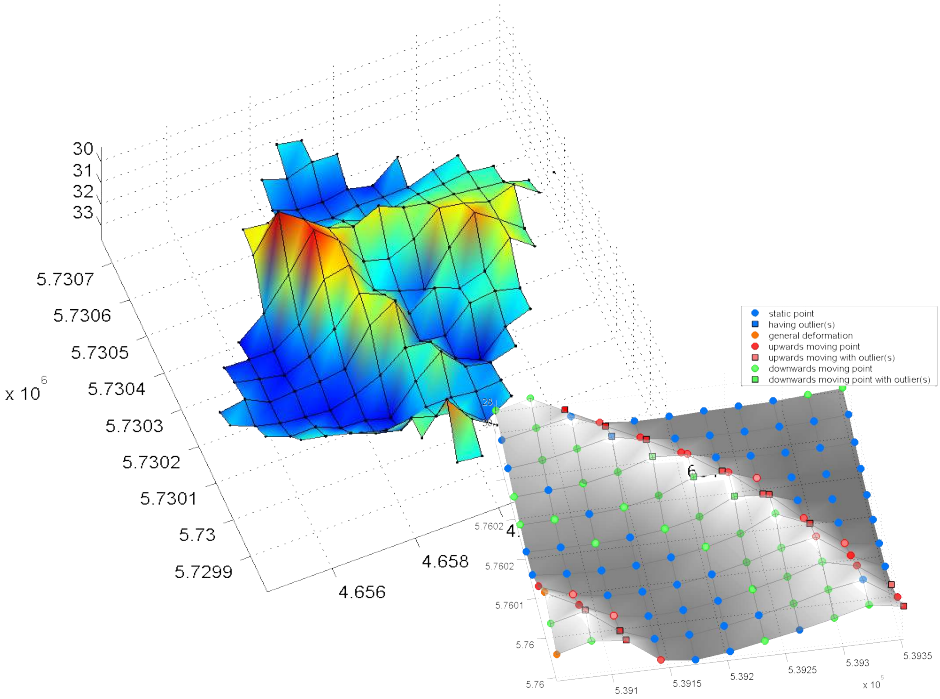
# Detection and prediction of sea floor dynamics




Peter Menting

Section of Mathematical Geodesy and Positioning
Faculty of Civil Engineering and Geosciences
Delft University of Technology

TU Delft
Delft University of Technology

# Detection and prediction of sea floor dynamics

M.Sc. thesis

Peter Menting

September 2004

Section of Mathematical Geodesy and Positioning
Faculty of Civil Engineering and Geosciences
Delft University of Technology

| | |
|---|---|
| Professor: | Prof. dr. ir. P .J. G. Teunissen |
| Supervisor Delft University of Technology: | Dr. R. C. Lindenbergh |
| Supervisor Hydrographic Service: | Ir. L. L. Dorst |
| Supervisor North Sea Directorate: | Ir. J. C. Wüst |

# Preface

This thesis is the result of my graduation research project for a M.Sc. in Geodetic Engineering at the Delft University of Technology. From November 2003 until September 2004 I have been working on detection and prediction of sea floor dynamics. This research is performed at the Section of Mathematical Geodesy and Positioning, which is part of the Faculty of Aerospace Engineering of the Delft University of Technology. Besides, two other authorities were involved in this research: the Hydrographic Service of the Royal Netherlands Navy and The North Sea Directorate of Directorate General of Public Works and Water Management.

For the realisation of this project, I would like to thank some people. First of all, I would like to thank my graduation professor Peter Teunissen for his comments on my research. I especially would like to thank my supervisor Roderik Lindenbergh for always being available to give support and advice about this project. Furthermore, I would like to thank Leendert Dorst of the Hydrographic Service and Hans Wüst of the North Sea Directorate for giving me the opportunity to use their programs and data, and for giving useful comments on this thesis. Finally I would like to thank the members of the MGP section, and especially my room mate Iris Valen, for helping me with various problems and for the company during work and lunch breaks.

Delft, September 15, 2004

Peter Menting

# Summary

The Dutch part of the North Sea is being mapped by two authorities, the Hydrographic Service of the Royal Netherlands Navy and the North Sea Directorate, part of the Directorate-General of Public Works and Transport. Monitoring the sea depth in the Southern North Sea is essential because it is a rather shallow sea and it is heavily used by shipping. To ensure that the main ports, such as Rotterdam, remain accessible, reliable depth information is needed. The sea depth is however not constant, due to e.g. sedimentation the sea floor rises. Another problem is the presence of sand waves. These are regular wave patterns on the sea bed, which have amplitudes of up to 10 meters. Besides, they tend to migrate, so a sand wave entering a sea channel might have severe consequences for the depth. It is however very expensive to survey large parts of the sea on a regular basis, therefore more insight in the dynamic behaviour of the sea floor is required. These two authorities both have developed a strategy to monitor the behaviour of the sea floor by using time series of bathymetric data. However, they do this with different methods and purposes. In this project, both methods are compared and it is investigated whether the method can be combined to benefit from the strengths of both.

The Hydrographic Service is amongst others responsible for the production and upkeep of nautical charts. In order to make a more efficient planning of the surveys, they want to get more insight in the dynamics of the sea floor. Therefore, they have developed a method for analysing time series, based on geodetic deformation analysis. The core of this method is a testing procedure to determine whether the sea floor is static or contains some kind of dynamics. Three different kinds of dynamics are considered: single deviating surveys, general deformation and trends. This procedure consists of three parts: a test for single grid points, a test for a larger area and a prediction for the future, based on the results of the point test. Furthermore, sand wave parameters and sand wave migration are estimated with the area test. This method proves to work very good for the detection of dynamics. The prediction to the future however, is less accurate. Essential here are the Minimal Detectable Biases, the minimal error that can be found with a certain probability. Because the MDB for a single point is rather high, smaller deviations and trends are ignored in the point test and the prediction. For the area test the MDB's are smaller, so the results are more accurate. The advantage of a point test is however the higher resolution: local dynamics can be found as well.

The North Sea Directorate is responsible for maintenance of the sea channels, like the Euro Channel to the port of Rotterdam. For this channel a nautical guaranteed depth is defined, so when the channel becomes too shallow it has to be dredged. To predict the moment when the depth in the channel will rise above a critical depth, they have developed a trend analysis model based on Kalman filtering. In this method a linear growth model is attributed to every point. This growth model consists of two components, a depth value and a linear trend. These depths and trends are used to make a prediction for the future sea depth. The Kalman filter is however rather sensitive for outlying values, because no testing procedure is implemented. Furthermore, the results are very much dependent on the calibration parameters and initial values of the model.

To gain more insight in the performance of both methods and to be able to compare them, the methods are tested on simulated data. Different situations are simulated, e.g. a static sea floor, deviating surveys, trends and migrating sand waves. These data sets are processed with both methods and the results are compared. From these simulations, it can be concluded that the deformation analysis gives better results

when outliers are present, because no outlier detection is included in the Kalman filter. The Kalman filter however gives better results in case of trends, because it takes all dynamics into account, whereas in deformation analysis dynamics smaller than a certain magnitude are missed. Furthermore, neither of the methods can deal with migrating sand waves, because only linear trends are estimated.

To overcome these problems, a combination of both methods is proposed. First a deformation analysis is applied to detect outlying surveys and determine sand wave migration parameters. The detected outliers are used to correct the input data of the Kalman filter. Furthermore, the Kalman filter is extended with a local testing procedure and a sand wave propagation model, based on the parameters found in the deformation analysis.

This combined method is tested on simulated data as well as on real data. On simulated data it proves to work well. For every simulated situation, it results in a correct prediction. On real data it proves to work well in general, although local differences between the predicted depths and the measured depths remain. Main reason for these differences is the used model to describe the sand wave: in this model the sand wave is described by a sine, although the sand wave shape is more complex in reality. But despite these local differences, a combination of both methods proves to give better results than both methods individually.

# Samenvatting

Het Nederlandse deel van de Noordzee wordt gekarteerd door twee instanties, namelijk de Hydrografische Dienst van de Koninklijke Marine en de Directie Noordzee van Rijkswaterstaat. Aangezien het zuidelijk deel van de Noordzee relatief ondiep is en daarnaast druk bevaren, is het essentieel een goed beeld van de diepte te hebben. Om te verzekeren dat belangrijke havens, zoals Rotterdam, toegankelijk blijven is betrouwbare informatie over de diepte nodig. De diepte is echter niet constant, als gevolg van factoren als sedimentatie kan de zee ondieper worden. Een ander probleem zijn zandgolven. Dit zijn regelmatige golfpatronen op de zeebodem, die een amplitude tot 10 meter kunnen hebben. Daarnaast verplaatsen ze zich, dus een bewegende zandgolf kan vergaande consequenties voor de diepte in de vaargeulen hebben. Het is echter erg duur om grote delen van de zee regelmatig op te nemen. Daarom is er behoefte aan meer inzicht in de zeebodemdynamiek. De twee voorgenoemde autoriteiten hebben beide een methode ontwikkeld om de zeebodemdynamiek te monitoren door middel van het analyseren van tijdreeksen van bathymetrische data. Ze doen dit echter met een verschillende methode en ook met een verschillend doel. In dit project worden beide methodes met elkaar vergeleken en wordt onderzocht of het mogelijk is beide methodes te combineren, zodat geprofiteerd kan worden van de sterke punten van beide methodes.

De Dienst der Hydrografie is onder andere verantwoordelijk voor de productie en bijwerking van zeekaarten. Om tot een meer efficiënte planning van de opnemingen te komen, is er behoefte aan meer inzicht in de zeebodemdynamiek. Daarom hebben ze een methode ontwikkeld voor het analyseren van tijdreeksen, gebaseerd op geodetische deformatieanalyse. De kern van deze methode is een toetsingsprocedure om te bepalen of de zeebodem statisch is, danwel bepaalde dynamiek vertoont. Drie verschillende soorten dynamiek worden beschouwd: afwijkende opnemingen, algehele deformatie en trends. De procedure kan uitgesplitst worden in drie delen: een test per gridpunt, een test voor een groter gebied en een voorspelling, die gebaseerd is op de test per gridpunt. Daaarnaast worden in de test per gebied zandgolfparameters en -verplaatsingen geschat. Dit blijkt een goede methode te zijn voor het detecteren van dynamiek, de voorspelling voor de toekomst is echter minder nauwkeurig. Een belangrijke rol spelen de grenswaarden (Minimal Detectable Biases), dit zijn de kleinste afwijkingen die gevonden worden met een zekere kans. Aangezien de grenswaarden voor een enkel punt vrij hoog zijn, worden kleine afwijkingen en trends niet gevonden bij de test per punt en dus ook niet meegenomen in de voorspelling. De grenswaarden voor de gebiedstest zijn kleiner, dus meer dynamiek kan gevonden worden. Het voordeel van een punt test boven de gebiedstest is echter dat ook lokale afwijkingen gevonden worden.

De Directie Noordzee is verantwoordelijk voor het onderhoud van de vaargeulen, zoals de Eurogeul naar Rotterdam. In deze geul is een nautisch gegarandeerde diepte vastgesteld: indien de geul te ondiep wordt, dient er gebaggerd te worden. Om het moment waarop de nautisch gegarandeerde diepte overschreden gaat worden te voorspellen heeft Directie Noordzee een trendanalysemodel ontwikkeld, gebaseerd op Kalman filtering. In dit model wordt in ieder gridpunt een diepte en een lineaire trend geschat. Deze diepten en trends worden gebruikt om de toekomstige diepte van de zee te voorspellen. In het Kalman filter is echter geen toetsingsprocedure opgenomen, waardoor het model relatief gevoelig is voor afwijkende waarnemingen. Daarnaast zijn de resultaten sterk afhankelijk van de gemaakte keuzen voor calibratieparameters en initiële waarden.

Om meer inzicht te krijgen in de prestaties van beide methoden en een vergelijking mogelijk te maken, zijn beide methodes getoetst op gesimuleerde data. Verschillende situaties zijn gesimuleerd, zoals een

statische zeebodem, afwijkende opnemingen, trends en verschuivende zandgolven. Deze data sets zijn met beide methodes verwerkt en de resultaten vergeleken. Op basis van de simulaties kan geconcludeerd worden dat de deformatieanalyse betere resultaten geeft in het geval van afwijkende opnemingen, doordat er in het Kalman filter geen toestingsprocedure opgenomen is. Het Kalman filter geeft echter betere resultaten in het geval van trends, omdat alle dynamiek meegenomen wordt, terwijl in de deformatieanalyse dynamiek kleiner dan de grenswaarde niet gevonden wordt. Voor beide methoden geldt daarnaast dat ze niet goed omgaan met zandgolfverschuivingen, omdat in beide gevallen slechts een lineaire trend geschat wordt.

Om deze problemen te verhelpen, wordt een combinatie van beide methode voorgesteld. Eerst wordt de deformatieanalyse procedure doorlopen om afwijkende opnemingen te detecteren en de zandgolfparameters te schatten. De gevonden afwijkingen worden vervolgens gebruikt om de data te corrigeren, alvorens ze in het Kalman filter gaan. Verder is het Kalman filter uitgebreid met een zandgolfvoortplantingsmodel, gebaseerd op de parameters die in de deformatieanalyse gevonden zijn.

Deze gecombineerde methode is zowel op gesimuleerde als op echte data getest. Het blijkt zeer goed te werken op gesimuleerde data, in alle gesimuleerde gevallen geeft deze methode een juiste voorspelling. Op echte data werkt het over het algemeen goed, alhoewel er lokaal wel aanzienlijke verschillen tussen de voorspelde en de echte data kunnen optreden. De belangrijkste reden hiervoor is het feit dat het model dat de zandgolf beschrijft niet precies past. Hier wordt de zandgolf beschreven door een sinus, terwijl de zandgolf in werkelijkheid een complexere vorm heeft. Maar ondanks deze lokale verschillen blijkt de combinatie van beide methodes betere resultaten te leveren dan beide methodes afzonderlijk.

# Contents

# List of Symbols

In this list all symbols that are used in this thesis are summed in alphabetical order, first the Roman alphabet, then the Greek alphabet and finally the other symbols. Some symbols have different meanings in different contexts.

| | |
|---|---|
| $a$ | range of a covariance function |
| $A$ | model matrix for least squares estimation; amplitude of a sand wave |
| $b$ | sill of a covariance function |
| $c$ | speed of sound in water |
| $C, c$ | design matrix (vector) of a model extension (Chapter 3); covariance matrix (Appendix A) |
| $C(s)$ | covariance model |
| $d$ | depth |
| $D$ | covariance matrix for Kriging |
| $D\{\}$ | variance |
| $\underline{e}$ | vector of residuals |
| $E\{\}$ | expectation |
| $\overline{h}$ | mean amplitude of the heave |
| $H_0, H_a$ | hypotheses in the testing theory |
| $i[k]$ | vector with a one at position $k$ and zeros elsewhere |
| $I$ | identity matrix |
| $J$ | number of support points in the Kalman model |
| $k$ | epoch/survey number |
| $k_\alpha$ | critical value |
| $K$ | number of epochs/surveys (Chapter 3); Kalman gain matrix (Chapter 4) |
| $m$ | number of observations |
| $\underline{m}$ | disturbance vector in the Kalman model |
| $MDB$ | Minimal Detectable Bias |
| $N$ | number of blocks in block discount method |
| $p$ | point number; probability |
| $P$ | number of points |
| $q$ | degrees of freedom of a model extension |
| $Q$ | variance matrix of a stochastic variable |
| $r$ | radius of a circle |
| $s$ | horizontal distance |
| $t$ | time |
| $\underline{t}^k$ | LS test statistic |
| $\underline{T}_q$ | test statistic |
| $\underline{T}^k$ | LOM test statistic |
| $u, v$ | parametrisation of a sand wave |
| $\underline{v}$ | vector or predicted residuals |
| $w$ | interpolation weights |
| $\underline{w}$ | one dimensional test statistic |
| $x$ | unknown parameter; first axis of a frame |
| $\underline{\hat{x}}$ | estimator for $x$ |

| | |
|---|---|
| $y$ | second axis of a frame |
| $\underline{y}$ | vector of observations |
| $z$ | third axis of a frame; arbitrary (unknown) point |
| $\alpha$ | level of significance, type I error |
| $\overline{\alpha}$ | mean slope of the sea floor |
| $\beta$ | beamwidth of a transducer (Chapter 2); type II error (Chapter 3) |
| $\delta$ | discount factor |
| $\Delta$ | difference, deviation |
| $\Delta T$ | travel time of an acoustic pulse |
| $\frac{\Delta}{\Delta t}$ | velocity |
| $\gamma$ | power of a test |
| $\lambda$ | wavelength |
| $\lambda_0$ | non-centrality parameter |
| $\mu$ | distance between support points (Chapter 4); Lagrange parameter (Appendix A) |
| $\sigma_{i,j}$ | covariance between two stochastic variables |
| $\sigma_p$ | standard deviation of $p$ |
| $\sigma_p^2$ | variance of $p$ |
| $\sigma_t$ | standard deviation of the tidal correction |
| $\phi$ | phase of a sand wave |
| $\Phi$ | transition matrix |
| $\chi^2$ | chi-squared probability density function |
| $\psi$ | beam angle (Chapter 2); distance between sand wave crest an a certain reference point (Chapter 3) |
| $^0$ | approximate value |
| $\partial$ | partial derivative |
| $\nabla$ | model extension; Minimal Detectable Bias |
| $\forall$ | for every |

# Chapter 1

# Introduction

The North Sea is heavily used by ships, because of the presence of some of the major harbours of the world, e.g. Rotterdam, Antwerpen and Hamburg. Therefore reliable information on the depths of the shipping routes is necessary for their safe navigation. However, the depth of the sea is not constant: it can rise or fall locally by, for example, sedimentation or migration of sand waves. It is important to monitor these changes as the shipping channel might get obstructed by the rise of the sea floor. This is especially important in the North Sea because of the intense shipping and the fact that it is a rather shallow sea. It is on the other hand very expensive to survey large parts of the sea floor on a regular basis. Therefore it makes sense to monitor and predict the changes in depth, in order to plan the depth measurements more efficiently.

The Dutch part of the North Sea is being mapped by two authorities: the Hydrographic Service of the Royal Netherlands Navy (Dienst der Hydrografie van de Koninklijke Marine) and the North Sea Directorate (Directie Noordzee), which is part of the Directorate General of Public Works and Transport (Rijkswaterstaat). They both use time series of acoustic measurements to monitor the behaviour of the sea floor, but they do this with a different goal and method.

The Hydrographic Service is responsible for the production and upkeep of nautical charts and other nautical publications. Therefore they have to measure the depth of the sea floor regularly. Three criteria are used for measurement planning: depth, number of ships passing and sea floor dynamics. In order to get more insight in the last criterion they have started an investigation on the analysis of time series of bathymetric data on basis of geostatistics and deformation analysis [Dorst, 2003], [Dorst, 2004c]. In order to get data on a regular grid, the data is interpolated using a geostatistical method. After this, the data from different epochs are compared to each other, using deformation analysis. The emphasis hereby is analysis of what has happened in the past.

The main task of the North Sea Directorate is the maintenance of the sea channels, especially the Eurochannel to the port of Rotterdam. Due to migration of sand waves and sedimentation, the channels become shallower and have to be dredged to prevent obstruction of the shipping. They try to predict the behaviour of the sea floor in order to know which area has to be dredged. The prediction method of the North Sea Directorate is based on Kalman filtering. The essence of this method is to predict the future depth on basis of time series of measurements [Wüst, 2003], [Wüst, 2004]. Whenever new data come available, the predicted depth can be updated.

The goal of this thesis is to compare the performance of both methods and to investigate whether it is possible to combine both methods to benefit from the strengths of both of them. Therefore both methods are tested on equal datasets. Two (small) datasets of the Hydrographic Service and one dataset of the North Sea Directorate are available. The datasets from the Hydrographic Service consist of 6 epochs, within the period from 1991 until 2003. The North Sea Directorate dataset consists of 9 epochs, in the period from 1991 until 2001. At both authorities some Matlab-scripts to analyse these data are written;

these scripts are also used in this project.

However, real data is not the most suitable to compare both methods. Therefore the programs are first tested on simulated data. This gives more insight in the performance and shortcomings of the methods. The simulated datasets are analysed with both methods and the results are compared to each other. It is tested whether both methods lead to the same results, if e.g. a local rise of the sea floor can be discovered with both methods. Another aspect is whether a prediction is correct, with both methods a prediction for five years ahead is made, and it is checked whether the results are correct. After the tests with simulated data, a combination of both methods is proposed. This method is first tested on the simulated data and thereafter on real data as well.

After this introduction, Chapter 2 will give some additional background information about sea floor mapping, such as a description of the authorities involved in this process, the echosounding technique and the characteristics of the sea floor. In the next chapters both methods are described, in Chapter 3 the method of the Hydrographic Service and in Chapter 4 the method of the North Sea Directorate. In order to gain more insight in both methods and make a good comparison, some simulations are carried out. The results of that are presented in Chapter 5 and the results of that comparison lead to a suggestion for a combination of the methods in Chapter 6. Finally, conclusions and recommendations will be drawn in Chapter 7.

# Chapter 2

# Sea floor mapping

Although 70 % of the earth surface is covered with water, the sea floor topography is much less known than the land surface topography. The use of accurate nautical charts is however necessary for safe navigation. A big difference between nautical charts and land maps is the possibility of verifying the maps: anyone can check whether a land map gives the correct information, but this is not very easy at sea. Therefore sea maps have to be very reliable; it is invisible to see what is happening. Moreover, a failure in updating or navigation can have enormous consequences for economy and environment, think of accidents. In this chapter a brief introduction to the sea floor mapping process will be given, with a focus on the North Sea. In Section 2.1 the regulations and authorities involved with mapping the sea floor in the Netherlands are introduced. The echo sounding technique, used to obtain depths, is described in Section 2.2 and finally Section 2.3 will focus on the shape of the sea floor, with the emphasis on sand waves.

## 2.1 Sea floor mapping in the Netherlands

The North Sea is very important for the Dutch economy. A lot of economic activities have a direct relationship with the sea. Some examples are fishery, shipping, exploration of oil and gas, tourism and recreation. Besides this, the North Sea is also used for military training, pipelines and cables, extraction of sand and a lot of other businesses. In the future a windmill park and maybe even an airport will be developed in the North Sea.

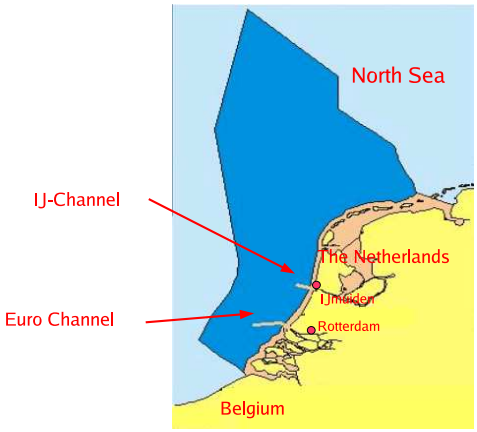



**Figure 2.1:** The NCP (Dutch Continental Shelf) and the two main approach channels. The area in blue is surveyed by the Hydrographic Survey, the area in brown by the North Sea Directorate and the coastal directorates of Rijkswaterstaat.

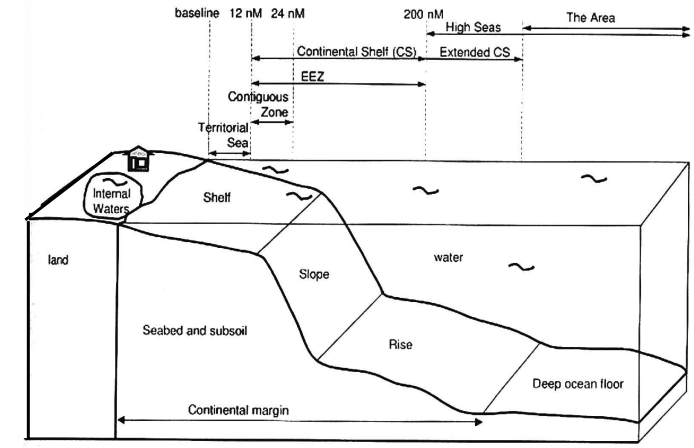



**Figure 2.2:** The various maritime zones used in the law of the sea. Source: [De Jong et al., 2002]

.

The Netherlands take control of a part of about 57,000 square kilometres of the North Sea, about one and a half times the size of the country. This area is called Nederlands Continentaal Plat (Dutch Continental Shelf), abbreviated as NCP, see Figure 2.1.

Although the seas and oceans do generally not belong to a single state, but are in benefit of everyone, there are some restrictions. The international law of the sea is regulated by the United Nations Convention on the Law of the Sea. This convention was first held in 1958, with 86 participating countries. Nowadays the law of the sea is defined by the UNCLOSIII. UNCLOS means United Nations Convention on the Law of the Sea. This convention has come into operation in 1994. Currently 145 states have signed this convention, among which also some landlocked countries like Switzerland and Afghanistan [UNCLOS, 1982].

The main aspect of UNCLOSIII is the division of the sea in several zones where different rights and duties are in force. The width of the zones is measured from the baseline. This is normally the low water line along the coast, although there are some exceptions, for example in case of bays and archipelagoes. In Figure 2.2 the various zones are illustrated. A more extensive description of the law of the sea is given in [De Jong et al., 2002].

The most important zones are the territorial sea and the Exclusive Economic Zone (EEZ). The territorial zone extends to at most 12 nM (1 nM = 1.852 km) from the baseline. In this zone, the coastal state has full sovereignty over the sea floor, subsoil, water and airspace. However, foreign vessels have the right of innocent and transit passage. The Exclusive Economic Zone (EEZ) extends to a limit of 200 nM from the baseline, this generally overlaps with the continental shelf. Within the EEZ, the coastal state has sovereign rights of exploration, exploitation, conservation and management of all resources, living (e.g. fish) as well as non-living (e.g. gas). The exploitation of the sea also includes hydrographic surveying. Furthermore, the coastal state has jurisdiction on the establishment and use of structures, artificial islands and research. The main difference with other zones is that the EEZ has to be claimed.

Because the width of the North Sea is less than 400 nM, the sea is divided among several countries. In case of the Netherlands, the EEZ coincides with the NCP, the continental shelf. As was mentioned before, the Netherlands therefore have full sovereign rights of exploration of this part of the North Sea, which includes the hydrographic surveying. In the Netherlands, the Hydrographic Service ("Dienst der Hydrografie") of the Royal Netherlands Navy is responsible for charting the North Sea. Parts of the North Sea are however surveyed by the North Sea Directorate ("Directie Noordzee") and the coastal directorates of the Directorate-general of Public Works and Water Management ("Rijkswaterstaat"). These

organisations are united in the NHI, the Dutch Hydrographic Institute.

The coastal directorates of Rijkswaterstaat are surveying the coastal area. They approximately survey the area until a depth of 12 m below NAP. The main task of the North Sea Directorate is the maintenance of the sea ways, for example dredging. There are two main sea ways in the Dutch part of the North Sea, the Euro Channel and the IJ-Channel, see Figure 2.1. Rijkswaterstaat make use of a multibeam echo sounder for acquiring depth information of the sea. The characteristics of such a system are discussed in Section 2.2.

The Hydrographic Survey surveys the remaining part of the NCP, so everything deeper than 12 m below NAP and outside the main sea ways. Their main tasks are mapping the sea floor and publication of sea maps of the NCP and the waters around the Netherlands Antilles and Aruba. Because it is part of the Royal Netherlands Navy, they also have specific military tasks. Until recent the Hydrographic Survey only had a single beam echo sounder available, but they have started to make use of a multibeam system.

## 2.2 Echo sounding technique

Echo sounding is a technique to measure depth from e.g. a ship. It is based on the observation of the travel time of acoustic waves. The acoustic pulse is transmitted from the ship and travels through the water. When the pulse arrives at the sea floor, it gets reflected and the pulse returns to a receiver on the ship. From the observed travel time of the pulse the depth can be calculated as:

$$d = c \cdot \frac{\Delta T}{2}, \tag{2.1}$$

where $d$ is the depth, $c$ is the speed of sound in water and $\Delta T$ is the time delay between the transmission and reception of the acoustic pulse.

An echo sounder usually consists of the following components, as illustrated in Figure 2.3:

- A transmitter, which generates the sound pulses;
- A transmitter/receiver switch, which passes the power from the transmitter to the transducer;
- A transducer, which converts the electrical power into acoustic power, sends and receives the signal and converts the acoustic returned signal back into an electrical signal;
- A receiver, which amplifies the returned signal;
- A recorder, which measures the two-way travel time, stores the data and calculates the range from the time interval.

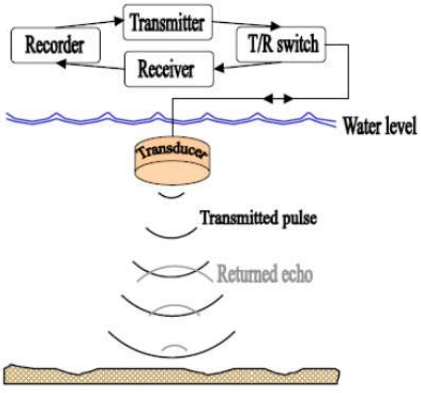



**Figure 2.3:** Overview of an echo sounding system. Source: [Hennis, 2003].

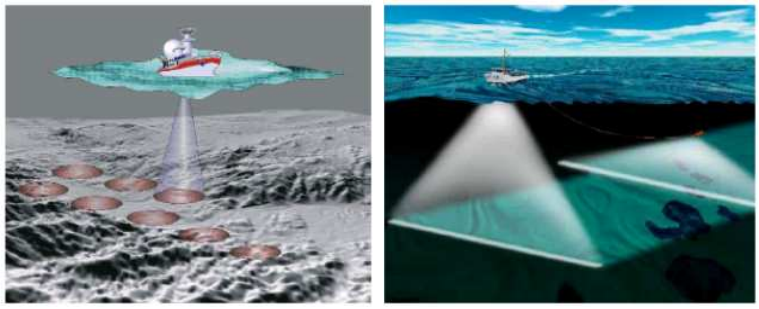


**Figure 2.4:** Echo sounding systems. On the left: Single beam echo sounding. On the right: Multibeam echo sounding and a side scan sonar mounted on a towfish, behind the vessel.

Three different kinds of echo sounding systems are currently in use: single beam echo sounders (SBES), side scan sonars (SSS) and multibeam echo sounders (MBES). An illustration of these concepts is shown in Figure 2.4.

### Side Scan Sonars

The SSS system can be compared with a Synthetic Aperture Radar system, the only difference being the use of sonar (Sound Navigation and Ranging) instead of radar (Radio Navigation and Ranging). It is mainly used for detection of objects and structures on the sea floor, e.g. shipwrecks. Usually the SSS is mounted on a towfish, a small 'fish' which is towed by the vessel. In principle Side Scan Sonar cannot be used for bathymetry because only the range is measured, no direction. This can however be solved by using SSS in an interferometric way.

### Single Beam Echo Sounders

As will be obvious, the Single Beam Echo Sounder (SBES) system consists of a single acoustic beam. The transducer is usually mounted on the hull of the ship and is directed vertically down. Therefore, only a narrow profile right under the ship will be recorded. The size of the footprint of a single beam depends on the beamwidth and the depth: $r = d\beta$, where $\beta$ is the beamwidth of the transducer. So with a beamwidth of 5 degrees and a depth of 30 m (which is the average depth of the southern North Sea), the size of the footprint is about 2.6 m.

### Multibeam Echo Sounders

Two types of Multibeam Echo Sounders (MBES) exist: sweep systems and swath systems. A sweep system consists of an array of single beam echo sounders, mounted on booms on the sides of the ship. Most currently operating systems however, are swath systems. A swath system transmits an acoustic pulse that is very wide in the across-track direction and narrow in the along-track direction. The transducer, which receives the backscattered echo, segments the echo into multiple smaller beams. The width of these beams is equivalent to the beamwidth of a single beam echo sounder. Therefore the measurement accuracy of a multibeam echo sounder is not better than the accuracy of a single beam echo sounder, in fact it is often worse, because of the oblique beams. The main difference is the amount of data that is acquired at one moment. For a single pulse transmission (which is called "ping"), a broad band of depths is obtained instead of a single value. The coverage (footprints) of a single ping in a multibeam system is illustrated in Figure 2.5. The MBES can be mounted on the hull of the ship, as well as on a remotely operated vehicle.

In a multibeam system the depth is calculated as:

$$d = c\frac{\Delta T}{2}\cos\psi. \tag{2.2}$$

Compared to Equation 2.1 the difference is the additional measurement of $\psi$, the beam angle. This is the angle between the vertical and the actual beam direction. As can be seen from Figure 2.5, the footprint

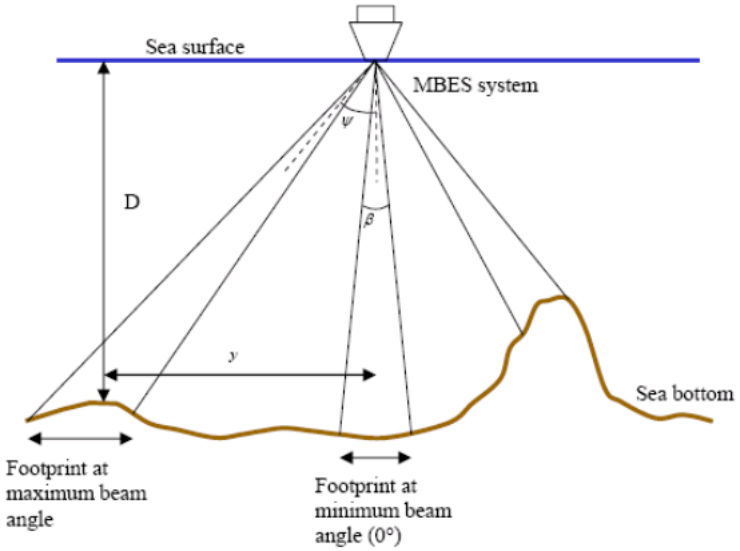



**Figure 2.5:** Footprints of a multibeam system. Source: [De Jong et al., 2002].

is not constant over a ping, but depends on the beam angle ($\psi$), the beamwidth ($\beta$) and the depth of the water column ($d$).

An extensive list of error sources for multibeam measurements can be made. Among the most important are the vessel's roll, pitch, yaw (heading) and heave parameters. They are measured as well, because they are an important error source. Roll is the rotation around the along-track axis (x-axis) of the vessel, pitch is the rotation around the across-track axis (y-axis), yaw the rotation around the vertical axis (z-axis) and heave is the vertical motion, compared to a reference plane. These concepts are illustrated in Figure 2.6. The attitude parameters are measured in real time by a motion sensor. Besides the attitude errors, other error sources exist such as system electronics, uncertainty in the sound speed, the tidal reduction, measurement errors due to the beamwidth and beam angle errors. In Paragraph 3.1.1 the effect of these error sources to the depth measurements is elaborated for the surveys of the Hydrographic Service.

## 2.3 The sea floor morphology

The bed of many shallow seas with a sandy bed, such as the southern North Sea, is not flat but shows various regular patterns such as sandbanks and sand waves [Németh, 2003]. In Table 2.1, some characteristics of these patterns in the North Sea are given. These patterns have in common that they are approximately harmonic in one direction, while elongated in the other direction. The actual seabed is sometimes a complex structure of combinations of those patterns.

The largest structures are the tidal sandbanks, which can have wavelengths of several kilometres and heights up to the water depth. The crest of the tidal sandbanks are generally oriented at about 30°

| | **wavelength [m]** | **max. ampl. [m]** | **migration rate** | **dir. w.r.t. tide [deg]** |
|---|---|---|---|---|
| Rippels | ± 1 | ± 0.01 | 1 m/hour | |
| Mega-rippels | ± 10 | ± 0.5 | 1 m/day | 0-20 |
| Sand waves | ± 550 | ± 10 | 10 m/year | 70-90 |
| Long bed waves | ± 1,500 | ± 5 | 1 m/year | 50-60 |
| Tidal sandbanks | ± 5,000 | ± 10 | 1 m/year | 0-30 |

**Table 2.1:** Typical orders of magnitude of the various patterns on the bed of the North Sea [Knaapen, 2004]

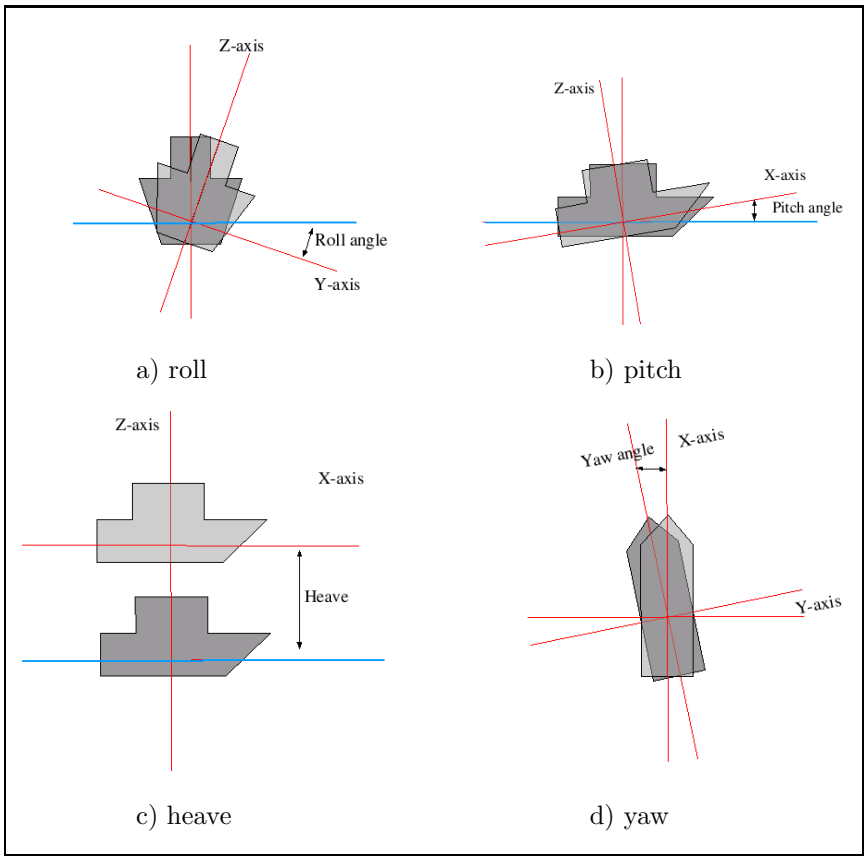



**Figure 2.6:** The four attitude parameters of the vessel.

respective to the tidal current. Various factors, such as the horizontal tidal current and sediment transport account for their existence. Furthermore it is believed that their location is also controlled by older morphological features, which act as a core for the tidal sandbank growth [Knaapen et al., 2001].

Long bed waves are first mentioned by [Knaapen et al., 2001] and have wavelengths between the sandbanks and sand waves. Long bed waves are about 1,600 m long and are oriented at about 60° with respect to the tidal current axis (which is in this area approximately oriented from South West to North East). The mechanism that generates these waves is not yet known.

Sand waves are probably the most well-known structures, almost the entire southern part of the North Sea is covered with sand waves [Hulscher and van den Brink, 2001]. They have a wavelength of several hundreds of meters and amplitudes of up to 10 m. Sand waves are generally oriented perpendicular to the tidal current, which is also the main mechanism in the generation of sand waves.

The sand waves are often covered with mega-rippels. With an amplitude of max. 0.5 m and a wavelength of about 10 m, these features are rather small-sized. They are however moving quite fast, whereas sand waves are relatively stable. The highest mega-rippels are found at the sand wave crests, with amplitudes of up to 2 m. In the troughs the amplitudes are close to zero. The direction of the crests of the mega-rippels is often perpendicular to the direction of the sand wave crests [Lindenbergh, 2004].

A large part of the southern North Sea is covered with these patterns, especially sand waves. These sand waves are not stable, but can migrate at a rather short time. The height of a sand wave can be up to 30 % of the water depth, so migrating sand waves significantly decrease the water depth. Because sand waves

are present in the main shipping routes, this can be a serious threat to the safety of shipping. Another problem is the exposure of pipelines. A pipeline has to be below the sea floor, but with a migrating sand wave they may become exposed and this may cause the pipeline to buckle or break [Németh, 2003]. Therefore monitoring the evolution and migration of sand waves is necessary.

Various models are (being) developed to predict the behaviour of sand waves. In general two approaches are possible: the emphasis can be lead on measurements or on the mechanisms that generate the sand waves. An example of the first approach is described in [Knaapen, 2004], where the sand wave migration rate is determined from the change in position of the crest. This is determined by comparing a series of bathymetric data. An example of the second approach is described in [Németh, 2003]. Here a model is developed to predict sand wave behaviour based on the physical mechanisms that may cause sand waves to migrate, The main mechanism are tidal currents, which causes sediment transport and evolution of sand waves.

# Chapter 3

# Method 1: Deformation analysis

In this chapter, the method of the Hydrographic Service of the Royal Netherlands Navy will be described. The main task of the Hydrographic Service is the ensurance of safe navigation at sea, by mapping of the sea floor and the production of nautical publications, like charts. For the planning of the surveys, they use three criteria:

- The sea floor depth,
- The frequency of ships passing,
- The dynamics of the sea floor.

The main area of interest is the so called "selected track", which consists of the seaway to the port of Rotterdam and the anchorages. In the selected track some areas are defined as "critical areas" which are the shallowest parts, see Figure 3.1. In this project, the data of the critical areas B and G are analysed.

In order to make the planning of surveys more efficient, the Hydrographic Service has started an investigation of the use of time series of bathymetric data [Dorst, 2003], [Dorst, 2004c]. The idea is that by analysing time series of bathymetric data, it is possible to gain more insight in the last criterion, the dynamics of the sea floor. The method for analysing the changes between different epochs is called deformation analysis. This method will be explained in Section 3.2. Before this, some preprocessing steps have to be performed. This consists of calculation of the standard deviations of the measurements, determining covariance functions and interpolation to a regular grid. These steps will be described in Section 3.1. An example of the deformation analysis will be given in Section 3.3. Finally in Section 3.4 a summary of this method will be given, together with a short evaluation of the method.

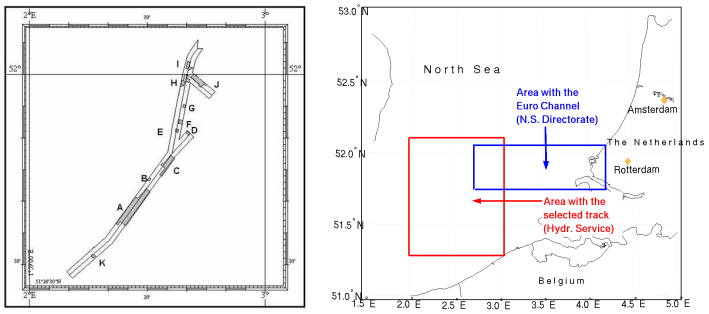



**Figure 3.1:** On the left side the selected track with the critical areas A until K. On the right side the the location of the selected track in the North Sea. Also illustrated is the Euro Channel, where the North Sea Directorate is active (Chapter 4). The left graphic is taken from [Dorst, 2004a].

## 3.1 Statistics of the sea floor

In this section three preprocessing steps that are necessary to perform the deformation analysis will be described.

- First, the calculation of the standard deviation of the measurements is considered in Paragraph 3.1.1. The deformation analysis is based on the principle that a measurement never gives the exact value, but has a certain standard deviation. It is tested whether or not a difference in depth between two epochs can be explained by a real change in depth, or is simply a result of the measurement inaccuracy.
- Secondly the determination of the covariance function is described in Paragraph 3.1.2. The covariance function is needed for two purposes: for the interpolation of the surveys and to determine the sand wave length, in the case that sand waves are present.
- Finally in Paragraph 3.1.3 a short description of the interpolation of the surveys is given. The interpolation is necessary because in general, results of a single beam survey are not on a regular grid, but follow the pattern of the surveyed tracks. However, to perform the analysis of the time series the depth value of the same point has to be known in several epochs. For multi beam data this step is not necessary, because multi beam data usually already fully covers the sea floor. However, for the purpose of data reduction an interpolation step can still be useful.

### 3.1.1 The variance of sea floor measurements

To test whether a difference in depth is significant, the variance of a single depth measurement has to be estimated. Several factors have to be taken into account to get a realistic value for the variance. In Table 3.1 the various factors are listed, together with their values. These values are based on experience, research and documentation of the Hydrographic Service and only valid for the single beam surveys of the Hydrographic Service. For a more thorough explanation see [Dorst, 2004c]. According to this table, the total variance $\sigma_p^2$ of a depth value at point $p$ can be expressed in as the sum of those components:

$$\sigma_p^2 = 1.4 \cdot 10^{-6} d_p^2 + 5 \cdot 10^{-5} d_p + 0.01 \overline{h}^2 + \sigma_t^2 + (\overline{\alpha}/100\%)^2 \cdot 6.25 + 0.0015\ m^2, \tag{3.1}$$

where $d_p$ is the depth at point $p$, $\overline{h}$ the mean amplitude of the heave and $\overline{\alpha}$ the mean slope of the sea floor in the analysed area. The precision of the tidal correction is expressed by $\sigma_t^2$. This correction is either based on tidal measurements or on a tidal prediction model. The value of $\sigma_t^2$ is based on experience and can be used for both methods if no actual values are available. The mean heave $\overline{h}$ is fixed at 0.6 m, because the actual heave values are not known for all measurements. When the heave, roll and pitch are too high, due to bad weather, no measurements will be taken. The variance of the horizontal positioning is a standard value for the use of DGPS receivers, propagated to the depth.

