Quantifying parameter uncertainty in predictions of coastal mega-nourishments

A case study on the Sand Engine at the Dutch coast

J.C. Christiaanse

MSc Thesis

TU Delft

Quantifying parameter uncertainty in predictions of coastal mega-nourishments

by

J.C. Christiaanse

Delft University of Technology

Faculty of Civil Engineering and Geosciences

Department of Hydraulic Engineering

Student number:	4372719		
Project duration:	May 1, 2020 – January 29, 2021		
Graduation committee:	Dr. ir. M.A. de Schipper,	TU Delft,	Chair
	Dr. ir. A.P. Luijendijk,	TU Delft & Deltares,	Supervisor
	Ir. J. Kroon,	TU Delft & Svašek Hydraulics	
	Dr. ir. R.C. Lanzafame,	TU Delft	
	Prof. dr. R. Ranasinghe,	IHE Delft, Deltares & Universiteit Twente	

An electronic version of this thesis is available at http://repository.tudelft.nl/

Cover image credit: Rijkswaterstaat/Joop van Houdt

TU Delft Delft University of Technology

Deltares Enabling Delta Life

Executive summary

Continuous sea level rise and growing environmental awareness have led to increasing implementation of nature-based solutions to counter coastal erosion. An example in the Netherlands is the Sand Engine—a mega-scale sand nourishment designed to feed the Dutch coast over a period of 20 years. For such designs to work, a good understanding of the governing natural processes is paramount. The behaviour of the sandy coast, however, is subject to natural variability in future weather and uncertainties in the interactions between sand dynamics and hydrodynamic forcing. To predict the evolution of coastal systems, engineers apply numerical models. Next to uncertainty due to variability in natural forcing, such models introduce a series of model-related uncertainties, which are often not consistently included in predictions. With limited knowledge of the magnitude of these uncertainties, the long-term strategy development and design of projects are impeded. Parameter uncertainty denotes our limited knowledge of the values of free model parameters and is an important source of overall model uncertainty. This study aims to quantify parameter uncertainty in process-based coastal area predictions by analysing uncertainty bounds for morphodynamic predictions of the Sand Engine.

The analysis of parameter uncertainty is carried out for a study period of 14 months, from August 2011 (directly after construction of the Sand Engine) until October 2012. Using advanced numerical acceleration techniques, a synthetic dataset of 1024 morphological Delft3D predictions is generated, each with an identical hindcast period but different model parameter settings. First, a sensitivity analysis (elementary effects method) is performed to find the most influential parameters on three morphological indicators: cumulative volume change, shoreline position, and bed level change. Based on the sensitivity analysis, five parameters are selected for an uncertainty analysis. Subsequently, parameter uncertainty is quantified by the Generalised Likelihood Uncertainty Estimation (GLUE). To increase the convergence speed of the samples, quasi-random Sobol sampling is applied. To validate the applied parameter ranges and assess model performance, the predictions are compared to observations. Using GLUE, uncertainty bounds for the morphological indicators and posterior likelihood distributions of the five parameters are derived from the 1024 model runs. Finally, through a simplified uncertainty comparison, the relative influence of parameter uncertainty and wave climate variability is examined.

A first key conclusion is that uncertainty in input parameters translates to significant uncertainty in predictions: cumulative erosion volumes for the peninsula show a spread of $1.3\ -0.4/\ +\ 0.7$ million m^3 after 14 months. Extended to bed level changes in 2D, the uncertainty bounds result in spatial uncertainty maps, which reveal that uncertainty is highest (5–6 m spread) in the northern area of the peninsula, where a sand spit develops. Most uncertainty ($\approx 90\%$ for volume changes) develops in the first seven months, during which uncertainty growth correlates strongly to periods of high morphological activity ($r > 0.95$ for volume changes). Further, posterior likelihood distributions of the parameters indicate that, of the five selected parameters based on the sensitivity analysis, only three (f_{sus}, γ, and d_{50}) significantly contribute to parameter induced uncertainty, while the other two (α_{rol} and θ_{sd}) are considerably less influential. This contrast in results is attributed to a resolution problem in the sensitivity analysis. Optimised parameter sets derived from the posterior likelihood distributions result in similar model skill as the best GLUE simulations ($BSS \approx 0.8$), implying that they may be close to the maximum achievable model skill within the examined parameter ranges. Finally, a simplified estimate of the variation due to wave climate variability is in the same order of magnitude as the parameter induced uncertainty bounds, implying that both uncertainty sources form significant contributions to overall prediction uncertainty.

The presented results have an impact on two key levels. First, they can be used to communicate and address uncertainty in predictions of coastal change. For example, the spatial uncertainty maps can let stakeholders understand the potential range of outcomes for a certain design. Second, the results, combined with the created dataset, can provide valuable information for future morphodynamic studies. This work contributes to justifying the need for stochastic simulations, which are expected to be increasingly used for many coastal engineering and management purposes in the years to come.

Preface

This thesis marks the end of my time as a student at Delft University of Technology and represents my final efforts to obtain the degree of Master of Science in Civil Engineering. Although difficult at times, the past nine months have been a highly interesting, educational and satisfying period of my life. Without the help of others, however, this thesis would not have been possible.

First of all, I would like to thank my academic supervisor, Dr. Matthieu de Schipper, for his enthusiasm, helpfulness, and feedback. His routine of challenging me to think beyond numbers and figures and put my work into a broader context helped me a lot in gaining new insights. I would like to thank Dr. Arjen Luijendijk, my daily supervisor, for the many meetings, feedback sessions, and for sharing his vast modelling experience, which has often inspired me to dig deeper. His frequent reminders to 'be pragmatic and move on' have eventually saved me a lot of time and taught me a more efficient working approach. Next, I would like to thank Ir. Anna Kroon, for her enthusiasm, the many meetings, and her extensive feedback on my draft reports. Our numerous discussions on the various results and further research ideas have inspired large parts of this thesis and I hope they don't end here. Finally, I would like to thank my remaining committee members, Dr. Robert Lanzafame for his positive feedback, fresh inputs during the progress meetings, and easygoing communication, and Prof. Rosh Ranasinghe for sharing his inspiring views and knowledge over the final two months. In addition, I would like to express my gratitude to Deltares for giving me access to company resources.

Furthermore, I would like to thank Ir. Paul van Riel for the inspirational discussions and feedback during the progress of writing my report. Also, thank you to my friends from Het Waterbouwdispuut for the many enjoyable moments over the course of the Master's, my roommates in Rotterdam, who have been part of my entire graduation work at home, and all other friends who supported me during the past nine months.

Last but not least, I would like to thank my family, especially my parents, for their unwavering support and continuous encouragement throughout my entire time as a student. This accomplishment would not have been possible without them.

J.C. Christiaanse
Feldis, Switzerland, January 2021

Contents

List of Figures

List of Tables

Glossary

2DH	2-dimensional horizontal
BF	Brute force
BFFC	Brute force filtered & compressed
BFM	Brute force merged
BSS	Brier Skill Score
BwN	Building with Nature
CDF	Cumulative distribution function
CI	Confidence interval
EE	Elementary effect
GLUE	Generalised Likelihood Uncertainty Estimation
HLS	Highest likelihood simulation
IPCC	International Panel on Climate Change
LDS	Low discrepancy sequence
LH	Latin Hypercube
LST	Longshore sediment transport
MAE	Mean absolute error
MC	Monte Carlo
MSE	Mean squared error
MSL	Mean sea level
OAT	One at A Time
OPS	Optimal parameter set
PBM	Process based model
PDF	Probability density function
SA	Sensitivity analysis
SE	Sand Engine
SLR	Sea level rise
UA	Uncertainty analysis

Nomenclature

α_{bn}	Transverse bed slope coefficient
α_{bs}	Streamwise bed slope coefficient
α_{rol}	Wave energy dissipation coefficient
β_{rol}	Mean roller slope under breaking waves
ΔX	Change in X
Δ	Increment value
γ	Wave breaker parameter
γ_{max}	Maximum breaker parameter
μ	Mean
μ^*	Mean of absolute values
ν_h	Horizontal eddy viscosity [m^2/s]
ρ	Density of sea water [1024 kg/m^3]
σ	Standard deviation
θ_{sd}	Dry cell erosion factor
$\vec{x}$	Vector x
A	Area [m^2]
BSS_Y	Brier Skill Score measured to output Y
C	Chézy roughness coefficient [$\sqrt{m}/s$]
CL	Combined likelihood
D_h	Horizontal eddy diffusivity [m^2/s]
d_{50}	Median grain diameter [μm]
E	Mean wave energy density per unit horizontal area [J/m^2]
EE	Elementary effect
f_s	Sampling resolution
$f_{bed,w}$	Scaling factor for wave-related bed load transport
f_{bed}	Scaling factor for current-related bed load transport
$f_{sus,w}$	Scaling factor for wave-related suspended sediment transport
f_{sus}	Scaling factor for current-related suspended sediment transport
Fac_{DSS}	Factor for initial suspended sediment diameter
g	Gravitational acceleration [9.81 m/s^2]

H	Wave height $[m]$
h	Water depth $[m]$
H_s	Significant wave height $[m]$
H_{m0}	Significant spectral wave height $[m]$
k	Number of dimensions of the parameter space
L	Likelihood
Med	Median
$morfac$	Morphodynamic acceleration coefficient
N	Sample size
n	Number of data points
N_L	Number of likelihood measures
$O(\ldots)$	Order
p	Number of parameter levels (EE method)
r	Pearson's correlation coefficient, number of EE trajectories
t	Time index
UI	Uncertainty indicator
V	Volume $[m^3]$
$W_{CI90,max}$	Maximum width of the 90% confidence interval
W_{CI90}	Width of the 90% confidence interval
X	Model input parameter
x, y	XY coordinates in local system
$X_{CI90,max}$	Position of the maximum width of the 90% confidence interval
$Y(X)$	Model output
z	Bed elevation $[m + MSL]$

1

Introduction

Chapter outline
Chapter 1 provides an introduction to the topic and lays the foundation of the research. It starts with a brief motivation in Section 1.1, which sets the context of the research topic, why it was chosen and why it is relevant to society. Then follows a more detailed introduction to uncertainty in numerical modelling in Section 1.2, discussing various types and sources of uncertainty. The research objective of this thesis is discussed in Section 1.3: the current state of research is assessed to identify relevant knowledge gaps and formulate an objective with corresponding research questions. Lastly, the research approach is presented in Section 1.4, along with some considerations regarding the scope of this thesis.

1.1. Motivation

Our planet faces ever more complex challenges through the effects of climate change and anthropogenic impacts, leading to various issues in the engineering world and beyond. Sea level rise (SLR) is one of the main consequences of climate change. In fact, the Intergovernmental Panel on Climate Change (IPCC) estimates that global mean sea level has risen by 19 cm during the 20^{th} century, and predicts at least twice as much for the 21^{st} century—probably even more (Figure 1.1) (Church et al., 2013). This is a major threat to society, especially considering that two-thirds of the world's largest cities are situated along its coasts (The Economist, 2019). Obviously, coastal areas are among the most affected by SLR. Nearly 60 years ago, Bruun (1962) already observed a significant link between SLR and coastal erosion. The observed long-term shoreline retreat rate for beaches along the US east coast was found to be around 150 times that of SLR (i.e. for every centimetre of SLR, the shoreline can be expected to retreat by around 1.5 m) (Leatherman et al., 2000). The Netherlands are no strangers to coastal erosion either—the Dutch coast has been eroding for centuries (van Rijn, 1997; Wijnberg, 2002). Since the 1990s this has been counteracted by actively nourishing the coast, through beach and shoreface nourishments, although the volume of sand required to keep up with erosion has been steadily increasing. Stive et al. (2013) estimate that the 12 million cubic metres per year required in 2013 may increase to 80 million cubic meters per year by the end of the 21^{st} century, assuming a high-end climate change scenario (Kabat et al., 2009). The size and frequency of traditional shoreface nourishments would need to be increased significantly to meet this demand, leading not only to wide, unattractive beaches (Stive et al., 2013), but also increase dependence on sand—the world's second-most consumed natural resource, behind water, which we are facing an increasing shortage of (Beiser, 2019; Luijendijk, 2019).

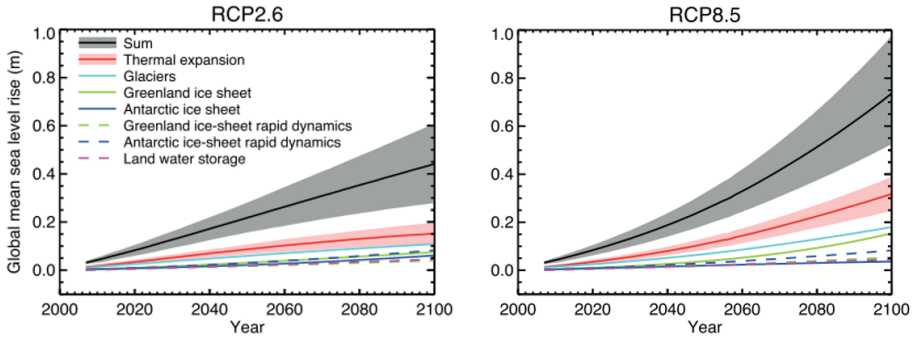

Figure 1.1: Overview of the lower (RCP2.6, left) and upper limit (RCP8.5, right) global SLR projections of the IPCC. The upper limit projects SLR of up to 1 m by the end of the 21^{st} century. Source: Church et al. (2013)

Along with growing environmental awareness and the resulting trend towards sustainable engineering practices, these considerations have led to a gradually increasing implementation of the *Building with Nature* (BwN) philosophy in engineering. This philosophy aims to design engineering solutions that combine the fulfilment of societal and structural requirements with natural development, by including natural processes within designs (de Vriend et al., 2015). An example in the Netherlands is the Sand Engine—a mega-scale nourishment designed to feed the Dutch coast over a period of 20 years. For such an approach to work, a good understanding of the governing processes is paramount. For example, predictions of beach width and volume change are crucial when assessing safety against flooding in many coastal regions worldwide. The behaviour of the sandy coast, however, is subject to natural variability in future weather and uncertainties in the interactions between sand dynamics and hydrodynamic forcing. Hence, making designs more dependent on natural processes also increases uncertainty compared with traditional hard structures (Kroon et al., 2020).

To predict the (long-term) behaviour of such solutions, engineers often use numerical models to simulate the development of the design in time. With the continuous advances in availability and efficiency of computational power, complex, process-based models are becoming increasingly popular for such purposes. Next to uncertainty arising from the variability in natural processes, these models introduce a series of model-related uncertainties. While natural variability in forcing conditions is regularly taken into account through probabilistic approaches, model uncertainty is often not consistently included in predictions. Yet, recent studies indicate that these uncertainties can contribute significantly to the overall prediction uncertainty, especially for simulations over large time-scales (e.g. Kroon et al., 2020; Ruessink, 2005; Simmons et al., 2017). With limited knowledge

of the magnitude of different uncertainties, the current long term strategy development and design of projects are impeded.

The main motivation behind this thesis is the fact that being able to better understand and quantify model uncertainty, will lead to more accurate and informed predictions (e.g. Plant & Holland, 2011). Consequently, the implementation of sustainable BwN designs is expected to improve. The presented study aims to contribute to this goal, by quantifying the effect of parameter uncertainty (an important source of model uncertainty) in process-based coastal area predictions.

1.2. Uncertainty in numerical modelling

Numerical models are used to predict how a natural system will change or develop over a simulated time span. Such predictions are based on assumptions, approximations, and simplifications of natural processes and, hence, accompanied by uncertainty. This section outlines the various types of uncertainty and how they arise. Parameter uncertainty is discussed in more detail as it is the main uncertainty source examined in this thesis.

1.2.1. Classification

Uncertainty in model predictions comes from various sources. In general, two main categories are considered (Figure 1.2): intrinsic and epistemic uncertainty (van Gelder, 2000; van Vuren, 2005). Natural processes such as ocean waves, tidal currents and weather conditions compose the forcing conditions to which our coastlines and beaches are exposed. The variability intrinsic to these processes is impossible to predict precisely, on a time scale exceeding hours to days, inevitably leading to uncertainty in model predictions. This is referred to as *intrinsic uncertainty* (Losada et al., 2011). It encompasses the randomness in the occurrence of natural processes and can manifest itself in space (e.g. turbulent fluctuations) and time (e.g. variations in the wave climate). Many studies have examined intrinsic uncertainty in the natural processes governing the coastal zone (e.g Callaghan et al., 2013; Vrijling et al., 1992). Within the hydraulic engineering world it is increasingly accounted for by the implementation of probabilistic approaches—substituting deterministic values of forcing parameters with probabilistic distributions (Jonkman et al., 2017).

Uncertainties that do not arise from the natural processes themselves, but are introduced by the limited ability of humans to approximate them, are categorised as *epistemic uncertainty*. They can be grouped into observation and model uncertainty (Kroon et al., 2020). Observation uncertainty includes errors in observations of natural processes and limitations of measuring instruments. Model uncertainty is divided into three subgroups. Parameter uncertainty, which denotes our limited knowledge of the values of so-called *free model parameters*, is probably the best known form of model uncertainty (Section 1.2.2). Model uncertainty can also arise from model inadequacy, which can be attributed to missing or simplified processes and aggregating complex, multidimensional physics (Blumberg & Georgas, 2008). The third subgroup of model uncertainty consists of numerical limitations. These include limitations in computational capacity, discretization techniques and applied numerical schemes.

1.2.2. Parameter uncertainty

Numerical models approximate natural processes through mathematical equations, that are solved in space and time. These equations include various parameters, which need to be specified in the model input. Parameter uncertainty stems from our limited knowledge of the values of these model parameters. Usually, model parameters represent calibration coefficients or physical quantities. Calibration coefficients are used to tune and calibrate the model to specific applications and often lack a physical basis, making it difficult to estimate their values (Ruessink, 2005). Parameters representing physical quantities are often site-specific, and can be difficult to measure, also resulting in uncertain estimations—grain size and bed friction are typical examples in coastal engineering (Berends et al., 2019; Kroon et al., 2020). Even parameters representing easily measurable quantities (e.g. wave height) can induce uncertainties due to measurement errors.

Engineers usually deal with parameter estimation during the model calibration—running the model for varying parameter settings to optimise the prediction compared to observation data. Continuous advances in numerical computing power have led to increasing popularity of complex process-based models (PBM) in morphological modelling. The advantage of PBM's is that they represent the underlying physics more accurately and provide detailed outputs on the governing processes in two or even three spatial dimensions (van der Wegen & Jaffe, 2013). On the other hand, they are often accompanied by simulation times in the order of several days or even

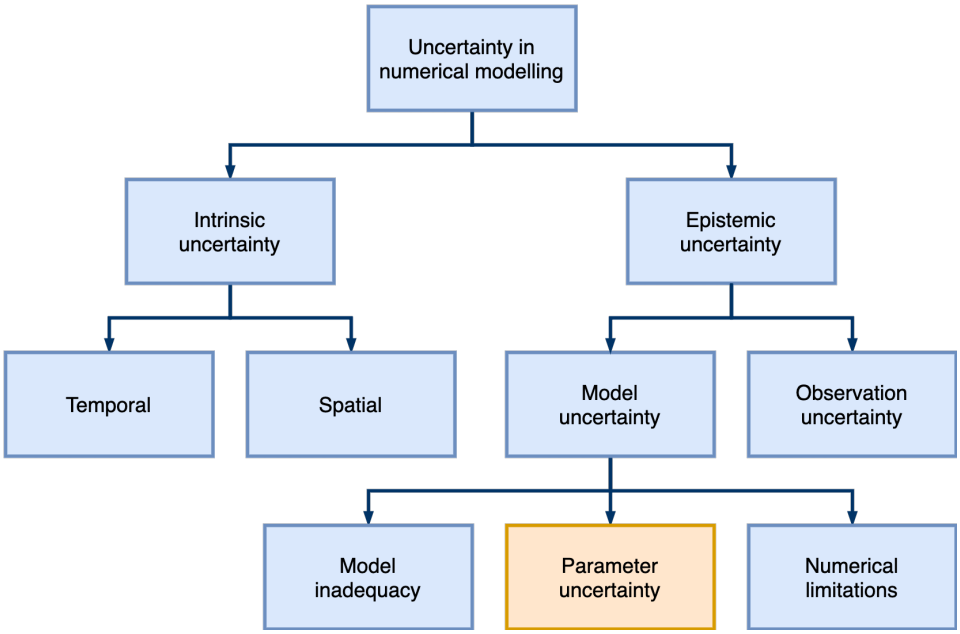

Figure 1.2: Classification of uncertainty in numerical modelling based on (van Gelder, 2000; Kroon et al., 2020; van Vuren, 2005). This study focuses on parameter uncertainty, which is highlighted in orange.

weeks (Luijendijk et al., 2019; Ranasinghe, 2020). Moreover, the large number of processes included in PBM's often introduces a wide range of input parameters. This makes it virtually impossible to explore a significant part of the parameter space during a traditional calibration process, especially since the parameters are usually not independent (i.e. the influence of one parameter on the model prediction depends on its own value, as well as the values of other parameters). The consequence is that the estimation of parameter values is strongly based on expert judgement and/or default values provided with the model. This results in a subjective, time-consuming process, which does not provide any information on whether the optimal parameter set has in fact been found. Especially for models with many parameters, the latter is highly unlikely (Oreskes et al., 1994; Ruessink, 2005).

1.2.3. Historical focus

In the past decades, research has focused on improving our ability to include intrinsic uncertainty (i.e. variability in natural forcing conditions) in model predictions. For example, instead of using a deterministic value of the significant wave height, variability in wave heights can be represented by extreme value distributions (Menéndez et al., 2009). The underlying assumption of the focus on natural variability, however, is that intrinsic uncertainty is more important than epistemic uncertainty. It has recently been shown that the validity of this assumption is, at a minimum, debatable (Kroon et al., 2020). The results of Callaghan et al. (2013), for example, show that even when applying a probabilistic framework to account for intrinsic uncertainty, storm erosion predictions are strongly influenced by model uncertainties. Arendt et al. (2012) caution that being able to quantify model uncertainty is paramount to reliable decision-making for engineering designs.

1.3. Research objective

In this section, the current state of research is used to identify relevant knowledge gaps, which are translated into a research objective for this thesis. The objective is then split in several research questions.

1.3.1. Academic setting

This study aims to build on recent modelling studies that have analysed the role of model uncertainty in coastal predictions in detail (Kroon et al., 2020; Ruessink, 2005; Simmons et al., 2017; van der Wegen et al., 2013). The consequences of parameter uncertainty are also discussed by Brière et al. (2011), who propose an efficient calibration tool to find optimal parameter settings in Delft3D, a process based numerical model. While their tool makes the calibration less resource intensive, it does not quantify parameter uncertainty itself. Ruessink (2005)

was the first to quantify parameter uncertainty in coastal predictions, using Unibest-TC, a cross-shore profile model. He found the influence of parameter uncertainty to be largest near the sandbar crests of the cross-shore profile, but rapidly decreasing in other areas. Kroon et al. (2020) further build on this, assessing the relative importance of model uncertainty compared to natural variability in a one-line coastline model. They observe that the total prediction uncertainty increases significantly when including model uncertainties next to intrinsic uncertainty, and that model uncertainty even becomes the dominant form of uncertainty for multi-year time scales. Simmons et al. (2017) quantify parameter uncertainty in XBeach, a process-based storm erosion model, for cross-shore beach profile predictions. They demonstrate that quantifying parameter uncertainty leads to significant performance improvements and allows the modeller to communicate the expected uncertainty. Van der Wegen and Jaffe (2013) quantify uncertainty in morphological Delft3D predictions of San Pablo Bay in California. They find the predicted morphology in the bay to be largely determined by its plan form and less dependent on parameter variation, but argue that this does not necessarily hold for freer, less confined systems—such as the open coast. Hence, most studies indicate that model uncertainty is indeed a significant contributor to prediction uncertainty and that only accounting for intrinsic uncertainty would lead to overconfident predictions.

1.3.2. Knowledge gaps

Based on a review of relevant literature, the following knowledge gaps have been identified:

- **From simple to complex models**
  Currently, research on parameter uncertainty depends on the possibility to carry out large numbers of model simulations. Complex models are more resource intensive and sustain a larger parameter space. Hence, most studies have focused on computationally efficient models. With the exception of Brière et al. (2011) and van der Wegen and Jaffe (2013), the studies mentioned in the previous section have used one-dimensional models. Kroon et al. (2020) identify the application of a more complex model as a logical next step and suggest that it might reduce overall model uncertainty as it approximates the natural dynamics more realistically. On the other hand, this advantage of complex models might be counteracted by the extra uncertainty introduced due to the use of a larger number of model parameters to represent such processes (van der Wegen & Jaffe, 2013). To assess when the added complexity is justified, research comparing simple and complex models is essential (Ranasinghe, 2020; de Schipper et al., 2020).

- **Application to complex nature-based designs**
  Almost all of the aforementioned studies have examined the role of uncertainty in predictions of existing coastlines and/or natural phenomena. Only Kroon et al. (2020) have applied an uncertainty quantification to a nature-based design (the Sand Engine) to assess the impacts in the modelling of human interventions, but with a simple one-line model. Kroon et al. (2017) do apply a complex model for a nature-based design (the Hondsbossche Dunes), but carry out a simplified uncertainty quantification. Human interventions disturb the equilibrium state of the coastline and might therefore induce large changes in the system, which could affect the relative influence of uncertainty sources. Especially for complex nature-based designs, the need to understand the influence of model uncertainty on predictions is strong, as they include many natural processes and are often designed for large temporal and spatial scales (de Schipper et al., 2020). To date, little has been reported about the influence of parameter uncertainty on future modelling studies and design processes of nature-based solutions.

- **Spatial variation of parameter uncertainty**
  Another aspect that has not yet been researched extensively is spatial variability in parameter uncertainty. Especially for large-scale BwN designs there can be significant spatial variation in the relative importance of physical processes and forcing conditions. Model calibration parameters are often applied as scaling or multiplication coefficients within physical equations, so this spatial variation can be expected to carry over to parameter uncertainty as well. Understanding and visualising the spatial variation in uncertainty makes it possible to identify locations of weak or strong predictive ability. Ruessink (2005) observed significant variation in the influence of parameter uncertainty along the modelled cross-shore profiles, but the variation is expected to be more complex in 2D. Van der Wegen and Jaffe (2013) plot uncertainty indicators over their 2D model domain of San Pablo Bay, but the influence of parameter uncertainty is insignificant in their predictions. They suggest more pronounced results might be found for freer systems (e.g. the open coast).

1.3.3. Objective and research questions

From the literature review and identified knowledge gaps, the following research objective has been formulated for this thesis:

Investigate how, and to what extent, parameter uncertainty influences coastal area predictions of mega-nourishments.

This objective is split up in four research questions:

1. Which model parameters contribute most to the total parameter uncertainty?
2. How does the uncertainty in the selected parameters translate to uncertainty bounds in the predictions?
3. How does the parameter induced uncertainty vary over the study area?
4. How does the parameter induced uncertainty compare to uncertainty induced by variability in the wave climate?

1.4. Approach and scope

This section briefly outlines the approach followed in the thesis. The methodology is fully described in Chapter 3. Additionally, some considerations regarding the scope of the study are discussed and the outline of the thesis is presented.

1.4.1. Approach

To achieve the objective, this thesis aims to quantify parameter uncertainty in morphological Delft3D predictions of the Sand Engine (SE), a mega-nourishment constructed in 2011 at the Dutch North Sea coast (see Section 2.2). The SE is chosen as case study for this thesis because it has been monitored extensively and therefore a lot of data is available (de Schipper et al., 2016). For all simulations, the Delft3D model by Luijendijk et al. (2017) and Luijendijk et al. (2019) is used. The study period consists of the first 14 months after construction of the SE, from August 2011 to October 2012. The research approach is divided into the following four steps (Figure 1.3):

The first step is to determine the model outputs for which the uncertainty is analysed, and to select an initial set of model parameters, as candidates for variation (Sections 3.1 and 3.2). This is done based on literature and expert judgement, resulting in a shortlist of 16 Delft3D parameters. The second step comprises a sensitivity analysis of the 16 parameters, to determine which parameters are most influential and can therefore be expected to compose the main sources of parameter uncertainty (Section 3.3). Based on the sensitivity analysis, a subset of $O(5)$ parameters is selected for the uncertainty analysis—the third and main step of the thesis. Hereby, the uncertainty in the selected parameter space is analysed by generating a synthetic dataset of $O(1000)$ Delft3D predictions with identical simulation periods, but varying parameter sets. This dataset is analysed by means of the *Generalised Likelihood Uncertainty Estimation* (Section 3.4) (Beven & Binley, 1992), to derive uncertainty bounds and posterior parameter distributions. In the fourth step, the goal is to get an indication of the relative importance of parameter uncertainty, by comparing it to uncertainty due to natural variability in the wave climate (Section 3.5). This is done by running additional simulations ($O(10)$), using a fixed parameter set and varying wave conditions.

1. **Model outputs and initial parameters**
   - Define the model outputs to which sensitivity and uncertainty will be measured.
   - Select an initial set of parameters and ranges based on literature and expert judgement.

2. **Sensitivity analysis**
   - Determine the relative influence of the parameters on the defined model outputs.
   - Reduce the initial parameter set to the most influential parameters ($O(5)$).

- Analyse the model performance to refine the parameter ranges for step three.

3. **Uncertainty analysis**
   - Create a large sample of parameter sets ($O(1000)$) with the selected ranges from step two.
   - Derive uncertainty bounds and posterior parameter distributions.
   - Analyse the spatial and temporal variation of the uncertainty bounds.

4. **Uncertainty comparison**
   - Simulate different wave years using a fixed parameter set.
   - Compare the uncertainty bounds due to parameter uncertainty and wave climate variability.

