, a qualitative risk assessment is based on the descriptive form of the variables [31].

For example, assuming that it is required to assess the risk of someone to lose its life on a car accident in 2019. Using QRA or PRA we need to know the exact number of how many people died in 2019 due to a car accident (consequence) and the number of car accidents in 2019 or the probability that a number of people will die due to a car accident that happened in 2019 (likelihood of the consequence to happen). Using qualitative approach we can define the number of car accidents or people died with a scale of 'low', 'medium' and 'high'.

In bathymetric applications, QRA is needed to validate the chosen resurvey frequencies of the NCS. It can be used to supply the need of a scientific approach in the decision-making upon the resurvey frequencies and to optimize the design of the bathymetric plan. However, the problem gets more complicated because the resurvey policy is based on the five factors, described before in Section 2.2, and each one of them introduces different risk alarm levels, consequences and probabilities.

Going a step further, even if we apply QRA for each factor separately another issue arises, which is how to combine the derived risk alarm levels to conclude to a final value of risk. The aforementioned challenges will be further analysed in Chapter 4, where the chosen methodology to answer the research question of this study is applied and justified.

3 CASE STUDY AREAS & DATA USED

The following chapter introduces the selected areas of interest, the reason that they were chosen and their special characteristics. A detailed presentation of the input bathymetric data is done and the chosen spatial scale (nodal) of the analysis is justified. The time-series formulation is explained and finally, discussed in the last section of the chapter to point out some of the special characteristics of the nodal time-series such as their variant length and the precision of the observations.

3.1 STUDY AREAS

The study areas were chosen depending on the research objectives and the goal of this project. Hence, the most interesting areas for this project is assumed to be the ones that help answer the research question and investigate different aspects of it. These areas are: the highly prioritized areas, according to the current survey policy of the NCS, and the areas that seabed dynamics and especially sand waves are present. Consequently, the following rule was applied to select between areas of interest. Possible study areas include:

1. Shallow areas (depth values below 40m)
2. Areas where sand waves are present
3. Location of shipping lanes (indicating high shipping intensities, increased grounding probability)
4. Areas influenced by human intervention (like dredging)

This methodology has been followed from Huizenga in [13], and almost applied the same way by Dorst in [5]. Hence, it is reasonable some specific areas, that fulfill the above, to be repeatedly selected among others and studied over the years. This is additional reason for choosing these study areas as the knowledge, already obtained could help to evaluate the results of a new approach developed or to support further research.

Another important factor is to choose areas that will supply in the interpretation of the results. This study will focus on shallow areas with sand waves and which are busy shipping routes. Hence, the selected study areas are two in order to compare the results and these are; the regions of **West of IJmuiden (IJgeul area)** and **West of Rotterdam**. These areas can be seen in Figure 3.1 and are prioritized as critical in the NCS current plan, with a survey frequency, once every 2 years. Research has been done before in these areas by vanDijk in [11], by Dorst in [5] and by [6]. Different methodologies have been applied and from different scopes; morphological modelling or data-driven approaches. The special characteristics of each area will be discussed in the following Subsections: 3.1.1 for West IJmuiden and 3.1.2 for West Rotterdam.

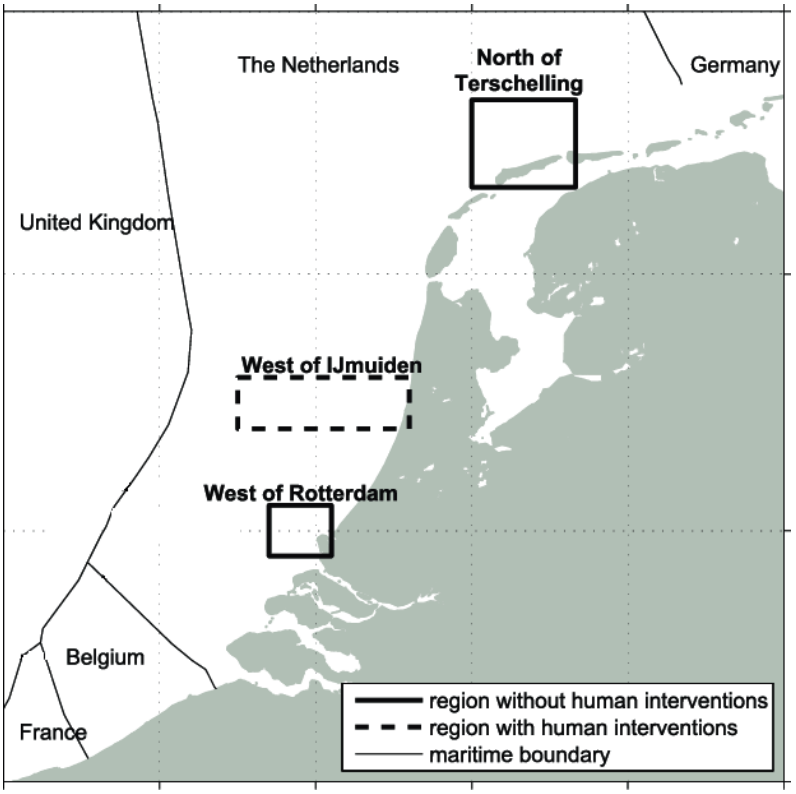
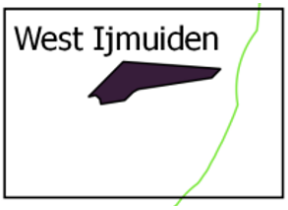


Figure 3.1: The locations of the study areas of West Ijmuiden and West of Rotterdam are noted, [5].

3.1.1 West of IJmuiden (IJgeul Area)



The area of IJgeul is a smaller part of the region West of IJmuiden. This area is shallow and sand waves are present. It is assumed to be a very dynamic area with complicated sand wave patterns that are associated with various risks. Therefore, its depth should be maintained. Human intervention is needed to avoid shallow areas to become even more shallow. Dredging is implemented to keep a fixed depth value but there is no further information regarding this [5].

This study area, for the purpose of this research, was subdivided in 30 subareas for processing. It is an area estimated to be around $150x150km$, as every subarea is almost $5x5km$ large and its mean depth is almost $24.5m$. It consists of around 2 million nodes, as a nodal time-series analysis was performed, which will be further explained in the following Subsection 3.2.1. The total number of surveys available is 10 and they start from 1989 and end up to 2013.

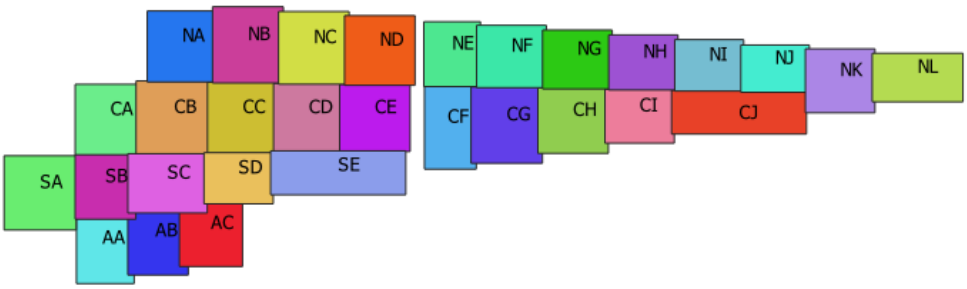


Figure 3.2: Overview of the complete area of West IJmuiden and its subdivision to 30 subareas.

3.1.2 West of Rotterdam

This study area is located close to the west part of the port of Rotterdam. It is relatively shallow, sand waves occur and migrate in its seafloor and it is also an important shipping route. The so-called area of West Rotterdam is located very close to the coast. RWS is responsible for the dredging of the area and the neighboring areas in order to maintain the depth in a safe scale. According to Dorst ([5]), dredging activities took place in an area very close to our study area in 1991, 1996, 2002, and 2004.

3.2 DATA INPUT FOR TIME-SERIES ANALYSIS

Hydrogpraphic surveys of the NCS were done through the years, following the annual survey plan and the defined survey frequencies, by operating SBES and MBES sonar systems. The data acquired from the NLHO and RWS have been archived and published by Deltares. Hence, for each location of the NCS a time-series of depth values and their associated uncertainties through time, is obtained. The time-series analysis is implemented on a zero dimension (nodal analysis) [6]. The choice of the spatial analysis and the process of the time-series formulation will be further described in the next Subsections 3.2.1 and 3.2.2, respectively.

3.2.1 Nodal Analysis

The spatial scale of the analysis done, should be specified. Different approaches have been adopted in bathymetric studies [11], [13], [5], [32] and [6], and according to Pluymaekers [33], they can be summarized to the following:

1. **Zero dimensional/Nodal analysis (0D):** The centre of every grid cell is represented by a node (point).
2. **One dimensional/Line analysis (1D):** A grid line is used as the spatial unit for the analysis.
3. **Two dimensional/Regional/Planar analysis (2D):** A plane/complete grid is analysed.

The choice of the spatial approach is mainly result-based. 1D and 2D will involve more parameters of interest, as it is expected to need more complicated functional models to describe them. 1D is more suitable for studying and defining the morphology of seabed dynamics. However, both of them use uniform area shape, as when applied it is already assumed that sand waves are present in the area and in fact they function better when the sand waves follow regular patterns [13].

The 0D analysis was used in the aforementioned studies. It is relevant, that it is also successfully used using InSAR techniques [34],[35] in other applications like studying land subsidence. 0D is more suitable for analysing vertical dynamics without making any assumptions regarding the natural processes taking place in the study area. Hence, the mathematical models used could be more simple and less parameters of interest are necessary. For the purpose of this study, a nodal analysis was chosen and is based on the study of Toodesh and Verhagen, [6], which was used as a background study.

3.2.2 Time-series formulation

The input data consist of time-series of depth values of nodes. Each node has fixed position (x, y) and a depth value (d) for each survey moment. For each new available survey, a depth value is added in the time-series for that node and the length of time-series increases, which will increase the confidence in the time-series analysis and will supply in more reliable predictions. The measured depth values are accompanied by their variances, that describe the precision of each measurement. Hence, the time-series formulated include the following information for each node:

1. Date of each available survey moment, (T_i).
2. Horizontal coordinates (x, y) fixed for all survey moments.
3. Measured depth values at every survey moment, D_i.
4. Variances of depth values at every survey moment, σ_i.

Figure 3.3 shows the time-series for three example nodes. Their observed depth values for each available survey can be seen with scatter points. The lines connecting the scatter points are only drawn for a better understanding of the depth trend of each node and they do not indicate any mathematical model. Additionally, for one of the three nodes (yellow scatter points) the variances of the observed depths are included in the form of error-bars (vertical lines to the scatter points). It can be seen visually that the variances of the last five surveys (after 2005) are much more accurate than the previous ones and this is because of the higher accuracy of MBES compared to the SBES sonar system.

During the time-series formulation, it became clear that the nodes have a different number of available observed depth values (length of time-series). For example, the node represented with blue scatter points in Figure 3.3 can prove this, as the first available measurement for this node is at 1996, which means that the length of its time-series is smaller by 2 observations. The issue raised will be further discussed in Subsection 3.3 as it was both a challenge and an opportunity for the analysis.

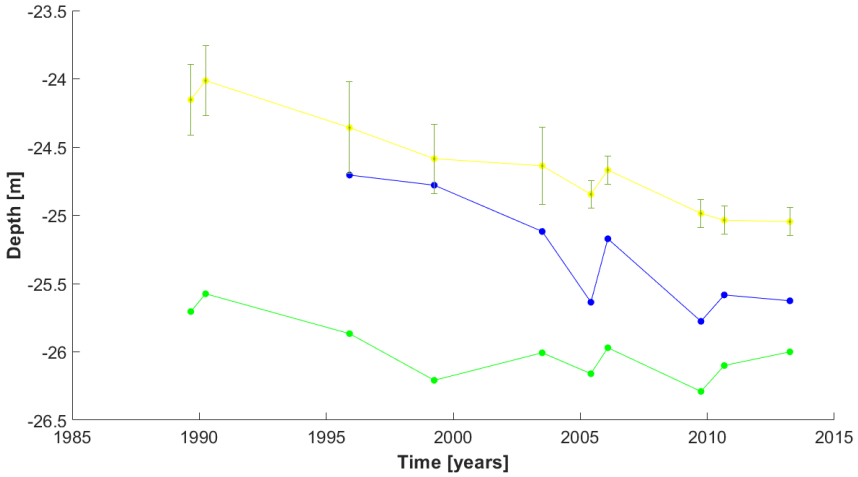


Figure 3.3: Example of the nodal depth time-series of three different nodes. The scatter points indicate the observations of each node and the connected line is drawn only to distinguish the time-series of each node (they do not represent functional models).

3.3 DISCUSSION

Regarding the selection of the study areas and time-series formulation a list of issues needs further attention and discussion. These issues need further investigation, which can be used as recommendations for future related research. Basically, they involve; the large amount of data and large size of the NCS compared to the desired level of detail and the adjustment of the analysis in the variant length of the time-series used.

Performing a time-series analysis for all the areas in NCS in order to conclude to their resurvey frequencies is a very time-consuming process. For this reason, every process done in this study is automated to serve fast application to as many parts of the NCS as possible. The amount of data is huge and the computational complexity of the analysis is also high. Therefore, the various regions of NCS should be separated into smaller subareas, which compared to the full size of NCS are very small. This as a fact, makes it difficult to get a feeling of the level of desired detail in relation to the decision-making. For example, in the study area of IJmuiden, which was divided in 30 subareas, the size of each subarea should always be compared to the distance that MBES measures and the size of areas which are decided to be surveyed by the authorities. This is a prerequisite in order to derive logical and reliable conclusions regarding whether a subarea will be resurveyed or not.

Considering the length of the time-series of the nodes, for this analysis, the nodes that have more than three observations were used. If this number increases and the condition becomes stricter, the number of nodes used will decrease significantly which is not desirable as the level of detail decreases and the predictions could be biased. Not only the length of the time-series but also the exact time of the available surveys play a role, as it is necessary to be included during the prediction process.

Variable length of time-series for the nodes, means gaps in the time-series of some nodes. This is due to the fact that: either some nodes have not been surveyed at all, compared to their neighboring ones, or the measurements were treated as outliers and excluded from the final products of a survey. The gaps in the time-series of a large amount of neighboring nodes can also be explained by the fact that the survey plan is not always successful and some areas are missed. This creates a large source of uncertainty and makes the process complicated. A future recommendation could be to give weights to the nodes depending on their time-series length. For example, nodes with very limited number of observations should be weighted less, compared

to the ones that have full length time-series. On the other hand, using only nodes with full length time-series, entails loss of crucial detail.

Finally, from Figure 3.3 it becomes obvious that the surveys after 2005 that were done with MBES sonar system are of better accuracy. This should be taken into account when performing a time-series analysis and especially when this includes predictions. Hence, another recommendation could be to assign different weights to the observations of the last surveys in order to obtain more precise predictions.

4 METHODOLOGY

This chapter describes in detail the chosen methodology to answer each of the sub-question which by themselves lead to the answer of the main research question. Firstly, the pre-processing of the data is presented by the chosen quadtree approach to create a reference grid for the data. Secondly, the way that estimations were made from the time-series is discussed and the use of deformation analysis. The chosen functional models are reported and sketched. Next, it follows the predictions estimation and mainly the method used to define their accuracy, quality and reliability. The latter, is included in a separate section and the selected validation method is justified. Going a step further to answer the research question, in the next section focus is given in explaining the way that PRA is being carried out for this project. Finally, the last section includes the method used to go from the nodal to the regional analyis.

4.1 PRE-PROCESSING THE DATA

Data pre-processing and preparation is one of the most challenging processes in bathymetric projects. SBES and MBES data use different technologies and this influences the available data which are of different spatial resolution and uncertainty sources. The data should be combined in an optimal and friendly-user way [32].

It is relevant to select an appropriate type of grid to reference the data and interpolate them. Two different methodologies have been applied for gridding the data, considering the studies done before. There is one option, to use a regular uniform grid, for example a 5x5m grid cell applied by Dorst in [5] or to use a new introduced quadtree approach implemented in [32]. An example of the latter method can be seen in Figure 4.1.

Using a quadtree approach provides the user with a list of advantages and therefore the products of this griding option were used for the analysis done here. Its main concept is to generate a multi-resolution DEM of the seafloor. This very detailed DEM, focuses on critical areas, like the crests and troughs of the sand waves by creating smaller grid cells in these areas of high variability. Hence, the original data are exploited in an optimal way regarding their spatial resolution. For each location different level of information is represented (multi-resolution) depending on users decision, which could be for example a proof regarding the presence of sand waves in a specific area of the NCS. It also deals with the uncertainties of the measurements interactively with the user [32]. Hence, one of its main advantages is that it lets the user to take a location-based decision regarding DEM's desired level of detail. For the interpolation of SBES data Universal Kriging was used like in [5] and the horizontal coordinates of all nodes were projected using UTM (Zone 32N) and WGS84 ellipsoid.

4.2 ESTIMATION THEORY: DEFORMATION ANALYSIS

According to Lindenbergh [12], deformation analysis was firstly introduced in the seventies at the geodesy department of Delft University of Technology [36], [37] and combined with Delft testing method [16], [17], it was used to study land deforma-

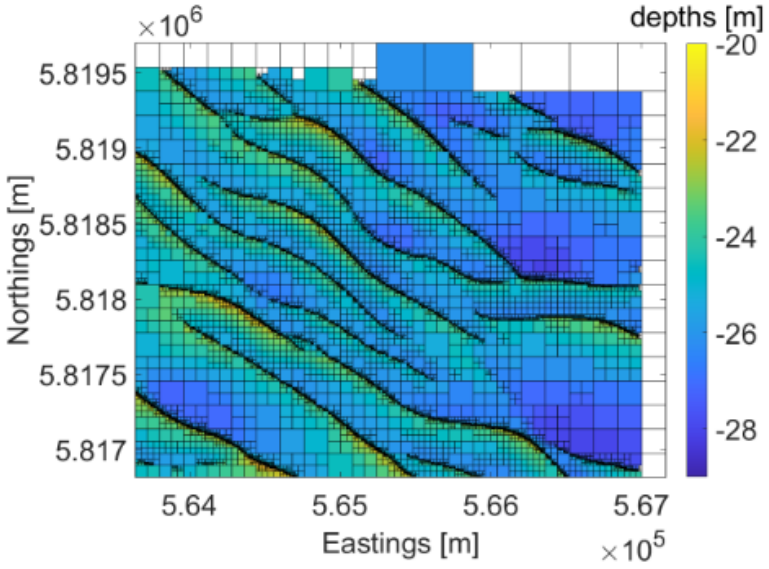


Figure 4.1: Quadtree output DEM, [6]

tion with data available from leveling measurements. Nowadays, with the advancements in surveying techniques, huge amounts of acquired data are available and it remains a challenge to best process these data in a sufficient and automated way. Deformation analysis can serve these demands and succeed in detailed and accurate results. It is applied in many bathymetric studies like [12], [5], [6], [11], [13] and it proved to be useful in modeling the morphology of sand waves and studying vertical dynamics on the sea floor.

A nodal deformation analysis is used to study the evolution of the seafloor in the chosen study areas and to predict the future depth values. Estimations were made using a least-squares approach, depending on the observations and the chosen functional models. Extending these estimations in the future, predictions were made. The prediction values were then evaluated and used as input data for the PRA, together with their associated uncertainties. Deformation analysis combined with estimation of their variances according to Dorst in [5], is a prerequisite to obtain high quality predictions that could serve a relevant risk assessment.

The main focus of this project, is firstly to evaluate the quality of predictions derived by the deformation analysis, [6] and then use them as an input for PRA. However, for evaluating the predictions, it is necessary to present the library of functional models used and an overview of the hypothesis testing approach used to select the most likely model that describe the bathymetric time-series of each node.

Stochastic model is used to describe the precision of the observations and is represented by the variance covariance matrix of the observations Q_{yy}. Functional models are used to mathematically describe the physical phenomenon taking place. They relate the observations $(y = \underline{d})$ with the unknown parameters of interest x. Hence, assuming m is the number of available observations for each node, this equals to m number of observation equations [38]:

$$\underline{d} = Ax + \underline{e} \tag{4.1}$$

where A is the design matrix and $\underline{e}$ are the residuals. Assuming, $E\{\underline{e}\} = 0$, the expected mean value of the residuals equals to zero, means that the observations are only influenced by random errors and they follow a normal distribution. Therefore:

$$E\{\underline{d}\} = Ax; \tag{4.2}$$

A set of functional (mathematical models) were used to estimate the depth values, depending on the vertical dynamics of the seafloor and these models are presented in Figure 4.2 and Table 4.1):

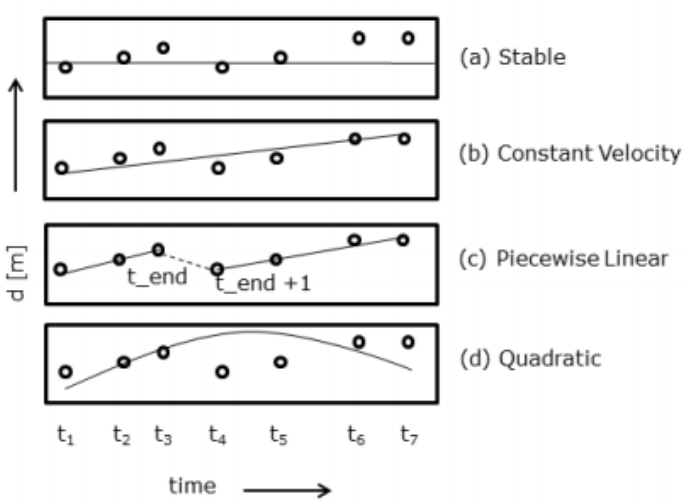


Figure 4.2: Sketch of the form of the functional models used in this analysis to derive estimations of depth values, [6]

Table 4.1: Mathematical models used and their main characteristics.

Model type	Parameters of interest	Observation equations
1. Stable	d_0	$y = d_0$
2. Constant velocity	d_0, v_0	$y = d_0 + v_0 t$
3. Piecewise linear	d_0, v_0, d_1, v_1	$y = d_0 + v_0 t$ $y = d_1 + v_1 t$
4. Quadratic	α, d_0, v_0	$y = \alpha t^2 + v_0 t + d_0$

From Table 4.1, it is observed that the number of parameters of interest for the various models ranges from 1 to 4. This indicates the level of complexity of each model. However, a very complicated model is not necessarily the best option. Regarding the parameters of interest used, d_0 is the depth of the first survey moment and v_0 is the respective slope, depending on this initial depth (d_0). The slope physically represents the vertical velocity of the depth value of a node, and is measured in [m/yr].

The stable model assumes a fixed depth value (the mean depth) for the node through the years. A constant velocity model, is one of the most often-used linear models and is very important for the purpose of this project. Among its advantages, which are related to the information it adds regarding vertical seabed dynamics, is its ability to detect upward or downward trends of the depth values of each node. Processing the direction according to which the seafloor changes can be a valuable

variable for the risk assessment and it is assumed to be a contribution of this study, although constant velocity model was also used by Dorst in [5]. The quadratic model has three parameters of interest the acceleration a, the initial depth d_0 and the slope v_0. Finally, the piecewise linear model is constructed to detect significant depth changes in the time-series and it is a new introduced mathematical model by Toodesh in [6]. It uses the initial depth (d_0) and the slope (v_0) and creates a straight line similar to the constant velocity model, until it detects a break-point (large change in depth) and uses this new location to create a new straight line with the same type of parameters of interest as before but taking the values of the new location (d_1) and (v_1). The latter model is valuable to detect critical parts of the sand waves like the crests and troughs, as these locations experience large depth changes in relatively small time intervals. Additionally, it could also be useful in detecting changes due to human intervention like dredging activities. It is worth mentioning, that other models can also be used to describe the behaviour of the sand waves. As the length of the time-series increases with new available surveys, the functional models are expected to improve.