In principle, it is assumed that the measurements are normally distributed. However, the Hydrographic Service does not store the average values of a grid cell, but the shallowest value. Therefore the assumption of normally distributed data does not hold anymore, a value that is too shallow will be present more often than a value that is too deep. The expectation will be too shallow, so a systematic error will be introduced. When the measurement accuracies are comparable, this systematic error will however be the same for all epochs. If the other effects on the variance (see Table 3.1) are the same for all epochs as well, this will not be a problem. The depths from several epochs can still be compared. The advantage of using the shoalest value in a grid cell is that the real depth will never be less than the depth that is stored. When storing the average value, the actual depth could be less than the depth that is given on a chart, which may cause a danger.

Applying Equation 3.1 with a mean depth of about 35 m for the North Sea, a heave of 0.6 m and a mean slope of 5 %, the variance of a measurement can be calculated as:

$$\sigma_p^2 = 0.034\ m^2. \tag{3.2}$$

| Stochastic variable | Variance ($\mathbf{m^2}$) |
|---|---|
| Measuring instrument (SBES) | $(0.001\ d_p + 0.025)^2$ |
| Digibar | $4 \cdot 10^{-7} d_p^2$ |
| Heave compensator | $(0.1 \cdot \overline{h})^2$ |
| Draught | $6.25 \cdot 10^{-4}$ |
| Velocity dependent sinking | $2.25 \cdot 10^{-4}$ |
| Tidal correction | 0.01 |
| Pitch and roll | negligible |
| Slope of the sea floor | negligible |
| Idealisation of the sea floor | negligible |
| Change of depth during the measurement | negligible |
| Horizontal positioning | $(\overline{\alpha}/100\%)^2 \cdot 6.25$ |

**Table 3.1:** Influences on the variances of a depth value $d_p$ at position $p$. Here $\overline{h}$ is the mean heave amplitude and $\overline{\alpha}$ is the mean slope of the sea floor. These values are only valid for the single beam surveys of the Hydrographic Service. When other equipment is used, these values will be different.

The standard deviation of a depth value is the square root of the variance and thus is about 18 cm. This accuracy is bound to the S44 standards of the International Hydrographic Organisation [International Hydrographic Organization, 1988]. For areas like this, the minimal standard deviation is defined as:

$$\sigma_p^2 = \sqrt{0.5^2 + (0.013 * d_p)^2}. \tag{3.3}$$

For a sea floor model, that is an interpolated sea floor, the requirement is lower. In that case the S44 standard is defined as:

$$\sigma_p^2 = \sqrt{1 + (0.026 * d_p)^2}. \tag{3.4}$$

For a depth of 35 m this results in a minimal standard deviation of 1.35 m, so the measurement accuracy of the Hydrographic Service is well below the S44 standard. Note however that Equation 3.1 is only valid for the single beam surveys of the Hydrographic Service, other equipment and corrections yield different values.

### 3.1.2 The covariance function

The covariance function is needed to determine the sand wave length and is a tool in the interpolation procedure. Mathematically, covariance between two stochastic variables $\underline{z}_i$ and $\underline{z}_j$ is defined as [Polman and Salzmann, 1996]:

$$\sigma_{i,j} = E\left\{(\underline{z}_i - E\{\underline{z}_i\})(\underline{z}_j - E\{\underline{z}_j\})\right\}, \tag{3.5}$$

where $E\{\underline{z}_i\}$ and $E\{\underline{z}_j\}$ are the expectations of the variables $\underline{z}_i$ and $\underline{z}_j$. The underscore denotes that the variables are stochastic.

A sea floor can be described by a set of data points $z_i$ with an $x$,$y$ and $d$ value, where $x$ and $y$ denote the position in the horizontal plane and $d$ is the depth at point $(x, y)$. In this case the covariance can be used to describe the correlation between two data points $z_i$ and $z_j$, separated from each other by a horizontal distance $s$. By calculating the covariance at specified distance intervals, an empirical covariance function is obtained. This results in a discrete function, with covariance values only at discrete distances. For practical use however the covariance at all distances is needed. To solve for this, an analytical covariance function is fitted through the discrete covariance values. The analytical covariance function can be characterised by the following parameters:

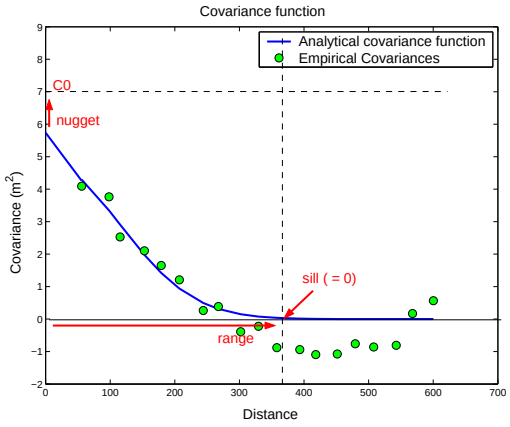



**Figure 3.2:** An example of a covariance function. The green circles illustrate the empirical covariance function, the blue line is the analytical model that is fit to the empirical covariances. The covariance function can be described by the parameters C0, range and sill.

- *Nugget* or *C0*: This provides a discontinuity at the origin. In the nugget, the measurement inaccuracies are expressed.
- *Range*: This provides the distance beyond which the covariance remains essentially constant. Therefore the range can be used to determine the search neighbourhood for interpolation. If the covariance does not reach a constant value, there is still a trend in the data.
- *Sill*: This is the covariance value at range distance.
- *Wavelength*: In case of a periodic signal, the wavelength can also be determined from the covariogram because the covariance is not only dependent on the distance but on the phase of the function as well.

When the sea floor is not isotropic, that means the sea floor does not have the same variability in all directions, two covariance functions are calculated, because the correlation is not the same is all directions. First the direction with the highest variability is determined by analysing the gradients of the data [Dorst, 2004c]. The direction with the highest gradient shows the highest variability, which is often the sand wave direction. Then one covariance function is calculated in the direction with the highest variability and one perpendicular to that direction. Through the deterministic covariance values, an analytical function is fitted. Two different models are used, first a *Gaussian model*, which is defined as [Chilès and Delfiner, 1999]:

$$C(s) = b \cdot exp\left\{-\left(\frac{s}{a}\right)^2\right\}, \tag{3.6}$$

where $s$ is the distance, $a$ the range and $b$ the sill.

Second a so called *hole-effect model* is used. This is a sinelike function and is used in case of sand waves. It is defined as [Chilès and Delfiner, 1999]:

$$C(s) = b \cdot exp\left\{-\frac{s}{a}\right\} \cos\left(\frac{s}{\lambda} + d\right), \tag{3.7}$$

with $b$ the sill, $a$ the range, $\lambda$ the wavelength and $d$ an additional scaling parameter.

In this study two critical areas are used, area B and area G. Area B does not have a clear sand wave pattern but is not isotropic either. Area G does show a sand wave pattern. The Matlab software for calculating the covariance function is obtained from [Gratton and Lafleur, 2001]. In Figure 3.3 the covariance functions of the areas B and G are shown. For area B it shows that the pattern is not entirely

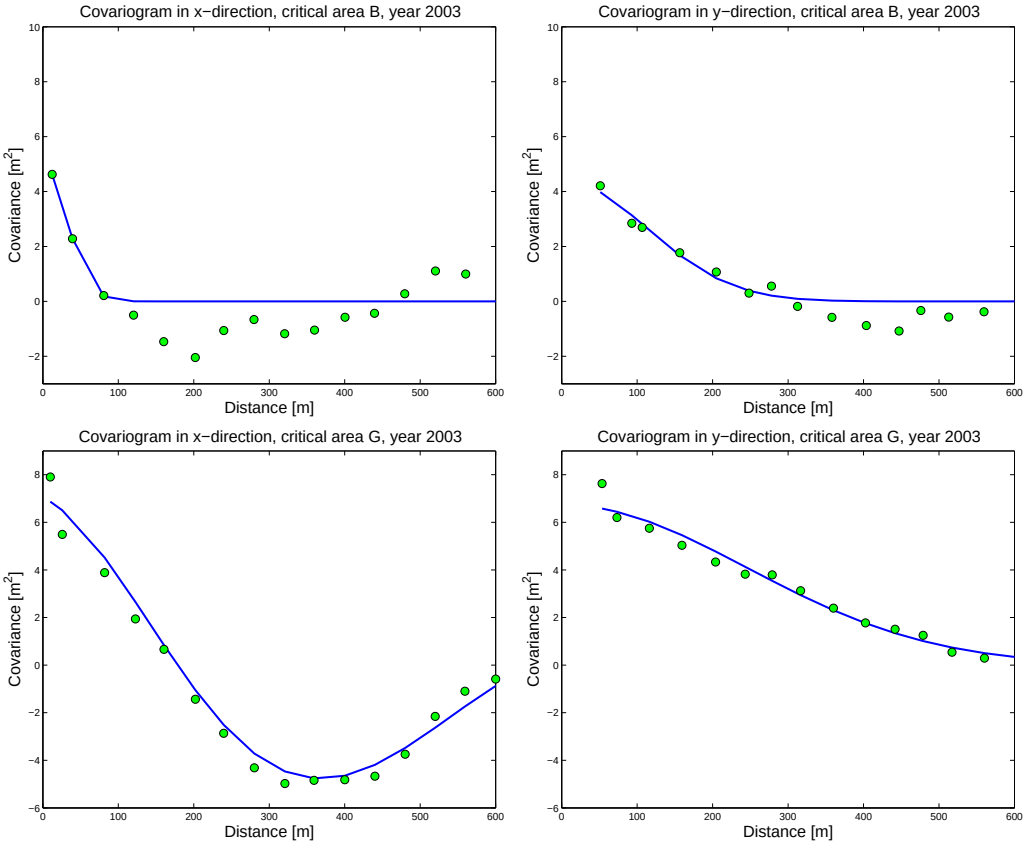



**Figure 3.3:** The covariance functions in $x$- and $y$-direction of the critical areas B (top) and G (bottom). The green circles are the deterministical calculated covariances, the blue line is the estimated analytical covariance function. For area B the same covariance function is estimated in both directions, for area G the function on the left is in sand wave direction, the function on the right is perpendicular to that direction.

the same in both directions, but does not show a clear sand wave either. The covariance function in $y$-direction fits well to the data, but in $x$-direction the fitting performs worse. This is caused by the fact that although no sand waves are present, the sea floor is not isotropic either. Therefore the Gaussian model that is used in $x$-direction is not really suitable for this area. For area G two different covariance functions are estimated. The covariogram in $x$-direction is in the sand wave direction, the $y$-direction is perpendicular to that direction. From the covariance function also the sand wave length is estimated, for this area this is about 800 m. In the next section the results of the interpolation of these areas are shown.

### 3.1.3 Interpolation of the surveys

In general, results of a single beam survey are not on a regular grid, but follow the pattern of the surveyed tracks. For many purposes however, it is more convenient when the data is gridded. This also holds for the surveyed depths in this case, because to perform the analysis of the time series, the depth value of the same point has to be known in several epochs. Therefore the data has to be interpolated. Several interpolation methods are available. The most simple way is nearest neighbour interpolation, where the unknown point simply gets the same value as the nearest data point. A some more advanced method is inverse distance weighting, where the value of the unknown point is determined by using the values of the surrounding points, with a certain weight. The method used by the Hydrographic Service is called Ordinary Kriging and is based on this latter principle, where the weights of the data points are obtained from a covariance function. This method is described in Appendix A, here only some results of

this procedure will be shown. The main advantage of this method, above other interpolation methods, is the fact that it is based on statistics and therefore variances of the interpolated data are calculated as well.

In Figure 3.4 the results of the Kriging process are shown. The interpolated depths together with the standard deviations of the interpolated points are given. The $x$- and $y$-coordinates are given in the UTM-31 ETRS89 reference frame. The depths are given w.r.t. the Mean Low Lower Water Spring (MLLWS), which is defined as the average of the lower low water heights over a certain period [De Jong et al., 2002]. The standard deviations consist of two aspects, the measurement accuracy and the interpolation accuracy. The accuracy is getting worse on the edges of the grid. This is because at the edges less data points are available to contribute to the interpolation, or they are distributed poorly around the grid points. But besides these edge points, the standard deviation of most of the interpolated points does not exceed 0.5 m. This is a lot more than the standard deviation of a measured point, compare with Equation 3.2. It is however still below the S44 standard, compare with Equation 3.4.

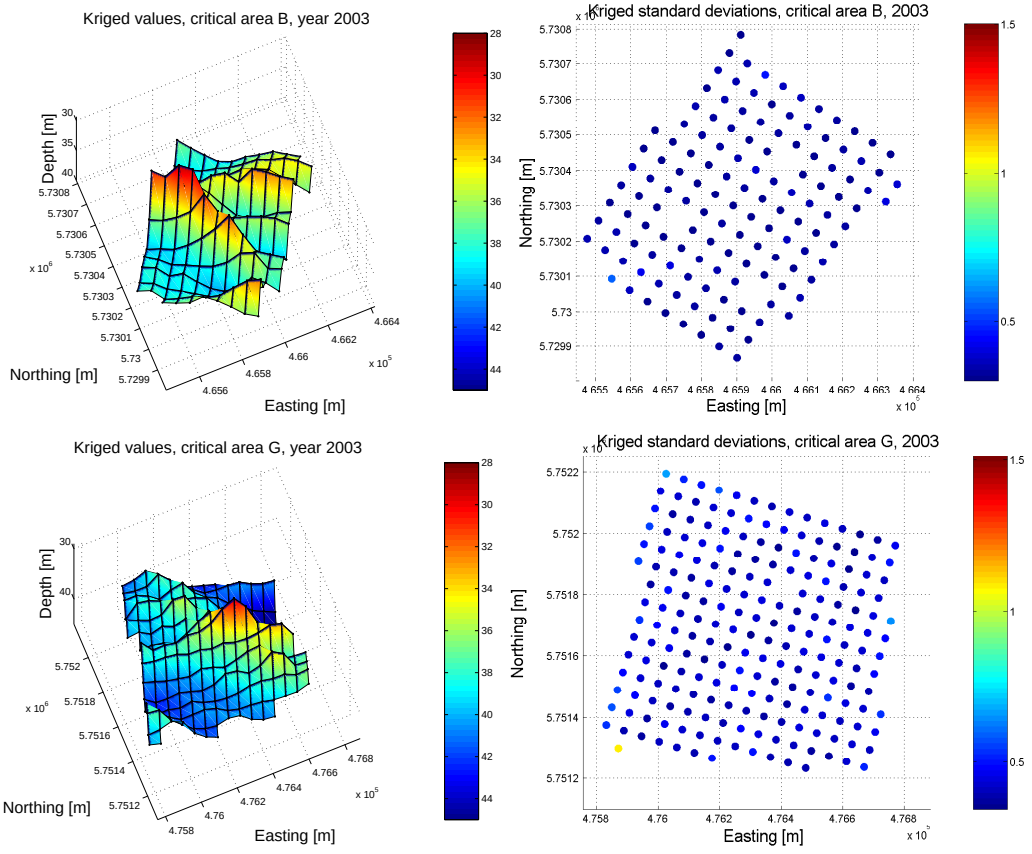



**Figure 3.4:** The results of Kriging of area B (top) and area G (bottom). On the left side the interpolated depths, on the right side the standard deviations of the interpolated points, in m. The $x$- and $y$-coordinates are given in UTM-31 ETRS89, the depths are given w.r.t. the Mean Low Lower Water Spring (MLLWS). The depth axis is exaggerated a hundred times, compared to the coordinate axes.

## 3.2 Deformation analysis

When the data are available on a regular grid and the variances and covariance function are known, the actual analysis of the time series can be performed. The question is whether a conclusion about the dynamics of the sea floor can be drawn from the available data points. It is investigated whether the differences between depth values of various epochs are due to a real change of the sea floor and not a result of measurement inaccuracies. This can be tested by applying the theory of deformation analysis, [Verhoef, 1997], which is a combination of adjustment and testing theory. First, a brief overview of these theories will be given in Paragraph 3.2.1, based on [Teunissen, 2000a] and Paragraph 3.2.2, based on [Teunissen, 2000b]. Hereafter, in Paragraph 3.2.3, a mathematical model of the sea floor is formulated. This is subsequently applied to detect dynamics (Paragraph 3.2.4) and make a prediction for the future (Paragraph 3.2.5).

### 3.2.1 Adjustment theory

Due to various reasons, measurements always contain a level of uncertainty. In case of depth measurements this is for example due to:

- Circumstances of the measurement are not always controllable and descriptive, for example the heave, roll and pitch parameters.
- A mathematical model is adapted to calculate the actual depth for the measurement, but some differences between the model and the real world remain. Consider for example the model for the tidal correction or the sound speed.
- The measuring is always bound to the accuracy of the instruments. This may be improved by using better instruments, but the measurement noise will never disappear completely.
- Some measurements may be wrong due to systematic or gross errors, consider e.g. reflections on fish. For the data used in this project, this does not play a role, because the gross errors have already been removed.

To deal with this, the measurements are supposed to be stochastic instead of deterministic. The unknown parameters are then estimated from the available data using adjustment theory. Usually, the mathematical model expresses a linear relationship between the unknown parameters $x$ and the observations $\underline{y}$ with help of a model matrix $A$ [Teunissen, 2000a]:

$$\underline{y} = Ax + \underline{e}. \tag{3.8}$$

The additional term $\underline{e}$ expresses the stochastic influence in the measurements. This relationship can also be expressed in terms of mathematical expectation and dispersion as:

$$E\{\underline{y}\} = Ax; \qquad D\{\underline{y}\} = Q_y, \tag{3.9}$$

since the expectation of the measurement errors $\underline{e}$ equals zero. The operator $E\{\}$ denotes the mathematical expectation and $D\{\}$ the variance-covariance matrix ($D$ denotes dispersion). Equation 3.9 is called the *model with observation equations* or *A-model*. It can be proved that the unknown parameters $x$ can be estimated as a Best Linear Unbiased Estimation (BLUE) with a least-squares method as, see [Teunissen, 2000a]:

$$\underline{\hat{x}} = (A^* Q_y^{-1} A)^{-1} A^* Q_y^{-1} \underline{y}. \tag{3.10}$$

Here, and in the following, $^*$ means the transpose. Applying the propagation laws of covariances gives:

$$Q_{\hat{x}} = (A^* Q_y^{-1} A)^{-1}. \tag{3.11}$$

With the estimate for the unknown, an adjusted value for the observation can be calculated as:

$$\underline{\hat{y}} = A\underline{\hat{x}}, \tag{3.12}$$

with variance

$$Q_{\hat{y}} = AQ_{\hat{x}}A^*. \tag{3.13}$$

Finally, the difference between the observations and their estimated values is:

$$\underline{\hat{e}} = \underline{y} - \underline{\hat{y}}, \tag{3.14}$$

with variance

$$Q_{\hat{e}} = Q_y - Q_{\hat{y}}. \tag{3.15}$$

As can be seen from Equation 3.9, the A-model presumes a linear relationship between the observations $\underline{y}$ and the unknowns $x$. Of course, this is not always the case, but this can be solved by linearising the observation equations with a Taylor expansion [Teunissen, 2000a]:

$$E\{\underline{y}\} = A(x^0) + \partial_x A(x^0)\Delta x; \qquad D\{\underline{y}\} = Q_y, \tag{3.16}$$

where $x^0$ is an approximate value for $x$. By denoting $\underline{\Delta y} = \underline{y} - A(x^0)$ the linearised model of observation equations becomes:

$$E\{\underline{\Delta y}\} = \partial_x A(x^0)\Delta x; \qquad D\{\underline{\Delta y}\} = Q_y. \tag{3.17}$$

By replacing $A$ with $\partial_x A(x^0)$ in the Equations 3.10 until 3.15, the estimate $\underline{\Delta\hat{x}} = \underline{\hat{x}} - x^0$ can be calculated. The value $\underline{\hat{x}}$ then replaces the initial value $x^0$ and the procedure can start again with the improved approximate value. In this way the nonlinear model can be solved iteratively until the difference between the old and new approximate value becomes negligible.

### 3.2.2 Testing theory

Now that a mathematical model has been adapted to adjust the measurements and estimate the unknowns, the question is whether this model is a valid representation of the reality. Apart from the fact that a model is always a simplification of the reality, some other factors can cause a discrepancy between the model and the reality:

- Wrong assumptions about the functional relations,
- Single errors in measurements (blunders),
- Systematic errors in the measurements, e.g. due to instrumental errors,
- Wrong stochastic description.

The aim of testing theory is checking whether the model is valid, and if not: what causes the model to be invalid. Therefore two hypotheses are tested against each other, [Teunissen, 2000b]:

$$\begin{array}{lll} H_0: & E\{\underline{y}\} = Ax & \text{The model is correct,} \\ H_a: & E\{\underline{y}\} = Ax + C\nabla & \text{The model needs a certain extension } \nabla. \end{array} \tag{3.18}$$

Here $C$ is a the design matrix of the model extensions. These hypotheses are called *null hypothesis* and *alternative hypothesis* respectively.

The test quantity for this model is:

$$\underline{T}_q = \underline{\hat{e}}^* Q_y^{-1} C (C^* Q_y^{-1} Q_{\hat{e}} Q_y^{-1} C)^{-1} C^* Q_y^{-1} \underline{\hat{e}}. \tag{3.19}$$

Under $H_0$, this test quantity has a central chi-squared distribution $\chi^2(q, 0)$ and a non-central $\chi^2(q, \lambda_0)$ under $H_a$, see Figure 3.5. Here, $q$ is the number of degrees of freedom, which is the number of model extension parameters. $\lambda_0$ is the non-centrality parameter which indicates how much the alternative hypothesis deviates from the null hypothesis. The test now reads:

$$\text{reject } H_0 \text{ if } \underline{T}_q > k_\alpha, \tag{3.20}$$

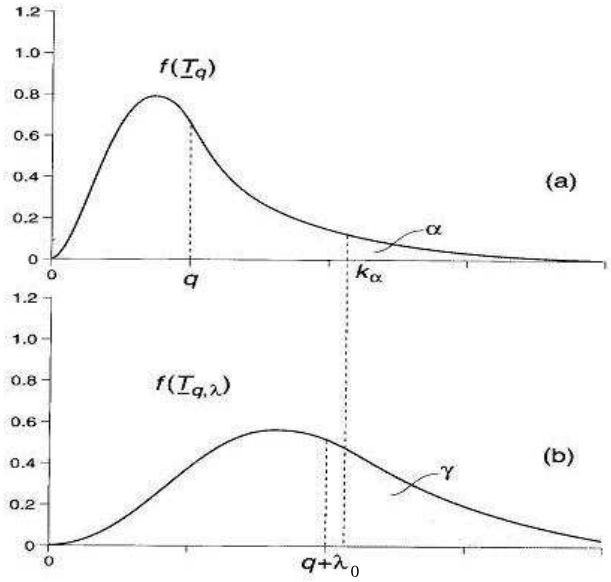



**Figure 3.5:** The expected value for the test quantity $\underline{T}_q$, under the $\chi^2$ distribution . Case (a): $H_0$ is true, case (b) $H_a$ is true. $k_\alpha$ is the critical value, $\alpha$ the level of significance and $\gamma$ the power of the test. Also, the parameters $q$ and $\lambda_0$ are illustrated. Source: [Polman and Salzmann, 1996].

where $k_\alpha$ is the critical value. The critical value can be derived from $q$ and $\alpha$, where $\alpha$ is the level of significance. This is the *type I error*, the chance that $H_0$ is rejected, while $H_0$ is true, so in fact it means a false alarm. Another error is the *type II error*, $\beta$, which is the chance that $H_0$ is accepted where $H_a$ is true, so this is a missed alarm. Usually the power of the test $\gamma$ is used, defined as $\gamma = 1 - \beta$. An overview of these errors is given in Table 3.2.

Often, only one-dimensional alternative hypotheses are used ($q = 1$). In that case the matrix $C$ takes the form of a vector $c$. The square root of the test statistic $\underline{T}_q$ is used, which is called $\underline{w}$:

$$\underline{w}_i = \frac{c_i^* Q_y^{-1} \underline{\hat{e}}}{\sqrt{c_i^* Q_y^{-1} Q_{\hat{e}} Q_y^{-1} c_i}}, \tag{3.21}$$

where $c_i$ is the $c$-vector corresponding with observation $\underline{y}_i$. Hence, the $\underline{w}$-test tests for an error in the $i$-th observation. This test has a standard normal distribution. The $\underline{w}$-tests are used in a data snooping procedure: the test statistic is calculated for every observation. The largest $\underline{w}_i$ then points at the most probable error. This observation can be removed or adjusted and the $\underline{w}$-test statistics are calculated again, until the null hypothesis is accepted.

Another important parameter is the *Minimal Detectable Bias*, which denotes the smallest error that can be discovered with the test:

$$|\nabla| = \sqrt{\frac{\lambda_0}{c_i^* Q_y^{-1} Q_{\hat{e}} Q_y^{-1} c_i}}, \tag{3.22}$$

| | $H_0$ **accepted** | $H_a$ **accepted** |
|---|---|---|
| $H_0$ **true** | correct, chance $1 - \alpha$ | type I error, chance $\alpha$ |
| $H_a$ **true** | type II error, chance $\beta$ | correct, chance $\gamma$ |

**Table 3.2:** Type I and type II errors.

so an error smaller than $|\nabla|$ can be discovered with less than $\gamma$ % probability. The minimal detectable bias depends on $\lambda_0$: a high value for $\lambda_0$ results in a high value for $\nabla$, so the model is then quite tolerant in accepting $H_0$.

### 3.2.3 Mathematical model of the sea floor

In Paragraph 3.2.1 the mathematical model with observation equations (A-model) is derived, see Equation 3.9. In this section, based on [Dorst, 2003], a mathematical model of the sea floor will be derived. Dependent on the type of sea floor, several models with varying complexity can be set up.

#### The sea floor as a flat plane

In this case the expected depth value in a point $p$ is simply equal to the mean depth over the whole area:

$$E\{\underline{d}_p\} = \overline{d}, \tag{3.23}$$

where $\overline{d}$ is the mean depth. Compared to the A-model (Equation 3.9), the observations $\underline{y}$ are expressed here as $\underline{d}_p$ and the unknowns $x$ are expressed as $\overline{d}$.

#### The sea floor as a sloping plane

In case of a sloping plane, the depth in a point $p$ is not only equal to the reference depth $\overline{d}$, but also contains a term for the slope, based on the location of $p$:

$$E\{\underline{d}_p\} = \begin{bmatrix} 1 & x_p & y_p \end{bmatrix} \begin{bmatrix} \overline{d} \\ \overline{\alpha}_x \\ \overline{\alpha}_y \end{bmatrix}, \tag{3.24}$$

where $x_p$ and $y_p$ are the horizontal coordinates , respective to a reference point and $\overline{\alpha}_x$ and $\overline{\alpha}_y$ the mean slopes in $x$- and $y$-direction. These unknowns are not independent, because the slope directions and the reference depth are both dependent on the coordinate system.

#### The sea floor as a sloping plane with sand waves

In case of sand wave presence, these sand waves have to modelled as well. In general, a one-dimensional wave can be described in terms of amplitude $A$ and initial phase $\phi$ as:

$$A\cos(\frac{2\pi x_p}{\lambda} + \phi). \tag{3.25}$$

This is however not a linear model. When substituting $u = A\cos(\phi)$ and $v = A\sin(\phi)$, the amplitude and phase can be calculated from:

$$A\cos(\frac{2\pi x_p}{\lambda} + \phi) = \cos(\frac{2\pi x_p}{\lambda}) \cdot u - \sin(\frac{2\pi x_p}{\lambda}) \cdot v, \tag{3.26}$$

with $\lambda$ the wavelength of the sand wave. Hence, the sea floor model can then be extended as:

$$E\{\underline{d}_p\} = \begin{bmatrix} 1 & x_p & y_p & \cos(\frac{2\pi x_p}{\lambda}) & -\sin(\frac{2\pi x_p}{\lambda}) \end{bmatrix} \begin{bmatrix} \overline{d} \\ \overline{\alpha}_x \\ \overline{\alpha}_y \\ u \\ v \end{bmatrix}. \tag{3.27}$$

This equation only holds when the wavelength is known. The wavelength can be derived from the covariance function, see Paragraph 3.1.2. Otherwise it has to be estimated as well, but then this model is not linear anymore, so it has to be solved iteratively. Note that $A$ and $\phi$ cannot be estimated independently because of the parametrisation by $u$ and $v$.

Summarising, the model of the sea floor can be expressed as:

$$E\{\underline{d}_p\} = \overline{d} + x_p \cdot \overline{\alpha}_x + y_p \cdot \overline{\alpha}_y + A\cos(\frac{2\pi x_p}{\lambda} + \phi). \tag{3.28}$$

### 3.2.4 Test for sea floor dynamics

The model derived in Paragraph 3.2.3 assumes the sea floor to be static, no changes in time occur. The aim of the deformation analysis is to test whether this model is valid by comparing it to alternative time-dependent models. The testing theory of Paragraph 3.2.2 is applied with several alternative hypotheses, that are tested simultaneously. Equation 3.28 is extended by time dependent factors. The following alternatives are modelled:

1. Deviation in a single survey
2. General deformation
3. Trends

General deformation means that the differences between the surveys are too large to be explained by noise, but no trend or deviating survey can be found. A deviating survey is a single survey that does not correspond with the others. This can be due to a measurement error or a real sea floor change. The tests can be performed per point or per area. In the following, the various alternative hypotheses will be formulated. In Figure 3.6, some examples of these alternative hypotheses are illustrated.

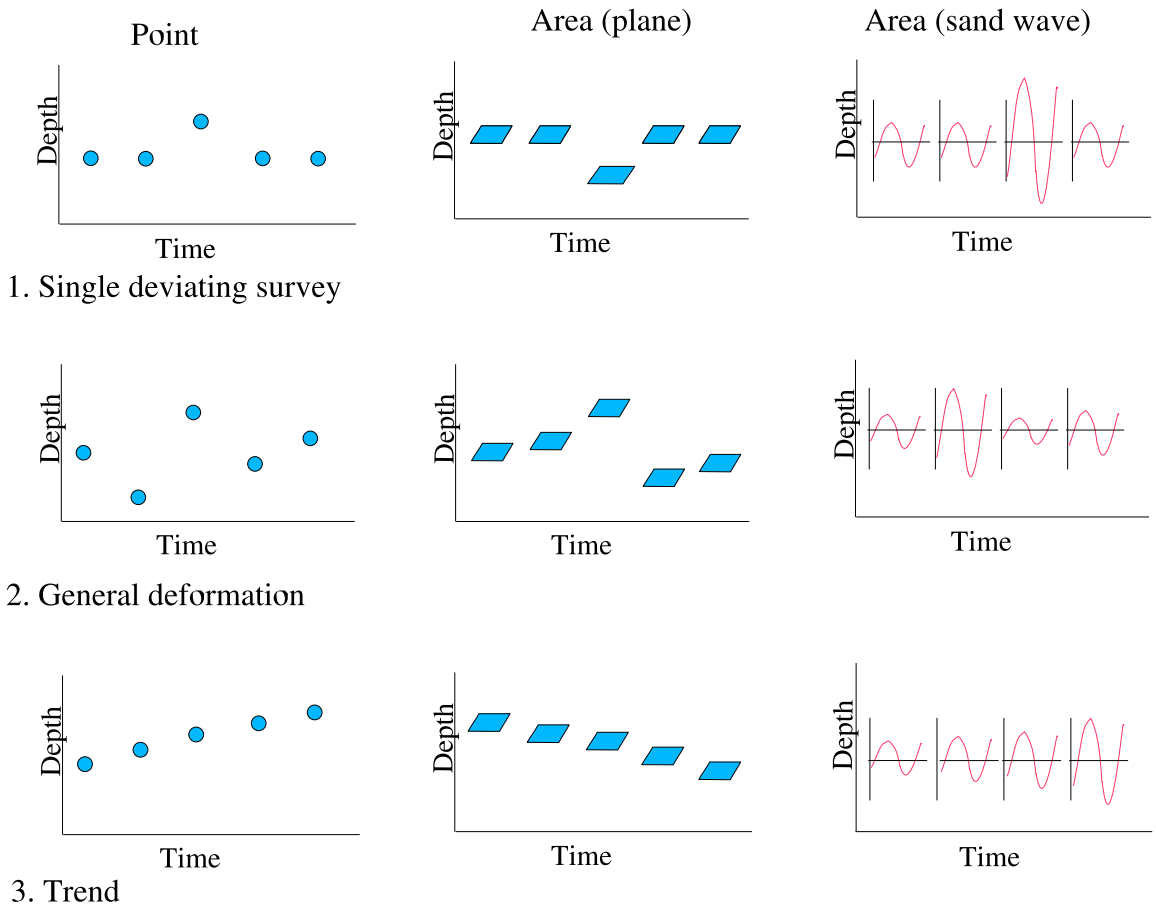



**Figure 3.6:** A graphical representation of some examples of the three different alternative hypotheses. Illustrated are the tests per point, the plane tests per area and the sand wave tests per area. So, for example, the alternative hypothesis for a deviation in survey 3 is illustrated in the upper left corner. Note that for the sand wave tests, only changes in amplitude are depicted, however changes in phase exist as well. The tests per plane all show a change in depth, changes in slope are not illustrated.

**Tests per point**

In this case, the alternative hypotheses will only be tested for the individual points, so the model of Equation 3.23 is the $H_0$-hypothesis. One or several extensions $\nabla$ will possibly be added to the model, as in Equation 3.18.

ad 1. In case of a single deviating survey, the alternative hypothesis is:

$$H_{dev} : E\{\underline{d}_p[k]\} = d_p + i[k] \cdot \Delta d_p, \tag{3.29}$$

where $k = 1 \dots K$ denotes the epoch (survey number) and $i[k]$ points at the deviating survey. It contains of a vector with a one at survey $k$ and zeros elsewhere. $\Delta d_p$ is the deviation in point $p$. This alternative hypothesis is applied for every point.

ad 2. In case of general deformation, the alternative hypothesis reads:

$$H_{gen} : E\{\begin{bmatrix} \underline{d}_p[1] \\ \underline{d}_p[2] \\ \vdots \\ \underline{d}_p[K] \end{bmatrix}\} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} d_p + \begin{bmatrix} 0 & 0 & \dots & 0 \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix} \begin{bmatrix} \Delta d_p[1] \\ \Delta d_p[2] \\ \vdots \\ \Delta d_p[K] \end{bmatrix}, \tag{3.30}$$

where $\Delta d_p[k]$ is the difference between survey $k$ and the reference survey (survey 1). Note the difference between 3.29 and 3.30: in the first case only an error is assumed in a single survey, whereas in the second case for all surveys except the reference survey an error is estimated.

ad 3. The alternative hypothesis for a linear trend is:

$$H_{trend} : E\{\underline{d}_p[k]\} = d_p + (t_k - t_0)\frac{\Delta d}{\Delta t}, \tag{3.31}$$

where $\frac{\Delta d}{\Delta t}$ is the mean change in depth per time unit, say per year, and $(t_k - t_0)$ the time difference between survey $k$ and the reference survey.

**Tests per area**

Instead of testing a single point, the test can also be performed per area, so for all points together instead of for single points. This is however only useful if it is assumed that the dynamics are the same for the complete area. The advantage of a test per area is that also changes in slopes or sand waves can be detected and the number of observations is larger, so a small change is more likely to be detected. For the tests per area the same three types of alternative hypotheses as for the tests per point can be modelled, but they can be subdivided in changes in the plane and changes in the sand wave, see Figure 3.6.

ad 1. The alternative hypothesis for a deviating plane is:

$$H_{dev} : E\{\underline{d}[k]\} = \overline{d} + i[k]\Delta d_k + x \cdot (\alpha_x + i[k]\Delta\alpha_{x,k}) + y \cdot (\alpha_y + i[k]\Delta\alpha_{y,k}), \tag{3.32}$$

$d$, $x$ and $y$ are $P \times 1$ vectors, where $P$ is the number of points. Compared to Equation 3.29, there are two differences: in the first place a change in depth for the entire area ($\Delta d_k$) is calculated instead of a change in depth for a single point ($\Delta d_p$). Second: changes in the slope of the plane, $\Delta_\alpha$ can be detected as well. Note that in this case all surveys at all points $1 \dots P$ together are used as observations to detect the unknown change parameters.

Besides a change in the depth or slope of a plane, a change in the sand wave can be detected as well. In this case an additional term for the sand wave is added to the alternative hypothesis $H_{dev}$:

$$\cos(\frac{2\pi x}{\lambda}) \cdot i[k]\Delta u_k - \sin(\frac{2\pi x}{\lambda}) \cdot i[k]\Delta v_k. \tag{3.33}$$

So the change in sand wave is parameterised by changes in $u$ and $v$, where $u = Acos\phi$ and $v = Asin\phi$, see Section 3.2.3. Only changes in amplitude and phase are considered, the wavelength is assumed to be constant.

ad 2. In case of general deformation, the alternative hypothesis $H_{gen}$ is

$$E\{\begin{bmatrix} \underline{d}[1] \\ \underline{d}[2] \\ \vdots \\ \underline{d}[K] \end{bmatrix}\} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \overline{d} + \begin{bmatrix} 0\,0\,0 & & & 0\,0\,0 & \dots & 0\,0\,0 \\ 0\,0\,0 & i[2]\ xi[2]\ yi[2] & & & & \\ & & & & \ddots & \\ & & & & & i[K]\ xi[K]\ yi[K] \end{bmatrix} \begin{bmatrix} \Delta d[1] \\ \Delta\alpha_x[1] \\ \Delta\alpha_y[1] \\ \Delta d[2] \\ \Delta\alpha_x[2] \\ \Delta\alpha_y[2] \\ \vdots \\ \Delta d[K] \\ \Delta\alpha_x[K] \\ \Delta\alpha_y[K] \end{bmatrix} . \tag{3.34}$$

In case of an dynamic sand wave the alternative hypothesis $H_{gen}$ is extended with:

$$\begin{bmatrix} 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ & & \omega \cdot i[2] & \theta \cdot i[2] & & & \\ & & & & \ddots & & \\ & & & & & \omega \cdot i[K] & \theta \cdot i[K] \end{bmatrix} \begin{bmatrix} \Delta u[1] \\ \Delta v[1] \\ \Delta u[2] \\ \Delta v[2] \\ \vdots \\ \Delta u[K] \\ \Delta v[K] \end{bmatrix}, \tag{3.35}$$

where $\omega = \cos(\frac{2\pi x}{\lambda})$ and $\theta = -\sin(\frac{2\pi x}{\lambda})$. Just like the tests per point, the alternative hypotheses for general deformation are a generalisation of the hypotheses for a deviating survey: all surveys are assumed to be different instead of a single survey. In the models 3.34 and 3.35, survey 1 is taken as reference survey.

ad 3. A trend in the sea floor is modelled as:

$$H_{trend} : E\{\underline{d}[k]\} = \overline{d} + (t_k - t_0)\frac{\Delta d}{\Delta t} + x \cdot \left(\overline{\alpha}_x + (t_k - t_0)\frac{\Delta \alpha_x}{\Delta t}\right) + y \cdot \left(\overline{\alpha}_y + (t_k - t_0)\frac{\Delta \alpha_y}{\Delta t}\right). \tag{3.36}$$

This can be seen as an extension of Equation 3.31 with two factors for the change in slope of the plane. $(t_k - t_0)$ is the time difference between the reference survey $t_0$ and the survey $k$, just like in Equation 3.31.

In case of a sand wave $H_{trend}$ is extended with:

$$\cos(\frac{2\pi x}{\lambda}) \cdot (t_k - t_0) \cdot \frac{\Delta u}{\Delta t} - \sin(\frac{2\pi x}{\lambda}) \cdot (t_k - t_0) \cdot \frac{\Delta v}{\Delta t}. \tag{3.37}$$

The trend in a sand wave can consist of a trend in amplitude (a growing sand wave), a trend in phase (a migrating sand wave), or a combination of both. The changes in amplitude and phase are parametrised by $\Delta u$ and $\Delta v$ and are therefore inseparable.

Now that the alternative hypotheses are formulated, they can be tested against the null hypothesis by applying the hypothesis snooping procedure, see Figure 3.7. Before applying this, it has to be defined whether sand waves are present, otherwise no test on sand waves is performed. For every survey the according test quantity $\underline{T}_q$ is calculated and divided by the critical value $k_a$. When this result is larger than 1, i.e. the test quantity is larger than the critical value, the null hypothesis is rejected. The most likely alternative hypothesis, which is the one with the largest test quotient, is accepted. This becomes the new null hypothesis and the test quantities are calculated again, until all remaining alternative hypotheses are rejected.

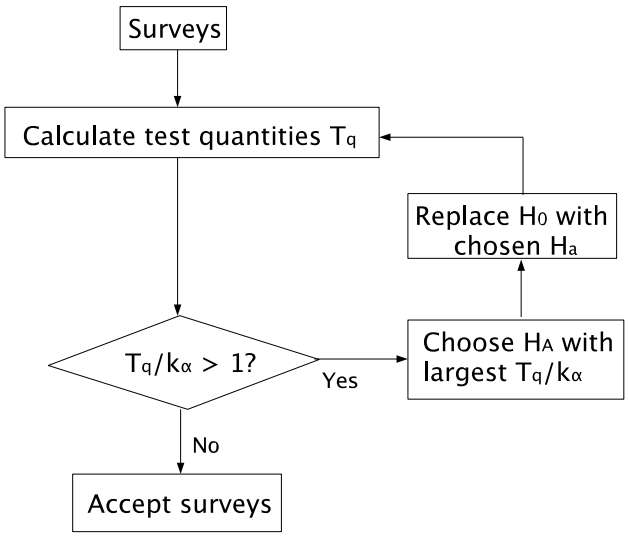



**Figure 3.7:** Overview of the hypothesis snooping procedure.

### 3.2.5 Prediction of the points on the sea floor

When the hypothesis snooping procedure is finished, the extended mathematical model can be used to make a prediction of the behaviour of the sea floor in the future. The aim of the whole deformation analysis procedure was to be able to make a more efficient planning of the surveys and a prediction might help to decide when to survey again. For the prediction only the results of the point test are used. The most simple case occurs when the sea floor is accepted as being static. In this case it does not matter at what time the depth is measured, the predicted depth value is simply the same as the adjusted depth value from the previous epochs. Therefore, the mathematical model reads:

$$E\{\underline{\hat{d}}_{K|K}\} = I \cdot d_{pred}, \tag{3.38}$$

with $I$ the $P \times P$ identity matrix, where $P$ denotes the number of points. $\underline{\hat{d}}_{K|K}$ means the adjusted depth, base on surveys $1 \dots K$, with survey $K$ the last survey. The predicted depth is time independent, it does not make a difference whether the prediction is one year ahead or e.g. 10 years ahead. The model tells that the old situation is the same as the new situation. Here, the vector of estimated depth values serves as the observation vector in the model of observation equations of Equation 3.9. Therefore the predicted values can be calculated as in Equation 3.10 and the corresponding variances as in 3.11. In this case, with an identity matrix as design matrix this of course reduces to:

$$\underline{\hat{d}}_{pred} = \underline{\hat{d}}_{K|K} \quad ; \quad Q_{\hat{d}_{pred}} = Q_{\hat{d}_{K|K}}. \tag{3.39}$$

In case of a deviating survey the predicted values and variances are calculated as:

$$\begin{aligned} \underline{\hat{d}}_{pred} &= p_{outlier} \cdot \underline{\hat{d}}_{outlier} + (1 - p_{outlier}) \cdot \underline{\hat{d}}_{normal} \\ Q_{\hat{d}_{pred}} &= p^2_{outlier} \cdot Q_{\hat{d}_{outlier}} + (1 - p_{outlier})^2 \cdot Q_{\hat{d}_{normal}}, \end{aligned} \tag{3.40}$$

where $p_{outlier}$ is probability of an outlying measurement, so the number of points containing an outlier divided by the total number of points. In case more than one outlying survey is present, the predicted depth is a summation of Equation 3.40 for all outlying surveys.

When a point on the sea floor is generally deformed, the predicted depth is the same as the weighted average of the depth measurements. Therefore this is the same model as for a static sea floor, see Equation 3.38. However, in case of general deformation, only the variance of the predicted estimator is not enough to describe the variance of the prediction, the variability of the sea floor also has to be taken into account as well. Because this is not known beforehand, the variance is calculated afterwards from the data:

$$\hat{Q}_{\hat{x}_{pred}} = \frac{1}{K}\sum(\hat{\underline{d}}_k - \hat{\underline{d}}_{pred}), \tag{3.41}$$

with $K$ the number of surveys. This term defines the standard deviation of the sea floor model parameters, which is a combination of the accuracy of the measurements, the accuracy of the interpolation and the estimated variability of the sea floor.