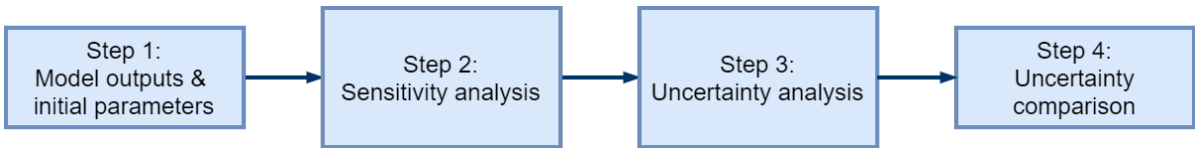

Figure 1.3: Flow chart showing the methodology for this thesis, divided in four steps. Steps two and three are represented by slightly larger boxes to show that they form the main part of the thesis work.

1.4.2. Modelling infrastructure

All simulations are carried out on the Cartesius supercomputer of SurfSara [1]. Individual tasks on its server are restricted to 120 consecutive computational hours. This poses a limit on the time period that can be simulated by the model. In theory, one could apply successive runs, using the results of one simulation as the input for the next, in order to simulate longer time periods. Within the scope of this thesis, however, all simulations will be limited to 120 hours, which allows a study period of roughly 14 months.

1.4.3. Limitations

There are certain limitations to the scope of this research, the most important are:

- Working with a 2DH-model requires large amounts of computational time, especially when many simulations need to be carried out. The amount of required simulations is strongly linked to the number of model parameters which are examined during the uncertainty quantification. The more parameters are analysed, the larger the parameter space in which the model operates. Hence, the study will be limited to a selected number model parameters for which the uncertainty will be quantified. The aim is to use the most influential parameters, so the remaining model parameters will be fixed at constant values, inevitably leading to residual uncertainty.
- While the study focuses on parameter uncertainty, other sources of model uncertainty will still be able to affect the model results. To isolate and analyse only parameter uncertainty itself, one would need to use a model which describes physical processes perfectly, has no numerical limitations and uses error-free observation data as forcing conditions. This is practically impossible, leading to further residual uncertainties.

1.4.4. Thesis outline

The thesis is divided in seven chapters. Chapter 1 has given an introduction to the topic and defined the scope and objective of the research. Chapter 2 provides additional background information on coastal erosion, the Sand Engine, the concepts of sensitivity and uncertainty analysis, and the applied Delft3D model. In Chapter 3 the exact methodology is laid out: the steps of the thesis are explained in detail, along with the applied methods. Chapters 4 and 5 present the results of the sensitivity and uncertainty analyses respectively. This is followed by a discussion of the results in Chapter 6. Finally, Chapter 7 summarises the conclusions of this study, along with several recommendations for future research.

[1] See SurfSara website: https://userinfo.surfsara.nl/systems/cartesius

2

Background

Chapter outline
Chapter 2 aims to summarise background information relevant to this research. First, the reader is introduced to the history of coastal erosion and what has been done to counteract it, specifically in the Netherlands (Section 2.1). This includes the recent application of mega-nourishments. Section 2.2 introduces the Sand Engine, which forms the case study for this thesis. The following sections treat the representation and quantification of uncertainty in numerical modelling in more detail: The concepts of sensitivity analysis (Section 2.3) and uncertainty analysis (Section 2.4) are explored, along with previous applications, specifically in coastal engineering. Special attention is given to the *elementary effects method* and *Generalised Likelihood Uncertainty Estimation* which are the main tools used in this thesis to quantify parameter uncertainty. Finally, the Delft3D model used for the simulations in this thesis is introduced in Section 2.5.

2.1. Coastal erosion

This section gives an introduction to coastal erosion and how it is being mitigated by sand nourishments. In this project we focus on wave-dominated sandy coasts.

2.1.1. The value of coastal zones

Coastal zones have long attracted human activity due to numerous factors, mainly the many ecosystem services they supply. It is therefore no surprise that, over time, coastal zones all around the globe have become subject to strong human population and development (see also Figure 2.1) (Small & Nicholls, 2003). In fact, 75% of cities with 10+ million inhabitants are located along its coastlines (Luijendijk et al., 2018). Climate change impacts, most notably SLR, are expected to continuously increase coastal erosion, especially of sandy beaches. According to various estimates, global SLR during the 21^{st} century may be up to eight times higher than during the 20th century (Church et al., 2013; Ranasinghe et al., 2009, and references therein). Many of the worlds coastlines have been eroding for decades to centuries already. An estimated third of the global ice-free shoreline is considered to be sandy and around 15% of which has regressed by an average of a metre or more per year in the last several decades (Luijendijk et al., 2018). Traditionally, coastal erosion has been counteracted by local, small-scale sand nourishments or so-called hard structures (i.e. revetments, groynes, seawalls etc.). But the combination of the immense socio-economic value of coastal zones and the uncertainty surrounding SLR estimates and corresponding levels of coastal recession, require a new approach to protect our natural sandy coastlines and low-lying communities in the long-term.

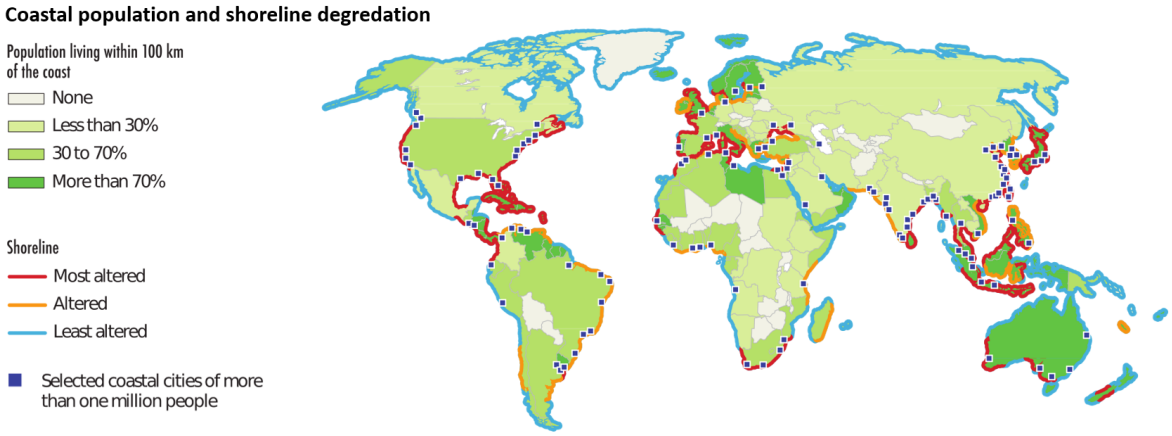

Figure 2.1: Estimated percentages of coastal population as of 2001, along with shoreline degradation levels and selected large coastal cities. By Bounford.com and UNEP/GRID-Arendal (2006) (Original source: Harrison & Pearce, 2001)

2.1.2. Sand nourishments

Sand nourishment forms a common strategy against erosion of the sandy coast in many parts of the world (e.g. Hanson et al., 2002; Valverde et al., 1999). It consists of increasing the sand volume of the beach using off-site sand (e.g. dredged at an inland location or further offshore). Next to protecting the hinterland from flooding and erosion, beach nourishment enhances recreational values (e.g. for tourism or surfing) and provides an opportunity to reuse dredged sediment from elsewhere. Moreover, they are less disruptive of the natural sediment transport patterns than hard structures such as sea walls or breakwaters (Hanson et al., 2002; de Schipper et al., 2020).

Traditionally, two types of sand nourishment are considered (Figure 2.2). The first is beach nourishment, where the sand is placed on the dry part of the beach, near the water line and/or near the dunes. The second type is shoreface nourishment, where the sand is placed in the submerged near-shore part of the beach (Hanson et al., 2002). In the Netherlands, shoreface nourishments have become increasingly popular as an alternative to beach nourishments since the late 1990s, based on two suggested advantages (Mulder et al., 1995): First, they reduce the nearshore water depth, inducing wave breaking earlier and damping the incoming wave energy. This is especially important during high wave energy events (i.e. storms), as earlier wave breaking reduces the

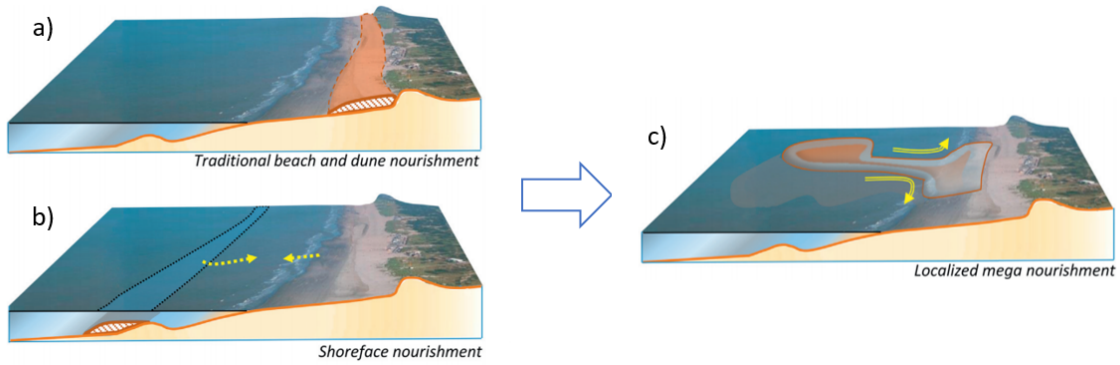

Figure 2.2: Overview of nourishment types: a) traditional beach nourishment (subaerial), b) shoreface nourishment (submerged), c) localised mega-nourishment. Adapted from Stive et al. (2013) and Radermacher et al. (2018)

alongshore current, leading to deposition of sediment. Secondly, they stimulate a gradual onshore movement of sand, resulting in a seaward movement of the shoreline.

Traditional beach and shoreface nourishments range in size between around 100 to 400 m^3 per metre alongshore. At locations that experience large annual sand losses, however, such smaller-scale nourishments are insufficient in the long-term, leading to frequent re-nourishments. These can have a detrimental effect to the natural system in which they are placed (de Schipper et al., 2016). Over the past two decades, a trend towards more integrated nourishment approaches has emerged. Such approaches often include social, environmental, and ecological aspects.

2.1.3. Mega-nourishments

A so-called mega-nourishment presents an alternative to traditional beach and shoreface nourishments, which incorporates the aspects of BwN (Figure 2.2c). It is essentially a mega-scale sand nourishment ($> 500m^3$ per metre alongshore). Mega-nourishments can come in a localised form (e.g. Sand Engine, approximately 21 million m^3 over 2 km alongshore) or distributed over a larger part of the beach (e.g. Hondsbossche dunes, approximately 35 million m^3 over 12 km alongshore). Here, we focus on localised mega-nourishments, for which Stive et al. (2013) describe four expected advantages over traditional nourishments: (1) The required nourishment frequency is decreased to 10–20 years (compared to the current $\pm$5-year frequency). (2) The localised nourishment will slowly diffuse due to natural processes (wind, waves, and currents) over a $O(10km)$ coastline stretch. (3) It considerably increases the locally available space for recreation and environment over the first stages of its life cycle. (4) While the ecological disturbance due to construction is strong, this is confined to the initial location ($O(2km)$), whereas the adjacent coastline is then slowly nourished by the natural diffusion of the nourishment. The first localised mega-nourishment is the SE, which is discussed in the next section.

2.2. The Sand Engine

This section introduces the case study of this thesis, the Sand Engine. The concept, design, initial evolution and dominant processes are discussed in more detail.

2.2.1. Concept and design

The Sand Engine (Dutch, *Zand Motor*) is a localised mega-nourishment constructed in 2011. It is located at the Delfland coast in the Netherlands, facing the North Sea between Hoek van Holland and Scheveningen (Figure 2.4a). This 17 km coastline stretch, bordered by the ports of Rotterdam and Scheveningen, has regressed by roughly 1 km in the period since 1600 a.d. and has been nourished with sand since the 1970s (de Schipper et al., 2016). The idea behind the SE is that it feeds the adjacent coastline over 20 years, as the sand is gradually redistributed through natural processes (wind, waves and currents).

The original nourishment comprised a large, hook-shaped sand peninsula, flanked by two shore-face nourishments (Figure 2.3). The total nourishment volume consisted of around 21.5 million m^3, divided over the penin-

sula ($\approx$ 17 million m^3) and shore-face nourishments ($\approx$ 4.5 million m^3). The peninsula covered 2.4 km in alongshore direction and reached around 1 km into the the sea. During the design phase, many different stakeholders were involved to combine engineering aspects with environmental, ecological and social considerations into one integrated coastal management system. Consequently, the shape of the peninsula was inspired by the potential for natural and recreational development (Stive et al., 2013). For instance, between the northern edge of the 'hook' and the beach, a sheltered tidal lagoon was included in the design to create a new habitat.

(a) (b)

Figure 2.3: Aerial photos of the SE, taken from the North in (a) July 2011 (shortly after construction) and (b) July 2012 (after the first year). Source: Rijkswaterstaat and van Houdt (2012)

2.2.2. Initial evolution

The evolution of the SE is being monitored since construction finished in August 2011. During the first 17 months, bathymetry surveys were conducted on a near-monthly basis (de Schipper et al., 2016). The shore normal of the Delfland coast is oriented at approximately 311 $^{\circ}N$, but the survey data was converted to a shore-orthogonal coordinate system, with its origin at the 'Schelpenpad' beach entrance (Figure 2.4).

As of 2016, the outer perimeter has regressed by up to 300 m, while the adjacent coast has progressed by up to 200 m. The initially asymmetric shape was reworked during the first 18 months to an almost symmetric shape along the coast (Figure 2.4). A sand spit developed at the northern edge of the peninsula, squeezing the lagoon entrance, but maintaining an active tidal channel. Observed morphological changes in the area were strongest in the first six months (including the storm season of December and January 2011/12), drastically reducing the curvature and cross-shore beach slope (de Schipper et al., 2016). After this 'initial response phase', the morphological changes became slower and more nuanced.

Cumulative volume changes for three areas (South section, peninsula, and North section) show that, of the volume loss measured around the peninsula over the first 17 months ($\approx$ 1.65 million m^3), around 72% was compensated for by accretion in the adjacent sections, confirming the natural diffusion of the localised nourishment (de Schipper et al., 2016). A larger portion of the redistributed sand volume settled in the North section ($\approx 60\%$ in the final months of 2012).

The morphological changes at the SE were strongly correlated to the incoming wave climate, i.e. months with high average wave heights resulted in larger changes and vice versa (de Schipper et al., 2016). This suggests that the waves and specifically the wave driven alongshore current form the dominant natural process governing sediment transport in this area. The majority of high-energy wave events after construction of the SE came from the southern / western sectors (de Schipper et al., 2016), leading to a Northward alongshore current. A sensitivity analysis of a Delft3D model of the SE also identified the wave climate as the clearly dominant forcing condition, followed by the vertical tide (Luijendijk et al., 2017). Other forcing conditions, such as the horizontal tide, surge levels and wind, had only minor effects according to the model results.

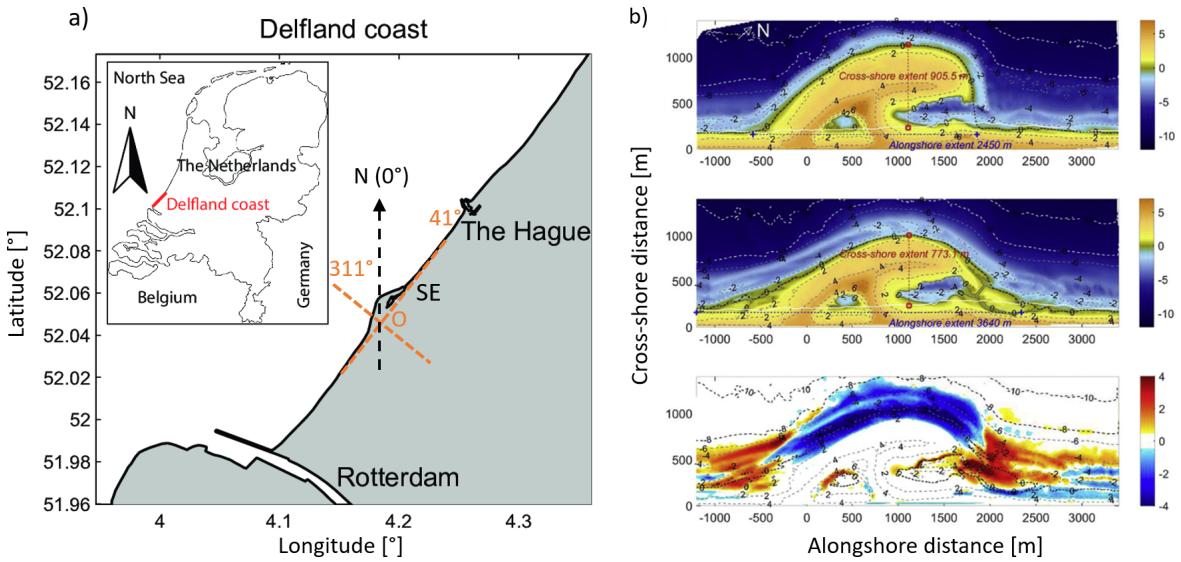

Figure 2.4: a) Geographical location of the Delfland coast and SE in the Netherlands. The orientation of the shoreline is visualised by the orange dashed lines, which represent the axes of the rotated, shore-orthogonal coordinate system, with origin (O) at the Schelpenpad beach entrance. Adapted from Radermacher et al. (2018). b) Measured bathymetries of the SE in August 2011 (first survey, top panel) and December 2012 (middle panel). The colors indicate the bed elevation with relation to MSL. The lower panel shows the measured bed level changes over the 17-month period in meters (red and blue colours indicate accretion and erosion respectively). Source: (de Schipper et al., 2016)

2.3. Sensitivity Analysis

Our ability to predict the behaviour of complex coastal interventions through numerical models is accompanied by many uncertainties (see Section 1.2). When it comes to quantifying and representing uncertainty in numerical modelling, two main tools exist, *Sensitivity Analysis* (SA) and *Uncertainty Analysis* (UA). Both tools examine how a model behaves when its inputs and parameters are varied over their range of validity. However, SA aims to identify the relative importance of model parameters and assumptions, while the goal of an UA is to quantify the actual uncertainty in the model predictions (Saltelli et al., 2019). This thesis utilises both tools. Saltelli (2002) defines SA as "the study of how the uncertainty in the output of a model (numerical or otherwise) can be apportioned to different sources of uncertainty in the model input". The SA in this study is carried out to determine which input parameters are expected to contribute most to the overall parameter uncertainty.

2.3.1. Overview

From a modelling perspective, SA is usually applied during model calibrations, to find out to which parameters the considered model output is most sensitive. The basic idea is to vary the input parameters of the model one at a time (local SA) or at the same time (global SA) and examine the changes in model outputs (Saltelli et al., 2008). SA is a large field on its own, and numerous methods exist. The most commonly applied methods include the simple One at A Time (OAT) method, which considers a base case of parameter values. For all subsequent model runs, only one parameter is changed from the base case. So-called screening methods increase in complexity from the OAT method by either changing multiple parameters simultaneously (e.g. factorial sampling) or considering multiple base cases (e.g. Morris, elementary effects) (van Arkel, 2016). Finally, probabilistic methods can be applied to change parameters in a stochastic manner (e.g. Monte Carlo or Latin Hypercube sampling). Which method is most suitable for a SA is case specific and depends on the goals of the study. Delft3D is a model with many different model parameters and the simulations in this study take 100–120 computational hours. Hence, the elementary effects (EE) method has been selected as it is considered an effective screening method for models with many parameters, while requiring a relatively low amount of computations. The concept of the method is explained in the following section. A thorough reference on SA in general, including the EE method can be found in Saltelli et al. (2008).

2.3.2. The elementary effects method

The concept of elementary effects was first proposed by Morris (1991), as a tool to find a model's input parameters, whose effects could be considered (a) negligible, (b) linear and additive or (c) nonlinear or involved in interactions with other parameters. The elementary effect EE_i of input parameter X_i is given by:

$$EE_i = \frac{Y(X_1, \dots, X_i + \Delta, \dots X_k) - Y(X_1, \dots, X_k)}{\Delta}, \tag{2.1}$$

where (X_1,...,X_k) are the input parameters, k represents the number of varied parameters, and Δ is a normalised increment value by which one of the input parameters is changed (i.e. Δ is a value in the interval $[0, 1]$ which is transformed to the ranges of the parameter: 0 and 1 correspond to the lower and upper limits respectively). $Y(X)$ is the model evaluation, i.e. the considered output variable. In short, an EE represents the difference between the model evaluations of two simulations, which differ in only one input parameter.

The parameter space is discretised into a p-level grid (i.e. every parameter can vary across p levels). Starting from a random base vector $\vec{x}^*$ of length k (containing one value for each parameter), every input parameter is changed by Δ one after another, with the condition that every parameter is changed exactly once and the order is randomised (Δ needs to be chosen such that it fits the p-level grid). This results in a random trajectory through the k-dimensional parameter space. Since every input parameter is changed exactly once during the course of the trajectory, one EE per parameter can be computed. This process is repeated r times ($r = O(10)$), resulting in r trajectories and r EE's per parameter (Figure 2.5). The model sensitivity to a certain input parameter can then be examined by computing the mean, μ (Equation 2.2), and standard deviation, σ (Equation 2.4) of that parameter's EE's. Hereby, μ_i represents the overall influence of parameter X_i on model output Y, while σ_i gives an indication of the ensemble of the parameter's effects—a high value of σ_i would imply that the EE's of X_i vary strongly among the trajectories, i.e. the model's response to the parameter is non-linear and/or the parameter interacts with other input parameters (see also Figure 3.4). Hence, the method provides insight into the influence as well as the dependence and non-linearity of the input parameters.

A potential disadvantage is that the application of μ and σ might render the method prone to type II errors (i.e. failing to identify an influential parameter) (Campolongo et al., 2007). This may occur when the model is non-monotonic, meaning the EE's for a certain parameter can be positive and negative, cancelling each other out during the calculation of μ and resulting in a low average. Morris (1991) proposed to handle this problem by considering the combination of μ and σ, as a low value for μ due to sign-switching EE's would still yield a high value of σ. Yet, especially for models with many parameters and outputs, this can become rather complicated. Campolongo et al. (2007) therefore suggest the application of μ^* (Equation 2.3), denoting the mean of the *absolute* values of a parameter's EE's, which solves the problem of sign-switching EE's. Moreover, they found that μ^* provides a good proxy of the total variance-based sensitivity index, making it a suitable measure for a parameter's absolute influence. The corresponding loss of information on the EE's signs can be prevented by simply computing all three sensitivity measures (μ, μ^*, and σ), providing as much information as possible.

$$\mu_i = \frac{1}{r} \sum_{j=1}^{r} EE_i^j, \tag{2.2}$$

$$\mu_i^* = \frac{1}{r} \sum_{j=1}^{r} \left| EE_i^j \right|, \tag{2.3}$$

$$\sigma_i = \sqrt{\frac{1}{r} \sum_{j=1}^{r} \left(EE_i^j - \mu_i \right)^2}, \tag{2.4}$$

where EE_i^j denotes the elementary effect corresponding to model parameter X_i and trajectory j.

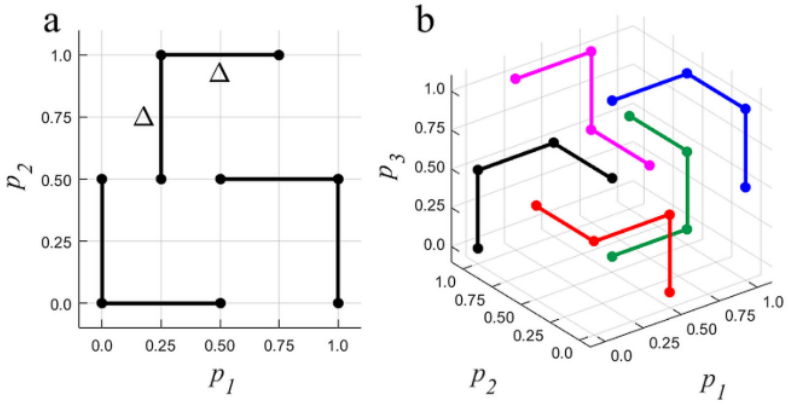

Figure 2.5: Example of r EE-trajectories through the k-dimensional parameter space of a model for a) $k = 2$ and $r = 3$ and b) $k = 3$ and $r = 5$. Source: Likhachev (2019)

2.4. Uncertainty Analysis

This section gives an introduction to uncertainty analysis. The concept is explained, followed by a brief overview of methods used to analyse and quantify uncertainty. The GLUE method is elaborated in detail as it will be used in this thesis.

2.4.1. Overview

Uncertainty analysis concerns the actual quantification of the uncertainty in model predictions. The resulting prediction uncertainty is commonly visualised in the form of confidence intervals (Baart et al., 2011), i.e. by giving a range of plausible outcomes, with a corresponding confidence level. Most uncertainty quantification methods rely on some sort of Monte Carlo (MC) approach. In fact, MC sampling is currently considered the only universal way to quantify model output uncertainty (Berends et al., 2019; Ranasinghe, 2016). Specific quantification methods do exist, however, depending on the considered source of uncertainty.

Intrinsic uncertainty is increasingly accounted for through probabilistic methods. A well-known procedure is to approximate the parameters representing the forcing conditions (e.g. wave climate) with probabilistic distributions, rather than choosing one design value (deterministic approach). These distributions can be derived from observation data, for example through an extreme value analysis. The applied distributions can then be used to draw a large amount of samples (usually through MC sampling), each of which is evaluated and the combined results visualised by confidence intervals. In case of dependence between forcing conditions, the applied distributions can be combined with correlation coefficients to derive the probability of occurrence for any set of forcing conditions (Jonkman et al., 2017). The main difficulty in applying probabilistic methods to estimate parameter uncertainty is that the distributions are based on observation data and/or physical theory behind the parameters. Since these are often lacking for model parameters, it is difficult to derive accurate distributions beforehand.

The quantification of epistemic uncertainty has received less attention in academic research. Most of the available studies have been carried out in hydrological and environmental modelling and focus on parameter uncertainty. Teng et al. (2017) provide an overview of the representation of uncertainty in flood-inundation modelling and distinguish two main types of uncertainty estimation. Both depend on model runs sampled from prior parameter distributions (commonly through MC sampling) to compute a parameter response surface which captures the uncertainty. Bayesian approaches use formal probability distributions and likelihood measures and have the potential to provide statistically rigorous confidence intervals (e.g. Hall et al., 2011). On the other hand, there are less formal methods, most prominently the *Generalised Likelihood Uncertainty Estimation* (GLUE) (Beven & Binley, 1992). GLUE depends on more flexible, subjective likelihood measures to divide the parameter space into areas of acceptable and non-acceptable model behaviour (Teng et al., 2017). In theory, formal Bayesian methods would be preferred, as subjectivity introduces bias in the results. A problem with Bayesian methods, however, is that they assume the prior parameter distributions to be known, which is usually not the case in practice. Therefore, through the selection of the prior parameter distributions, they ultimately depend on subjective choices as well.

A significant limitation to all quantification methods relying on an ensemble of model runs is the required computational capacity to run sufficient simulations. This is especially true for models with a large number of input parameters and/or long running times, which is often the case in coastal area modelling (Ranasinghe, 2016). Efforts to develop more efficient approaches have mainly focused on two branches: using alternative sampling techniques over MC or applying so-called surrogate modelling. Alternative sampling strategies commonly have one of two goals. They can be applied to distribute the samples more uniformly over the parameter space, for example using Latin Hypercube sampling (e.g Uhlenbrook & Sieber, 2005) or low-discrepancy sequences (e.g. Sobol's sequence, Sobol', 1967). On the other hand, the aim can be to achieve a higher sampling density in areas of the parameter space with higher likelihoods, for example through importance sampling (e.g. Sun et al., 2013) or Markov Chain MC methods (e.g. Blasone, Vrugt, et al., 2008; Vrugt et al., 2003). In surrogate (or emulation) modelling (Castelletti et al., 2012), a less resource intensive *surrogate model* is used to mimic the results of the complex model. Examples of surrogate models include Bayesian networks (e.g. Kroon et al., 2017; Plant et al., 2011), artificial neural networks (e.g. Khu & Werner, 2003; Ruessink, 2006), a model based on a simpler structure, or a low-fidelity version of the same structure as the original model (e.g. Berends et al., 2019). The idea is to use a limited amount of complex model runs to calibrate the surrogate model, which is subsequently used to carry out a large amount of runs at a lower computational cost. Hence, the applied sampling strategy and model structure play an important role in UA as they influence the number of possible samples and corresponding distribution over the parameter space.

GLUE is generally applied more often than its formal, Bayesian counterparts, mainly because it is simple to implement and does not require prior assumptions on the error structure (as formal Bayesian methods do) (Blasone, Madsen, et al., 2008). Although GLUE has been used and researched mainly in hydrological applications, it has also been successfully applied to simple coastal morphology problems (Kroon et al., 2020; Ruessink, 2005; Simmons et al., 2017). Therefore, it has been selected as the uncertainty quantification method for this thesis. The method is described in more detail in the following section, while the exact application is discussed in Section 3.4.

2.4.2. The GLUE method

The Generalised Likelihood Uncertainty Estimation was developed by Beven and Binley (1992) to examine uncertainty in predictions of hydrological models. The fundamental idea is to assign prior distributions to model parameters—uniform distributions can be applied in case of little or no prior knowledge on the parameter value—and subsequently drawing a large sample of parameter sets, originally through MC sampling (Figure 2.6). For each set, the resulting model prediction is evaluated against observation data using a skill score. A user-defined threshold is applied to the skill score, below which a model prediction is deemed *non-behavioural*, i.e. it is not accepted. The parameter sets resulting in *behavioural* predictions (a skill score above the threshold) are further evaluated using a likelihood measure, related to the achieved skill score. Non-behavioural runs receive likelihood scores of zero. The result is a likelihood range for each model parameter, resembling a probability density function (PDF). Hence, GLUE lets the user search the model parameter space for acceptable parameter sets, by assigning a non-zero likelihood to all sets with a prediction skill above a predefined threshold (Beven & Binley, 1992). Uncertainty bounds resulting from the ensemble of behavioural runs can then be used to derive an observation-based value of parameter uncertainty.