In estimation theory, least-squares adjustment is an often-used technique to link the functional models with the stochastic model and derive estimations of the unknowns $\left(\widehat{d}\right)$ [38]. Depth observations are assumed to be unbiased which means that they are normally distributed. Additionally, the functional models used here are linear models. Hence, the requirements of linearity and unbiasedness are satisfied and therefore, Best Linear Unbiased Estimation (BLUE) was applied. BLUE general solution of the unknown parameters $(\widehat{x})$, adjusted observations $(\widehat{y})$ and residuals $(\widehat{e})$ can be derived by:

$$\widehat{x} = (A^T Q_{yy}^{-1} A)^{-1} A^T Q_{yy}^{-1} d \tag{4.3}$$

$$\widehat{d} = A\widehat{x} \tag{4.4}$$

$$\widehat{e} = d - \widehat{d} \tag{4.5}$$

This version of least-squares estimation ensures that the estimators $(\underline{\widehat{x}})$, also, follow a normal distribution and hence they are unbiased. However, regarding the reliability of the functional models, BLUE is not an adequate tool for a safe conclusion. Therefore, hypothesis testing was implemented to specify the most likely model to describe the time-series of each node [38].

To determine the most appropriate characterisation of the nodal time-series, the hypothesis testing procedure applied, was developed based on the study [6]. For statistical hypothesis testing different hypotheses were tested with respect to the functional models. In general, hypothesis testing theory consists of the null hypothesis H_0, which is tested over the alternative hypothesis H_a.

$$H_0 : E\{\underline{d}\} = Ax \tag{4.6}$$

and

$$H_a : E\{\underline{d}\} = Ax + C_d \nabla \tag{4.7}$$

The approach adopted here differs, in a way that all hypothesis are treated equally during testing procedure [6]. This means that there is no assumption that one hypothesis is more likely than the other [6]. The hypothesis constructed are based on the functional models and are the following:

1. H_1: stable model

2. H_2: constant velocity model
3. H_3:quadratic model
4. H_4: piecewise linear model, where, $H_{4_{1...m}}$ represents different piecewise linear models depending on the length of the time-series.

Piece-wise linear model involves 4 parameters of interest which actually represent a break-point, where there is a significant change in depth. This means, that there are many possible forms of this model for the same time-series of a node and, hence, the hypothesis testing also includes these different formations of piecewise linear model as different hypothesis. The number of hypotheses increase depending on the number of observations (m) that are available for each node $(H_{4_{1...m}})$. The hypothesis testing used a combination of the Overal Model Test (OMT) and the General Likelihood Ratio Test (GLRT) to select the most likely model for each node [6].

4.2.1 Precision of estimators

Before performing the predictions, it is important to explain how uncertainties propagate from observations to estimations. This process will be discussed in this subsection in detail and the respective derivations will be presented.

During the time-series formulation, for each depth measurement, its variance is also provided. The stochastic model consists of the precision of the observations and mathematically this is expressed by the variance covariance matrix of the observations (Q_{yy}). After BLUE was applied the precision of the estimators $(\widehat{x})$, $(\widehat{y})$, $(\widehat{e})$ was calculated.

The precision of estimated unknowns is given by the covariance matrix $Q_{\widehat{x}\widehat{x}}$ and it is a measure of the quality of the estimated parameters:

$$Q_{\widehat{x}\widehat{x}} = (A^T Q_{yy}^{-1} A)^{-1} \tag{4.8}$$

$Q_{\widehat{x}\widehat{x}}$ is dependent on the precision of the observations, Q_{yy}, and the design matrix, A. The precision of the estimated observations $Q_{\widehat{y}\widehat{y}}$ is:

$$Q_{\widehat{y}\widehat{y}} = A Q_{\widehat{x}\widehat{x}} A^T = A(A^T Q_{yy}^{-1} A)^{-1} A^T \tag{4.9}$$

Combining the above, the precision of the estimated residuals $Q_{\widehat{e}\widehat{e}}$ is derived by:

$$Q_{\widehat{e}\widehat{e}} = Q_{yy} - Q_{\widehat{y}\widehat{y}} = Q_{yy} - A(A^T Q_{yy}^{-1} A)^{-1} A^T \tag{4.10}$$

$Q_{\widehat{e}\widehat{e}}$ is used for the model selection during testing and therefore it is of great importance. To summarize, with BLUE the best estimators among all linear unbiased estimators can be computed and best equals to the most precise ones. The measure of precision is the variance covariance matrix of each estimated quantity that was presented above.

4.3 PREDICTIONS

The predictions of the depth values are made using the parameters of interest of the most likely selected model for each node, resulting in either the null assumption H_0 or one of the alternative hypothesis H_a. Nodes for which H_0 is accepted, satisfy the constant velocity model whereas nodes for which H_a is accepted, satisfy one of the piecewise linear models. For the latter model the predictions are made using the second segment of the piecewise linear model and hence depending only on d_1 and v_1.

The length of the time-series of the observed depths for each node influences the process of making predictions, as explained in Chapter 2. A series of prediction moments, T_p is generated and it depends on the last survey moment, T_{last} of each node, which means that despite its fixed length (=9), the exact prediction time is different for every node.

$$T_p = [T_{last}, T_{last}+1, T_{last}+2, T_{last}+3, T_{last}+4, T_{last}+5, T_{last}+10, T_{last}+15, T_{last}+30] \tag{4.11}$$

$$T_p = \left[T_{p_0,} T_{p_1,} T_{p_2,} T_{p_3,} T_{p_4,} T_{p_5,} T_{p_6,} T_{p_7,} T_{p_8,}\right] \tag{4.12}$$

The numbers in T_p represent the years of prediction after the last survey moment. For example, the last survey moment (T_{last}) should be equal to the first predicted moment $T_{p_0,}$ and $T_{p_3,}$ equals to the prediction made 2 years after the last survey moment. Hence, the length of the prediction time-series (T_p) is fixed and equal to 9, but the prediction moments could be different for each node.

Consequently, depending on the selected model and the corresponding hypothesis, the estimations of the predicted depth values $(\widehat{d_p})$ are computed by:

$$\widehat{d_p} = A_p \widehat{x} \tag{4.13}$$

The design matrix, A_p, for constant velocity model is constructed from the estimated parameters of interest $\widehat{x} = \widehat{d_0}, \widehat{v_0}$ and for the piece-wise linear model it uses only the $\widehat{x} = \widehat{d_1}, \widehat{v_1}$. This is a powerful method suggested for generating predictions as it deals with the problem of gaps in the time-series and it provides the user with more precise predictions as the variances are also estimated depending on the last survey moment, T_{last}, of each node.

4.3.1 Accuracy of Predictions

The precision of the predictions is calculated by the covariance law and expressed by their variance covariance matrix $Q_{\widehat{d_p}\widehat{d_p}}$, which is calculated as follows:

$$Q_{\widehat{d_p}\widehat{d_p}} = A_p Q_{\widehat{x}\widehat{x}} A_p^T \tag{4.14}$$

The standard deviation of each prediction value, σ_{d_p} is given by the square root of the diagonal line of the variance covariance matrix and is measured in meters.

$$\sigma_{d_p} = \sqrt{diag(Q_{\widehat{d_p}\widehat{d_p}})} \tag{4.15}$$

The interpretation of the precision of the predictions is expressed by the concept of CI. It is known that for normally distributed random variables, it is possible to define probabilistic intervals around their expectation value. The probability that $\widehat{d_p}$ deviates less than $2\sigma_{\widehat{d_p}}$ from its expectation value $E(\widehat{d_p})$ is given by:

$$P(\underline{d_p} - E(\underline{\widehat{d_p}}) \in [-2\sigma_{\underline{\widehat{d_p}}}, 2\sigma_{\underline{\widehat{d_p}}}]) = P(\left|\underline{d_p} - E(\underline{\widehat{d_p}})\right| < 2\sigma_{\underline{\widehat{d_p}}}) = 0.954 \tag{4.16}$$

which equals to the following statement:

"There is 95.4% chance that $\underline{\widehat{d_p}}$ *is in-between the interval of* $[E(\underline{d_p}) - 2\sigma_{\underline{\widehat{d_p}}}, E(\underline{\widehat{d_p}}) + 2\sigma_{\underline{\widehat{d_p}}}]$*"*

The concept of CI is mostly applied for the estimations of predictions $\underline{\widehat{d_p}}$. For the normally distributed random variables, it applies that $\widehat{d_p} = E(\underline{\widehat{d_p}})$ (unbiasedness) and the CI are, now, not applied for the expectation value but for the realisation of the estimation. This means:

$$P(d_p - 2\sigma_{\underline{\widehat{d_p}}} < \underline{\widehat{d_p}} < \underline{d_p} + 2\sigma_{\underline{\widehat{d_p}}}) = P(\underline{d_p} \in [\underline{\widehat{d_p}} - 2\sigma_{\underline{\widehat{d_p}}}, \underline{\widehat{d_p}} + 2\sigma_{\underline{\widehat{d_p}}}]) = 0.954 \quad (4.17)$$

The new CI used is:

$$S = [\underline{\widehat{d_p}} - 2\sigma_{\underline{\widehat{d_p}}}, \underline{\widehat{d_p}} + 2\sigma_{\underline{\widehat{d_p}}}] \quad (4.18)$$

and it is commonly used to express the chance or the probability that the true unknown value of the variable ($\underline{d_p}$ in this example) is inside S interval. For example, here, there is a low probability that d_p is outside the limits of S. The left part of S is usually called *lower* CB and the right part *upper* CB. The mostly-used CI are 95% and 99%.

The CI of 95% and 99% were calculated for the predicted depth values and they correspond to:

$$[\underline{\widehat{d_p}} - 1.96\sigma_{\underline{\widehat{d_p}}}, \underline{\widehat{d_p}} + 1.96\sigma_{\underline{\widehat{d_p}}}] \quad (4.19)$$

$$[\underline{\widehat{d_p}} - 2.58\sigma_{\underline{\widehat{d_p}}}, \underline{\widehat{d_p}} + 2.58\sigma_{\underline{\widehat{d_p}}}] \quad (4.20)$$

The variances of predictions and the fitted CB were combined with the estimations and the original observations to get a complete visual verification of the method used here. These will be presented and interpreted in Chapter 5.

4.4 VALIDATION OF PREDICTIONS

Defining the precision and quality of the prediction values does not add information regarding how powerful are the models used, considering their ability to predict. For this reason, a validation process is necessary to test the performance of the chosen models in practice and hence, the reliability of predictions. During the validation process, the model is tested over an arbitrary dataset that was not used during model generation process, to check how well it performs [39].

CV technique is a well-known technique when performing machine learning. It is mostly used by the form of *k-fold* CV or *leave-one-out* CV method. In general, when applying a validation algorithm, the available dataset is divided in [40]:

1. The training dataset
2. The validation dataset
3. The test dataset

The training dataset is used to derive the model used. The validation dataset is used to get a first impression of the validation results, but its main use is to actually tune model parameters. The test dataset is used to test model's performance. Test dataset should be independent and totally separated from the training dataset [39], [41].

For time-series validation there are some limitations and the CV methods of k-fold CV and leave-one-out should be adjusted. These limitations have mainly to do with the time dependency, as it is impossible to have in validation or test dataset, observations that are before or in-between the time interval in which the training data are referred to. In other words, the model's ability to predict makes sense to be tested only in future time [39].

K-fold CV does not seem to be suitable as the dataset gets shuffled and divided into small parts (k-folds) to get tested and this can not be applied for time-series [40]. On the other hand, leave-one-out CV can be applied but with a time-dependent condition; only for the last survey moments. In general, assuming that m number of observations, when performing leave-one-out method, in the first iteration the first observation will be excluded from the training dataset to be used as test dataset. In the second iteration, the second observation will be excluded, and used as test dataset, but the rest observations (including the first one) will be used as a training dataset [40]. For this analysis, it was decided to apply only two iterations due to the time-dependency and the small length of the obsevred depth time-series (around 9). For these two iterations the datasets were divided as follows:

Table 4.2: Datasets for leave-one-out CV method for each iteration

Iterations	Observations	Training dataset	Test dataset
1	m	$m-1 = m - T_{last}$	$1 = T_{last}$
2	m	$m-2 = m - T_{last}$-$T_{(last-1)}$	$2 = T_{last}$+$T_{(last-1)}$

Figure 4.3: Representation of the dataset division for each iteration for an example time-series of 9 observations.

The validation error was then calculated by computing the absolute difference between the original measurement (test dataset) and the corresponding prediction value. This was done for every nodal time-series and the mean value of the validation errors is the performance of the validation (CV_e).

$$CV_e = \frac{1}{n}\sum_{n=1}^{n} |d_{last} - d_p| \tag{4.21}$$

The same process was followed in the second iteration.The absolute value is used here, as we are mainly interested to test how close the predictions are to the real measurements. As a step further, the algorithm of CV could be adjusted to test the model's prediction performance with respect to the direction that seafloor moves.

Another element that adds complexity to the problem is the gaps in the time-series data. The CV_e is computed referring to different survey moments for every node that does not have full length time-series. However, what remains fixed is the time-step (for example 2,3,4 years after the last survey moment) and this works as the time unit in this approach. Hence, the results are more precise, as the conclusions are made upon years of prediction and not fixed future survey moments and all the available nodes were used, without lowering the spatial resolution.

4.5 PROBABILISTIC RISK ASSESSMENT OF DEPTH CHANGES

Current resurvey policy is based on the 5 factors of the minimum depth, draught, shipping intensity, human interventions and sea floor dynamics as discussed in Chapter 2. These factors are highly variable depending on the area of NCS. Hence, they are associated to different weights and risk alarm-levels. Quantifying the risk-alarm levels will help on optimizing the resurvey frequencies of the NCS in a scientifically proven way. Especially regarding seabed dynamics, the NLHO implemented a risk assessment study, [42], which ensured that information regarding the seafloor dynamics can be valuable for the resurvey policy of the NCS, something which was also stated in [5].

The derived predictions of seafloor depth values together with their variances are used as input data to perform PRA, considering the risks associated to significant depth changes. If the predictions are precise and reliable this risk assessment, can be used by the authorities as a decision-making tool to further inform or optimize the resurvey plan. However, it could also stand alone, as it is fully-automated, and different sets of prediction values can be used as input data. The main objective of this study is to *to determine the risk of the seafloor to change significantly* . Focusing on this and on PRA principles (analysed in Chapter 2, Section 2.4) the following should be specified and estimated:

1. The consequences related to this risk
2. The probability of occurrence of these consequences

4.5.1 Risk consequences: Definition & Quantification

Assigning numerical values to the consequences, requires to firstly, define them. As this is a nodal analysis, the consequence resulted from the risk of the seafloor to change significantly can be defined as:

The number of nodes for which quantities related to their vertical dynamics are exceeding a certain threshold value.

The input dataset for this risk analysis consists of prediction values and their variances and the depth values are not adequate by themselves to derive conclusions regarding seafloor evolution. Therefore, we need to define quantities related to the evolution of each node in the vertical dimension in order to quantify the consequences. These quantities are referred to as *risk indicators*, because they give a sense of the risk's consequences and hence, risk severity. The risk indicators used are:

1. The Relative Depth Change RDC
2. The Rate of ChangeRoC

They are estimated per node and they are used to compare the depth evolution in different locations of the NCS.

4.5.2 Relative Depth Change (RDC)

This indicator was decided because the change in depth, by itself, can add information regarding the seafloor terrain evolution but, in order to make it valuable for decision-making, it should be combined with the absolute depth value of the node, to indicate the significance of the change. This becomes clear, taking into account the first factor of prioritizing areas in NCS; the minimum depth value. For example, assuming a node with a depth of 70m to experience a change depth of 2m after one year and a node with a depth of 20m to experience a change in depth of 0.5m in the same time interval. These changes are associated to different risk alarm-levels as a half meter depth change in one year for a shallow area is expected to be more significant than a larger depth change in a deeper area. Hence, RDC indicates the change in depth of a node as a percentage of its absolute depth value.

RDC is defined by the depth difference of a node between its depth value at a selected prediction moment, d_{p_i}, and the depth value at the last survey moment, d_{last}, divided by the absolute depth value at the last survey moment, d_{last}, and multiplied by 100%.

$$RDC_n = \left(\frac{d_{p_i} - d_{last}}{d_{last}} \right) 100\% \tag{4.22}$$

The variable n refers to the number of nodes and i to the prediction moments used to compute RDC. RDC is a dimensionless number, a ratio and it is expressed as a percentage [%]. This risk indicator results from the mathematical relationship of two normally distributed variables; depth value at the last survey moment, d_{last} and depth value at a selected prediction moment, d_{p_i}. However, as it is a ratio and hence, linearity is not satisfied, it can not be treated as a normally distributed variable. This information is relevant because probability estimation will be applied only for normally distributed variables in this analysis. It would be possible to estimate the probability of RDC by defining its distribution or by applying a simulation, however it was decided to use only RoC for probability estimation. Therefore the usage of this indicator, RDC, regarding PRA is limited.

On the other hand, RDC is an informative variable for interpreting vertical seabed dynamics in various areas of the NCS. Knowing only this single value of RDC for each node, we can compare the depth change in different areas of the NCS. That is why, it was estimated and used in this project, but not for the probability estimation. To overcome this issue another risk indicator was defined which is very similar to RDC but it can be used for probability estimation and this is RoC.

4.5.3 Rate of Change (RoC)

RoC is the second indicator used to quantify the nodal vertical dynamics and it was also used by Dorst in [5]. It follows the same concept of RDC and it is defined; as the depth difference of a node between its depth value at a selected prediction moment, d_{p_i}, and the depth value at the last survey moment, d_{last}, divided by the corresponding time interval, $T_{p_i} - T_{last}$ (time difference between the selected prediction moment and the last survey moment).

$$RoC_n = \frac{d_{p_i} - d_{last}}{T_{p_i} - T_{last}} \tag{4.23}$$

RoC is equivalent to the slope value v, which is a parameter of interest for both functional models used in the hypothesis testing procedure. It shows how fast (rate) the depth of a node is changing with respect to time. Physically it is the vertical velocity of the node and is measured in $\left[{}^{m}/_{yr} \right]$. This risk indicator is also

a fraction, but the denominator refers to time which is a deterministic variable and not a random one. Therefore, it can be used in the probability estimation and, PRA implemented in this project is based on this indicator.

4.5.4 Direction of change

Both the risk indicators of RDC and RoC can be used for studying the *direction* of change in the depth time-series of a node. In this study, the change in depth is estimated and not the difference in depth, which means that the signs of the depth values were kept during the calculations and they can be used for valuable conclusions. By keeping the sign of numerator we conclude that:

1. Positive sign (+): node gets shallower
2. Negative sign (-): node gets deeper

This information can be used as a step further and quantify risks related to maritime activities, such as shipping or establishment of pipeline and cable network. Areas that are getting deeper are more interesting for the industry of pipeline and cable networking, as the risks of exposure increase. At the same time, information regarding areas that get shallower is important for the shipping sector, in these areas there is a higher probability of a ship to run aground. These will be further analysed in Chapter 5.

4.5.5 Uncertainties of Risk Indicators

Given the variances of the last survey moment, $\sigma^2_{d_{last}}$, and the variances of the prediction moments, $\sigma^2_{d_{p_i}}$, the uncertainties of the risk indicators were computed by applying Taylor series expansion. The computations for each indicator are presented here separately. Hence, for RDC it is obtained:

$$RDC = \frac{d_{p_i} - d_{last}}{d_{last}} = \frac{d_{p_i}}{d_{last}} - 1$$

$$\sigma^2_{RDC} = \left(\frac{\partial RDC}{\partial d_{p_i}}\right)^2 \sigma^2_{d_{p_i}} + \left(\frac{\partial RDC}{\partial d_{last}}\right)^2 \sigma^2_{d_{last}} \tag{4.24}$$

$$\sigma^2_{RDC} = \left(-\frac{1}{d_{last}}\right)^2 \sigma^2_{d_{p_i}} + \left(\frac{d_{p_i}}{{d_{last}}^2}\right)^2 \sigma^2_{d_{last}} \tag{4.25}$$

$$\sigma^2_{RDC} = \sqrt{\left(-\frac{1}{d_{last}}\right)^2 \sigma^2_{d_{p_i}} + \left(\frac{d_{p_i}}{{d_{last}}^2}\right)^2 \sigma^2_{d_{last}}} \tag{4.26}$$

For RoC the following computations were done:

$$RoC = \frac{d_{p_i} - d_{last}}{T_{p_i} - T_{last}} = \frac{d_{p_i}}{dT}$$

$$\sigma^2_{RoC} = \left(\frac{\partial RoC}{\partial d_{p_i}}\right)^2 \sigma^2_{d_{p_i}} + \left(\frac{\partial RoC}{\partial d_{last}}\right)^2 \sigma^2_{d_{last}} \tag{4.27}$$

$$\sigma^2_{RoC} = \left(-\frac{1}{dT}\right)^2 \sigma^2_{d_{last}} + \left(\frac{d_{p_i}}{dT}\right)^2 \sigma^2_{d_{last}} \tag{4.28}$$

$$\sigma^2_{RoC} = \sqrt{\left(-\frac{1}{dT}\right)^2 \sigma^2_{d_{last}} + \left(\frac{d_{p_i}}{dT}\right)^2 \sigma^2_{d_{last}}} \tag{4.29}$$

The uncertainties of the risk indicator RoC were computed also using a level of significance, α, of 0.05 and 0.01 for the respective 95% and 99% CI.

4.5.6 Estimation of the probability of the consequences to occur.

After defining the consequence of the seafloor to change significantly and computing the risk indicators, as the measure of risk severity, it is necessary to estimate the probability of this consequence to occur. This is defined in general as:

- The probability that the quantities related to the vertical dynamics of a node are exceeding a threshold value.

Taking into account that only RoC will be used for probability estimation, the following probability should be computed:

- The probability that the RoC of a node is exceeding a threshold value.

$$P\left([RoC_n > \varepsilon_k]\right) = (1-\alpha)\cdot 100\% \tag{4.30}$$

The estimation of probability is done interactively with the user. The threshold values (ε_k) indicate how extreme a depth change could be and α is the level of significance applied. The NLHO provided a set of threshold values that are used as extreme values and show how significant a change can be, and they are presented in Chapter 5. However, these thresholds are derived based more on empirical studies and hence they should be optimized or justified. In this analysis a set of probabilities and ε_k threshold values are computed iteratively and automatically, based on a probabilistic approach and considering the direction of change (positive or negative). Both the properties of automation and direction can add value to the results and contribute to the research currently done.

A set of varying ε_k values is defined based on the RoC values for each area, and the corresponding probabilities are estimated. The user can either decide a fixed ε value and estimate the respective probability, or inverse the problem and give the probability a fixed value and define the corresponding ε value. A flow chart included in Figure 4.4 represents the process in detail.