A fourth possibility is a trend in the sea floor. In this case, a time dependent parameter $\frac{\Delta\hat{d}}{\Delta t}$ has to be included. The model becomes:

$$E\{\begin{bmatrix} \hat{\underline{d}}_{t_0} \\ \frac{\Delta\hat{d}}{\Delta t} \end{bmatrix}\} = \begin{bmatrix} I & \Delta t \cdot I \\ 0 & I \end{bmatrix} \begin{bmatrix} d_{pred} \\ \frac{\Delta d}{\Delta t} \end{bmatrix}. \tag{3.42}$$

Applying the formulae of the A-model yields expressions for the predicted values and variances as:

$$\begin{aligned} \hat{\underline{d}}_{pred} &= \hat{\underline{d}}_{t_0} + \frac{\Delta\hat{d}}{\Delta t} \cdot \Delta t \\ Q_{\hat{d}_{pred}} &= Q_{\hat{d}_{t0}} + \Delta t^2 \cdot Q_{\frac{\Delta d}{\Delta t}}. \end{aligned} \tag{3.43}$$

## 3.3 Deformation analysis applied to a data set

In Section 3.2 the theory of the deformation analysis of the Hydrographic Service has been derived. In this section it is applied to a practical example. Therefore, the two critical areas of Section 3.1 are used again. The analysis can be split up in two parts, the point analysis and the area analysis, which will be described below. In both cases, the analysis is based on 6 epochs: for area B data is available from the years 1991, 1995, 2000, 2001, 2002 and 2003; for are G data is available from 1991, 1995, 1998, 2001, 2002 and 2003.

### 3.3.1 Point stability of the critical areas

This test yields an analysis for the stability of all individual points. According to the hypothesis snooping procedure, see Figure 3.7, this is done by comparing a test quantity with a certain critical value. These can be calculated from the $\chi^2_{q,\lambda}$ distribution. Two parameters are important here, $\alpha$ and $\gamma$. $\alpha$ is the level of significance or the type I error (false alarm), see Paragraph 3.2.2. Ideally this should be as small as possible, but on the other hand, the type II error $(1-\gamma)$ should be minimised too. Moreover, minimisation of the type II error (missed alarm) is in fact even more important for this application. First consider the type I error: applied to sea floor dynamics this means that e.g. a trend is detected, while in fact the sea floor is static. Because a trend is detected, the survey frequency should be increased. However because the sea floor is static, this is not necessary, so the surveys become less efficient. But now consider a type II error: this means that the sea floor is assumed to be static, while in fact e.g. a trend is present. Because no dynamic behaviour is detected, the survey frequency might be decreased, resulting in a situation that a significant rise of the sea floor is not detected in time. Therefore, a large type II error could lead to dangerous situations.

Usually, first a value for $\gamma$ is chosen and then an optimal value for $\alpha$. The effects of varying those parameters are illustrated in [Dorst, 2003]. Here, also a suggestion for the levels of significance for the

| | |
|---|---|
| Test on a deviating sand wave in one survey | 1 % |
| Test on a deviating plane in one survey | 1 % |
| Test on a trend in a sand wave | 10 % |
| Test on a trend in the plane | 10 % |
| Test on general deformation in the sand wave | 5 % |
| Test on general deformation of the plane | 5 % |

**Table 3.3:** Levels of significance for the different types of dynamics.

| | **Area B** | **Area G** |
|---|---|---|
| Point stability analysis: | | |
| Static points | 83.1 % | 80.9 % |
| General dynamic points | 2.9 % | 1.9 % |
| Downwards moving points | 5.9 % | 10 % |
| Upwards moving points | 0.7 % | 3.8 % |
| Outlying values | 7.3 % | 3.4 % |

**Table 3.4:** The results of the point stability analyses for the areas B and G. The percentages of static points, dynamic points etc. are shown.

different tests in this program is done, based on experience. These can be found in Table 3.3. So the test on trends has the highest probability: the test will be more likely to assume a trend than e.g. a single deviating survey. That is because it has more severe consequences to miss a trend than to miss e.g. an outlier. In this case, a $\gamma$ of 80 % is chosen, see [Dorst, 2004a]. With Equation 3.22 the minimal detectable biases can be calculated. These are illustrated in Figure 3.8, for different point standard deviations. It can be seen that much smaller changes can be detected with the tests per area than with the tests per point.

The results of the tests per point for the critical areas B and G are given in Figure 3.9 and Table 3.4. Four different types of points are discerned: static points, general dynamic points, upwards moving points and downwards moving points. Furthermore outlying values are detected. It can be seen that most points are static, although there are some points that show a upward or downward movement, especially near the sand wave crests.

### 3.3.2 Area stability of the critical areas

For an analysis of a single point only a few observations are available and therefore the redundancy is not very high. This means that sometimes there may not be enough observations to determine dynamic behaviour. An analysis for the entire area is therefore more suitable, because many points contribute to the solution and the redundancy is high. The disadvantage of analysing the entire area is of course that it is not possible to determine local changes. For the area stability analysis the same levels of significance are used as for the point analysis, see Table 3.3. The adjusted values for area B are given in Figure 3.10. It shows an upwards trend in the depth with 4 cm per year and a change in the slopes in both directions. Furthermore two of the six surveys are classified as outlier. Especially survey 1 shows a large deviation. This however also indicates that the estimated trend is based on only a few surveys, and could easily be different in reality. A trend through surveys 1, 4 and 6 is for example also possible. This also depends on the grid distance, when another grid distance is chosen, the mean depth may be estimated slightly different. This could result in different dynamics being found. The standard deviations are indicated as well. It can be seen that for surveys at the beginning or end of the time series the estimation accuracy of dynamics is getting worse. The same holds for outlying surveys, their accuracy is also worse than those of normal surveys. Compare this for example with the results of area G, Figure 3.11. The estimated values for the depths there show almost the same pattern, but here a downward trend is estimated.

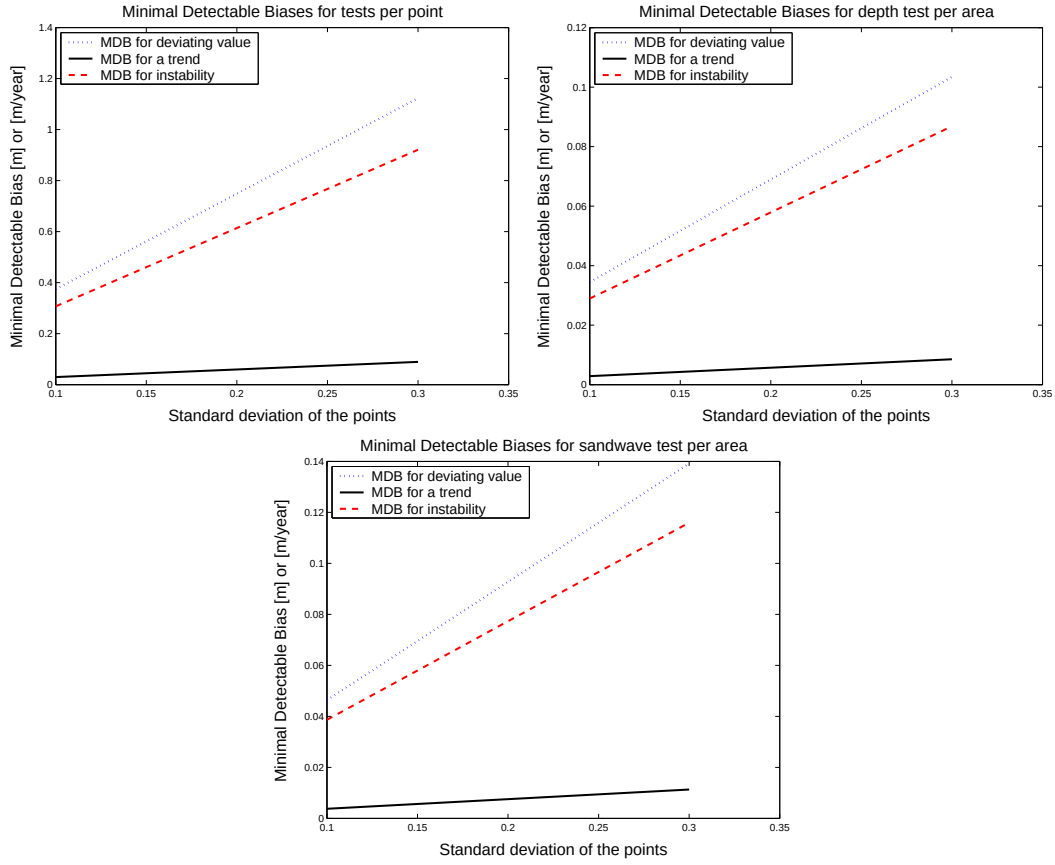



**Figure 3.8:** Minimal Detectable Biases for the tests per point, the test per area for depth and the test per area for sand waves. The MDB's are dependent on the standard deviations of the data points. These MDB's are valid for the critical areas B and G, with 6 surveys in 13 years.

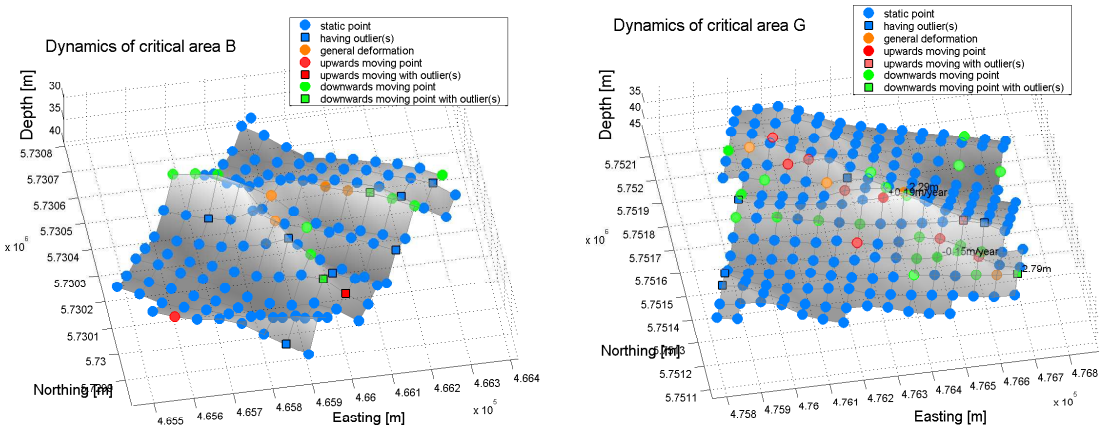



**Figure 3.9:** Graphical representation of the point analysis of the critical areas B (left) and G (right). It is illustrated whether the points are static, dynamic, downward moving or upward moving. Having outliers means that an outlying value is found for that point in one or more epochs.

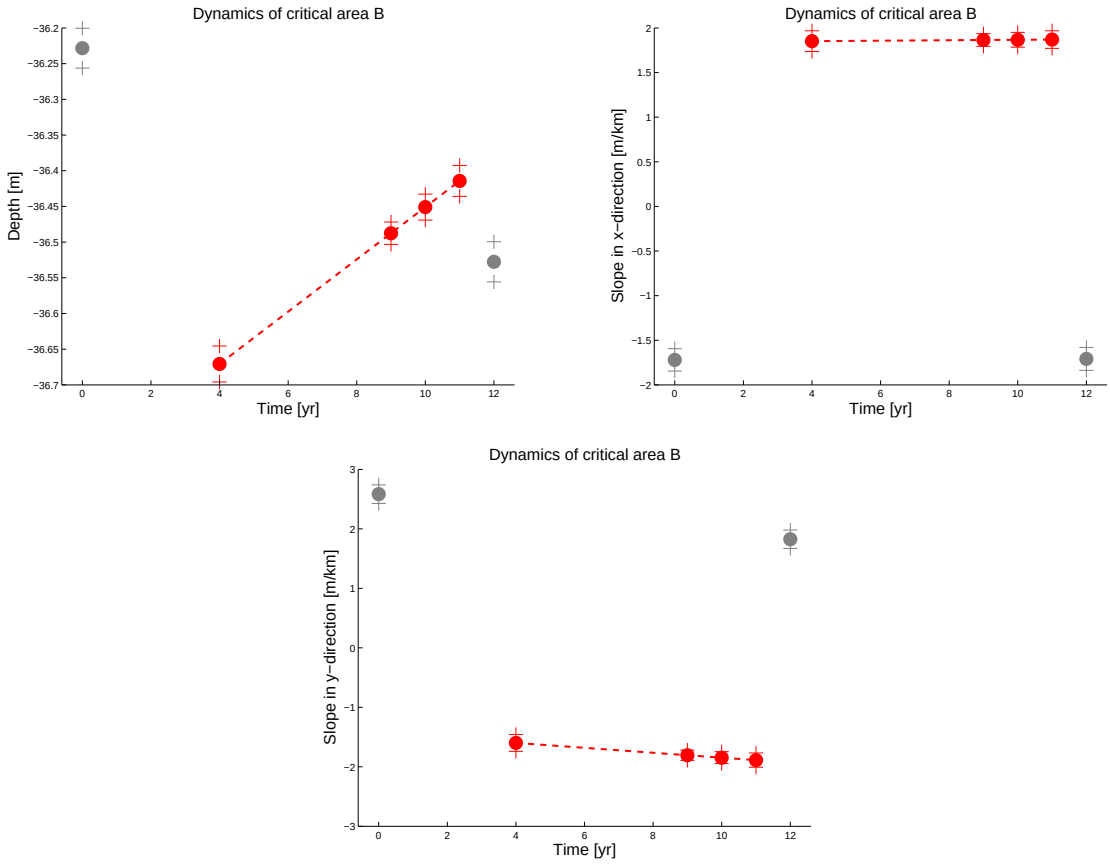



**Figure 3.10:** Area stability of critical area B, the adjusted depths and slopes. The fact that most points are coloured red means that a upward trend is present, analogous to the colours in Figure 3.9, the grey points are outlying values. The "+" marks above and below the values indicate the standard deviations of the estimated sea floor parameters. It can be seen that the outlying values show a much larger standard deviation than the other points.

The stability of area G is given in Figure 3.11. For this area the amplitude and crest distance of the sand wave are analysed as well. The crest distance ($\psi$) is defined as in terms of the phase of the sand wave:

$$\psi = \left(1 - \frac{\phi}{2\pi}\right)\lambda, \tag{3.44}$$

so this defines the distance from the sand wave crest to a certain reference. This can be chosen arbitrarily, because only relative dynamics are considered. It is shown that the sand wave is migrating and there is a very small decrease in amplitude, but this is almost negligible. Besides this, the average depth is increasing as well. In Table 3.5 an overview of the results of the point and area stability analyses of both critical areas are given.

### 3.3.3 Predicted depths for the critical areas

Finally, the results of the stability analysis are used to predict the state of the sea floor a few years ahead. The results for the prediction for 2008 (5 years ahead) are given in Figure 3.12, together with the predicted standard deviations. The prediction is based on the analysis per point, so only the points that show dynamic behaviour have changed w.r.t. the last survey, compare with Figure 3.9. It can also be seen that the standard deviation of these points is worse than the static points. It must be remarked that

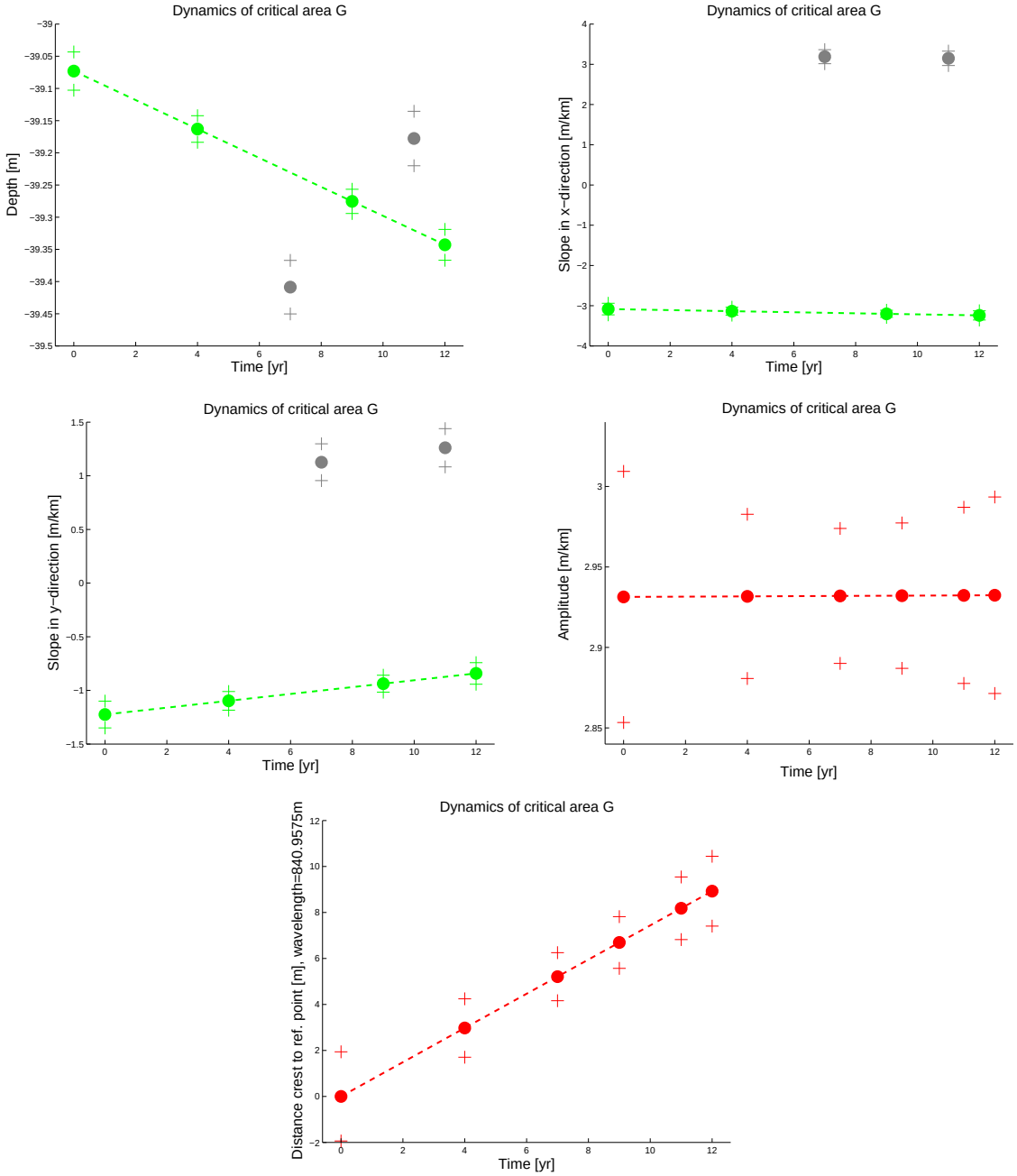



**Figure 3.11:** Area stability of critical area G, the adjusted depths, slopes and sand wave parameters. The green points indicate that a downwards trend is present, the grey points are outlying values. The "+" signs indicate the standard deviations.

| Area B | | Area G | |
|---|---|---|---|
| Dev. plane in survey 1 | -0.59 ± 0.05 m | Trend in depth (deepening) | +0.02 ± 0.003 m/yr |
| Trend in plane (shoaling) | -0.04 ± 0.005 m/yr | Dev. plane in survey 3 | +0.18 ± 0.04 m |
| Dev. plane in survey 6 | +0.15 ± 0.04 m | Migrating sand wave | +0.74 ± 0.26 m/yr |
| | | Dev. plane in survey 5 | -0.12 ± 0.04 m |

**Table 3.5:** The accepted alternative hypotheses for the areas B and G, with their magnitudes and standard deviations.

this prediction is only on basis of the detected dynamics, no physical theories are applied. The dynamics are detected on only a few epochs, so it is not certain that these movement parameters describe the total behaviour. Furthermore, it assumes that this movement remains the same between the last survey and the prediction time, which is of course not the case in reality.

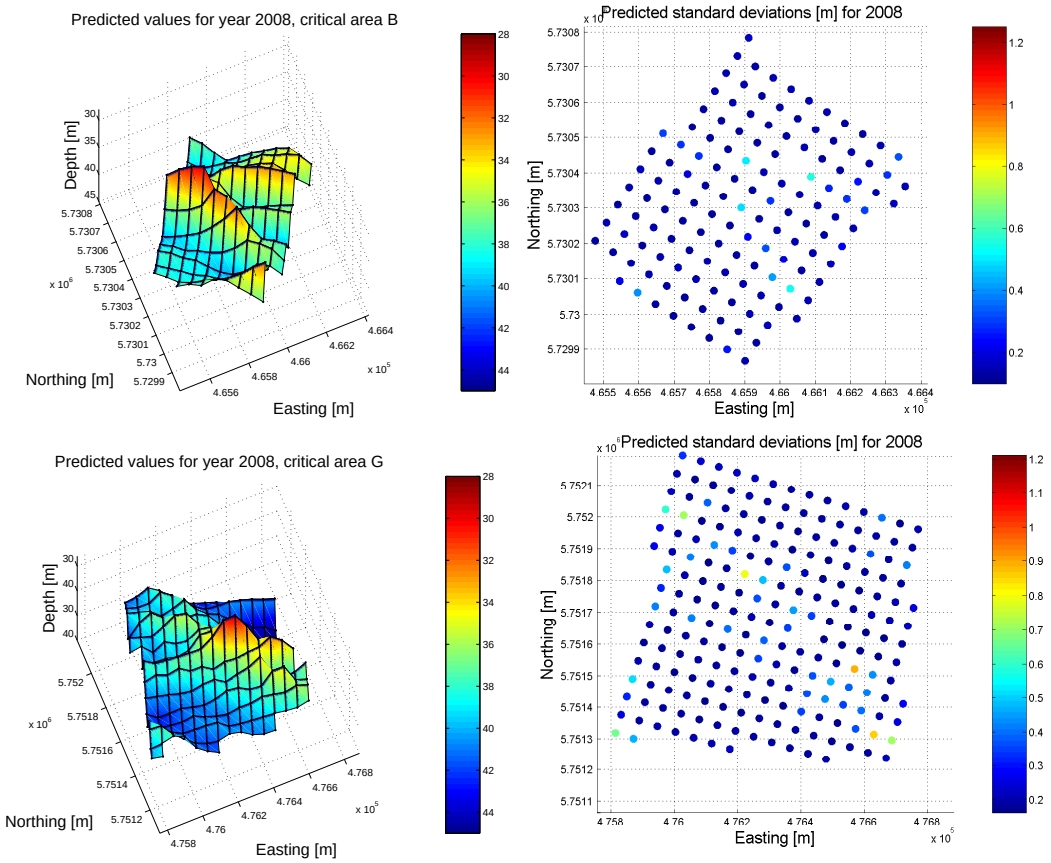



**Figure 3.12:** Predicted depths and standard deviations of the critical areas B and G for the year 2008, based on surveys until 2003.

## 3.4 Overview and discussion

The method of the Hydrographic Service consists of two phases, statistics and deformation analysis. The first phase is necessary, because the same data points have to be available in every epoch, otherwise a horizontal bias would be introduced. Therefore the data are interpolated to a regular grid. Furthermore, the variances have to be known and in case of sand waves the wavelength and direction of the sand wave have to be known, in order to be able to perform the second part. Analysis of the original data is also impractical because of the large amount of data that has to be processed. Furthermore, the geostatistics phase can be divided in several steps, which will be described below.

- Reading in of the data. The data is given in a Easting-Northing system (UTM-31 ETRS89) and a depth value with respect to the Mean Lower Low Water Spring. The surveys of the Hydrographic Service are usually carried out perpendicular to the sand wave direction. Therefore the tracks do not lie in E-N direction, but in an arbitrary direction. The variances of the measurements are calculated with Equation 3.1.

- Preprocessing. The data are reduced to a representative value per grid cell. The Hydrographic Service are working with the shallowest value of a 5×5 meter grid cell. The practical purpose of this choice is the advantage that one is always on the safe side. However the choice for the shallowest value is mathematically not very practical, because then the assumption that the data are normally distributed does not hold anymore. A systematic error is introduced, however when the standard deviations are of the same magnitude, this systematic error will be the same for all epochs and the depths can still be compared.

- Calculation of the variances of the measurements. This is done by applying Equation 3.1.

- Calculation of the covariance function. From the unequally distributed depth values, deterministic covariance can be derived. Then, an analytical covariance function is fitted, see Paragraph 3.1.2. Two covariance functions are calculated, in different directions. This is because the sea floor is usually not isotropic but there are for example sand waves. To determine the directions in which the covariance functions are calculated, the direction of highest variability is calculated by means of the maximum gradients. In case of presence of sand waves this means that one covariance function is calculated along the crests and one perpendicular to the sand waves. This also has consequences for the analytical function that has to be fitted: in case of a sea floor without sand waves a Gaussian model is chosen, in case of a sea floor with sand waves a hole effect model is chosen in the direction of the highest variability and a Gaussian model is chosen in the other direction.

- Interpolation. With help of the covariance functions, the irregularly spaced data are interpolated to a regular grid with the Kriging method. Together with the Kriged values, also standard deviations are calculated. These are a combination of the measurement accuracy and the interpolation accuracy.

In the deformation analysis, the points for the various epochs are compared to each other and tested for deviations or trends. The two main factors concerning the data are the number of epochs and the quality of the (interpolated) data. The deformation analysis procedure tests whether a difference between two or several epochs is a real deformation or just a consequence of the stochastic nature of the data. When more epochs come available, a trend will therefore become more obvious and easier to detect. When the standard deviations of the data are very low, a difference between two epochs is more likely to be a real deformation than in case of very high standard deviations. Because the data of the Hydrographic Service were acquired with a single beam echo sounder, not many points are available to contribute to the interpolation, which results in a lower quality. This of course has also consequences for the deformation analysis.

The second phase consists of the following steps:

- Model extending for point stability. First, the alternative hypotheses for the individual points are tested, see 3.2.4, following the procedure of Figure 3.7. This results in conclusions whether the

individual points are static, dynamic or having outliers. A disadvantage of this test is that it is only valid for a single point. One would expect that if a point is moving downwards, the neighbouring points would also show some dynamics. This is however not always the case, maybe because not enough epochs are available to detect a trend.

- Model extending for area stability. After the point test, the entire area is tested at once. This makes it more easy to detect e.g. an upwards movement of the area. However, the trend detection is usually based on only a few surveys, slight differences in the mean depth may cause other dynamics being found.

- Prediction of future values. When it is determined whether the sea floor is dynamic or not, this knowledge can be used to predict what the situation will be in a few years. When for example an upwards trend of 3 cm/year is detected, one can predict what the depth will be a few years later. It is however very questionable whether this prediction is right, because of the assumption that the trend will be constant. To be able to make a reliable prediction, one needs to know more about the physical theories. Alternatively, this method could be used, but the precision of the predicted depths should increase with time. Therefore not too much confidence much be laid in this prediction, it can however be useful for planning purposes. Another aspect is the fact that only trends higher than the minimal detectable bias are accepted with a certain probability. If a trend is present, but it does not exceed this limit, it is missed and not taken into account in the prediction. On the other hand when a trend is detected, it is always questionable whether this trend is truly present and what is the size of the trend. Summarising, this analysis very much depends on the choices for the levels of significance and the precision of the measurements.

# Chapter 4

# Method 2: Kalman filtering

In Chapter 3, the Dutch Hydrographic Service' method to detect changes of the sea floor with time series of bathymetric data was described. They are however not the only authority in the Netherlands that is surveying the North Sea. The North Sea Directorate, part of the Ministry of Public Works and Transport also has this responsibility, see Section 2.1.

The North Sea Directorate has three main areas of interest, namely:

- "Living Sea": This comprises the management of the various activities at sea, spatial planning and maintenance of the eco-system.
- Maintenance of the sea channels. The North Sea is rather shallow, so the sea channels to the ports have to be dredged so that the ports remain accessible. Another aspect is the marking of the channels by means of buoys and lighthouses.
- Safety of the shipping. This is connected with the maintenance of the sea channels, but also includes for example salvage of shipwrecks and combat of chemicals and other garbage.

The most important channel in the Dutch part of the North Sea is the Euro Channel to the port of Rotterdam, see Figure 4.1. Besides the actual channel it also consists of some approach areas, where the ships can stay for a short period until they are allowed to enter the port. The Euro Channel and its approaches are maintained at a depth of 23.4 m, such that ships with a draught of up to 22.5 m are able to enter the port of Rotterdam. The other channel in the Dutch part of the North Sea, the IJ-Channel (see Figure 2.1) to the ports of IJmuiden and Amsterdam is maintained at a depth of 19 m for ships with a draught of up to 16.5 m [Directie Noordzee, 2004]. In this project only data from the Euro Channel are used, see Figure 4.2. It is an area of approximately 3,000 × 400 m, somewhere halfway the channel,

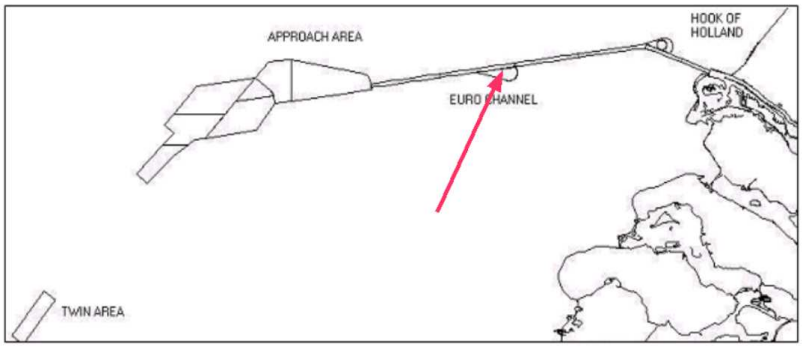



**Figure 4.1:** The Euro channel and its approaches. The red arrow indicates the location of the data set used in this project. Source: [Wüst, 2004]

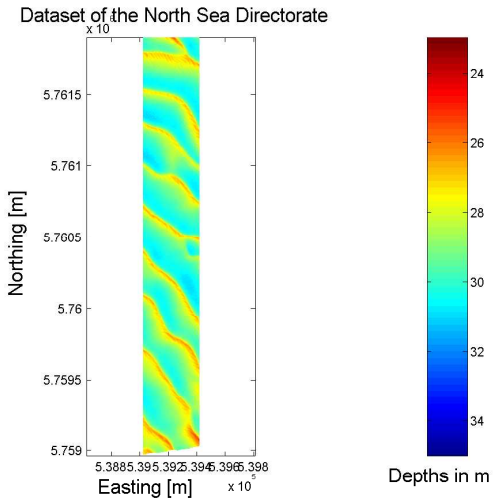



**Figure 4.2:** The data set of the North Sea Directorate used in this project. It is part of the Euro Channel. The coordinates are given in Easting and Northing, w.r.t to the ED-50 reference system. The depths are given w.r.t. the Mean Water Level.

indicated in Figure 4.1 by the red arrow. The need for dredging the channels is obvious if you look at the depths, at some places the sand wave crests are just below or maybe even above the maintained depth of 23.4 m.

In order to gain more insight in the dynamics of the sea floor and to predict when the maintained depth will be overstepped, the North Sea Directorate have developed a trend analysis model. This model will be described in this chapter. First in Section 4.1 a more general introduction on dynamic models will be given, followed by the description of the model by the North Sea Directorate in Section 4.2. After this, a practical example of the analysis of a data set will be given in Section 4.3. This chapter will end with a summary and discussion of this method in Section 4.4.

## 4.1 Dynamic models

In Paragraph 3.2.1 the A-model was derived to estimate unknown parameters from a number of observations, by means of least-squares estimation. This is a time dependent solution, the unknown parameters are calculated afterwards, based on the measurements up to a certain moment. This form of estimation is also called batch processing. This theory will now be extended to a recursive model, where the parameters are time-varying. This allows sequential processing of data from several epochs. A typical example of recursive estimation is navigation, where an estimate of the actual position has to be calculated instantly.

The essence of a recursive model is that it gives a relationship between data from several epochs. When using a recursive model, the parameter values are updated when new measurements come available. This update is only based on the current parameter value and the new measurements, so not all measurements have to be stored.

Three types of recursive estimations can be discriminated [Teunissen, 2001]:

- Filtering: This occurs when the time at which the parameter is estimated coincides with the last measurement time.
- Smoothing: This is the adjustment of the past estimates, based on the new measurement data.
- Prediction: This is calculation of a future parameter value based on the past measurements.

When data from several epochs are available, an A-model for all epochs together can be defined as:

$$E\{\begin{bmatrix} \underline{y}_0 \\ \underline{y}_1 \\ \vdots \\ \underline{y}_K \end{bmatrix}\} = \begin{bmatrix} A_0 & & & \\ & A_1 & & \\ & & \ddots & \\ & & & A_L \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_K \end{bmatrix} \; ; \; D\{\begin{bmatrix} \underline{y}_0 \\ \underline{y}_1 \\ \vdots \\ \underline{y}_K \end{bmatrix}\} = \begin{bmatrix} Q_0 & & & \\ & Q_1 & & \\ & & \ddots & \\ & & & Q_K \end{bmatrix} \tag{4.1}$$

Here $x_0 \dots x_K$ denote the parameters $x$ at the epochs $1 \dots K$. This vector is often called *state vector*. Here all epochs are treated separately so the estimates for the unknowns are calculated every epoch. It is assumed that no correlation between the various epochs exists.

This model makes it possible to estimate $x$ at the moments of the measurements $t_0$ until $t_K$, but it is still not possible to estimate $x$ for all $t$. Therefore it is assumed that there is some functional relation between the parameters $x$ at different epochs:

$$x_t = \Phi_{t,k} x_k, \tag{4.2}$$

where $\Phi_{t,k}$ is the so-called transition matrix. This matrix links the parameters at time $t$ with the parameters at a previous epoch $k$.

Because the inverse of a transition matrix $\Phi_{t,k}$ equals $\Phi_{k,t}$ [Teunissen, 2001], Equation 4.2 can also be written as $x_k = \Phi_{k,t} x_t$. By combining this result with Equation 4.1 it is now possible to estimate the parameter values for an arbitrary moment $t$, based on the observations for every epoch:

$$E\{\begin{bmatrix} \underline{y}_0 \\ \underline{y}_1 \\ \vdots \\ \underline{y}_K \end{bmatrix}\} = \begin{bmatrix} A_0\Phi_{0,t} \\ A_1\Phi_{1,t} \\ \vdots \\ A_K\Phi_{K,t} \end{bmatrix} x_t \; ; \; D\{\begin{bmatrix} \underline{y}_0 \\ \underline{y}_1 \\ \vdots \\ \underline{y}_K \end{bmatrix}\} = \begin{bmatrix} Q_0 & & & \\ & Q_1 & & \\ & & \ddots & \\ & & & Q_K \end{bmatrix}. \tag{4.3}$$

The least squares solution of this model can be calculated using the formulae of Paragraph 3.2.1, and should be denoted as $\underline{\hat{x}}_{t|k}$. Note that when $\Phi_{t,k}$ is the identity matrix, the usual A-model is obtained again. The first index $(t)$ refers to the moment to which the parameter value corresponds and the second index $(k)$ indicates that the estimate is based on the observations $1 \dots k$. So when the estimate at time $k$ is based on observations until $k-1$, the predicted estimator reads:

$$\begin{aligned} \underline{\hat{x}}_{k|k-1} &= \Phi_{k|k-1} \underline{\hat{x}}_{k-1|k-1} \\ Q_{\hat{x}_{k|k-1}} &= \Phi_{k|k-1} Q_{\hat{x}_{k-1|k-1}} \Phi^*_{k|k-1} \end{aligned} \tag{4.4}$$

The equations of 4.4 are often referred to as the *time update*, or prediction. With help of the transition matrix $\Phi_{k|k-1}$ the predicted estimator $\underline{\hat{x}}_{k|k-1}$ can be calculated from the filtered estimator $\underline{\hat{x}}_{k-1|k-1}$. Both estimators are based on the same set of observables, namely until $t = k-1$.

After this, the prediction can be corrected (filtered) when new measurements come available. These relationships are called *measurement update* and read:

$$\begin{aligned} \underline{v}_k &= \underline{y}_k - A_k \underline{\hat{x}}_{k|k-1} \\ Q_{v_k} &= Q_k + A_k Q_{\hat{x}_{k|k-1}} A_k^* \\ K_k &= Q_{\hat{x}_{k|k-1}} A_k^* Q_{v_k}^{-1} \\ \underline{\hat{x}}_{k|k} &= \underline{\hat{x}}_{k|k-1} + K_k \underline{v}_k \\ Q_{\hat{x}_{k|k}} &= [I - K_k A_k] \, Q_{\hat{x}_{k|k-1}}. \end{aligned} \tag{4.5}$$

In these equations $K_k$ is the gain matrix. This gives an indication of the improvement in $\underline{\hat{x}}$ when new measurements come available. By substituting $Q_{v_k}$ the expression for the gain matrix becomes:

$$K_k = Q_{\hat{x}_{k|k-1}} A_k^* (Q_k + A_k Q_{\hat{x}_{k|k-1}} A_k^*)^{-1}. \tag{4.6}$$

From this, it can be seen that the gain matrix depends on:

- $Q_{\hat{x}_{k|k-1}}$: the precision of the previous estimator. When the elements of $Q_{\hat{x}_{k|k-1}}$ are small, the confidence in the old estimator is high and therefore the new estimator should not differ too much from the previous one. This results in a small gain matrix.
- $A_k$: the relation between the model and the added observable.
- $Q_k$: the precision of the added observable. When the elements of $Q_k$ are small, the confidence of the added observable is high and therefore the old estimator should gain a lot from the added one. This results in a gain matrix with large entries.

This system however assumes that the transition matrix $\Phi$ is known for all instances, so the knowledge of $x$ at one particular moment is sufficient for determining the complete history of $x$. In practical cases however, this seems a bit unrealistic, or even impossible. Therefore the transition model is taken stochastic as well by introducing the concept of filter divergence. In this case an additional error vector $\underline{m}_k$ is modelled in the time update equations:

$$\begin{aligned} \underline{\hat{x}}_{k|k-1} &= \Phi_{k|k-1}\underline{\hat{x}}_{k-1|k-1} + \underline{m}_k \\ Q_{\hat{x}_{k|k-1}} &= \Phi_{k|k-1}Q_{\hat{x}_{k-1|k-1}}\Phi^*_{k|k-1} + Q_{m_k}. \end{aligned} \tag{4.7}$$

The vector $\underline{m}_k$ denotes the error in the transition model. It is assumed that the predicted estimator cannot be determined exactly by applying the transition model. Because of the assumption that the difference vector $\underline{m}_k$ grows as the time span gets larger, this causes filter divergence. The filter divergence is a sort of measure for the durability of the model; when the filter divergence is 0, the model is applicable for all moments. When the filter divergence is large however, it indicates that the model becomes unreliable with progressing time. The matrix $Q_{m_k}$ is the covariance matrix of this model error. The second step in the process is the measurement update, herefore still the Equations 4.5 can be used.

A problem now is how to determine the filter divergence. Because this is often very difficult, a variance discounting method is proposed [West and Harrison, 1997]. Here, the system variance is derived from the measurement variance as:

$$Q_{m_k} = Q_{\hat{x}_{k|k-1}}(1-\delta)/\delta, \tag{4.8}$$

where $\delta$ is a certain discount factor ($0 < \delta < 1$). Note that in case $\delta = 1$, the filter divergence is zero and the model of Equation 4.7 reduces to that of Equation 4.4, which means that the model perfectly describes the situation. A $\delta$ close to zero can be interpreted as a model that is almost random. The actual value for $\delta$ has to be chosen beforehand; a typical value of the discount factor lies between 0.9 and 0.99 [West and Harrison, 1997].

From Equation 4.5 it becomes obvious that for every time step the matrix $Q_{v_k}$ has to be inverted. As this matrix is of size $m \times m$, where $m$ is the number of observations, this matrix can become very large and high computational effort is needed to calculate the measurement update. However, to make this more efficient the measurement update can also be computed in blocks, as described by [Anderson and Moore, 1979]. The measurements are divided in $N$ blocks with $m/N$ measurements and so sequentially all $N$ blocks of measurements are processed to compute the measurement update. The equations for the measurement update then become:

$$\begin{aligned} K_{k_n} &= Q_{\hat{x}_{k|k-1}}A_{k_n}(Q_{k_n} + A_{k_n}Q_{\hat{x}_{k|k-1}}A^*_{k_n})^{-1} \\ \underline{\hat{x}}_{k|k} &= \underline{\hat{x}}_{k|k-1} + K_{k_n}\underline{v}_{k_n} \\ Q_{\hat{x}_{k|k}} &= [I - K_{k_n}A_{k_n}]\,Q_{\hat{x}_{k|k-1}}, \end{aligned} \tag{4.9}$$

with $n = 1 \ldots N$. As can bee seen in Equations 4.9, an $m/N \times m/N$ matrix has to be inverted, instead of an $m \times m$ matrix. This has to be done for all $N$ blocks.

To start the Kalman filtering procedure, an initial value $\underline{\hat{x}}_{0|0}$ and a corresponding variance $Q_{\hat{x}_{0|0}}$ need to be known. Then the first step is the time update to predict the estimator $\underline{\hat{x}}_{1|0}$ and its variance. When the measurements at time $t = 1$ come available, the estimator can be corrected by applying the measurement update and the process repeats.

## 4.2 Trend analysis model of the North Sea Directorate

The North Sea Directorate has developed a trend model for the sea channel depths. Goal of the project is to gain experience in the prediction of the sea floor depth in the maintenance areas of the North Sea Directorate. This can be used to assist in the decisions concerning the planning of dredging and other maintenance activities, such as seaway marking. In this section a description of this method will be given, based on [Wüst, 2003] and [Wüst, 2004].

The model consists of a set of spatially distributed support points. The support points are a weighted average of all data points. The set of support points defines the interpolation grid, so the support points can be seen as grid points. The weights are calculated with a Gaussian kernel function, based on the distance between the data points and the location of the support points [Lee et al., 2002]:

$$\tilde{w}_{ij} = \frac{exp\{-\frac{1}{2}s_{ij}\mu^{-1}s_{ij}^{*}\}}{2\pi\mu}, \tag{4.10}$$

where $s_{ij}$ is the horizontal distance between the data point $y_i$ and the support point $x_j$ and $\mu$ is distance between the support points. To ensure that the sum of the weights per data point equals 1, the weights are normalised by dividing by the sum of the weights per data point:

$$w_{ij} = \frac{\tilde{w}_{ij}}{\sum_{j=1}^{J}\tilde{w}_{ij}}, \tag{4.11}$$

with $J$ the number of support points.

At each support point a transition model is attributed, consisting of two components: a depth and a trend. Therefore the measurement equations are:

$$\begin{bmatrix} \underline{y}_1 \\ \underline{y}_2 \\ \vdots \\ \underline{y}_I \end{bmatrix} = \begin{bmatrix} w_{11} & \dots & w_{1J} & 0 & \dots & 0 \\ w_{21} & \dots & w_{2J} & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ w_{I1} & \dots & w_{IJ} & 0 & \dots & 0 \end{bmatrix} \begin{bmatrix} d_1 \\ \vdots \\ d_J \\ \frac{\Delta d_1}{\Delta t} \\ \vdots \\ \frac{\Delta d_J}{\Delta t} \end{bmatrix}, \tag{4.12}$$

where $y_1 \dots y_I$ are the measurements (soundings) and $d_1 \dots d_J$ the support points. For every support point two parameters are estimated, namely the depth ($d_j$) and a trend ($\frac{\Delta d_j}{\Delta t}$).

Next step in the model is the time update. This is defined as:

$$\begin{bmatrix} \underline{\hat{d}}_{k|k-1} \\ \frac{\Delta \underline{\hat{d}}_{k|k-1}}{\Delta t} \end{bmatrix} = \begin{bmatrix} I_{J\times J} & I_{J\times J} \\ 0_{J\times J} & I_{J\times J} \end{bmatrix} \begin{bmatrix} \underline{\hat{d}}_{k-1|k-1} \\ \frac{\Delta \underline{\hat{d}}_{k-1|k-1}}{\Delta t} \end{bmatrix} \tag{4.13}$$

This is mainly an identity matrix, the only off-diagonal elements represent the trend component. So the actual change in depth can be calculated from the previous depth and the trend. The model uncertainty is defined by the matrix $Q_{m_k}$, which has the following structure:

$$Q_{m_k} = \begin{bmatrix} Q_{\hat{x}_{k|k-1}}\frac{1-\delta_d}{\delta_d} & \\ & Q_{\hat{x}_{k|k-1}}\frac{1-\delta_t}{\delta_t} \end{bmatrix}, \tag{4.14}$$

where $\delta_d$ and $\delta_t$ are the discount factors for the change in depth and the trend respectively.

Now that the model is defined, it has to be initialised by choosing the initial values $\underline{\hat{x}}_{0|0}$ and $Q_{\hat{x}_{0|0}}$ as:

$$\underline{\hat{x}}_{0|0} = \begin{bmatrix} \overline{d} \\ \vdots \\ 0 \\ \vdots \end{bmatrix} \quad ; \quad Q_{\hat{x}_{0|0}} = \begin{bmatrix} \sigma_d^2 & & & \\ & \ddots & & \\ & & \sigma_{trend}^2 & \\ & & & \ddots \end{bmatrix}. \tag{4.15}$$

| parameter | value |
|---|---|
| support point distance | 17.5 m |
| cut off limit for the interpolation support point | 3.5 × support point distance |
| depth discount factor | 0.93 |
| trend discount factor | 0.93 |
| measurement noise | 0.23 $m^2$ |

**Table 4.1:** Model and calibration parameters of the trend analysis model

The initial values for the depth are the mean depth, derived from the first survey. The initial values for the trend are set to zero. Their variances are $\sigma_d^2$, which is the variance of the first survey, and $\sigma_{trend}^2$, a chosen variance for the trend. In [Wüst, 2004], these variances of the trend is chosen as $\sigma_{trend}^2 = 0.1$ $(m/s)^2$.