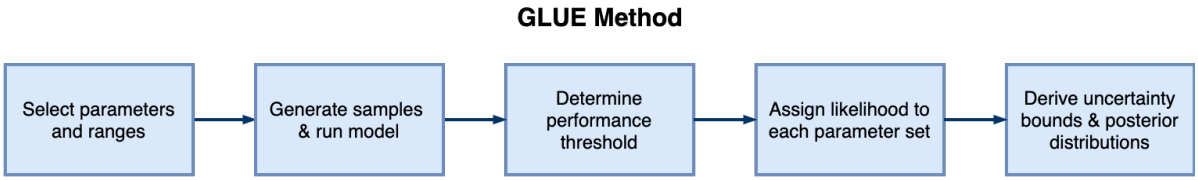

Figure 2.6: Flowchart showing the process of the GLUE method for uncertainty analysis, based on Beven and Binley (1992) and Simmons et al. (2017).

Another important aspect of the GLUE method is the concept of *equifinality* (Beven, 1993). Equifinality denotes the possibility that multiple parameter sets can result in a model prediction of equal skill, i.e. a unique optimal parameter set is non-existent (hence also called *non-uniqueness*). Equifinality can be caused by parameter interdependence and/or model insensitivity (Ruessink, 2005). Consider, for example, a model that is insensitive to a certain parameter. Consequently, changing that parameter's value will not (significantly) influence the model skill, yet it is a different parameter set. It is important to note, however, that equifinality is bound to the calibra-

tion space, as equifinal parameter sets may respond differently outside this space (van Maren & Cronin, 2016). For instance, an optimal parameter set determined for a certain forecast period is unlikely to remain optimal when the forecast period is extended (or shortened). One way to find a reliable parameter set, is to look for the best combined likelihood over several different time periods or locations (Simmons et al., 2017).

While the GLUE method has been implemented in a variety of applications, it has also received criticism. Mainly because the application of the method involves subjective decisions that need to be made by the user (i.a. the choice of the skill score and behavioural threshold). This may appear contradictory to the aim of objective science, hence, GLUE has been criticised for not implementing a formal statistical framework (e.g. Mantovan et al., 2006; Stedinger et al., 2008). Beven and Binley (1992) already acknowledge this issue in their original paper, stating that "the importance of an explicit definition of the likelihood function is then readily apparent as the calculated uncertainty limits will depend on the definition used. The modeller can, in consequence, manipulate the estimated uncertainty of his predictions by changing the likelihood function used." They argue, however, that this would not necessarily be unreasonable, provided the likelihood definition is explicit. Beven and Binley (2014) further argue that for some applications it is even better not to apply a formal statistical framework due to the presence of uncertainties that result from a simple lack of knowledge rather than random variability and would therefore obstruct such a formal framework, leading to bias in the results (e.g. when estimating model calibration parameters). A flexible method like GLUE, then at least gives the possibility to exclude unacceptable model runs from the analysis.

2.4.3. GLUE in coastal modelling

Ruessink (2005) was the first to apply GLUE to a coastal modelling study. He examined parameter uncertainty using the cross-shore profile model Unibest-TC for predictions of the beach profile at Egmond aan Zee in the Netherlands. He examined four model parameters using the GLUE method and observed that, indeed, a wide range of parameter sets resulted in acceptable model predictions (equifinality), mainly due to parameter interdependence and model insensitivity. The study finds GLUE to be an effective method of uncertainty estimation, yet argues it is also inefficient, as a large amount of computational time is essentially wasted on non-behavioural simulations. In this light Ruessink also touches on potential improvements by applying more efficient sampling methods, for example Latin Hypercube (LH) sampling or Markov Chain MC methods. The subjectivity of GLUE, is also discussed, as the width of the uncertainty bounds depends directly on the choice of the behavioural threshold. This choice is highly influential on the results, as a high threshold will lead to unrealistically narrow bands, while a low threshold would result in uncertainty bounds so wide, that one could essentially dismiss the model altogether. Furthermore, some model parameters may turn out to be unidentifiable by GLUE, a problem described as parameter identifiability (i.e. the method does not succeed in identifying a likely range of values for a certain parameter) (Ruessink, 2005). This is, to some extent, linked to the concept of equifinality and also caused by parameter interdependence and model insensitivity. For example, two of the four parameters Ruessink examines turn out to be strongly correlated, undermining the GLUE methods efforts to find optimal values.

Simmons et al. (2017) apply the GLUE method, to examine parameter uncertainty in storm erosion modelling at the Emilia-Romagna coast in Northern Italy. First, 15.000 simulations are carried out to analyse six model parameters. The results are used to asses parameter sensitivity and refine the parameter distributions before carrying out a second GLUE analysis for the four most sensitive parameters, with adjusted ranges. Prediction accuracy is found to be greatly improved compared to a traditional calibration process. Similar to Ruessink (2005), Simmons et al. (2017) recognise GLUE as an effective method for uncertainty estimation and give some recommendations for implementation. For example, when using multiple sets of observation data or model outputs, it is important to ensure that the same dominant physical processes occur across them—if different physical processes dominate, behavioural parameter values might differ among these processes, impeding the search for reliable parameter values for the entire dataset. The criticism directed towards the subjectivity of GLUE is also discussed, but it is argued that this does not devalue its use, as it should be seen as a method that provides valuable uncertainty estimates rather than an absolute uncertainty. This allows the modeller to make decisions based on the insight from these estimates.

Finally, Kroon et al. (2020) use GLUE to evaluate model uncertainty in a simple one-line model with a single parameter, to simulate the evolution of the SE. The focus lies on examining the relative importance of model uncertainty and intrinsic uncertainty, by combining GLUE with a probabilistic bootstrapping method to represent wave climate variability. By using a simple model with only one model parameter and calibrating to uncorrected observation data (prone to measurement errors), the study is not limited to parameter uncertainty but also

includes model inadequacy, numerical uncertainty and observation uncertainty (i.e. model uncertainty as a whole). The study is carried out in two steps (similar to Simmons et al. (2017)): first, a GLUE analysis with 200 simulations is carried out to evaluate the model parameter range. The resulting posterior distribution is used for a second GLUE consisting of 12.000 model runs. These are combined with 12.000 wave-series obtained from the bootstrapping method. The results show that including model uncertainty in the predictions leads to a significant increase in overall prediction uncertainty. Furthermore, while wave climate variability dominates in the first year, model uncertainty becomes the main uncertainty source thereafter. While not focusing on the suitability of the GLUE method itself, Kroon et al. (2020) reckon that the full strength of the method could be exploited by applying it to a more complex model.

2.4.4. Spatial variation of uncertainty

One of the research questions focuses on how parameter uncertainty varies over the SE area. Few studies actively discuss spatial variation in the degree of uncertainty over the model domain. It is logical that the uncertainty will vary to some degree, as forcing conditions and site specific characteristics also vary spatially (e.g. local sand bars). For one-dimensional models it is easier to identify and visualise such spatial variation as the confidence bounds are single lines. Ruessink (2005), for example, observes the width of the uncertainty bounds to be largest at the bar crests of the cross-shore profile, while parameter uncertainty diminishes in other areas.

Spatial variation becomes more complex with additional dimensions. For 2DH predictions (depth averaged) uncertainty may vary in both horizontal dimensions. The uncertainty bounds are then surfaces instead of lines. Van der Wegen and Jaffe (2013) look at spatial variation in morphological (2DH) Delft3D simulations of San Pablo Bay in California. They do not focus on parameter uncertainty, but examine overall prediction uncertainty in bed level changes. A confidence index represents the degree of uncertainty for the bed level changes at grid cell level. This confidence index is derived from the mean and standard deviation of an ensemble of model simulations. Interpolating the confidence index over the model domain, gives a visual representation of uncertainty throughout the study area. At San Pablo Bay, variations in model input only translate to small variations in the model outcome, however, van der Wegen and Jaffe (2013) suggest this may be different if more model runs were carried out. Furthermore, they argue that the system is governed mainly by the plan form of the bay and the hydrodynamic interactions in the model, but that freer, less confined systems may be subject to more uncertainty due to parameter variation. A system such as the SE, which lies at the open coast of the North Sea, and is constantly forced by waves and currents can certainly be considered free and unconfined.

2.5. Delft3D model

The existing Delft3D model by Luijendijk et al. (2019) will be used for all simulations in this thesis. This model is based on Luijendijk et al. (2017), who used Delft3D to hindcast the initial evolution of the sand engine, from 2011–2014. Luijendijk et al. (2019) implemented morphodynamic acceleration techniques in order to allow simulations over larger time-scales. This section elaborates the model configuration used in this thesis in more detail, while an overview of the governing equations is given in Appendix A. A full conceptual description of Delft3D can be found in Deltares (2011a) and Lesser et al. (2004).

2.5.1. Basic configuration

The model uses a coupled setup of Delft3D-FLOW and Delft3D-WAVE. The FLOW module computes the hydrodynamic forcing (tides and currents) and the sediment transport. It is the main component which ultimately computes the bed level changes in the model domain (Figure 2.7a). The WAVE module carries out the wave computations, using SWAN (Booij et al., 1999). The two modules work together in an online coupling setup, meaning that the WAVE module computes the wave conditions in certain intervals and feeds these back to the FLOW module, which includes them in the sediment transport computations. Additionally, the FLOW model provides an optional roller model to include the roller energy, transferred by breaking waves, in the computations (Reniers et al., 2004). Including the roller model led to more accurate results at the SE, albeit at the cost of extra computational time (Luijendijk et al., 2017).

2.5.2. Model domain & boundary conditions

This subsection gives a brief overview of the applied model domain, computational grid, and boundary conditions, but is not intended as a complete description. The reader is referred to Luijendijk et al. (2017) and

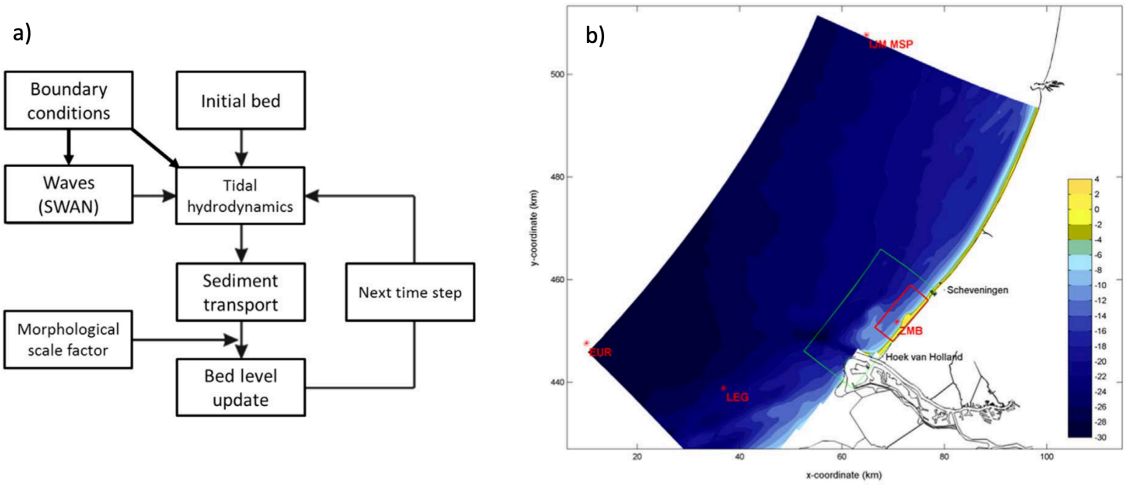

Figure 2.7: a) Flowchart of the model structure applied in Delft3D. b) Model domain applied in the SE model of Luijendijk et al. (2017), showing the nested wave grids (entire domain and green box) and the hydrodynamic grid (red box). The model applied in this thesis uses a larger hydrodynamic grid, as described in Luijendijk et al. (2019). The measurement stations from which the boundary conditions stem are also shown: Europlatform (EUR, waves), Ijmuiden (IJM MSP, waves) and Lichteiland Goeree (LEG, wind). Source: Luijendijk et al. (2017)

Luijendijk et al. (2019) for a detailed description and corresponding model validation. The model is applied on a 2DH computational grid. Hence, the grid covers the model domain in the horizontal dimensions, but vertically all quantities are averaged over the water depth, which is defined for each grid cell. The model domain of the Delfland coast is schematised by a curvilinear grid, which covers an area of 26 km in alongshore and 15 km in cross-shore direction (see Figure 2.7b). This grid is used for all hydrodynamic and morphological computations by the FLOW module. The resolution varies from 35 m to 500 m and is highest in the SE area, decreasing with distance from there. The bathymetry in the SE area is based on the first measurements taken after construction on 2 August 2011 (de Schipper et al., 2016). The remaining domain (beyond the 10 m depth contour) is based on Echo-sounding surveys conducted by Rijkswaterstaat.

To propagate the offshore wave conditions into the hydrodynamic domain, two additional nested grids are applied in the WAVE module (the larger one covering an area of 79 km alongshore and 42 km cross-shore). The largest grid is forced by time series of wave heights, periods, and directions from two offshore locations (Europlatform and Ijmuiden), combined with a uniform wind based on the measured time series at Lichteiland Goeree. The Europlatform conditions are used to force the southern boundary and the Ijmuiden conditions to force the northern boundary. The western boundary is forced by either location, depending on the incoming wave direction. The WAVE module propagates these conditions through the nested grids into the hydrodynamic domain, where they are included in the FLOW computations.

The tidal boundary conditions originate from nesting in a large-scale model for the Dutch Continental Shelf. The obtained tidal information is converted into astronomical components for the offshore boundary of the hydrodynamic model domain and used to derive zero-gradient water level conditions at the lateral boundaries. Finally, surge levels and wind speeds observed at Hoek van Holland are added to the tidal water level.

2.5.3. Morphodynamic acceleration

As the timescales of hydro- and morphological processes vary greatly—morphological time scales are usually a factor $O(100 - 1000)$ larger—it can be computationally very inefficient to make morphological simulations if a small time step is applied for stability of hydrodynamic computations (see also Ranasinghe, 2020). Therefore, morphodynamic acceleration techniques have been developed. The morphological acceleration factor by Lesser et al. (2004) (henceforth referred to as *morfac*) is one of the most commonly applied methods. The idea is to upscale the effects of hydrodynamic processes on the morphology, by multiplying the computed bed level changes by a factor equal to the *morfac* (denoted as Morphological scale factor in Figure 2.7). For example, a simulation over one tidal cycle (12 h) with a *morfac* of 10 would equal 120 h of morphological change (Luijendijk et al., 2019). The model used in this thesis applies a *morfac* equal to 3.

Input reduction techniques are also commonly applied in morphological modelling, and are normally required in

combination with the *morfac*. The idea is to reduce the computational effort by applying a reduced set of forcing conditions, which still yields satisfactory results. Luijendijk et al. (2019) examined and tested several methods for input reduction on the Delft3D model of the SE. The two main approaches are outlined below:

- **Brute Force (BF)**
  The key concept of BF methods is that they are based on a realistic time series, rather than a statistical representation of the forcing conditions (in this case the wave climate). If the original time series is used directly it is simply called BF. However, the time series can also be filtered and compressed with a time-compression factor (Li et al., 2018), resulting in the Brute Force Filtered & Compressed method (BFFC). When used in combination with the $morfac$, the time compression-factor and $morfac$ should be equal. Finally, the BFFC series can be split up into multiple time series, which are processed in parallel, giving the Brute Force Merged (BFM) technique. The computed bed level changes are then merged at each time step by taking a weighted average, which is returned to all processes to continue with the next time series. An important assumption with this is that the hydrodynamic conditions vary much faster than morphological development. In other words, if the period in which different hydrodynamic conditions appear is short compared to the morphological time scale, then they can be computed at the same time.

- **Wave reduction based on longshore transport**
  Longshore sediment transport (LST) based approaches do not apply a time series for the wave climate, but rather derive a set of wave conditions from the LST, which is computed for the entire wave climate. A reduced set of conditions ($O(10)$) is determined, which results in a similar LST. This can be a difficult process, however, especially if there is temporal variability in the wave climate (e.g. seasons). Another limitation is that the depth of closure should be embedded in the wave climate. But this is very difficult as the full outer range of wave heights would have to be represented, requiring too many conditions.

The best overall method for medium to long-term simulations is the BFM, which includes the full wave climate variability and therefore most physical processes, but is still computationally efficient due to parallel computing (Luijendijk et al., 2019). The time-scale for the simulations of this thesis is in the order of one year, however. Merging the time-series would in that case mean that seasonal variations are computed at the same time, which might impair the results. On the other hand, BF results in too large computational times. Therefore, the BFFC method is used as input reduction technique for this study. The LST methods are only viable for larger time-scales (several years to decades). They are computationally very efficient, but the physical processes are less represented because of the reduced wave conditions. Hence, they will not be considered further. The applied BFFC wave series is described in more detail in the next section.

2.5.4. Forcing conditions

The model is forced by four mechanisms: waves, tide, surge levels and wind. Hereby, the incoming waves represent the dominant mechanism (Luijendijk et al., 2017). The model is forced by a time series of wave conditions at each grid boundary, which are propagated through the model domain. However, the wave time step is larger than the hydrodynamic time step (the wave field is updated and propagated by the WAVE module after a certain amount of hydrodynamic time steps, at which the tidal forcing and sediment transport are computed by the FLOW module). In other words, the two time steps are decoupled (Luijendijk et al., 2019).

During the BFFC approach, all waves lower than 1 m or directed away from the coast are filtered from the series for computational efficiency, as these do not lead to significant morphological changes (Luijendijk et al., 2019). However, the model is forced by two wave time series (Europlatform and Ijmuijden), both consisting of identical time entries, in 10 minute intervals. The two series cannot both be filtered, as that would lead to inconsistent time entries. As the Western sector contains the dominant long-term wave directions, the Europlatform series is selected as BFFC series. After the filtering, the remaining series contains only 48% of the original wave conditions. The same time entries filtered from the Europlatform series are also removed in the Ijmuiden series, to ensure consistency in time. The gaps that appear at the filtered time entries are eliminated (i.e. the time intervals between the data points of the remaining series are restored to ten minutes). Because of this, the difference in 'real' time between two data points might differ significantly from the time interval of ten minutes (hence, the time series is distorted). The model will simply run through the distorted wave time-series, which means that in real time it skips the periods with negligible wave conditions. Finally, the filtered wave series is compressed by a factor equal to $morfac = 3$. This means that the time step in the wave series, and with it the computational time, is divided by three. This is compensated for by the *morfac*, which multiplies all computed

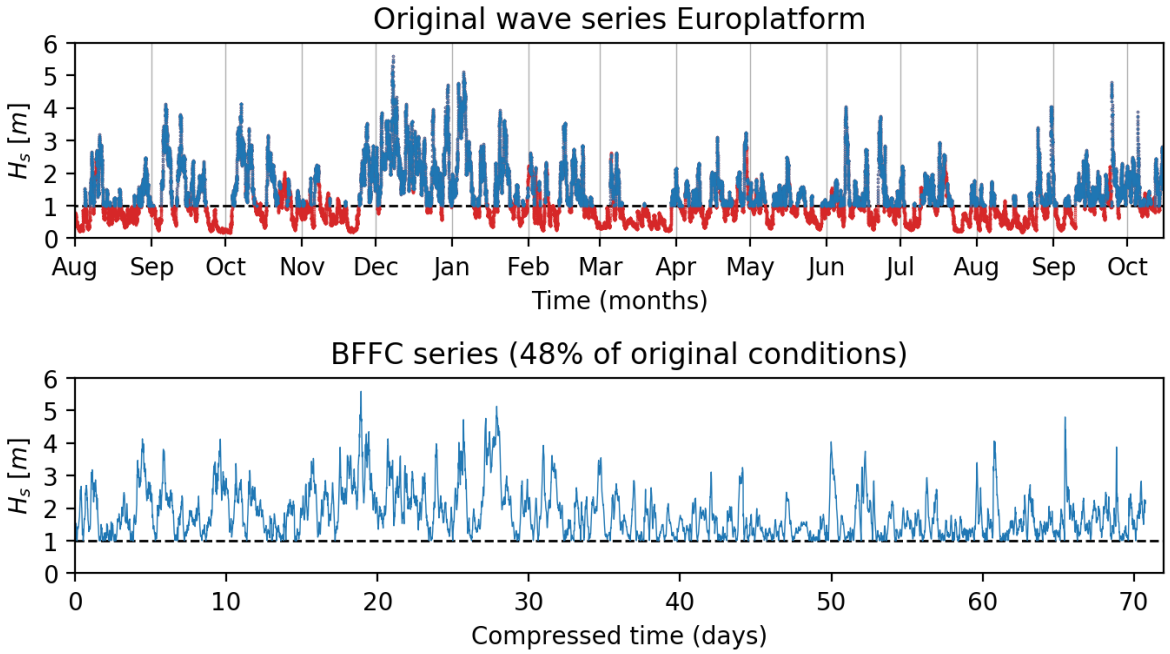

Figure 2.8: Time series of the significant wave height (H_s) at the Europlatform for the study period of August 2011 to October 2012. The upper panel shows the original time series, while the filtered and compressed time series (BFFC) is presented in the lower panel. To obtain the BFFC series, all waves lower than 1 m or directed away from the coast are removed (red in the upper panel), leaving only 48% of the original wave conditions (blue in the upper panel). The remaining series is compressed in time by a factor $morfac = 3$, resulting in the BFFC series (lower panel), which is only $\approx$ 72 days long, compared to the original 442-day series. The black dashed lines represent the limit $H_s = 1\ m$.

bed level changes by three. The resulting filtered and compressed (BFFC) series is only 72 days long (compared to the original 442 days), so the simulation time has been reduced by a factor ≈ 6 (see Figure 2.8 for a comparison between the original and the BFFC time series).

For the tidal forcing, the local tidal components corresponding to the geographic location are used. However, the tidal time series cannot be coupled to the BFFC wave series. This is because the main forcing mechanisms of the tide are the water level and flow velocity. Compressing the tidal water level variations might lead to unrealistic behaviour of horizontal flow velocities, which would carry over into the sediment transport. Thus, one of the main assumptions behind the BFFC approach is that the tidal time step can be decoupled form the wave time step. The computed sediment transport is still multiplied by the *morfac*, so the time compression is compensated. However, the filtered time periods in the BFFC series are not compensated for with regard to the tidal forcing. This means that the model is only forced by the tide for approximately half the simulated period ($\approx$ 216 out of 442 days). Hence, this assumption is only valid if the waves are the dominant forcing mechanism, which is the case for the SE (except for the sheltered lagoon) (Luijendijk et al., 2017).

Finally, the forcing by wind and surge levels occurs on the same time scale as the wave forcing. Hereby, the model is simply forced by the wind speed and surge levels corresponding to the same original timestamp as the current wave condition. Wind speed and direction are assumed constant throughout the model domain, while the surge level is added to the tidal water level.

3

Methodology

Chapter outline
This chapter outlines the methods and tools used to work towards the research objective of this thesis. Hereby, the four steps defined in Section 1.4 are elaborated in detail. Section 3.1 discusses which output variables are used to assess the model performance. In Section 3.2, the 16 parameters forming the initial selection are introduced. Next, the two main steps of this thesis are explained: The sensitivity analysis by means of the Elementary Effects method (Section 3.3) and the uncertainty analysis through the GLUE method (Section 3.4). Finally, in Section 3.5, the uncertainty comparison is elaborated.

3.1. Model performance and outputs

To gauge the model's predictive ability, it is important to clearly define how the performance is measured and which output variables are examined. This study focuses on morphological modelling of a mega-nourishment, hence it is of most interest to see whether Delft3D is able to accurately predict the redistribution of sediment from the peninsula to the adjacent coastline. This section outlines how the model performance is determined and describes the output categories in more detail: cumulative volume changes, shoreline position and bed level changes.

3.1.1. Model performance

During the GLUE method, the various simulations need to be graded, so the definition of the performance measure forms an integral part of the method. For the SA this is not necessarily required, as the model performance itself plays no part in the EE method. Nonetheless, the SA provides the first simulations, which form a useful dataset to analyse before the GLUE analysis.

In general, a model's performance may be assessed through three criteria: bias, accuracy and skill. Bias denotes the difference in central tendencies, while accuracy represents the error between predictions and observations. Skill, on the other hand, measures the accuracy of a prediction relative to a baseline prediction, which is often more useful than an indication of pure accuracy or bias (Sutherland et al., 2004). In coastal modelling, a skill score is the most commonly applied performance measure.

A suitable measure of skill in morphological modelling is the Brier Skill Score (BSS) (e.g. Luijendijk et al., 2017; van Rijn et al., 2003; Roelvink et al., 2020; Simmons et al., 2017). The BSS compares predictions to observations by means of the mean squared error (MSE) (Equations 3.1 and 3.2). The so-called 'zero change scenario' is commonly used as baseline prediction. The score returns a skill value with an upper bound of 1 (representing perfect agreement between simulation and observation data). A value of 0 would indicate identical performance to the 'zero change' scenario, while a negative score represents a simulation that approximates the observed data worse than the baseline prediction (note that the BSS does not have a lower bound, meaning it can reach large negative values, especially when the denominator is small). Van Rijn et al. (2003) found the BSS to be particularly useful for morphological predictions, using the initial bathymetry as baseline prediction (i.e. the zero change scenario). This setup will also be applied in this thesis.

$$\mathrm{MSE}(Y,\hat{Y}) = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2 \tag{3.1}$$

$$BSS = 1 - \frac{MSE(Y,\hat{Y}_m)}{MSE(Y,\hat{Y}_0)}, \tag{3.2}$$

where Y represents the observed values, $\hat{Y}$ the predicted values by the model (subscript m) and the baseline prediction (subscript 0), and n the number of evaluated data points.

3.1.2. Output categories

During the thesis work, the simulations were evaluated starting with simpler output variables and then gradually increasing in complexity and dimensions. This approach has worked well and the idea is to reflect this build-up in complexity in the output variables used for this thesis. The aim is to use output variables which reflect the focus on the redistribution of the sediment around the SE (see Section 1.4). However, the broad range of output data from the Delft3D model can be used to define numerous different outputs. Therefore, the definition of the examined model outputs will always remain subjective to some degree. The chosen output variables are divided in three categories, with inspiration taken from Luijendijk et al. (2017) and de Schipper et al. (2016): cumulative volume changes, cross-shore shoreline position and cumulative bed level changes. Each output variable is evaluated over the entire SE area, but also specific control areas (the coming sections explain the exact definitions of these areas).

Two areas of particular interest are the spit development at the northern end of the peninsula and the erosion of the head of the peninsula (the most seaward part of the peninsula). The spit evolution is a highly three-dimensional feature, governed by many physical processes, which is hard to predict for the model. This makes it a point of interest for this study. Ideally, the spit evolution would be reflected in an output variable that includes

these 3D processes, however, finding a suitable morphological indicator is difficult. Comparing bed levels per grid-cell is not an option as this quickly results in a double penalty problem (i.e. if the spit is modelled well, but deviates slightly in exact location from the observation, it will result in a very low score as the bed levels are compared at identical positions). Another idea would be to use a hypsometric curve, which looks at the fraction of the surface area located at each elevation. In theory, this should eliminate problems arising from deviations in the exact spit location. Unfortunately, when applied to this study, a hypsometry criterion still failed to identify the simulations which show good spit evolution. Finally, it has been decided to use the shoreline position in the spit area (see Section 3.1.4), as this was the best metric to identify spit evolution.

The erosion of the head section provides a lot of information on the diffusion process of the initial coastline perturbation by the SE. It is less dependent on the local 3D morphology and hence the double penalty problem is less pronounced. Therefore, the head section was included in the bed level category.

3.1.3. Cumulative volume changes

One of the easiest ways to measure the performance of the model regarding the redistribution of sediment is to look at changes in sediment volumes. Four control areas are used to compute the volume changes, following de Schipper et al. (2016): the middle section of the SE (encompassing the original peninsula), the adjacent sides to the North and South, and the combined SE area, composed of the three sections together (Figure 3.1). The middle section is adjusted for this study, to exclude the permanently dry part of the peninsula (see lower bound of the red polygon in Figure 3.1). This is because Delft3D does not include aeolian transport, which has in practice eroded sediment from the dry part of the peninsula. To prevent this from affecting the model performance too much, the adjusted middle section is used.

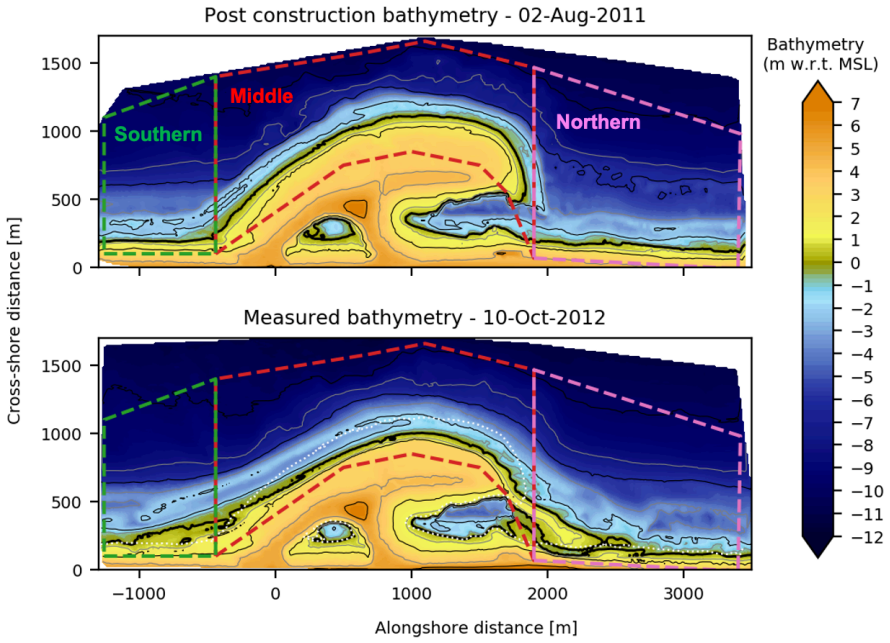

Figure 3.1: Bathymetry plot obtained from the measured bed levels by de Schipper et al. (2016), shortly after construction (2 August 2011, upper panel) and after around 14 months (10 October 2012, upper panel). The three polygons show the control areas for which the volume changes are calculated. The fourth area simply consists of the entire area enclosed by the three polygons together.