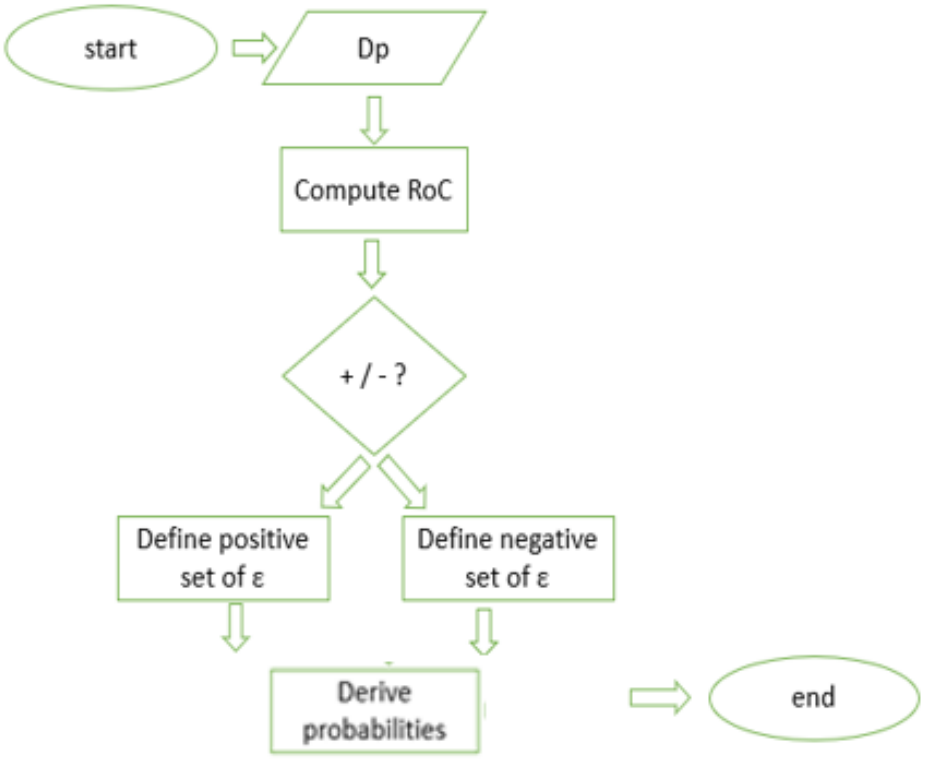


Figure 4.4: Flow chart that describes the way that the probability of RoC to exceed a certain threshold value is computed per node, considering the direction of change.

4.6 GOING FROM NODAL TO REGIONAL

The chosen nodal analysis has many advantages and its relevance is described in detail in Chapter 2. Performing PRA on a nodal scale is not adequate, unless it is accompanied by a generalisation from nodal to regional scale. The resurvey plan should be designed considering that survey vessels measure on a regional scale and not specific nodes. Finally, the conclusions should be made regarding the study areas as complete entities.

Estimators, predictions, risk indicators and probabilities were computed for each node in the chosen study areas. These nodal results were used to, firstly, visualise the seafloor terrain with bathymetric maps and also to create thematic maps of the study areas, considering the decision-making. Three different levels of detail were applied during map generation:

1. Nodes (points) representing each location of the study area
2. Polygons (tiles) representing the boundaries of the subareas
3. Uniform entity (either raster or vector) representing the complete study area

Especially, the second level of detail is relevant; the polygons (tiles). Decisions regarding the time of next survey can be taken for each subarea separately and automatically in a greater level of detail compared to the initial complete study areas. In general, the maps were geo-referenced and the same geodetic characteristics of the bathymetric data were used; the projection of UTM zone 32N and (EPSG:25832 - ETRS89).

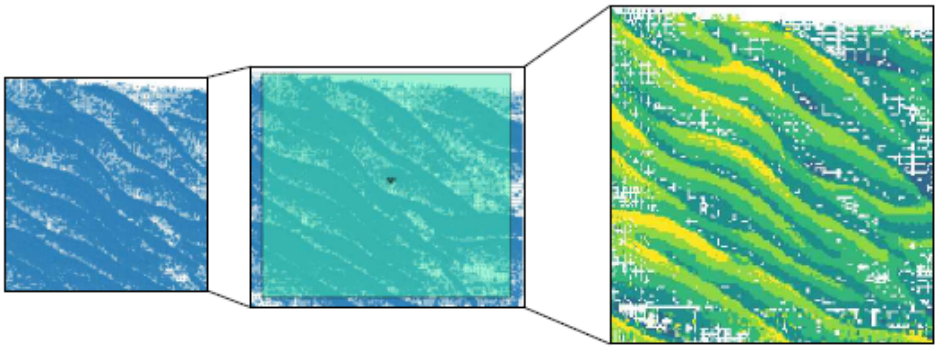


Figure 4.5: From nodal to regional; set of nodes (left), set of nodes and an overlaying polygon of a subarea (middle), raster representation of the subarea as part of the complete study area (right).

5 RESULTS AND DISCUSSION

The results of the analysis will be presented in the following sections. Each section includes graphs, maps and interpretation of results for each one of the chosen study areas; West IJmuiden and West Rotterdam. The content of the subsections, together with the results of each analysis step, will follow a symmetric visualisation process for both study areas. To validate the existing methodology, an adjustment of the results with the standards of the NLHO is also provided. The discussion of the results and the method applied follows in the final section of this chapter.

5.1 CASE STUDY I: WEST IJMUIDEN

The west region of IJmuiden is the first study area that will be discussed. Some of its main characteristics were presented in Chapter 3. Figure 5.1 is a bathymetric map of the complete study area at the last survey moment, which is for most of the nodes 2013.

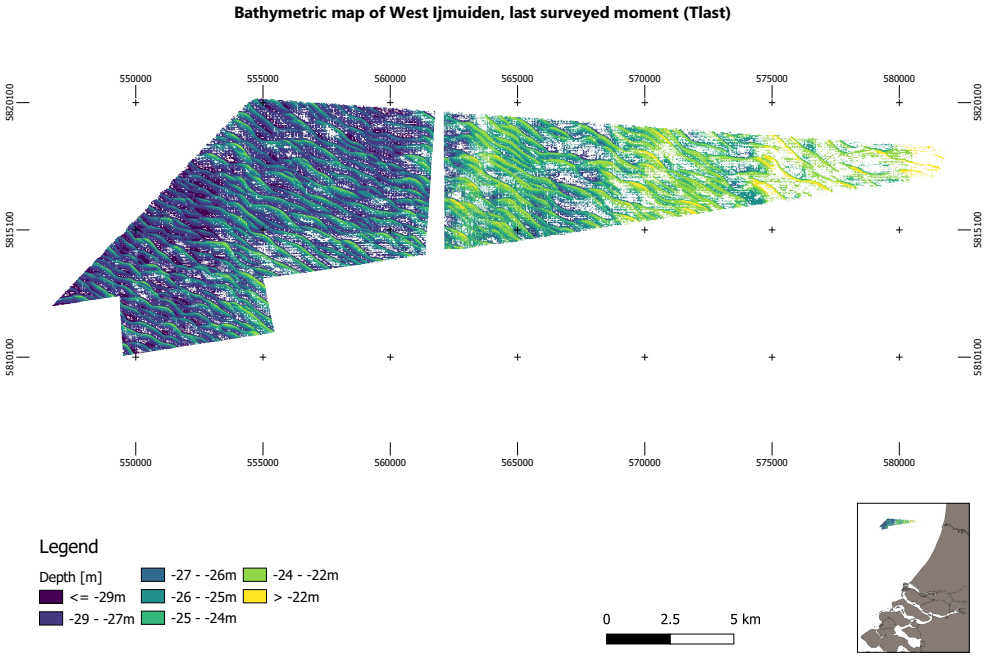


Figure 5.1: Bathymetric map of the complete area of West IJmuiden at the last surveyed moment. Depth varies from areas deeper than 30m (deep blue) to areas shallower than 22m (yellow).

However, as this study area was very large for a nodal analysis ($>$ 2 million nodes), it was divided into smaller subareas, which were analysed separately, and it was finally reunited, in order to produce maps, like the one in Figure 5.1. To have a compact representation of the results some of them will be presented in the scale of a subarea (NF) or even in the scale of a node of this subarea. Thus, it was

considered necessary to introduce some of the characteristics and statistics of this subarea.

5.1.1 Statistics of NF subarea

Before getting deeper in the results, it is important to get a sense of the data of NF subarea, as it will be used as an example area for the results. The size of NF area is around $2.5 \times 2.3km \simeq 5.7km^2$ and it consists of 80813 nodes. The presence of sand waves is obvious, as the crests and troughs can be seen in Figure 5.2, which is a bathymetric map of NF for the last survey moment (2013).

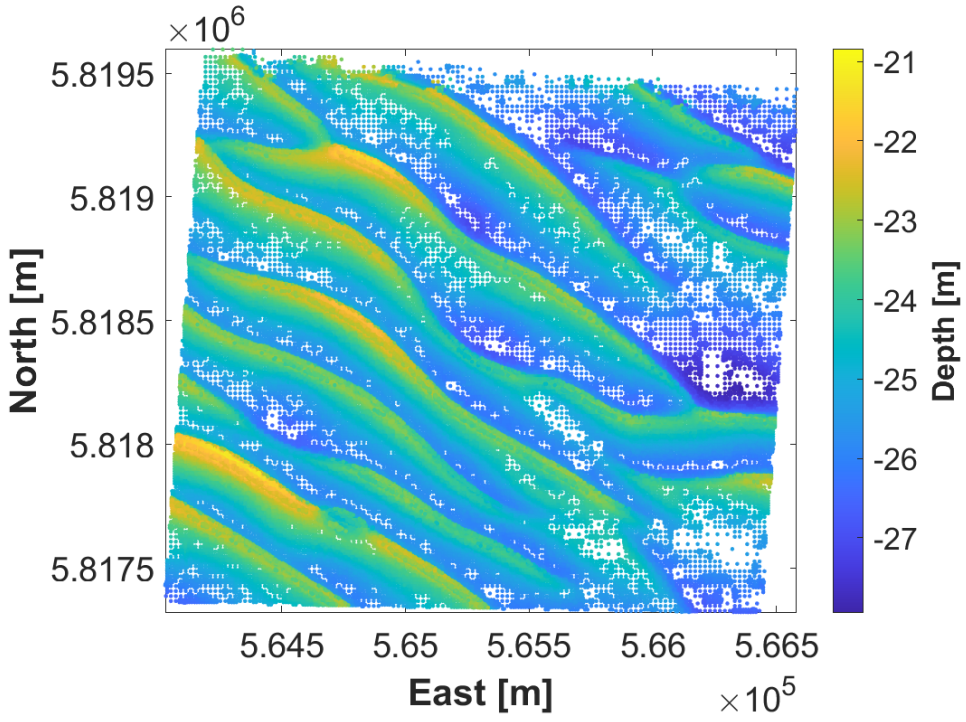


Figure 5.2: A view of the NF subarea and its depth at the last surveyed moment.

Statistics regarding the nodal timeseries of NF subarea are presented in Table 5.1. The statistics were calculated using only the full-length timeseries, in order to include all the available surveys. The results are representative as the 98% of the nodes of NF subarea have full-length time-series, which means that only 2% of the nodes have gaps in the time-series. Taking a more closer look, it can be determined that NF is a shallow area with a mean depth value of around 24.5m. The deepest parts measured to be around 28m and the shallowest around 21m. The precision of the measurements until 2005 varies between 15 to almost 30cm and after 2005 it drops significantly to 8cm. This is explained by the fact that after 2005 MBES sonar systems operated for surveying the NCS.

Table 5.1: Statistics of timeseries of NF subarea.

Survey	$d_{\max}$	$d_{\min}$	$\overline{d}$	$\overline{\sigma}$	Δd
1989	-27.609m	-21.413m	-24.765m	0.14m	
					0.135m
1990	-27.479m	-21.273m	-24.630m	0.26m	
					0.005m
1995	-27.919m	-21.430m	-24.636m	0.33m	
					0.046m
1999	-27.866m	-21.504m	-24.590m	0.25m	
					0.146m
2003	-27.894m	-21.256m	-24.443m	0.28m	
					-0.113m
2005	-27.810m	-20.840m	-24.556m	0.08m	
					0.098m
2006	-27.645m	-20.615m	-24.458m	0.08m	
					-0.233m
2009	-27.940m	-20.980m	-24.691m	0.08m	
					0.207m
2010	-27.817m	-20.900m	-24.484m	0.08m	
					0.002m
2013	-28.100m	-20.850m	-24.482m	0.08m	

5.1.2 Estimations, Predictions & Uncertainties

The method of deformation analysis, as explained before in Chapter 4, was applied and estimations of depth values were computed depending on different functional models. The results of the hypothesis testing included two different hypotheses, H_0 and H_a. When H_0 is accepted, the constant velocity model is selected as the most likely model to depict the nodal time-series. The parameters of interest for this model are the depth at the first survey moment (d_0) and the respective slope (v_0). Figure 5.3, shows the fit of the constant velocity model with the original observations, as well as, the predictions and their associated uncertainties. The quality of the predictions was defined by applying the CI of 95% and 99%. These CI were constructed using the variances of the predictions, which were computed using the variance covariance law (explained in Chapter 4).

From Figure 5.3, we firstly, obtain that the precision of the original observations increases after 2005. The estimations seem to fit visually well with the observations and as expected, there is a high correlation between the prediction time and the precision of the predicted values. As the prediction moment increases, the CB deviate more from the prediction trend. Taking this into account, the user can decide how many predictions can be trusted and thus which of them can be used depending on the desired maritime application. To get a better understanding of this, we can use Figure 5.3, to obtain an order of magnitude of predictions precision and quality. The depth at the last survey moment (original observation, 2013) is -24.55m and its prediction value at 2015, after 2 years, is -24.79m $\pm$ 0.4m with a 99% CI. The prediction value after 10 years (2023) is estimated to be -25.31m $\pm$ 0.71m with a 99% CI.

Almost the same rule applies for the other model used, the piecewise linear model. Figure 5.4 shows how this model fit looks like and the corresponding prediction values. This model is suitable to detect relatively large depth changes. The predictions are made depending only on the second segment of the fitted line. This entails higher uncertainties of the predictions done. Prediction trend is drawn based on less original observations. Additionally, the CB deviate more from the predicted

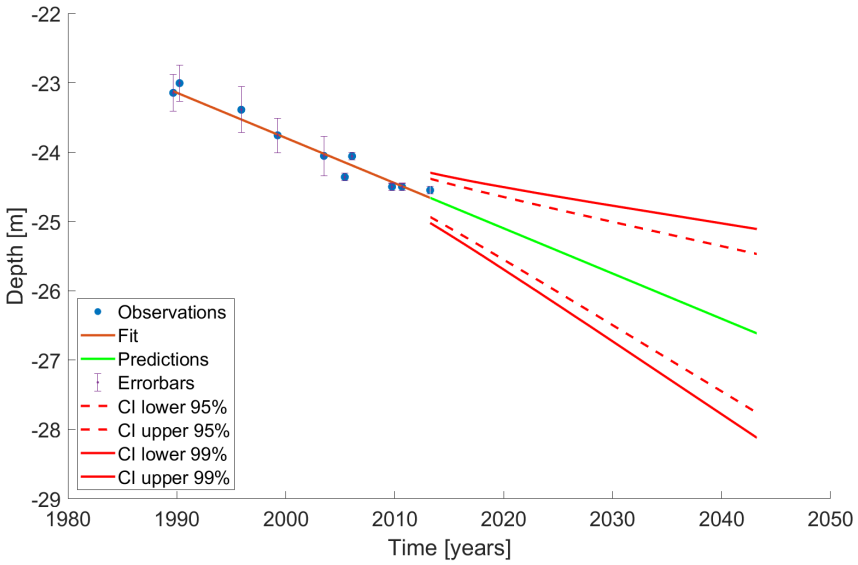


Figure 5.3: Example node for Ho accepted: Original observations (blue scatter points) with their precision (errorbars), estimations done using constant velocity model (red line, fit), prediction trend (green line) with the associated quality of the predictions illustrated by the CI of 95% and 99%.

line as the prediction time increases, but with a higher rate compared to the constant velocity model. As before indicative values of the precision and quality of the predictions are addressed. The depth at the last surveyed moment (original observation, 2013) is -23.67m and its prediction value at 2015, after 2 years, is -23.82m $\pm$ 0.5m with a 99% CI. The prediction value after 10 years (2023) is estimated to be -24.45m $\pm$ 1.7m with a 99% CI.

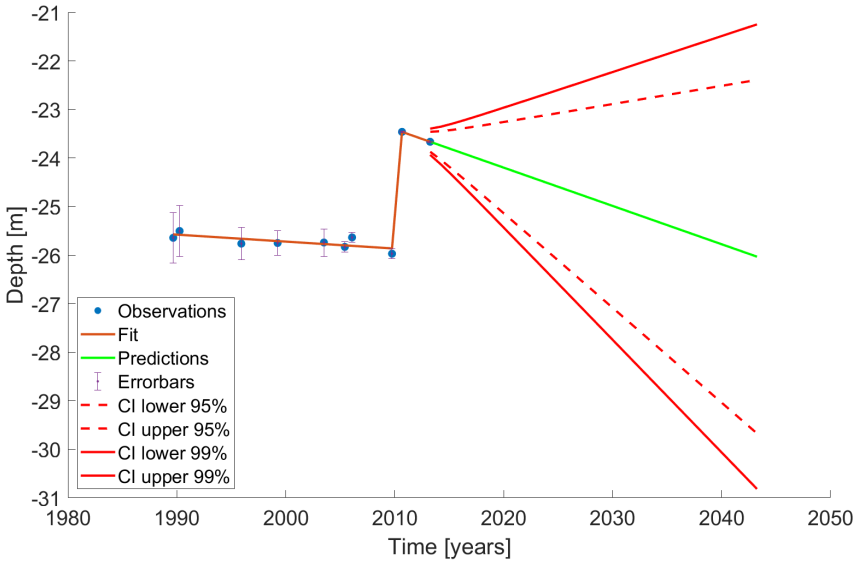


Figure 5.4: Example node for Ha accepted: Original observations (blue scatter points) with their precision (errorbars), estimations done using piecewise linear model (red line, fit), prediction trend (green line) with the associated quality of the predictions illustrated by the CI of 95% and 99%.

It is important to show the spatial locations of the nodes considering the chosen models. Therefore, a binary map shown in Figure 5.5 was generated. Blue nodes

represent the nodes for which H_0 was accepted and hence, the constant velocity model was chosen as the most likely model to describe their vertical dynamics. Red nodes represent H_a and the forms of piecewise linear model. From Figure 5.5, it becomes obvious that the crests are mostly represented with red nodes. This goes along with the large changes that piecewise model should be able to detect. Blue nodes are the 'in-between' parts of the sand waves, which we expect to experience smoother changes and this is ensured by the selection of constant velocity model. Finally, the percentage of nodes for which H_a is accepted, reach a 7% of all the nodes of NF and they can be assumed to be the most critical ones. The rest 93% represent the nodes for which H_0 is accepted.

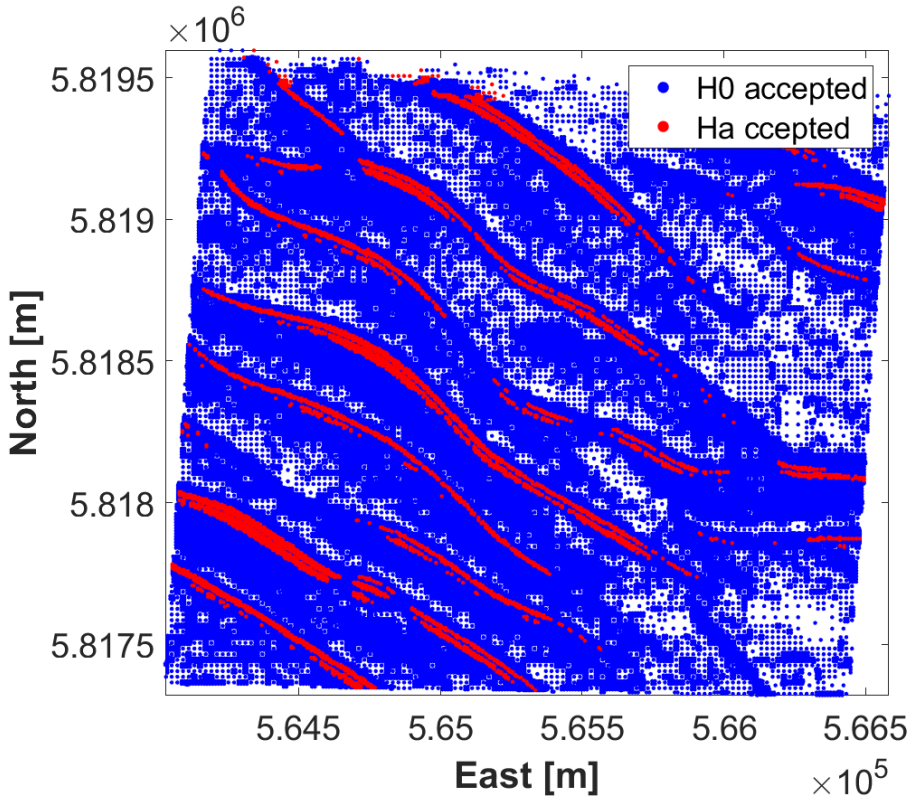


Figure 5.5: Binary map of NF subarea showing the locations of nodes for each hypothesis-functional model selected: Ho accepted (blue nodes), Ha accepted (red nodes).

5.1.3 Validation of Predictions

The validation of predictions was done, by adjusting the leave-one-out CV method to the time-series analysis, as explained in Chapter 4. Considering the first iteration, only the last survey moment was excluded form the training dataset and used for generating the test dataset. Due to the gaps in time-series and to avoid reducing spatial resolution by keeping only nodes with full-length timeseries, the training dataset does not have a fixed number of observations for each node. The mean validation error (CV_e) was estimated to be:

$$CV_e \simeq 0.36m$$

This error resulted by comparing the original observation and the predicted value. The time difference between them was 2 years and thus it concerns a two year prediction error. The error value (36cm), compared to the stochastic model of the observations (33cm - 8cm), the uncertainties of predictions (4cm) and the limited length of timeseries (around 10 observations), it can be assumed to be sufficient as it is in the same order of magnitude.

To further validate the quality of predictions values and the models used, the CB of 95% and 99% CI were constructed (from the predictions done using the training dataset). Then, the percentage of nodes from the test dataset (original observation of last survey) that are inside the CI was computed. This adds information to our validation, as with (CV_e) value we check how close is the original observation to the predicted value, but this way we can test the probability that the prediction would be inside the CI. This information is summarized in Figure 5.6. Consequently, the percentage of nodes of NF subarea that are inside the CI of 95% and 99% (Figure 5.6) equals to 85% and 89%, respectively.

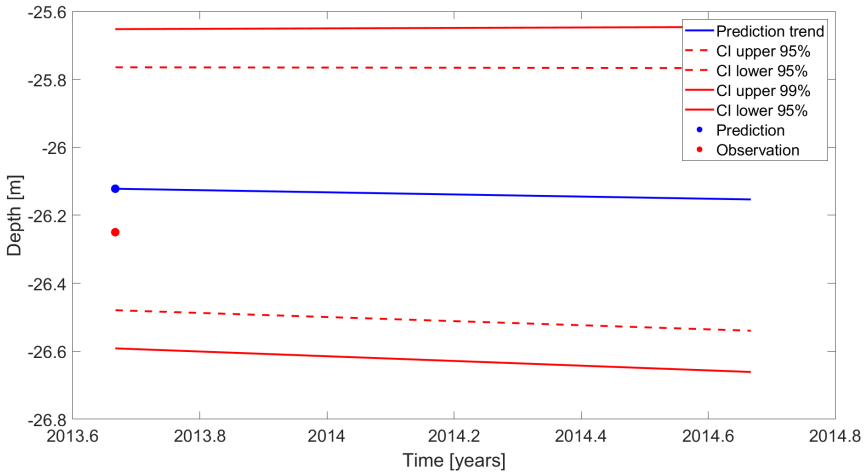


Figure 5.6: Original observation (test dataset, red scatter point), prediction trend (blue line) of training dataset with the corresponding CB (red lines).