Besides the initial state vector and initial variance, some model parameters and calibration parameters have to be defined as well. Examples are the discount factors $\delta$ and the distance between the support points. These are defined as in Table 4.1. The last parameter in this table is the measurement noise. As observations in this model the average depth values of a grid cell are taken. The variances of these observations consist of two parts, namely the standard deviation for the determination of the mean value in the grid cell and the measurement noise. The standard deviation of the grid cell value is determined by:

$$\underline{\hat{\sigma}}_{cell} = \frac{\sum_1^N (\underline{d}_n - \overline{d})}{N}, \tag{4.16}$$

with $N$ the number of points in the grid cell. The second part, the measurement noise is not precisely known, in this case a value of 0.23 $m^2$ is chosen. The measurement noise consists of the accuracy of the observations and some unpredictable defectivenesses in the model, for example the movement of small morphological features such as mega-rippels. This value is chosen after a calibration procedure. It is chosen rather high in order to make sure that the accuracy of the prediction is not exaggerated.

After applying the dynamic model in the support point position, the value of $\underline{\hat{x}}$ can be used to adjust the depth values by applying the measurement equations again:

$$\underline{\hat{y}}_k = A\underline{\hat{x}}_{k|k}. \tag{4.17}$$

These values are now the new depth values at every grid point at a certain moment $k$.

Another application of this model is the determining of the moment when the depth exceeds a certain limit. Three depth limits are defined, with respect to the mean water level:

- the nautical guaranteed depth: 23.4 m,
- the extra depth: 25.7 m,
- the dredging depth: 26.2 m.

The nautical guaranteed depth is the upper limit, guaranteed by the authorities, because the port of Rotterdam has to remain accessible to ships with a draught of 22.5 m. To secure that the channel will always be deeper than this limit the dredging depth is somewhat lower, so it is defined that the channel has to be dredged when the depth is less than 26.2 m.

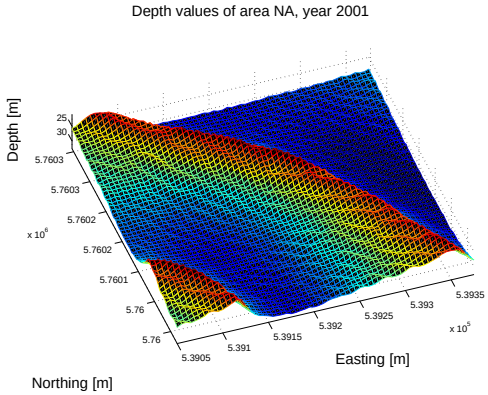



**Figure 4.3:** Depth values of the data set NA, measured in 2001.

## 4.3 Example of the trend analysis model

In this section an example of the trend analysis model will be given. Therefore an area of 330×330 m is selected. The data are given as the average of a 5×5 m grid cell, so 4,356 data points are available. Those are interpolated to 324 support points, the interpolation grid is 18×18 points. Data are available from 9 years, from 1991 until 2001, except for the years 1992 and 1998. The original data are the mean values of 5×5 meter grid cell. A plot of the original data is given in Figure 4.3.

The method of the North Sea Directorate is implemented in a few Matlab scripts, where the Kalman filter is the main part. The output of the program gives, besides the data values, various plots which will be shown below. First output are the profile plots, see Figure 4.4. In the input stage, some lines can be defined along which a depth profile will be calculated. The situation of these profiles is shown in the upper left figure in the plot. The white line indicates the situation of the profile shown below, the red lines indicate the other profiles that are calculated. Shown are the predicted depths for 2002, based on surveys until 2001, together with the standard deviation. Besides the predicted depths the three depth limits (nautical guaranteed depth, extra depth and dredging depth) are depicted.

The next plot gives time series for some selected points. By default, time series are plotted for the fastest rising point, the fastest subsiding point, the most stable point, the shallowest point, the point with the largest predicted standard deviation and the first point that oversteps the nautical guaranteed depth. Here the plot of the time series for the fastest rising point is shown, see Figure 4.5. The fastest rising point is determined by comparing the measured depth of the last survey and the predicted depth for the first year after the last survey. Shown are the soundings for the different years and the predictions. Besides this also profiles in North-South and East-West directions through this point are shown.

The third plot gives the moments when the different depth limits are overstepped. In Figure 4.6 the predicted moments when the dredging depth is overstepped are shown. These moment are predicted with a 50 % probability. It can be seen that the dredging depth is predicted to be overstepped for the first time in the first quarter of 2005.

The last plot is the diagnosis plot, which gives some information about the quality of the prediction. The lower left figure of the diagnosis plot in Figure 4.7 gives the differences between the measured depths and the predicted depths for the last year.

In the lower right figure a QQ-plot (quantile-quantile plot) of the uncertainty is given. This is a graphical technique for determining whether two data sets have the same distribution [Isaaks and Srivastava, 1989]. It plots the quantile from the one data set against the other one, where a quantile is the fraction of points that fall below a certain value. For example the 30 % quantile is the point where 30 % of the data fall

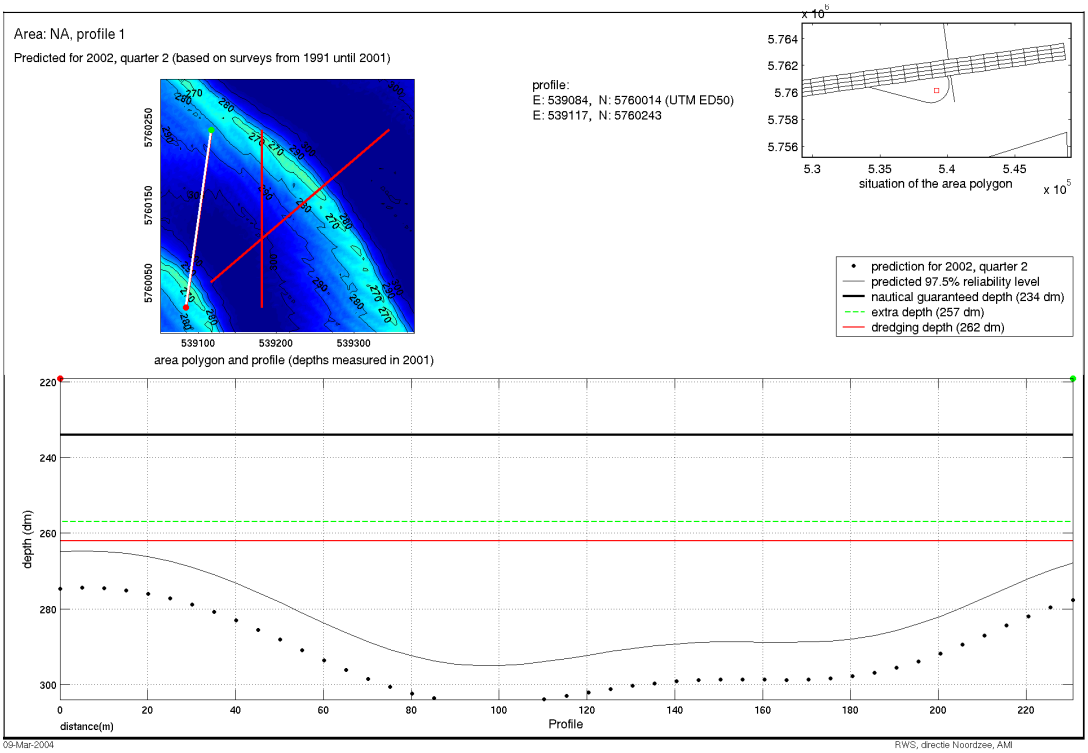



**Figure 4.4:** A profile plot for the area NA. The situation of this profile is indicated by the white line, in the overview figure in the upper left corner.

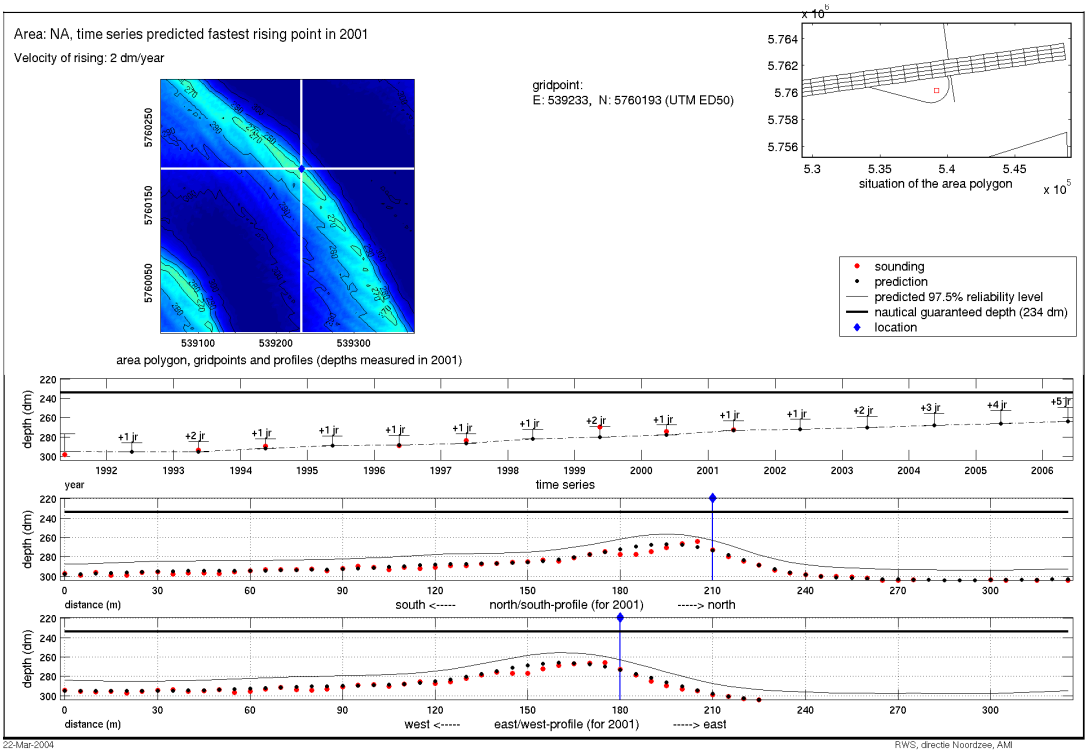



**Figure 4.5:** The time series plot for the fastest rising point in area NA. This point is located at the blue diamond in the upper left figure.

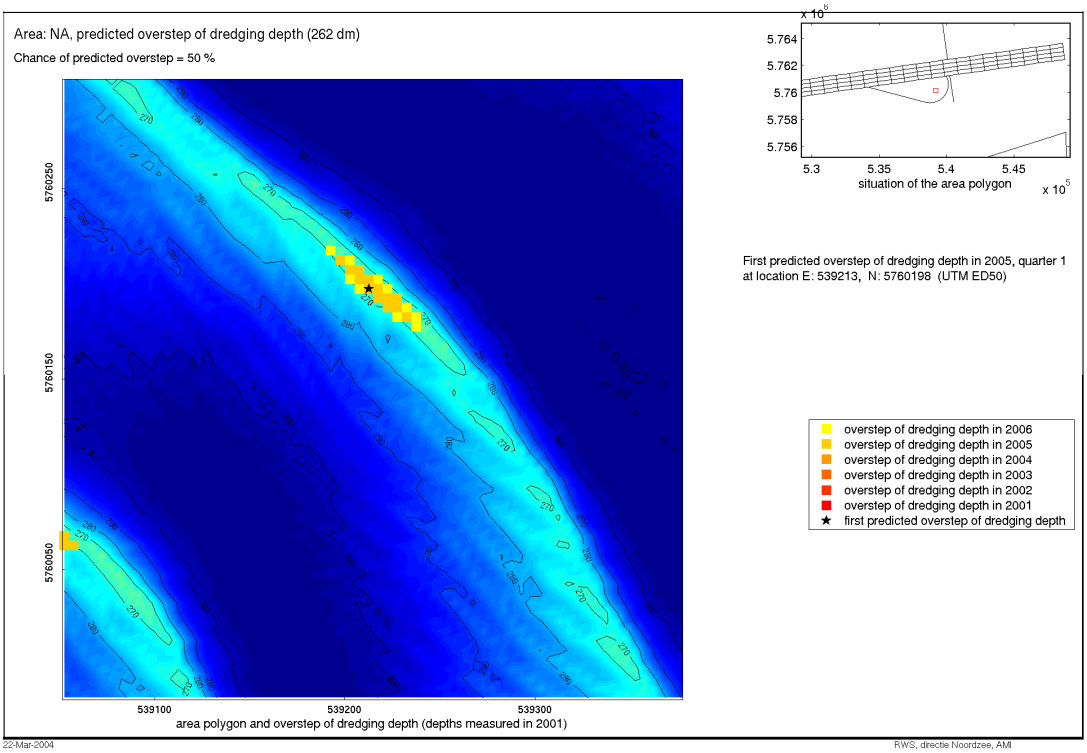



**Figure 4.6:** The prediction of the moment when the dredging depth (26.2 m) is overstepped, with a 50 % chance.

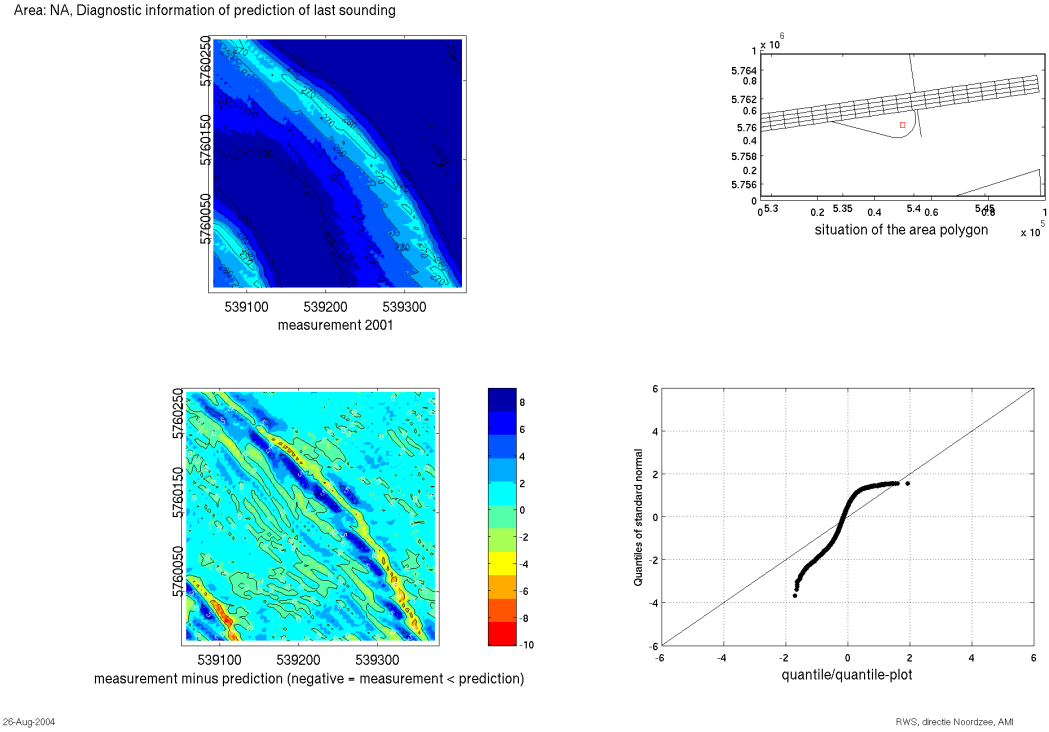



**Figure 4.7:** The diagnosis plot which gives information about the quality of the prediction.

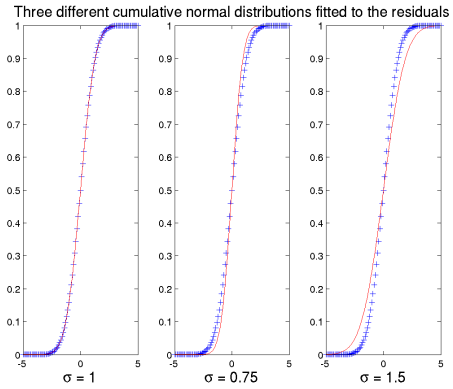



**Figure 4.8:** Example of residuals and three different cumulative normal distributions. The blue + signs are the residuals, the red line the cumulative normal distribution. In the left figure, the normal distribution fits the residuals exactly. In the middle figure, the standard deviation of the normal distribution is too small, resulting in a probability density function that is too narrow. In the right figure the standard deviation is too high, resulting in the probability density function to be too wide.

below that value and 70 % fall above that value. [NIST/SEMATECH, 2003]. When the two data sets come from the same distribution, the points will approximately fall on a 45° reference line; if not, the points will deviate much from this line.

In this case, the QQ-plot is used as a probability plot: instead of plotting two data sets against each other the residuals of the predicted depths are plotted against the normal distribution. So when the dots fall on the reference line, the uncertainty of the prediction is normally distributed. If however the line bends to the right of the reference line, to many points fall outside the predicted distribution. This means that the predicted distribution is too narrow, the predicted standard deviation is too small so the quality of the prediction is lower than expected. On the other hand, when the line bends to the left, too many residuals fall within the the distribution so the quality of the prediction is higher than expected.

In Figure 4.8 these situations are illustrated. The blue + marks denote the residuals, the red line shows the cumulative normal distribution for three different standard deviations. In the left figure, the predicted standard deviation is right, therefore the points in the QQ-plot should be around the reference line. In the middle figure, the situation that the QQ-plot bends to the right is illustrated. The predicted standard deviation is too high. The fact that the line bends to the right can be seen by comparing the values of the residuals and the predicted distribution at the equal $y$-level. For $y$ higher than 0.5, the residual value is higher than the cumulative distribution value, so if these values are plotted against each other the line bends to the right. In the right figure, the case that the standard deviation is too high is illustrated.

## 4.4 Summary and discussion

The core of the method of the North Sea Directorate is the interpolation of the data to several support points (support points) and attributing a linear growth model to every support point. These local growth models consists of two components: a depth component and a trend component. The following steps in the procedure can be discerned:

- Reading in of the data. The data are given as average values of a grid cell on a regular North-South oriented grid. Besides the coordinates (in Easting-Northing, ED-50) and the depth value, also a standard deviation per grid cell is given. This is the standard deviation of all points within a grid cell, used for calculation of the average value. Therefore this value does not completely describe the precision of the measurements, because it also contains the natural variability of the sea floor. It gives however an indication of the precision of the measurements. All data are acquired with

multibeam, so a complete sea floor coverage is present.

- Initialisation. Before applying the model, some initial values and calibration parameters have to be defined. These parameters are obtained from an experimental calibration procedure. Therefore the outcome of the model very much depends on the choice for the initial values and calibration parameters. Three different types of parameters have to be defined:

  1. Model parameters: these are both the discount factors and the measurement noise, see Table 4.1. The first two parameters represent the uncertainty of the trend model, the last one represent the uncertainty in the measurements. This is not only the accuracy of the measurement itself, but also some variability of the sea floor, especially presence of mega-rippels. The mega-rippels are dynamic features, they vary quite a lot through, or within the years. Therefore they are considered measurement noise, added to the actual sea floor model.
  2. Initial state: In order to apply a Kalman filter, some initial values have to be chosen. In this case, initial values have to be chosen for the depth and the trends. The initial values for the depth are the mean value of the first survey, the trend is set to zero. Furthermore, initial values have to be chosen for their variances. The variance of the depth is derived from the data as well, the variance of the trend is set to a predefined value.
  3. Model structure: these parameters define the location and number of support points. This is defined by the grid distance and the support point width.

- Processing of all epochs. This can again be divided into two steps, the measurement update and the time update:

  1. Interpolation of the data points to the grid with support points, see Equation 4.12. This is in fact the measurement update, so this gives estimates for the depths and trends, based on observations until the epoch that is being processed.
  2. Applying the time update in each grid point, see Equation 4.13. For each grid point now a prediction for the change in level and a trend, as well as the average depth of the area is calculated, based on all available epochs. The trend model is only calculated in the grid points, for each grid point now predictions are available.

  When applying this model for the first time, all epochs have to be processed. All epochs are treated equally, there is no testing on deviating surveys. This means that once a (systematic) error is present in one of the surveys, this automatically effects all other estimates and therefore the model performance. At every epoch, estimates for the depth and trends at that time, and predicted values for one year ahead, are calculated. When however new measurements come available, the estimation and prediction can be made on basis of only the last epoch. Not all epochs have to be processed again. The interpolation of the data points to the set of support points is done by a Gaussian model. That is the interpolation weights are only based on the distance between the data point and the grid point. No spatial correlation has been taken into account, although the sea floor is generally not isotropic. The result of this interpolation may be an oversmoothing because important features, such as sand waves may not be preserved. Instead of using this method, also a Kriging method can be used, which does not assume isotropy.

- Prepare output. This also consists of several items:

  1. Output on the original grid: for each data point the estimated and predicted value is calculated. This is done with the same interpolation method, now applied to obtain a new value for each data point, based on the values of the support points.
  2. Output on profiles: for several predefined profiles the depths are calculated.
  3. Prediction of the moment when the various nautical limits are overstepped. This is done by means of the predicted time series for every data point.

- Plot analysis. The various plots, described in Section 4.3, which show some profiles and time series. Only the results of the prediction in the various grid points are given. No conclusions about whether a point is moving or not are made and the output is only given for individual points.

# Chapter 5

# Simulations

In Chapters 3 and 4 the methods of the Hydrographic Service and the North Sea Directorate were described. To illustrate the procedure some examples with real data were given. However the use of real data is not very suitable for comparing the two methods because it is unknown what the result should be. Therefore, the two methods are tested on simulated data. In this case the outcome is know beforehand, so it can be checked whether the results are correct. In this chapter different situations are simulated to test the performance of the two methods:

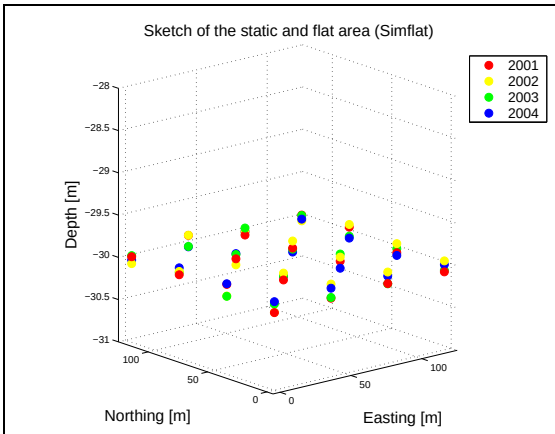
- A static and flat area (Section 5.1),
- A static and flat area with an outlying survey (Section 5.2),
- A flat area with a single outlying point in one survey (Section 5.3),
- An area with a static slope (Section 5.4),
- A flat area with an upwards trend (Section 5.5),
- An area with a static sand wave (Section 5.6),
- An area with a migrating sand wave (Section 5.7).

This chapter will be concluded in Section 5.8 with a summary of the conclusions that can be drawn from the simulations.

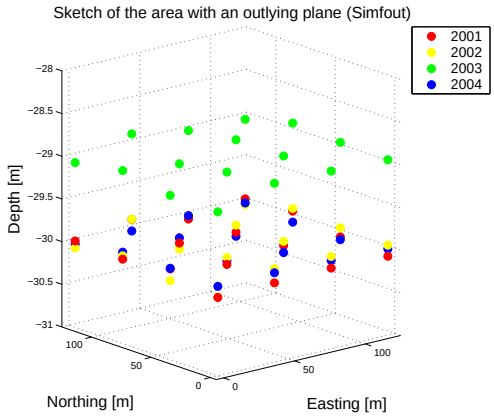
In each section a data set is created, which is first tested on the Hydrographic Service' method and then on the North Sea Directorate method. An illustration of the first five cases is given in Figure 5.1. The last two cases deal with sand waves, they are illustrated in the corresponding sections. In order to make the simulations easy to understand, only very small data sets are simulated. Besides this, the different deformations were chosen to be large, to make the results more obvious. Furthermore, it is assumed that the data are gridded and the variances and covariance function are known.

## 5.1 Simulation of a static flat area

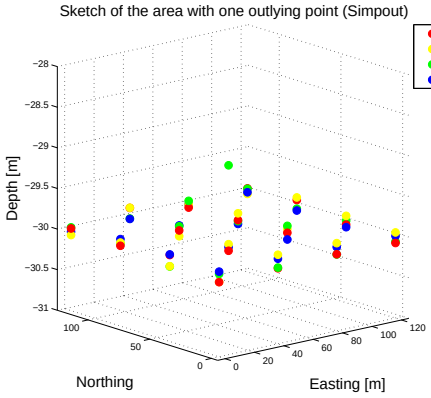
In this section, a flat area of 4×4 points is simulated, with a mean depth of 30 m. It is assumed that it is "measured" in 4 epochs. To make sure that the data sets resemble the reality, some random noise is added. Furthermore, it is assumed that all points have an equal standard deviation of 20 cm. The grid size is set to 40×40 m. The simulated data set is given on page 47. This data set will be the starting point of all simulations, in further sections different kinds of deformation will be added to this data set. All depth values in this chapter are given in meters, all variances in $m^2$.



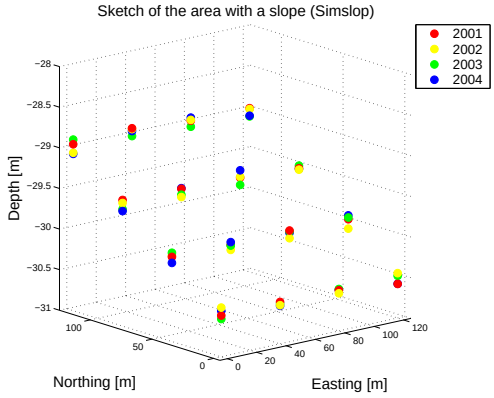

a) Static and flat



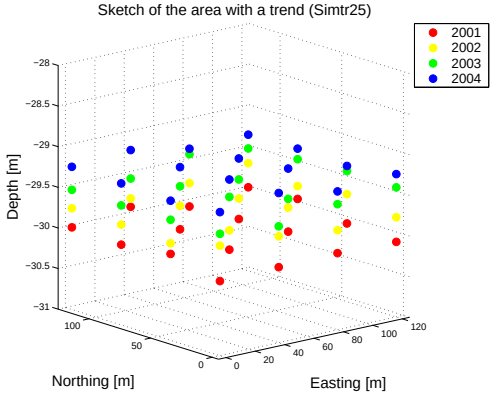

b) Outlying survey




c) One outlying point




d) Slope




e) Upwards trend

**Figure 5.1:** The various cases that are simulated.

$$\underline{y}_{2001}: \begin{bmatrix} 29.9972 & 29.9037 & 30.0584 & 29.9811 \\ 30.0214 & 29.9913 & 30.0231 & 29.9353 \\ 29.9462 & 30.0524 & 29.9889 & 30.0476 \\ 30.0900 & 30.0783 & 30.0643 & 30.0844 \end{bmatrix}$$

$$\underline{y}_{2002}: \begin{bmatrix} 30.0787 & 29.9020 & 30.0208 & 30.0494 \\ 29.9821 & 30.0626 & 29.9397 & 29.9031 \\ 30.0834 & 29.9706 & 29.9406 & 29.9398 \\ 30.0871 & 29.9116 & 29.9278 & 29.9544 \end{bmatrix}$$

$$\underline{y}_{2003}: \begin{bmatrix} 29.9837 & 30.0344 & 29.9759 & 29.9858 \\ 29.9932 & 29.9405 & 30.0363 & 30.0419 \\ 30.0864 & 30.0050 & 29.9039 & 30.0006 \\ 29.9890 & 30.0692 & 30.0676 & 30.0664 \end{bmatrix}$$

$$\underline{y}_{2004}: \begin{bmatrix} 30.0364 & 30.0396 & 30.0187 & 30.0290 \\ 29.9387 & 29.9302 & 30.0707 & 30.0643 \\ 29.9379 & 30.0083 & 30.0720 & 30.0800 \\ 29.9609 & 29.9606 & 29.9757 & 29.9993 \end{bmatrix}$$

### 5.1.1 Test with deformation analysis

First, this data set is evaluated with the deformation analysis procedure as in Section 3.2. To illustrate this procedure, the calculation is fully elaborated in Appendix B.

In Table 5.1 the test results for this area are given. Because an alternative hypothesis is accepted when the test quantity $\underline{T}_q$ is larger than the critical value $k_\alpha$, it can be seen that no alternative hypothesis is accepted. Therefore all points are classified as static points, see Figure 5.2. This is of course right, because a static plane was modelled. The Minimal Detectable Biases for the simulation are given in Table 5.2. These values are valid for all simulations in this chapter, except the simulations of a sand wave. Because all measurements in all epochs are assumed to have the same standard deviation, the sea floor depth is simply estimated as the mean of all epochs:

$$\underline{\hat{d}} = \begin{bmatrix} 30.0240 & 29.9699 & 30.0184 & 30.0113 \\ 29.9838 & 29.9812 & 30.0175 & 29.9861 \\ 30.0135 & 30.0091 & 29.9764 & 30.0170 \\ 30.0318 & 30.0049 & 30.0088 & 30.0261 \end{bmatrix} \quad ; \quad \sigma^2_{\hat{d}} = 0.01. \tag{5.1}$$

Note that the grid point (1,1) coincides with matrix element (4,1). These values are also used for the prediction, see Figure 5.3. Because no deformation is present, the predicted values are simply the same as the adjusted values, and the accuracy is the same as the measurement accuracy.

| | **Point test** | | | **Area test** | | |
|---|---|---|---|---|---|---|
| **Type of deformation** | $\underline{\mathbf{T}}_\mathbf{q}$ | $\mathbf{k}_\alpha$ | $\underline{\mathbf{T}}_\mathbf{q}/\mathbf{k}_\alpha$ | $\underline{\mathbf{T}}_\mathbf{q}$ | $\mathbf{k}_\alpha$ | $\underline{\mathbf{T}}_\mathbf{q}/\mathbf{k}_\alpha$ |
| Outlying value in survey: 1 | 0.1132 | 6.63 | 0.0171 | 0.4913 | 11.34 | 0.0433 |
| 2 | 0.1021 | 6.63 | 0.0154 | 0.9904 | 11.34 | 0.0873 |
| 3 | 0.0610 | 6.63 | 0.0092 | 0.0987 | 11.34 | 0.0087 |
| 4 | 0.1674 | 6.63 | 0.0252 | 0.7386 | 11.34 | 0.0651 |
| General deformation | 0.3325 | 7.81 | 0.1087 | 1.7393 | 16.92 | 0.1028 |
| Trend | 0.2945 | 2.71 | 0.0426 | 0.5653 | 6.25 | 0.0904 |

**Table 5.1:** The test quantities, critical values and normalised test quantities of the point test for point (1,1) and the area test of the example data set.

| | Point test | Area Test |
|---|---|---|
| Outlying survey | 0.789 m | 0.2270 m |
| Trend | 0.223 m/yr | 0.0663 m/yr |

**Table 5.2:** Minimal Detectable Biases for the simulations in this chapter (except for the simulations with a sand wave).

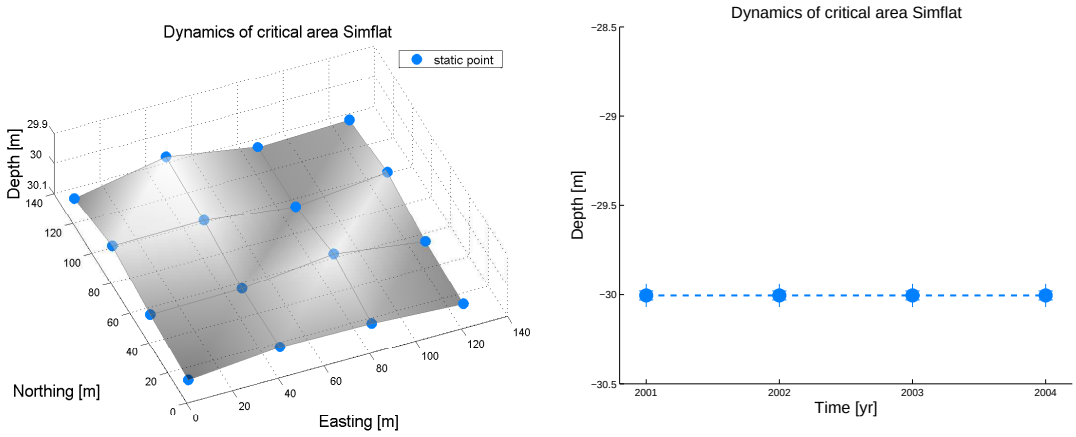



**Figure 5.2:** Results of the deformation analysis for the simulated data of a static and flat area. On the left the results for the point test, on the right the results for the depth in the area test. It is obvious that no deformation is detected with the point test as well as with the area test.

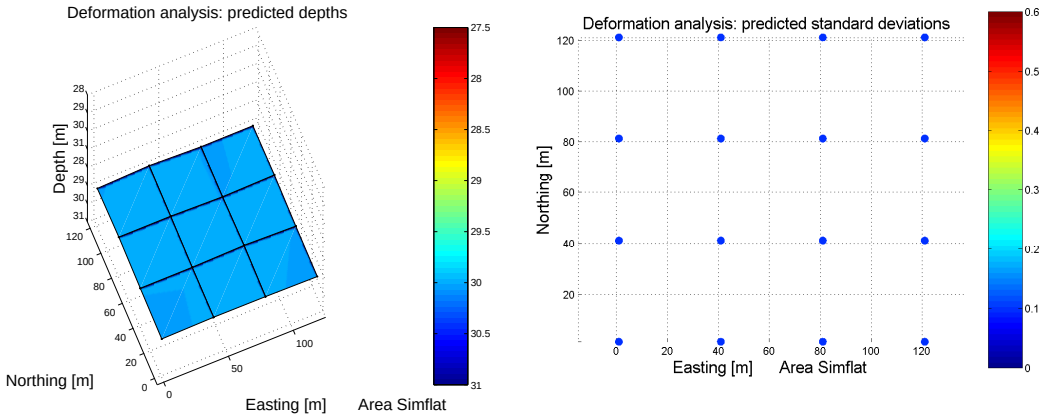



**Figure 5.3:** Predicted values and standard deviations for the static and flat area. Because no deformation is detected the predicted values are the same as the adjusted values and the predicted standard deviations are the same as the standard deviations of the measurements.

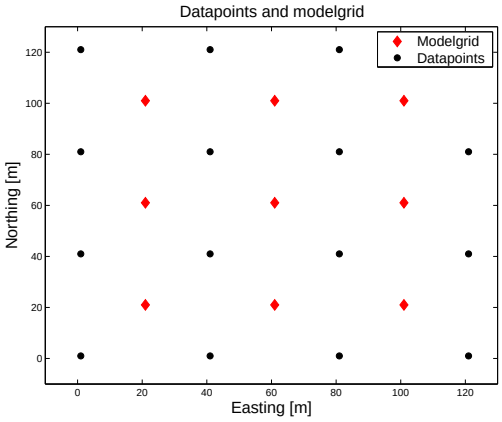



**Figure 5.4:** The data points and model grid for the simulated data.

### 5.1.2 Test with Kalman filtering

Now the same area serves as input for the Kalman filtering model. This calculation is elaborated in Appendix C. The area consists of 4×4 equally spaced data points. The model grid, with the interpolation support points, is defined as a grid with 3×3 points in between the data points, see Figure 5.4. The Kalman equations are only calculated in the support points. The use of a model grid is convenient when a large data set is used, because it reduces the number of equations to be calculated. In this example it is in fact not necessary, but it is interesting to see the effects of this interpolation.

As described in Section 4.2, the procedure starts with the choice of initial values for the depth and trends. In this example the initial values are:

$$\underline{\hat{d}}_{00|00} = \begin{bmatrix} 30.0165 \\ 0 \end{bmatrix} \quad ; \quad Q_{\hat{d}_{00|00}} = \begin{bmatrix} 0.0032 & 0 \\ 0 & 0.1 \end{bmatrix}, \tag{5.2}$$

because the mean depth is 30.0165 m and the standard deviation of the depth values is 0.0032 m. This standard deviation is calculated a posteriori from the measurements. Then by turns the time update and measurement update are calculated. After 4 years, the depth values and variances are:

$$\underline{\hat{d}}_{04|04} = \begin{bmatrix} 30.0128 & 30.0105 & 30.0211 & 30.0393 \\ 29.9869 & 29.9920 & 30.0140 & 30.0421 \\ 29.9769 & 29.9803 & 30.0031 & 30.0340 \\ 29.9877 & 29.9812 & 29.9937 & 30.0189 \end{bmatrix}, \tag{5.3}$$

$$Q_{\hat{d}_{04|04}} = \begin{bmatrix} 0.0177 & 0.0103 & 0.0103 & 0.0177 \\ 0.0103 & 0.0060 & 0.0060 & 0.0103 \\ 0.0103 & 0.0060 & 0.0060 & 0.0103 \\ 0.0177 & 0.0103 & 0.0103 & 0.0177 \end{bmatrix}. \tag{5.4}$$

These results can be compared with the results of the deformation analysis, see Equation 5.1. In Figure 5.5 the time series of point (1,1) are given and the predicted values for 2009. From the time series, it can be seen that a very small upwards trend is estimated for this point, although the area is static. This is because the data are treated deterministically, whereas in fact they are stochastic. In Figure 5.6 the predicted depths and standard deviations for 2009 are given. When these results are compared to the results of the deformation analysis it can be seen that the predicted depths are equal, at about 30 m. However, in the deformation analysis the predicted standard deviation is the same as the measurement accuracy, while in Kalman filtering the standard deviation of the prediction grows with time, compare with Figure 5.3.

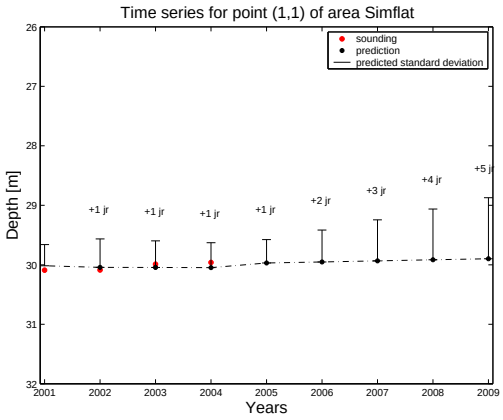



**Figure 5.5:** The time series for the first point, at location (1,1), of the simulated static flat area.

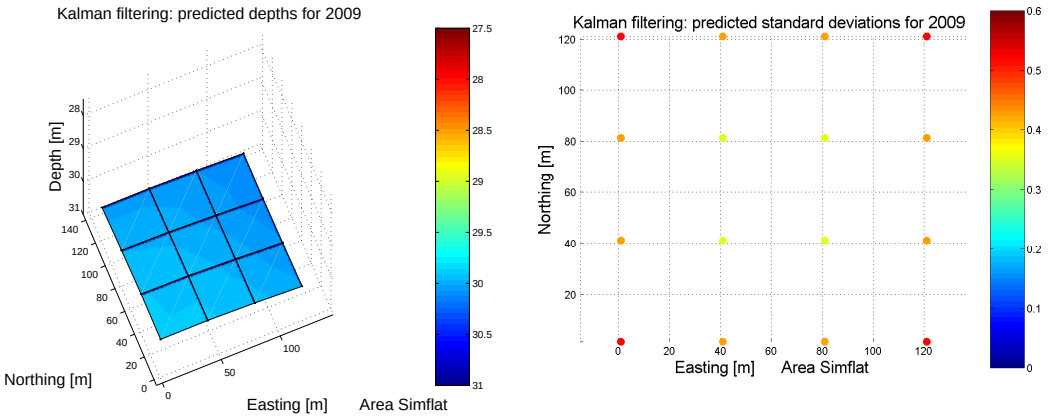



**Figure 5.6:** Predictions for the static flat area, obtained with Kalman filtering. On the left the predicted depths for 2009. On the right the predicted standard deviations for 2009.

### 5.1.3 Conclusions

Two conclusions can be drawn from these two results:

- Although the precise output values of both methods are not exactly the same, they both result in a flat area of 30 m with some random noise of only a few centimetres. This is about the same as the input, only the noise has been reduced with a factor 4. That is of course also what one would expect, because no changes in depth or any other effects have been modelled.
- The average variance of the output of both methods is about the same, but in the deformation analysis all points share the same variance, whereas with the Kalman filtering method the points on the edges get a larger variance. This is due to the interpolation of the support points which results in a smoothing over the area. Besides, in Kalman filtering the variances grow then the time span gets larger, whereas in deformation analysis the variances remain the same as the measurement accuracy.

## 5.2 Simulation of a flat area with an outlying survey

Now consider the situation that one survey deviates from the others. Due to, for example a systematic error, all points of that year are about 1 meter too shallow. The same data set as on page 47 is used, but survey $\underline{y}_{2003}$, the third epoch, is replaced by:

$$\underline{y}_{2003} = \begin{bmatrix} 29.0787 & 28.9020 & 29.0208 & 29.0494 \\ 28.9821 & 29.0626 & 28.9397 & 28.9031 \\ 29.0834 & 28.9706 & 28.9406 & 28.9398 \\ 29.0871 & 28.9116 & 28.9278 & 28.9544 \end{bmatrix}.$$

### 5.2.1 Test with deformation analysis

The data are evaluated with the same procedure as in Paragraph 5.1.1, but in this case the residuals are much higher, because of the outlying survey. The estimate for the first point is:

$$\hat{\underline{d}}(1,1) = 29.8063 \quad ; \quad Q_{\hat{d}} = 0.01. \tag{5.5}$$

The test statistics for this case are given in Table 5.3.

The test quantities of the area test are much larger, because the minimal detectable bias for the area test is much smaller, see Figure 3.8. In fact all normalised test quantities in the area test are larger than 1. However, after accepting the hypothesis of an outlying survey in survey 3, adjusting the model and recalculating the test quantities, they all drop below 1. The largest test quantities are $\underline{T}_{q,gen}$, but because these tests also have larger critical values than the outlier tests the alternative hypothesis of an outlying survey in survey 3 is accepted, see Figure 5.7. The predicted values are shown in Figure 5.8. It can be seen that the predicted depths are about the same as the adjusted depths, at about 30.25 m,

| | **Point test** | | | **Area test** | | |
|---|---|---|---|---|---|---|
| **Type of deformation** | $\underline{\mathbf{T}}_\mathbf{q}$ | $\mathbf{k}_\alpha$ | $\underline{\mathbf{T}}_\mathbf{q}/\mathbf{k}_\alpha$ | $\underline{\mathbf{T}}_\mathbf{q}$ | $\mathbf{k}_\alpha$ | $\underline{\mathbf{T}}_\mathbf{q}/\mathbf{k}_\alpha$ |
| Outlying value in survey: 1 | 2.6836 | 6.63 | 0.4048 | 39.0846 | 11.34 | 3.4466 |
| 2 | 2.6285 | 6.63 | 0.3965 | 30.1567 | 11.34 | 2.6593 |
| 3 | 17.2412 | 6.63 | 2.6005 | 311.4731 | 11.34 | 27.4668 |
| 4 | 0.7971 | 6.63 | 0.1202 | 36.7705 | 11.34 | 3.2425 |
| General deformation | 17.5128 | 7.81 | 2.2424 | 313.1137 | 16.92 | 18.5055 |
| Trend | 2.4058 | 2.71 | 0.8877 | 21.6616 | 6.25 | 3.4659 |

**Table 5.3:** The test quantities, critical values and normalised test quantities for the point test for point (1,1) and the area test of the flat area with an outlying survey.

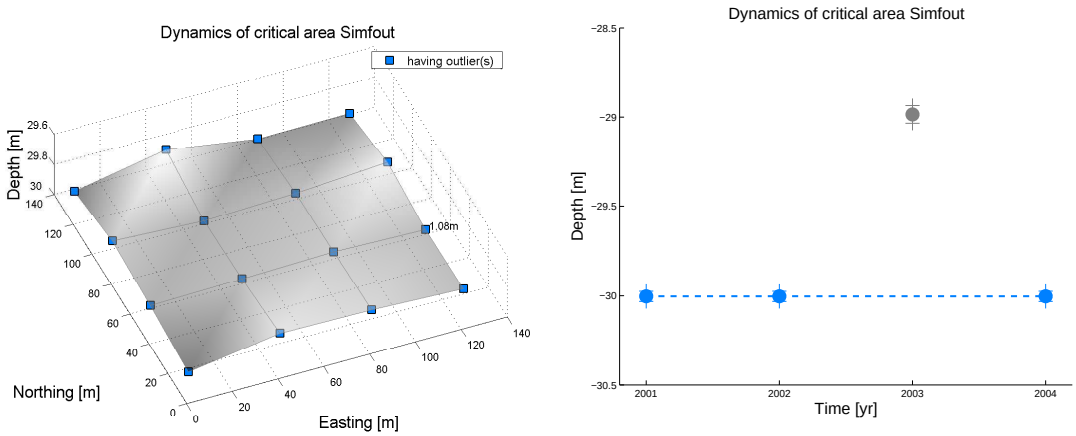



**Figure 5.7:** Results of the deformation analysis for the simulated data of a flat area with an outlying survey in survey 3. On the left the results for the point test, on the right the results for the depth in the area test.