The cumulative volume changes can be computed at every output time step by multiplying the cumulative bed level change for each grid cell with the corresponding cell area to get the volume change per grid cell. The sum of the volume changes of all cells within the corresponding control area gives the cumulative volume change (Equation 3.3).

$$\Delta V_t = \sum_{m}^{M} \sum_{n}^{N} \Delta z_{t,m,n} \cdot A_{m,n} \tag{3.3}$$

where ΔV_t denotes the cumulative volume change (in m^3) at time step t, $\Delta z_{t,m,n}$ the cumulative bed level change at t, $A_{m,n}$ the cell area, the subscripts m,n the indices of the current grid cell, and M,N the indices of the final grid cell within the considered control area.

The BSS for the volume changes ($BSS_{\Delta V}$) is computed on an equal weight basis for the four control areas. The monthly observed bathymetries at the SE can be transformed to a time series of cumulative volume changes from August 2011 to October 2012. For each timestamp, the chronologically closest model output is used to compute the predicted volume change. Then the computation of the BSS is simply done through the MSE comparison of the obtained time series.

3.1.4. Shoreline position

Another measure which relates directly to the transport of sediment, is the change of the shoreline position. As sand is eroded from the middle section and moved towards the adjacent coast, the shoreline position will change towards a flatter curve. Here, the shoreline is defined as the zero-meter-contour during MSL, which is computed by interpolation from the output bathymetry, similar to de Schipper et al. (2016).

As the output data is transformed to the shore-orthogonal coordinate system, the cross-shore position gives a direct indicator of the shoreline position. To compare different shorelines using a similarity measure as the MSE, the cross-shore position should be defined over constant alongshore points. This can be achieved by interpolating the observed and predicted shorelines over equidistant alongshore points. Every alongshore position can only correspond to one cross-shore position, however, which means that, in case the $0m$-contour crosses an alongshore point multiple times (e.g. the lagoon), the most seaward position is taken. In other words, the shoreline does not include the lagoon, but 'jumps' from the northern end of the peninsula to the beach (Figure 3.2). The MSE can then be computed over any alongshore interval, whereby the baseline prediction is the initial shoreline (August 2011). It should be noted, however, that the inter-tidal morphology may affect the exact position of the shoreline. Especially for the observations this will inevitably lead to some residual uncertainty (de Schipper et al., 2016).

The shoreline is used to derive two BSS's: one for the overall shoreline position (BSS_{sl}, computed over the alongshore limits used for the volume changes as well, $x = [-1200, 3300]$) and one for the shoreline position in the spit area ($BSS_{sl,spit}$, computed over $x = [1500, 2800]$), used as metric for spit evolution. Each observed shoreline is compared with the chronologically closest model output to compute the BSS. However, the first two observations (August and September 2011) are excluded. This is because the BSS can lead to very low, unrepresentative scores in case the difference between the observation and the baseline prediction (the denominator in Equation 3.2) is small. This is the case after the first month, when only marginal changes from the initial shoreline are observed.

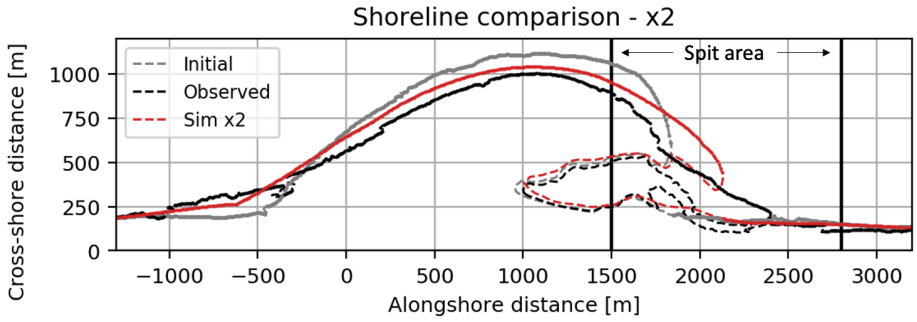

Figure 3.2: Observed (black) and predicted (red) shorelines for an arbitrary simulation (x2), in October 2012. The initial shoreline from August 2011 is also shown (grey). The solid lines show the interpolated shorelines, while the dashed lines complete the entire 0m-contours, (i.e. including the lagoon area). The two vertical black lines denote the limits of the spit area, used for the computation of $BSS_{sl,spit}$. Note that only the interpolated shorelines are used in the BSS computation.

3.1.5. Bed levels

The final output category consists of bed level changes. This might be the most straightforward way to compare simulations with observations, as the bed levels (i.e. the bathymetry) are a direct output of the model. Yet, it

forms the most complex output variable for this thesis, as it is in 2D. The performance is measured by interpolating the observed bathymetries onto the model grid and comparing the observed and predicted bed levels at each grid point. This is done for different sections of the SE. As a first indication of model performance, a bed level comparison of the entire SE area appears a logical choice. As explained, a bed level comparison is not ideal for areas with complicated three-dimensional processes, such as the spit, because of the double penalty problem. For the head section (the top part of the peninsula), however, this is less of a problem and bed levels provide a suitable performance indicator. Hence, two BSS measures are derived from the bed levels. One for the entire SE area (BSS_z), identical to the combined control area for the volume changes, and one for the head section ($BSS_{z,head}$), defined by the pink polygon in Figure 3.3.

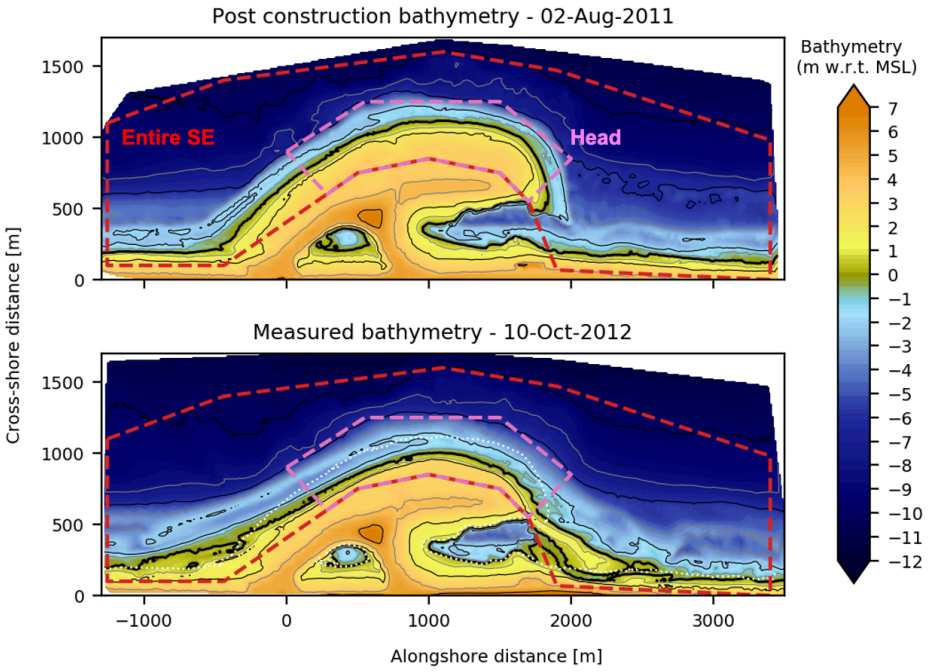

Figure 3.3: Polygons used for the computation of the bed level performance indicators. The red polygon represents the entire SE area (same as the combined polygon for the volume changes in Figure 3.1). The pink polygon is used to compute the bed levels in the head section of the peninsula only.

3.1.6. Combined performance

Next to the individual performance indicators discussed in the previous sections, it can be useful to define a combined performance score to gauge the model's overall skill. Accordingly, an average score (BSS_{tot}) is defined (Equation 3.4). Hereby, each of the three output categories is given equal weight (1/3), Hence, as the shoreline position and bed level categories consist of two BSS values, the score for the volume changes is counted twice.

$$BSS_{tot} = \frac{1}{6}\left(2 \cdot BSS_{\Delta V} + BSS_{sl} + BSS_{sl,spit} + BSS_z + BSS_{z,head}\right), \tag{3.4}$$

Using a combined performance score in such a way gives more weight to the areas of the SE which are of particular interest to this research (the spit and the head section). This is a subjective choice that will influence the results. However, this does not need to impede the findings in any way. In fact, traditional calibration processes always involve subjective choices by the modeller.

3.2. Initial parameter selection

This section introduces the initial parameter selection. The model by Luijendijk et al. (2019) is used as reference model for this study, hence, with the exception of the parameters examined in steps 2 and 3, the model configuration is equivalent to the reference configuration for all simulations.

Delft3D contains over 50 model parameters, which could in theory all be candidates for variation. As this is not feasible, the parameter set needs to be reduced. The following is an overview of 16 parameters, which have been selected for the SA based on expert consultation and literature (Table 3.1 gives a compact overview). It does not aim to give a complete overview of Delft3D parameters. Each parameter is briefly described along with the selected ranges, while Appendix A gives a more detailed description of the implementation of each parameter in the model. The parameters are categorised in four classes—hydraulic, wave, sediment, and morphology parameters, according to their implementation in the model. General information on the parameters is taken from the Delft3D-FLOW user manual (Deltares, 2011a).

3.2.1. Hydraulic parameters (H)

The hydraulic parameters describe or influence hydrodynamic processes, such as flow velocities and shear stresses. Here three parameters are considered.

- *Bottom roughness*
  The bottom roughness is specified through Chézy's coefficient, C. The coefficient can be specified in longitudinal and transverse directions, but is assumed uniform over the entire domain in this thesis. The default value is 65 $\sqrt{m}/s$ and the applied range is set to 50–80 $\sqrt{m}/s$, based on expert judgement.

- *Viscosity*
  Delft3D applies a turbulence model which requires an eddy viscosity as input. This is specified through two parameters, namely the horizontal eddy viscosity, ν_h, and the horizontal eddy diffusivity, D_h. These are sums of a constant part and a user-defined background value (ν_h^{back} and D_h^{back}, respectively). These background values are varied in this study. Their default values are set at zero, however, results from previous coastal modelling studies (e.g. Luijendijk et al., 2017) have shown that non-zero values lead to better results. The applied range has been set to 0.1–1 m^2/s for both. For the remainder of this thesis, ν_h^{back} and D_h^{back} will be denoted simply by ν_h and D_h.

3.2.2. Wave parameters (W)

The wave parameters describe processes or characteristics of the incoming waves, which force the coastline. They are specified separately for the WAVE module (SWAN) and the roller model. This thesis focuses on parameters of the FLOW module of Delft3D, hence only the roller parameters are considered for variation. Four of which are examined in this study.

- *Wave breaking*
  To model depth induced breaking, Delft3D applies the breaker index γ, which denotes the critical wave height to water depth ratio (H/h), beyond which waves start to break. As the hydrodynamic time step is smaller than the wave time-step, the water depth may change in-between wave time steps (only after a fixed amount of hydrodynamic time steps is the wave field updated). Hence, to ensure the wave height to depth ratio does not become too large, a maximum allowed value γ_{max} is applied which enforces wave-breaking on the hydrodynamic time-step level. The default value of γ is 0.55 but higher values have been used as well. Based on expert judgement the range has been set to 0.55–0.8. The same range is applied for γ_{max}, because of its physical similarity to γ and no literature is available that indicates otherwise.

- *Roller energy*
  Two additional parameters are considered, which are used in the computation of the roller energy. The wave energy dissipation coefficient, α_{rol}, and the mean slope under the roller, β_{rol}. α_{rol} is a calibration coefficient of $O(1)$ which directly scales the energy dissipated by a breaking wave, derived from the analogy between breaking waves and bores (see also Reniers et al., 2004; Roelvink, 1993). Most studies have adhered to its default value of 1, while some have examined ranges from 0–2. β_{rol} largely determines the energy transfer to and from the roller (default 0.1) (Brière et al., 2011). Both parameters directly influence the energy dissipation due to wave breaking and the roller formation. Zero-values for the two parameters also result in zero energy dissipation, and initial test simulations have confirmed that zero-values often lead to an unstable model. Therefore, the ranges have been set to 0.1–2 for α_{rol} and 0.01–0.2 for β_{rol}.

3.2.3. Sediment parameters (S)

Sediment parameters directly influence the computed sediment transport by the model. This is the largest group, with six parameters.

- *Sediment transport scaling*
  Several multiplication factors are used to scale the computed sediment transport. They are defined for suspended and bed load sediment transport, due to currents (f_{sus} and f_{bed}) and waves ($f_{sus,w}$ and $f_{bed,w}$). The factors are directly multiplied with the computed sediment transport components, hence their default values are equal to 1. It should be noted that $f_{sus,w}$ aims to represent the effect of asymmetric wave orbital motions on the sediment transport within 0.5 m of the bed, which is actually included in the bed load transport. Sediment transport due to the wave-induced alongshore current falls under the suspended transport, which is scaled by f_{sus} (see also Appendix A). Predicted sediment transport is often over-estimated and artificial erosion/sedimentation (i.e. factors larger than 1) may result in an unstable model. Therefore, the range for all four factors has been set to 0.1–1 (a value of 0 is not suitable, as it would nullify sediment transport all-together).

- *Grain size*
  The grain size of the seabed is represented by the median grain diameter, d_{50}. For morphological modelling, the value of d_{50} is of great importance, as it directly affects the sediment transport (see Appendix A). The natural d_{50} at the Delfland coast (i.e. before construction of the SE) is estimated at $\approx 250\ \mu m$ Wijnberg (2002). An analysis during construction of the SE showed an average d_{50} of 281 μm (de Schipper et al., 2016). However, Huisman et al. (2016) report significant spatial variation in the d_{50} around the SE, specifically a coarsening of sediment in front of the peninsula (+90 to +150 μm) and fining of the sediment in the adjacent sections (up to 50 μm). For this thesis, however, the d_{50} is assumed uniform over the model domain. Tonnon et al. (2018) also made Delft3D predictions of the SE and used a value of 200 μm. This seems rather low compared to the above values, so the applied range has been set to 200–350 μm.

- *Suspended sediment size*
  Next to the grain size of the bed, the representative diameter for suspended sediment is determined through the factor Fac_{dss}, which is multiplied with the d_{50} of the bed. The default value is 1 (i.e. the suspended grain size is equal to that of the seabed). It seems a reasonable assumption that the suspended grain size is not larger than the bed grain size, as this would imply that the flow is strong enough for these large grains to leave the bed and enter suspension. On the other hand, an extreme difference in grain size would also be unexpected. The applied range has been set to 0.6–1.

3.2.4. Morphology parameters (M)

Morphology parameters describe effects of the morphology (i.e. the existing bed) on the sediment transport. Three parameters are considered.

- *Bed slope*
  Delft3D applies two factors to scale the effect of streamwise (α_{bs}, default 1) and transverse (α_{bn}, default 1.5) bed level gradients on the bed load transport. For modelling bar dynamics, realistic values of α_{bs} range between 1–5 (Brière et al., 2011). In river engineering, which often includes steep banks, much higher levels are routinely used (especially for α_{bn}), but for coastal purposes the parameters are generally considered less important. To include the possibility of larger values in the SE model, the range has been set to 1–25 for both parameters. In case higher values appear to be required, this might indicate that a layered model (3D instead of 2DH) would be needed.

- *Dry cell erosion*
  Delft3D enables the erosion of dry cells (cells with a water depth below a certain threshold), by distributing (part of) the computed erosion for a wet cell over adjacent dry cells. The fraction transferred to the adjacent dry cells is specified through the dry cell erosion factor θ_{sd}. This appeared to be one of the most important model features for the SE model and the best results were acquired for $\theta_{sd} = 1$ (Luijendijk et al., 2017). The logical range for this parameter, which will be applied here, is between 0 and 1.

Cat.	Parameter	Symbol	Def.	Ref.	Range (SA)			
H	Bed roughness (Chézy) $[m^{1/2}/s]$	C	65	65	50	60	70	80
	Horizontal eddy viscosity $[m^2/s]$	ν_h^{back}	0	1	0.1	0.73	1.37	2
	Horizontal eddy diffusivity $[m^2/s]$	D_h^{back}	0	1	0.1	0.73	1.37	2
W	Breaker index	γ	0.55	0.73	0.55	0.63	0.72	0.8
	Breaker index limiter	γ_{max}		0.8	0.55	0.63	0.72	0.8
	Wave dissipation coefficient	α_{rol}	1	1	0.1	0.73	1.37	2
	Mean roller slope	β_{rol}	0.1	0.1	0.01	0.07	0.137	0.2
S	Suspended transport scaling (current)	f_{sus}	1	0.5	0.1	0.4	0.7	1
	Suspended transport scaling (waves)	$f_{sus,w}$	1	0.2	0.1	0.4	0.7	1
	Bed-load transport scaling (current)	f_{bed}	1	0.5	0.1	0.4	0.7	1
	Bed-load transport scaling (waves)	$f_{bed,w}$	1	0.2	0.1	0.4	0.7	1
	Median grain diameter $[\mu m]$	d_{50}		300	200	250	300	350
	Suspended grain size factor	Fac_{DSS}	1	1	0.6	0.73	0.87	1
M	Streamwise bedslope coefficient	α_{bs}	1	10	1	9	17	25
	Transverse bedslope coefficient	α_{bn}	1.5	15	1	9	17	25
	Dry cell erosion	θ_{sd}	0	1	0	0.33	0.67	1

Table 3.1: Overview of the initial parameter shortlist, showing the 16 selected parameters, default values (if available), applied values in the reference configuration (Luijendijk et al., 2019), and the ranges for the SA, divided into four levels (see Section 3.3).

3.3. Sensitivity analysis (EE method)

This section describes the application of the SA, by means of the EE method (see Section 2.3 for the concept). This is done to examine the relative importance of the 16 parameters presented above, and further reduce the parameter set to the most influential parameters for the UA.

3.3.1. Application

Several choices need to be made when applying the EE method, namely the number of parameters (k), the range of values for each parameter, the amount of levels for each parameter range (p), the normalised step value Δ, and the number of trajectories (r). The parameters and their respective ranges and levels are summarised in Table 3.1. p and Δ are directly related, as Δ needs to fit the applied levels. A convenient way to ensure that the sampling process guarantees equal probability for each level is to assign an even value to p combined with $\Delta = p/(2(p-1))$ (Saltelli et al., 2008). The choice of r depends on several factors and typically lies between 10 to 50 (Campolongo et al., 2007). On the one hand, a larger r is desirable as it increases the sampling resolution of the parameter space. On the other hand, each additional trajectory increases the amount of simulations by $k+1$ (the total amount of simulations is given by $r(k+1)$). Moreover, the choice of r is strongly linked to p. A larger value of p divides the parameter space into more levels, inherently increasing the resolution. However, the effective sampling resolution only improves if the increase in p is combined with a higher value of r, as a small amount of trajectories would simply leave a large part of the parameter space unexplored. In short, the choice of r and p depends on the desired sampling resolution and the available computational budget. Previous studies have indicated that values of $p = 4$ and $r = 10$ lead to reliable results (Campolongo et al., 2007; Saltelli et al., 2008, and references therein). Unfortunately, however, there is little information in literature on generally applicable combinations of r and p, given k. This study examines $k = 16$ parameters—less than some of the experiments described in Saltelli et al. (2008), which report reliable results for $p = 4$ and $r = 10$. Still, considering the computational budget, it has been decided to use $p = 4$ and $r = 20$, to achieve a higher

sampling resolution. This gives a step value of $\Delta = p/(2(p-1)) = 2/3$. Note that this means that every time a parameter is in- or decreased by Δ, it effectively skips one level (it hops from the first to the third level, for example). The 20 trajectories consist of $k + 1 = 17$ simulations each, resulting in a total of 340 simulations. The applied settings are summarised in Table 3.2.

Parameter	Symbol	Value
Number of parameters	k	16
Number of trajectories	r	20
Number of levels	p	4
Normalised step value	Δ	2/3
Normalised levels		[0, 0.33, 0.67, 1]
Total amount of simulations	$r(k+1)$	340

Table 3.2: Summary of the settings for the EE method applied in this thesis.

The python package SALib (Herman et al., 2019) is used to draw 20 random sampling points in the 16-dimensional unit hypercube, their coordinates all equal to one of the four levels ($[0, 0.33, 0.67, 1]$). These form the base vectors of the 20 trajectories. The components of each base vector are then in- / decreased by Δ one-by-one, in randomised order, resulting in 16 additional vectors per trajectory, whereby any two consecutive vectors differ in exactly one component. Once all the trajectories are created, the vector components (representing coordinates in the 16-dimensional unit hypercube) are transformed to the respective parameter ranges, to give the parameter sets for the simulations (Table 3.1).

3.3.2. Computation of the EE's

Once all the simulations are done, the EE's of each parameter can be computed for the different model outputs. Since the normalised value of Δ is constant for all parameters and the prior parameter distributions are all uniform, the inclusion of Δ in the EE computation (Equation 2.1) does not affect the relative influence of the parameters. Hence, it is eliminated from the computation of the EE's. The exact computation differs slightly per output category. For the volume changes, the EE is computed for every individual control area, and only the final time step is used. This results in one value for the volume change of each control area, which can be subtracted between subsequent simulations to give the EE. The exact formula then depends on whether the corresponding parameter between two simulations has been in- or decreased by Δ (Equation 3.5).

$$EE_i^{\Delta V,j} = \begin{cases} \Delta V_j\left(\vec{x}^{(l+1)}\right) - \Delta V_j\left(\vec{x}^{(l)}\right) & \text{in case of increase by } \Delta \\ \\ \Delta V_j\left(\vec{x}^{(l)}\right) - \Delta V_j\left(\vec{x}^{(l+1)}\right) & \text{in case of decrease by } \Delta, \end{cases} \tag{3.5}$$

where $EE_i^{\Delta V,j}$ denotes the EE associated with parameter X_i and control area j, ΔV_j denotes the cumulative volume change for control area j, and $\vec{x}^{(l+1)}$, $\vec{x}^{(l)}$ are the vectors representing the two parameter sets between which X_i was in- / decreased by Δ (note that, because of the randomised order, i and l are not necessarily the same). Hence, for a strictly monotonic model response, the EE's would not switch signs, regardless of an in- or decrease of the parameter (in which case μ and μ^* would be equal).

The computation of the EE's for the shoreline position and bed level change requires a different approach, as the outputs are defined over multiple positions in the model domain. The EE computation requires a single metric representing the model output, which can be compared between two simulations. Simply choosing one position for the shoreline position / bed level change (analogous to taking the final time step for the volume changes) would not make sense, however. Therefore, it has been decided to represent the EE's by the mean absolute error (MAE) between the shoreline positions / bed level changes of two simulations (Equations 3.6 and 3.7). This allows the outputs to be represented by a single number and is the closest such metric to the EE definition in Equation 2.1 (the MSE, for example, would inflate the EE's for more influential parameters because of the squaring of the errors).

$$\mathrm{MAE}(Y_1, Y_2) = \frac{1}{n}\sum_{i=1}^{n}\left|Y_{1,i} - Y_{2,i}\right| \tag{3.6}$$

$$EE_i^{Y,j} = \mathrm{MAE}\left(Y_j\left(\vec{x}^{(l)}\right), Y_j\left(\vec{x}^{(l+1)}\right)\right) \tag{3.7}$$

where $Y_{1,2}$ denote two datasets which are compared, n denotes the amount of data points in each set, and $EE_i^{Y,j}$ denotes the EE associated with parameter X_i and output Y_k (shoreline position or bed levels). The application of the MAE makes a distinction between μ and μ^* irrelevant (the MAE already takes the absolute values of the errors, hence $\mu = \mu^*$). Still, for the sake of consistency, the notion of μ^* will be used for all EE comparisons.

3.3.3. Evaluation

Once the EE's have been computed, the sensitivity measures can be derived and the parameters can be compared (see Section 2.3.2). Hereby, the most important measure is μ^*, representing the absolute influence of the parameter on the model output (Campolongo et al., 2007). This will be the main selection criterion to determine the parameters for the GLUE method. Hereby, the value of μ^* itself is not so important, but rather its relative value for each parameter. The values of σ give insight into a parameters non-linearity or interaction with other parameters. The examined parameters are plotted in μ^*-σ space, showing the parameters' influence on the x-axis and non-linearity / interdependence on the y-axis (Figure 3.4).

It is difficult to draw absolute conclusions from the σ values. A large σ would imply interaction effects with other parameters or a non-linear model response, but does not provide any specific information on these effects. Plotting the variables in $\mu - \sigma$ space, Morris (1991) suggested drawing two lines corresponding to $\mu \pm 2$ standard errors of the mean, to form a 'wedge'. Parameters inside this wedge can be considered to interact significantly with other parameters. This does not hold when using μ^*, however. An alternative criterion in $\mu^* - \sigma$ space would be to use the $\mu^* = \sigma$ line (Khare et al., 2019). All parameters above the line can be considered to be involved in significant interactions or non-linearities, while the ones below the line are not. The $\mu^* = \sigma$ line should not be seen as a hard criterion, however, but rather as a guideline (e.g. parameters that are close together, but just on different sides of the line, will have similar interaction effects). Another use of σ lies in the identification of unstable simulations. An unusually high value for σ may be caused by an outlier among the EE's (e.g. due to an unstable simulation). Hence, when encountering high σ-values (far above the $\mu^* = \sigma$ line), it is useful to check the corresponding EE's for outliers.

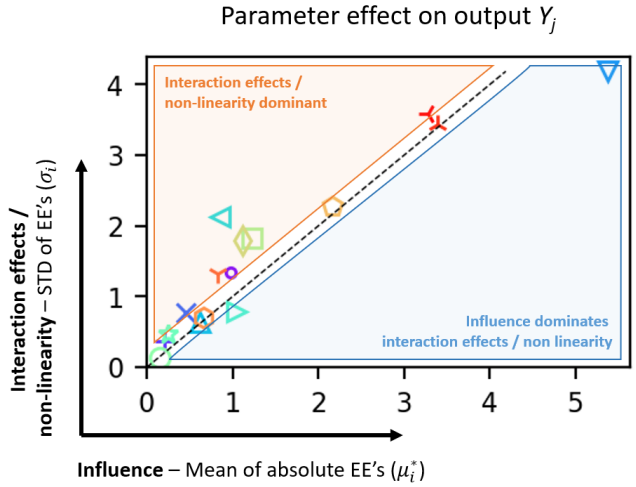

Figure 3.4: Example plot of a parameter comparison by the EE method, plotting the mean (μ^*) and standard deviation (σ) of the EE's for every input parameter, measured to output variable Y_j. The dashed line represents the $\mu^* = \sigma$ line which can be used as a criterion for strong interaction effects (Khare et al., 2019).

3.3.4. Performance analysis

The aim of the SA is to select a reduced parameter set for the GLUE analysis. But in addition, the data of the 340 simulations can be used to carry out an initial performance analysis, to get an idea of the how the applied parameter ranges translate to model performance. This information can be used to potentially adjust the ranges of the parameters selected for the GLUE analysis and determine the fixed values of the remaining parameters. Therefore, the BSS indicators outlined in section 3.1 are computed for each simulation, to analyse their relation to the corresponding parameter values. It should be noted, however, that this should only be used as an initial indication. For any conclusions from this performance analysis to be statistically rigorous, the sample size is too small and unevenly distributed over the parameter space.

3.4. Uncertainty analysis (GLUE)

The main step of this research is the uncertainty analysis, which will be carried out by means of the GLUE method. The concept of this method is explained in Section 2.4.2 but its application requires several user-defined decisions, elaborated in this section.

3.4.1. Overview

The steps encompassing the GLUE method can be summarised as below (see also Figure 2.6). The components of each step are explained in more detail in the coming sections. Step 1 will not be addressed in this section as it essentially encompasses the SA, which is explained in Section 3.3.

1. Select parameters and assign prior distributions
2. Generate a large number of unique parameter sets and run simulations
3. Determine a likelihood measure and set a behavioural threshold
4. Compute the likelihood scores for each parameter set
5. Derive uncertainty bounds and posterior parameter distributions

3.4.2. Sampling strategy

An important aspect of the GLUE method is the sampling strategy, by which the parameter sets are drawn. The accuracy of the GLUE method depends on the sampling resolution, which is determined by the size of the parameter space and the amount of samples drawn from it. The possible amount of samples is determined by the available computational capacity for the model runs, which is often a problem for models with large simulation times. In the original GLUE the samples are drawn by MC sampling (Beven & Binley, 1992). However, MC leads to so-called 'clusters and gaps' in the sample (i.e. spots with high / low density of sample points respectively). To obtain a representative uncertainty estimation, it is desirable to efficiently sample the parameter space, within the limits of the available computational capacity. This dilemma has led to several studies examining possibilities to use the computational capacity more efficiently (see also Section 1.2). While there are different approaches (e.g. Berends et al., 2019; Teng et al., 2017, and references therein), this thesis focuses on improving the sampling strategy, to optimise the distribution of the samples over the parameter space.