The leave-one-out CV method was applied for the complete area of West IJmuiden, to check if and how the CV_e varies. Therefore, a thematic map showing the CV_e values for each subarea of West IJmuiden was created (Figure 5.7). It is observed that CV_e) ranges from 0.18 to 0.66m.

Only another iteration was preformed due to the limitation of time dependency and the small length of time series. In every iteration, a new survey is excluded from the training dataset and added in the test dataset and hence the length of training dataset changes and reduces. Therefore, the second iteration was done mostly to ensure the previous result. Now, the test dataset consists of the last two surveyed moments and the training dataset of the rest. The CV_e is expected to be a

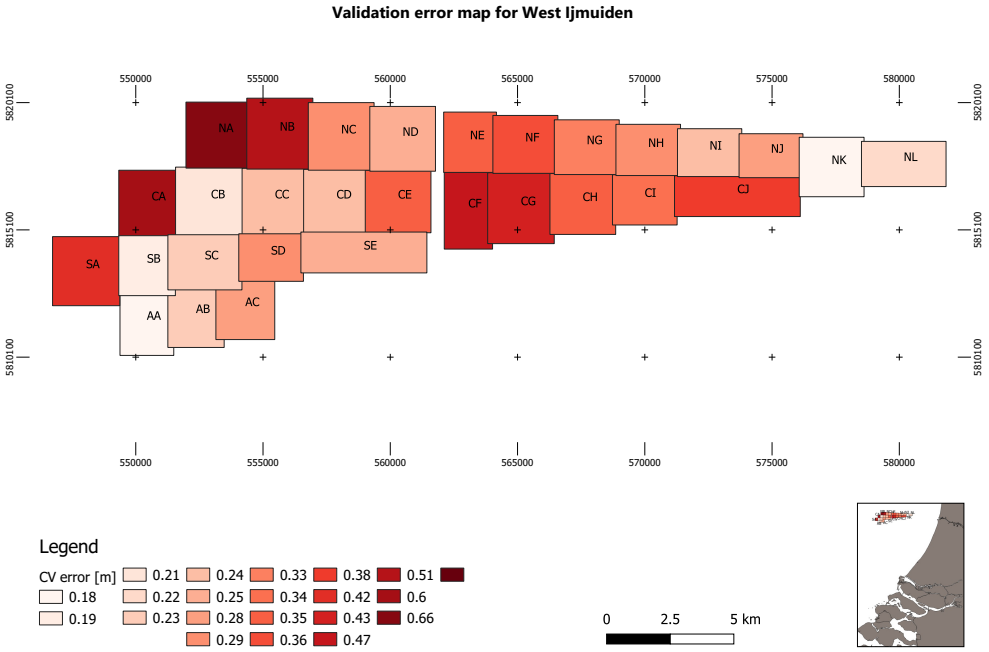


Figure 5.7: Thematic map of CV_e per subarea of the West IJmuiden. Areas with deep red color have higher errors compared to light pink areas.

bit higher as the training dataset reduced and it is:

$$CV_e \simeq 0.40m$$

The same steps, as before, were applied. The percentage of 'powerful' predictions, depending on the CB of the new training dataset was computed. They are also very close to the corresponding ones from the first iteration, as expected, and in particular: 80% for 95% CI and 85% for 99% CI.

In general, considering the full-length of the timeseries (10) and the precision of the original observations (33cm-8cm), the results of the validation show that the predictions for close prediction time intervals can be highly trusted. Hence, the chosen functional models are assumed to perform sufficiently, regarding their ability to predict how the seafloor changes through time. Additionally, this means that the predictions can be used as a powerful input for PRA.

5.1.4 Risk Indicators

The risk indicators of RDC and RoC, were defined in Chapter 4. They are used to quantify risk's impact, consequences, and RoC is used for the probability estimation, which is necessary for the risk assessment. RDC and RoC were estimated for all the prediction moments, but here only the ones using the predictions after two years of the last survey moment, $T_p = 3$, will be presented. This prediction interval was decided because West IJmuiden is categorized to be resurveyed every 2 years. The maps of RDC and RoC follow in Figures 5.8 and 5.9, respectively.

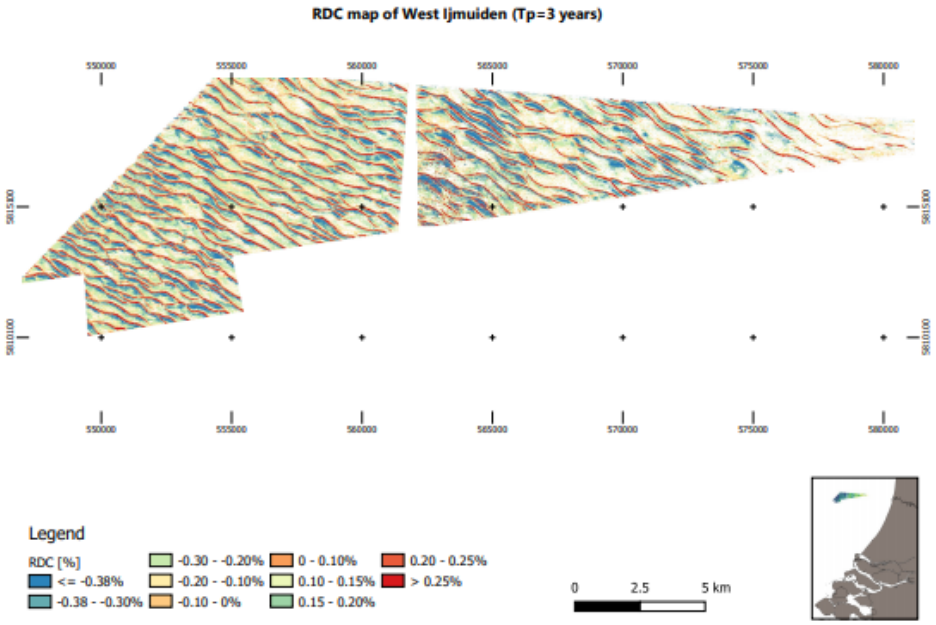


Figure 5.8: Map of the RDC of West Ijmuiden for $T_p = 3$. Red locations have the highest positive RDC, whereas blue locations have the highest negative RDC. Regions with light orange have RDC values close to zero.

Map of the RDC in Figure 5.8, is created for both directions (positive: shallower and negative: deeper). Red nodes have the largest RDC values, they are the most critical locations and they form the crests. For a better representation, the values of RDC were scaled between -0.38% and 0.25%. However, the maximum and minimum values differ and they are larger, which is expected as RDC is not normally distributed. RDC indicator is a ratio expressed in percent [%] and therefore, sometimes it is difficult to interpret it. A single RDC value is assigned to each node and it can be used to inform the user regarding the change in depth, without any other additional information needed, as for example; the average depth value of the node, or its categorization according to NLHO's factor of minimum depth.

RoC map in Figure 5.9, follows the same spatial pattern as the RDC map. Green nodes are getting shallower with a RoC value of more than 0.06m/yr (order of cm) and purple nodes are getting deeper with a RoC of more than -0.05m/yr. The range of RoC shows that it is more symmetrical for both directions compared to RDC, as expected as we assume that it is influenced only by random errors.

Direction of change is an important factor for this analysis, as explained in Chapter 4. A binary RoC map of West IJmuiden, indicating the direction of nodes, was created and is presented in Figure 5.10. Negative change equals to locations going deeper and these are the zero values (blue locations). Red nodes, that form lines, with values equal to one have a positive change and thus, they are getting shallower. From Figure 5.10, it is clear that the crests of the sand waves are getting shallower and the troughs and 'in-between' parts of them are getting deeper.

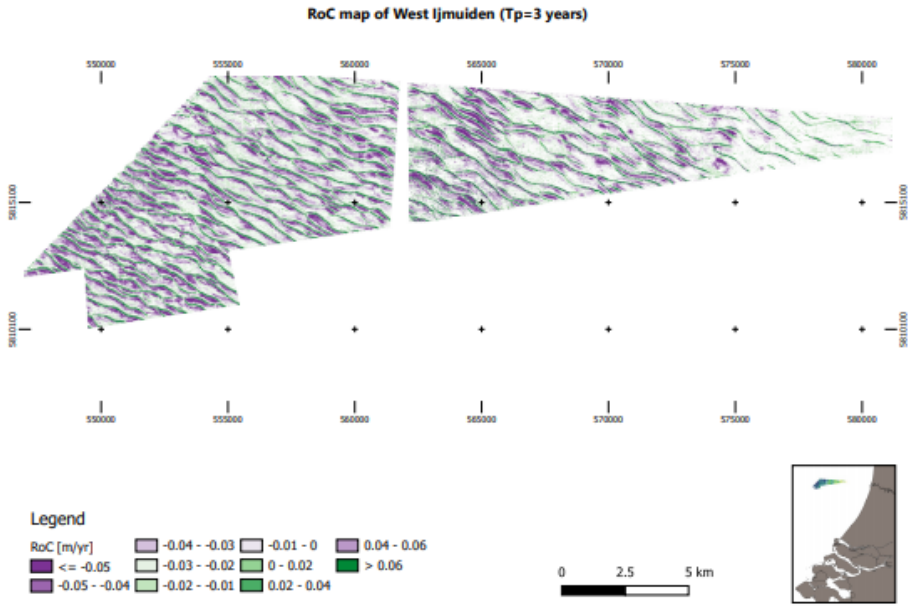


Figure 5.9: Map of the RoC of West IJmuiden for $T_p = 3$. Green locations have the highest positive RoC, whereas purple locations have the highest negative RoC. Regions with light purple and light green have RoC values close to zero.

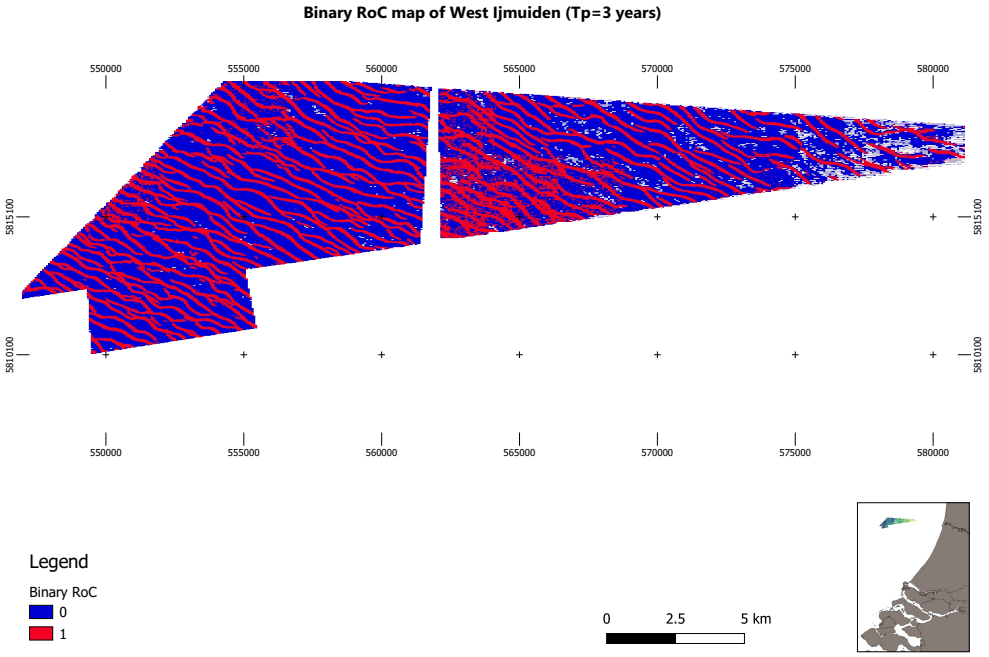


Figure 5.10: Directional binary RoC map of West IJmuiden. Blue regions, that are indicated with 0, show areas that are getting deeper, while red regions, that have value 1, represent the areas that are getting shallower.

5.1.5 Uncertainties of the Risk Indicators

The variances of the risk indicators of RDC and RoC were computed as explained in Chapter 4. For the RDC the variances were computed and are presented in Figure 5.11. For the RoC apart from the variances, CB were constructed for a 95% and 99% CI and they are presented in Figure 5.12 and in the Appendix A.

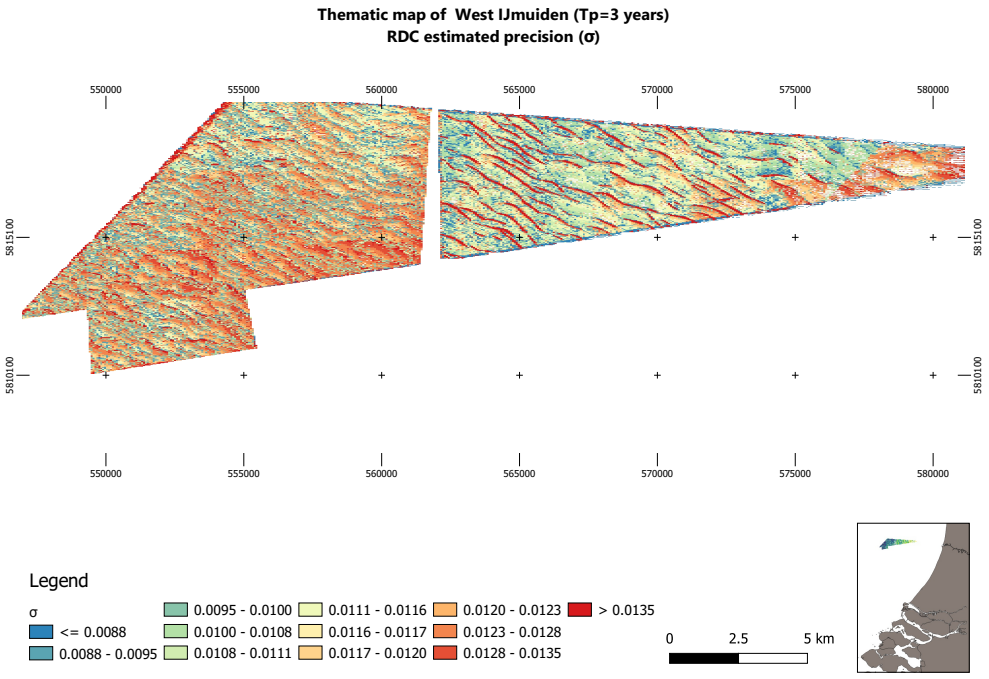


Figure 5.11: Map of the West IJmuiden with the uncertainties for the RDC values in a prediction moment equal to 2 years. Red values show locations with large uncertainties, while blue locations show areas with low uncertainties.

Figure 5.11 shows the uncertainties of the RDC risk indicator for the complete area of the West Ijmuiden in a prediction moment of 3 years. In general, the precision values range from 0.0088 to 0.0135 [%]. The red areas have high variances and thus, low precision while the blue areas have low variances and hence, high precision. It could be observed that the areas close to the boundaries of the study area have higher variances and this is linked to possible errors introduced during the gridding process. Additionally, what is more interesting is that the crests of the sand waves are coloured red, which means that they are estimated with lower precision. This was expected as the piece-wise linear model is used for these areas, which although it performs well when detecting large depth changes, it uses less observations for the predictions and this results in higher variances. Another important remark that should be made is that on the East side of the map there is a sudden diagonal line and a transition from green to red. This can be explained by the fact that for the subarea on the East side less observations are available compared to the other subareas and this influenced the uncertainties for the complete subarea. The latter can be proved by comparing this map with other previous presented maps of the area, where gaps in the data of the subarea can be observed. Consequently, Figure 5.11 indicates that the variances are in general relatively low for the complete area and that critical areas such as the crests of the wand waves can be detected and separated using the piece-wise model, however the predictions done are of lower precision.

The uncertainties of RoC indicator follow the same spatial pattern as before and a map of the variances of RoC for the complete area is presented in the Appendix A. Here, it was decided to present the uncertainties of RoC in NF subarea, (Figure 5.12) using σ and 1.96σ values. It becomes clear that the RoC values in the crests and troughs, especially when applying a 95% CI are of lower precision than the rest of the nodes in the area.

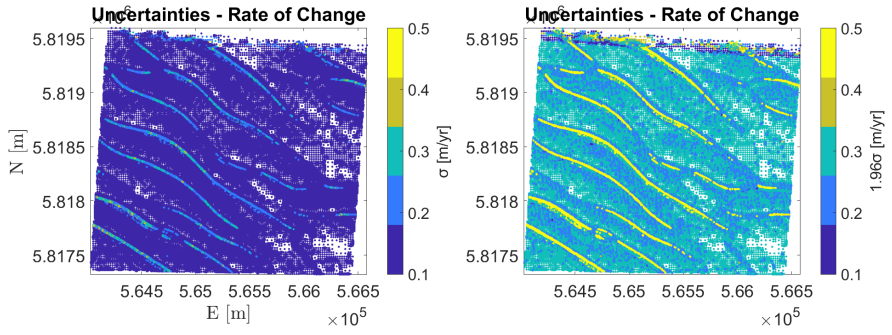


Figure 5.12: Overview of the uncertainties of RoC values for NF subarea (left) and uncertainties at a 95% CI, 1.96σ (right). Same scale was used in order to be able to compare the graphs and it is obvious that the right graph is associated to larger uncertainties.

5.1.6 Probability estimation

Considering the relationship between risk indicators and consequences, as defined in Chapter 4, the following probability was estimated:

The probability that the RoC of a node is exceeding a set of threshold values (ε_k)

The process of probability estimation was done for a set of ε_k, based on the original RoC values of each node. Hence, the process is automated and iterative. Additionally, probability estimation was done separately for positive and negative depth changes. This decision was made because the range of RoC shows different results for negative and positive direction, which should be pointed out as they are associated to different applications and risks. Hence, different thresholds should be applied in the same area for the RoC of the nodes depending on the application and the risk assessment. If, for example, nodes in NF area experience much larger changes, when getting deeper and there is a pipeline or cable network established in this location, then different ε_k should be applied (than for the nodes getting shallower) to indicate significant depth changes. Finally, sand waves do not grow symmetrically and hence, the depth changes they cause should be associated to different ε_k values for a proper PRA. The probability estimation of RoC is presented for both positive and negative direction:

Figure 5.13 illustrates the relationship between probability and ε_k values. The spatial pattern of the curve shows that as the threshold values (ε_k) increase the probability that a node will exceed them decreases. The maximum positive RoC value is around 0.45 [m/yr], as from the graph we obtain that the probability of a node to experience a larger RoC than 0.45 [m/yr] is almost zero. However, from the curve it is obvious that most of the nodes have a RoC value smaller than 0.15 [m/yr]. The histogram presented in Figure 5.14 is equivalent with the graph in figure 5.13 and is presented here for a better understanding. It can also be used as a tool to calculate and evaluate the consequence, when a particular value of ε_k is defined. For example in this case $\varepsilon_k = 0.17[m/yr]$; the sum of red bins is the number of nodes exceeding this particular threshold value, which seems to be a small amount of nodes of NF area.

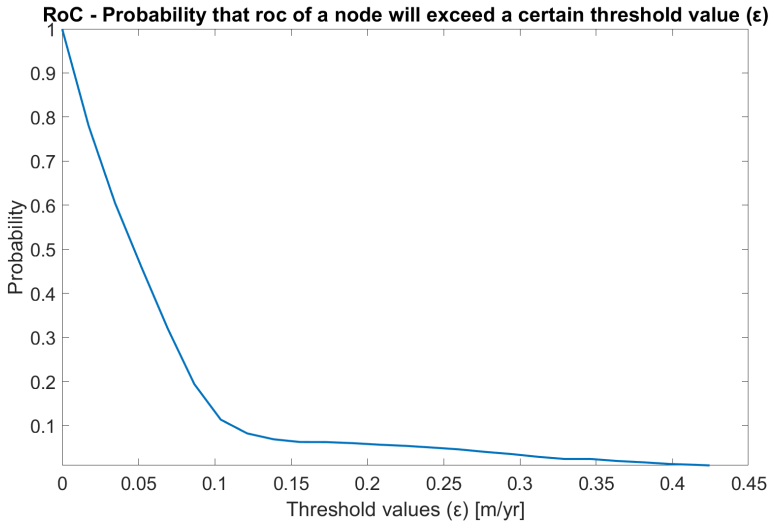


Figure 5.13: Grapth with the probability of the positive RoC of a node to exceed a set of ε_k values for NF subarea.

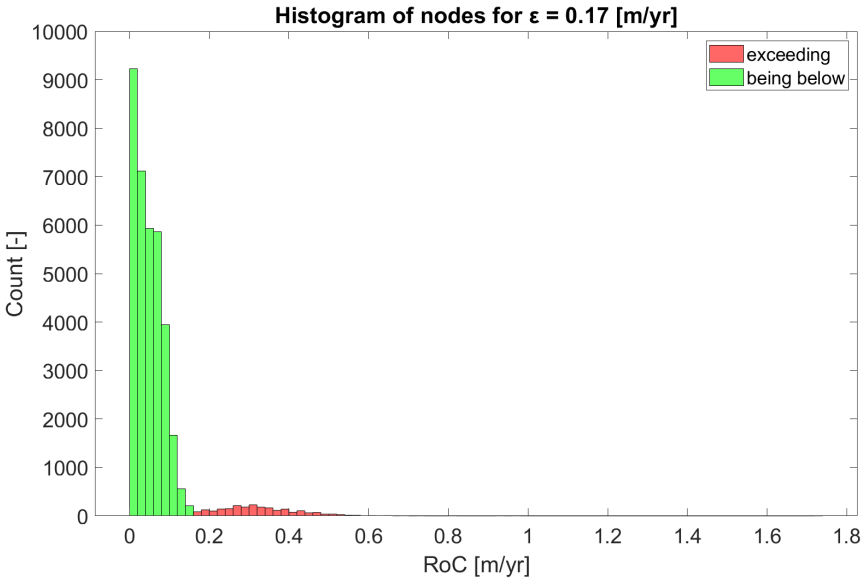


Figure 5.14: Histogram of the positive RoC of a node to exceed 0.17[m/yr] (red bins). Green bins show the number of nodes moving with a RoC smaller than below 0.2[m/yr].

The same steps were implemented to estimate the respective probability of RoC risk indicator for the negative direction, when the nodes are getting deeper. The curve and histogram are inverse and the threshold values (ε_k) get different values.

From Figure 5.13 it is obtained that the minimum value for the negative direction is -0.15 [m/yr]. The minimum value in the negative direction is equivalent to the maximum value in the positive direction and they are assumed to be the extreme threshold values ε_k. Hence, the nodes that are getting deeper experience depth changes slower with respect to time compared to the nodes getting shallower. Most of the nodes have a RoC value larger (closer to zero) than -0.04 [m/yr].

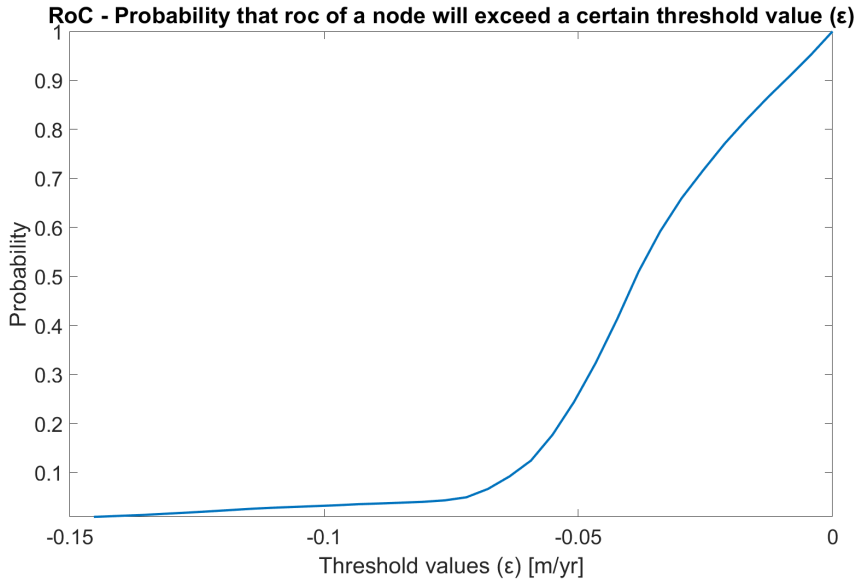


Figure 5.15: Probability of the negative RoC of a node to exceed a set of ε_k values for NF subarea.