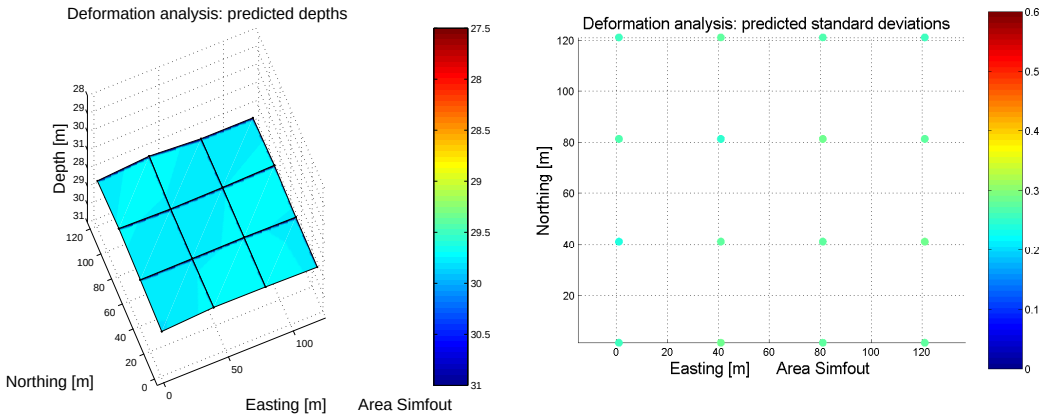



**Figure 5.8:** Predicted values and standard deviations for the flat area with an outlying survey, obtained with the deformation analysis.

so the outlying survey is taken into account. This is because the predicted depths are calculated as, see Paragraph 3.2.5:

$$\underline{\hat{d}}_{pred} = p_{outlier} \cdot \underline{\hat{d}}_{outlier} + (1 - p_{outlier}) \cdot \underline{\hat{d}}_{normal}, \tag{5.6}$$

where $p_{outlier}$ is the number of outlying measurements divided by the total number of measurements. In this case one entire survey is classified as outlier, so $p_{outlier}$ is 0.25. Therefore in this case the magnitude of the predicted depths is about the same as the average of the four epochs. The outlying survey is still taken into account because it might be the result of a real change in the sea floor at some time. For example after a storm the sea floor can change considerably and these effects may repeat at the same place at some moment. Therefore, it is also acceptable to have a lower confidence in the depth at a point where an outlier is detected, so the standard deviation of the prediction will increase.

Next the situation that the fourth survey is deviating, instead of the third, is considered. The results of this simulation are quite interesting, see Figure 5.9 and Table 5.4. In this case none of the points are correctly classified as static points with an outlier. Instead, all points are having a trend. In fact, two different situations are present: points with a rather large upwards trend, about 30 cm/year and points with a very small trend (upwards or downwards) and an outlier. To explain this, the time series of points

| | | Point 1 | | Point 3 | | Area test | |
|---|---|---|---|---|---|---|---|
| **It.** | **Type of deformation** | $\underline{T}_q$ | $\underline{T}_q/k_\alpha$ | $\underline{T}_q$ | $\underline{T}_q/k_\alpha$ | $\underline{T}_q$ | $\underline{T}_q/k_\alpha$ |
| I | Outlying value in survey: 1 | 2.5389 | 0.3829 | 2.6751 | 0.4035 | 36.0625 | 3.1769 |
| | 2 | 2.4853 | 0.3749 | 0.7182 | 0.1083 | 28.4493 | 2.5088 |
| | 3 | 1.0206 | 0.1539 | 2.7386 | 0.4131 | 34.4569 | 3.0385 |
| | 4 | 17.4733 | 2.6355 | 17.1224 | 2.5826 | 292.7953 | 25.8197 |
| | General deformation | 17.6385 | 2.2585 | 17.4408 | 2.2331 | 293.7960 | 17.3638 |
| | Trend | 11.9976 | 4.4272 | 10.2256 | 3.7733 | 176.8585 | 28.2974 |
| II | Outlying value in survey: 1 | 2.9681 | 0.4477 | 1.7698 | 0.2669 | 47.2958 | 4.1707 |
| | 2 | 0.4982 | 0.0752 | 0.0000 | 0.0000 | 4.3733 | 0.3857 |
| | 3 | 3.8865 | 0.5862 | 6.5925 | 0.9943 | 92.6991 | 8.1745 |
| | 4 | 5.6032 | 0.8451 | 6.8970 | 1.0403 | 115.9936 | 10.2287 |

**Table 5.4:** The test quantities and normalised test quantities for the point test for point 1 (1,1), point 3 (81,1) and the area test of the data set with the outlying plane in survey 4. The first two iterations are shown, the test for point 1 is accepted after the first iteration, the test for point 3 and the area test after the second iteration.

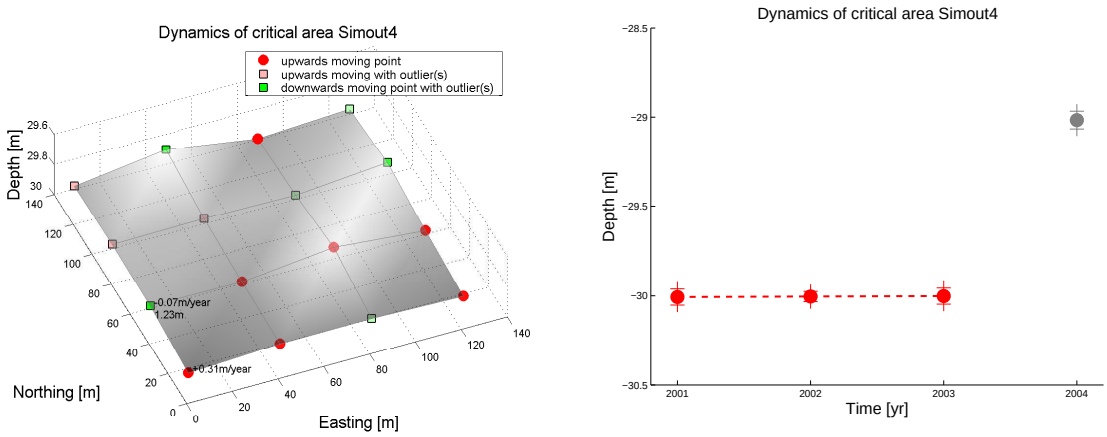



**Figure 5.9:** Results of the deformation analysis for the simulated data of a flat area with an outlying plane in survey 4. On the left the results for the point test, on the right the results for the depth in the area test. In both cases, first a trend is accepted, thereafter in some cases the outlying survey as well.

(1,1) and (81,1) are illustrated in Figure 5.10.

Point (1,1) is an example of a point where only an upwards trend is estimated. First the measurements are adjusted, resulting in the blue line. Because this line does not fit the data, the model is extended with a trend. When adjusting the measurements with the extended model this results in the red line, which fits the data well enough. So the procedure is finished and the model is only extended with a trend.

Point (81,1) is an example where a trend and an outlier are detected. Again the measurements are adjusted with the initial model, resulting in the blue line. Then the model is extended with a trend, resulting in the red line. However this model still does not describe the data accurately, so it is extended with an outlier in survey 4. When adjusting the model with a trend and the outlier, the trend is only estimated through the first three surveys and almost completely disappears. This is indicated by the green line.

So the problem is that in the first iteration a trend is estimated, while in the second iteration, when the outlier is detected, the trend is not significant anymore. The trend is however still visible in the output, the points are still coloured red or green, although this trend is negligible. This can be solved in two ways:

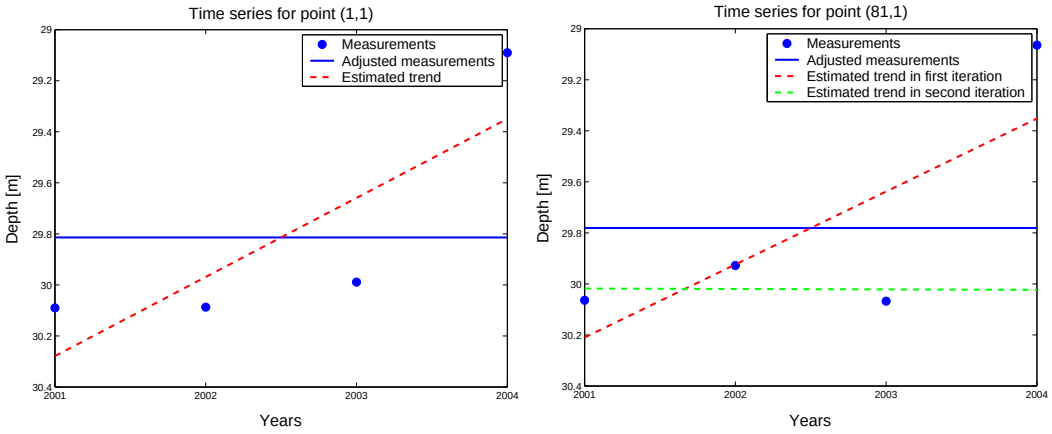



**Figure 5.10:** Illustration of the deformation analysis procedure for two points. On the left the estimation of point (1,1), resulting in an upwards trend of about 30 cm/year is shown. On the right the estimation of point (81,1), resulting in a downwards trend and an outlying value.

| **Type of points** | $\alpha = 0.01$ | $\alpha = 0.02$ | $\alpha = 0.03$ | $\alpha = 0.04$ | $\alpha = 0.05$ |
|---|---|---|---|---|---|
| Moving upwards | 7 | | | | |
| Moving upwards with an outlier | 3 | 10 | 10 | 3 | |
| Moving downwards with an outlier | 6 | 6 | 3 | | |
| Static points with an outlier | | | 3 | 13 | 16 |

**Table 5.5:** Results of the point stability test for various levels of significance of the outlier test. Displayed are the number of points (out of 16) that are classified as a certain type. Only when $\alpha$ for outliers is set to 0.05 or more all points are correctly classified.

- The procedure could be extended, such that in case an outlier is detected, it is tested whether the trend is still significant. If not, the trend extension could be removed again.
- To this point, the levels of significance for an outlier and a trend are 1% and 10% respectively, so a trend is far more likely to be accepted than an outlier. The level of significance of an outlier could be increased, such that an outlier is more likely to be accepted.

When an accurate sea floor model is required, the trend should in fact not be detected at all, because only an outlier is present. Therefore the second solution is preferred. An insignificant trend might still be found, but this is less likely to occur. In Table 5.5, the results of a test with various values for the level of significance of an outlier ($\alpha_{outlier}$) are given. The initial value for $\alpha$ is 0.01. When doubling the value of $\alpha$ to 0.02, the outlying value is detected in every point, but only when $\alpha$ is set to 0.05 (or more) all points are correctly classified as static points with an outlying value. With the area test an $\alpha$ of 0.02 already results in the correct hypothesis of a static plane with an outlying survey. This is confirmed by [Dorst, 2004b], where for new surveys an $\alpha$ of 0.05 is proposed. For planning purposes however, it makes sense to choose a trend, because it is not sure whether it is an outlier or a trend.

### 5.2.2 Test with Kalman filtering

This area is also tested using the same procedure as in Paragraph 5.1.2. In Figure 5.11 the time series for the same point as in Figure 5.5 are presented. It can be seen that the outlying survey causes the trend line to bend upwards. Although the trend line recovers itself with the last (correct) sounding, an upwards trend remains, which is in fact incorrect. Therefore the depth values after four years are:

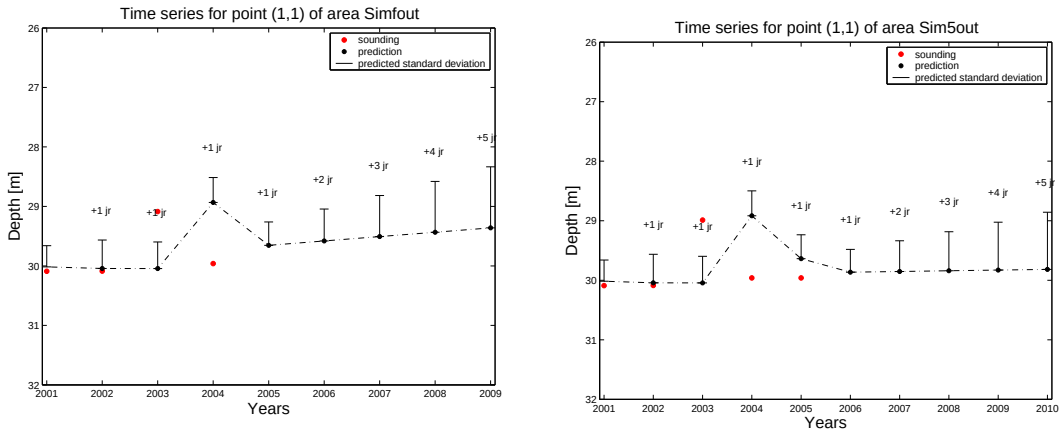



**Figure 5.11:** The time series for the first point, at location (1,1), of the simulated flat area with an outlying plane in survey 3. On the left the time series when 4 epochs are available, on the right the case when 5 epochs are available.

$$\underline{\hat{d}}_{04|04} = \begin{bmatrix} 29.7513 & 29.7524 & 29.7621 & 29.7757 \\ 29.7437 & 29.7517 & 29.7641 & 29.7748 \\ 29.7343 & 29.7319 & 29.7377 & 29.7490 \\ 29.7277 & 29.7025 & 29.6958 & 29.7109 \end{bmatrix}. \tag{5.7}$$

In this case the mean predicted depth is about 29.75 m, whereas in fact they should be the same as in Equation 5.3. The standard deviations are still the same as in Equation 5.4, because they are not affected by the outlying survey:

$$Q_{\hat{d}_{04|04}} = \begin{bmatrix} 0.0177 & 0.0103 & 0.0103 & 0.0177 \\ 0.0103 & 0.0060 & 0.0060 & 0.0103 \\ 0.0103 & 0.0060 & 0.0060 & 0.0103 \\ 0.0177 & 0.0103 & 0.0103 & 0.0177 \end{bmatrix}. \tag{5.8}$$

The effect of the erroneous upwards trend line becomes more obvious when a larger time interval is evaluated. In Figure 5.12 the predicted values and standard deviations for 2009 are given. In this case it is predicted that the area will shoal about 0.5 m whereas in fact it remains static. This could result in a decision to survey or dredge the area, whereas in fact it is not yet necessary. When however more correct surveys are available, the effect of the outlier decreases. In the right graph of Figure 5.11 the time series for the same point are given, but with 5 surveys. The upwards bending of the trend line significantly decreases.

When the data set with the outlying fourth survey is processed, the upwards trend becomes even more obvious, see Figure 5.13. In Figure 5.14 the prediction for 2009 is given. As a result of the large upwards trend, the predicted depth is about 28 m, which is about 2 m too shoal.

### 5.2.3 Conclusions

Summarising this section:

- With the deformation analysis an outlying value is in principle detected very well, except for outliers in the last survey. The influence of an outlier in the last survey may however decrease when more epochs are available. However, the outlying values still influence the prediction. When the outlier is a result of an error in the measurements, it should be removed. However, the outlier can also be a real deformation, therefore it is still included in the prediction. For the standard deviations however it seems right to attach a higher standard deviation when outliers are found. These outlier could also be the results of a real deformation, so less confidence should be laid in areas where outliers are found.

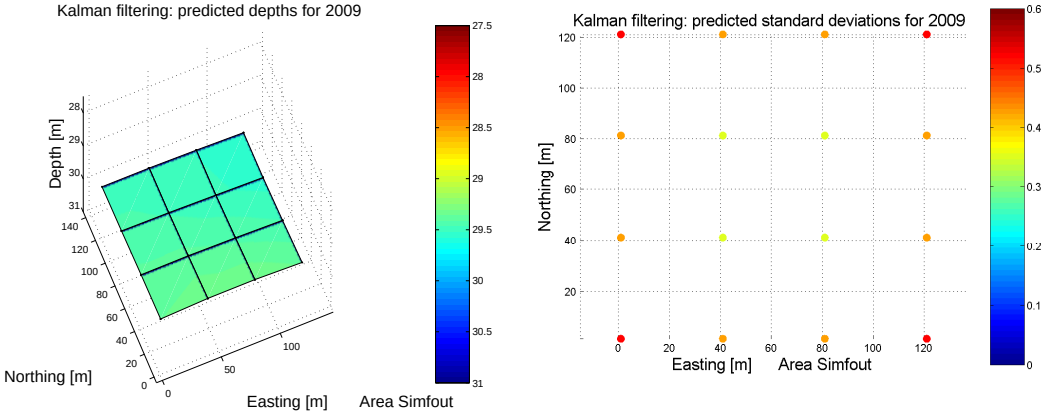



**Figure 5.12:** Predictions of the area with an outlying plane in survey 3, obtained with Kalman filtering. On the left the predicted depths for 2009. On the right the predicted standard deviations for 2009.

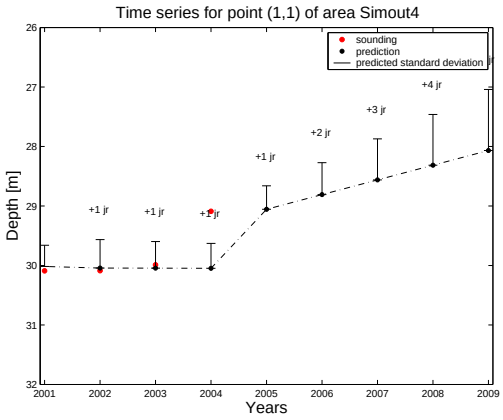



**Figure 5.13:** The time series for the first point, at location (1,1), of the simulated flat area with an outlying plane in survey 4. Because the $4^{th}$ survey differs a lot from the others, the predicted trend line bends up.

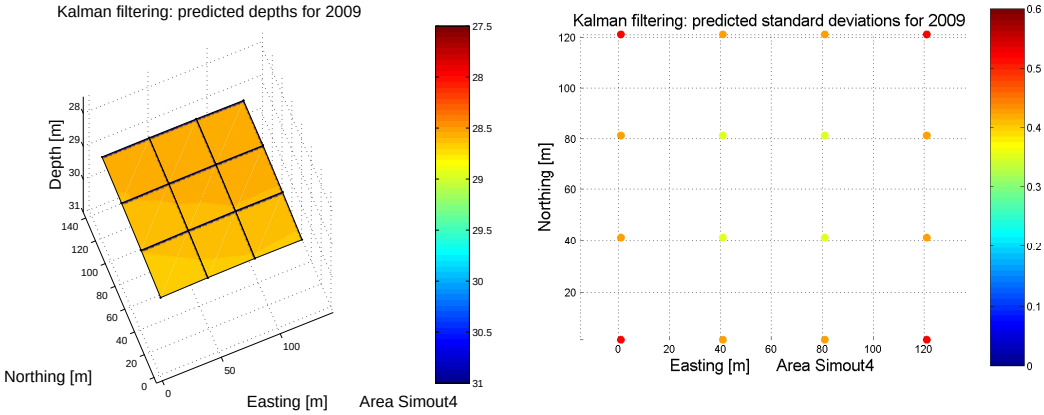



**Figure 5.14:** Predictions of the area with an outlying plane in survey 4, obtained with Kalman filtering. On the left the predicted depths for 2009. On the right the predicted standard deviations for 2009.

- With the Kalman filtering an outlier causes an artificial trend to be estimated. Therefore predicted values will to be too deep or too shoal. An outlier in the last survey has more influence than an outlier in another survey, but it recovers when more epochs come available. The standard deviations of the points remain the same, whatever type of area is considered. This seems however not right, because more confidence should be laid in a stable area than in an area where dynamic behaviour is found.

## 5.3 Simulation of a flat area with one outlying point

Now consider the case there is no systematic error, but some outlying points. In this section the situation that a single value is different is tested. Therefore in survey 3 of the data set of page 47 the value of point (41,41) is replaced by 29.

### 5.3.1 Test with deformation analysis

The depth values for this model are almost the same as for the flat model of Section 5.1, except for the point (41,41). In this point the depth is equal to 29.7578. Because an outlier is present in only one point, it is only found in the point test. For the area test no alternative hypotheses are accepted, see Figure 5.15. This also holds for the prediction, see Figure 5.16. Only the outlying point is affected and gets a shoaler depth and a higher standard deviation. All other points get the almost the same values as in the static data set of Section 5.1, the mean depth is however slightly less than 30 m.

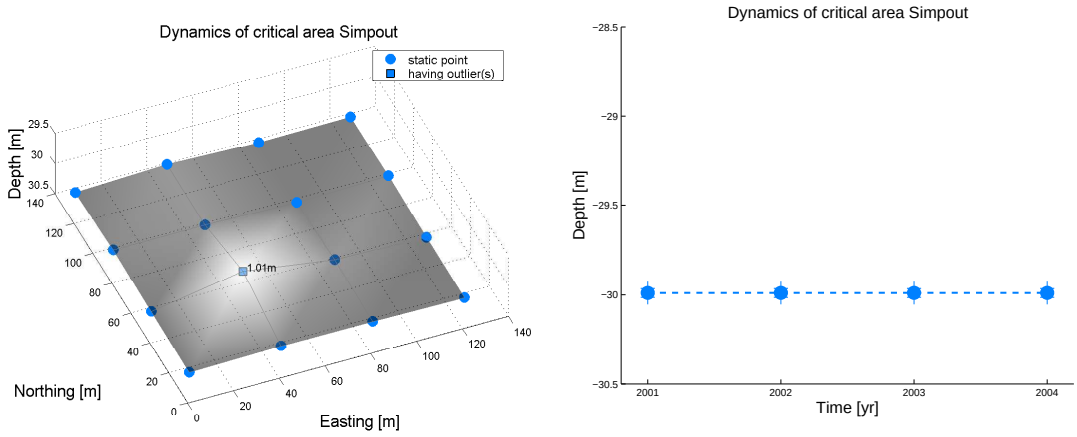



**Figure 5.15:** Results of the deformation analysis for the simulated data of a flat area with a single outlying point. On the left the results for the point test, on the right the results for the depth in the area test. The outlying point is detected in the point test, but does not really affect the area test.

### 5.3.2 Test with Kalman filtering

The time series of two points, using the Kalman filtering, are presented in Figure 5.17. The time series for the ordinary point are constant. The time series for the outlying point shows an upwards trend line, just like in case of an outlying survey but to a lesser extent. This is because of the interpolation to the model grid, which acts as a smoothing filter. So the effect of an outlier is spread out to the surrounding points.

The predicted values and standard deviations are illustrated in Figure 5.18. Here the effect of the interpolation becomes more obvious. Not only the point with the outlying value shows an upward movement, but also, to a lesser extent, the surrounding points. So the interpolation performs as a smoothing filter. The effect of the incorrectly estimated trend does also effect other points, whereas

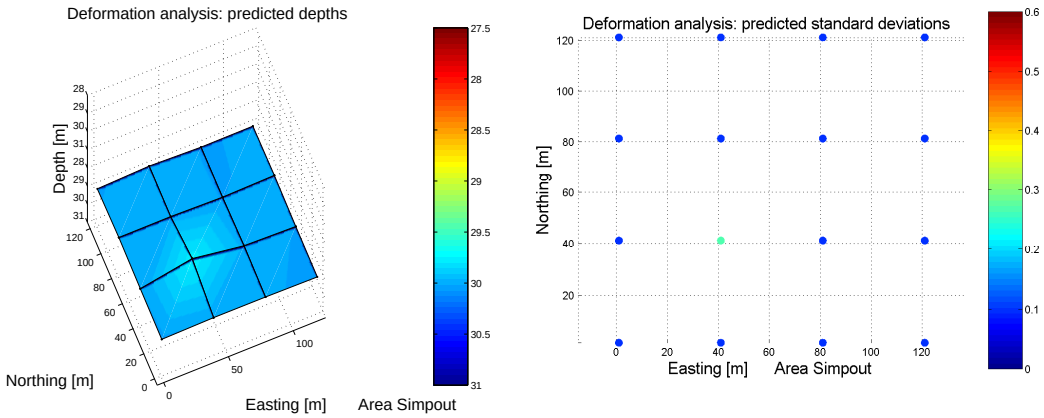



**Figure 5.16:** Predicted values and standard deviations for the area with an outlying point, obtained with deformation analysis. Only the outlying point is affected and gets another depth and standard deviation. All other points are the same as in the simulation of a static area, see Figure 5.3.

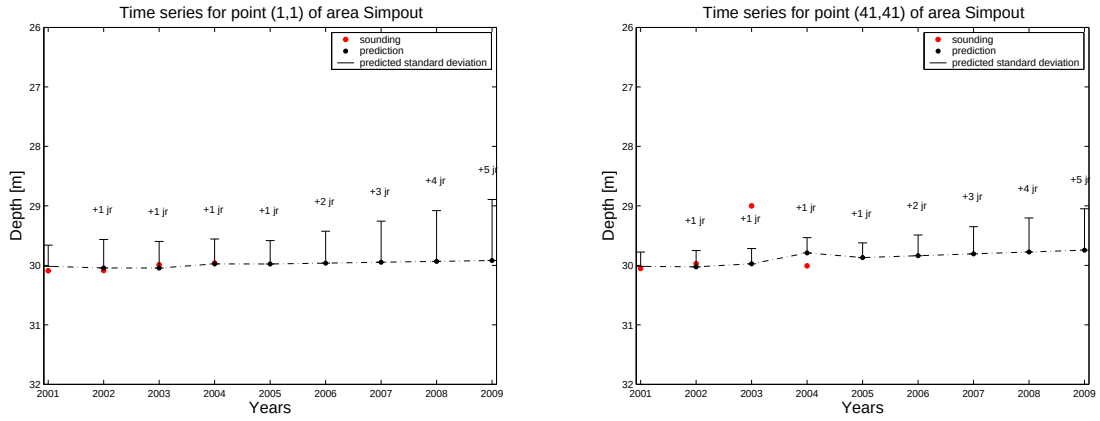



**Figure 5.17:** Time series for the flat area with a single outlying point, using Kalman filtering. On the left the time series for an ordinary point, on the right the time series for the point with an outlying value.

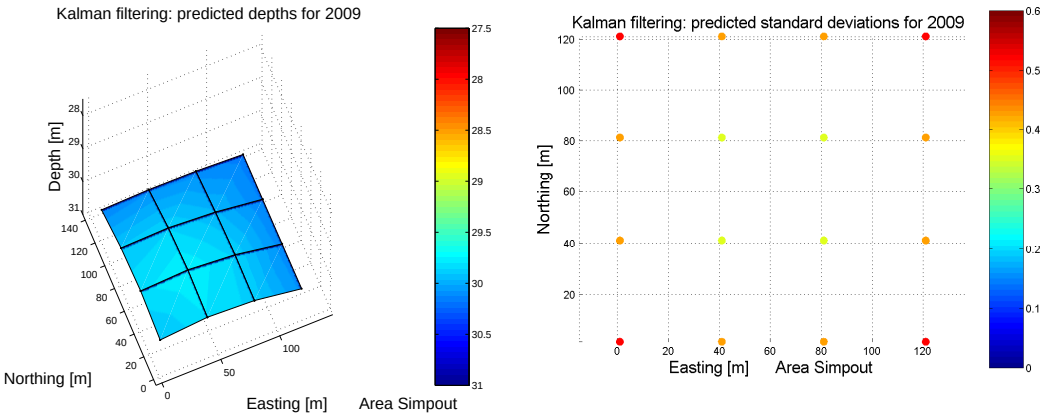



**Figure 5.18:** Predictions of the area with a single outlying point, obtained with Kalman filtering. On the left the predicted depths for 2009. On the right the predicted standard deviations for 2009. The presence of an outlying value also effects the surrounding points.

in the deformation analysis only the point with an outlying value shows a trend, see Figure 5.16. The standard deviations are still the same, whatever type of area is considered.

### 5.3.3 Conclusions

- In the deformation analysis the outlier is detected very well. In this point the predicted value differs from the others and this point gets a larger standard deviation. All the other points remain static. The area test is slightly affected, because the mean depth is influenced by the outlier. This is hardly visible in case of a single outlier, but could be significant in case of many outliers because in that case the influence on the mean depth is larger.

- In the Kalman analysis procedure the effect of an outlier is spread out to the neighbouring points as well. Therefore the a single outlier results in an upwards trend line in many points, although the magnitude is reduced.

## 5.4 Simulation of a sloping area

In this section the area is not flat anymore, but contains a slope in North-South direction of about 12.5 m/km. The shallowest points are at 29 m, the deepest points at 31.5 m, see Figure 5.19.

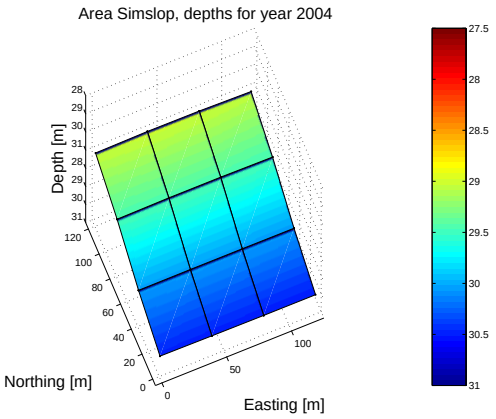



**Figure 5.19:** Area with a slope in North-South direction. The average slope is 12.5 m/km (1.25 % ).

### 5.4.1 Test with deformation analysis

The results for the deformation analysis of this area are illustrated in Figure 5.20. All points are correctly classified as static points and the area test also results in a static slope. Because all points are static, the prediction is the same as the mean of the data. So it does not matter whether the area is flat or contains a slope. This also holds when outliers are present in the data. The effect of outliers, as described in Sections 5.2 and 5.3 is the same for an area with a slope.

The adjusted values for this area are:

$$\hat{d} = \begin{bmatrix} 29.0022 & 28.9580 & 28.9952 & 29.0449 \\ 29.5290 & 29.5206 & 29.4976 & 29.5420 \\ 29.9702 & 29.9891 & 29.9954 & 29.9936 \\ 30.4730 & 30.5201 & 30.5124 & 30.5301 \end{bmatrix} \quad ; \quad \sigma^2_{\hat{d}} = 0.01. \tag{5.9}$$

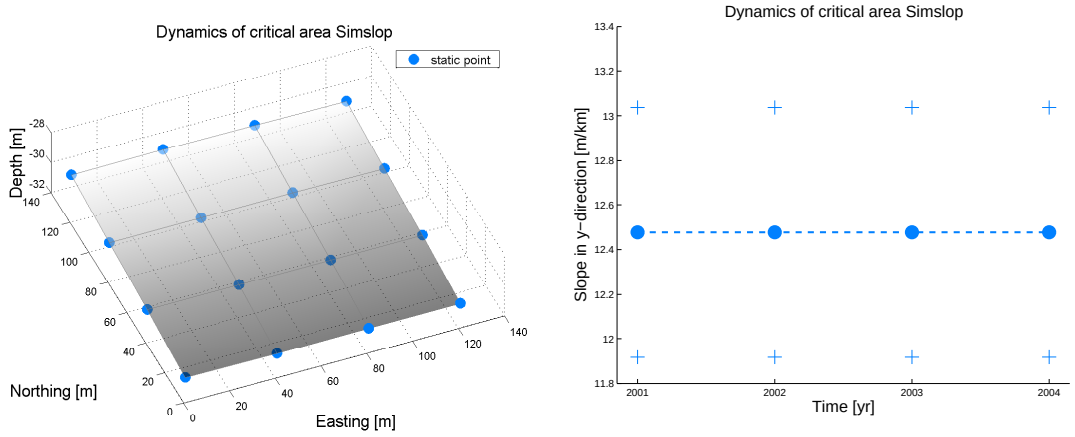



**Figure 5.20:** Results of the deformation analysis for the simulated data of an area with a slope. On the left the results for the point test, on the right the results for the slope in the area test.

### 5.4.2 Test with Kalman filtering

In Figure 5.21 the time series for point (20,20) is illustrated. The prediction behaves quite well, although the initialisation differs from the first measurement. This is because the initial value is the mean depth of the area, about 29.75 m, while the actual depth in this point is 31 m. But because more emphasis is laid on the measurements than on the initialisation, this is corrected by the first measurements. Hereafter, the trend line remains static, so also for Kalman filtering it does not make any difference whether the area is flat or not.

However, when the measurements are less accurate the initialisation becomes relatively more important, which results in trends being found. For example when the standard deviation of the measurements is doubled, to 40 cm, the points in the North get an upwards trend and the points in the South a downwards trend due to the big difference between the initialisation and the first measurements. On the right graph of Figure 5.21 the time series of the same point is given, but with a standard deviation that is twice as high. Therefore in case the variance of the measurements is good compared to the initial variance, the Kalman filtering method also produces good results for the sloping area. When it is not certain whether the initial value is a good approximation, this could be ensured by taking a very high value for the initial variance, so that the influence of the initialisation is very small.

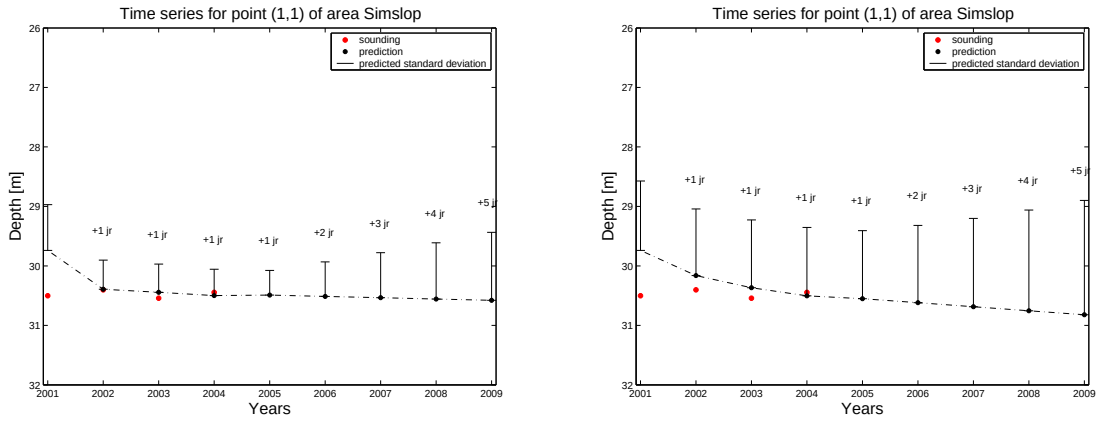



**Figure 5.21:** The time series for the area with a slope, using Kalman filtering. On the left the time series of a point with a standard deviation of 20 cm, on the right the standard deviation is 40 cm. So when the measurement accuracy is low, the initialisation has more influence and causes the prediction to be less accurate.

### 5.4.3 Conclusions

- In the deformation analysis it does not make any difference whether the area is flat or contains a slope. The same results as for a flat area hold, also in case of outliers.
- The results for the Kalman filter in case of a slope are in principle the same as the results for a flat area. However, when the accuracy of the measurements is low the initialisation causes some erroneous effects. This can however be solved by decreasing the variance of the initialisation, so the accuracy of the measurements must be high compared to the initial variance.

## 5.5 Simulation of an area with a trend

In this section the simulation of an area with a trend in the depth is described. Starting point is a flat plane at about 30 m depth, then an upwards trend with various velocities is modelled.

### 5.5.1 Test with deformation analysis

In the first simulation a large upwards trend of 0.25 m/year is modelled. In Figure 5.22 the results of the deformation analysis are shown. The trend is found in all points as well as in the area test, see Table 5.6, but the magnitude of the trend varies from point to point. The point with the largest upwards trend is indicated in the left figure. In this point the trend is estimated as 0.30 m/year, whereas a trend of 0.25 m/year was modelled. This is a result of the stochastic character of the data. Therefore in the individual points the differences between the surveys are not exactly 0.25 m/year, but vary somewhere between 15 and 35 cm. When considering the entire area, this effect disappears. The trend estimation is more accurate for the area test, the estimated standard deviation of the trend is 0.022 m. The magnitude of the trends for the individual points are:

$$\begin{bmatrix} 0.2691 & 0.2256 & 0.2502 & 0.2463 \\ 0.2864 & 0.3003 & 0.2537 & 0.2340 \\ 0.2601 & 0.2436 & 0.2471 & 0.2489 \\ 0.2869 & 0.2409 & 0.2191 & 0.2124 \end{bmatrix}. \tag{5.10}$$

These differences in trends also have implications on the prediction, as the prediction is based on the point test. In Figure 5.23 the results of the prediction are given. Because a different trend value is estimated in the individual points, the predicted depths do not lie in a plane but show some depth differences. The

| | $\underline{T_q}/k_\alpha$ for point test | | $\underline{T_q}/k_\alpha$ for area test | |
|---|---|---|---|---|
| **Type of deformation** | 25 cm/year | 10 cm/year | 25 cm/year | 10 cm/year |
| Outlying value in survey: 1 | 1.0954 | 0.2231 | 7.7184 | 1.2535 |
| 2 | 0.0046 | 0.0162 | 0.6482 | 0.0748 |
| 3 | 0.0669 | 0 | 0.9755 | 0.0969 |
| 4 | 0.7322 | 0.1052 | 6.7832 | 1.2077 |
| General deformation | 1.2091 | 0.2196 | 8.1055 | 1.3235 |
| Trend | 3.3394 | 0.4784 | 21.7671 | 3.4990 |

**Table 5.6:** The normalised test quantities for the point test for point (1,1) and the area test of the data sets with a trend of 0.25 m/year and 0.10 m/year.

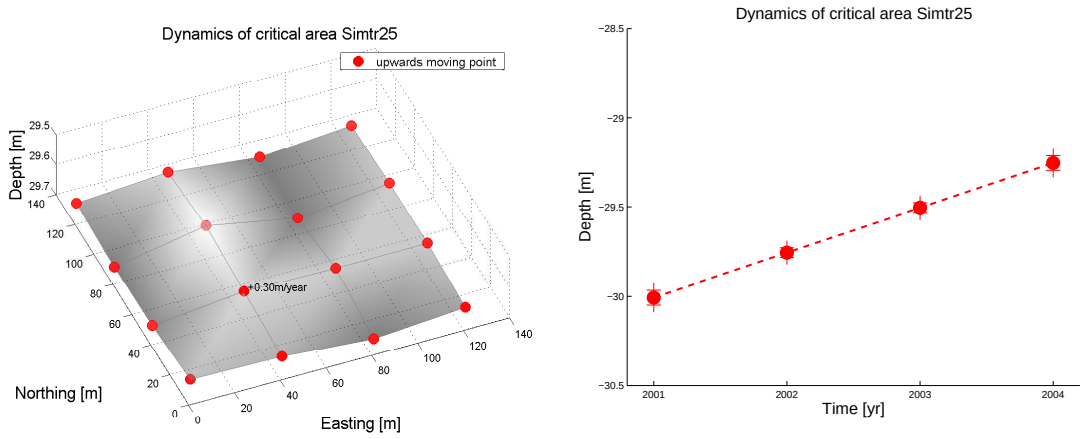



**Figure 5.22:** Results of the deformation analysis for the simulated data of an area with a large upwards trend (0.25 m/year). On the left the results for the point test, on the right the results for the area test.

predicted depth values after 5 years of prediction are somewhere between 27.6 and 28.2 m. The prediction of the standard deviation is independent of the magnitude of the trend.

Now consider the case a smaller trend is present, about 0.10 m/year. For most points in this data set the trend is not found anymore. However, the area test still performs very well, see Figure 5.24. This is even more obvious when looking at the prediction, see Figure 5.25. Only the depth of the point where a trend was found is updated, in the other points the predicted depth is the mean of the four surveys, which is about 29.85 m. However the fact that a trend is present should result in a predicted depth of 29.2 m in 2009. So by ignoring the trend in almost all points a prediction error of 65 cm is made. According to Equation 3.22, with the values $\alpha = 0.10$ and $\gamma = 0.8$, the Minimal Detectable Bias of this situation is calculated as $\nabla = 0.22$ m/year. This means that a trend of 0.22 m/year can be found with a probability of 80 %. In this case still a trend is found in one point, but trends of this magnitude are in fact not very likely to be found. However, when more epochs are available, a small trend will be more likely to be detected. When e.g. six surveys are available, the MDB drops to 0.12 m/year.

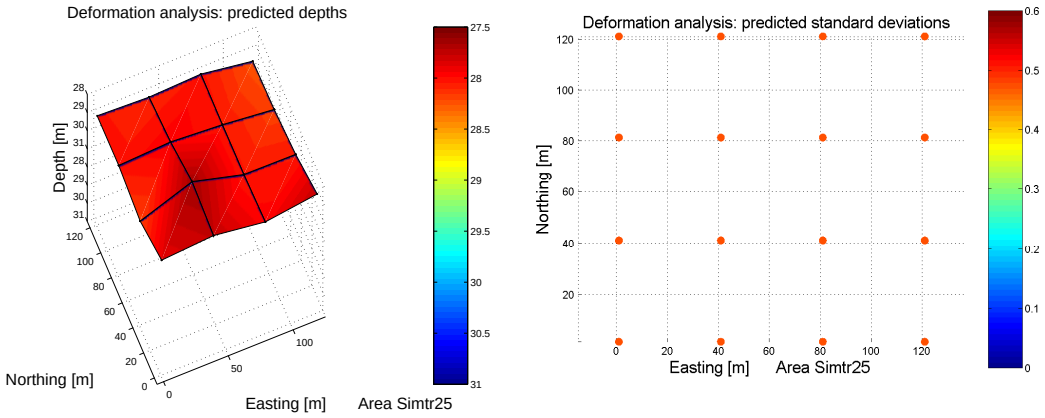



**Figure 5.23:** Predicted values and standard deviations for 2005, for the area with an upwards trend of 25 cm/year, obtained with deformation analysis.

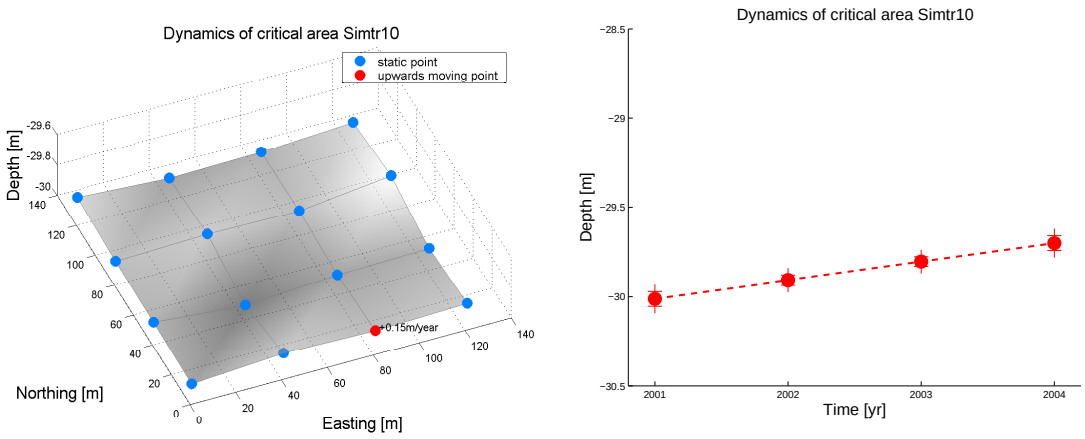



**Figure 5.24:** Results of the deformation analysis for the simulated data of an area with a smaller upwards trend (0.10 m/year). On the left the results for the point test, on the right the results for the area test.

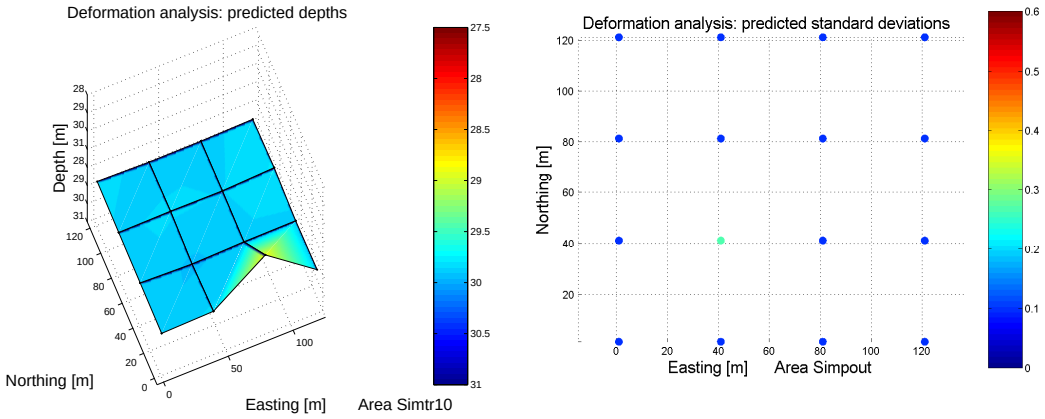



**Figure 5.25:** Predicted values and standard deviations for the area with an upwards trend of 0.10 m/yr, obtained with deformation analysis.

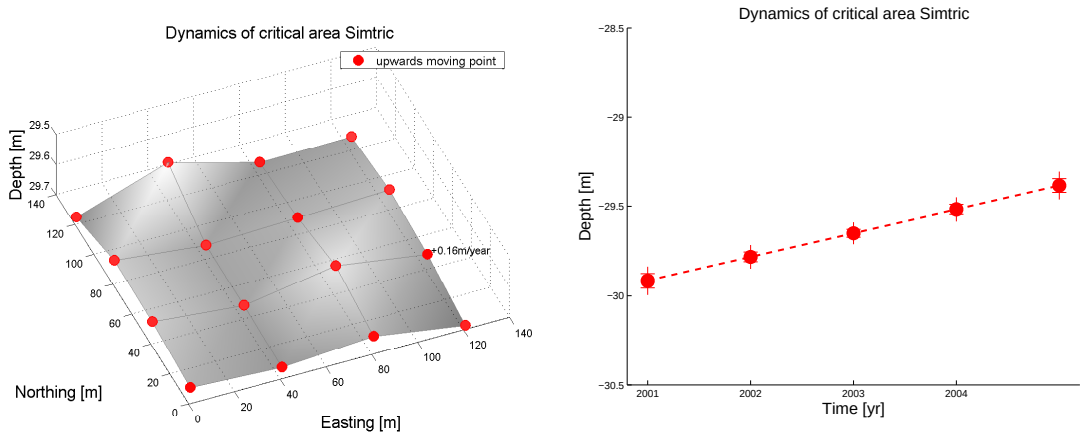



**Figure 5.26:** Results of the deformation analysis for the simulated data of an area with an varying trend. On the left the results for the point test, on the right the results for the area test.