All three GLUE studies in coastal modelling use MC sampling (Table 3.3). Ruessink (2005) applies a simple approach, using a theoretical sampling resolution to determine the required amount of simulations: to sample the k-dimensional parameter space with an average resolution of $f_s = 1/a$ of each parameter range, one needs to perform a^k model runs (in his case, a desired resolution of $1/8$ with four parameters gives $8^4 \approx 4100$ simulations). This is, however, only an average sampling resolution over the entire parameter space—as MC sampling is random, the resolution may show significant local deviations (clusters and gaps of sample points). Simmons et al. (2017) propose to test convergence of parameter statistics, by plotting the mean and variance of each parameter against the number of simulations. Once these statistics converge, a sufficient amount of simulations is carried out. These studies, however, use computationally cheaper models with the capacity to run multiple thousands of simulations.

Uhlenbrook and Sieber (2005) apply Latin Hypercube (LH) sampling, which appears to be very efficient, as the study examines 37 parameters with only 400 simulations. They describes LH as an effective method for a GLUE

study with many parameters but limited computational capacity. LH divides the parameter space in a predetermined amount of sub-spaces, from which the samples are drawn. This leads to a more uniform sampling resolution than MC, but decreases the 'randomness' of the sample. The downside of this is that it is not possible to derive Sobol' indices (Saltelli et al., 2008), as in Kroon et al. (2020), which is required for additional work that the here presented computations will be used for, outside this thesis.

Field	Source	Model	Sampling	k	N
Coastal	Ruessink (2005)	Unibest-TC	MC	4	4100, 1001
	Simmons et al. (2017)	Xbeach	MC	6, 4	15000, 15000
	Kroon et al. (2020)	One-line model	MC	1	200,12000
Hydrology	Uhlenbrook and Sieber (2005)	TAC^D	LH	37	400

Table 3.3: Overview of several past studies applying GLUE. The table shows the applied sampling method, the number of varied parameters (k) and the sample size (N). Note that this is not a complete list of studies which have applied GLUE—especially in hydrological modelling, GLUE has been applied in numerous additional studies. The research by Uhlenbrook and Sieber (2005) is included here as its sampling method is of interest to this study.

The sampling strategy used in this thesis is based on the average sampling resolution used by Ruessink (2005), but not combined with MC sampling. To minimise local deviations from the average resolution, the samples are drawn by a quasi-random sampling method: the Sobol' low-discrepency sequence (Sobol', 1967), which does allow the computation of Sobol' indices. Low-discrepancy sequences (LDS), as the name indicates, are sequences with a low *discrepancy*—a measure for the deviation of sampled points from the uniform distribution (Kucherenko et al., 2015). In other words, they are quasi-random sampling algorithms, which distribute the samples more evenly over the parameter space. Because of this they have a convergence rate of up to $O(N^{-1})$, whereas for MC this is $O(N^{-1/2})$. Of the few well-known LDS, the Sobol' sequence has been proven by several studies to be superior to other LDS (e.g. Glasserman, 2003; Tuffin, 1996). For a more detailed description of the concept of LDS and the Sobol' sequence in particular, the reader is referred to Kucherenko et al. (2015) and Sobol' (1967).

The computational budget for this study allows ± 1000 model runs for the GLUE analysis. This limit is used to find a feasible combination of sampling resolution and number of parameters (Table 3.4). On the one hand, the more parameters are varied, the lower the sampling resolution. The results will then capture a larger part of the total parameter uncertainty (more degrees of freedom), albeit at the cost of less detail in the parameter space. At the other extreme, a very high resolution combined with only few parameters will lead to a more accurate uncertainty estimation per parameter, but with less degrees of freedom in the model it is more difficult to capture the total parameter uncertainty. Moreover, it introduces additional bias as model parameters that are not varied are assigned fixed, user-defined values. Ultimately, a decision was also made under consideration of the SA results presented in Chapter 4: 1024 samples are drawn for five parameters, resulting in an average sampling resolution of $1/4$.

# of sims		# of parameters (k)				
		3	4	5	6	7
sampling res.	1/3	27	81	243	729	2187
	1/4	64	256	1024	4096	16384
	1/5	125	625	3125	15625	78125
	1/6	216	1296	7776	46656	279936
	1/7	343	2401	16807	117649	823543
	1/8	512	4096	32768	262144	2.10E+06

Table 3.4: Overview of the required amount of simulations for a given combination of the number of parameters and the average sampling resolution. The GLUE analysis in this study is limited to ± 1000 model runs. Green cells indicate possible combinations, while red combinations are ruled out. Note that the sampling resolution is not discrete, i.e. if more or less simulations are carried out than the value in the table, the resolution will lie in-between the given values.

One condition for the improved convergence rate of the Sobol' sequence is that the sample size must be sufficiently large. There is no generally applicable number for 'sufficiently large', but there are several tests to confirm the superiority of a give Sobol sample. These have been checked and confirmed for a set of 1024 samples drawn with the Sobol' sequence against ten MC samples of the same size (see Appendix C for the computation).

3.4.3. Likelihood measure

Once the simulations are graded by the BSS metrics, the behavioural threshold determines if they are behavioural or not (see next section). The BSS's of the behavioural simulations are then transformed into a likelihood measure, representing each parameter set's likelihood of being the optimal set (for a certain output variable). All non-behavioural runs automatically receive a likelihood of zero. A practical likelihood measure, which rescales the BSS to sum to one is given by (Simmons et al., 2017):

$$L_i^Y = \frac{BSS_i^Y}{\sum_{i=1}^{n} BSS_i^Y}, \tag{3.8}$$

where L_i^Y and BSS_i^Y are the likelihood measure and BSS of simulation i for output variable Y and n is the total number of behavioural runs. Additionally, it can be useful to combine likelihood measures if different output variables are considered. A combined likelihood then provides a way to find the most reliable parameter sets over all output variables. Simmons et al. (2017) suggest a formula to compute a combined likelihood CL, which results in 0 for every simulation that is non-behavioural in any of the considered outputs.

$$CL_i = \left(\prod_{Y=1}^{N_L} L_i^Y \right)^{\frac{1}{N_L}}, \tag{3.9}$$

where N_L denotes the total number of likelihood measures (i.e. the number of considered output variables). Hence, any likelihood measure $L_i^Y = 0$ for a given simulation i will lead to $CL_i = 0$. As the combined likelihood provides a more practical way to assess a simulation's likelihood over the different outputs, BSS_{tot} is not used for the GLUE analysis.

3.4.4. Behavioural threshold

Another integral part of the GLUE method, which strongly influences the results, is the determination of the behavioural threshold. This is another subjective choice that the user of the GLUE method needs to make. The goal is to separate *good* or *acceptable* simulations from those deemed unacceptable. There are proposed classifications to grade morphological simulations based on the achieved BSS (e.g. van Rijn et al., 2003; Sutherland et al., 2004), but the general applicability of these classifications to specific output variables is debatable. This is because the BSS not only depends on the prediction accuracy of the considered simulation, but also on the difference between the observations and the baseline prediction (the denominator in the BSS computation). Additionally, output variables that are derived by aggregation of data over different dimensions (e.g. the cumulative volume changes) tend to result in a higher BSS than output variables which are compared point-wise (e.g. the spatially distributed bed levels). In other words, a certain BSS value does not necessarily imply the same prediction skill for different outputs. In this study, for example, the volume changes result in much higher scores than the bed levels (Table 3.5). This makes it undesirable to apply the same BSS threshold to all considered output variables. Fore example. a threshold of 0.6 would render 779 runs behavioural for the volume changes, but only 92 for the bed levels.

Ruessink (2005) finds that GLUE may be inefficient, as much computational time is wasted on non-behavioural simulations. However, setting the threshold so low that (nearly) all simulations become behavioural would undermine the purpose of the GLUE method to find the most likely parameter sets, as all values would be accepted to some degree. Hence, unless the modeller knows the 'real' parameter distributions beforehand, it seems that the GLUE method depends on a certain amount of model runs being discarded as non-behavioural. A different approach to finding a suitable threshold would be to look at the BSS from a more relative perspective, and determine a desired fraction of behavioural model runs. This is particularly useful when a meaningful threshold for multiple output variables is non-existent (Simmons et al., 2017). Especially when computing combined likelihoods over different output variables by multiplication (e.g. using Equation 3.9), care must be taken that

Classification	Range	BSS_{tot}	$BSS_{\Delta V}$	BSS_{sl}	$BSS_{sl,spit}$	BSS_z	$BSS_{z,head}$
Max		0.81	0.95	0.79	0.74	0.64	0.88
Min		0.13	0.08	0.1	0.14	0.1	0.02
Mean		0.59	0.74	0.54	0.52	0.44	0.59
Excellent	0.8–1	8	486	0	0	0	168
Good	0.6–0.8	576	293	455	292	92	414
Fair	0.3–0.6	381	227	493	670	773	334
Poor	0–0.3	59	18	76	62	159	108
Bad	< 0	0	0	0	0	0	0

Table 3.5: Overview of the BSS performance of the 1024 GLUE simulations. The first three rows show the maximum, mean, and minimum BSS for the respective output variables. The following rows give an overview of the amount of simulations per grading category, based on the classification by van Rijn et al. (2003). Not that this merely serves as an indication of how the BSS values are distributed and how they would compare with the applied classification. It is not intended as a concluding classification of the results in this study.

the thresholds for individual output variables are not too strict, since a single zero-likelihood also results in a combined likelihood of zero. Therefore, in this study the behavioural threshold for each output is set such that around half the simulations (500–600) are deemed behavioural. The resulting thresholds and corresponding amount of behavioural runs are summarised in Table 3.6.

Output	Score	Threshold	n
Cumulative volume changes	$BSS_{\Delta V}$	0.75	587
Shoreline position	BSS_{sl}	0.55	586
Shoreline position spit area	$BSS_{sl,spit}$	0.52	589
Bed levels, entire SE	BSS_z	0.43	574
Bed levels, head section	$BSS_{z,head}$	0.6	588

Table 3.6: Overview of the applied behavioural threshold for each output variable and the resulting amount of behavioural runs (n).

3.4.5. Evaluation

The computed likelihoods for each parameter set are used to analyse the GLUE simulations statistically. The results are divided in three categories: Uncertainty bounds, spatial uncertainty maps, and parameter optimisation. The following paragraphs explain each category in more detail.

Uncertainty bounds

By combining the results of all behavioural model runs, uncertainty bounds can be derived for the various model outputs specified in Section 3.1. Uncertainty bounds are often visualised using a confidence interval (CI). A CI presents a range of plausible values for a stochastic variable, and is accompanied by a confidence level. For example, the 90% CI gives a range of values for a stochastic variable, with a confidence level of 90% that the true value lies within this range. A CI can be computed using the empirical data directly, or by deriving it from a statistical distribution fitted to the dataset. As a general rule, the larger the dataset, the better the estimate of the CI computed from it.

If the empirical dataset is small, uncertainty is often assumed to be normally distributed. This makes it easy to compute any desired CI, with only the mean and standard deviation of the samples. In this study, however, it is assumed that the GLUE dataset is large enough to compute the CI's empirically. This is done by constructing an empirical cumulative distribution function (CDF) from the simulation outputs, weighted by the likelihood scores (e.g. the bed level changes at a certain location and time). The empirical CDF is then interpolated onto the desired percentiles, to give the bounds of the corresponding CI. For example, the upper and lower bounds of the 90% CI are obtained from the 95^{th} and 5^{th} percentiles of the empirical CDF, respectively. The example in Figure 3.5 shows the weighted histogram and empirical CDF for the cumulative bed level change at a location in

the spit area. As can be seen, the distribution differs significantly from a normal distribution. This has also been checked for other locations and outputs, confirming that the uncertainty is not normally distributed.

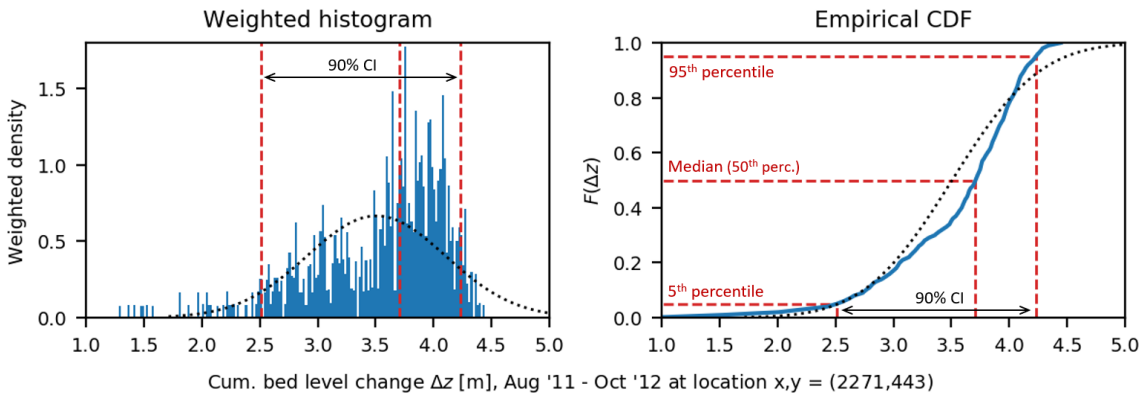

Figure 3.5: Weighted histogram (left panel, blue bars) and empirical CDF (right panel, blue line) derived from the 1024 GLUE simulations for cumulative bed level changes from Aug '11 to Oct '12 at an arbitrary location in the spit area. The red, dashed lines show the the median and 90% CI from the empirical CDF. The histogram is created from the behavioural GLUE simulations, weighted by the corresponding likelihood scores. The corresponding normal distribution is plotted as a reference in both panels (black, dotted line).

Spatial uncertainty maps

One way to communicate the impacts of prediction uncertainty and consider it in decision-making processes, would be to visualise where (geographically) the predictions are the most uncertain. Here, this is done by looking at the uncertainty in the predictions of bed level changes throughout the SE area. This is essentially an extension of the uncertainty bounds in 2D. Often, uncertainty is expressed as the spread of a variable's change relative to the magnitude of the expected change. For example, an uncertainty indicator (UI) could be based on the mean and standard deviation of the sample (Equation 3.10) (van der Wegen & Jaffe, 2013). If the standard deviation is larger than the mean for a certain area, UI will be negative and the predictions in this area are deemed uncertain.

$$\mathrm{UI} = 1 - \frac{\sigma^2}{\mu^2} \tag{3.10}$$

Such approaches are useful if one wishes to identify areas where the variability of the predicted change is large compared to the mean. This, however, results in extreme uncertainty in areas with little expected change, but relatively high variability. On the other hand, areas where the variability is in the same order as the expected change, will be identified as more certain, even though the absolute variability might be significantly larger than for areas identified as uncertain. The following gives a more intuitive example. Say we have two distinct areas around the SE. In area 1, the expected bed level change (mean) is 5 m and the standard deviation is also 5 m. In area 2, the expected change is 0.005 m (5 mm) and the standard deviation is 0.05 m (5 cm). Using an uncertainty indicator as described above, area 2 would be identified as highly uncertain (the standard deviation is 10 times larger than the mean), whereas area 1 would be significantly less uncertain in comparison (the mean and standard deviation are equal). The fact remains, however, that the actual bed level change in area 1 might deviate from the expected value by a factor of $O(100)$ more than in area 2. Hence, for stakeholders and decision-makers, the possibility of a large absolute deviation from the expected outcome (area 1) might be of more interest than that of a large relative deviation (area 2).

For this study, it is assumed to be of more interest to identify areas with a large absolute variability. To create the spatial uncertainty maps, the median and 90% CI of predicted bed level changes are computed at grid cell level. This is done using the behavioural simulations, determined by the threshold for BSS_z. The width of the 90% CI (W_{CI90}) is used as uncertainty indicator. The computed W_{CI90} per grid cell can then be interpolated over the SE area, resulting in the spatial uncertainty map.

Parameter optimisation

The likelihood scores of the parameter sets can be used to transform the prior (uniform) parameter distributions to marginal posterior distributions. This is done by creating a weighted histogram of parameter values based on

the likelihood values corresponding to each simulation (similar to the computation of the uncertainty bounds, but for the parameter values instead of the output variables). The result is a likelihood distribution for each parameter, representing the likelihood of each value in the range to be optimal, based on the GLUE analysis. The posterior distributions can then be compared to the prior distributions and default or reference values for the parameters. Moreover, they can be used to estimate an optimal parameter set (OPS), which represents the parameter set with the highest likelihood. The OPS can then be simulated as well, to check if the parameter set is indeed optimal.

3.5. Uncertainty comparison

The aim of the uncertainty comparison is to compare the quantified parameter uncertainty with uncertainty arising from variability in the wave climate (intrinsic uncertainty), to get an idea of the relative importance of the two. To this end, eight extra simulations will be carried out, using the OPS estimated from the parameter optimisation. The wave forcing, however, will be varied for each simulation. The spread in the results of these simulations can then be compared to the spread in the GLUE results. Two main approaches present themselves to simulate variations in the wave climate.

1. **Select alternative years from historical data**
   Here, historical wave data is used to simulate alternative wave climates. In other words, the simulations represent scenarios in which the SE would have been constructed in a different year.

2. **Bootstrapping**
   In this approach, a statistical analysis of the available wave data is carried out, after which the statistics are used to create a synthetic dataset of wave conditions. This is commonly applied when a large data set is desired (i.e. more samples than the available data provides), because it allows the creation of unlimited samples from the obtained wave statistics (e.g. Kroon et al., 2020).

It is important that only the wave conditions are varied among the simulations—all other model settings should be identical. This means that the relative simulation period should remain equal (14 months, from August until October of the next year) and the same morphodynamic acceleration technique should be applied (the BFFC approach combined with *morfac* = 3). Filtered and compressed wave series from the offshore locations (Europlatform and IJmuiden) are available for 25 years, from 1989 to 2014. Thus, the dataset is large enough to select eight wave years and bootstrapping is not required.

One of the eight simulations will use the original wave year (August 2011 – October 2012), as the fixed parameter is not among the GLUE simulations. For the remaining simulations, seven different wave years are selected from the available 25-year span. Ideally one would like to simulate years with different characteristics, to get a more representative estimate of the possible spread in the results. The main driving force behind the morphological changes at the SE is the alongshore current due to wave forcing, which mainly depends on the wave energy and direction (Luijendijk et al., 2017; de Schipper et al., 2016). Hence, not only the total amount of wave energy can play a role, but also the dominant wave direction. Here, two main directional sectors are distinguished, separated by the shore normal at $311^{\circ}N$. Waves from the northern sector ($\approx 311^{\circ} - 41^{\circ}$) lead to a Southwest current, whereas waves from the western sector ($\approx 221^{\circ} - 311^{\circ}$) lead to a Northeast current. All wave conditions outside these two sectors move away from the SE and are neglected.

A comparison of the wave energy characteristics for each year is made, using a time series of transformed nearshore wave conditions at the 10 m depth contour, close to the SE. A detailed description of how the energy characteristics are computed is given in Appendix B, along with Figures of the full 25-year wave series (Figure B.1) and corresponding energy characteristics (Figure 3.6). The energy characteristics are compared over each 14-month time period, distinguishing between the energy coming from the northern and western sectors. Note that each period is denoted by the year in which the simulation starts (e.g. 1989 represents the wave conditions from August 1989 to October 1990). Additionally, for each year, the same comparison is made for the first seven months only, from August until February. As will be shown in Chapter 5, most morphological changes occur in this time period, whereas significantly less happens in the period from February to October. Hence, it is also useful to compare the characteristics for the first seven months.

Each year is classified based on its total wave energy density and dominant direction. The energy is classified as high (energy density close to maximum), medium (close to mean) and low (close to minimum). The dominant direction is classified as West, North or equal (=), based on the proportion of energy coming from each sector.

Note that, in general, the dominant wave direction in this area is from the Western sector, especially during storm conditions (Wijnberg, 2002). Therefore, only few years have a larger proportion of wave energy coming from the North sector. The classification is used to select seven wave years with differing characteristics: 1989, 1990, 1995, 2000, 2003, 2004, 2007, and the original year, 2011. The energy characteristics and classification of the selected wave years are presented in Table 3.7 (Appendix B shows the same for the full 25 year series).

Note that the GLUE analysis with 1024 simulations provides a more complete statistical representation of parameter uncertainty than the eight wave year simulations do for intrinsic uncertainty. Therefore, this comparison should be seen as a first qualitative indication of the relative importance of the two uncertainty sources, not as an absolute quantitative comparison.

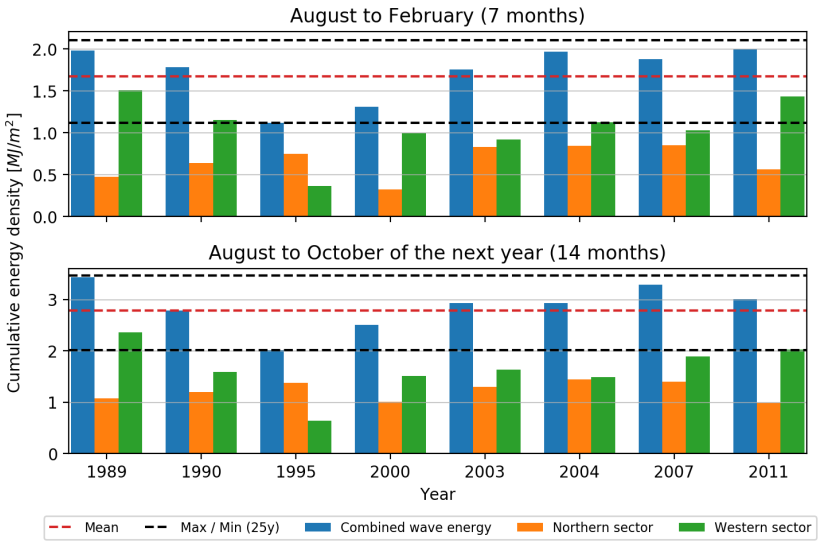

Figure 3.6: Bar plots showing an estimate of the cumulative, three-hourly averaged wave energy density per selected wave year (see Appendix B for the computation). The values are computed over the first seven months (Aug–Feb, upper panel) and the total 14-month study period (Aug–Oct, lower panel). The colours represent the energy coming from the combined directional sectors (blue), the northern sector (orange), and the western sector (green). The dashed lines show the maximum, minimum and mean (red) values for the combined conditions. Note that 2011 represents the original wave series (Aug 2011 – Oct 2012). Figure B.2 shows the same comparison for the full 25-year series.

Year	Simulation period	Energy		Dominant direction	
		Aug – Feb	Aug – Oct	Aug – Feb	Aug – Oct
1989	Aug 1, 1989 - Oct 16, 1990	high	high	West	West
1990	Aug 1, 1990 - Oct 16, 1991	medium	medium	West	West
1995	Aug 1, 1995 - Oct 16, 1996	low	low	North	North
2000	Aug 1, 2000 - Oct 16, 2001	low	medium	West	West
2003	Aug 1, 2003 - Oct 16, 2004	medium	medium	=	=
2004	Aug 1, 2004 - Oct 16, 2005	high	medium	West	=
2007	Aug 1, 2007 - Oct 16, 2008	medium	high	=	West
2011	Aug 1, 2011 - Oct 16, 2012	high	medium	West	West

Table 3.7: Overview of the selected wave years for the uncertainty comparison (including the original year, 2011). The years are classified based on total energy (low, medium or high) and dominant wave direction (West, North or equal).

4

Results – Sensitivity Analysis

Chapter outline
This chapter describes the results of step 2 of this thesis, the sensitivity analysis. The goal is to determine the most influential model parameters using the EE method. In Section 4.1 the model results from the 340 simulations are examined, giving a first indication of the model sensitivity to the 16 parameters. Then follows a detailed evaluation of the EE method in Section 4.2, to determine which parameters have the largest influence on the selected model outputs. Finally, in Section 4.3 the parameter selection for the GLUE method is discussed. This includes a performance analysis of the 340 model runs, which is used to adjust some of the selected parameter ranges and the fixed values of the parameters which will not be varied in the GLUE method.

4.1. Model results

This section provides an initial analysis of the spread in the results of the 340 SA simulations. Hereby, it is important to note that the presented bandwidths are purely based on literature and expert judgement and not constrained by the observations (as they will be during the GLUE analysis). Hence, to avoid confusion, the bandwidths presented in this chapter will be labelled as sensitivity, not uncertainty. This sensitivity can offer valuable information. For instance, it can be checked if the observations fall within the sensitivity bounds, to confirm that the selected parameter ranges include reasonable values. Moreover, it facilitates the identification of general model tendencies and potential outliers or unstable simulations.

4.1.1. Volume changes

First, the volume changes are considered. The empirical CDF of the 340 simulations is be used to derive the median and the 90% and 50% CI's for each control area (Figure 4.1). Note that the different areas are plotted on the same scale, to visualise the differences in the magnitudes of predicted changes. In fact, the predicted changes in the middle section (upper right panel) are up to seven times larger than for the South section (lower left panel).

An encouraging sign is that all of the observations fall within the 90% CI of the model runs, with the exception of one observation for the entire area (upper left panel). Most points, except for the entire area, also fall within the 50% CI. Judging from the median, the trends of the simulations appear to follow the trend of the observations well, as the main morphological activity occurs in the storm season of December and January (again, less so for the entire area). These observations suggest that the selected parameters and ranges include values that allow a reasonable approximation of the measurements.

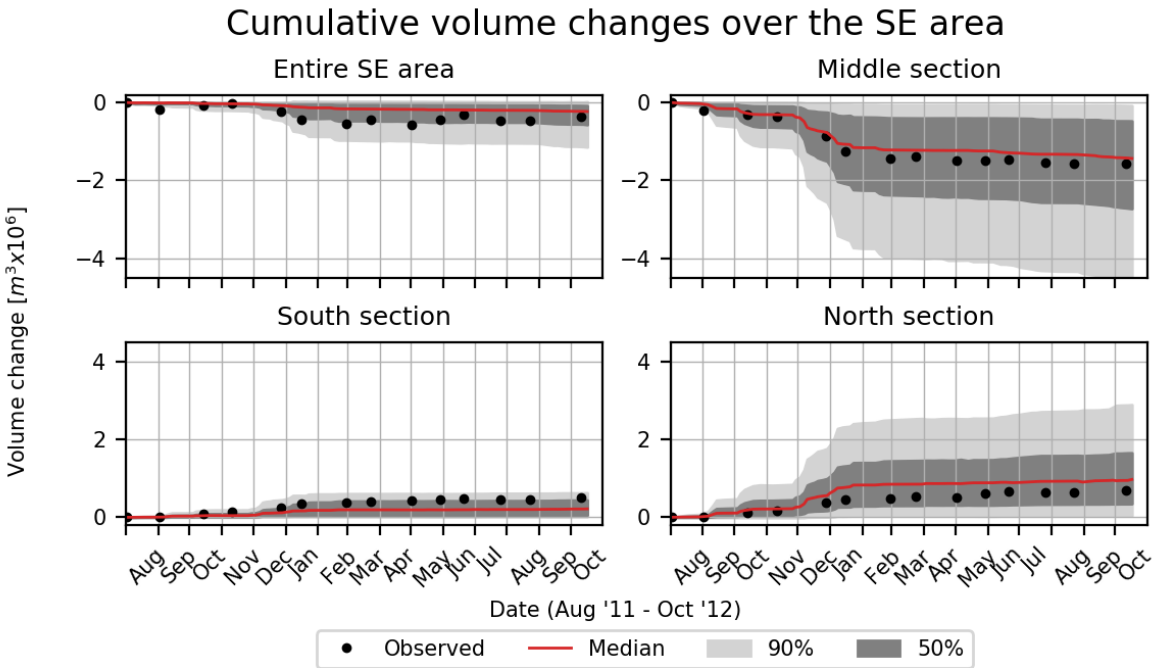

Figure 4.1: Time series of cumulative volume changes around the SE from August 2011 to October 2012, for each of the four control areas (in million m^3). The ensemble of 340 SA simulations is represented by the median (red line) and 90% and 50% CI's (light and dark grey areas respectively), along with the observed volume changes (black dots) (de Schipper et al., 2016).

The middle section appears to be modelled best: the median of the model runs follows the observations closely. For the North and South sections, the median follows the trend of the observations, but is less accurate. It appears the model has a general tendency to overestimate accretion in the North and underestimate it in the South, which can be derived from the relative position of the observations within the CI's: the 95^{th} percentile (upper bound of the 90% CI) overestimates the accretion by a factor ≈ 4 in the North compared to ≈ 1.2 in the South. This also follows from the width of the 90% CI (W_{CI90}), which represents the absolute sensitivity, and shows that the variation is by far largest in the middle section ($W_{CI90} \approx 4.5$ million m^3), followed by the North

section ($W_{CI90} \approx 2.9$ million m^3). This is expected to some degree, as these are the main areas of morphological change (erosion off the peninsula and the spit formation in the North section). Yet, the relative magnitude of the observed changes in the North and South sections differs significantly less than for the predicted changes. Hence, the model struggles to correctly predict the relative distribution of sediment volumes over the North and South sections.

Next to the absolute sensitivity, we can also consider the ratio between W_{CI90} and the median ($Med_{\Delta V}$) to get an indication of the relative model sensitivity around the median (Table 4.1)—the higher the ratio, the higher the relative model sensitivity. The reative sensitivity is highest for the entire area, whereas it is approximately equal for the three individual sections. In general, the model struggles most with the predictions for the entire SE. This is in line with the findings of Luijendijk et al. (2019), and might be partially caused by omitted processes, such as aeolean transport and sediment sorting.

A final, interesting observation is that the CI's are asymmetric, implying significant deviation from the normal distribution. This supports the use of the empirical CDF to compute the CI's. Moreover, for all areas except the entire SE, the 5^{th} percentile does not cross the zero-line. This is encouraging, as it implies that unrealistic simulations, which predict erosion in the South / North sections or accretion in the middle section, are either non-existent or limited to small numbers.