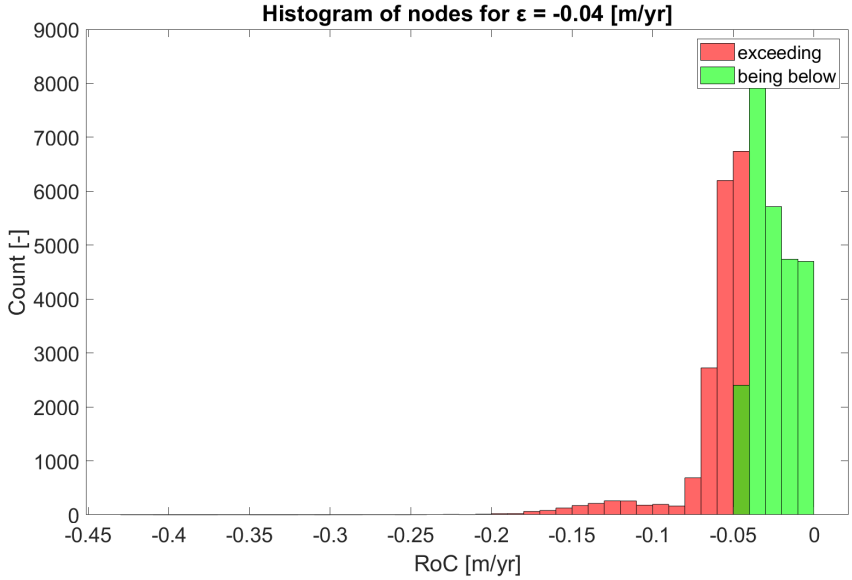


Figure 5.16: Histogram of the negative RoC of a node to exceed -0.04 [m/yr] (red bins). Green bins show the number of nodes moving with a RoC larger (closer to zero) than -0.04 [m/yr] for NF subarea.

5.1.7 Risk maps

Two thematic risk maps of the complete area of West IJmuiden were constructed. They are used here as an example for the visualisation purpose. Map in Figure 5.17 represents the different probability values estimated for each subarea, when using a fixed threshold value for RoC: $\varepsilon = 0.07[m/yr]$. Areas of deep green color have a relatively high probability (>35%) that the RoC of their nodes will exceed $\varepsilon = 0.07[m/yr]$.

Following the same concept, figure 5.18 shows the corresponding threshold values evaluated (ε_k) for RoC of a node when keeping fixed the probability value equal to 5% for all the subareas of West IJmuiden. The probability value is adaptable and relatively low to show an extreme scenario and emphasize the corresponding ε_k

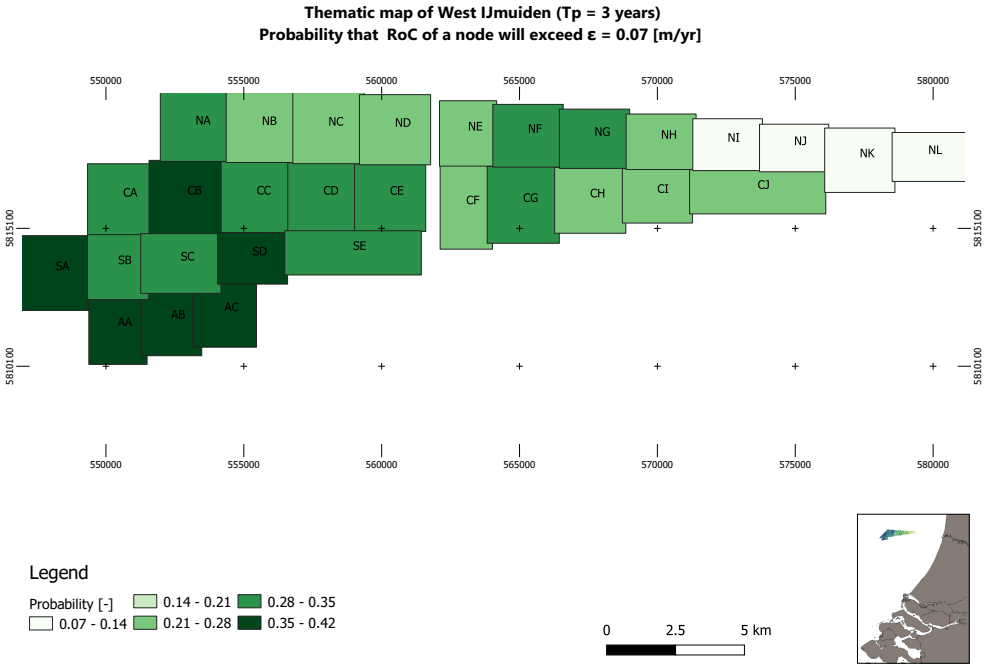


Figure 5.17: Thematic risk map of West IJmuiden showing the probability that the RoC of a node exceeds fixed $\varepsilon = 0.07[m/yr]$. A probability value is assigned for every subarea. Areas with deep green color have the highest probability (0.42), whereas areas with white color the lowest probability (0.07)

values that the RoC can take for each subarea. This can be proved from the results as ε_k ranges from 0.07 to 0.22 [m/yr].

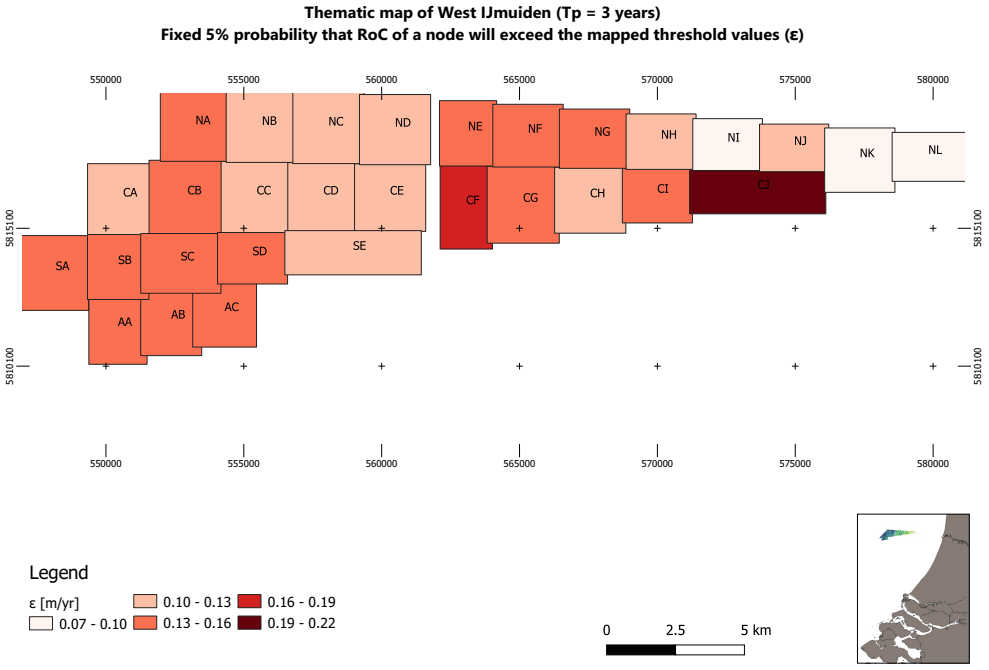


Figure 5.18: Thematic risk map of West IJmuiden showing ε values when the probability of the RoC of a node is fixed and equal to 5%. An ε value is evaluated and assigned to each subarea. For regions with red color a high value of ε should be applied (0.2), whereas for regions with white color a lower value of ε should be applied (0.07).

Thematic maps, like the ones presented here, can be generated for every chosen threshold and probability value. These maps are an example of how the user can apply the algorithm and decide upon different values. They can be summarized in one risk probability map which was also constructed. However, the risk maps presented here show the spatial distribution of the values and hence, they are more representative for illustrating the risk of the seafloor to change significantly. Additionally, they can be used, combined with other background maps, to better interpret the vertical dynamics of the study areas.

5.2 CASE STUDY II: WEST ROTTERDAM

West Rotterdam is the second area of interest that will be discussed in this study. It is a smaller area compared to IJmuiden and therefore there was no need to divide it into smaller subareas. It consists of almost 100,000 nodes and its area size is around $2.5 \times 4.5km \simeq 11.25km^2$. It is close to the coast, sand waves are moving and it is a main shipping route of NCS. Its special characteristics are discussed in Chapter 3 and some basic statistics will be presented in Table 5.2. To get a sense of how the area looks like, a bathymetric map was created using the last surveyed moment (Figure 5.19).

Table 5.2: Basic statistics of West Rotterdam.

Survey	$d_{\min}$	$d_{\max}$	$\overline{d}$	$\overline{\sigma}$	Δd
1990	-15.213m	-18.645m	-15.963m	$0.77m^2$	
					0.277m
1994	-15.340m	-18.889m	-16.240m	$0.47m^2$	
					0.023m
1996	-15.486m	-18.782m	-16.263m	$0.75m^2$	
					0.128m
2000	-15.261m	-18.703m	-16.135m	$0.70m^2$	
					0.021m
2000	-14.423m	-19.177m	-16.114m	$0.65m^2$	
					-0.033m
2001	-15.315m	-18.694m	-16.147m	$0.77m^2$	
					0.950m
2006	-13.685m	-19.720m	-17.096m	$0.06m^2$	
					0.481m
2007	-13.280m	-19.320m	-16.615m	$0.06m^2$	
					0.354m
2011	-13.300m	-19.300m	-16.260m	$0.07m^2$	
					0.580m
2014	-13.320m	-19.136m	-15.680m	$0.07m^2$	

From table 5.2 and figure 5.19, it is observed that it is a very shallow and dynamic area with a mean depth value being around 16m. The maximum depth values in the whole area are bellow 20m and the minimum values reach 13m. Additionally,the precision of measurements differs significantly throughout the years. Before 2005 is relatively large around $0.77m^2$ and after 2005 it drops to $0.6m^2$. This large difference, we expect to affect the results of deformation analysis and the predictions. A w-test was performed to find possible outliers, however we can not be sure if these measurements are outliers or the values are due to a physical process happening in the area. Therefore, they were not excluded from the analysis. Finally, from the gridding process and the way that nodes are organised in figure 5.19, it can be seen which are the most interesting locations of the study area, as the ones with smaller grid size have higher variability. These location have denser information which equals to more nodes.

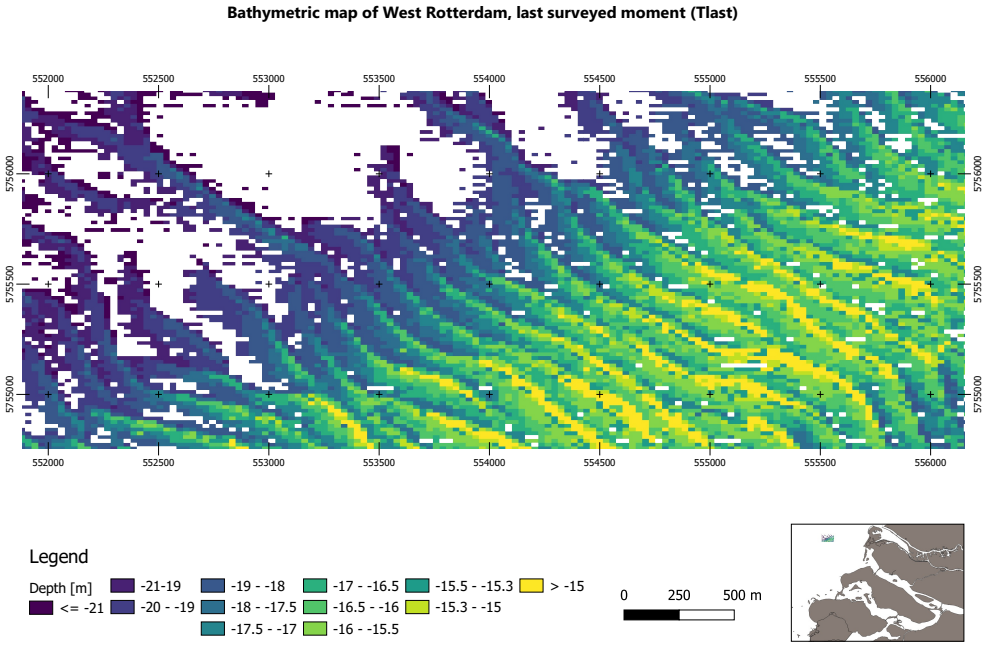


Figure 5.19: Bathymetric map of the West Rotterdam at the last surveyed moment. Depth varies from areas deeper than 20m (deep blue) to areas shallower than 15m (yellow).

5.2.1 Estimations, Predictions & Uncertainties

The same functional models (constant velocity model and piecewise linear model) are used to describe the seafloor evolution of West Rotterdam. Figures 5.3 and 5.4 show the observations, the model fit and the predictions together with their uncertainties for each hypothesis accepted, H_0 and H_a, respectively.

Figure 5.3 shows an example node for which H_0 was accepted. Original observations have large errorbars, until 2005, and hence poor precision. The precision of the estimations is also influenced as the errors propagate. Therefore, the estimations do not fit that well to the real observations (fit line), compared to West IJmuiden. We also expect the uncertainties of predictions to be larger. To get a sense of the quality of predictions the following have been estimated:

- Depth of the node at the last surveyed moment (original observation,2014) is -19.14m.
- Its prediction depth value after 3 years is 19.27 $\pm$ 0.9m with a 99% CI.
- Its prediction depth value after 10 years is 19.41 $\pm$ 1.1m with a 99% CI.

Figure 5.4 refers to the nodes for which the piecewise linear model is chosen as the most likely model and H_a is accepted. The fitted line seems to mimic well the original observations and the quality of predictions made with this model seem to be almost equivalent with the predictions made with the constant velocity model. Hence, we expect uncertainties of prediction to be almost the same as for IJmuiden. The quality of predictions can be described as follows:

- Depth of the node at the last surveyed moment (original observation) is -15.81m.
- Its prediction depth value after 3 years is 14.41 $\pm$ 1.1m with a 99% CI.
- Its prediction depth value after 10 years is 11.15 $\pm$ 3.5m with a 99% CI.

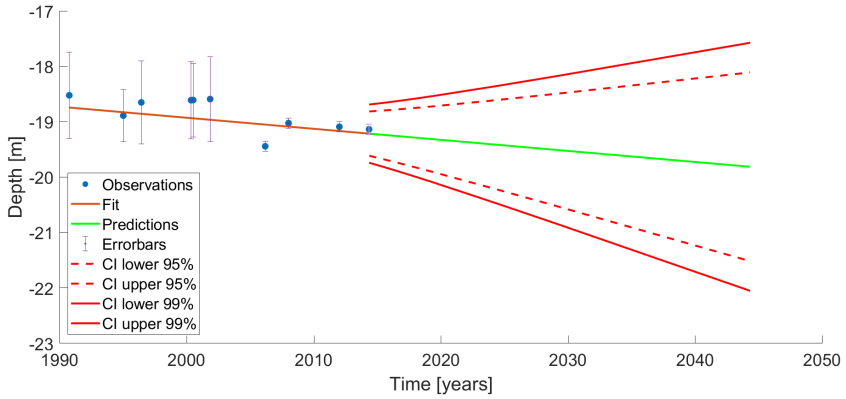


Figure 5.20: Example node for Ho accepted: Original observations (blue scatter points) with their precision (errorbars), estimations done using constant velocity model (red line, fit), prediction trend (green line) with the associated quality of the predictions illustrated by CI of 95% and 99%.

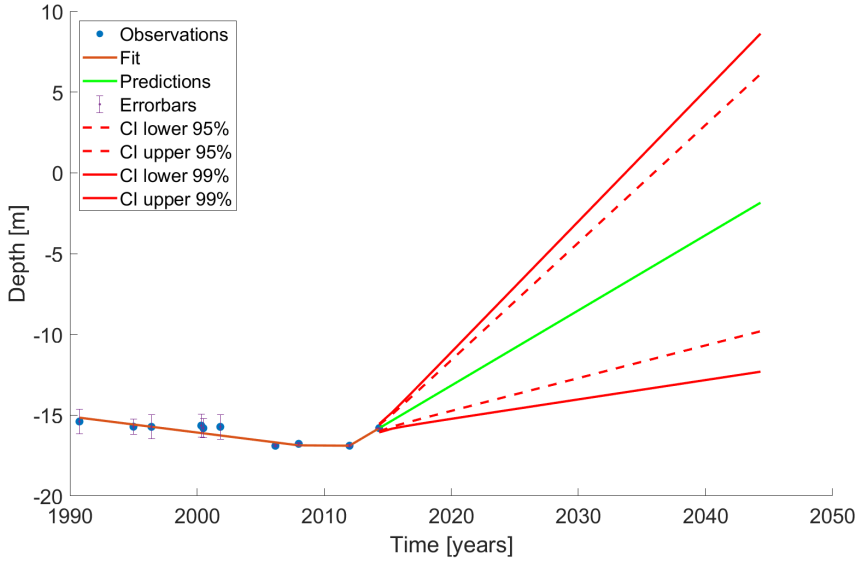


Figure 5.21: Example node for Ha accepted: Original observations (blue scatter points) with their precision (errorbars), estimations done using piecewise linear model (red line, fit), prediction trend (green line) with the associated quality of the predictions illustrated by CI of 95% and 99%.

It is obvious from the results and the graphs that the precision of observations for West Rotterdam is lower compared to the precision of of observations of West IJmuiden. This influenced the prediction values and their quality. What is interesting, is the relatively large uncertainties coming from the constant velocity model, as it was expected to provide more precise predictions and will be discussed further in the following sections.

Finally, a binary map of the area depending on the hypothesis testing is presented in Figure 5.22. From this we obtain that in West Rotterdam there is greater amount of nodes for which H_a hypothesis is satisfied, compared to IJmuiden. Hence, piecewise model is used, for many nodes as the most likely model to describe their evolution through time. To be more precise, almost half of the nodes (50%) use piecewise linear model and this proves that West Rotterdam is a very dynamic and critical area of NCS.

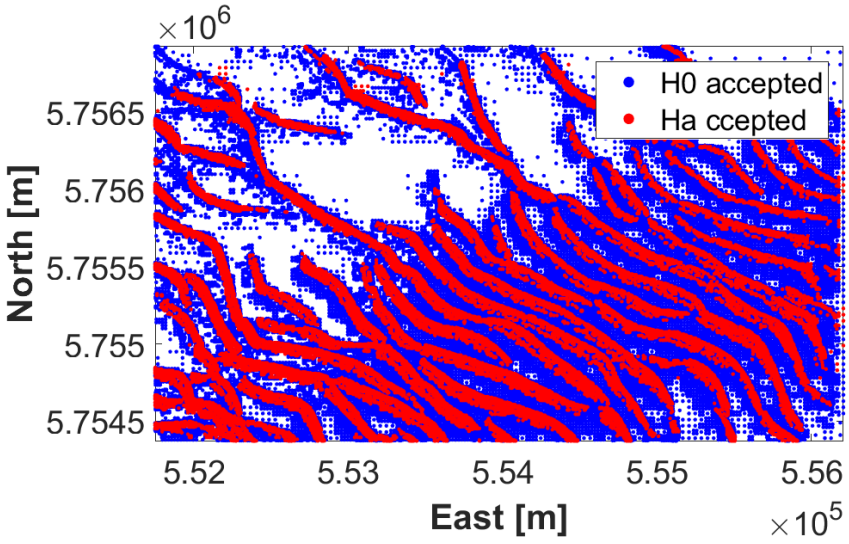


Figure 5.22: Binary map of West Rotterdam. Locations of nodes based on hypothesis testing: Ho accepted (blue nodes), Ha accepted (red nodes).

5.2.2 Validation

Validation results for West Rotterdam are relevant to conclude regarding the performance of the models and how useful the predictions could be for a further analysis. Leave-one-out CV method was applied for West Rotterdam as well and it resulted in a mean validation error equal to:

$$CV_e \simeq 0.98m$$

It should be mentioned, that West Rotterdam has a large amount of nodes with gaps in the time series and some of them have gaps of large time intervals. This is a factor that influences the quality of the results, as with small timeseries it is difficult to obtain high quality estimations and hence predictions. Considering also the level of detail, which is necessary to be at a high level in order to map the area, it was decided to use almost all the nodes even if the length of timeseries was short. Nodes with more than 3 observations were used.

The error value was expected from the previous results to be a bit higher than of West IJmuiden ($\approx 0.38m$) as the accuracy of the predictions is lower. Considering the intense natural process taking place in the area it can be assumed that this value is sufficient. The number of nodes inside the CB (95% and 99% CI) of prediction trend (training dataset) is: around 76% of nodes for 95% CI and 78% of nodes for 99% CI. An example of this can be seen in Figure 5.23. These results show that a relatively large amount of predictions can be trusted and hence they are useful for a further analysis.

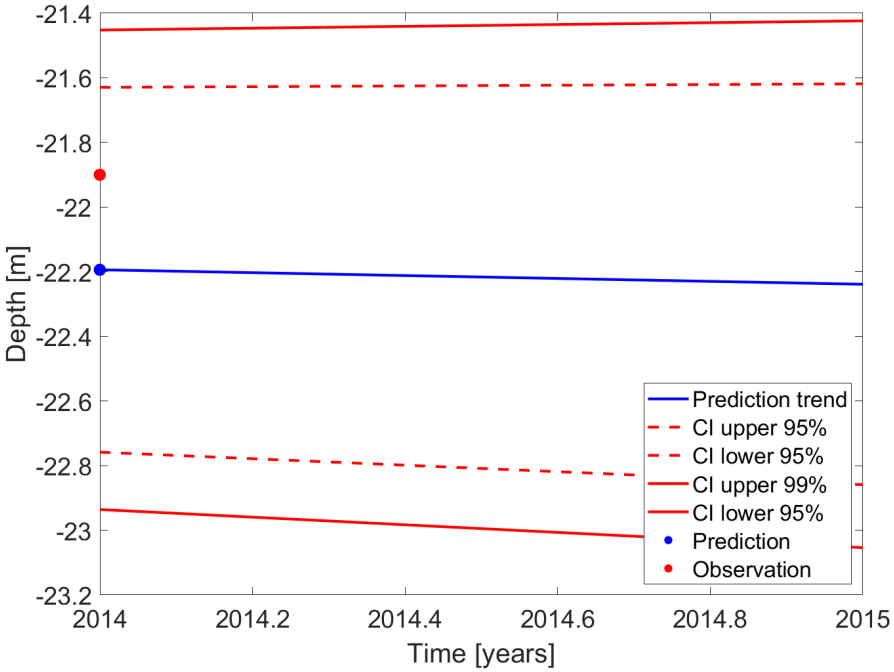


Figure 5.23: Original observation (test dataset, red scatter point), prediction trend (blue line) of training dataset with the corresponding CB (red lines).

5.2.3 Risk Indicators

The risk indicators of RDC and RoC were also computed for each node of West Rotterdam. Maps of them were constructed and can be seen in Figure 5.24 and 5.25.