Finally the situation is considered that initially a trend is present, but after a while it stabilises. For example the mean depths are 30, 29.75, 29.5, 29.5 and 29.5 m respectively (for this case an extra year is added to make the situation more obvious). The deformation analysis method only handles one, constant trend so an average trend is fitted through the data, see Figure 5.26. All points have a constant trend of about 12.5 cm/year. This trend remains so the prediction is also based on this trend, resulting in an error of about 0.5 m after 5 years.

### 5.5.2 Test with Kalman filtering

The simulated data with a trend are also processed with the Kalman filtering. First the large trend of 0.25 m/yr is considered. From Figure 5.27 can be concluded that this trend is estimated very well. After the initialisation the trend is first set to zero, because an estimated trend in the first epoch is just a result of the initialisation, rather than a real trend. Because of the smoothing effect of the interpolation support points, the predicted values are more accurate than those of the deformation analysis, see Figure 5.23.

The results for the smaller trend of 10 cm/year are also very good. The magnitude of the trend does not make any difference, the performance of the method is the same, see Figures 5.29 and 5.30.

Finally, the data set with the changing trend is processed. The results are visualised in Figures 5.31

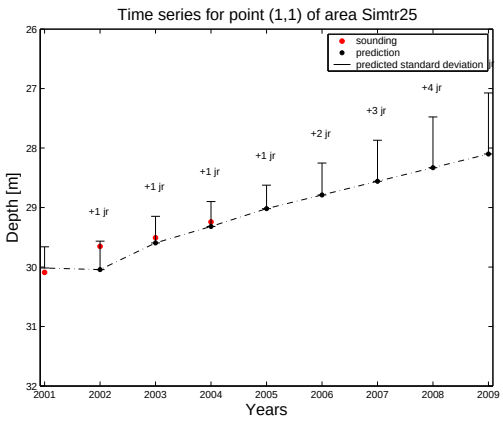



**Figure 5.27:** Time series for an area with an upwards trend of 25 cm/year.

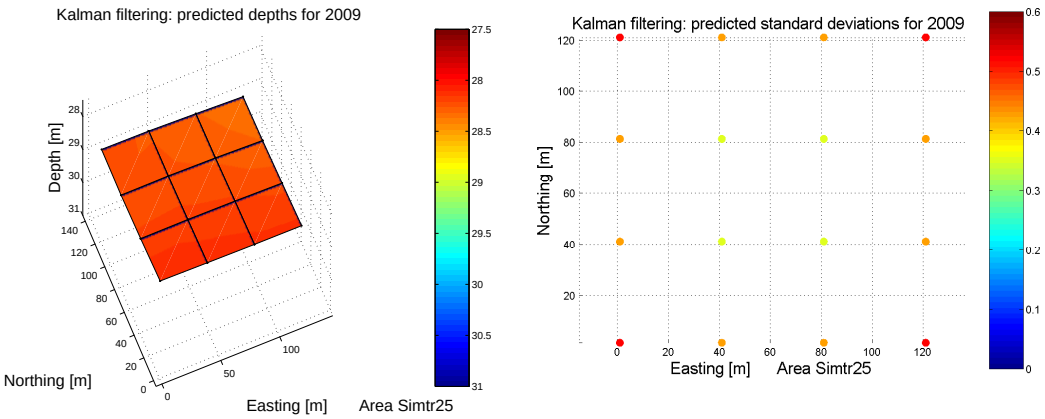



**Figure 5.28:** Predictions of the area with an upwards trend of 25 cm/year, obtained with Kalman filtering. On the left the predicted depths for 2009. On the right the predicted standard deviations for 2009.

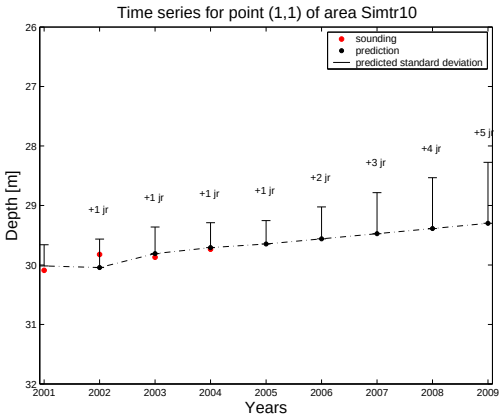



**Figure 5.29:** Time series for an area with an upwards trend of 10 cm/year.

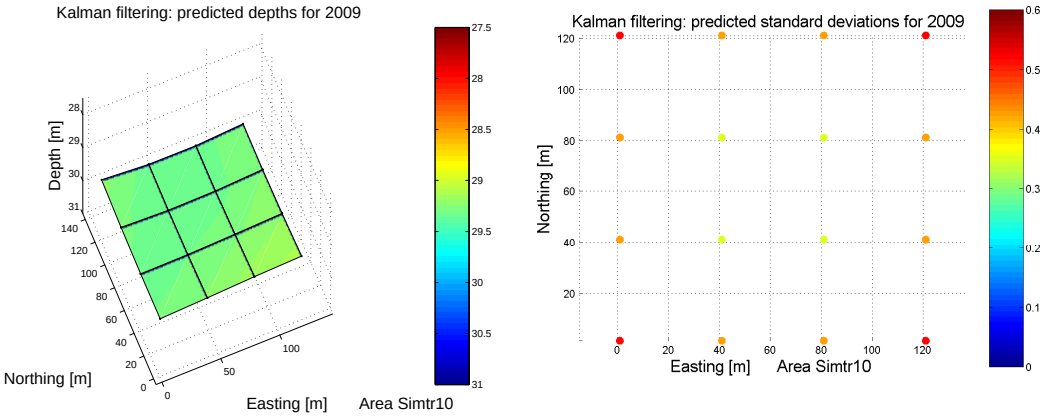



**Figure 5.30:** Predictions of the area with an upwards trend of 10 cm/year, obtained with Kalman filtering. On the left the predicted depths for 2009. On the right the predicted standard deviations for 2009.

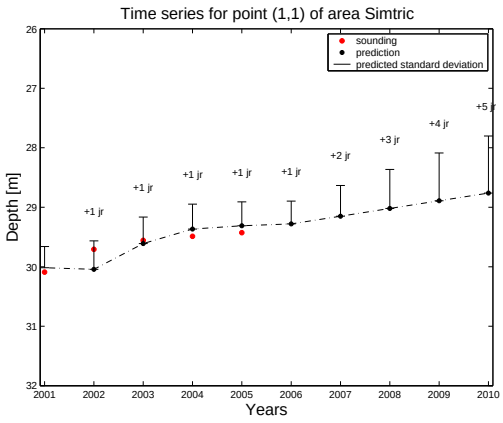



**Figure 5.31:** Time series for an area with a changing trend. Initially a rather large trend is present, but later on it should stabilise. The upwards trend however remains present.

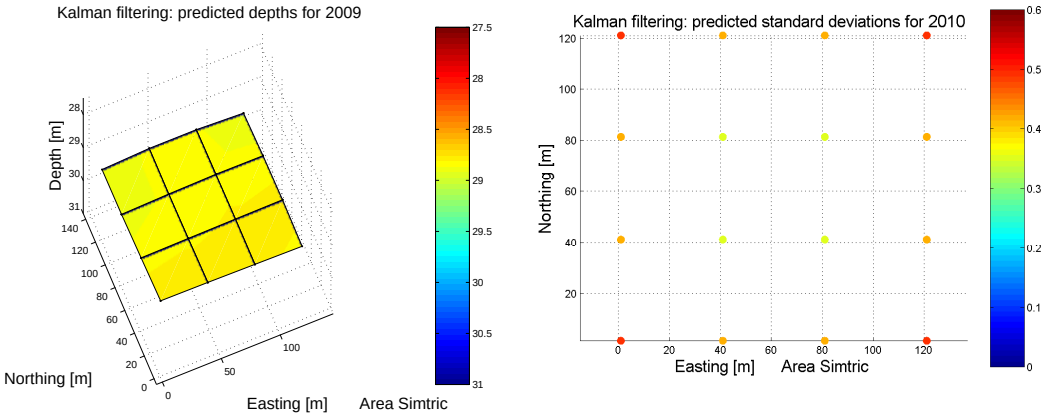



**Figure 5.32:** Predictions of the area with the inconstant trend, obtained with Kalman filtering. On the left the predicted depths for 2009. On the right the predicted standard deviations for 2009.

and 5.32. The predicted trend line is adjusted to the data set, but an upwards trend remains. Because the estimated trend only slowly disappears, the trend is still present when the last data set is processed. After the last survey no measurements become available to contribute to the trend estimation, so the trend value remains constant in the further predictions. When more epochs were available, the value for the trend would eventually become zero. So although the Kalman filtering gives better results for this simulation, compare with Figure 5.26, some more epochs are needed to acquire the right solution.

### 5.5.3 Conclusions

Summarising this section:

- In the deformation analysis procedure, the trend can be detected very well with the area test, but for the individual points only trends larger than the Minimal Detectable Bias can be detected with a certain probability.
- Because of the stochastic character of the data the magnitude of the trends found varies from point to point. In the deformation analysis this might give significant errors in the prediction, in the Kalman filtering these errors are smaller because of the smoothing effect of the interpolation.
- In the Kalman filtering the trend line adjusts itself after a while when the speed has changed, in the deformation analysis only a single constant trend is estimated.

## 5.6 Simulation of an area with a static sand wave

In the previous sections only areas without sand waves were simulated. Most of the Southern part of the North Sea is however covered with sand waves [Németh, 2003]. Therefore it is interesting to see whether the present waves can be detected with these methods, and what the effect of sand waves on the prediction will be. First a static sand wave is simulated, i.e. a sand wave that is not moving. To represent the wave better the simulated area is enlarged to 8×8 grid points, the grid distance is 20 m resulting in a 140×140 m area. Hereafter, the depth is defined as:

$$d(x, y) = 30 + \sin\left(\frac{2\pi x}{300}\right), \tag{5.11}$$

where $x, y$ denote the horizontal position in the area. In this way, a part of a sand wave with amplitude 1, initial phase 0 and a wavelength of 300 m is generated, see Figure 5.33. On top of this wave some random noise is added. However, real sand waves are not exactly harmonic, so in reality the sine model will never fit the sand wave accurately, but it is a suitable approximation.

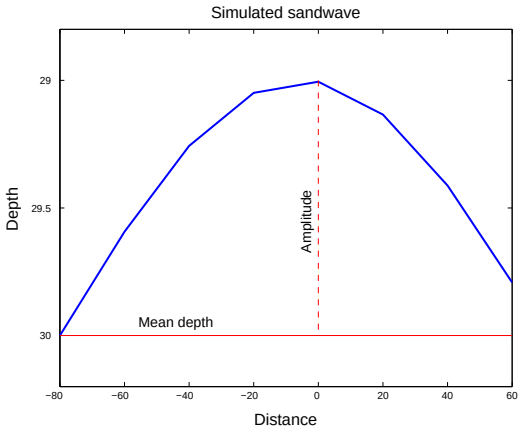



**Figure 5.33:** Profile of the simulated sand wave with amplitude 1 and wavelength 300.

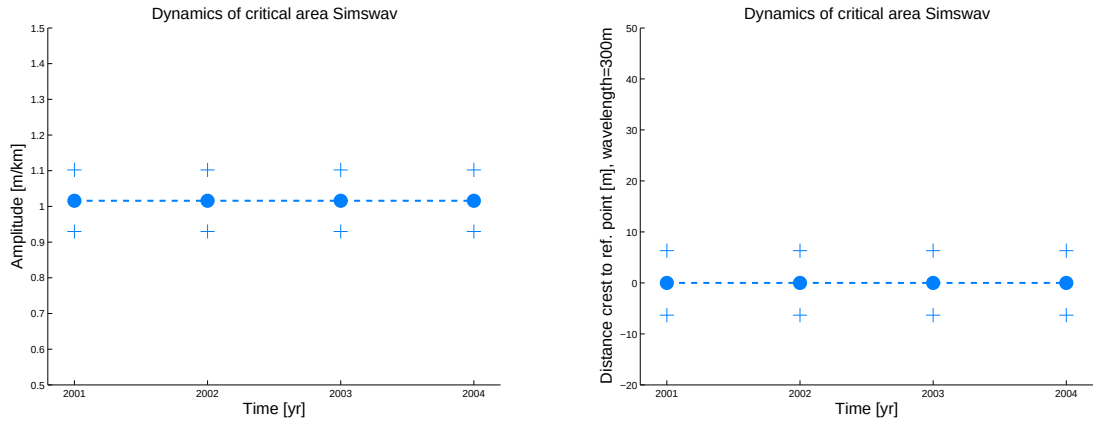



**Figure 5.34:** Results of the sand wave test for a static sand wave. On the left the amplitude, on the right the distance from the sand wave crest to a reference point.

### 5.6.1 Test with deformation analysis

First, this area is tested with the deformation analysis procedure. Using the area test the amplitude and phase of the wave are now determined as well. In Figure 5.34 the results of this test are shown. Due to the noise, the estimated value for the amplitude is not exactly 1, but slightly higher. The estimate for the phase is expressed in terms of the crest distance, which is the distance from the sand wave crest to a reference point. The reference point is chosen as the location of the sand wave crest in the first survey, so this parameter indicates the distance the sand wave has moved over the years. The crest distance is calculated with Equation 3.44.

This wave can be detected very well. Essential in this detection is however the knowledge of the wavelength. If the wavelength is not accurately determined, this influences the determination of the other parameters as well. This mainly results in a wrong value for the amplitude, but also the phase, the mean depth and the slope are influenced.

Because in this data set a static sand wave is simulated, the prediction is the same as the mean value of the four epochs and the standard deviation is one fourth of the original standard deviation, just like in case of a flat area (Section 5.1).

### 5.6.2 Analysis with Kalman filtering

Using Kalman filtering, no wave detection is performed, so the prediction is just based on the time series of the individual points. Because these data sets are assumed to be static, the predicted values are the the mean value of the four epochs. In Figure 5.35 the time series is given. The results of the prediction are influenced by the initial values, just like in case of the area with the slope (Section 5.4). This can be solved by enlarging the initial standard deviation.

### 5.6.3 Conclusions

- A static sand wave can be detected by the deformation analysis procedure. It is however essential that the wavelength and the direction of the sand wave are known beforehand.

- The results of the prediction are the same for all static areas, whether they are flat, having a slope or are covered with sand waves. This holds for the deformation analysis as well as for the Kalman filtering.

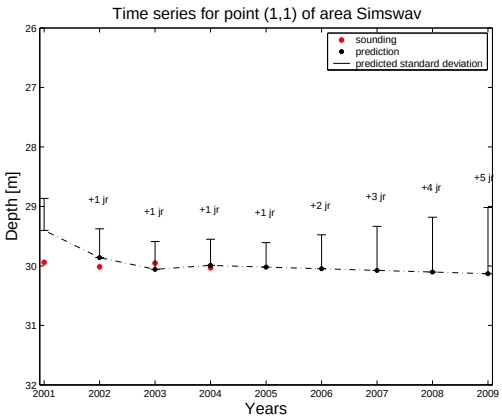



**Figure 5.35:** Time series for point (1,1) of the data sets with a static sand wave.

## 5.7 Simulation of an area with a migrating sand wave

In the previous section, it was shown that a static sand wave can be detected by the deformation analysis procedure. In this section, it is tested whether a migration of those sand waves can be detected. Therefore the same wave is simulated, but every year the crest is shifted 10 m, see Figure 5.36.

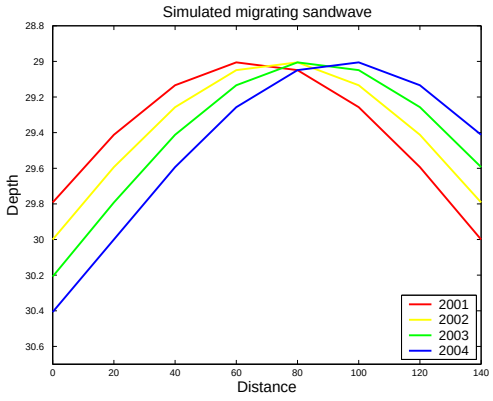



**Figure 5.36:** Profile of the migrating sand wave in the 4 epochs.

### 5.7.1 Test with deformation analysis

The results for the area test for this simulation are shown in Figure 5.37. The estimation of the crest distance clearly shows that the sand wave is moving with a speed of 10 m/year. The amplitude estimation is not entirely correct, but the error is still smaller than the standard deviation. However, when the migration rate of the waves is higher, the estimation deteriorates. Consider e.g. the case that the same wave is moving with 20 m/year. In that case the estimation of the crest distance (phase) is still quite good, but the amplitude estimation shows large errors, see Figure 5.38.

These were the results for the area test, but a point test is performed as well. In Figure 5.39, the results of the point test and the predicted values for 2009 are given. The results of the point test indicate the migration of the sand wave, because in general points on the left are going down and points on the right are going up. For the central points the differences are not big enough to be significant so they are assumed to be static. Furthermore, due to the stochastic character of the data, the pattern is not entirely

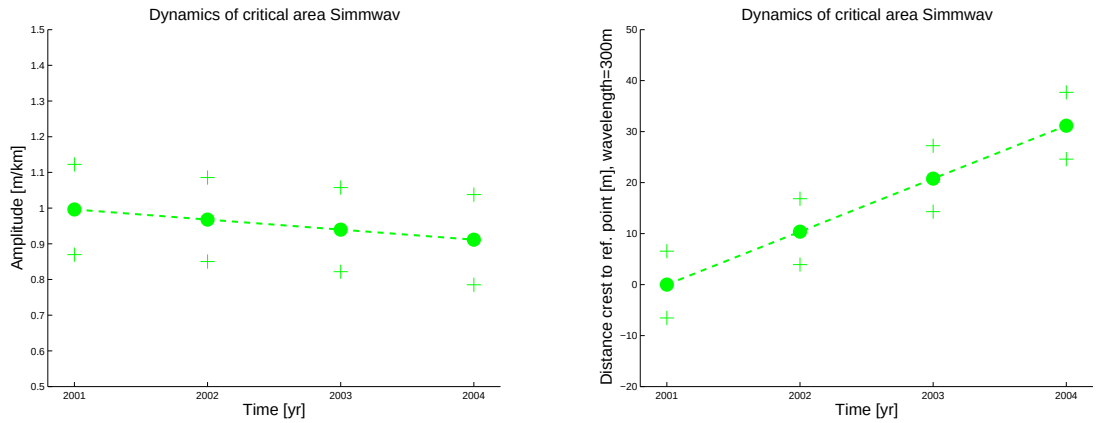



**Figure 5.37:** The result of the area test with a sand wave with a migration rate of 10 m/year. On the left, the amplitude of the wave, on the right the distance from the wave crest to a reference point.

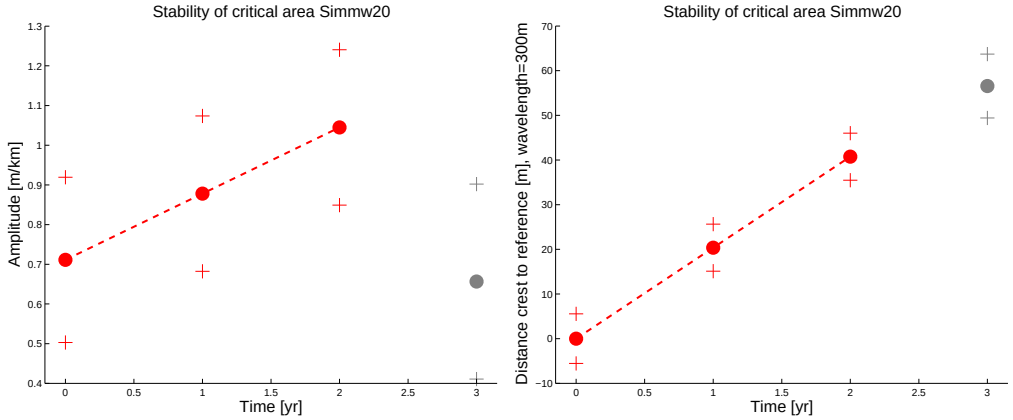



**Figure 5.38:** The result of the area test for the sand wave with a migrating rate of 20 m/year. On the left the amplitude of the wave, on the right the distance from the wave crest to a reference point.

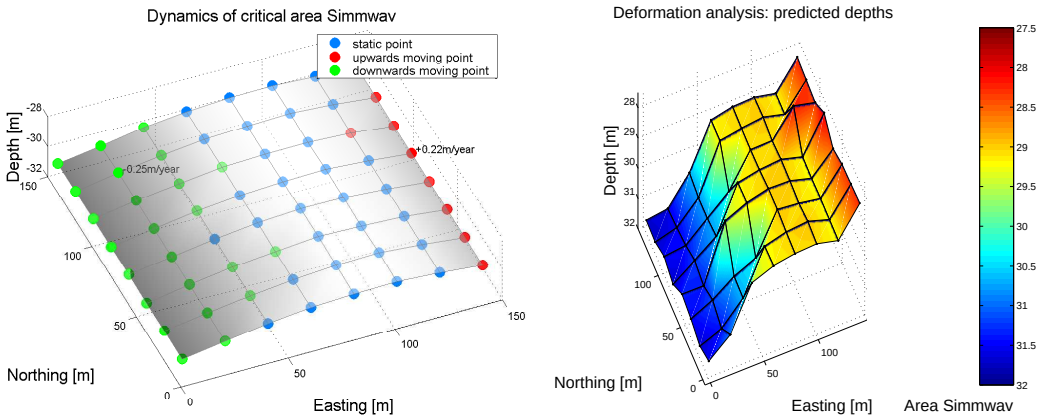



**Figure 5.39:** On the left the result of the point test for the migrating sand wave. On the right the predicted values 5 years ahead.

regular, but some other exceptional points occur as well.

The results of this point test are also used for the prediction, so the trends that are found are propagated into the future. Therefore, the results of the prediction are totally wrong, because the waveform is not taken into account. It is assumed that when a point has an upwards trend it remains rising, but in reality when the sand wave crest has passed the point will go down again. Therefore the predicted shape of the area is wrong.

### 5.7.2 Test with Kalman filtering

Because the Kalman procedure only aims at predicting the depth, by means of a constant velocity parameter this gives the same problems as the deformation analysis method. The points that are initially rising, before the sand wave crest has passed, keep rising and the points that are initially shoaling remain shoaling. In Figure 5.40 the time series for two points are given. Point (1,1) is situated on the left of the area, so this point is initially shoaling, until the next sand wave will come. The predicted trend line therefore goes down, and remains going down. Point (81,1) is situated near the sand wave crest, so the crest is passing within this observation time. Therefore the depth differences between the epochs are very small and the trend line is almost horizontal. These trends result in a prediction as in Figure 5.41.

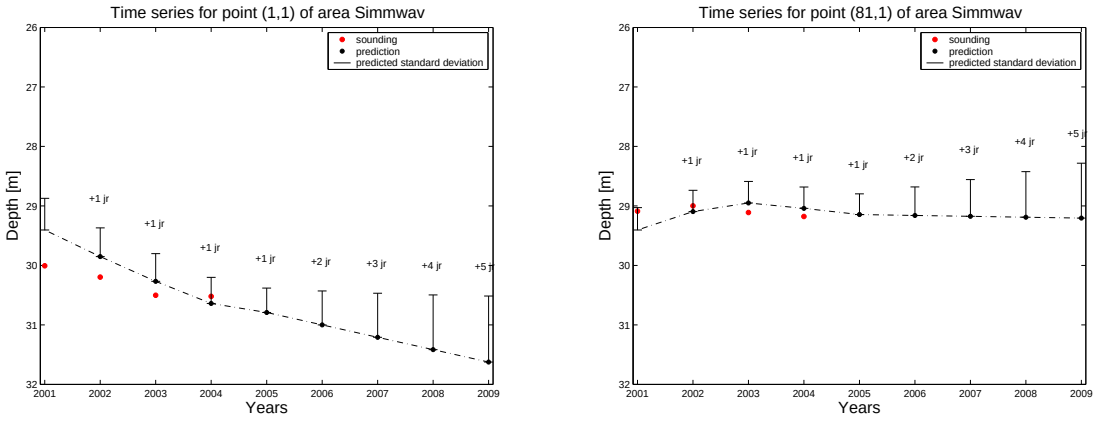



**Figure 5.40:** Time series for two points of the data set with the migrating sine wave. On the left the point (1,1), which is initially shoaling, on the right point (81,1) on the sand wave crest.

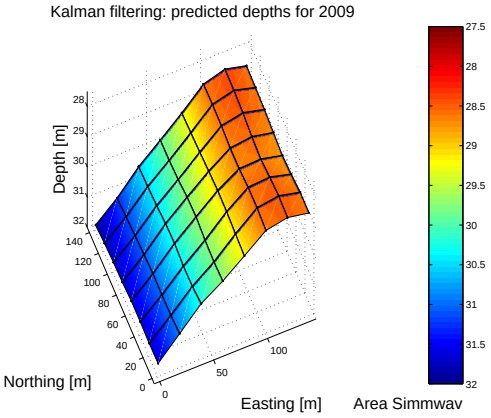



**Figure 5.41:** Predicted depths 5 years ahead, for a migrating sand wave.

### 5.7.3 Conclusions

- A migrating sand wave can be detected by the deformation analysis procedure, provided that the migration is not too large compared to the wavelength and the grid size.
- Time series for a moving sand wave cannot be used for prediction. In this case a point that is initially rising, might go down again after the sand wave crest has passed. Both methods can however only handle a constant trend, otherwise the predicted values are incorrect.

## 5.8 Conclusions of the simulations

In this section the conclusions of the previous sections are summarised. In the following table, the behaviour of both methods in different situations are compared to each other.

| **Type of area** | **Performance with deformation analysis** | **Performance with Kalman filtering** |
|---|---|---|
| Static areas | Unimportant whether the area is flat, has a slope or a sand wave. The prediction is correct, noise is reduced. | The initialisation might cause the trend line to bend up- or downwards. Otherwise the prediction is correct, noise is reduced, because of the smoothing effect of the interpolation the variances of the edge points are worse. |
| Outlying surveys | Outliers in the central surveys can be handled very well, if an outlier is present in the first or last survey sometimes a trend is preferred. The prediction is influenced by the presence of an outlying value. | Any outlier influences the prediction, the extent depends on the magnitude of the error and the number of epochs. Furthermore, an error in the last epoch has more influence than errors in previous epochs. |
| Single outlying values | The conclusion about systematic errors also holds for outlying values, except that the effects are only locally. | Because of the smoothing an outlying point also influences the neighbouring points, the effect of the outlier is spread out. |
| Trends | Only trends larger than a certain value are accepted, smaller trends are ignored. Therefore in the prediction significant errors are made by ignoring smaller trends. | All differences between the epochs are taken into account, so also very small trends are still visible in the prediction. |
| Sand waves | A detection of the sand wave and estimation of the amplitude and phase is carried out. This works very well in case of a static wave, reasonably well in case of a migrating wave. Essentially hereby is the correct determination of the wavelength and direction. The static wave does not influence the prediction, a migrating wave causes the prediction to go wrong, because the prediction is only performed locally and only uses constant trends. | No detection of sand waves is performed, just a local prediction. For the static wave the prediction might be influenced by the initialisation, a migrating wave gives erroneous results. |

# Chapter 6

# Combination of both methods

In Chapters 3 and 4, the methods of the Hydrographic Service and the North Sea Directorate to analyse time series of bathymetric data were described. Furthermore, these methods were tested on simulated data in Chapter 5. The two main conclusions about the performance of both methods are:

- The deformation analysis procedure is very suitable for detecting dynamics in the past, while the Kalman filter works better in propagating the present state into the future.
- In both methods sand wave migration has major impact on the prediction results.

These methods are however not in contrast with each other, but rather supplementary. For example an outlier detection by the deformation analysis can be used to obtain better results in Kalman filtering. The same with sand wave migration: the trend in a sand wave (change in phase and amplitude) can be detected quite well with deformation analysis. If this trend would be included in the Kalman filter, the results can be expected to be more accurate. In this chapter a new method is described, which combines the deformation analysis with the Kalman filter. The Kalman filter is extended with a testing procedure and a sand wave propagation model. Section 6.2 gives an approach to deal with outliers in the Kalman filter and in Section 6.3 the inclusion of wave propagation parameters in the Kalman filter is described. In Section 6.4 the new method is tested on some of the simulated data sets. A complete overview of this new method is given in Section 6.5 and this new method is tested on real data in Section 6.6. Finally in Section 6.7 the possibilities and limitations of this method are discussed. This chapter will however start with a review of the differences between both methods in Section 6.1.

## 6.1 Comparison of both methods

From the previous chapters, it will be clear that there are many differences between the methods of the Hydrographic Service and the method of the North Sea Directorate. In Table 6.1 a review of the main differences are summarised. These differences follow from the overviews of both methods, in Section 3.4 and 4.4 respectively. The differences treated here are mainly differences in the implementation and practical issues of the methods. Conclusions about the performance of the methods result from the tests on simulated data and are given in Section 5.8.

| | **Hydrographic Service** | **North Sea Directorate** |
|---|---|---|
| Method | Deformation analysis. | Kalman filtering. |
| Type of method | Batch method. | Recursive method. |
| Software | Scripts in Matlab and the Matlab Kriging toolbox of [Gratton and Lafleur, 2001]. | Scripts in Matlab. |
| Main goal of the analysis of time series. | Insight in dynamics of the sea floor | Prediction of oversteps of nautical limits. |
| Data | Shoalest depth in a 5×5 m grid cell is used. Up to now only single beam data is available. | Mean depth in a 5×5 m grid cell is used, all data are obtained with multibeam. |
| Coordinate system | Horizontal coordinates w.r.t. the ETRS89 frame, depths w.r.t. Mean Low Lower Water Spring. | Horizontal coordinates w.r.t. ED50 frame, depths w.r.t. Mean Sea Level. |
| Quality estimation | Theoretical standard deviations, based on various factors, are calculated for every depth value. | Standard deviations are calculated from the measurements, a fixed value for the measurement accuracy is adapted. |
| Interpolation / smoothing | Measurements are interpolated with Kriging, the data have to be on a regular grid (e.g. 60×60 m) to perform time series analysis. | Gaussian interpolation within Kalman filter, this works essentially as smoothing. A typical grid distance is 17.5×17.5 m. |
| Input parameters | Flag whether sand waves are present or not, levels of significance for the tests. | Initial values, model parameters (e.g. grid distance), calibration parameters (e.g. discount factors). |
| Necessary preprocessing steps | Calculation of standard deviations, estimation of covariance function , determination of sand wave length, interpolation. | |
| Resolution of the analysis | Point-wise analysis as well as area-based analysis. | Point-wise prediction, but because the Kalman model is applied in the interpolation kernels, spatial correlation exists. |
| Type of dynamics that can be found | Stability, outliers, instability, linear trends, sand wave dynamics (amplitude, phase, trend in sand wave). | Only linear trends. |
| Prediction | Prediction is only point-wise and based on detected dynamics, outliers are taken into account. Only points where dynamics larger than a certain boundary are detected, get a different predicted depth. | Prediction based on the estimated trend. The prediction is sensitive for outliers. |
| Output | Plots of the accepted dynamics, for individual points and for the entire area, plots of predicted depths and standard deviations. | Plots with time series, profiles, oversteps and a diagnosis plot. |

**Table 6.1:** Differences between the methods of the Hydrographic Service and the North Sea Directorate.

## 6.2 Outlier detection in the Kalman filter

As was concluded in Chapter 5, the Kalman filter is rather sensitive for outlying values and especially for systematic errors. In Paragraph 3.2.2, a testing procedure was described, based on the residuals:

$$\underline{\hat{e}} = \underline{y} - A\underline{\hat{x}}. \tag{6.1}$$

For a recursive model, a similar procedure exists, called *DIA (Detection, Identification and Adaptation)*, [Salzmann, 1992]. This procedure is based on the predicted residuals:

$$\underline{v}_k = \underline{y}_k - A_k\underline{\hat{x}}_{k|k-1}. \tag{6.2}$$

As obvious, the DIA procedure consists of three steps:

1. Detection. The detection phase is the overall model test. This gives a first indication whether the mathematical model fits to the data. Two different kinds of overall model test statistics are available: The *Local Overall Model (LOM)* test statistic and the *Global Overall Model (GOM)* test statistic. The difference between these two is that the local test takes only one epoch into account, whereas the global test takes more than one epoch into account [Teunissen and Salzmann, 1989]. So, the local test only compares the predicted state at time $k$ with the observations at this moment. Therefore this test can be performed in real-time.

   The LOM test statistic at epoch $k$ is given as [Salzmann, 1992]:

   $$\underline{T}^k = \frac{\underline{v}_k^* Q_{v_k}^{-1} \underline{v}_k}{m_k}, \tag{6.3}$$

   with $m_k$ the number of observations in epoch $k$. Under the null-hypothesis, this test statistic is distributed as $\chi^2(m_k, 0)/m_k$. Because in this case only errors in a single epoch have to be detected, only the local tests are used.

2. Identification. When the overall model test is rejected, the error has to be identified. For this purpose the *Local Slippage (LS)* test statistic and the *Global Slippage (GS)* test statistic are used.

   The LS test statistic at epoch $k$ is given as [Salzmann, 1992]:

   $$\underline{t}^k = \frac{c_{v_k}^* Q_{v_k}^{-1} \underline{v}_k}{\sqrt{c_{v_k}^* Q_{v_k}^{-1} c_{v_k}}}. \tag{6.4}$$

   The vector $c_{v_k}$ points at the observation where an error is assumed. So, the LS test statistic is calculated for every observation. When one or more test statistics are larger than a critical value, the observation with the largest test statistic is rejected and the procedure repeats until all observations are accepted. This can be compared to the data snooping procedure as described in Paragraph 3.2.2. The LS test statistic has a $N(0,1)$ distribution under the null-hypothesis.

3. Adaptation. After an error has been found, the Kalman filter has to be corrected so that this error does not influence the predictions anymore. This is called adaptation. Therefore the magnitude of the model error has to be estimated:

   $$\underline{\hat{\nabla}}^k = \frac{c_{v_k}^* Q_{v_k}^{-1} \underline{v}_k}{c_{v_k}^* Q_{v_k}^{-1} c_{v_k}}. \tag{6.5}$$

   When this error has been estimated, the state vector can be corrected:

   $$\underline{\hat{x}}_{k|k}^{new} = \underline{\hat{x}}_{k|k}^{old} - K_k c_{v_k} \underline{\hat{\nabla}}^k. \tag{6.6}$$

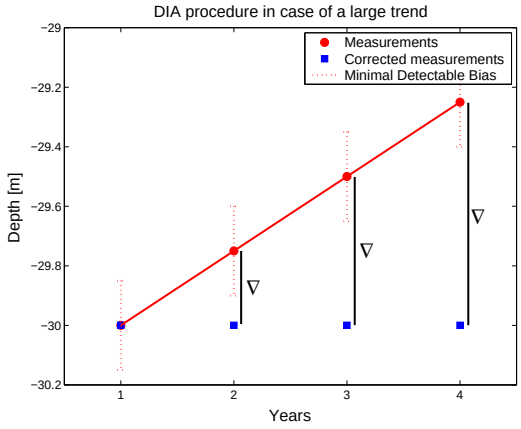



**Figure 6.1:** Illustration of the DIA procedure in case a trend, larger than the Minimal Detectable Bias is present. It is clear that the difference between the first and second survey (red dots) is larger than the MDB. The error is estimated as $\nabla$. Therefore the second survey is corrected (blue square). As a result the difference between the second and third survey is too large as well, so the third survey is corrected as well and so on. The result is that the present trend is completely ignored.

With the DIA procedure, errors in the Kalman model can be detected and corrected. This is done on basis of the predicted residual $\underline{v}$ and its variance $Q_v$, which come available automatically with every measurement update, see Equation 4.5. Therefore, the DIA procedure is easy to implement in the Kalman model. Two different model errors are considered:

1. An outlying value. In this case a deviation is assumed in one or several points, but the assumed error is different for each point. The vector $c_{v_k}$ now takes the form $(0 \dots 1 \dots 0)^*$ where the "1" points to the observation where an error is assumed.

2. A single outlying survey. In this case a systematic error is assumed, so the error will be the same in all points. Therefore the vector $c$ takes the form $(1\,1 \dots 1)^*$.

The implementation of the DIA procedure shows that it works well for the first type of errors. In Paragraph 6.4.2, it is shown that a single outlying value is detected and corrected. For the second case however this may lead to problems. If e.g. a trend larger than the minimal detectable bias is present, the surveys might be rejected and corrected. Hereafter, the following survey is rejected as well, resulting in the trend to be ignored, see Figure 6.1. In this case, the difference between the surveys is too big compared to the Minimal Detectable Bias, so all surveys are rejected and the trend is completely removed.

One possibility to solve this would be increasing the MDB, by changing the levels of significance. When the level of significance for an outlier is lowered, a single outlier would be less likely to be detected. Instead, a trend would become more likely to be detected. In that case however, the filter is less sensitive for outliers. Another possibility would be the use of global testing. However, another approach is chosen here. In Chapter 5 it has already been demonstrated that the deformation analysis behaves very well in case of outliers, so why not use this information to improve the Kalman filter performance. In case an outlying survey is detected with the area deformation test, the estimated deviation can be used to correct the survey before starting the Kalman filter. Because the deformation analysis is a batch processing procedure, it will be clear whether a difference between two surveys is caused by an outlier or by a trend. In Paragraph 6.4.1 it is shown that this approach indeed leads to correct results. It is assumed that a deviation is caused by an error and therefore has to be corrected. A deviation can however also be caused by real deformation, therefore the user has to decide whether the outlier is corrected or not.

In this case, the input data have already been corrected for outlying surveys. Therefore the predicted depths will not be influenced by outlying surveys anymore. Still, single outlying points can be present. Therefore, the Kalman filter could be extended with a LS test to correct for these outliers as well. The

MDB for a single outlier is however much higher than the MDB for an outlying survey, therefore a situation as sketched in Figure 6.1 is less likely to occur.

Summarising, the DIA procedure does not work very well for outlying surveys. Because the MDB for an outlying survey is rather small, the possibility that a trend is ignored is rather high. A situation might occur that every survey is rejected and corrected, resulting in a removal of the trend. Outlying surveys are better detected by the deformation analysis procedure, therefore the data can be corrected for outlying surveys before starting the Kalman filter. However, besides an outlying survey, single outlying points might still be present. Therefore, the DIA procedure can be applied only for single points. In that case the MDB is much higher, so the probability of ignoring a trend is not that high.

## 6.3 Wave propagation in the Kalman model

In the Kalman model, as described in Section 4.2, the sea floor is considered as a horizontal plane: no sand waves are taken into account. However, large parts of the North Sea are covered with sand waves and they have a major impact on the predictions in the Kalman model, see the simulations of Section 5.7. If however some knowledge about the magnitude and migration of the sand wave is available, this can be included in the Kalman model to improve the predictions.

In Paragraph 3.2.3 a mathematical description of the sea floor was given. The sea floor is expressed in terms of average depth, slope and a sand wave:

$$E\{\underline{d}_p\} = \left[\begin{array}{ccccc} 1 & x_p & y_p & \cos(\frac{2\pi x_p}{\lambda}) & \sin(\frac{2\pi x_p}{\lambda}) \end{array}\right] \left[\begin{array}{c} \overline{d} \\ \overline{\alpha}_x \\ \overline{\alpha}_y \\ u \\ v \end{array}\right]. \tag{6.7}$$

In this equation the last two terms, $u$ and $v$ describe the sand wave. These are calculated from the more common wave parameters amplitude ($A$) and phase ($\phi$) as:

$$\begin{aligned} u &= A\cos\phi \\ v &= A\sin\phi. \end{aligned} \tag{6.8}$$

The advantage of using $u$ and $v$ instead of the amplitude and the phase is that now a linear relationship exists, see Equation 3.26. Inversely, the amplitude and phase of the wave can be calculated from $u$ and $v$ as:

$$\begin{aligned} A &= \sqrt{u^2+v^2} \\ \phi &= \arctan\left(\frac{v}{u}\right). \end{aligned} \tag{6.9}$$

To include these parameters in the Kalman model, the transition matrix $\Phi$, which defines the evolution of the parameters $u$ and $v$ from one epoch to the next has to be known. In formulae this can be written as:

$$\left[\begin{array}{c} u_0 \\ v_0 \end{array}\right] = \left[\begin{array}{c} A_0\cos\phi_0 \\ A_0\sin\phi_0 \end{array}\right], \tag{6.10}$$

at epoch $t_0$ has to be evolved to epoch $t_1$.

When the velocity of the sand wave is known, expressed as a change in amplitude $\Delta A$ and in phase $\Delta\phi$ per epoch, $u_1$ and $v_1$ can be expressed as:

$$\left[\begin{array}{c} u_1 \\ v_1 \end{array}\right] = \left[\begin{array}{c} (A_0+\Delta A)\cos(\phi_0+\Delta\phi) \\ (A_0+\Delta A)\sin(\phi_0+\Delta\phi) \end{array}\right] = \left[\begin{array}{c} \Phi_{u,v} \end{array}\right] \left[\begin{array}{c} u_0 \\ v_0 \end{array}\right], \tag{6.11}$$

The unknown matrix $\Phi_{u,v}$ defines the transition between two epochs. With Equations 6.10 and the expressions

$$\begin{aligned} \cos(\phi_0 + \Delta\phi) &= \cos\phi_0 \cos\Delta\phi - \sin\phi_0 \sin\Delta\phi \\ \sin(\phi_0 + \Delta\phi) &= \sin\phi_0 \cos\Delta\phi - \cos\phi_0 \sin\Delta\phi, \end{aligned} \tag{6.12}$$

$u_1$ and $v_1$ can be written as:

$$\begin{aligned} u_1 &= \frac{A_0 + \Delta A}{A_0} \cos\Delta\phi \cdot u_0 - \frac{A_0 + \Delta A}{A_0} \sin\Delta\phi \cdot v_0, \\ v_1 &= \frac{A_0 + \Delta A}{A_0} \sin\Delta\phi \cdot u_0 + \frac{A_0 + \Delta A}{A_0} \cos\Delta\phi \cdot v_0. \end{aligned} \tag{6.13}$$

Analogous to the original model, where a depth and a trend were estimated, trends in $u$ and $v$ are added as well. Hence, from Equations 6.13 the matrix $\Phi_{u,v}$ can be calculated as:

$$\Phi_{u,v} = \begin{bmatrix} \frac{A_0+\Delta A}{A_0}\cos\Delta\phi & \frac{A_0+\Delta A}{A_0}\cos\Delta\phi & -\frac{A_0+\Delta A}{A_0}\sin\Delta\phi & -\frac{A_0+\Delta A}{A_0}\sin\Delta\phi \\ 0 & 1 & 0 & 0 \\ \frac{A_0+\Delta A}{A_0}\sin\Delta\phi & \frac{A_0+\Delta A}{A_0}\sin\Delta\phi & \frac{A_0+\Delta A}{A_0}\cos\Delta\phi & \frac{A_0+\Delta A}{A_0}\cos\Delta\phi \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{6.14}$$

The transition matrix $\Phi$, as defined in Section 4.2, Equation 4.13 has to be extended with the transition matrix $\Phi_{u,v}$. Therefore the total transition matrix becomes:

$$\Phi = \begin{bmatrix} I_{J\times J} & I_{J\times J} & & \\ & I_{J\times J} & & \\ & & & \Phi_{u,v} \end{bmatrix}. \tag{6.15}$$

This is a square matrix with dimension $2J+4$, where $J$ denotes the number of support points in the Kalman filter model.

Besides the transition matrix, the state vector has to be changed as well:

$$x = \begin{bmatrix} d_1 \\ \vdots \\ d_J \\ \frac{\Delta d_1}{\Delta t} \\ \vdots \\ \frac{\Delta d_J}{\Delta t} \end{bmatrix} \quad \text{is now replaced by:} \quad x = \begin{bmatrix} d_1 \\ \vdots \\ d_J \\ \frac{\Delta d_1}{\Delta t} \\ \vdots \\ \frac{\Delta d_J}{\Delta t} \\ u \\ \frac{\Delta u}{\Delta t} \\ v \\ \frac{\Delta v}{\Delta t} \end{bmatrix}. \tag{6.16}$$

Finally, the initial values for the model are:

$$\underline{\hat{x}}_{0|0} = \begin{bmatrix} \overline{d} \\ \vdots \\ 0 \\ \vdots \\ u_0 \\ 0 \\ v_0 \\ 0 \end{bmatrix} ; \; Q_{\hat{x}_{0|0}} = \begin{bmatrix} \sigma^2_{\overline{d}} & & & & & & & \\ & \ddots & & & & & & \\ & & \sigma^2_{trend} & & & & & \\ & & & \ddots & & & & \\ & & & & \sigma^2_{u_0} & & & \\ & & & & & \sigma^2_{t,u0} & & \\ & & & & & & \sigma^2_{v_0} & \\ & & & & & & & \sigma^2_{t,v0} \end{bmatrix}. \tag{6.17}$$

The initial values $u_0$ and $v_0$ can be calculated with Equation 6.10, the initial variances $\sigma^2_{u_0}$ and $\sigma^2_{v_0}$ can be calculated by propagating the variances of $A_0$ and $\phi_0$:

$$\begin{aligned} \sigma^2_{u_0} &= \cos^2\phi_0\sigma^2_{A_0} - A_0^2\sin^2\phi_0\sigma^2_{\phi_0} \\ \sigma^2_{v_0} &= \sin^2\phi_0\sigma^2_{A_0} + A_0^2\cos^2\phi_0\sigma^2_{\phi_0}. \end{aligned} \tag{6.18}$$

The initial variances for the trends in $u$ and $v$ are the same as the initial variances for the trend in depth.