Section	W_{CI90}	$Med_{\Delta V}$	$\frac{W_{CI90}}{Med_{\Delta V}}$
SE	1.1	0.2	5.5
Middle	4.5	1.4	3.2
South	0.6	0.2	3
North	2.9	1	2.9

Table 4.1: Overview of the width of the 90% CI (W_{CI90}), the median volume change ($Med_{\Delta V}$), and the ratio between the two. Values are given in million m^3.

4.1.2. Shoreline position

The shoreline position adds another spatial dimension as it is defined over multiple alongshore positions. This makes it more difficult to visualise a time series of shorelines. Since the results presented here are only used as a first indication, only the final time step (October 2012) is considered, for now.

A first look at the 340 predicted shorelines (Figure 4.2, upper panel) offers the possibility to identify potential outliers. The majority of predictions appear to stay within a reasonable range, however (say, within 200–300 m of the observations). This is not necessarily true in the spit area, but in this area more variation is expected due to the complex processes governing the spit formation. Some simulations stand out as they predict the shoreline to extend seaward over the 14 months, which is uncharacteristic (grey lines in the upper panel, which extend further seaward than the initial shoreline).

Variation is largest in the spit area ($x \approx [1700, 2500]$), where the 90% CI covers around 700 m in cross-shore position, at its widest point. Hence, it appears the spit formation is a complex process, which is sensitive to the parameter values. The variation is also quite large in the head section of the peninsula, with W_{CI90} covering up to 350 m. The South section (up to $x \approx -450m$) is the least sensitive, the CI's being significantly narrower there (100–200 m). The observed shoreline is generally more retreated than for most simulations and, in some areas, falls outside the 90% CI, suggesting that the model generally diffuses the shoreline too little. This may be (partly) explainable by the fact that Delft3D still omits some processes, such as aeolian transport and potential 3D cross-shore processes (e.g. sediment sorting) (Huisman et al., 2016; Luijendijk et al., 2017).

Overall, it can be concluded from the model results that the selected parameters and ranges should include reasonable values, as the derived CI's capture the majority of observations. This is encouraging and suggests the ranges are sufficiently broad. On the other hand, there are also some unrealistic simulations (e.g. predicting a seaward movement of the shoreline in the middle section, or near zero volume change after 14 months). This reflects the fact that the parameter ranges are based purely on literature and expert judgement and may also include unrealistic values.

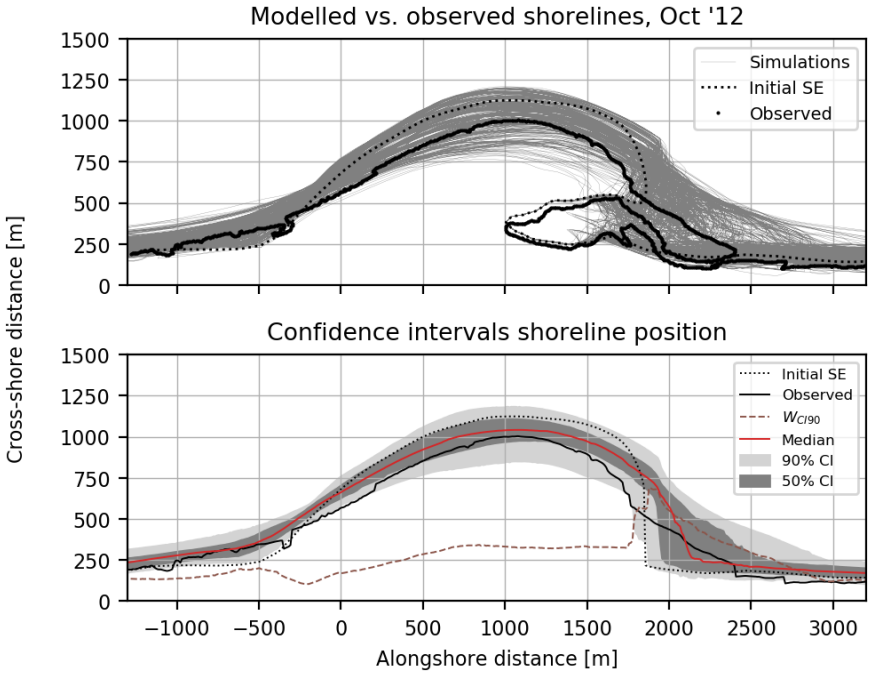

Figure 4.2: *Upper panel:* Computed $0m$-contours for each of the 340 simulations on 12 October 2012 (grey lines) and observed $0m$-contours derived from measurements shortly after construction (2 August 2011, black, dotted line) and on 10 October 2012 (black, dashed line). *Lower panel:*: Median (red line) and 90% / 50% CI's of the interpolated, predicted shorelines (grey shaded areas), observed shoreline (black line) and cross-shore width of the 90% CI (brown, dashed line). The area with the greatest variance is just north of the lagoon, where the spit forms.

4.2. Results EE method

This section describes the main step of the SA, the results of the EE method. The relative influence of the 16 parameters is analysed and transformed into a ranking. Potential interaction effects and non-linearity of parameters is also considered. This is done by plotting the the parameters in $\mu^* - \sigma$ space, based on the sensitivity measures derived by the EE method (see Section 3.3). The results are discussed for each output variable.

4.2.1. Volume changes

The EE analysis has been carried out for the cumulative volume changes in the four control areas (Figure 4.3). It stands out that f_{sus} is the most influential parameter, as it has the highest influence (μ^*) for all four areas. It is followed by the d_{50} and the Fac_{DSS}, which are close to each other throughout the four areas. It is not surprising that these three parameters are all influential as they are connected to each other. Especially the Fac_{DSS} is directly related with the d_{50} as the product of the two parameters represents the suspended sediment diameter (see AppendixA). Furthermore, the grain size is a direct input in the computation of the sediment transport (see Appendix A), which is multiplied with the scaling factor f_{sus}.

After this 'top three', a second group of parameters shows similar degrees of influence across the four areas (see red circles in Figure 4.3). This group consistently includes C, $f_{sus,w}$, α_{rol}, β_{rol}, γ, γ_{max}, and θ_{sd}. Their relative influence shows variations among the four areas, but they are generally close. α_{rol}, γ, and θ_{sd} generally lead this group. In the South section, d_{50} and Fac_{DSS} also belong in this second group. The remaining parameters are consistently at the low end of the μ^*-axis implying low influence on volume changes.

With the exception of f_{sus}, all parameters are located relatively close to the $\mu^* = \sigma$ line. Specifically, not a single other parameter is consistently located far off the line. This suggests that interaction effects are not extreme, yet exist for most parameters. Note that the relatively low value of of σ for f_{sus} is by no means an indication that it does not interact with other parameters or that the model response to this parameter is linear. It still has the largest absolute σ of all parameters. Rather, its position below the dashed line suggests that the interaction / non-linear effects are overshadowed by the overall influence of the parameter.

Originally, $f_{sus,w}$ and α_{rol} showed an uncharacteristically large σ-value for the entire SE area (see Figure D.1),

which pointed to a potential outlier among the EE's. Upon further analysis it was found that the simulations of one of the EE trajectories showed unrealistically strong erosion patterns and that this caused the high σ values. As this impedes the consistency of the EE results, this trajectory is ignored for the EE analysis (i.e. the results are based on 323 of the 340 simulations).

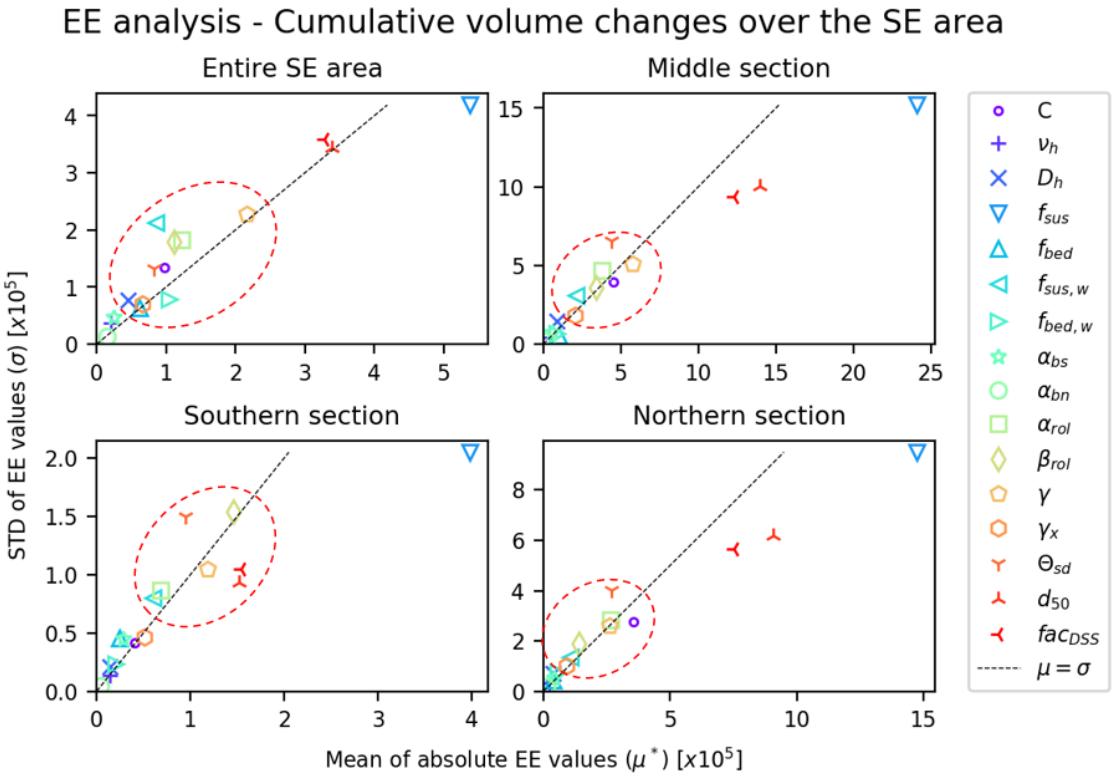

Figure 4.3: Results of the EE analysis for the cumulative volume changes in the four control areas. On the x-axis of each plot, the mean of the absolute EE's (μ^*) is plotted while the y-axis represents the standard deviation (σ) of the EE's. The $\mu^* = \sigma$ line is plotted as well (dashed). The red circles represent the 'second group' behind the leading parameters.

4.2.2. Shoreline position

Although f_{sus} is again the most influential parameter for the shoreline position (Figure 4.4, upper panels), the d_{50} and Fac_{DSS} are less influential than for the volume changes (relative to the other parameters). The 'second group' now clearly consists of 7 parameters (red dashed circles): d_{50}, Fac_{DSS}, $f_{sus,w}$, α_{rol}, β_{rol}, γ, θ_{sd}, as C falls out of this group.

The differences between the entire shoreline and the spit area are limited. $f_{sus,w}$ appears more influential for the entire shoreline, but the second group is clustered together relatively closely for both areas. There are, however, some significant differences in σ. For example, θ_{sd} has the highest σ while α_{rol} appears to interact less (or enforce a more linear model response). Almost all parameters are below, or around, the $\mu^* = \sigma$ line. This, however, is also influenced by the fact that the MAE was used to compute the EE's for the shoreline and bed levels (see Section 3.3). Hence all EE's are positive, inevitably leading to lower σ values than for the volume changes. Therefore, it might be more informative to consider the relative σ values between the parameters rather than the location relative to the $\mu^* = \sigma$ line.

4.2.3. Bed levels

The bed levels paint a very similar picture to the shoreline position. f_{sus} remains dominant over the other parameters, indicating that it is by far the most influential parameter for the applied model. Then follow the same seven parameters as for the shoreline. This group is a bit more spread out and, similar to the volume changes, led by d_{50} and Fac_{DSS}. In the next section, the discussed results will be transformed into a ranking of

the parameters.

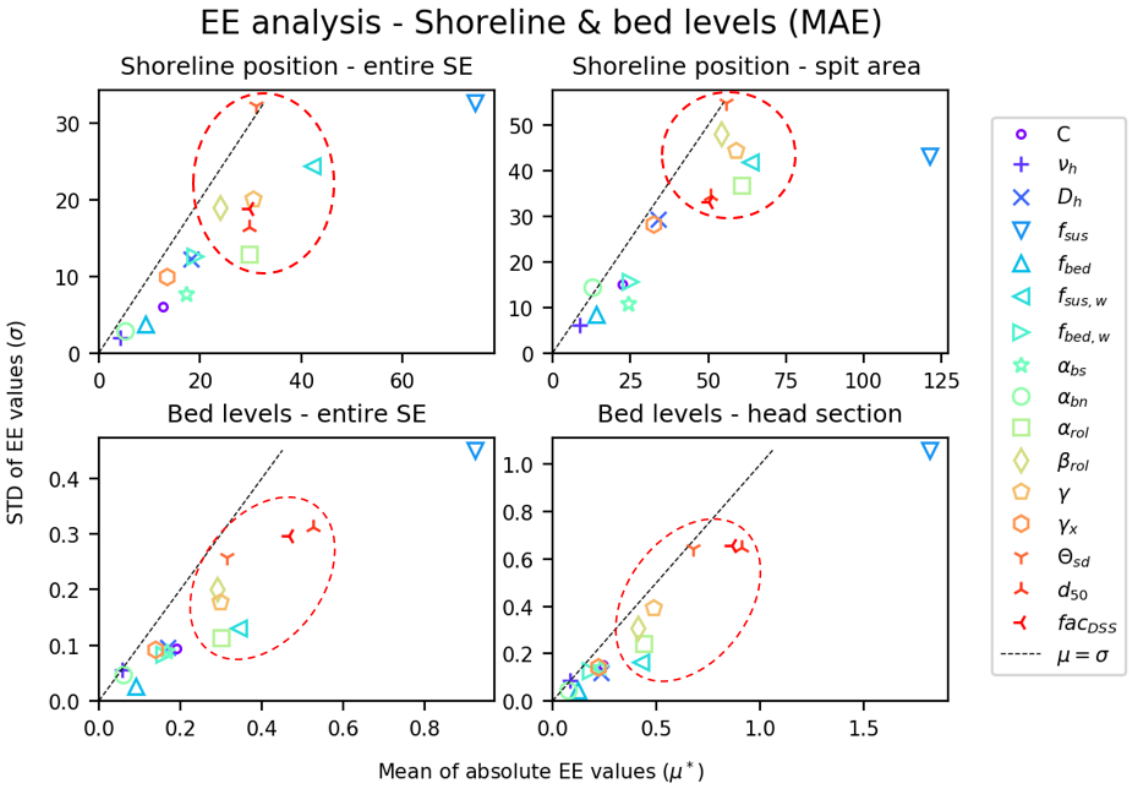

Figure 4.4: Results of the EE-analysis for the shoreline position of the entire SE (upper left) and the spit area (upper right) and the bed levels for the entire SE (lower left) and the middle section (lower right). The red, dashed circles represent the 'second group' behind the leading parameters.

4.2.4. Final ranking

Based on the results of the previous sections, a final ranking of the parameters is created. The parameters are ranked for each of the three output categories (Volume changes, shoreline position and bed levels), based on their corresponding μ^* value. For each output category, the components are given equal weight (e.g. for the volume changes, the parameters are ranked by their average rank for the four control areas). Note that this is a simple ranking, meaning that it does not represent the absolute difference in influence between parameters. Hence some parameters are tied for the same rank.

The ranking (Table 4.2) shows the dominance of f_{sus}, which ranks first for each category. d_{50} and Fac_{DSS} rank second and third, respectively, for the volume changes and bed levels, but only tied-sixth for the shoreline. Note, however, that for the shoreline, ranks two to seven have similar μ^* values and mainly differ in σ (Figure 4.4). γ, α_{rol}, β_{rol}, θ_{sd} and $f_{sus,w}$ all consistently feature in the top seven as well. In fact, the top eight of each category consists of the same eight parameters (except for C, which takes the place of $f_{sus,w}$ for the volume changes). These eight parameters form a more influential group, whereas the remaining eight parameters hardly feature in the top ten for any output category.

4.3. Parameter selection

This section describes the selection process for the parameters that will be examined during the GLUE analysis. The selection is based on three criteria:

1. **Absolute influence** (represented by μ^*)
   This is the main selection criterion and is based on the ranking shown in Table 4.2. Ultimately, the parameters with the most influence are expected to be the largest sources of parameter uncertainty.

Rank	Volume changes	Shoreline position	Bed levels
1	f_{sus}	f_{sus}	f_{sus}
2	d_{50}	$f_{sus,w}$	d_{50}
3	Fac_{DSS}	$\alpha_{rol}, \gamma, \theta_{sd}$	Fac_{DSS}
4	γ	-	θ_{sd}
5	α_{rol}	-	$f_{sus,w}$
6	β_{rol}	d_{50}, Fac_{DSS}	α_{rol}, γ
7	C, θ_{sd}	-	-
8	-	β_{rol}	β_{rol}
9	$f_{sus,w}$	D_h	C
10	γ_{max}	$f_{bed,w}$	D_h

Table 4.2: Top 10 influential parameters, ranked by the μ^* values obtained from the EE method. The ranking is given for each output category (volume changes, shoreline position and bed levels). The empty ranks (-) are a result of two or more parameters ranking equally above.

2. **Interaction effects**
   Interaction effects between parameters impede the the GLUE analysis as they cause equifinality (e.g. if two parameters are each others reciprocals the GLUE analysis will struggle to identify optimal values for them). It is difficult to find a concrete selection criterion for this, as σ does not say anything about which parameters interact with each other, or how it is influenced by non-linearity. However, σ may give a first indication, as does the conceptual implementation of the parameters in the model (see Appendix A). Finally, parameters from the same class (hydraulic, waves, sediment and morphology) are also expected to interact more with each other.

3. **Performance analysis**
   Analysing the model performance in relation to the parameter values can give additional insight into how the parameter range affects model performance. This is mainly used to analyse the applied ranges for the parameters, which is useful for the setup of the GLUE method.

4.3.1. Performance analysis

The performance analysis focuses on the eight most influential parameters from the EE analysis (f_{sus}, $f_{sus,w}$, α_{rol}, β_{rol}, γ, θ_{sd}, d_{50} and Fac_{DSS}, see Table 4.2). For each simulation, the BSS is computed for the different output variables (see Section 3.1). In this section, the focus will be on BSS_{tot} as indicator of overall model performance, which is plotted against each parameter range. But the individual outputs have also been considered and are presented Appendix D.3.

Looking at the 35 best simulations (given by a threshold of $BSS_{tot} > 0.6$) several trends can be observed (Figure 4.5). f_{sus} shows better results for the higher values (0.7 and 1) while for $f_{sus,w}$ the lowest value (0.1) gives the best model performance. α_{rol} shows a trend towards high values, whereas γ leans towards lower values. For the other parameters the BSS values are more evenly distributed over the range. We can also look at low BSS values, to see what parameter values might cause poor performance. Looking at all simulations resulting in a negative BSS_{tot} (Figure 4.6), similar trends can be observed as for the good simulations. For f_{sus}, α_{rol} and d_{50}, most simulations return a negative BSS_{tot} for the lowest parameter value, while this occurs for higher values of $f_{sus,w}$ and γ. Note that the amount of total simulations is not uniformly distributed over the parameter values (see also Figure D.2). Hence, the fraction of simulations resulting in a negative BSS is considered, not simply the amount of negative simulations.

It is important to note that the model is influenced (to some degree) by all parameters, thus absolute conclusions on the distribution of model performance over a single parameter range are impossible. That is also why, for some parameters, no conclusion can be drawn from the performance analysis (e.g. β_{rol}). Thus, this analysis should be seen as a first qualitative indication only.

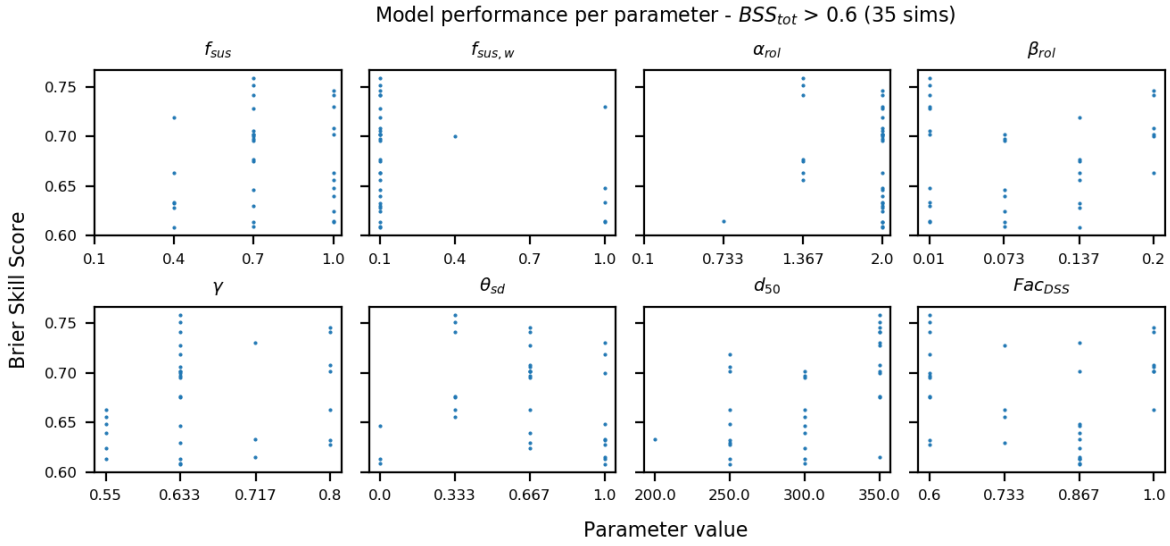

Figure 4.5: Scatter plots showing the BSS_{tot} against the parameter value for each of the eight considered parameters. Only simulations with a $BSS_{tot} > 0.6$ are plotted here (35 simulations in total).

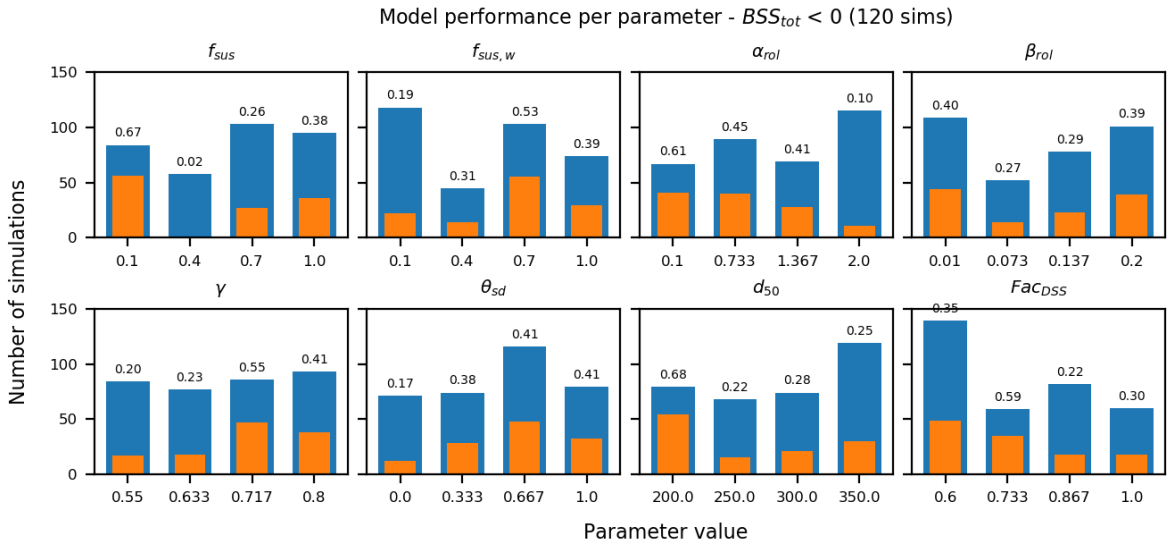

Figure 4.6: Bar plot showing the total amount of simulations per parameter value (blue), the amount of simulations with a negative BSS_{tot} (orange), and the ratio between the two given by the black number above each bar (i.e. the fraction of simulations that results in a negative BSS_{tot}). In total 120 of the 340 simulations resulted in a negative BSS_{tot}.

4.3.2. Final selection and GLUE setup

The final selection for the GLUE method consists of the following five parameters: f_{sus}, d_{50}, γ, α_{rol}, and θ_{sd} (Table 4.3). The selection of each parameter is briefly discussed below.

- *Suspended sediment transport scaling factor*, f_{sus}
  This parameter is the most obvious choice as it is by far the most influential parameter on every output variable, according to the EE results.

- *Median grain diameter*, d_{50}
  The d_{50} is generally the second-most influential parameter, albeit it is closer to the rest of the parameters than to f_{sus}. Moreover, grain size is possibly the most important physical quantity in sediment transport

dynamics. One could argue that there interaction effects with f_{sus} can be expected, as both parameters directly influence the suspended sediment transport. Based on their importance, however, not selecting either of the two would be hard to justify.

- *Wave breaker index*, γ
  According to the EE results, the breaker index γ is the most influential wave parameter (slightly ahead of α_{rol}, based on μ^*). It effectively determines at which water depth waves start to break, hence directly influences the amount of energy transferred by breaking waves, which was also found to be an important aspect of the SE model by Luijendijk et al. (2017).

- *Wave energy dissipation coefficient*, α_{rol}
  α_{rol} is another influential wave parameter. It also shows interesting results in the performance analysis, as the best value appears to be its upper limit (2). This could be an indication that the full range of suitable values has not yet been applied for this parameter, and that it might be even more influential. Finally, it has relatively low σ values compared to other parameters, making it less likely to interact with other selected parameters.

- *Dry cell erosion factor*, θ_{sd}
  θ_{sd} is not necessarily an obvious choice among the remaining parameters based on absolute influence (that would be Fac_{DSS}). However, the above four parameters already include two sediment parameters and two wave parameters. θ_{sd} adds a morphology parameter to this group (while the remaining three options would add another wave or sediment parameter). Although it does show high σ values, its simple implementation in the model appears less likely to interact strongly with other parameters (see Appendix A). Finally, Luijendijk et al. (2017) identified dry cell erosion as an important feature in the SE model.

Next follows a short explanation why the remaining three parameters of the influential group have not been selected.

- *Wave-affected sediment transport scaling*, $f_{sus,w}$
  This parameter is highly influential for the shoreline position (ranks second), but less so for the other outputs (e.g. ninth for the volume changes). Additionally, the performance analysis suggests that the lowest value is by far the best (30 of the 35 best scoring simulations), making it easier to find a fixed value for the parameter. One might argue that this parameter should be included as the waves form the dominant forcing mechanism. However, as explained in Appendix A, the main wave affected sediment transport (caused by the wave-induced alongshore current) falls under the suspended transport, which is scaled by f_{sus}, not $f_{sus,w}$.

- *Mean roller slope*, β_{rol}
  Of the three considered wave parameters β_{rol} ranks lowest in absolute influence. Only for the volume changes in the South section it is more influential than γ and α_{rol}, but the South section is the least uncertain (hence, less interesting for the GLUE analysis). Moreover, the implementation of α_{rol} and β_{rol} as scaling coefficient of two dissipation components in the roller energy balance (see Appendix A) suggests they may function as each other's reciprocals. Hence, adding β_{rol} instead of θ_{sd} would significantly increase the chance of strong interaction effects.

- *Factor for suspended sediment size*, Fac_{DSS}
  Overall, Fac_{DSS} is the third or fourth most influential parameter. However, its direct connection to d_{50} (Appendix A) suggest interaction effects between the two are high (similar to α_{rol} and β_{rol}). Based on these observations and expert consultation, the assumption is made that part of the uncertainty in Fac_{DSS} can be captured by varying d_{50}.

For some of the selected parameters, the range has been adjusted from their SA values, based on the results from the performance analysis in the previous section (Table 4.3). The reason for this is that the GLUE analysis is limited to an average resolution of 1/ 4 per parameter range. If it appears likely that a certain part of a parameter range leads to poor model performance, that range can be narrowed to indirectly improve the sampling resolution for that parameter. This is not possible for all parameters, however. In fact, the range for α_{rol} was shifted and even increased slightly.

Cat.	Parameter	Symbol	Range		Comment / Change
W	Breaker parameter	γ	0.55	0.8	same range as for SA
	Roller dissipation coefficient	α_{rol}	1	3	range shifted upwards from [0.1, 2]
S	Suspended transport scaling	f_{sus}	0.2	1	lower limit raised from 0.1
	Median grain diameter [µm]	d_{50}	200	350	same range as for SA
M	Dry cell erosion factor	θ_{sd}	0.1	1	lower limit raised from 0

Table 4.3: Overview of the five selected parameters for the GLUE analysis and their respective sampling ranges.

5

Results – Uncertainty Analysis

Chapter outline
Chapter 5 describes the results of the uncertainty analysis using the GLUE method, which comprises the main step of this thesis. It is divided in five main sections. Section 5.1 discusses the derived uncertainty bounds for the volume changes and the shoreline position. In Section 5.2 these bounds are extended to the bed levels in 2D, to give a spatial uncertainty map. This is followed by a temporal analysis in Section 5.3, which examines the development of the uncertainty bounds over time. Next, the parameter optimisation is discussed in Section 5.4. Finally, in Section 5.5 the results of the uncertainty comparison between parameter uncertainty and wave climate variability are presented.

5.1. Uncertainty bounds

This section discusses the first category of GLUE results, the derived uncertainty bounds. These will be analysed for the volume changes and the shoreline position, similar to Section 4.1. Here, the CIs represent the induced parameter uncertainty from the five selected parameters which is constrained by the observations (the empirical CDFs of the outputs are weighted by the likelihood values).