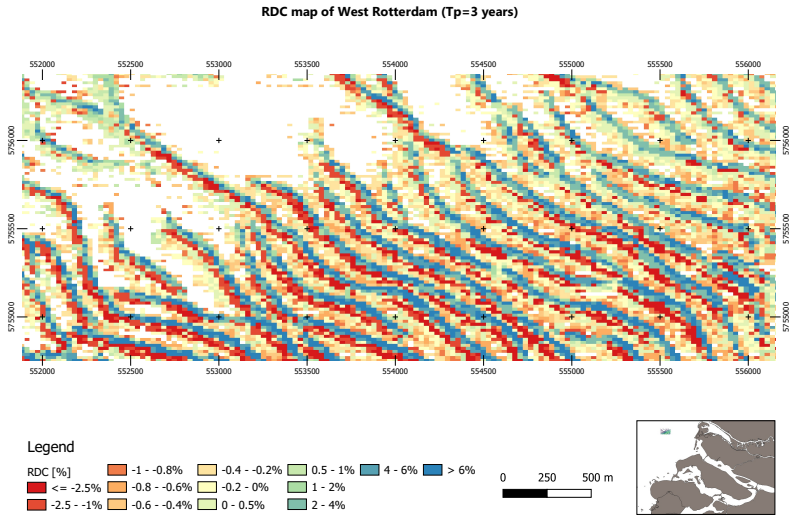


Figure 5.24: Map of the RDC of West Rotterdam for $T_p = 3$. Blue regions have large positive RDC values (greater than 6%), while red regions have large negative RDC values (lower than -2.5%). Areas with light yellow and light green color have RDC values close to zero.

The spatial resolution of the maps is affected by the small number of nodes (around 100000) but still it is obvious that the area is influenced by intense natural processed. RDC values for a prediction moment after 2 years from the last surveyed moment, range from -10% to over 20% which are slightly extreme values,

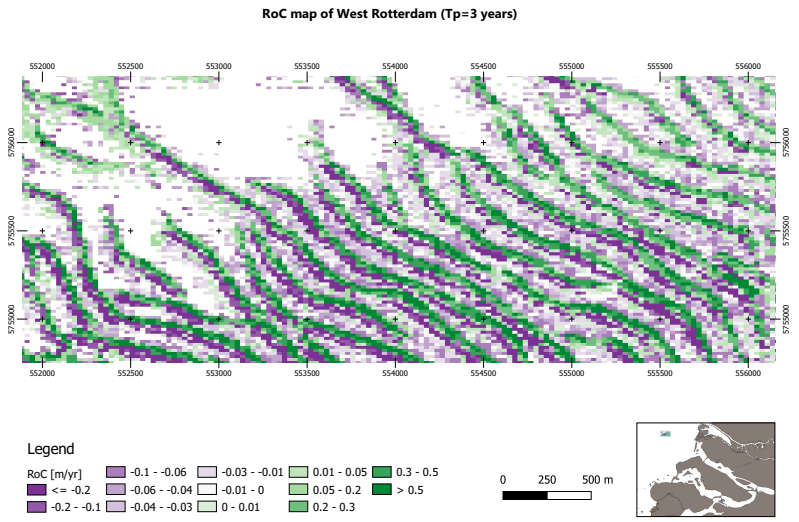


Figure 5.25: Map of the RoC of West Rotterdam for $T_p = 3$. Deep green regions have large positive RoC values (greater than 0.5m/yr), while purple regions have large negative RoC values (lower than -0.2%). Areas with light green color have RoC values close to zero.

compared to West IJmuiden. The same applies for RoC values that vary from -1 to 1.5 m/yr. The maps in figures 5.24 and 5.25 are scaled in order to be representative.

Additionally, as in West IJmuiden, RDC values do not grow symmetrically in the positive and negative direction. The nodes getting shallower (green and blue nodes) experience much larger depth changes than the ones getting deeper (orange to red nodes), as illustrated in 5.24. RoC values are more symmetrical with respect to direction and comparing figure 5.24 and 5.25 we obtain that nodes that getting shallower faster than the nodes getting deeper. Considering that these nodes are the crests and troughs of the sand waves we conclude that these are the most critical locations.

A directional binary map was also created to indicate the direction of change for the nodes of West Rotterdam (Figure 5.26).

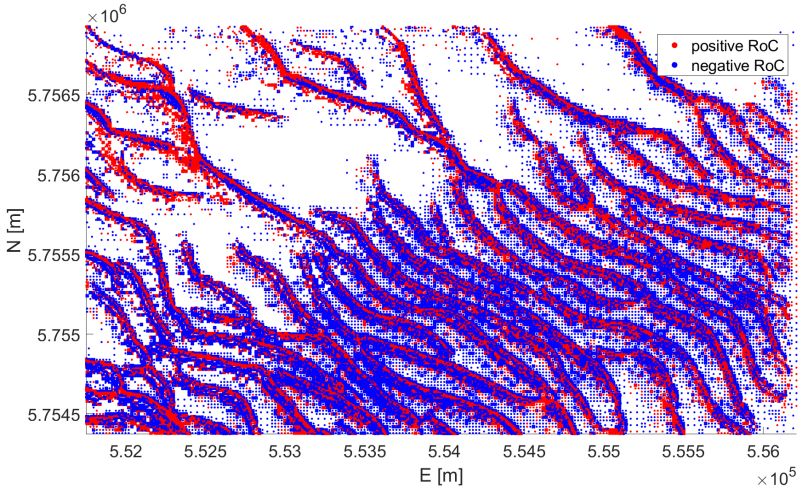


Figure 5.26: Directional binary RoC map of West Rotterdam. Blue regions, indicate areas that are getting deeper, while red regions, represent the areas that are getting shallower.

5.2.4 Uncertainties of Risk Indicators

The uncertainties for the risk indicators of RDC and RoC were computed for West Rotterdam and are presented in this section. For the RDC a map of the variances per node is presented in Figure 5.27. It becomes obvious that the uncertainties for this study area are larger and vary from 0.05 to 0.1 [%]. As in West IJmuiden the crests of the sand waves have larger uncertainties and especially on the edges.

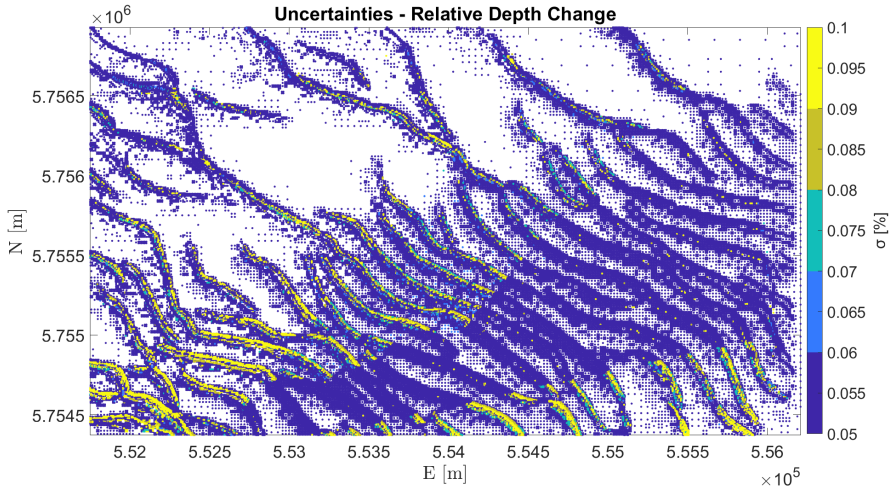


Figure 5.27: Map with the uncertainties, σ, of RDC values for West Rotterdam. Yellow areas have larger uncertainties and thus, lower precision.

For the RoC the variances were computed and in Figure 5.27 two maps are presented; on the left side there is a map with the variances of the RoC and on the right side a map with the variances computed with a 95% CI, 1.96 σ. The same locations as in the RDC indicator seem to have larger variances and these are the crests of the sand waves. This becomes more obvious on the right map of Figure 5.28. A map with the estimated variances for the RoC indicator with a 99% CI is included in the Appendix A.

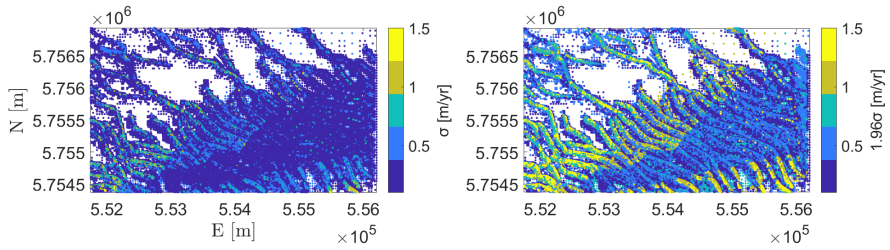


Figure 5.28: Overview of the uncertainties of RoC values for West Rotterdam (left) and uncertainties at a 95% CI, 1.96σ (right). Same scale was used in order to be able to compare the graphs and it is obvious that the right graph is associated to larger uncertainties.

5.2.5 Probability Estimation

The same concept as in West IJmuiden was applied to define the probabilities that the RoC of a node is exceeding a set of threshold values (ε_k). The probability estimation for the positive direction is presented in Figure 5.29 as well as the corresponding histogram in Figure 5.30.

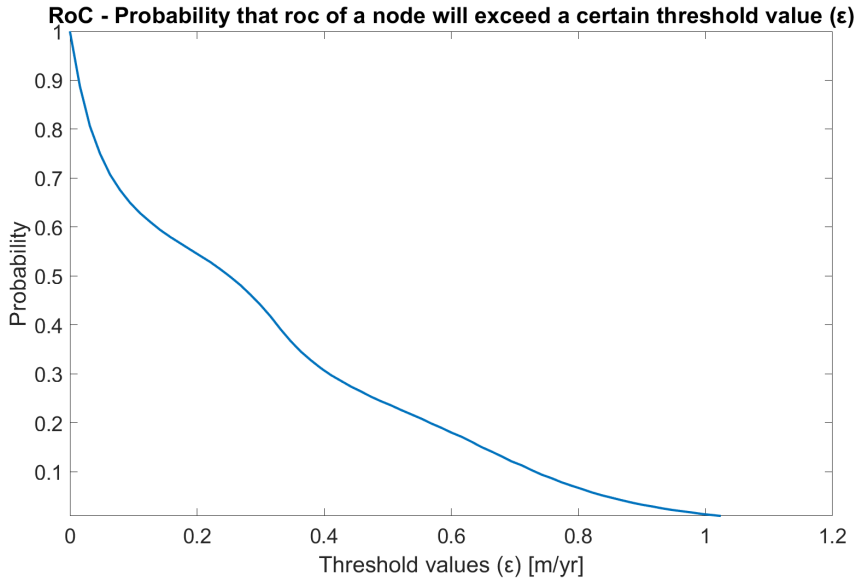


Figure 5.29: Probability of the positive RoC of a node to exceed a set of ε_k values for NF subarea.

It can be seen from figure 5.29 that the curve is not that smooth, which can be translated to more unexpected values regarding this risk indicator.This applies for areas that their depth is reducing and hence they are getting shallower with a RoC that can reach up to 1m/yr.

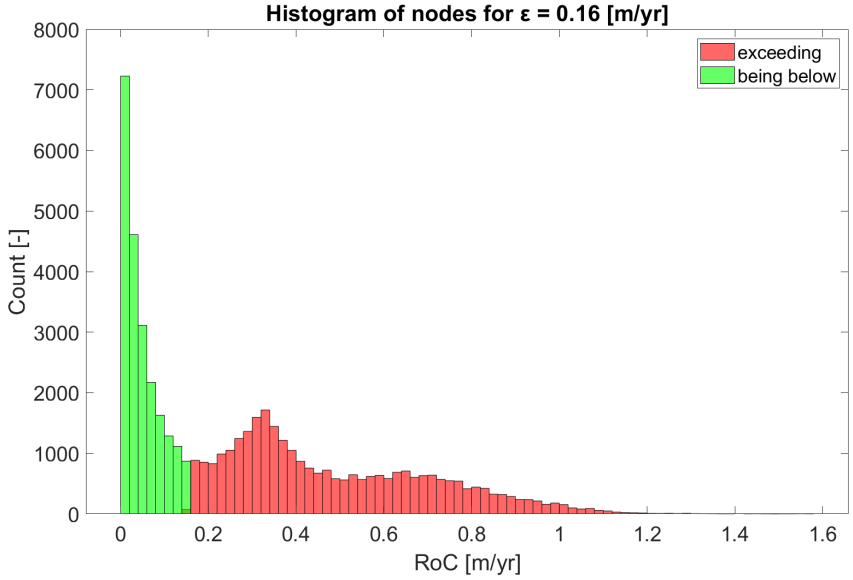


Figure 5.30: Histogram of the positive RoC of a node to exceed 0.2[m/yr] (red bins). Green bins show the number of nodes moving with a RoC smaller than below 0.2[m/yr].

The histogram in figure 5.30 is created considering a fixed RoC threshold value of $\varepsilon = 0.16[m/yr]$. The number of nodes with RoC values, in the positive direction, exceeding this threshold value is indicated by the sum of the red bins. Additionally, the histogram ensures what was discussed previously; as the probability curve is not that smooth, the number of nodes is not always reducing with respect to the increase of RoC threshold values (ε_k). This as fact influences the risk of the seafloor to change significantly in West Rotterdam.

Studying the probability pattern of RoC in the negative direction the following Figures 5.31 and 5.32 are presented:

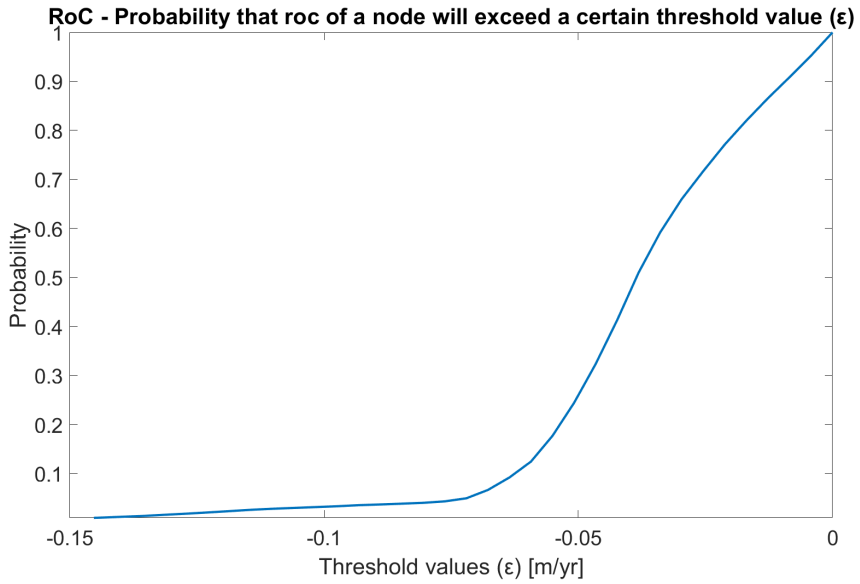


Figure 5.31: Probability of the negative RoC of a node to exceed a set of ε_k values for West Rotterdam.

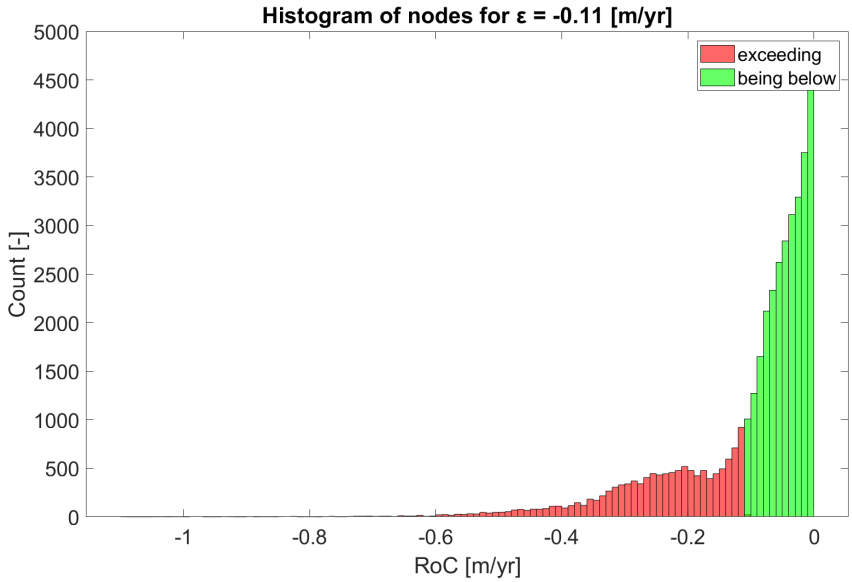


Figure 5.32: Histogram of the negative RoC of a node to exceed 0.1[m/yr] (red bins). Green bins show the number of nodes moving with a RoC smaller than below 0.1[m/yr] for West Rotterdam.

5.2.6 Risk maps

West Rotterdam is not subdivided in smaller areas and hence the corresponding risk maps are presented per node. In Figure 5.33 a threshold value for RoC indicator is fixed $\varepsilon = 0.82[m/yr]$ and the corresponding probability is estimated and is equal to 0.06, which equals to 6% of nodes (red nodes).

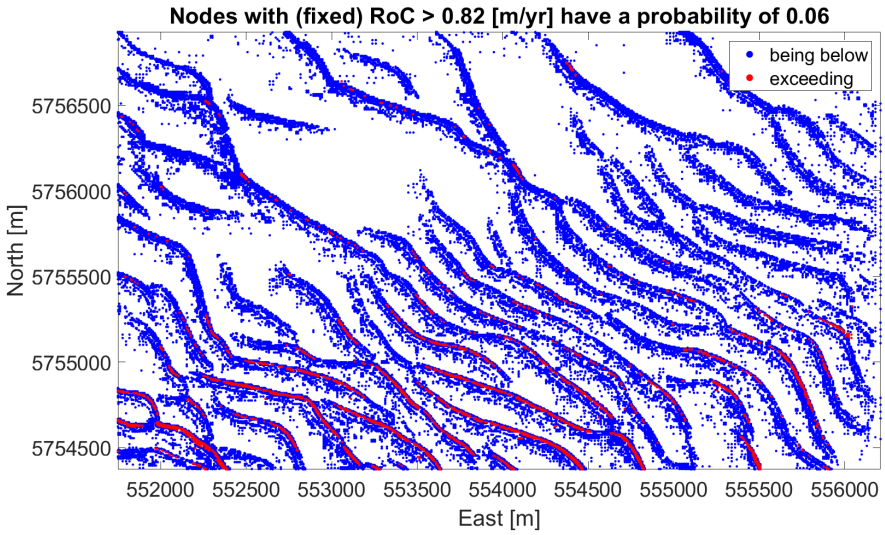


Figure 5.33: Probability that RoC of a node exceeds 0.82 [m/yr]. Blue locations indicate the nodes that their RoC is bellow 0.82 [m/yr] and red areas, the areas that their RoC exceeds this threshold value and thus, an alarm level should assigned for these areas.

On the other hand, in Figure 5.34, the probability is fixed and the respective threshold value for RoC is estimated. Hence, it is observed that 0.32 [m/yr] is the threshold value ε that should be used to obtain a probability of around 0.42. This is equivalent to 42% of nodes of West Rotterdam have a RoC value that exceeds 0.32 [m/yr].

This iterative method gives the user the opportunity to take the decision regarding; either fixing a probability value and find the corresponding threshold value, ε, or define a threshold value ε, and estimate the corresponding probability. Therefore, the process of defining significant depth changes, can be translated to assigning extreme threshold values to RoC values and the probability of them to occur for each area can be estimated.

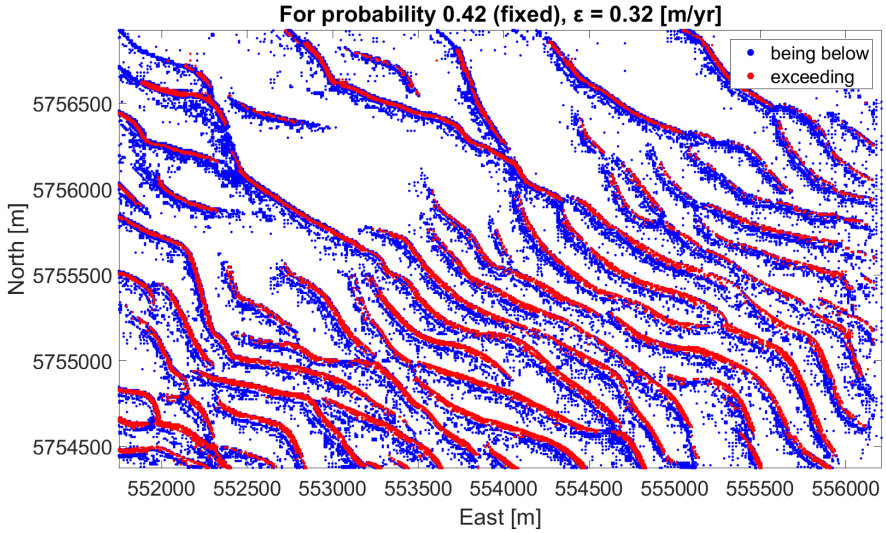


Figure 5.34: An ε value of RoC for fixed probability equal to 0.42. Blue nodes represent the nodes that considering the fixed probability equal to 0.42 and the resulted $\varepsilon = 0.32$ m/yr, their RoC is below this threshold value value. The same concept applies for the red nodes that exceed this threshold.

5.3 ADJUSTMENT WITH THE STANDARDS OF THE NETHERLANDS HYDROGRAPHIC OFFICE

It is relevant to ensure that the method applied here can contribute to the undergoing research regarding the optimization of the resurvey plan of the NCS. Therefore, it was necessary to get in contact with the NLHO in order to adjust their standards in the method applied in this study. This process can be seen as a validation of the results and as an opportunity to optimize the approach suggested.

The current survey plan of NLHO (until 2019) regarding the NCS, on a global scale, is based on the following:

1. Every 2 years: Anchorage areas, adjacent to shipping routes to major North Sea ports
2. Every 3 years: Areas adjacent to shipping routes to major North Sea ports
3. Every 5 years: International Maritime Organisation (IMO) ships routing
4. Every 10 years: From 10m depth contour to approximately 30m contour
5. Every 15 years: From 30m depth contour to approximately 40m contour
6. Every 25 years: deeper than 40m depth contour and no dynamic seabed

Until now, approximately 90% of the NCS has been surveyed using an MBES and mapped by generating a digital terrain model with at least 1 observation per 5m. The remaining 10% needs to be monitored to know if there are any objects above the sea floor that may cause a danger for navigation.

NLHO has the ambition to maintain its quality level of bathymetric surveys at Order 1 according to the categorization of S44 IHO standards. For this reason it uses a set of indicative threshold values for the depth change (defined again by S44, IHO standards) to depict the shallowest point in an area that may cause a danger for navigation. The threshold values for the vertical changes are the following:

1. For a depth range from 0 to 20m a threshold value rounded up to nearest 0.1m.
2. For a depth range from 21 to 30m a threshold value rounded up to nearest 0.5m.
3. For a depth range greater than 31m a threshold value rounded up to nearest 1.0m.

These threshold values were adjusted to the risk indicators of RDC and RoC. As explained before, the method suggested in this study suggests a set of threshold values that indicate significant depth changes. Applying the thresholds given from the NLHO, two new indicators are created; The Maximum Allowed Change (MAC) and the Next Survey Moment (NSM). These two indicators are used to derive conclusions regarding the percentage of the nodes that should be resurveyed in an area as well when the area should be resurveyed.

The steps done to compute MAC are:

1. The change in depth was computed per node at a decided prediction time T_p (as done for the nominator of the RDC and the RoC).
2. Within the provided depth intervals, the change in depth was divided by the corresponding threshold values. For defining the depth intervals, d_{last} was used for each node.
3. The resulted ratio is MAC and it varies between 0 and > 1.