The planar part of the Kalman filter model remains the same. So, when the amplitude, phase and the trends in amplitude and phase are known, the sand wave can be included in the prediction. Fortunately, these parameters can be calculated quite well in the deformation analysis. So, first a part of the deformation analysis is performed, to determine the sand wave parameters and they serve as input for the Kalman model. However, these wave parameters need to be known with sufficient accuracy, otherwise the predicted results are still not accurate. Another possibility would be to estimate the initial magnitude and velocity of the sand wave from other sources, such as physical sand wave models, see e.g. [Németh, 2003] or by determining the crest positions, see e.g. [Knaapen, 2004].

## 6.4 Test of the extended Kalman model on simulated data

In the two previous sections it was described how the Kalman filter can be improved, such that the problems found in Chapter 5 are solved. In this section it is tested whether these adaptations really lead to the right conclusions. Therefore, three of the data sets used in Chapter 5 are tested with the new method. These data sets are:

- The data set with an outlying survey,
- The data set with a single outlying value,
- The data set with the migrating sand wave.

### 6.4.1 The data set with an outlying survey

In this paragraph, the data set with an outlying plane in survey 3 is processed. This outlying plane can e.g. be caused by a systematic error. The results can be compared with those of Section 5.2. In that section, it was shown that the error in survey 3 could be detected with the deformation analysis procedure. However the prediction is still influenced by this outlier. On the other hand, the Kalman filtering is affected even more, resulting in an upward trend to be estimated. In Figure 6.2, the results with the extended Kalman filter are illustrated for this area. So with this method outliers are first detected by the deformation analysis procedure. Hereafter the measurements can be corrected with the estimated deviation. In reality however, an outlying survey can either be due to a measurement error or due to some real deformation. Therefore the user has to decide whether the outlying survey has to be corrected or not. In this example however, it is assumed that the outlying survey is due to a systematic error in this survey, so the measurements of survey 3 are corrected. Thereafter, the Kalman filter is applied. The left picture of Figure 6.2 shows the predicted depths five years ahead. When comparing this figure with the results of the deformation analysis (Figure 5.8) and Kalman filtering (Figure 5.12), it is obvious that the results are correct. The predicted depth is about 30 m, exactly what it should be. On the right, the time series of the point (1,1) is illustrated. Here it is clearly visible what the result of this procedure is: the outlying value in the third survey is corrected and the trend line remains static. The predicted depths are not exactly 30 m but vary a little bit, but this is of course due to the stochastic character of the data.

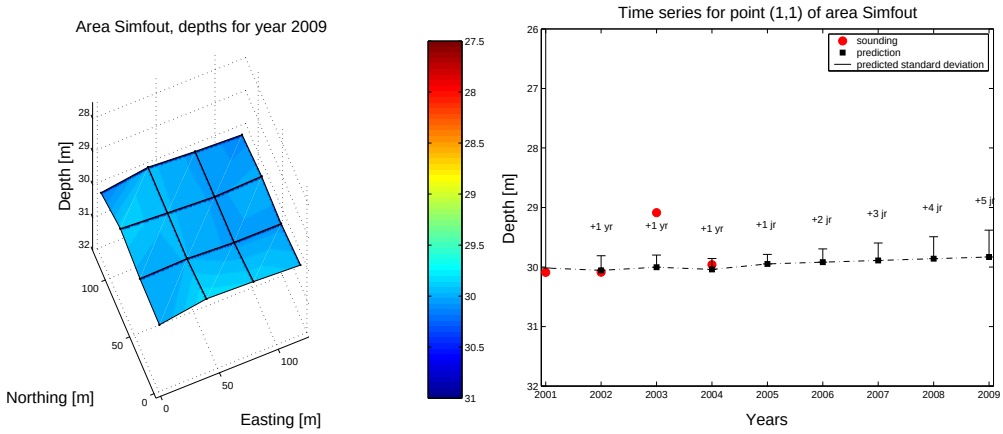



**Figure 6.2:** Results for the area with an outlying plane in survey 3. On the left the predicted depths for 2009, on the right the time series of the point (1,1).

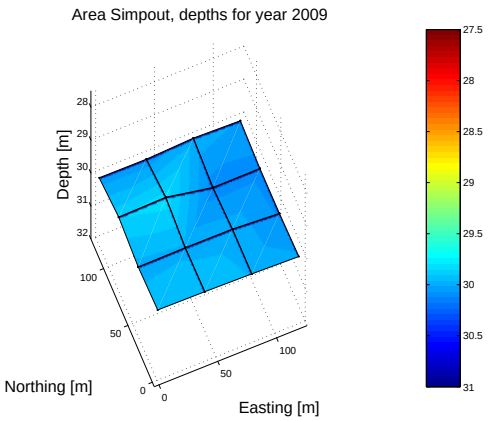



**Figure 6.3:** Results for the area with one outlying point (41,41).

### 6.4.2 The data set with a single outlying value

In this paragraph, the data set with a single outlying value is processed. These results can be compared with those of Section 5.3. In that section, it was shown that an outlying value influences the prediction of both the deformation analysis and the Kalman filter. In deformation analysis however, only the outlying point is influenced whereas with the Kalman filter the effect is spread out to several points. The predicted depths of the extended Kalman filter for this data set are shown in Figure 6.3. Here, it can be seen that the outlying point is corrected, such that the predicted depths all remain stable. This correction is a result of the DIA procedure.

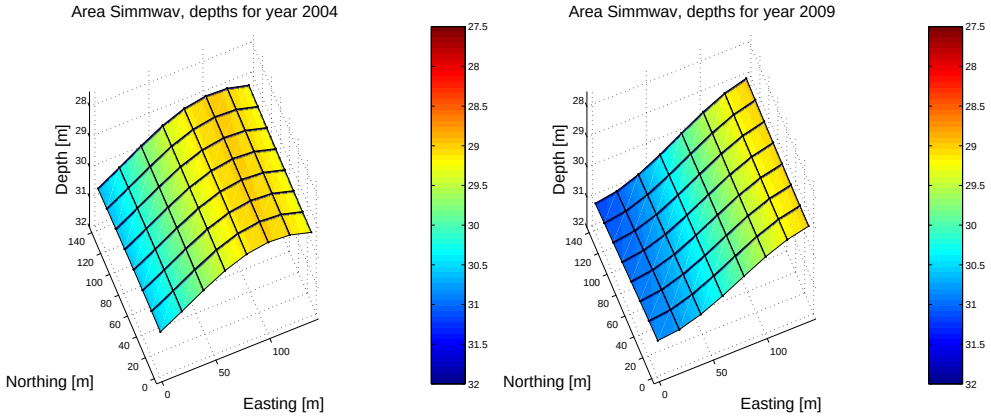



**Figure 6.4:** Predicted depths for the migrating sand wave. On the left the sand wave in 2004, with the crest at $x = 100$ m. On the right the sand wave in 2009, where the crest has moved to the right and is now at about $x = 150$ m.

### 6.4.3 The data set with a migrating sand wave

The results of this paragraph can be compared with those of Section 5.7. In that section it was shown that the wave parameters can be determined quite well. The migration rate of the wave was determined as well as the amplitude and phase of the wave in the first survey, by the method described in Section 3.2. For this area, the following parameters were estimated:

$$\begin{aligned} A_0 &= 0.9961 \text{ m} \\ \frac{\Delta A}{\Delta t} &= -0.0282 \text{ m/year} \\ \phi_0 &= 3.1207 \\ \frac{\Delta \phi}{\Delta t} &= -0.2175. \end{aligned}$$

Another important parameter is the wavelength $\lambda$, in this case this is defined as 300 m. The estimated migration rate of the sand wave can be calculated from the trend in phase and the wavelength as:

$$\frac{d\psi}{dt} = -\frac{\Delta \phi}{\Delta t} * \frac{\lambda}{2\pi}, \tag{6.19}$$

with $\psi$ the crest distance, the distance that the crest has move w.r.t. to the first survey. So, this results in an estimated migration rate of 10.3849 m/year. The simulated velocity was 10 m/year, the difference is a result of the additional noise. The same holds for the small trend that is found for the trend in amplitude. In the deformation analysis part, no other trends or deviations were found, so the mean depth and slope are accepted as static.

The predicted depths, for the deformation analysis as well as for the Kalman filtering, did not take the waveform into account. This results in errors in the predicted depths, see Figures 5.39 and 5.41. In this paragraph, the results for the new method are shown. Here, the waveform is taken into account in the prediction with the Kalman filter.

In Figure 6.4, the predicted depths for 2004 and 2009 are given. The dimension of this area is 140×140 m. In 2004, the sand wave crest is located at $x = 100$ m. In 2009 however, the crest has shifted to the right, just outside the area. The sand wave shape however is still the same. When counting the distance that the sand wave has migrated in these five years, it can be concluded that it has migrated about 50 m. This is correct, because with Equation 6.19 the migration rate was calculated as 10.38 m/year.

In Figure 6.5 the time series of two points are shown, the same points as in Figure 5.40. So, the sand wave propagation works very well for this example. For practical use however, it is necessary that the

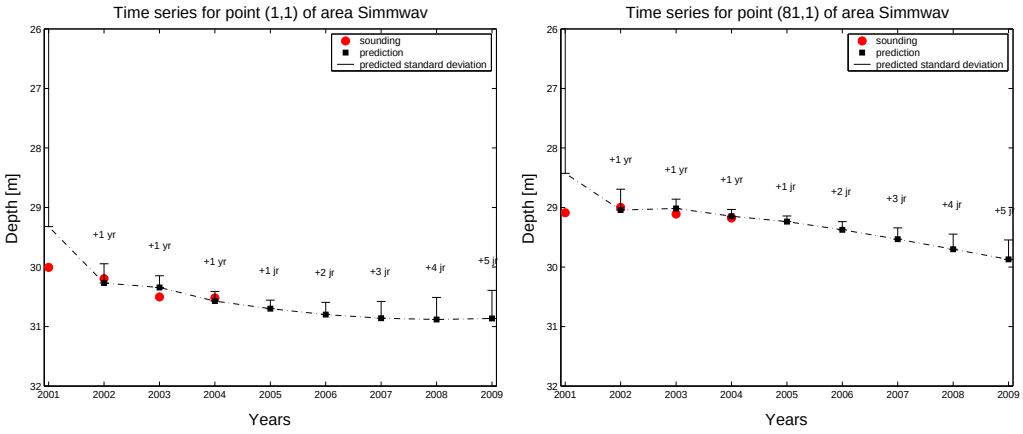



**Figure 6.5:** Time series for two points on the migrating sand wave. Point (1,1) is located on the left of the area and is deepening. Point (81,1) is initially on the sand wave crest, but is deepening after the crest has passed.

sand wave parameters are estimated well.

The three data sets shown in this section show that the adaptations in the Kalman filter, as described in the first two sections of this chapter, give good results for the simulated data sets. The prediction gives good results in all three cases. For the other simulated data sets, the Kalman filter gave already good results. This has not changed because these adaptations can be skipped when not necessary. If no sand waves are present, the Kalman model is not extended with the sand wave part. When a sand wave is present, but does not change the estimated trends in amplitude and phase are zero. Then the sand wave transition matrix $\Phi_{u,v}$ reduces to an identity matrix, see Equation 6.14. Therefore, it can be concluded that the predictions with this method are in principle correct.

## 6.5 Overview of the new method

In the previous sections it was shown that the principle of this combined method works well on the simulated data set. To get it work with an arbitrary data set, the following steps have to be carried out:

1. Reading of the data. For each measured point an $x$-coordinate, $y$-coordinate and a depth value are needed. Missing points have to be detected and identified with some dummy value.

2. Calculation of the standard deviation of the depth values. This is necessary for the testing whether a deviation is significant or not. The Hydrographic Service has derived a formula to calculate the theoretical standard deviation for every single point in a single beam data set, see Equation 3.1. For multibeam data however, this is more complicated. Some programs exist to calculate the theoretical standard deviation of multibeam data, see e.g. [De Koning and Valckenier von Geusau, 2000]. This is however not taken into consideration in this project. Therefore, for the standard deviation for multibeam data a value of 25 cm is chosen for each point. Of course, this is a simplification, but the actual calculation of the standard deviations goes beyond the scope of this project.

3. Calculation of the covariance function. The covariance function is necessary for two purposes. First a covariance function is needed if the data have to be interpolated with Kriging. The second reason is that the covariance function can be used to determine the wavelength of the sand waves. So even when the data do not have to be interpolated, the covariance function is useful, if sand waves are present. Besides this, the direction of the highest variability has to be calculated, for an accurate modelling of the sand waves. This is done by comparing the gradients of the data.

4. Kriging. In case of single beam data, the Kriging step is necessary to create a regular grid of data points. In case of multibeam data this is usually not necessary. Because of the large amount of

data, other less time-consuming methods can be applied to grid the data. However, the advantage of Kriging above other interpolation methods is the property that the spatial correlation of the sea floor is taken into account to obtain Best Linear Unbiased Estimations.

5. Deformation analysis. In this step the data of all epochs are analysed and tested on dynamic behaviour. A test per point and a test per area can be performed. The test per point is only for indication of the dynamics. The test per area is used to detect deviating surveys and sand wave parameters. These data are used in the following steps.

6. Kalman filtering. After detecting outlying surveys, they are either ignored or the error is estimated and the survey is corrected. The corrected data form the input of the Kalman filter. All epochs are processed and, based on that, a prediction is made. Within the Kalman filter an additional test for outlying points is performed.

7. Calculate oversteps. This option gives the possibility to predict when a certain limit, such as the Nautical Guaranteed Depth will be exceeded. Another possibility would be to predict the moment when the accuracy exceeds the S44 limit.

8. Give output. The output of the Kalman filter is written to files and plotted. Different plots are available: the depths and variances, a plot of the predicted oversteps, time series of different point and a plot of the predicted behaviour of the individual points.

## 6.6 Tests on real data

In this section, the new method is tested on two real data sets. Therefore two data sets that are already used in Chapters 3 and 4 are used again.

### 6.6.1 Data set NA

First, data set NA is processed. This data set was acquired by the North Sea Directorate, with a multi-beam echo sounder. As said before, the standard deviation of this data is not known. Therefore a value or 25 cm is assumed. In Figure 6.6 the results of the first survey (1991) of this data set are illustrated. In this data set, the sand wave is very obvious.

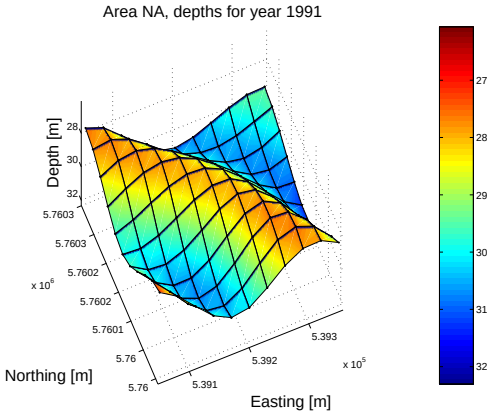



**Figure 6.6:** The first survey (1991) of the area NA.

As this data set is already gridded, the Kriging procedure can be skipped. Only the covariance function has to be calculated to estimate the sand wave length and the direction of the sand wave. These were estimated as 225 m and 43 degrees respectively. Next, the deformation analysis procedure is performed. The results of this analysis are shown in Figure 6.7. The upper left graphic shows the results of the point

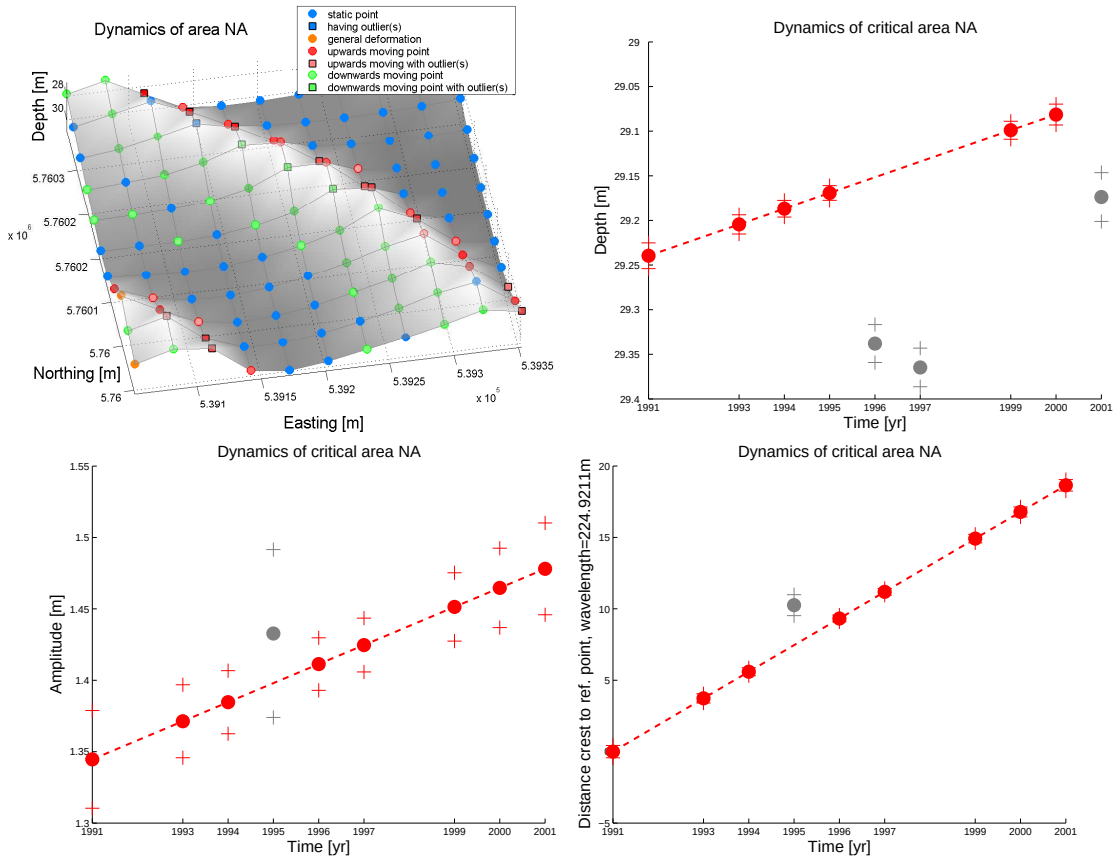



**Figure 6.7:** Results of the deformation analysis for area NA. In the upper left the results for the point test are shown, in the upper right the mean depth. The lower graphics show the behaviour of the sand waves. These are the adjusted depths, with the estimated trends.

test. This gives already an indication that the sand wave is moving, because upward and downward trends are estimated on opposing sides of the sand wave. The migration is confirmed by the lower graphics, where the trend in amplitude and crest distance are shown. The crest distance after 11 years is about 18 m. So in these 11 years, the sand wave has moved about 18 m in northeasterly direction, which is about 1.6 m/year. This is in accordance to the values found in the literature [Németh, 2003]. The amplitude is estimated at about 1.4 m, with a growth rate of 1.3 cm/year. This is a rather small amplitude for a sand wave, but when comparing this to Figure 6.6, this value seems to be right. A trend in amplitude which is found, could also be caused by the fact that not a complete sand wave is present in this area, but this is not further investigated. It is also not a big problem, because the Kalman filter will adapt to the real situation. The upper right graphic shows the mean depth. There is a small upwards trend and some outlying surveys are present.

In Figure 6.8 the predicted depths for 2006 are given. At first sight, these data seem rather good. Also when the predicted depths are compared to the filtered depths of the last survey, as shown in the left graphic of Figure 6.9, the prediction seems to work fine. The sand wave has clearly migrated a few meters in those years. However, during the calculation of this results, in every survey up to 20 points (out of 121) are rejected by the DIA procedure, even with a small $\alpha$ (0.001). Or in other words, up to 20 points are classified as outliers, with 99.9 % probability. When looking at Figure 6.6, it can be seen that the surface is very smooth. Therefore the number of outliers seems unrealistic.

In the right graphic of Figure 6.9, the difference between the predicted depths for 2001 and the actual measurements of that year is illustrated. Ideally this difference should be almost zero, only some

stochastic influence should be present. The predicted values are the result of the modelled sand wave propagation, and if the model is right, this should be acknowledged by the data. However, here this is clearly not the case. Especially around the sand wave crests differences up to 2 meters occur, so this explains why so many outliers are detected in the DIA procedure. Therefore, the amount of rejected points is not caused by real outliers, but rather by a model error.

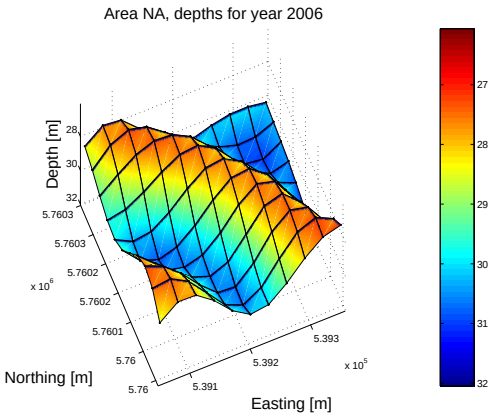



**Figure 6.8:** The predicted depths for 2006.

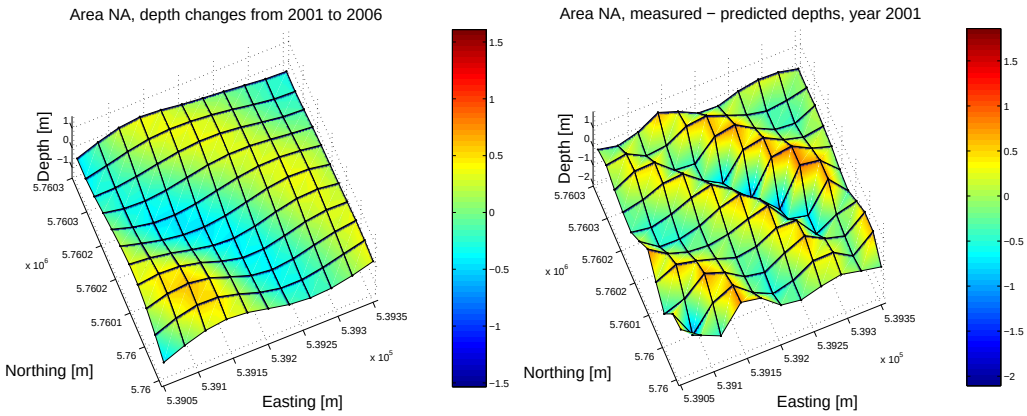



**Figure 6.9:** On the left the predicted change in depth from 2001 to 2006. On the right the difference between the predicted depths for 2001 (based on the measurements until 2000) and the measurements of 2001.

To have a closer look at the differences between the predicted depths and the measurements, two depth profiles are taken. The location of these profiles is illustrated in Figure 6.10. In Figure 6.11 the measured and predicted depths along profile 1 are illustrated. This is done for the first year (1991) and the last year (2001). In 1991 the predicted values do not fit very well to the measurements. This can partly be explained by the resolution of the real data. Because the data are on a 30×30 m grid, and the profiles are taken under an angle of approximately 45 degrees, the spacing between the data points in the profile is about 40 m. Apart from this, the sand wave model does not fit very well. Especially the amplitude of the sand wave is underestimated, because the actual shape of the sand wave is not a sine. Its crests are narrower than the troughs and the slope in the direction of the current is less steep than in the opposite direction. It must however be remarked that the predicted values for this first year are only initial values, so it is not yet a real problem when they do not fit very well, the Kalman filter will adapt to the data.

On the right, the profile for 2001 is illustrated. These predicted values are the result of running the Kalman filter for ten years. Therefore the estimate of the amplitude and phase of the sand wave should

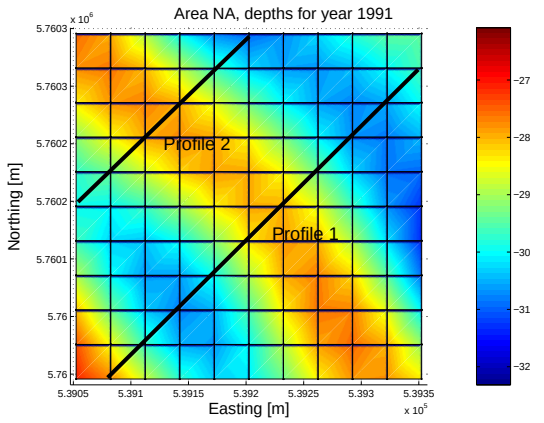



**Figure 6.10:** The location of the two profiles of data set NA.

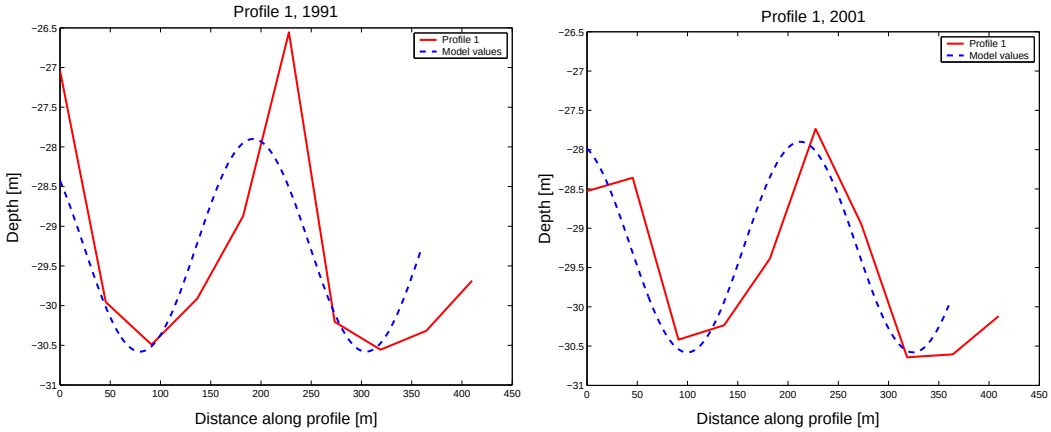



**Figure 6.11:** Profile 1 for the measurements and modelled depths. On the left the profile for 1991, on the right the profile for 2001.

be more accurate. Besides, the migration of the sand wave should be visible as well, compared to the 1991 profile. From the figure, it can be seen that the Kalman model has actually worked quite well. The modelled values and the measurements do not differ very much anymore. Still the shape of the sand wave is not a sine in reality, but the amplitude and phase are estimated quite well. Although the profile fits in general quite well, the local differences are however still rather high.

In Figure 6.12, the results for the second profile are given. This profile is located more to the North, and is also shorter than profile 1. Here the modelled data fit less well to the measurements than in the first profile. Especially for the 1991 profile, the sand wave crest is in reality located more to the left than in in the model. In the 2001 profile it fits better, but still the differences are high. An explanation herefore could be the fact that this profile is located on the edge of the data set, whereas profile 1 is located in the centre. For determination of the sand wave parameters a single wave model is used: it is assumed the the crest of the wave is a straight line. In Figure 6.10 it is obvious that the wave crest is curved. Therefore the modelled wave fits better in the centre than at the edges.

However, the difference between the actual sand wave shape and the modelled sand wave shape does not completely explain the differences between the predicted depths and the actual depths. In Figure 6.9 positive differences up to 2 meter occur. This means that in reality the sand wave is is about 2 meter shoaler then predicted. Negative differences of up to 1.5 m are present. In the profiles above however, these differences are not that large, especially the positive differences are a lot less than in Figure 6.9.

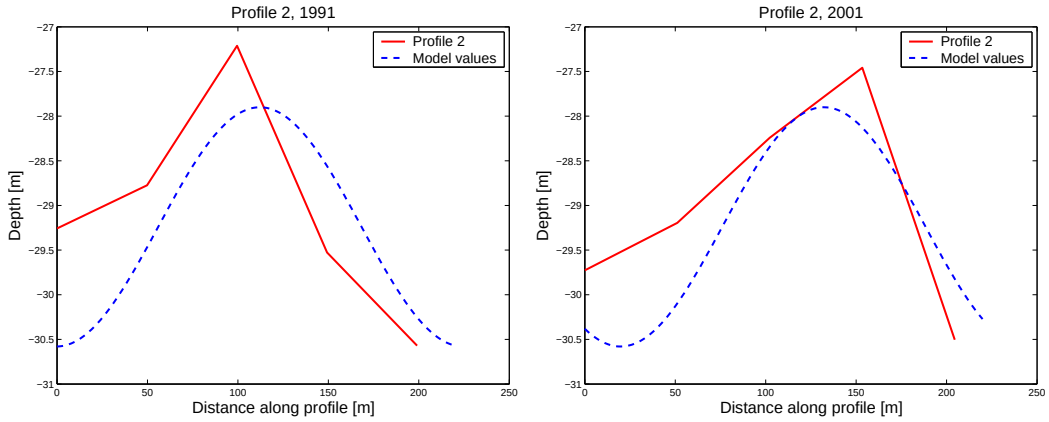



**Figure 6.12:** Profile 2 for the measurements and modelled depths. On the left the profile for 1991, on the right the profile for 2001.

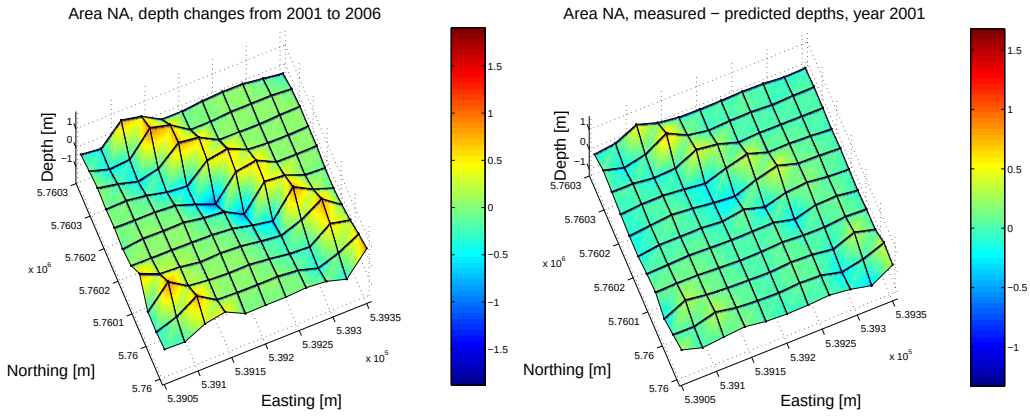



**Figure 6.13:** Results for the prediction of area NA, in case no interpolation is carried out. On the left: the difference between the (filtered) depths of 2001 and the predicted depths for 2006. On the right the difference between the predicted depths in 2001 and the measurements of the same year.

The origin of this effect can be explained by the way this Kalman filter works. As is described in Section 4.2 an interpolation is performed as well. This interpolation has two goals: first, the amount of data that has to be processed, is reduced. Second, the interpolation works as a smoothing filter, so the noise and the effect of outlying values is reduced. However, this interpolation is only based on the distance from the data point and the kernel point. The anisotropy of the sea floor is not taken into account, and therefore the shape of the sand wave might be partly filtered out.

In Figure 6.13 the results are shown for the case that the interpolation procedure is left out. The difference between the predicted depths and actual depths for 2001 are shown, as well as the predicted change in depth from 2001 to 2006. In this case the difference between the actual depth and the prediction is a lot less, compare with Figure 6.9, so the prediction works a lot better. The mean difference is 12 cm, with maximum values of +45 cm and -12 cm. Still the largest differences occur near the sand wave crest. From the predicted changes it can be seen that the sand wave propagation still works well, and when comparing this to Figure 6.8 the smoothing effect will be clear. Therefore it can be stated that the interpolation kernel has a deteriorating effect on the prediction. In this situation, the DIA procedure also gives better results, because the differences between the model and the real data are smaller. Only a few points are rejected, so these points can be assumed to be real outliers. However, the final predicted depths, as shown in Figure 6.14 look less well than the results in case an interpolation is performed. The

smoothing effect of the interpolation kernels results in a visually nicer representation of the area.

The right graphic of Figure 6.14 gives another useful visualisation of the prediction, namely the difference between the measurements of the last survey and the prediction one year ahead, for the individual points. This representation gives insight in the local dynamics and can be used for e.g. a survey planning. It shows the same pattern as the plot of the point deformation test, the upper left graphic in Figure 6.7. The plot of the point deformation test is however only based on past dynamics, whereas in this plot the predicted dynamics are shown.

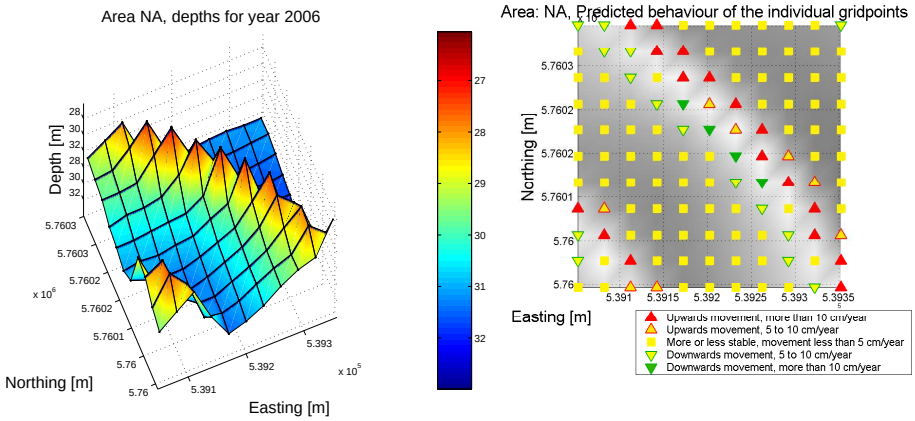



**Figure 6.14:** On the left the predicted depths for area NA, in case no interpolation is carried out. On the right the one year prediction for the individual points.

From these results, it can be concluded that the Kalman model gives promising results for the prediction of the sand wave migration. Although the mean difference between the actual and predicted values is not much, about 12 cm on average, still differences up to 45 cm exist between the mathematical model of the sand wave and the real data. In this case the standard deviation of the prediction is 17.4 cm, so the differences are still significant. The reason for these differences is that a sine wave was modelled, but in reality a sand wave is not a sine wave. When the shape of the sand wave is known, the model should fit better to the real data. This is however very difficult, if not impossible, because the sand waves are not very regular. Another problem are the curved crests and the varying direction of the waves. This effect could however be decreased when considering smaller areas.

### 6.6.2 Data set B

The second data set that was processed is data set B, which was acquired by the Hydrographic Service. This data set is obtained with a single beam echo sounder and therefore the data have to be interpolated first. Data are available for the years 1991, 1995, 2000, 2001, 2002 and 2003. The Kriged depths of this area are shown in Figure 3.4 in Chapter 3. This a difficult area: although the area is clearly not flat, it does not have an obvious sand wave either. Therefore no sand wave estimation is performed. As is known from Section 3.3 two outlying surveys and an upward trend of 4 cm were estimated.

The predicted depths for 2002 and the difference between prediction and measurement for 2002 are given in Figure 6.15. The difference between predicted depths and measured depths are somewhat larger than for the data set NA, but still the method performs reasonably well. The maximum differences are +0.76 m and -0.47 m. The one year prediction for the individual points is shown in Figure 6.16. This area is not very dynamic, in only a few points changes are predicted.

Although no obvious sand wave is visible in this area, a sand wave test is performed as well. That is because some structures are present, which could be sand waves as well. It is probably a junction of two

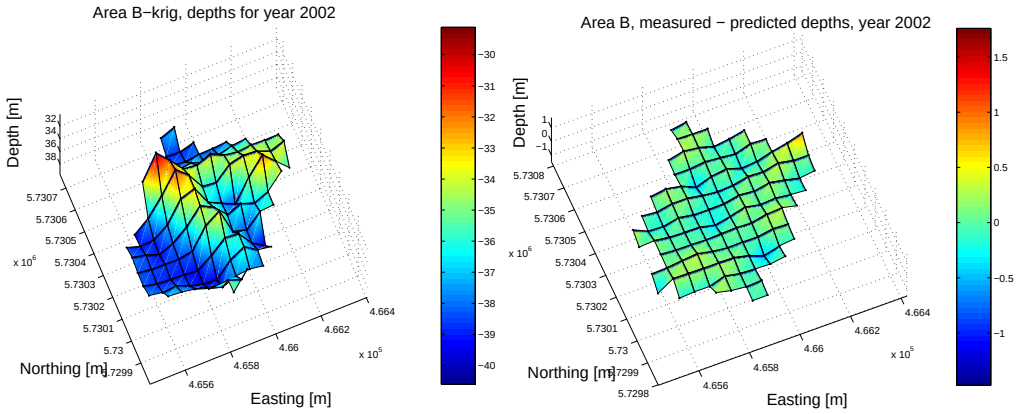



**Figure 6.15:** Results of the Kalman filter for area B. On the left the predicted depths for 2002, on the right the difference between the predicted depths and the measurements.

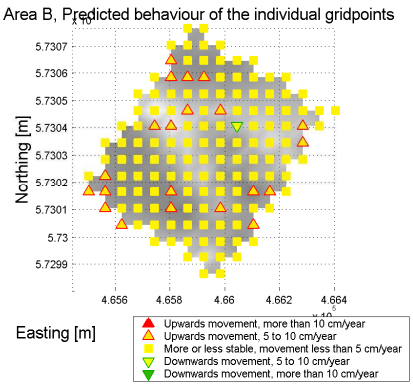



**Figure 6.16:** The one year prediction for area B.

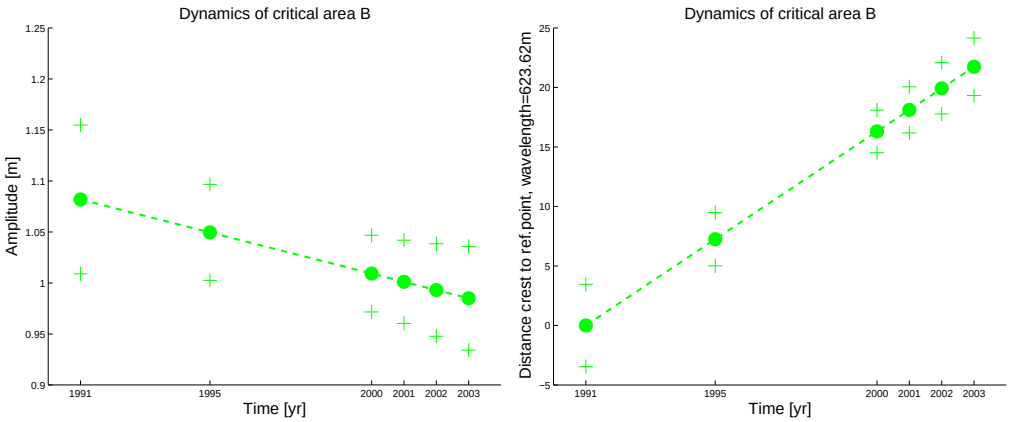



**Figure 6.17:** Deformation analysis for the sand wave parameters. On the left the amplitude, on the right the migration.

sand waves with different directions. The results of the sand wave test are illustrated in Figure 6.17. So a wave can be fitted to the data, but the results do not seem to be realistic. The amplitude is estimated as 0.72 m and the sand wave is assumed to be static. However from Figure 6.15 it is obvious that the height differences are more than 0.72 m. Therefore it can be concluded that the sand waves have to be regular. In an area like this, where different features are present, the detection does not work anymore. Because in this area no dynamic sand wave is detected, the results of the prediction are the same as in Figure 6.15.

## 6.7 Possibilities and limitations of this method

This method combines the performance of the deformation analysis with Kalman filtering. The deformation analysis is useful to detect changes that have happened in the past. Statistics are used to decide whether a change is significant or not, analyses are performed on various scales. The test per point is useful because it gives a good resolution. For every point it is determined whether it is dynamic or not. However, due to the low redundancy, the reliability is not very high. This is indicated by the large Minimal Detectable Biases. Hence, the main purpose of this point test is to give a general overview of the dynamics and it could be used to determine areas for further investigation.

The area test gives a better reliability, because the redundancy is much higher. The resolution is however lower, so local changes cannot be detected. The area test can however be used to detect outlying surveys and sand wave behaviour.

The sand waves can in principle be detected very well, as long as the wavelength and direction are known. These can be estimated from the covariance function and the gradients of the data respectively. However, the sand wave pattern has to be regular. When the crests are curved, or two sand waves in different direction are present, this does not work anymore. Therefore, this should be used only for small areas, such that the shape of the sand wave is as regular as possible.

The sand wave parameters and outlying surveys are used as input for a Kalman filter. With this method, the present state can be used for prediction of the depths a few years ahead. Therefore, the depth trends are analysed and the sand wave is propagated. This however only works at a general scale, the migration of the sand wave as a whole can be predicted, but in the individual points rather large differences occur. This is mainly due to the used model of a sand wave, which only approximately fits to the real sand waves. Therefore, this method is used to model dynamics, rather than modelling the sea floor.

Another limitation is a change in the natural process of sand wave migration. In this model it is implicitly assumed that a rather constant migration is present. If however sudden changes in the process occur, the model does not fit anymore. This sudden change can for example be due to extreme weather circumstances, like storms. Another aspect is human intervention: dredging causes a cut off of the sand wave crests. It has been proved that sand waves regenerate in a couple of years after dredging [Knaapen and Hulscher, 2002], but this causes a discontinuity in the trend model, because only one trend is estimated. A method to use a Kalman filter with more than one trend is proposed by [Calder and Mayer, 2003]. This is however not considered in this study.

The main purpose of the prediction is the planning of the surveys. The predicted depths can help to decide when or where to survey again. It can e.g. be predicted when the sea floor will become too shallow for shipping, or when the uncertainty about the depths becomes too large. Although considerable errors are present between the real depth and the predicted depth, this method can still be useful. Because the migration of the sand wave can be predicted rather well, it can be used to determine when for example a sand wave will enter a sea channel. Besides, the estimation of the sand wave amplitude seems to work well, so the predicted depths will not be shoaler than the real depths.

The main advantage of this model instead of the Kalman Model, as developed by the North Sea Directorate, is that in this case the shape of the sand waves are taken into account, whereas originally only

a constant trend could be estimated. Consider e.g. a point near a sand wave trough, using the original model it could be estimated that the depth at that point is increasing, so the point would be predicted to remain deepening. When however the trough has passed, it could be possible that the point will rise again. This model would be able to predict that, whereas in the original model this would not be the case.

So summarising, this method works in two ways. On the one hand it can be used to detect what has happened in the past. This is mainly used by the Hydrographic Survey, to determine which areas have priority to be surveyed. Areas that are highly dynamic are more important than rather stable areas. Hereby not too much confidence should be laid on the results for the individual points, but it should rather be used to give an overview. On the other hand it can be used to predict the depth a few years ahead. This is mainly used by the North Sea Directorate, to decide when the sea channel has to be dredged. This is for example the case when a sand wave enters the channel. Again, not too much confidence should be laid in the results for the individual points, but more on the general behaviour of the area, provided that the area is not too large.

# Chapter 7

# Conclusions and recommendations

As was pointed out in the introduction, the objective of this study is the evaluation, comparison and combination of two methods for analysing time series of bathymetric data. In the Chapters 3 and 4 these two methods were described and evaluated. In Chapter 5, a comparison between these two methods was performed, by means of simulated data. Finally, in Chapter 6 a proposal for the combination of these two methods was made. The conclusions of this study will be presented in Section 7.1. Based on these conclusions a number of recommendations for further research will be done in Section 7.2.

## 7.1 Conclusions

The conclusions of this study can be split up in four parts. Conclusions will be drawn on these four topics:

- Evaluation of the Hydrographic Service method,
- Evaluation of the North Sea Directorate method,
- Comparison of both methods by means of simulated data,
- Combination of both methods.

**Evaluation of the Hydrographic Service method**

For the Hydrographic Service the aim of developing a method to analyse time series of bathymetric data was the demand for knowledge on the dynamic behaviour of the sea floor. Therefore the method is mainly aimed at analysing the past. The Minimal Detectable Biases play an important role, they indicate the strength of the analysis and can be used for planning purposes. Three types of output are present:

- Deformation test on single points. This gives a high resolution picture with the dynamics of the individual points. It gives a good overview of what is happening, however the accuracy is not very high. Due to the low redundancy only large differences are detected. It is however useful for an indication whether the area is interesting for a closer study.
- Deformation test on areas. This gives a more accurate overview of the behaviour of a larger area. However, the detected dynamics are often based on only a few surveys, slightly different values might give different results. If sand waves are present, these can be detected as well and the amplitude and phase can be estimated accurately. The results of the area test give good results, however local changes cannot be detected anymore.
- Prediction of the depths. As the Hydrographic Service is mainly interested in the past, the prediction is not the most important part. The prediction is based on the point test, therefore the accuracy is not very good. Only large deformations, larger than the Minimal Detectable Bias are taken into account. The prediction is however useful to test whether a new survey is in line with the previous ones.