5.1.1. Volume changes

For the volume changes, a threshold of $BSS_{\Delta V} > 0.75$ is applied, leading to 586 behavioural simulations. The resulting CIs are significantly narrower than for the SA (Figure 5.1). This is expected for two main reasons: first, the BSS threshold and likelihoods constrain the uncertainty bounds, and secondly 11 of the initial 16 parameters were fixed. The black dashed lines show the 90% CI for all 1024 simulations, weighted equally (i.e. the prior, unconstrained distribution). As can be seen, the CIs obtained from the behavioural runs (weighted by their likelihoods) are narrower and shifted more towards the observations. Note that here, the figures for the different areas are not plotted on equal scale, to facilitate an easier visual analysis of the individual areas. However, the observed and predicted changes in the middle and North sections are up to five times larger than for the South section and the entire SE area. For reference, the same figure is plotted on equal scale in Appendix E, Figure E.1), to give an overview of the relative width of the uncertainty bounds.

The model struggles to correctly predict the volume changes over the entire SE, as practically all observations are outside the 90% CI. This was already visible in the SA results and is in line with Luijendijk et al. (2017). The middle and North sections are modelled better—nearly all observations fall within the 90% CI. Considerably less fall in the 50% CI, however, and the model still tends to overestimate the accretion in the North section and underestimate it in the South section (several observations lie outside both CIs). The fact that still struggles to correctly predict the distribution of sediment over the North and South sections, suggests this not caused by the considered parameters. This makes it difficult to find optimal simulations: more sediment transport leads to closer approximation of the observations in the middle and South sections, but move further away in the North section. This is also visible from the behavioural CI, which is shifted towards higher volume changes compared to the prior CI, which improves the predictions in all areas except the North section.

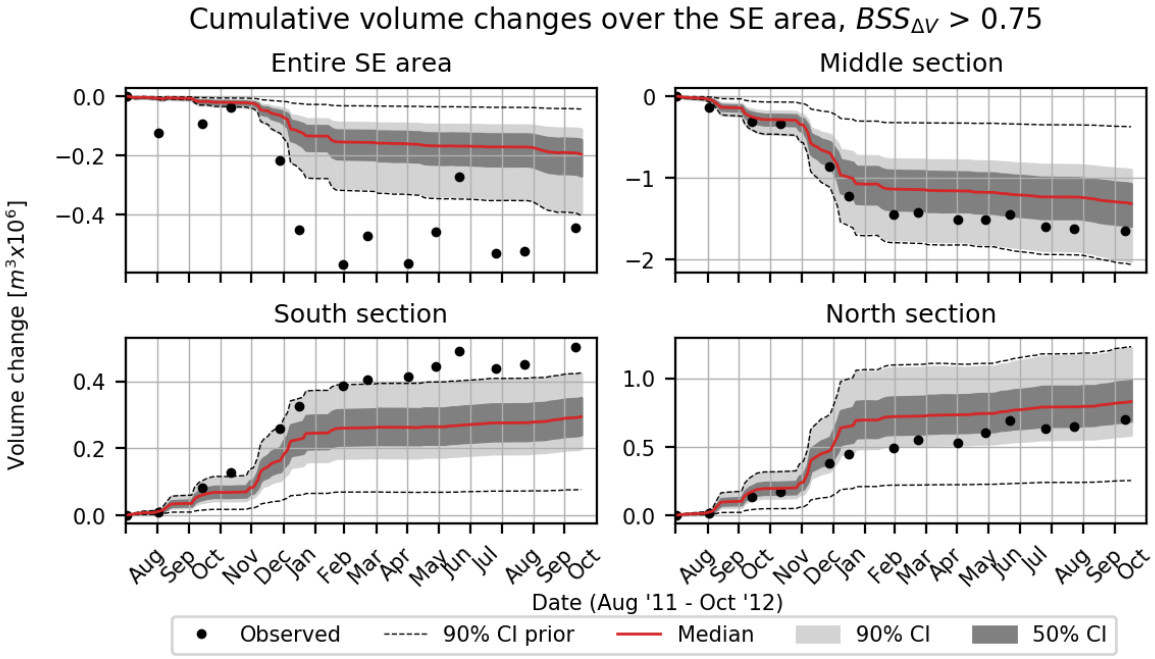

Figure 5.1: Time series of cumulative volume changes between August 2011 and October 2012 for the four control areas. The plots show the median (red) and CIs (grey areas) of the behavioural GLUE simulations (weighted by the likelihood scores), the 90% CI of the prior distribution (black, dashed), and the observed volume changes (black dots) (de Schipper et al., 2016). Note that the plots are not on equal scale. The observed and predicted changes in the middle and North sections are higher than for the South section and entire SE area. The same figure is plotted on equal scale in Appendix E, Figure E.1).

Section	$Med_{\Delta V}$	$CI_{90,lower}$	$CI_{90,upper}$	W_{CI90}	$\frac{W_{CI90}}{Med_{\Delta V}}$
Entire SE	0.2	0.11	0.4	0.29	1.45
Middle	1.32	0.92	2.04	1.12	0.85
South	0.29	0.2	0.42	0.22	0.76
North	0.83	0.6	1.23	0.63	0.76

Table 5.1: Overview of the uncertainty bounds for the cumulative volume changes in October 2012. Given is the median predicted volume change ($Med_{\Delta V}$), the upper and lower bounds of the 90% CI, the width of the 90% CI (W_{CI90}) and the ratio between W_{CI90} and $Med_{\Delta V}$ (all in million m^3). The higher the ratio, the larger the degree of uncertainty.

Erosion in the middle section shows a spread of 1.3 million m^3 $-400.000/+700.000$ m^3 after the 14 month study period. Corresponding accretion volumes in the adjacent South and North sections show spreads of $290.000\,m^3$ $-100.000/+120.000\,m^3$ and $830.000\,m^3$ $-230.000/+400.000\,m^3$, respectively. Hence, uncertainty is largest for the middle section ($W_{CI90} = 1.12$ million m^3), followed by the North section ($W_{CI90} = 0.63$ million m^3) (Table 5.1). An indication of the uncertainty relative to the predicted change is given by the ratio between W_{CI90} and the median of the behavioural runs ($Med_{\Delta V}$). From this ratio, the entire SE area appears to be most uncertain (W_{CI90} is larger than the median). For all three individual sections, the width of the uncertainty bounds is at least 75% of the median predicted changes.

What can be concluded from these results, is that parameter uncertainty translates to significant uncertainty in the predicted volume changes. This is in spite of constraining the uncertainty bounds by the observations through the likelihood scores. Moreover, we can see that most of the uncertainty appears to develop during the storm season of December and January (see also Section 5.3). Finally, we can see that the parameter uncertainty is not normally distributed. Rather, it is asymmetric towards higher magnitudes of change (i.e. more erosion/sedimentation).

5.1.2. Shoreline position

This section analyses the uncertainty in the predicted shoreline position in October 2012, while the entire time series is considered in the temporal analysis in Section 5.3. Afgain, the CIs are significantly narrower than for the SA, as expected. The applied threshold is $BSS_{sl} > 0.55$, resulting in 587 behavioural runs. The model predicts the evolution of a spit for virtually all behavioural simulations, but struggles to predict the correct spit location (Figure 5.2, upper panel). Similar conclusions can be drawn from the CIs (lower panel): the observed spit starts further landward (outside the 90% CI), is less curved, and crosses the CI at around $x = 2100\ m$. The median of the simulations also underestimates the length of the spit by around 250 m.

The spit area contains by far the most uncertainty, as W_{CI90} (brown, dashed line) reaches up to 350 m in cross-shore position (at $x \approx 2100\ m$). Uncertainty in the middle section has decreased significantly from the SA results, as the 90% CI now only covers $\approx 100\ m$ at $x \approx 800\ m$. The observed shoreline is still more retreated than the median of the predictions, and falls outside the 90% CI at several locations—most notably in the spit area between 1700 m and 2100 m alongshore. Apart from the spit area, however, the deviations are generally small ($O(10m)$). Finally, in the South section, there is very little uncertainty. The 90% CI only covers around 75 m at its widest point ($x \approx -600$). The median follows the observation well in this section. In short, parameter uncertainty in the shoreline predictions is mostly restricted to the spit area, and even there it is concentrated around the tip of the spit ($x \approx [2100, 2500]$).

In the spit area, the asymmetric distribution of the uncertainty stands out and is clearly non-normal, while the CIs are more or less symmetric for the remaining part of the shoreline.

5.2. Spatial uncertainty map

This section discusses the spatial uncertainty map, which is a geographical visualisation of uncertainty over the SE area and basically an extension of the uncertainty bounds in 2D. Hence, a spatially distributed output variable is required to measure the uncertainty: the bed level changes. The idea is that this can be used to communicate uncertainty to stakeholders in a clear and efficient way. The width of the 90% CI (W_{CI90}) will be used as uncertainty indicator (see Section 3.4).

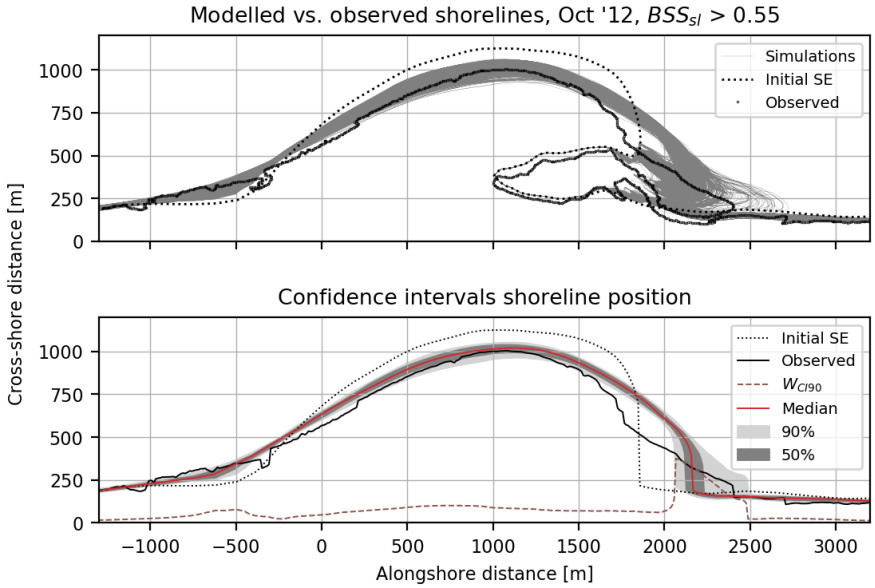

Figure 5.2: *Upper panel:* predicted shoreline in October 2012 for all behavioural runs (grey lines) compared to the observed (black, solid) and initial (black, dotted) shorelines. *Lower panel:* median (red line) and CIs (grey shaded areas) of the interpolated, behavioural shorelines, and width of the 90% CI (W_{CI90}, brown, dashed line).

The median and 90% CI are computed for the cumulative bed level changes over the 14-month study period, from August 2011 to October 2012. The spatial uncertainty map is a combination of five maps (Figure 5.3). The left column (three panels) represents the model prediction which received the highest likelihood for bed levels (i.e. the highest value for BSS_z). Hereafter, this simulation will be referred to as the highest likelihood simulation (HLS). The upper panel shows the initial bathymetry in August 2011, which is equal for all model runs. The two lower panels show the outcome of the HLS: the predicted cumulative bed level changes (middle panel; blue corresponds to erosion, red to accretion) and the final bathymetry in October 2012 (lower panel). Hence, these plots present the most likely outcome that a stakeholder can expect. The right column (two panels) represents the accompanying uncertainty: the median of the predicted bed level changes (middle row) and the width of the 90% CI (bottom row), both derived from the weighted behavioural simulations. For reference, the contour lines in all plots (except the initial bathymetry) correspond to the final bathymetry of the HLS, so the user knows see where to expect uncertainty.

The maps show an expected pattern: both the HLS, as well as the median of the behavioural runs predict erosion (blue) up to $4-5\ m$ in the middle section and accretion (red) up to $5\ m$ and $2.5\ m$ in the North and South sections, respectively. The predicted bed level changes by the HLS are similar to the median predicted changes. Again, uncertainty is highest in the spit area ($W_{CI90,max} \approx 5-6m$), followed by the head of the peninsula ($W_{CI90,max} \approx 4-5m$). High levels of uncertainty generally correlate with large expected bed level changes ($r = 0.65$ between the magnitude of the median bed level changes and W_{CI90}). There are also areas with high uncertainty that do not correspond to large expected bed level changes, however, especially in the spit area. The 'corridor' of low uncertainty along the spit edge indicates that most simulations predict the formation of a sand spit, while the high uncertainty around the tip of the spit suggests that the length and exact position of the spit differ among the predictions. Also, the uncertain area next to tip of the spit ($x \approx [2400, 2700]$), extends around 200–300 m further North (to the right alongshore) than the zone where high accretion is expected (red area around the spit in the median plot), showing that uncertainty can also be high in places of low expected changes.

The high uncertainty around the head section of the peninsula indicates that the predicted erosion varies significantly. Another interesting observation is the accumulation of uncertainty ($W_{CI90} \approx 1-2m$) along the $2m$-depth contour, extending from the head section into the South section. This might be an indication of uncertainty around the near-shore sandbar. The remainder of the South sections shows low uncertainty ($W_{CI90} < 1m$). Finally, beyond the $6m$-depth contour, uncertainty quickly decreases to negligible levels (as do the predicted changes).

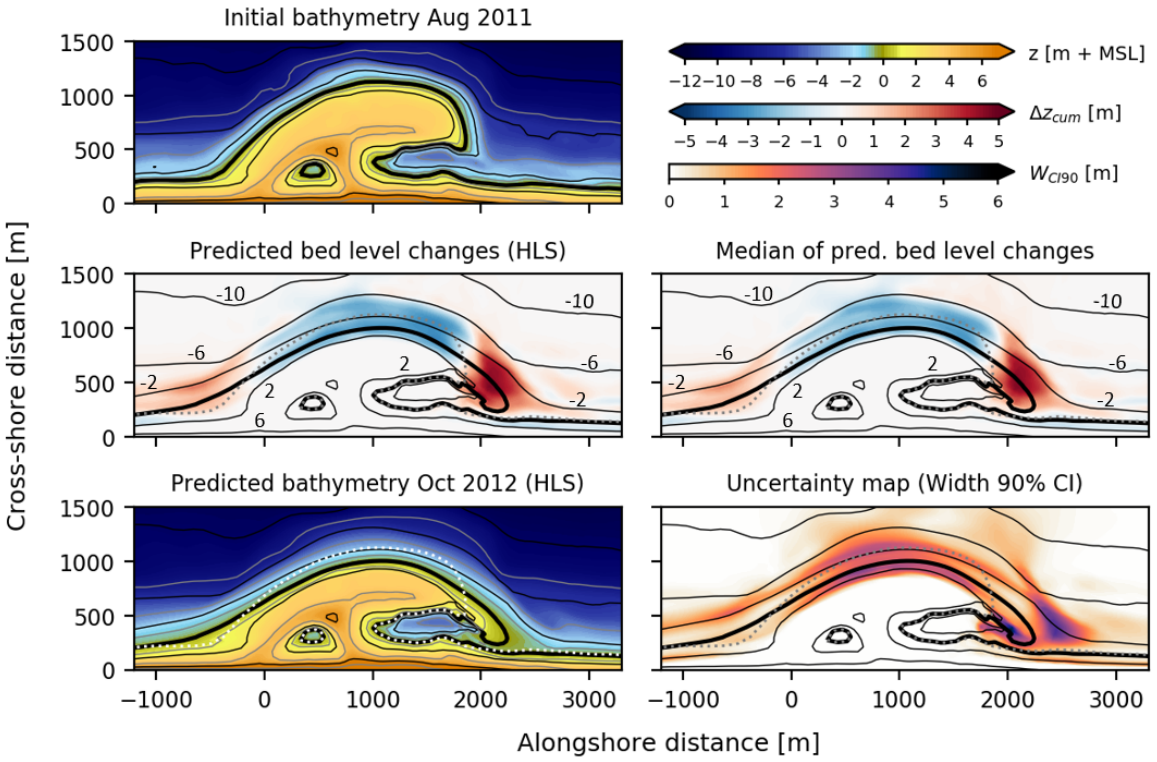

Figure 5.3: *Left column:* Initial bathymetry in August 2011 (top row), cumulative bed level changes predicted by the HLS over the 14 month study period (middle row), and final bathymetry in October 2012, predicted by the HLS (bottom row). *Right column:* Median of predicted, behavioural bed level changes (middle row, blue and red colours indicate erosion and accretion respectively) and width of the 90% CI (W_{CI90}) for cumulative bed level changes (bottom row, darker colours indicate high uncertainty). The black contour lines represent the bathymetry predicted by the HLS in all plots except the initial bathymetry. The bold contour represents the shoreline (0 m + MSL). In the middle row, the non-zero depth contours are labelled by their respective depth / elevation. Finally, the dotted lines represent the shoreline of the initial bathymetry.

5.3. Temporal analysis

Except for the cumulative volume changes, the results analysed in the previous sections only consider the parameter uncertainty at the end of the simulation period (October 2012). To find out how parameter uncertainty affects model predictions, however, it is of interest to examine its development over the considered study period. This may also reveal relations between the wave climate and the development of uncertainty.

The model returns outputs over 64 time intervals, equally spaced over the filtered and compressed wave series. These output intervals make up the temporal resolution for this analysis. Note that, because of the BFFC approach (Section 2.5.3), the output intervals are equally spaced over the BFFC time series, but not over the original time series, where they span varying amounts of time (i.e. if an equidistant time axis is used the output intervals are not equally spaced). To enable a direct comparison between wave climate and model output, the wave time series is converted to the same temporal resolution as the model outputs (i.e. 64 equally spaced intervals over the filtered time series). The wave climate is represented by the cumulative mean wave energy density (E) over each output interval (see Appendix B for the computation of E).

First, the development of the wave climate in time, is compared to the predicted volume changes and the uncertainty bounds (represented by W_{CI90})(Figure 5.4). Hereby, the focus is on the middle section, as this is the most representative section—the erosion starts in the middle section and is then distributed over the adjacent areas. Note that the data is plotted over the filtered time series, hence the time axis is distorted. For instance, in March 2012 there were very few waves above 1 m, so the majority of March has been filtered from the time series (see also section 2.5.3). Consequently the entire month of March is represented by two output intervals only. On the other hand, December is characterised by high wave heights and takes up nine output intervals, as hardly any wave conditions were filtered.

The magnitude of the predicted erosion volumes is strongly correlated to the wave energy ($r_E = 0.92$) over the 14-month study period. The correlation between the magnitude of the erosion and the change in W_{CI90} is

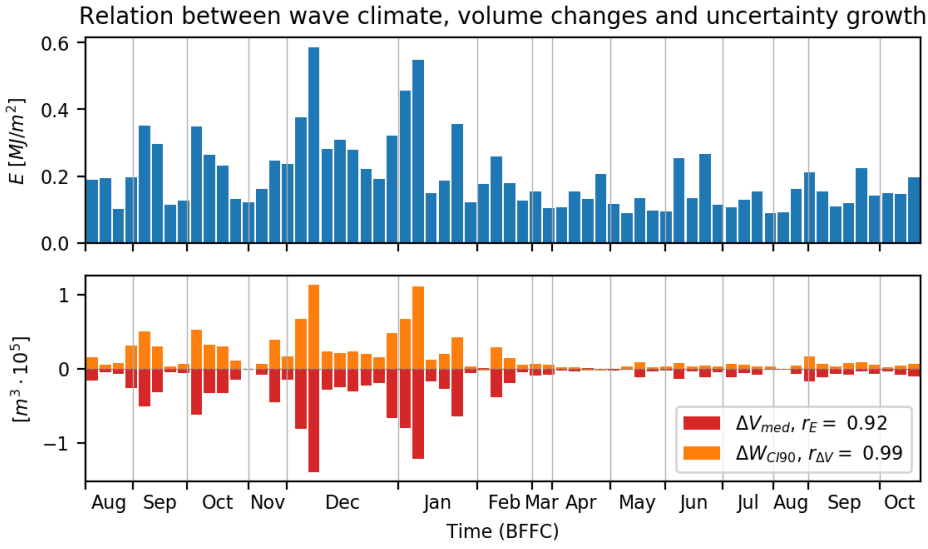

Figure 5.4: Bar plots showing the cumulative mean wave energy density (E, blue bars), the median predicted erosion volume in the middle section (ΔV_{med}, red bars) and the change in width of the corresponding 90% CI (ΔW_{CI90}, orange bars). All three quantities are given per model output interval (i.e. every bar represents one output interval). Due to the BFFC approach the output intervals span varying amounts of 'real' time, resulting in the distorted time axis.

Section		Aug - Oct (14 months)	Aug - Feb	Mar - Oct
Mid	r_E	0.92	0.94	0.51
	$r_{\Delta V}$	0.99	0.99	0.79
North	r_E	0.91	0.92	0.53
	$r_{\Delta V}$	0.97	0.97	0.56
South	r_E	0.92	0.94	0.55
	$r_{\Delta V}$	0.96	0.95	0.27
Combined	r_E	0.70	0.71	0.27
	$r_{\Delta V}$	0.99	0.99	0.98

Table 5.2: Overview of the correlation between the wave energy and the median predicted volume change (r_E) and between the median predicted volume change and the change in W_{CI90} ($r_{\Delta V}$). The correlation is given for each section, and over three time periods: the entire 14-month study period (Aug - Oct), the initial response phase (first seven months, Aug - Feb), and the final seven months (Mar - Oct).

almost perfect ($r_{\Delta V} = 0.99$). Hence, periods of high wave energy cause the most morphological activity, and with it the most uncertainty in the predicted erosion volumes. Figure 5.4 only shows the middle section, but the correlations for all areas are presented in Table 5.2. Another fact that stands out, is that both V_{med} and W_{CI90} change very little from March onwards, compared to the first seven months (see the height of the red and orange bars in Figure 5.4). In fact, at the end of February, W_{CI90} is already at $\approx 90\%$ of its final value in October 2012. Apparently, there is an initial response period from August until February, during which most of the morphological changes are predicted. This is similar to the observations of de Schipper et al. (2016). After the first seven months, the SE appears to reach a state in which morphological activity becomes slower, more nuanced, and less dependent on the wave climate. This is confirmed by the correlations over the two separate seven-month periods (Table 5.2). Especially r_E is significantly smaller over the final seven months, compared to the first seven months, for all sections. This implies that the correlation between wave climate and uncertainty growth only holds for the initial response phase. Consider, for example, the predicted erosion volumes corresponding to the two highest energy intervals in June and the highest in February (the blue bars in Figure 5.4, all three approximately equal). The expected volume change after the February event is about three times larger than after the two periods in June. In June, the SE has apparently reached a state in which the same amount of wave energy leads to significantly less morphological changes.

Similar observations can be derived from the shoreline changes. The Hovmöller diagram in Figure 5.5 shows

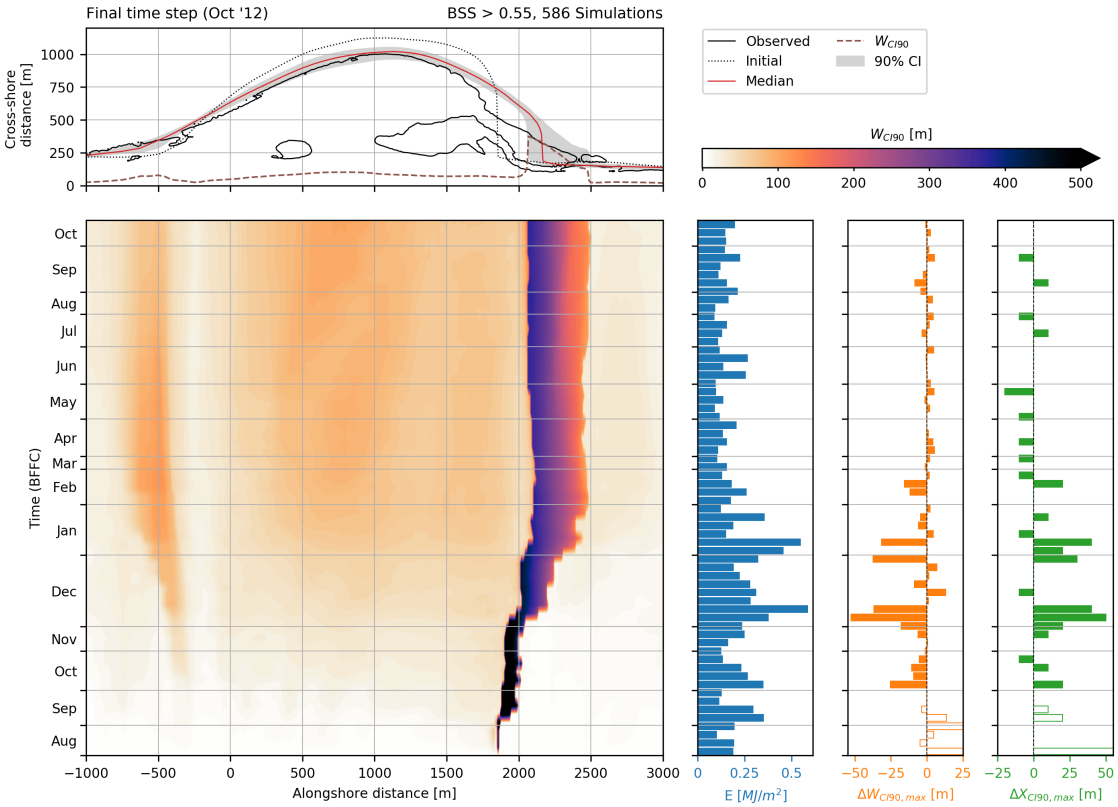

Figure 5.5: Hovmöller diagram showing the evolution of W_{CI90} in time, for the shoreline position. W_{CI90} is represented by the colour scale (darker colours indicate a wider 90% CI). The y-axis of the Hovmöller diagram represents time (again, distorted due to the BFFC series). The top panel shows the final situation in October 2012 (as in Figure 5.2). The three panels on the right show the wave energy (blue), change in the maximum width of the 90% CI (orange) and change in the alongshore position of the maximum width (green).

how the width of the 90% CI of the cross-shore shoreline position develops over time. For the volume changes, uncertainty continuously grows along with the actual volume changes, specifically in the first seven months. For the shoreline position this is a bit different. The cross-shore position of the shoreline is defined over equidistant alongshore points. Hence, the width of the CI is defined in cross-shore direction as well. As only the most seaward point counts as shoreline, at some point it jumps from the northern end of the peninsula (the spit) to the beach (see also Section 3.1). This jump is in cross-shore direction as well. Once the initial bathymetry starts to change during the simulations, the alongshore location of this jump will change for each simulation. The consequence is an extremely large but localised value of W_{CI90} in the early stages of the simulations (up to $W_{CI90} \approx 600m$ at $x \approx 1850m$). As the spit becomes longer and moves closer to the beach over time, the width of the CI decreases, along with the cross-shore distance between the spit and the beach. Hence, during the first time steps, the width of the CI is not directly governed by wave forcing, but rather by the initial spit formation. Therefore, the outputs over the first two months (empty orange and green bars in Figure 5.5) are not taken into account when looking at the correlation between the wave energy and the change in W_{CI90}. For the remaining period, the wave energy correlates with *absolute* changes in the maximum CI width $W_{CI90,max}$ ($r = 0.67$, correlation between the magnitude of the blue and orange bars in Figure 5.5). Similarly, changes in the alongshore position of the maximum CI width ($X_{CI90,max}$, green line) are also correlated to the wave climate ($r = 0.67$). Yet, these correlations are lower than for the volume changes. From the colour plot, it is clearly visible that, as the spit develops in time, it moves the area of highest uncertainty northward along the beach. This also shows the diffusion process of the initial coastline perturbation, which is slowly 'smeared out' over the alongshore distance. Similar to the volume changes, there is much less activity from March onward, when changes in $W_{CI90,max}$ and $X_{CI90,max}$ become significantly smaller and less correlated to the wave climate compared to the first seven months.

Finally, the temporal development of uncertainty in bed level changes over the SE can be analysed. As the spatial uncertainty map already consists of three dimensions without time, however, the development in time can not

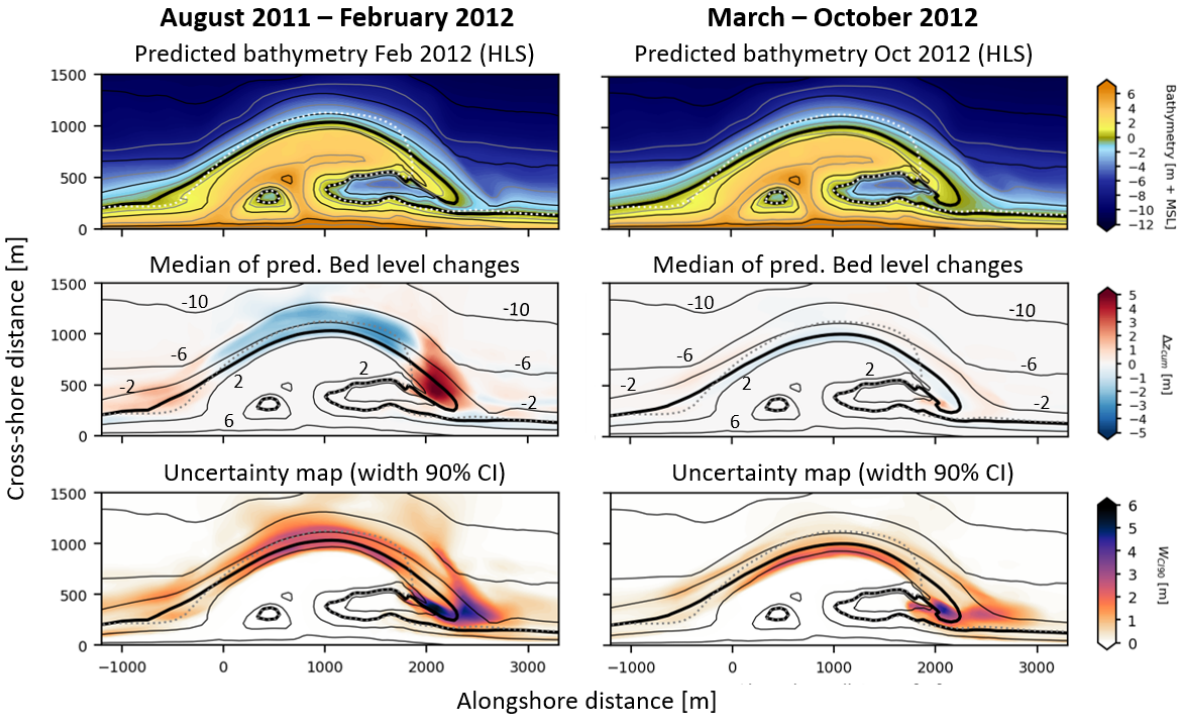

Figure 5.6: Spatial uncertainty maps for the periods of August - February (left) and March - October (right). Shown is the bathymetry at the end of the respective period, predicted by the HLS (top row), the median predicted bed level change (middle row) and the width of the 90% CI of bed level changes (bottom row).

be visualised as easily. Therefore, the study period is split in the two seven month periods for which the spatial uncertainty map is analysed—August until February and March until October (Figure 5.6). The maps confirm the decrease in morphological activity after the initial response phase (see the median of the predicted bed level changes). It is also confirmed that the majority of the uncertainty forms in the first seven months. However, the second period (Mar - Oct) also shows some areas of high uncertainty (e.g. $W_{CI90} \approx 4m$ at the tip of the spit), even though the median predicted bed level changes are very small ($O(1m)$). Hence, it seems that the here the impact of the initial response phase on uncertainty development is less pronounced.