After MAC was computed for each node, a rule was created; if MAC is greater than 1 then the node exceeds the defined maximum allowed depth change, if MAC is smaller than 1 then the node is below the defined maximum allowed change. Hence, the percentage of nodes with MAC greater than 1 was computed and these are the critical points that need to be resurveyed at T_p. Following this process, a map showing the critical nodes (red locations) of the West IJmuiden was created and is presented in Figure 5.35.

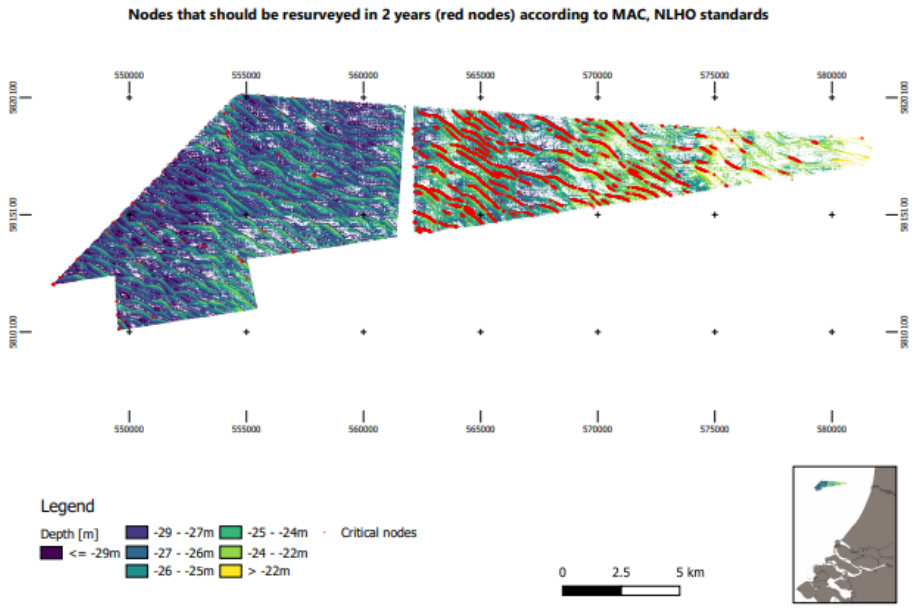


Figure 5.35: Critical nodes according to NLHO standards and MAC indicator, applied for West IJmuiden using as a background the bathymetric map of the last surveyed moment. The red nodes indicate the areas that should be resurveyed 2 years after the last surveyed moment.

The MAC of each node was computed for a prediction time equal to 2 years after the last surveyed moment. This was decided as the study area of West IJmuiden is categorised to be resurveyed every 2 years and therefore it makes sense to test the results in this time interval. The red nodes in Figure 5.35 are the ones which have $MAC > 1$. It is interesting that the critical locations create lines that cross the crests of the sand waves and their spatial pattern can be seen easily. This becomes more obvious, in the right part of the map where the presence of sand waves becomes more noticeable.

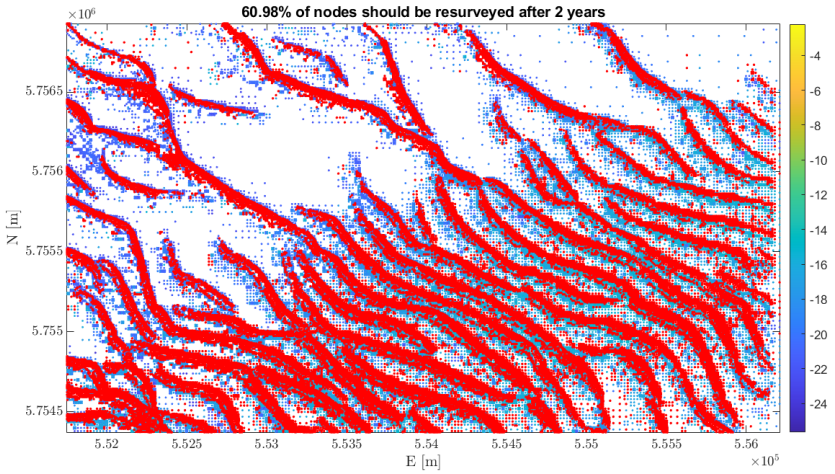


Figure 5.36: Critical nodes according to NLHO standards and MAC indicator, applied for West Rotterdam. The red nodes indicate the areas that should be resurveyed 2 years after the last surveyed moment.

For the West Rotterdam, the same process was followed and the MAC was computed per node. However as it is not divided in subareas the results are presented as follows in Figure 5.36. It can be observed, that more than 65% of the nodes should be resurveyed after 2 years. Hence,the assigned prioritisation and survey frequency of the are remain the same and are justified.

The resulted percentages of nodes that should be resurveyed were used to decide regarding the subareas. This is a relevant process, as the nodal analysis succeeds a high level of detail but the decision should be taken on regional scale as the vessel can not survey specific nodes but areas. Hence, the following map was generated in Figure 5.37, which shows the subareas tat should be resurveyed in 2 years. The half size of the area should be resurveyed and in particular 12 out of the 30 subareas of West IJmuiden.

The indicator of NSM was computed using MAC as follows:

1. Iterate through the MAC values at the prediction time series of each node and find the first index where $MAC > 1$, $indNSM$.
2. NSM will be equal to the difference between the prediction time of indNSM and the last surveyed moment; $NSM = T_p(indNSM) - T_{last}$

The NSM indicator was used to show when an area should be resurveyed. The values of NSM were categorized and the map in Figure 5.38 was generated. Three categories were used; NSM = 2 years (pink nodes), NSM = 5 years (red nodes) and NSM = 10 years (green nodes). Applying this indicator we can conclude when it is required to resurvey each area of the West Ijmuiden. For example, the subareas on the right part of the West Ijmuiden should be resurveyed in 2 years (as indicated from the previous maps), while the subareas in the left part of the study area should be resurveyed in 5 years.

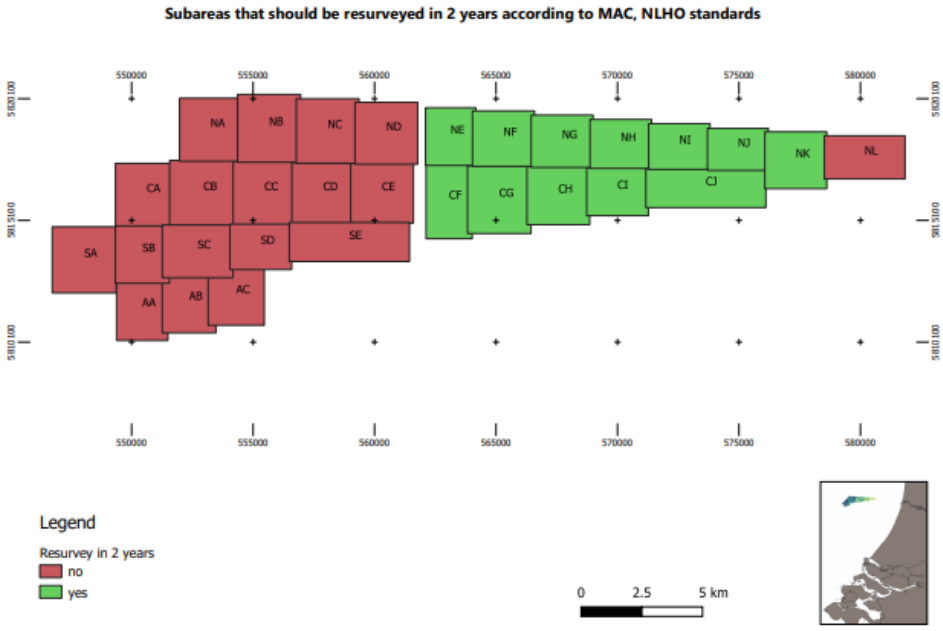


Figure 5.37: Critical subareas of IJmuiden according to NLHO standards and MAC indicator; Green tiles should be resurveyed in two years, whereas red tiles can surveyed in a future prediction moment.

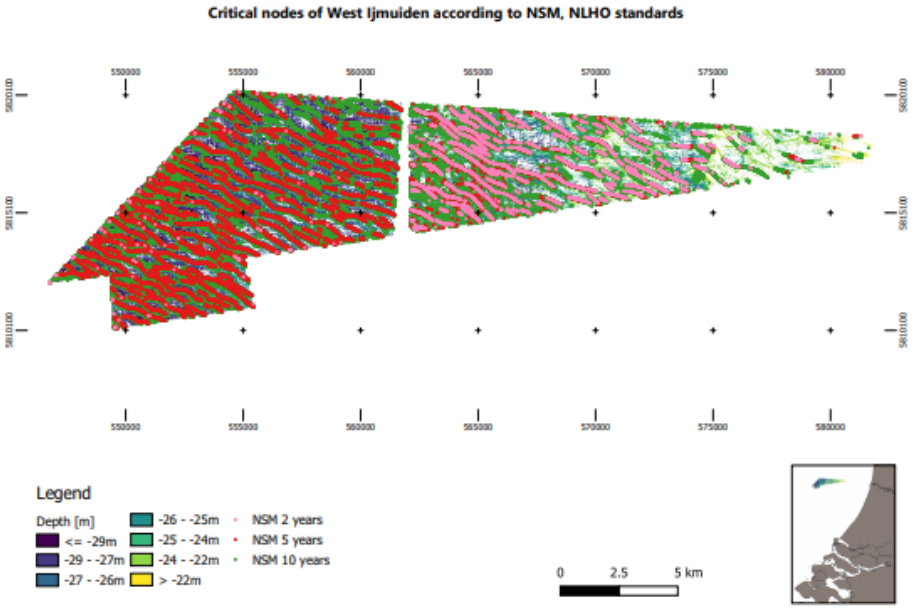


Figure 5.38: Critical nodes according to NLHO standards and NSM indicator, applied for West IJmuiden. Pink nodes should be resurveyed in 2 years, red nodes in 5 years and green nodes in 10 years.

5.4 DISCUSSION

The goal of this project is to perform a quantified assessment of the risk that the seafloor is changing significantly and to present the impact of this in the survey frequencies of the NCS. The results presented here include:

1. a quality and validation assessment of the prediction values that derived from the applied deformation analysis
2. a definition and quantification of the risk, its impact and the probability to occur
3. a validation of the applied method and results with the NLHO standards in order to give an example of the conclusions that could be made from this study regarding the survey frequencies.

The method suggested, is interactive and automated, which means that it could be adapted for all the possible areas of the NCS. During the project, a number of issues were listed to discuss either due to their relevance for future research and similar studies or because they were challenging topics and they need further investigation. In Chapter 3 the last section was devoted for discussing issues which arose regarding the length of the time-series and choice of the study areas. Here, the discussion of the results and the applied methodology will be done. This includes:

Performance of the piece-wise linear model

The piecewise linear model is designed to detect significant depth changes, which is translated to the peaks and drops of the nodal time-series. Despite it was not the main focus of this study, it is worth mentioning that its contribution is relevant. It is introduced by Toodesh and Verhagen in [6] and it adds valuable information regarding the seafloor evolution. It can also be used in dredged areas.

Process of the validation of predictions

For the validation of the prediction values a leave-one-out CV method is used and adjusted to the present time-series analysis. The method can also be applied for other set of prediction values, derived from different functional models. Finally, it deals with the issue of the time dependency, which greatly limits the available options regarding the validation properties.

Relevance of the RDC and direction of change as risk indicators/variables

RDC is a risk indicator and an informative variable for this analysis, despite that it was not possible to use it for the probability estimation during PRA. As already mentioned, it is a single value that could give a sense of the significance of a depth change. It can be computed per node, which means, regarding the relativeness of the depth, that it achieves a high level of detail in the results.

The direction of change is also introduced in this project and its contribution is assumed to be important. In general, the direction of change is associated with different maritime applications and risks; locations that are getting shallower are more interesting for the shipping sector and the probability of a ship to go aground, while the nodes that are getting deeper could cause risks to the pipeline and cable industry as pipelines and cables could be exposed and damaged. Therefore, it was used and included in the probability estimation of RoC and different threshold values are suggested for the positive and negative direction.

Estimate the depth changes as a result of the migrating sand waves

In this project, the vertical - depth changes of the seafloor are studied. There was no initial assumption regarding the presence of sand waves or any other physical phenomenon in the study areas. This means that the method applied - the functional models used, are designed to detect significant or smoother changes of the seafloor, without necessarily trying to physically represent the sand waves. For

example, dredging activities cause also significant depth changes in the seafloor, that could be detected using the piece-wise linear model. Hence, the main goal of this project is related to the depth changes and the risk that they introduce. Areas with the presence of sand waves were chosen as study cases as they are the most dynamic and highly prioritised.

It is known that the sand waves migrate and they cause changes to the submarine terrain and to the seafloor depths. Hence, it is obvious that depth (vertical) changes can be interpreted as a result of the migration of the sand waves. Considering the knowledge that is obtained about the physical modeling of the sand waves and the nodal vertical analysis of the seafloor depths, the relevance of combining the two approaches becomes a challenging but relevant topic.

To study the vertical depth change as a result of the horizontal movement, migration, of the sand waves it is necessary to firstly, define their migration rate and comment on the regularity of their spatial pattern. As it was explained in Chapter 2, considering the available literature, there are different values assigned to the migration rate of the sand waves. For this project, the migration rate of the sand waves in the NF subarea of West IJmuiden was computed based on the method of [25]. The migration rate for the crests and troughs is computed separately. **For the crests, it is estimated to be 3.59 m/year and for the troughs 3.44 m/year.** This difference shows that the crests are moving faster, as it is expected. It is worth mentioning, that for the crests the last 2 surveys are excluded based on the visual output of the results and they are assumed to be outliers.

Now that the migration rate of the sand waves is computed as the mean value and it is possible to estimate the horizontal change of the nodes for each prediction moment. For this purpose, the following equation was used based on the assumption that the migration rate is constant.

$$d_x = MR * dT_p \tag{5.1}$$

The d_x represents the horizontal change of the node due to the migration of the sand waves, MR represents the migration rate and dT_p the prediction time interval, which varies from 1 to 30 years. This was applied for NF subarea and the horizontal changes for the prediction moments are estimated and presented in Figure 5.39. It is obtained that the locations of the nodes that being part of migrating sand waves, it is expected to move horizontally 1m in 3.5yrs and 2m in 7yrs.

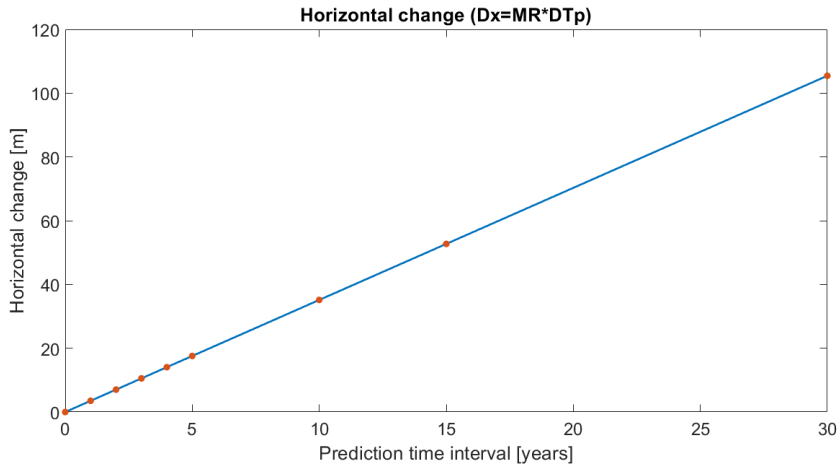


Figure 5.39: Horizontal change with respect to prediction time interval and migration rate of the NF subarea in West Ijmuiden.

After computing the migration rate and the resulted horizontal change per node, it is possible to estimate the depth change due to the migration. This could be compared with the depth change resulted from the real observations to either prove

that the sand waves are responsible for the dynamic areas or to test the regularity of their migrating pattern. However, determining the depth change as a result of the horizontal change can be a challenging and complicated process. The approach used here can be described by applying the computed horizontal change, d_x to the location of a node, shifting it back to the last survey moment and finally, assigning the respective depth value. This process can be seen in Figure 5.40.

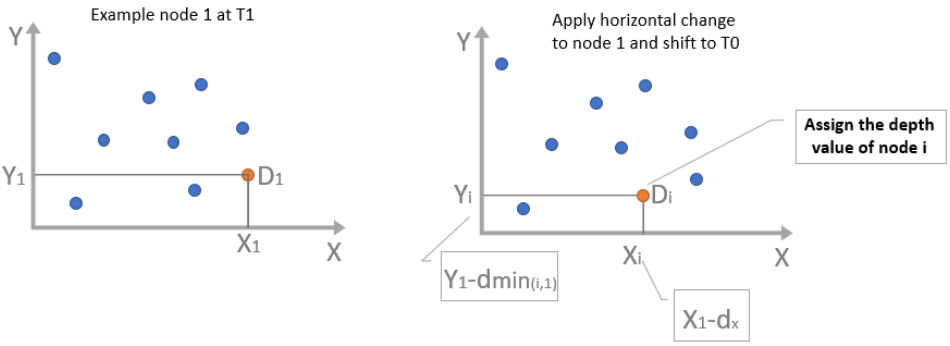


Figure 5.40: Graph representing the process of computing the depth value of an example node (1) as a result of its migration.

D_i is the depth value of the example node (1) at T_0 considering the migration of the sand waves and the corresponding horizontal change. This process is applied for each node and it can be used to focus on the seafloor depth changes as a result of the horizontal change.

6 CONCLUSIONS & RECOMMENDATIONS

The answers of the sub-questions are presented in the first section, which is a necessary step to answer the research question of this study. After this in the second section, the conclusions of the research question are presented and finally, in the third section recommendations are made regarding future related research.

6.1 ANSWERS TO THE SUB-QUESTIONS

To answer the main research question of this study, a set of intermediate sub-questions was addressed. Answering these sub-questions is necessary before concluding regarding the main research question.

1. What is the accuracy and quality of the depth prediction values, that resulted from deformation analysis?

The method of deformation analysis applied in this study, resulted, firstly, in the estimation of parameters of interest of the functional models used. The model selection was implemented by performing hypothesis testing between; the constant velocity model and different forms of the piece-wise linear model. A new nodal time-series of depth prediction values was the product of this procedure. Regarding the accuracy; the measurements assumed to be influenced only by random errors which means that they are normally distributed. The errors propagate from the observations (stochastic model) to the estimations (variance covariance matrix) and thus, the accuracy of the predictions could also be determined by variance covariance law. To assess the quality and reliability of the prediction values, CB were constructed depending on the CI of 95% and 99%.

Taking a closer look to the results, it can be observed, that depth measurements acquired before 2005 are of lower precision (around 30cm), while the observations after 2005 are of higher precision (around 8cm). This is due to the technological enhancement of the SBES to the MBES sonar system. The accuracy of the predictions and the quality assessment proved, that as expected, there is a strong correlation between the desired prediction moment and the precision of the prediction. This means that the far-future predictions are of worse precision and probably they are not that useful. As explained before, predictions were done depending on the 2 functional models used and regarding this, it should be mentioned, that the piecewise linear model provides worse predictions for far-future prediction moments, compared to the constant velocity model. For example, for the nodes presented in the results, it is obtained that at a prediction moment of 2 years after the last survey moment, both models perform with an accuracy of around 0.5m at a 99% CI. Increasing the prediction interval to 10 years, the piecewise model performs with a precision of 1.7m at a 99% CI, while constant velocity model with a precision of 0.7m at a 99% CI. This can be explained by the fact the piece-wise linear model uses only the second segment of the fitted line to perform predictions. It is designed to detect the large depth changes of a node and forecast based on these changes thus, the nodes for which this model is used are the most crucial parts of the study area and these are the crests and the troughs of the sand waves. The same results apply for both the study areas of West IJmuiden and West Rotterdam. The latter resulted

in a bit larger prediction errors (around 1m for 2 years prediction), which can be justified; from the worse precision of the observations until 2005 (80cm), the large and dense gaps in the time-series and the greater amount of nodes that piece-wise linear model was chosen as the most likely one.

A validation of the prediction values was also implemented by performing a leave-one-out CV approach adjusted to the time-series analysis. This was necessary to conclude regarding the reliability of the predictions and the ability of the used functional models to predict accurately. The results were for the West IJmuiden: $CV_e \simeq 0.4m$ and for the West Rotterdam: $CV_e \simeq 0.9m$. The validation errors seem to follow the same order of magnitude that was discussed before and therefore the results can be considered successful. Especially, when taking into account the percentages of the nodes of the training dataset that are inside the CB of the 99% CI, which were computed, and are for both the study areas above 75% of the nodes.

2. How can risk be defined using a probabilistic approach?

North Sea and especially the NCS are of great interest regarding their seafloor terrain and bathymetry. They fulfill a combination of shallow, dynamic and densely used waters. This entails high potential risk in the NCS, considering the occurrence of significant depth changes in relatively short-time scale. It is proved, that seabed dynamics and especially the form of sand waves, could cause large seafloor changes that should be mapped frequently to avoid the severe impact of them remain undetected. Therefore, a definition considering the risk assessed in this study, which is associated to the seabed dynamics should be given and this is:

The risk of the seafloor to change significantly.

However, this is a more general definition and using a probabilistic approach means that the risk should be quantified. Hence, the risk's impact/consequence should be defined as well as the probability of this consequence to occur. The consequence was defined as:

The number of nodes for which quantities related to their vertical dynamics are exceeding a certain threshold value.

The so-called "quantities related to the vertical dynamics" of a number of nodes, denote the measure of risk severity and they are analysed in detail in the answer of the next sub-question. The probability of the consequence to occur is defined as:

The probability that the quantities related to the vertical dynamics of a node are exceeding a threshold value.

These definitions were given to be able to quantify as a step further, the consequence and its probability to happen and finally, perform PRA. The risk of the seafloor to change significantly is closely related to additional risks involved in maritime applications. In this study, a data-driven approach is used and hence, the risk assessed is mainly about the evolution of the nodal time-series of the observed depth values.

3. Which is an appropriate measure of risk severity?

Measuring risk's severity is necessary to understand and compute its impact or consequence, which is a necessary process for the PRA. In this study, this becomes more complicated as the estimations of the prediction depth values are not adequate by themselves to measure depth changes and conclude regarding their significance. Therefore, quantities related to the vertical dynamics of the nodes were determined

and computed per node, these are the risk indicators of RDC and RoC. Additionally, the direction of change plays an important role and it was determined and included in both the aforementioned risk indicators.

RDC was used more as an informative variable, as it is not linear and thus, not normally distributed and it was not used during probability estimation. However, it adds value to the results as it is a single value per node, that can indicate the significance of a depth change without the need for any other quantity to be known. To overcome this issue, the RoC risk indicator was used as the main risk indicator to quantify the consequence and to estimate the respective probability. In general, the risk indicators were computed per node per prediction time as well as their precision and quality using Taylor series approximation.

4. How to go from nodal to regional analysis?

The spatial analysis chosen in this study is on a nodal scale. This was decided, in order to achieve the best possible level of detail and not to make any assumption regarding the presence of sand waves. Additionally, the time-series of depth observations for each node is exclusive and the nodal analysis was decided to be the most appropriate to study the depth evolution through time more precisely. Hence, estimations, predictions and the risk indicators are computed per node.