### Evaluation of the North Sea Directorate method

For the North Sea Directorate the aim of developing a method to analyse time series of bathymetric data was the prediction of the depths, in order to plan maintenance activities. Because the prediction is the main aim, a Kalman model is chosen. The following conclusions can be drawn about the performance of this method:

- The Kalman filter proves to be a suitable method for the prediction of the depth. The prediction for a few years ahead gives good results.
- The model is however rather sensitive for outlying values, because no testing procedure is implemented. For individual points this effect is small, because a smoothing filter is used, but outlying surveys, e.g. systematic errors, have more influence.
- The output of the Kalman filter is very much dependent on the calibration parameters and initial values of the model.
- Within the Kalman filter an interpolation is performed. This has the advantage that the amount of data to be processed is reduced. Furthermore, it acts as a smoothing filter, so the noise is reduced and the model is less influenced by outliers. The interpolation does however not preserve the real shape of the sea floor, sand waves might be partly filtered out.

### Comparison of both methods

From the tests performed on simulated data, the differences between the two methods became obvious. The following conclusions can be drawn about the performances:

- In static areas both methods perform well, it is unimportant whether the area is flat, sloping or covered with (static) sand waves; the predicted depths are correct.
- With the deformation analysis outliers can be detected very well. In Kalman filtering however no outlier detection is performed. An outlier causes a bending of the trend line, so the predicted results are much more influenced by outliers.
- In case of trends the predicted depths by the Kalman filter are much better than the prediction by the deformation analysis. This is because the Kalman filter takes all the differences between the epochs into account, whereas with deformation analysis the prediction is only based on accepted trends. Only trends larger than the Minimal Detectable Bias are detected with a certain probability.
- In case of migrating sand waves, both methods give false predictions. This is because only linear trends are estimated, the wave form is not used. Therefore, a point that has been rising is assumed to remain rising, while in case of sand waves a point will fall again when the sand wave crest has passed.

### Combination of both methods

The two methods can be combined to benefit from the strengths of both. Starting point of the combined method is the Kalman filter, where three extra features are added:

- The results of the outlier detection in the deformation analysis are used to correct the input for the Kalman filter.
- The DIA testing procedure is implemented in the Kalman filter, to detect and correct outlying measurements.
- The Kalman filter is extended with a sand wave propagation model.

These extra features proof to work out well. On the simulated data sets, this extended Kalman model always gives the right results for the prediction. This method is tested on real data as well, but some differences between the measured depths and the predicted depths are present. The sea floor is in reality much more complex than as described by this model. Therefore, the method is more suitable for modelling sea floor dynamics than for giving an accurate description of the sea floor.

## 7.2 Recommendations

- The used interpolation method in the Kalman filtering proves to deteriorate the prediction. This is mainly caused by the fact that the interpolation does not take the actual shape of the sea floor into account. It can however be advantageous to implement such an interpolation, because it works as a smoothing filter so it reduces the influence of irregularities and noise. A Kriging based procedure could give better results.

- To make a testing procedure possible, the theoretical standard deviations of the depth need to be known. This is the case for the Hydrographic Service' single beam data, for multibeam this has to be worked out. Some publications and programs exist to make this possible. The actual standard deviations however depend on the equipment and models at the surveying authority, so no general rule can be applied.

- In this study sand waves are modelled as a sine. Although this model gives rather good results, local deviations between the real data and the modelled sand wave exist. This is because in reality the sand wave is not a sine, but has a more complex shape and is not always regular. When however a better representation for the sand wave could be implemented in the Kalman filter, the prediction results would benefit.

- For predicting the sand wave migration in the Kalman model, the parameters sand wave direction, wavelength, migration rate and change in amplitude have to be known. In this study, these are estimated from the time series of data. There exist however physical models for sand wave prediction, that are based on the mechanisms that generate them. These models could be used to give a better estimate of these parameters, or to verify the results of the parameters estimation as it is performed in this study.

- In this study only small areas are considered, where in most cases only a part of a sand wave was visible. However when a larger area, with several parallel and straight sand waves is considered, the wavelength and migration rate estimation might work better. A condition is however that the sand waves must form a regular pattern.

#### Recommendation to the Hydrographic Service

The Hydrographic Service is interested in the dynamic behaviour of the sea floor. The deformation analysis procedure gives good insight in the past dynamics. In addition to that, a prediction for one or a few years ahead is made, to support the planning of the surveys. This prediction can be significantly improved by using this combined method. The point-wise predictions give an accurate description of the expected sea floor dynamics.

#### Recommendation to the North Sea Directorate

The North Sea Directorate are mainly interested in predicting the depths a few years ahead, The Kalman filter is suitable for trend analysis, the main drawback is that only linear trends are considered. Because this combined method does take the sand wave shape in account, a more realistic prediction can be made.

# Bibliography


Anderson, B. D. O. and Moore, J. B. (1979). *Optimal Filtering*. Prentice-Hall, New Jersey.

Calder, B. and Mayer, L. A. (2003). Automatic processing of high-rate, high-density multibeam echosounder data. *International Hydrographic Review*, 4(1):53–68.

Chilès, J. P. and Delfiner, P. (1999). *Geostatistics: modelling spatial uncertainty*. Wiley Series in Probability and Statistics. John Wiley & Sons, New York.

De Jong, C. D., Lachapelle, G., Skone, S., and Elema, I. A. (2002). *Hydrography*. Delft University Press, Delft.

De Koning, R. and Valckenier von Geusau, E. (2000). Gebruikershandleiding MEET. Technical report, Rijkswaterstaat, Meetkundige Dienst, Delft.

Directie Noordzee (2004). www.noordzee.org.

Dorst, L. L. (2003). Zeebodemmonitoring met geostatistiek en deformatieanalyse. Technical report, Koninklijke Marine, Dienst der Hydrografie, Den Haag. in Dutch.

Dorst, L. L. (2004a). Analyse critical areas in de selected track. Technical report, Koninklijke Marine, Dienst der Hydrografie, Den Haag. in Dutch.

Dorst, L. L. (2004b). More efficent and reliable sea floor mapping by geodetic interpretation of bathymetric data. In *Proceedings, ECUA 2004, Seventh European Conference of Underwater Acoustics*, Delft.

Dorst, L. L. (2004c). Survey plan improvement by detecting sea floor dynamics in archived echo sounder surveys. *International Hydrographic Review*, 5(2):49–63.

Gratton, Y. and Lafleur, C. (2001). Matlab kriging toolbox 6.1. University of Quebec. http://www.inrs-ete.uquebec.ca/activites/repertoire/profs/yg/krig.htm.

Hennis, N. (2003). Automatic outlier detection in multibeam data. Master's thesis, TU Delft.

Hulscher, S. J. M. H. and van den Brink, G. M. (2001). Comparison between predicted and observed sand waves and sand banks in the north sea. *Journal of Geophysical Research*, 106(C5):9327–9338.

International Hydrographic Organization (1988). IHO Standards for Hydrographic Surveys. Special publication no. 44, 4th edition.

Isaaks, E. I. and Srivastava, R. M. (1989). *Applied Geostatistics*. Oxford University Press, New York.

Knaapen, M. A. F. (2004). Measuring sand wave migration in the field. comparison of different data sources and an error analysis. In *Proceedings of International Workshop on Marine Sand Wave and River Dune Dynamics II, 2004*, Enschede.

Knaapen, M. A. F. and Hulscher, S. J. M. H. (2002). Regeneration of sand waves after dredging. *Coastal Engineering*, 46(4):277–289.

Knaapen, M. A. F., Hulscher, S. J. M. H., and de Vriend, H. J. (2001). A new type of sea bed waves. *Geophysical Research Letters*, 28(7):1323–1326.

Lee, H. K. H., Holloman, C. H., Calder, C. A., and Higdon, D. M. (2002). Flexible gaussian processes via convolution. Technical Report 02-09, Duke University, Durham.

Lindenbergh, R. (2004). Parameter estimation and deformation analysis of sand waves and mega ripples. In *Proceedings of International Workshop on Marine Sand Wave and River Dune Dynamics II, 2004*, Enschede.

Németh, A. A. (2003). *Modelling offshore sand waves*. University of Twente, Enschede.

NIST/SEMATECH (2003). *e-Handbook of Statistical Methods*. National Insitute for Standards and Technology, Gaithersburg. http://www.itl.nist.gov/div898/handbook.

Olea, R. A. (1999). *Geostatistics for engineers and earth scientists*. Kluwer Academic Publishers, Norwell.

Polman, J. and Salzmann, M. A. (1996). *Handleiding voor de technische werkzaamheden van het Kadaster*. Kadaster, Apeldoorn. in Dutch.

Salzmann, M. A. (1992). *Dynamische Gegevensverwerking II, Filter Analysis*, chapter Part III: Detection Identification Adaptation. Delft University of Technology.

Strang, G. (1988). *Linear Algebra and its Applications*. Harcourt Brace Jovanovich, San Diego.

Teunissen, P. J. G. (2000a). *Adjustment theory: an introduction*. Delft University Press, Delft.

Teunissen, P. J. G. (2000b). *Testing theory: an introduction*. Delft University Press, Delft.

Teunissen, P. J. G. (2001). *Dynamic data processing: recursive least-squares*. Delft University Press, Delft.

Teunissen, P. J. G. and Salzmann, M. A. (1989). A recursive slippage test for use in state-space filtering. *Manuscripta Geodaetica*, 14(6):383–390.

UNCLOS (1982). United nations convention on the law of the sea. http://www.un.org/Depts/los.

Verhoef, H. M. E. (1997). *Geodetische Deformatie Analyse*. Collegedictaat TU Delft, Delft. in Dutch.

West, M. and Harrison, J. (1997). *Bayesian forecasting and dynamic models*. Springer, New York.

Wüst, J. C. (2003). Ontwikkeling productieproces voor trendanalyse vaargeuldiepte. Technical reports. Rijkswaterstaat, Directie Noordzee. in Dutch.

Wüst, J. C. (2004). Data-driven probabilistic predictions of bathymetry. In *Proceedings of International Workshop on Marine Sand Wave and River Dune Dynamics II, 2004*, Enschede.

# Appendix A

# Kriging

In this appendix the theory of the Kriging interpolation procedure is described. This method is used by the Hydrographic Service to interpolate the single beam surveys to a regular grid. In general, the basic model for interpolation is:

$$\hat{z} = \sum_{i=1}^{N} w_i \cdot z_i, \tag{A.1}$$

where $\hat{z}$ is the unknown point, $w_i$ are the weights and $z_i$ are the data points. In case of Kriging the weights are determined by means of a covariance function. Many different varieties of Kriging can be distinguished, the most important are [Olea, 1999]:

- *Simple Kriging*: the mean of the data is known and constant throughout the study area (no spatial trend exists).
- *Ordinary Kriging*: the mean of the data is unknown, but constant over a local neighbourhood.
- *Universal Kriging*: the mean of the data is unknown and may vary over a local neighbourhood. A trend can be modelled as a combination of analytical functions.

In case of sea floor data, universal Kriging sounds the most applicable, because the mean depth is unknown and there may be a spatial trend in the data. The disadvantage of universal Kriging however is, that it is not possible to calculate the covariance function. That is because one of the constraints for calculating the covariance function is that the area is stationary, which does not hold anymore in case of a trend. Therefore, the trend has to be removed first and after that universal Kriging can be applied to the original data, or ordinary Kriging with the residual data. In this study ordinary Kriging is chosen, because it is somewhat easier than universal Kriging.

In case of ordinary Kriging the weights are calculated as [Isaaks and Srivastava, 1989]:

$$\sum_{j=1}^{n} w_j C_{ij} + \mu = C_{i0} \qquad \text{for } i = 1, .., n, \tag{A.2}$$

$$\sum_{j=1}^{n} w_j = 1. \tag{A.3}$$

In these equations, $\mu$ is the Lagrange multiplier and $C_{i,j}$ is a covariance matrix. The Lagrange multiplier can be interpreted as a scale factor to the weights, because of the constraint that the sum of the weights must equal 1. Equation A.2 and A.3 can together be written in matrix notation:

$$\begin{bmatrix} C_{11} & \dots & C_{1N} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ C_{N1} & \dots & C_{NN} & 1 \\ 1 & \dots & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} w_1 \\ \vdots \\ w_N \\ \mu \end{bmatrix} = \begin{bmatrix} C_{10} \\ \vdots \\ C_{N0} \\ 1 \end{bmatrix}, \tag{A.4}$$

or in short matrix notation:

$$C \cdot w = D. \tag{A.5}$$

To solve for the weights, Equation A.5 can be rewritten as $w = C^{-1}D$, because $C$ has to be square and invertible. The matrices $C$ and $D$ can be calculated from the covariance function: $C$ is the matrix with covariances between the data points, $D$ is the vector with covariances between the unknown point and the data points.

The estimation variance of the unknown point $\hat{v}$, using ordinary Kriging, is derived in [Isaaks and Srivastava, 1989] as:

$$\sigma^2_{\hat{v}} = C(0) - w \cdot D. \tag{A.6}$$

The main advantage of Kriging is that it is based on statistics and gives the variance of the interpolated points. The Kriging method is a so called *Best Linear Unbiased Estimator (BLUE)*. It is *linear*, because the estimates are linear combinations of the available data, see Equation A.1. The second property is *unbiasedness*. This is defined as, see [Teunissen, 2000a]:

$$E\{\underline{\hat{x}}\} = x, \quad \forall\, x. \tag{A.7}$$

This means that the expectation of an unknown parameter equals the parameter itself. From Equation A.1 it can easily be verified that this property holds, since the function is assumed to be stationary, i.e.:

$$E\{\hat{z}\} = E\{z_i\} \quad \text{and} \quad \sum_{j=1}^{n} w_j = 1. \tag{A.8}$$

The last property is *best*, which means that it aims at minimising the variance. In [Isaaks and Srivastava, 1989] it is proved that the error variance in Equation A.6 is indeed minimal. This proves that Kriging is a *BLUE*. The disadvantage of Kriging however is that it is a quite complex method and requires a lot of computational effort, because the covariances between all interpolation and data points have to be calculated. Therefore, points further away than a certain distance limit are often ignored.

To illustrate the Kriging procedure, the dataset of critical area B is shown. This area does not have a clear sand wave pattern but is not isotropic either. In Figure A.1, the 2003 survey of critical area B is illustrated. The left figure shows the $x$ and $y$ coordinates in the UTM-31 ETRS89 frame. The surveyed tracks are clearly visible. The distance between the tracks is 50 m, so the size of this area is about 700×700 m. In the right figure, a three dimensional point cloud of the surveyed points is shown. Note that the axes are not the same scale, the depth value is exaggerated a hundred times. The depths are given with respect to the *Mean Low Lower Water Spring (MLLWS)*, which is defined as the average of the lower low water heights over a certain period [De Jong et al., 2002]. The average depth is about 37 m, but varies between about 30 m and 40 m.

From the plot of the tracks, it is clear that the data points are not regularly spaced, so the data have to be interpolated (Kriged) to a regular grid. The first step in the Kriging procedure is to determine the covariance function. As was stated before, the sea floor is usually not a stationary function but may show a trend. So before calculating the covariance function, the trend is removed. This is done by fitting a plane through the data and proceeding with the residuals.

First, the covariance function has to be estimated. In this study two different covariance models are used: first a *Gaussian model*, which is defined as [Chilès and Delfiner, 1999]:

$$C(s) = b \times exp\left\{-\left(\frac{s}{a}\right)^2\right\}, \tag{A.9}$$

where $s$ is the distance, $a$ the range and $b$ the sill, see Figure 3.2.

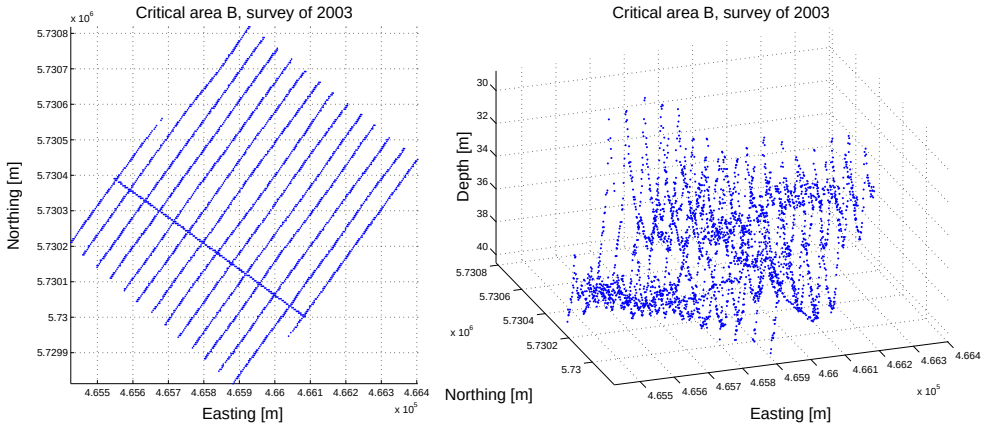



**Figure A.1:** The data of the survey of area B, performed in 2003. On the left side the tracks are visualised, on the right side a three dimensional plot of the depth values is shown. The $x$ and $y$ coordinates are in UTM-31 ETRS89, the depths are given w.r.t. the Mean Lower Low Water Spring (MLLWS).

The second model is a so called a *hole-effect model*. This is a sine like function and is used in case of sand waves. It is defined as [Chilès and Delfiner, 1999]:

$$C(s) = b \cdot exp\left\{-\frac{s}{a}\right\} \cos\left(\frac{s}{\lambda} + d\right), \tag{A.10}$$

with $b$ the sill, $a$ the range, $\lambda$ the wavelength and $d$ an additional scaling parameter.

These covariance functions are used to obtain the elements in the matrix $C$ of Equation A.5. The question is now whether these covariance function lead to a stable and unique solution of the system A.5. Since the matrix $C$ is square the existence and uniqueness of the solution are ensured when the matrix $C$ is *positive-definite* [Strang, 1988]. One of the tests for positive-definiteness reads:

$$\theta^* C\theta > 0, \tag{A.11}$$

for any vector $\theta$ with at least one nonzero element. Because of the exponential the Gaussian model always leads to positive values in the $C$-matrix. Therefore the test A.11 holds for this model and this is a suitable covariance model. The hole effect model is somewhat more difficult. Since there is a cosine term, the covariance values can become negative as well. To ensure that the $C$-matrix remains positive-definite, only the first part of the model, where the covariance is positive, is used. Therefore, points that are further apart than a certain limit are not used in the interpolation.

As was mentioned before, the depth values are interpolated using ordinary Kriging. In this case the original depths are interpolated to a 60×60 meter grid, see Figure A.2. Only at grid points within a certain distance to the original points, an interpolated value is calculated. In this case, the distance threshold is defined as two times the distance between the tracks. This seems reasonable because data points that are far away get a low covariance value and thus a very low weight. Therefore their contribution to the interpolation is negligible and restriction of the number of data points that have to be taken into account considerably speeds up the interpolation process. Besides, using too many data points in the interpolation may cause oversmoothing. Here a grid size of 60×60 meter is chosen. Of course it is possible to create a denser grid, but again at the cost of the processing time, so a 60×60 meter grid seems to be good enough to represent this data.

In Figure A.3 the result of the Kriging process is shown. The interpolated depths and the standard deviations of the interpolated points are given. The standard deviations are a combination of the measurement accuracy and the interpolation accuracy. The accuracy is getting worse on the edges of the grid. This is because at the edges less data points are available to contribute to the interpolation.

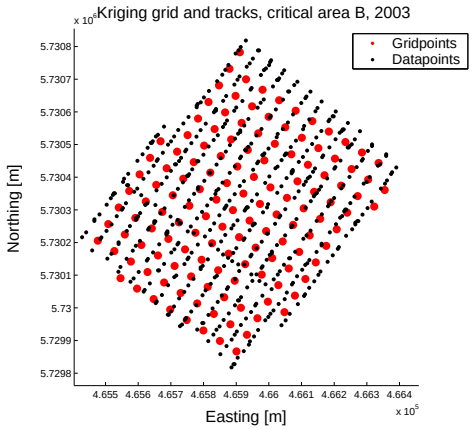



**Figure A.2:** Interpolation of the original data to a regular grid. The small black dots are the original measurements, they are interpolated to a 60×60 meter grid, represented by the large dots.

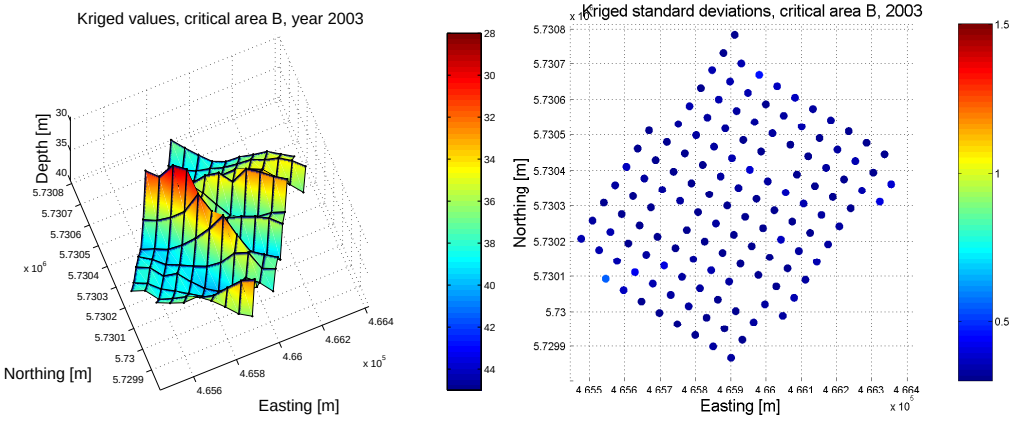



**Figure A.3:** The results of Kriging of area B. On the left side the interpolated depths, on the right side the standard deviations of the interpolated points, in m. The $x$- and $y$ coordinates are given in UTM-31 ETRS89, the depths are given w.r.t. the Mean Low Lower Water Spring (MLLWS).

# Appendix B

# Example of the deformation analysis method

In Chapter 3 the deformation analysis method of the Hydrographic Service was described. To gain more insight in the method a small example is elaborated in this appendix. Therefore the simulated dataset of Chapter 5 is used, with one outlying survey, as described in Section 5.2. This dataset looks as follows:

$$\underline{y}_{2001}: \begin{bmatrix} 29.9972 & 29.9037 & 30.0584 & 29.9811 \\ 30.0214 & 29.9913 & 30.0231 & 29.9353 \\ 29.9462 & 30.0524 & 29.9889 & 30.0476 \\ 30.0900 & 30.0783 & 30.0643 & 30.0844 \end{bmatrix}$$

$$\underline{y}_{2002}: \begin{bmatrix} 30.0787 & 29.9020 & 30.0208 & 30.0494 \\ 29.9821 & 30.0626 & 29.9397 & 29.9031 \\ 30.0834 & 29.9706 & 29.9406 & 29.9398 \\ 30.0871 & 29.9116 & 29.9278 & 29.9544 \end{bmatrix}$$

$$\underline{y}_{2003}: \begin{bmatrix} 29.0787 & 28.9020 & 29.0208 & 29.0494 \\ 28.9821 & 29.0626 & 28.9397 & 28.9031 \\ 29.0834 & 28.9706 & 28.9406 & 28.9398 \\ 29.0871 & 28.9116 & 28.9278 & 28.9544 \end{bmatrix}$$

$$\underline{y}_{2004}: \begin{bmatrix} 30.0364 & 30.0396 & 30.0187 & 30.0290 \\ 29.9387 & 29.9302 & 30.0707 & 30.0643 \\ 29.9379 & 30.0083 & 30.0720 & 30.0800 \\ 29.9609 & 29.9606 & 29.9757 & 29.9993 \end{bmatrix}.$$

It is assumed that all points have an equal standard deviation of 20 cm. First an estimate of the sea floor is calculated with the A-model, see Paragraph 3.2.1. For a single epoch, the parameters and observations are the same, there is no redundancy. When using different epochs an adjustment can be carried out. For every point $p$ the A-model reads:

$$E\{\begin{bmatrix} \underline{y}_p[2001] \\ \underline{y}_p[2002] \\ \underline{y}_p[2003] \\ \underline{y}_p[2004] \end{bmatrix}\} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} d_p, \tag{B.1}$$

with $\underline{y}$ the (vector of) observations and $d$ the (vector of) unknown parameters, in this case the depth. This model is the same for every point, because the points are treated independently. Therefore the number of observations is 4 and the number of unknowns is 1, resulting in a redundancy of 3. So, for

example for point (1,1) the A-model is:

$$E\{\begin{bmatrix} 30.0900 \\ 30.0871 \\ 29.0871 \\ 29.9609 \end{bmatrix}\} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} d_{1,1}. \tag{B.2}$$

In general, the estimator and its variance are calculated as:

$$\underline{\hat{x}} = (A^* Q_y^{-1} A)^{-1} A^* Q_y^{-1} \underline{y} \tag{B.3}$$

$$Q_{\hat{x}} = (A^* Q_y^{-1} A)^{-1}. \tag{B.4}$$

In this simple example the vector of unknown $\underline{\hat{x}}$ only consists of the depth and all measurements have equal standard deviations. Therefore, the estimator is simply the average of the four epochs and the precision of the estimator is the precision of the measurements divided by four:

$$\underline{\hat{x}}_{1,1} = 29.8063 \quad ; \quad Q_{\hat{x}} = 0.01. \tag{B.5}$$

Now this estimator can be used to detect deformation. Therefore some alternative hypotheses are formulated by extending the A-model with some extra parameters, see Paragraph 3.2.4.

1. For an outlying value, the model extensions are:

$$c_y = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \text{ and } \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \tag{B.6}$$

   for an outlier in survey 1,2,3 or 4 respectively.

2. The model extension for general deformation is:

$$C_y = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{B.7}$$

3. Finally, the model extension for a trend is:

$$c_y = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \end{bmatrix}. \tag{B.8}$$

Note that when the model extension is a vector, like in case 1 and 3, it is denoted by $c_y$, whereas in the second case it is a matrix, denoted by $C_y$.

Now it has to be tested whether the original model or an extended model fits the data better. This is done by comparing a test quantity with a critical value. The test quantities are calculated on basis of the residuals:

$$\underline{\hat{e}}_{1,1} = \underline{y}_{1,1} - A\underline{\hat{x}}_{1,1} = \begin{bmatrix} 30.0900 \\ 30.0871 \\ 29.0871 \\ 29.9609 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \cdot 29.8063 = \begin{bmatrix} 0.2837 \\ 0.2808 \\ -0.7192 \\ -0.1546 \end{bmatrix}, \tag{B.9}$$

| Type of deformation | | $\mathbf{T_q}$ | $\mathbf{k}_\alpha$ | $\mathbf{T_q/k}_\alpha$ |
|---|---|---|---|---|
| Outlying value in survey: | 1 | 2.6836 | 6.63 | 0.4048 |
| | 2 | 2.6285 | 6.63 | 0.3965 |
| | 3 | 17.2412 | 6.63 | 2.6005 |
| | 4 | 0.7971 | 6.63 | 0.1202 |
| General deformation | | 17.5128 | 7.81 | 2.2424 |
| Trend | | 2.4058 | 2.71 | 0.8877 |

**Table B.1:** The test quantities, critical values and normalised test quantities for the point test for point (1,1) of the example dataset, in the first iteration.

and the residual variance is:

$$Q_{\hat{e}} = Q_y - AQ_{\hat{x}}A^* = \begin{bmatrix} 0.03 & 0 & 0 & 0 \\ 0 & 0.03 & 0 & 0 \\ 0 & 0 & 0.03 & 0 \\ 0 & 0 & 0 & 0.03 \end{bmatrix}. \tag{B.10}$$

So with the residuals $\underline{\hat{e}}$, the residual variance $Q_{\hat{e}}$ and the model extensions $c_y$ or $C_y$, the test quantities are calculated as:

$$\underline{T}_q = \underline{\hat{e}}^* Q_y^{-1} C_y (C_y^* Q_y^{-1} Q_{\hat{e}} Q_y^{-1} C_y)^{-1} C_y^* Q_y^{-1} \underline{\hat{e}}. \tag{B.11}$$

These test quantities now have to be compared to a critical value in order to determine whether the differences are significant. The critical value is derived from the $\chi^2$ distribution and is dependent on the level of significance ($\alpha$) and the column dimension of the model extension ($q$). In this case the $\alpha$'s are 1%, 5% and 10% for an outlier, instability and a trend respectively and the $q$'s are 1, 3 and 1 respectively. In Table B.1 the test statistics for this example are presented.

An alternative hypothesis is accepted when the normalised test quantity $T_q/k_\alpha$ is larger than 1. If more than one alternative hypothesis is suitable, the one with the highest normalised test quantity is accepted. Therefore, in this case an outlying value in survey 3 is the most probable model extension. The $A$-matrix is extended with a $[0\ 0\ 1\ 0]^*$ column. The vector of unknowns now consists of two terms: the depth and the estimated deviation in survey 3. By an adjustment with the extended model this is calculated as:

$$\underline{\hat{x}}_{1,1} = \begin{bmatrix} 30.0460 \\ -0.9589 \end{bmatrix}. \tag{B.12}$$

So, the estimated deviation in the third survey is -0.9589 m. After this extension, the residuals are calculated again:

$$\underline{\hat{e}}_{1,1} = \underline{y}_{1,1} - A\underline{\hat{x}}_{1,1} = \begin{bmatrix} 30.0900 \\ 30.0871 \\ 29.0871 \\ 29.9609 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 1 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 30.0460 \\ -0.9589 \end{bmatrix} = \begin{bmatrix} 0.0440 \\ 0.0411 \\ 0 \\ -0.0851 \end{bmatrix}, \tag{B.13}$$

and the residual variance is:

$$Q_{\hat{e}} = Q_y - AQ_{\hat{x}}A^* = \begin{bmatrix} 0.0267 & 0 & 0 & 0 \\ 0 & 0.0267 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.0267 \end{bmatrix}. \tag{B.14}$$

Because the deviation in the third survey is calculated without redundancy, the variance of the adjusted depth in this survey is the same as the measurement accuracy. Therefore, the residual variance is zero for this survey. With these residuals, the test quantities are calculated again. The test quantities for the second iteration are presented in Table B.2. The residuals are not tested for a deviation in the third

| Type of deformation | $\mathbf{T_q}$ | $\mathbf{k}_\alpha$ | $\mathbf{T_q/k}_\alpha$ |
|---|---|---|---|
| Outlying value in survey: 1 | 0.0726 | 6.63 | 0.0110 |
| 2 | 0.0633 | 6.63 | 0.0095 |
| 4 | 0.2715 | 6.63 | 0.0410 |
| Trend | 0.2458 | 2.71 | 0.0907 |

**Table B.2:** The test quantities, critical values and normalised test quantities for the point test for point (1,1) of the example dataset, in the second iteration.

survey, because this is already accepted, and for general deformation, because this is not possible anymore.

It can be seen that no alternative hypothesis is accepted in the second iteration, so the model is accepted with an outlying value in survey 3.

In this case the predicted values are the mean of the 4 surveys, including the outlying values:

$$\underline{\hat{x}}_{pred} = \begin{bmatrix} 29.7481 & 29.7518 & 29.7014 & 29.7772 \\ 29.7239 & 29.7355 & 29.7433 & 29.7797 \\ 29.7155 & 29.7505 & 29.7617 & 29.6868 \\ 29.8063 & 29.7627 & 29.7310 & 29.7978 \end{bmatrix}. \tag{B.15}$$

The variances however are a lot higher than the measurement accuracy:

$$Q_{\hat{x}_{pred}} = \begin{bmatrix} 0.2548 & 0.2756 & 0.2674 & 0.2576 \\ 0.2643 & 0.2486 & 0.2815 & 0.2799 \\ 0.2424 & 0.2740 & 0.2788 & 0.2842 \\ 0.2549 & 0.2816 & 0.2791 & 0.2784 \end{bmatrix}. \tag{B.16}$$

Besides the point test, an area test is performed as well. In this case, the A-model is defined as in Paragraph 3.2.3:

$$E\{\begin{bmatrix} \underline{y}_{1,1}[2001] \\ \vdots \\ \underline{y}_{1,1}[2004] \\ \underline{y}_{2,1}[2001] \\ \vdots \\ \vdots \\ \underline{y}_{4,4}[2004] \end{bmatrix}\} = \begin{bmatrix} 1 & x_{1,1} & y_{1,1} \\ \vdots & \vdots & \vdots \\ 1 & x_{1,1} & y_{1,1} \\ 1 & x_{2,1} & y_{2,1} \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \\ 1 & x_{4,4} & y_{4,4} \end{bmatrix} \begin{bmatrix} \overline{d} \\ \overline{\alpha}_x \\ \overline{\alpha}_y \end{bmatrix}, \tag{B.17}$$

Here $x$ and $y$ are the coordinates of the points w.r.t the origin. By defining the origin as the centre of the dataset, the coordinates of the first point are (-60,-60). So for this example the model is:

$$E\{\begin{bmatrix} 30.0900 \\ \vdots \\ 29.9609 \\ 30.0783 \\ \vdots \\ \vdots \\ 30.0290 \end{bmatrix}\} = \begin{bmatrix} 1 & -60 & -60 \\ \vdots & \vdots & \vdots \\ 1 & -60 & -60 \\ 1 & -20 & -60 \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \\ 1 & 60 & 60 \end{bmatrix} \begin{bmatrix} \overline{d} \\ \alpha_x \\ \alpha_y \end{bmatrix}, \tag{B.18}$$

Now the number of observations is 4×16=64 and the number of unknowns is 3, so there is a redundancy of 61. The estimator equals:

$$\underline{\hat{x}} = \begin{bmatrix} 29.7483 \\ -0.0002 \\ 0 \end{bmatrix} \quad ; \quad Q_{\hat{x}} = \begin{bmatrix} 0.6250 & 0 & \\ 0 & 0.0003 & 0 \\ 0 & 0 & 0.0003 \end{bmatrix}. \tag{B.19}$$

| | Iteration 1 | | | Iteration 2 | | |
|---|---|---|---|---|---|---|
| **Type of deformation** | $\underline{\mathbf{T}}_\mathbf{q}$ | $\mathbf{k}_\alpha$ | $\underline{\mathbf{T}}_\mathbf{q}/\mathbf{k}_\alpha$ | $\underline{\mathbf{T}}_\mathbf{q}$ | $\mathbf{k}_\alpha$ | $\underline{\mathbf{T}}_\mathbf{q}/\mathbf{k}_\alpha$ |
| Outlying plane in survey: 1 | 39.0846 | 11.34 | 3.4466 | 0.7295 | 11.34 | 0.0643 |
| 2 | 30.1567 | 11.34 | 2.6593 | 0.9717 | 11.34 | 0.0857 |
| 3 | 311.4731 | 11.34 | 27.4668 | | | |
| 4 | 36.7705 | 11.34 | 3.2425 | 0.7598 | 11.34 | 0.0670 |
| General deformation | 313.1137 | 16.92 | 18.5055 | | | |
| Trend | 21.6616 | 6.25 | 3.4659 | 0.6863 | 6.25 | 0.1098 |

**Table B.3:** The test quantities, critical values and normalised test quantities for the area test of the example dataset, in first and second iteration

The model extensions for the area test are formulated as in Paragraph 3.2.4:

1. In case of an outlying survey:

$$C_y = \begin{bmatrix} 1 & -60 & -60 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & -20 & -20 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ \vdots & \vdots & \vdots \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 1 & -60 & -60 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & -20 & -20 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ \vdots & \vdots & \vdots \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & -60 & -60 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & -20 & -20 \\ 0 & 0 & 0 \\ \vdots & \vdots & \vdots \end{bmatrix} \text{ and } \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & -60 & -60 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & -20 & -20 \\ \vdots & \vdots & \vdots \end{bmatrix}. \quad \text{(B.20)}$$

2. In case of general deformation:

$$C_y = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -60 & -60 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -60 & -60 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -60 & -60 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -20 & -60 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -20 & -60 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -20 & -60 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}. \quad \text{(B.21)}$$

3. In case of a trend:

$$C_y = \begin{bmatrix} -1.5 & -90 & -90 \\ -0.5 & -30 & -30 \\ 0.5 & 30 & 30 \\ 1.5 & 90 & 90 \\ -1.5 & -90 & -90 \\ -0.5 & -30 & -30 \\ 0.5 & 30 & 30 \\ 1.5 & 90 & 90 \\ \vdots & \vdots & \vdots \end{bmatrix}. \quad \text{(B.22)}$$

With the model extensions and the residuals, calculated with Equations B.9 and B.10 the test quantities are calculated, see Table B.3. These test quantities are somewhat larger than the test quantities for the point test, but the critical values are larger too. So with the area test, the outlying survey is found as well.

# Appendix C

# Example of the Kalman filtering method

The Kalman filtering method of the North Sea Directorate was described in Chapter 4. In this appendix, a small example is elaborated to gain more insight in the method. Therefore, the simulated dataset of Chapter 5 is used. This data set looks as follows:

$$\underline{y}_{2001}: \begin{bmatrix} 30.0900 & 30.0783 & 30.0643 & 30.0844 \\ 29.9462 & 30.0524 & 29.9889 & 30.0476 \\ 30.0214 & 29.9913 & 30.0231 & 29.9353 \\ 29.9972 & 29.9037 & 30.0584 & 29.9811 \end{bmatrix}$$

$$\underline{y}_{2002}: \begin{bmatrix} 30.0871 & 29.9116 & 29.9278 & 29.9544 \\ 30.0834 & 29.9706 & 29.9406 & 29.9398 \\ 29.9821 & 30.0626 & 29.9397 & 29.9031 \\ 30.0787 & 29.9020 & 30.0208 & 30.0494 \end{bmatrix}$$

$$\underline{y}_{2003}: \begin{bmatrix} 29.9890 & 30.0692 & 30.0676 & 30.0664 \\ 30.0864 & 30.0050 & 29.9039 & 30.0006 \\ 29.9932 & 29.9405 & 30.0363 & 30.0419 \\ 29.9837 & 30.0344 & 29.9759 & 29.9858 \end{bmatrix}$$

$$\underline{y}_{2004}: \begin{bmatrix} 29.9609 & 29.9606 & 29.9757 & 29.9993 \\ 29.9379 & 30.0083 & 30.0720 & 30.0800 \\ 29.9387 & 29.9302 & 30.0707 & 30.0643 \\ 30.0364 & 30.0396 & 30.0187 & 30.0290 \end{bmatrix}.$$

No dynamic behaviour is present in this area, all points are static. This area consists of 4×4 equally spaced data points. It is assumed that all points have an equal standard deviation of 20 cm and the grid size is 40×40 m. In this case, the model grid, with the interpolation kernels, is chosen as a grid with 3×3 points in between the data points. The Kalman equations are only calculated in the model points.

The process starts with initialising the state vector and covariance matrix. The initial state vector can be seen as the depth prior to the measurements. Because it is assumed that the first measurements are from 2001, the initial state vector is the depth in 2000. This vector consists of 18 elements, for every support point a depth value and a trend value. Because it is assumed that no measurements before 2001 are available, the initial depth is chosen as the mean depth of the first survey. The initial depth variance is estimated from the measurements:

$$\hat{\underline{\sigma}}^2 = \frac{1}{N-1}\sum_{i=1}^{N}(d_i - \overline{d})^2, \tag{C.1}$$

with $N$ the number of points and $\overline{d}$ the mean depth. The initial variance for trends is 0.1 m$^2$, see also Section 4.2. Therefore, the initial state vector and initial variance matrix for a single support point are:

$$\underline{x}_{00|00} = \begin{bmatrix} 30.0165 \\ 0 \end{bmatrix} \quad ; \quad Q_{x_{00|00}} = \begin{bmatrix} 0.0032 & 0 \\ 0 & 0.1 \end{bmatrix}. \tag{C.2}$$

Now the measurement equations can be formulated as in Section 4.2:

$$\begin{bmatrix} \underline{y}_1 \\ \underline{y}_2 \\ \vdots \\ \underline{y}_{16} \end{bmatrix} = \begin{bmatrix} w_{11} & \dots & w_{19} & 0 & \dots & 0 \\ w_{21} & \dots & w_{29} & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ w_{91} & \dots & w_{99} & 0 & \dots & 0 \end{bmatrix} \begin{bmatrix} d_1 \\ \vdots \\ d_9 \\ \frac{\Delta d_1}{\Delta t} \\ \vdots \\ \frac{\Delta d9}{\Delta t} \end{bmatrix}, \tag{C.3}$$

Therefore, the interpolation weights are calculated using Equation 4.10. The weights are only dependent on the distance between the data points and the support points. For example, the weights for the first data point, w.r.t. the 9 support points are:

$$\begin{array}{lll} w_{11} = 0.4976 & w_{12} = 0.1830 & w_{13} = 0.0248 \\ w_{14} = 0.1830 & w_{15} = 0.0673 & w_{16} = 0.0091 \\ w_{17} = 0.0248 & w_{18} = 0.0091 & w_{19} = 0.0012. \end{array} \tag{C.4}$$

The first step now is the time update. The transition matrix is defined as in Section 4.2:

$$\Phi = \begin{bmatrix} I_{9\times 9} & I_{9\times 9} \\ 0_{9\times 9} & I_{9\times 9} \end{bmatrix}, \tag{C.5}$$

with $I$ the identity matrix. The model error matrix $Q_m$ is calculated using the discount method with discount factors 0.93 (see Eq. 4.8):

$$Q_{m_{00}} = \begin{bmatrix} \frac{1-0.93}{0.93} & 0 \\ 0 & \frac{1-0.93}{0.93} \end{bmatrix} Q_{x_{00|00}} = \begin{bmatrix} 0.00024 & 0 \\ 0 & 0.0075 \end{bmatrix}. \tag{C.6}$$

The time update is defined as:

$$\begin{aligned} \underline{\hat{x}}_{k|k-1} &= \Phi_{k|k-1} \underline{\hat{x}}_{k-1|k-1} \\ Q_{\hat{x}_{k|k-1}} &= \Phi_{k|k-1} Q_{\hat{x}_{k-1|k-1}} \Phi^*_{k|k-1} + Q_{m_k}, \end{aligned} \tag{C.7}$$

so for this example the first time update is:

$$\underline{\hat{x}}_{01|00} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \underline{\hat{x}}_{00|00} = \begin{bmatrix} 30.0165 \\ 0 \end{bmatrix} \tag{C.8}$$

$$Q_{\hat{x}_{01|00}} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} Q_{\hat{x}_{00|00}} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} + Q_{m_{0|0}} = \begin{bmatrix} 0.1034 & 0.1 \\ 0.1 & 0.1075 \end{bmatrix}. \tag{C.9}$$

Now, a prediction for the epoch 2001 is calculated, based on the (initial) values for 2000. In case measurements of the year 2001 are available, they can be used to correct those predictions. This step is called the measurement update and consists of the equations:

$$\begin{aligned} \underline{v}_k &= \underline{y}_k - A_k \underline{\hat{x}}_{k|k-1} \\ Q_{v_k} &= Q_k + A_k Q_{\hat{x}_{k|k-1}} A^*_k \\ K_k &= Q_{\hat{x}_{k|k-1}} A^*_k Q^{-1}_{v_k} \\ \underline{\hat{x}}_{k|k} &= \underline{\hat{x}}_{k|k-1} + K_k \underline{v}_k \\ Q_{\hat{x}_{k|k}} &= [I - K_k A_k] Q_{\hat{x}_{k|k-1}}. \end{aligned} \tag{C.10}$$

The measurement update is based on the differences between the prediction and the measurements:

$$\underline{v}_{01} = \underline{y}_{01} - A\underline{\hat{x}}_{01|00}, \tag{C.11}$$

which is in this case the difference between the actual value at every location and the mean depth. But in the first epoch, the difference between the actual value and the initial value is just a result of the initialisation and not the result of a trend. Therefore the trend elements are set to zero again after the first epoch. With the adjusted parameters $\underline{\hat{x}}_{01|01}$ in the model locations, the depth at the data points can be calculated again using the same interpolation matrix $A$:

$$\underline{\hat{y}}_{01|01} = \begin{bmatrix} 30.0429 & 30.0454 & 30.0477 & 30.0488 \\ 30.0227 & 30.0251 & 30.0271 & 30.0277 \\ 30.0007 & 30.0030 & 30.0049 & 30.0053 \\ 29.9875 & 29.9899 & 29.9919 & 29.9926 \end{bmatrix}, \tag{C.12}$$

$$Q_{\hat{y}_{01|01}} = \begin{bmatrix} 0.0117 & 0.0066 & 0.0066 & 0.0117 \\ 0.0066 & 0.0037 & 0.0037 & 0.0066 \\ 0.0066 & 0.0037 & 0.0037 & 0.0066 \\ 0.0117 & 0.0066 & 0.0066 & 0.0117 \end{bmatrix}. \tag{C.13}$$

Note that the variance matrix is symmetric and the variances in the centre are smaller than those on the edges. This is a result of the smoothing effect of the interpolation. The parameters $\underline{\hat{x}}_{01|01}$ are then used to make a prediction for 2002, $\underline{\hat{x}}_{02|01}$ and the process starts over again.

So after 4 years, the depth values and variances are:

$$\underline{\hat{y}}_{04|04} = \begin{bmatrix} 29.9877 & 29.9812 & 29.9937 & 30.0189 \\ 29.9769 & 29.9803 & 30.0031 & 30.0340 \\ 29.9869 & 29.9920 & 30.0140 & 30.0421 \\ 30.0128 & 30.0105 & 30.0211 & 30.0393 \end{bmatrix}, \tag{C.14}$$

$$Q_{\hat{y}_{04|04}} = \begin{bmatrix} 0.0177 & 0.0103 & 0.0103 & 0.0177 \\ 0.0103 & 0.0060 & 0.0060 & 0.0103 \\ 0.0103 & 0.0060 & 0.0060 & 0.0103 \\ 0.0177 & 0.0103 & 0.0103 & 0.0177 \end{bmatrix}. \tag{C.15}$$