However, an issue arises regarding the decision-making, which should be generalised on a regional scale. Considering also the expensive operation of the vessels to monitor the seafloor of the NCS, it is necessary to pass from the nodal analysis to a regional scale. Therefore, for the West IJmuiden, the nodal analysis was firstly, generalised in the size of each subarea and then finally, to the complete area as one entity. Maps were created as a final step, using both vector and raster graphics for visualisation purposes. The maps generated from the products of the deformation analysis and the PRA, focusing on the relevant quantities that needed to be mapped to derive conclusions for the study areas. This approach can be used as an example of how to treat large study areas, as the computational complexity of the nodal analysis is very high to perform it directly to large parts of the NCS and to visualize the results. Consequently, it is possible and necessary to pass from the nodal to regional analysis to visualise the complete results and to supply the decision-making.

5. How to validate the current survey frequencies with the results of the risk assessment?

The predictions of depth values derived from the deformation analysis are used as input for the risk assessment, which is based on the risk definition given, and the risk indicator of RoC. RDC and direction of change are used as additional informative risk indicators. The goal of this analysis is to optimize and validate the assigned resurvey frequencies at critical parts of the NCS such as the study areas used here.

It is necessary to combine and adjust the results with the NLHO standards. This is done by using a set of provided indicative threshold values regarding the vertical changes of the seafloor and are used to define and compute two new indicators, the MAC and the NSM. These quantities are computed for the same prediction time with the prioritisation of the areas, in order to compare the results. Maps of MAC and NSM were generated and are presented for the West IJmuiden and West Rotterdam. They provide information related to which areas experience significant depth changes and hence they should be resurveyed and when they should be resurveyed. For West IJmuiden, it was concluded that almost half of the area should be resurveyed according to the current assigned resurvey frequency and for the rest of it the survey frequency could be reduced. For the West Rotterdam the assigned resurvey frequency is justified and remain the same. This is, in general, an automated pro-

cess that could be used to validate the current survey frequency of every area of the NCS.

6.2 CONCLUSIONS REGARDING THE RESEARCH QUESTION

The research question of this study is:

"How to determine the risk of the seafloor to change significantly, and what is the impact on the survey frequencies in critical areas of the NCS?"

In this study, an attempt is made to define and quantify the risk related to significant depth changes of the seafloor in critical areas of the NCS. A data-driven approach is suggested to assess the risk. As input data, a set of prediction depth values are used and therefore, its precision and quality should be provided. The more reliable the predictions the better the performance of the risk assessment. However, this approach could stand alone and be used for other or additional set of prediction values.

The use of the piece-wise linear model proved to be relevant in studying the critical parts of the NCS. Areas with significant depth changes, like the crests and the troughs of the sand waves are indicated using different forms of the piece-wise linear model. Despite the higher variances it presents in far-future prediction moments, compared to the constant velocity model, it is able to detect the peaks or drop in the time-series. The validation of the prediction values tested their reliability and the performance of the models to model reality and to forecast. The results compared to the errors propagating from the original observations are assumed to be sufficient for both the study areas of the West IJmuiden and West Rotterdam. This as a fact, increases the confidence in the risk assessment, as explained before. The risk indicators used in this project can add information to the undergoing research and to the decision-making.

A probabilistic approach is used to determine the risk and supply the design of the survey plan of the NCS. Until now, the current survey policy of the NCS is based on more intuitive manners and the lack of evidence based results ends in areas being over-/under- surveyed and survey plans that are not always feasible to achieve. PRA can serve the optimization and validation of the current survey frequencies of the NCS. It is a rational approach, used to illustrate the impacts of seabed dynamics, the magnitude of depth changes happening, as well as, the probability that these changes will take place in the future. This is a fundamental knowledge in order to assign reliable resurvey frequencies in the different areas of the NCS and justify them.

The method suggested in this project, is automated and hence it could be adapted for all the areas of the NCS. Considering the size of the NCS, the need of updating the plan frequently and the huge amount of data derived from bathymetric surveys, an automated process is necessary. It is also an interactive method, as it gives the user the opportunity to tune the variables involved in PRA. The user can decide; either a desired probability value and compute the corresponding extreme threshold value that should applied to achieve this, or to fix a threshold value that indicates a significant depth change and estimate the corresponding probability that a node will exceed this extreme value in a specific location. To prove the impact of PRA on the survey frequencies, a set of indicative thresholds provided from NLHO is used. These thresholds highlight, what is considered significant depth change. By applying the method suggested, it is concluded that some parts of the West IJmuiden study area and namely 18 out of 30 subareas could be resurveyed with a reduced frequency than the assigned one ($> 2 years$). The depth changes occurring in these areas are below the maximum allowed change according to the threshold values of NLHO. On the other hand, for West Rotterdam it was concluded that half of its

nodes should be resurveyed in the assigned survey frequency and therefore, in this case the survey frequency was justified to kept the same.

The risk of the seafloor to change significantly is related to other additional maritime applications and risks in the NCS. Thus, this risk assessment is not only relevant for optimizing the current survey frequencies of the NLHO, but it can also be useful as a first step to quantify these additional risks. Issues such as the above will be further discussed in the next section as recommendations for future related research and work.

6.3 RECOMMENDATIONS ON FUTURE RELATED WORK

The following parts of the research could be further studied:

"Define the points in the nodal time-series of the prediction values where the direction is expected to change as a result of the migrating wand waves."

The predictions of the depth values follow a certain trend defined by the choice of the functional model for each node. It is proved that as the prediction time increases the prediction values have lower accuracy, which is reasonable. Considering that sand waves are mainly responsible for the depth changes resulted in an area and that there is knowledge regarding their migrating pattern, it would be valuable to define the *tipping points* in the nodal time-series of the prediction values. These are the points that the prediction trend will change direction, from trough to crest or from crest to trough. In Figure 6.1 graphs are made to understand which are these critical points.

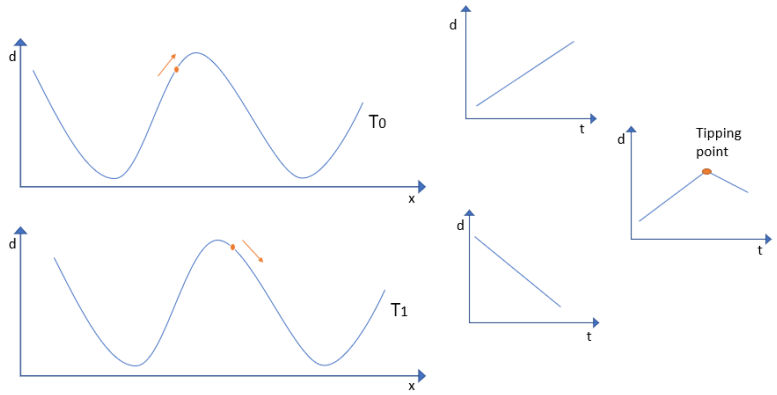


Figure 6.1: An example node at T_0 and T_1 (left), the predicted trends (middle) and the tipping point that should be defined in the prediction time-series of the node (right).

Defining these critical points in the prediction trend will add value to the reliability and use of the prediction values, especially regarding far future time moments. It is suggested to express the migration rate in (rad/year) instead of (m/year) and handle with the *phase* of the sand waves (Figure 6.2). Hence, it would be possible to depict the tipping points

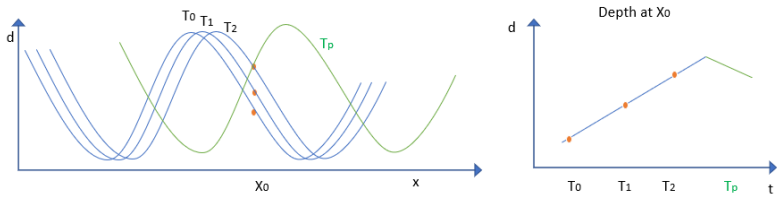


Figure 6.2: Graph that shows how the phase of the sand waves could be used to determine the tipping points of the prediction trend.

"Testing of additional functional models, during deformation analysis or assign different weights to observations and nodes."

The present functional models used, the constant velocity model and the piecewise linear model, provide powerful estimations and predictions of depth values and hence, they are considered sufficient. However, it is almost always possible to improve the mathematical model used in order to fit better the observations. For example, a functional model that includes the change in direction as a parameter of interest could be added or a more simplified version of the piecewise linear model that uses one slope value for the two segment lines could be also useful. Especially, in time-series analysis that new data will be added as the time passes, it becomes necessary to validate the chosen model's performance and tune its parameters if necessary.

Additionally, when having a longer timeseries, it is possible to weigh the observations depending on their precision. Bathymetric surveys done with MBES succeed a better stochastic model and thus, these observations should be weighted more in comparison to the observations before 2005 that are of lower accuracy.

Regarding the nodal time-series, a challenging issue arose was the variable length of the time-series due to the gaps that they have. Apart from the complexity that it adds to the processing of the data, it increases the uncertainty of the results. Nodes with very few past observations are treated the same as nodes with all the possible available observations used. This needs further research considering the prediction procedure. For this reason, it is suggested to investigate the potential to assign weights to the nodes depending on their time-series length.

"Derive additional risk indicators and use them in PRA or optimize the existing ones."

In the PRA of the seafloor to change significantly a definition is given in risk's consequence. For measuring the risk severity, specific risk indicators were determined to describe the vertical dynamics of the seafloor. The most important is the RoC, which is computed per node and used in the probability estimation. If further information regarding the evolution of the seafloor is necessary, as for example, the migration of sand waves or the horizontal uncertainties, new risk indicators can be added in the analysis and add information considering these aspects.

Additionally, direction of change is an important variable for this study and it is recommended to be used also in future studies. It can add valuable information regarding the risk of a ship going aground (positive direction = shallower areas) or the risk that cables or pipelines will be exposed (negative direction = getting deeper). Finally, the RDC is also relevant and further research should be done regarding its contribution and the potential to be used for probability estimation during PRA.

"Combine the risk related to the seabed dynamics with other risks associated to other factors of prioritising the NCS areas, for a comprehensive PRA."

The current survey policy is based on five factors and each of them introduces different risk-alarm levels. It would be of great value to assess each of them separately

and finally, incorporate the risks introduced for a comprehensive PRA. Especially, in the coastal areas this could be very useful, as most of them combine all the factors. They are very shallow and dynamic, they have high shipping intensity, strict draught limitations and are best candidates for offshore projects. Therefore, risks related to each factor should be discriminated and quantified.

"Investigate the probability to find high quality AIS data and combine them with the bathymetric data."

As already mentioned, the risk assessed in this project is tightly connected to other maritime applications such as the navigation of ships or the cable and pipeline industry. To ensure safe navigation the risk of a ship going aground should be quantified. Automatic Identification System (AIS) data could be very useful for this process. AIS data that include the shipping intensity and the draught should be provided or simulated per grid cell in order to use them together with the bathymetric data. Combining AIS data with the available bathymetric data and assessing the associated risks by overlaying the computed probabilities with the probability of a ship going aground, can provide valuable results regarding both the shipping sector and the design of the survey plan of the NCS.

"Integrate the algorithms and the process of map generation into a compatible software application."

Neighboring countries is expected to have the same bathymetry as the one in the NCS and the same issues to overcome. For example UK is one of those countries and it uses an advanced technology regarding the decision-making. It has improved a mapping tool which informs the user regarding the bathymetry of the chosen areas. This application can be used as an example on how to integrate the algorithms created, with a graphic environment.

BIBLIOGRAPHY

[1] P. Doornenbal and S. van Heteren, "Bathymetric range map of the Dutch Continental Shelf (NCP)," *Citeseer*, 2020.

[2] T. Song, A. Main, G. Scovazzi, and M. Ricchiuto, "The shifted boundary method for hyperbolic systems: Embedded domain computations of linear waves and shallow water flows," *Journal of Computational Physics*, 2018.

[3] N. Arjun, "Hydrographic Surveying – Methods, Applications and Uses." https://theconstructor.org/surveying/hydrographic-surveying-methods-uses/13838/, 1999. Accessed: 2020-12-09.

[4] M. Van Ledden, N. Van den Berg, M. De Jong, P. Van Gelder, C. Den Heijer, J. Vrijling, S. Jonkman, P. Roos, S. Hulscher, and A. Lansen, "An idealized meteorological-hydrodynamic model for exploring extreme storm surge statistics in the North Sea," in *ICCE 2014: Proceedings of 34th International Conference on Coastal Engineering, Seoul, Korea, 15-20 June 2014*, Coastal Engineering Research Council, 2014.

[5] L. L. Dorst, *Estimating sea floor dynamics in the Southern North Sea to improve bathymetric survey planning*. NCG, Nederlandse Commissie voor Geodesie, 2009.

[6] R. Toodesh, S. Verhagen, and A. Dagla, "Prediction of changes in seafloor depths based on time series of bathymetry observations: Dutch North Sea Case ," *Remote Sensing (Under Review)*, 2021.

[7] E. Union, "The 2018 annual economic report on eu blue economy." https://ec.europa.eu/maritimeaffairs/sites/maritimeaffairs, 2018. Accessed: 2020-06-20.

[8] EC, "Innovation in the Blue Economy: Realising the potential of our seas and oceans for jobs and growth," 2014.

[9] European MSP platform, "North Sea." https://www.msp-platform.eu/sea-basins/north-sea-0, 2010. Accessed: 2020-07-19.

[10] C. de Jongh and T. Caris, "Bathymetry as a keystone for the blue economy of the netherlands: The implementation of inspire at the netherlands hydrographic service," *Hydro17*, 2015.

[11] T. Van Dijk, C. Van der Tak, W. De Boer, M. Kleuskens, P. Doornenbal, R. Noorlandt, and V. Marges, "The scientific validation of the hydrographic survey policy of the Netherlands Hydrographic Office, Royal Netherlands Navy," *Deltares report*, 2011.

[12] R. Lindenbergh, "Parameter estimation and deformation analysis of sand waves and mega ripples," in *Second International Conference on Marine Sandwave and River Dune Dynamics (MARID2004), Univ. of Twente, Enschede, Netherlands*, Citeseer, 2004.

[13] B. Huizenga, "The interpretation of seabed dynamics on the Netherlands Continental Shelf (NCS)," Master's thesis, University of Twente, 2008.

[14] NLHO, "Hydrographic service." https://english.defensie.nl/organisation/navy/navy-units/hydrographic-service, 2020. Accessed: 2020-14-12.

[15] I. H. Organization, *International Hydrographic Organization Standards for Hydrographic Surveys, 6th Edition, IHO publication No S-44*. IHO, 4b, quai Antoine Ier B.P. 445 - MC 98011 MONACO Cedex Principauté de Monaco, 6 ed., 3 2020.

[16] P. Teunissen, *Adjustment theory: An introduction*. VSSD Press, 2000.

[17] P. J. Teunissen, *Testing theory*. VSSD Delft, 2006.

[18] NOAA, "What is bathymetry." https://oceanservice.noaa.gov/facts/bathymetry.html, 2018. Accessed: 2020-09-15.

[19] P. I. Elhassan, "Bathymetric techniques," *FIG Working Week 2015*, 2015.

[20] C. De Jong, I. Elema, G. Lachapelle, and S. Skone, *Hydrography*. DUP Blue Print, 2002.

[21] W. Xu and K. Ma, "Technology of Sound Velocity Correction for Multi-beam Bathymetry Sounding," in *2018 International Conference on Computer Science, Electronics and Communication Engineering (CSECE 2018)*, Atlantis Press, 2018.

[22] E. Salameh, F. Frappart, R. Almar, P. Baptista, G. Heygster, B. Lubac, D. Raucoules, L. P. Almeida, E. W. Bergsma, S. Capo, *et al.*, "Monitoring beach topography and nearshore bathymetry using spaceborne remote sensing: A review," *Remote Sensing*, 2019.

[23] L. Alexander and R. Ward, "Hydrographic products/services as a fundamental component of the e-navigation concept of operation," *Canadian Hydrographic Association*, 2010.

[24] L. L. Dorst, "Charting a dynamic seafloor: How an excellent survey becomes a poor data-set, and what to tell the mariner," *Lighthouse: Edition 82*, 2014.

[25] J. G. Scheurer, "Bachelor thesis: Algorithm development for detection of sand wave migration from bathymetry data," 2018.

[26] S. J. Wiarda, "Bachelor thesis: Investigating the vertical, nodal dynamics of a sand wave field using the empirical knowledge of sand wave migration rates," 2018.

[27] E. B. Macropedia, *Europe*, p. 832–835. 15th edition: Benton Foundation, 1985.

[28] M. Knaapen, "Measuring sand wave migration in the field. Comparison of different data sources and an error analysis," *Journal of Geophysical Research*, 2005.

[29] P. Blondeaux and G. Vittori, "A model to predict the migration of sand waves in shallow tidal seas," *Continental Shelf Research*, 2016.

[30] S. Jonkman, A. Vrouwenvelder, R. Steenbergen, O. Morales-nápoles, and J. Vrijling, "Probabilistic design: Risk and reliability analysis in civil engineering," *W a*, 2015.

[31] M. Stamatelatos, "Probabilistic risk assessment: What is it and why is it worth performing it," *NASA Office of Safety and Mission Assurance*, 2000.

[32] R. Toodesh and S. Verhagen, "Adaptive, variable resolution grids for bathymetric applications using a quadtree approach," *Journal of Applied Geodesy*, 2018.

[33] S. Pluymaekers, R. Lindenbergh, D. Simons, and J. de Ronde, "A deformation analysis of a dynamic estuary using two-weekly MBES surveying," in *OCEANS 2007-Europe*, IEEE, 2007.

[34] L. Chang and R. F. Hanssen, "A probabilistic approach for InSAR time-series postprocessing," *IEEE transactions on geoscience and remote sensing*, 2015.

[35] M. C. Cuenca, R. Hanssen, A. Hooper, and M. Arikan, "Surface deformation of the whole Netherlands after PSI analysis," in *Proceedings Fringe 2011 Workshop, Frascati, Italy*, 2011.

[36] J. Van Mierlo, "A testing procedure for analysing geodetic deformation measurements," in *2nd FIG Symposium on Deformation Measurements by Geodetic Methods, Bonn*, 1978.

[37] H. Verhoef, *Geodetic deformation analysis: Lecture notes belonging to GE235*. Delft University of Technology, 1997.

[38] A. Verhagen, *The GNSS integer ambiguities: Estimation and validation*. PhD thesis, Delft University of Technology, 2004.

[39] B. A. Swastanto, "Gaussian process regression for long-term time series forecasting," Master's thesis, University of Twente, 2016.

[40] R. Meijer, "Efficient approximate leave-one-out cross-validation for ridge and lasso," Master's thesis, Delft University of Technology, 2010.

[41] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT Press, 2016.

[42] D. der Hydrografie, "Risicoanalyse achterstanden hydrografische opnemingen, peildatum 1 oktober 2007," tech. rep., Tech. rep., Hydrographic Service of the Royal Netherlands Navy, The Hague, 2007.

A APPENDIX

UNCERTAINTIES OF RISK INDICATORS - WEST IJMUIDEN

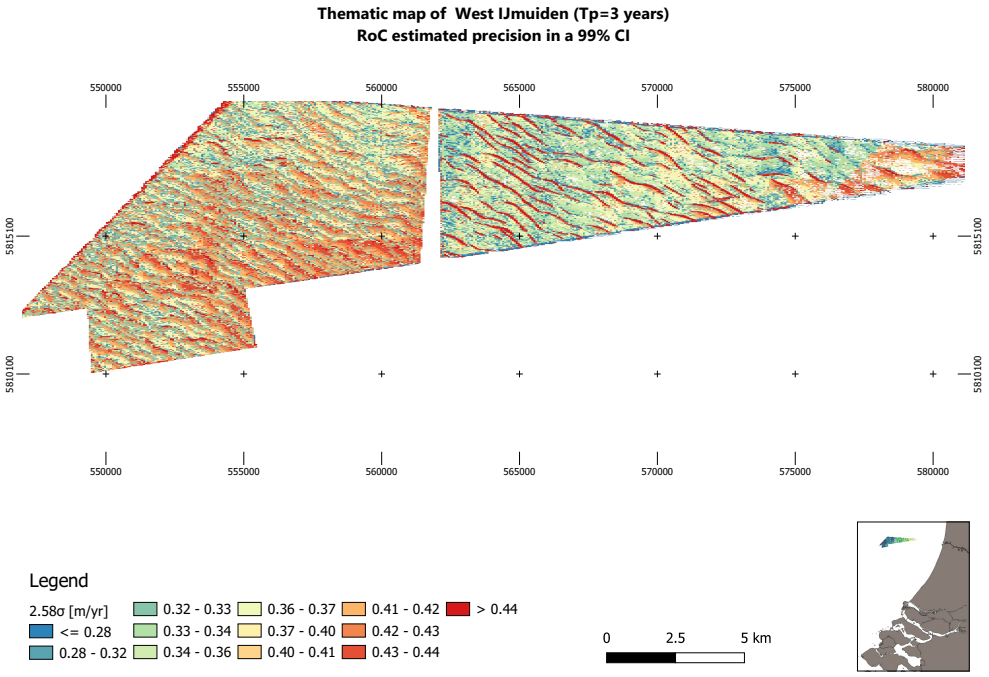


Figure A.1: Map of West IJmuiden with uncertainties for RoC values at a 99% CI.

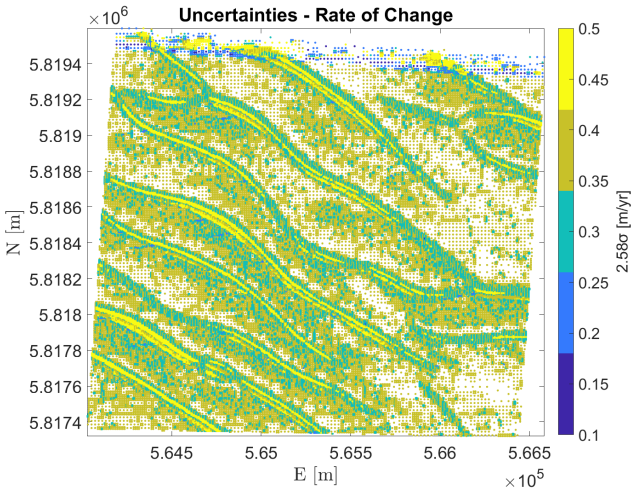


Figure A.2: Uncertainties of RoC values for NF subarea at 3σ, 99% CI.

UNCERTAINTIES OF RISK INDICATORS - WEST ROTTERDAM

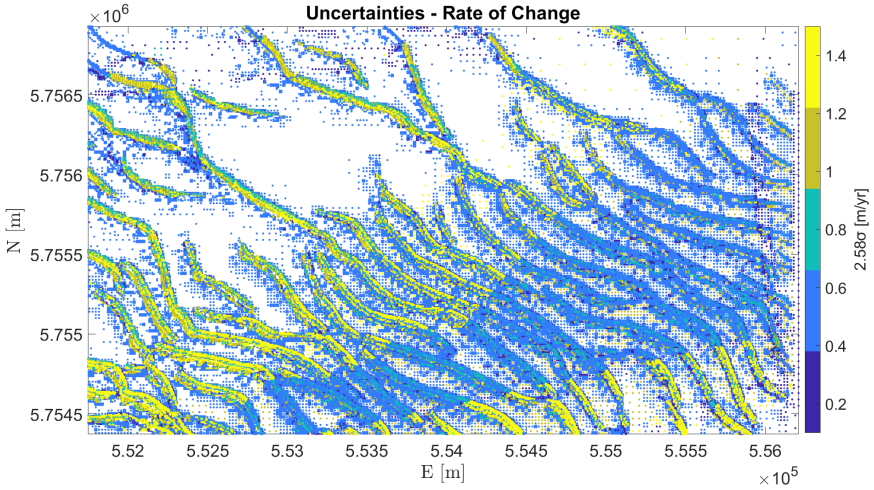


Figure A.3: Uncertainties of RoC values for West Rotterdam at 3σ, 99% CI.