To conclude, most of the morphological activity occurs in the initial response phase, from August until February. After this period, the predicted morphological activity decreases significantly, which agrees with the observations of de Schipper et al., 2016. This translates to the growth of the output uncertainty, which, over the first seven months, is strongly correlated to the morphological activity. In the period after February, the correlation decreases and less uncertainty is formed. Although this trend applies to all three output variables, it is strongest for the volume changes, for which only $\approx 10\%$ of the uncertainty forms after February. The bed level changes, on the other hand, still show relevant uncertainty development in the final seven months, albeit mostly limited to the area around the tip of the spit.

5.4. Parameter optimisation

In this section, the likelihood values corresponding to each parameter set are used to derive posterior likelihood distributions for the parameters. These distributions represent the likelihood of each parameter value to be optimal, based on the GLUE analysis. After deriving the distributions, they can be used to estimate an optimal parameter set. This can be done for every output variable separately, but here the focus will lie on the distributions derived from the combined likelihood (CL), which only includes parameter sets that are behavioural for all considered outputs. These can be used to find a reliable parameter set, suitable for all output variables.

5.4.1. Posterior distributions

The individual output variables each contain between 570 and 590 behavioural predictions (see Table 3.6). Based on CL, 507 parameter sets are behavioural, implying that most of the behavioural sets for a certain output are

also behavioural for the other outputs. In other words, there is general agreement among the different output variables. A likely reason for this is that the considered output variables are governed by similar dominant processes. This is unsurprising, as ultimately all three output categories are derived from the output bathymetry. It is therefore appropriate to seek common optimal values for the input parameters.

The posterior distributions are computed as a weighted histogram from the parameter sets, using the CL values as weights (Figure 5.7). They are then compared to the prior distributions and default / reference values provided with the model. The prior distribution is simply the distribution of all sampled parameter sets, which is near-uniform over the respective range (due to the applied Sobol' sampling). The more the posterior distribution of a parameter differs from the prior distribution, the more it affects the model skill (represented by CL), hence, the more uncertainty it introduces.

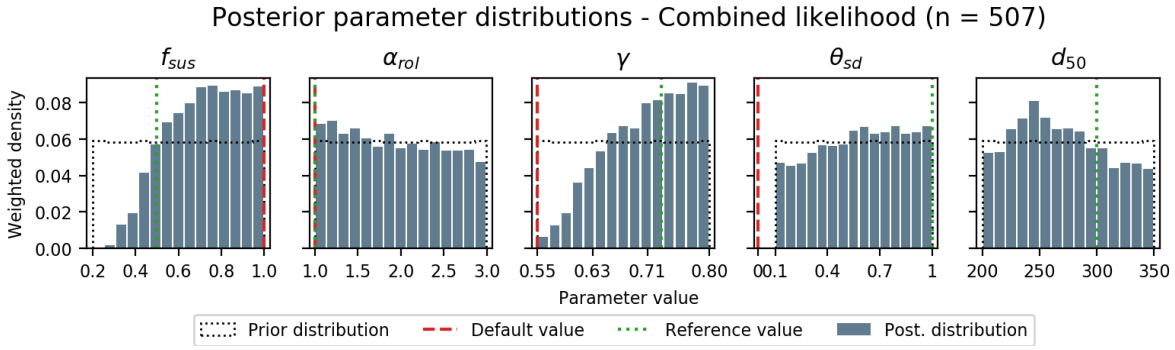

Figure 5.7: Prior (black, dotted line) and posterior distributions (histogram bars) for the five input parameter examined in the GLUE analysis. The Delft3D default values (red, dashed line) and reference values used by Luijendijk et al. (2019) (green, dotted line) are shown as well. The posterior distributions are weighted according to the combined likelihood, CL.

For f_{sus}, γ, and d_{50}, the posterior distribution shows a clear trend or peak. Optimal values for f_{sus} and γ lie in the higher range, indicating that the model performs better when sediment transport scaling is high and wave breaking is induced later (i.e. the waves can reach lower water depths before breaking). The distribution of d_{50} shows a peak at 240–250 μm, but is the most similar to the prior distribution of these three parameters. γ appears to differ most from the prior distribution, implying it might have the strongest effect on the model skill (although f_{sus} shows a similar deviation). This is interesting, as the SA deemed f_{sus} by far the most influential parameter. For α_{rol} and θ_{sd} the trends are less pronounced and the posterior distributions are similar to the prior ones. This implies that, based on the GLUE data, the model is more or less insensitive to these parameters. This is surprising, as both parameters appeared similarly influential to γ in the SA. Additionally, both the roller model as well as dry cell erosion were found to significantly improve Delft3D predictions of the SE (Luijendijk et al., 2017). Another surprising fact is that certain trends have reversed from the SA results. For instance, the posterior distribution of γ clearly favours the higher values, while the the performance analysis of the SA simulations (Section 4.3) implies a trend towards lower values. Similarly, for α_{rol} the SA simulations indicated a trend towards higher values, such that the range was shifted upwards for the GLUE analysis. The posterior distribution, however, is almost uniform, even showing a weak trend towards lower values.

While surprising, these observations do not make the obtained results less meaningful. Ultimately, the GLUE analysis provides a much more complete picture than the SA, which was carried out for a 16-dimensional parameter space, with three times less simulations. Hence, it appears that the SA simulations provided an incomplete picture (Chapter 6 offers a more detailed discussion on these differences). Still, the posterior distributions show clear differences between optimal values and the default Delft3D and reference values (Table 5.3). This shows the potential of the parameter optimisation to provide better parameter estimations.

Parameter set	f_{sus}	α_{rol}	γ	θ_{sd}	d_{50}
Default	1	1	0.55	0	-
Reference	0.5	1	0.73	1	300
Combined likelihood	0.75	1.2	0.78	0.8	245

Table 5.3: Default, reference and optimal parameter values (based on combined likelihood).

Sim	f_{sus}	α_{rol}	γ	θ_{sd}	d_{50}	BSS_{tot}	$BSS_{\Delta V}$	BSS_{sl}	$BSS_{sl,spit}$	BSS_z	$BSS_{z,head}$
OPS_0	0.75	1.2	0.78	0.8	245	0.71	0.82	0.67	0.57	0.58	0.81
OPS_1	0.75	1.3	0.76	0.73	330	0.79	0.93	0.72	0.68	0.63	0.85
OPS_2	0.78	1.5	0.78	0.73	325	0.80	0.94	0.74	0.70	0.64	0.86
OPS_3	0.72	2.5	0.76	0.85	245	0.78	0.94	0.74	0.67	0.59	0.82
OPS_4	0.85	1.5	0.78	0.9	295	0.79	0.92	0.74	0.68	0.63	0.85
Ref	0.5	1	0.73	1	300	0.72	0.94	0.58	0.52	0.53	0.80

Table 5.4: Overview of the five OPS estimated from the results of the parameter optimisation, as well as the reference simulation by Luijendijk et al. (2019). The table shows the parameter values for the GLUE parameters and the achieved model skill for each simulation.

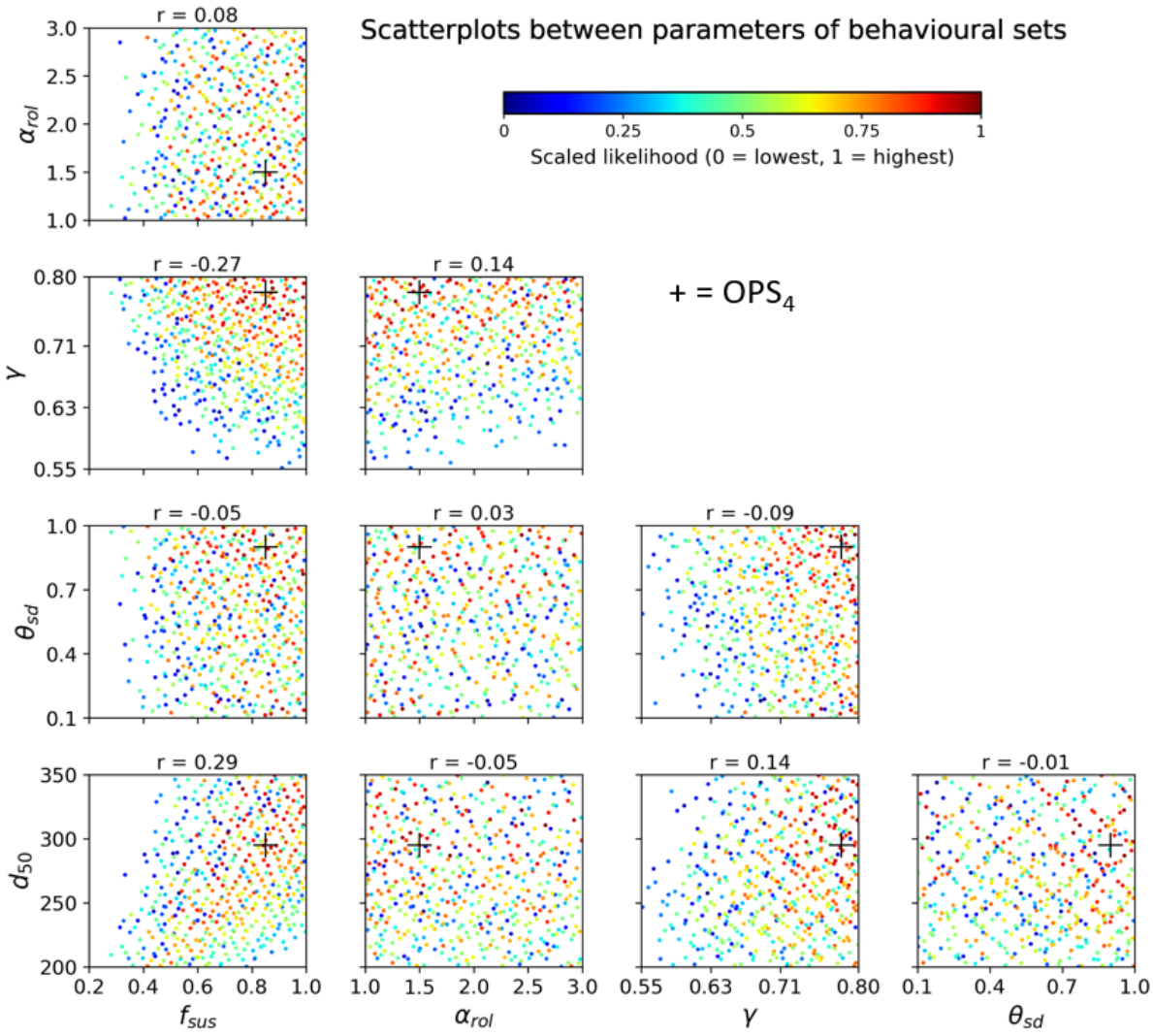

Figure 5.8: Overview of correlations (r) between each set of two parameters of the GLUE analysis. The correlations are based on the behavioural parameter sets for combined likelihood (507 simulations). The combined likelihood is represented by the colours of the dots (red implies a high likelihood simulation, blue a low likelihood). OPS_4 was selected based on a visual analyses of the correlation plots and is shown by the black cross in each plot.

5.4.2. Estimation of optimal parameter sets

In a next step, the posterior parameter distributions can be used to derive an optimal parameter set (OPS), which is a common goal of the GLUE method. The simplest way to do this is to take the value with the highest likelihood from the posterior distribution of each parameter. However, the posterior distributions are based on the likelihood values corresponding to individual parameter values. In other words, they do not take any interaction effects between parameters into account. Hence, such an OPS estimation can only be expected to work in case of independent and non-linear parameters. Between the five GLUE parameters, there are two notable correlations, between f_{sus} and d_{50} ($r = 0.29$) and between f_{sus} and γ ($r = -0.27$) (Figure 5.8). While not extreme, the correlations are existent and may play a role. However, they are based on simple correlations between the behavioural parameter sets, and do not take into account the model skill of each set. A way to combine model skill and parameter correlations in the estimation of the OPS would be to select a value for one parameter and create conditional posterior distributions for the other parameters, given the selected value for the first parameter. Theoretically, this process can be repeated after each parameter, creating new conditional distributions for the remaining ones. In practice, however, the sample size needs to be sufficiently large to create conditional distributions for multiple conditions.

For the parameter optimisation in this study, five OPS's have been estimated. The first set (OPS_0) is selected by simply taking the highest likelihood value for each parameter from the posterior distributions (Figure 5.7). Hence, it does not take the parameter correlations into account. The next three sets are estimated by selecting the highest likelihood value for f_{sus} (OPS_1), γ (OPS_2), and d_{50} (OPS_3), and estimating the remaining four parameters from conditional distributions based on the first parameter. To ensure a reasonable sample size (50–100 samples), the conditional distributions are created from all parameter sets, for which the value of the fixed parameter (f_{sus}, γ, or d_{50}) deviates by a maximum of 5% from the selected value (the 5% are relative to the applied parameter range). For example, for OPS_1, a value of 0.75 is selected for f_{sus} from its posterior distribution (Figure 5.7). The four other parameters are then estimated from a conditional posterior distribution, which considers only parameter sets for which $0.71 < f_{sus} < 0.79$. The derived conditional distributions are shown and explained in more detail in Appendix E.3.2. The final set (OPS_4) is estimated from a visual analysis of the correlation plots in Figure 5.8, which also show the model skill for each simulation. The parameter values are estimated by manually selecting areas in the correlation plots which contain high likelihood simulations (red colours). The selected values for the five OPS's are summarised in Table 5.4.

5.4.3. OPS evaluation

The five OPS's are evaluated by the same BSS metrics as the GLUE simulations (Table 5.4). BSS_{tot} is used as indicator of overall performance. Additionally, for each OPS simulation, the rank and percentile among the 1024 GLUE simulations is computed to assess the relative performance (see Appendix E, Table E.1). It follows that OPS_0, which was estimated without taking correlations into account, has the worst model performance of the OPS simulations ($BSS_{tot} = 0.71$). In fact, it only ranks in the 70^{th} percentile of all GLUE simulations, showing that it is important to consider the parameter correlations. The other four OPS's achieve similar model performances ($BSS_{tot} \approx 0.8$) and rank in the 95^{th} to 99^{th} percentiles. OPS_2, which is based on the conditional distributions for $\gamma \approx 0.78$, is the best scoring OPS for every considered output variable and the 6^th best simulation of all GLUE simulations (based on BSS_{tot}). This may further support the indication by the posterior parameter distributions that γ has the most influence on the model skill. Furthermore, all OPS's (except OPS_0) show an increase in model skill over the reference simulation by (Luijendijk et al., 2019). Especially for the shoreline position (BSS_{sl} and $BSS_{sl,spit}$), the model model skill for the OPS simulations is considerably higher (≈ 0.74) than for the reference simulation ($= 0.58$, see Table 5.4).

Notably, no OPS ranks in first place for any of the considered metrics. This is not necessarily surprising for the individual output variables, as the OPS's are based on the combined likelihood, which looks for the best combined model performance over all outputs. This is also reflected in that the OPS's generally rank higher for BSS_{tot} than for the individual outputs (Table E.1). Furthermore, the absolute difference in BSS_{tot} between the four OPS's (without OPS_0) and the best scoring GLUE simulations is in the order of 0.01, so it could be argued that the model performance is more or less equal. In that sense, it should not necessarily be seen as a failure that the parameter optimisation fails to identify a single absolutely optimal parameter set. In fact, due to equifinality, there exist multiple parameter combinations which lead to similar model skill and it is very likely that a single optimal set is non-existent (see also discussion in Section 6.3).

Even though the model skill of the four best OPS simulations is very similar, the results still differ for the various output variables. In fact, looking at the volume changes, the OPS simulations show significant variation among

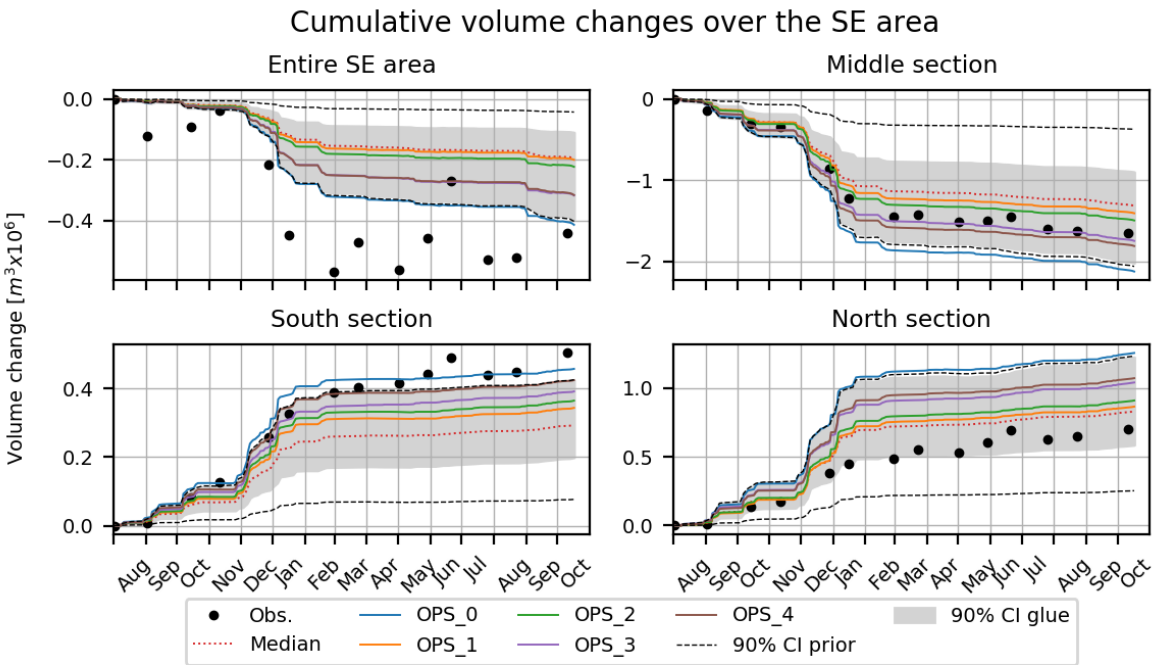

Figure 5.9: Time series of cumulative volume changes, showing the observations (black dots), 90% CI (grey area) and median (red, dotted line) of the behavioural GLUE simulations, prior 90% CI of all GLUE simulations (black dashed lines), and the OPS simulations (solid, coloured lines).

each other (Figure 5.9). The envelope created by all five OPS simulations covers more than half of the 90% CI for both the middle and North sections, where most of the morphological activity is predicted. Even disregarding OPS_0 the resulting envelope still covers around $1/3$ of the 90% CI. Considering that the parameter values of the OPS's are relatively close together, except for d_{50}—f_{sus} only varies between 0.75 and 0.85, γ between 0.76 and 0.78—the variation in the predicted volume changes is significant. For the shoreline position, the length of the spit also varies by up to 300 m in alongshore distance (see Appendix E Figure E.11). This shows that the simulations can achieve similar model skill with differing outputs, as shortcomings for one output category can be compensated by a better approximation for another output category.

Finally, we can consider the bathymetries for the OPS simulations to get a more visual overview of how they compare (Figure 5.10). Visually, the bathymetries differ mostly in the development of the spit. OPS_0 shows the most differences to the other simulations, with a strongly developed spit, which appears too large compared to the observation. OPS_1, OPS_2, and OPS_3 show very similar bathymetries, with only minor variations around the tip of the spit. These are also very similar to the highest BSS_z simulation (the HLS from Section 5.2), which could be expected from the similar model skill among the simulations. OPS_4 shows some more variation around the tip of the spit, while the reference simulation shows a thinner spit than the OPS simulations. Overall, the differences between the bathymetries of the simulations look less pronounced than for the volume changes, but this is based on a visual interpretation of the bathymetry plots only.

5.5. Uncertainty comparison

This section discusses the uncertainty comparison to get an indication of the relative importance of parameter uncertainty and wave climate variability. The spread in the results of the eight wave year simulations is compared to the uncertainty bounds from the GLUE analysis. For all eight simulations, OPS_0 is used (Table 5.3).

The cumulative volume changes of the wave climate simulations (Figure 5.11) show that there is considerable variation between the different wave years. Not only the final values of the volume changes differ, but also the morphologically active months vary significantly. This is because the time period with the highest wave energy (storm season) slightly differs among the years. Hence, the wave climate simulations induce variation at different points in the time series, as storms arrive earlier/later for some years than for others. This may also affect the length of the initial response phase (e.g. the year 2000 shows significant morphological activity in September 2001, 13 months into the simulation).

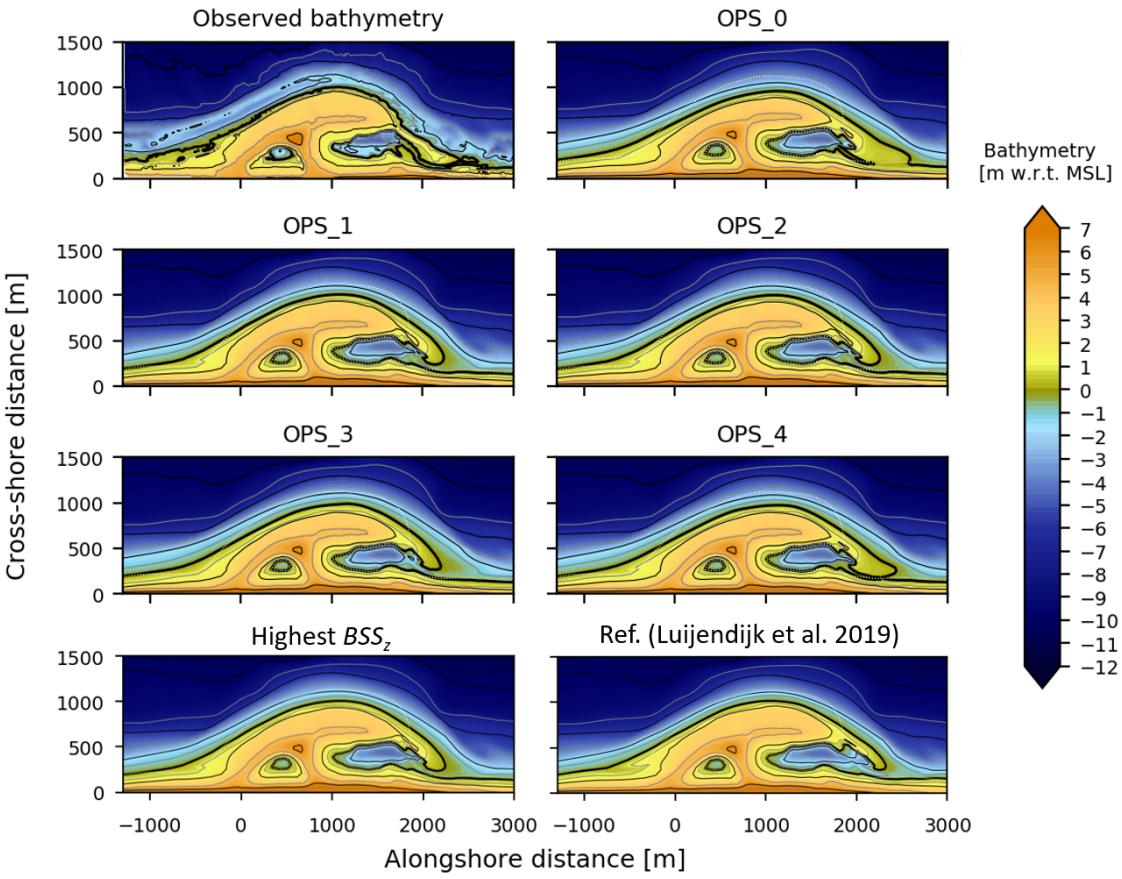

Figure 5.10: Observed bathymetry in October 2012 along with the predicted batymetries by the five OPS simulations, the GLUE simulation with the highest skill score for bed levels (BSS_z), and the reference simulation by Luijendijk et al. (2019).

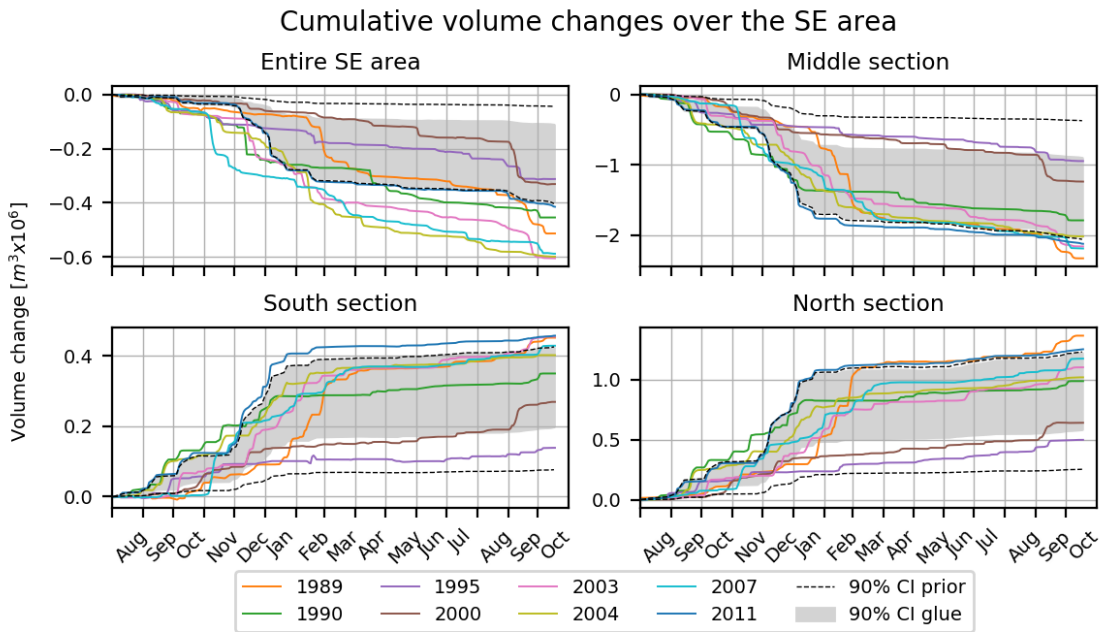

Figure 5.11: Time series of cumulative volume changes showing the 90% CI of the behavioural GLUE simulations (grey area) and all GLUE simulations (black dashed line) and the volume changes computed for the various wave climate simulations (coloured lines, see legend).

This is an important difference compared to parameter uncertainty—all GLUE simulations use identical wave conditions, hence time periods of high morphological activity are similar among the GLUE simulations. It is also visible when comparing the bandwidth of the two uncertainty sources (shown for the middle section in Figure 5.12). For simplicity, the bandwidth of the wave climate variability is defined as the envelope of the eight simulations (i.e. the maximum and minimum value at each point in time, considering all eight simulations). The bandwidth of the GLUE simulations is given by the 90% CI derived in Section 5.1. Uncertainty due to wave climate variability grows faster than parameter uncertainty. Especially in the first months (August to November), the ratio between the two bandwidths is unstable as it ranges between approximately 1.5 and 2.5 (i.e. the bandwidth of the wave climate is larger). As of December / January, the ratio appears to stabilise at around 1.25. Hence, over the first 14 months the bandwidth of the wave climate simulations is slightly larger than the 90% CI of the GLUE simulations, but in the same order of magnitude.

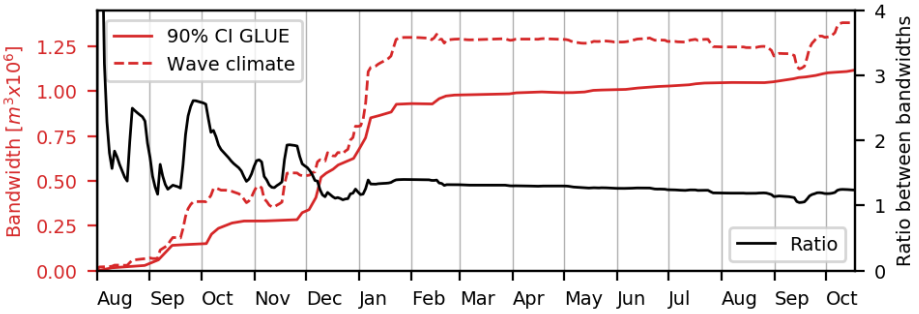

Figure 5.12: Time series showing the development in time of the bandwidth of the cumulative volume changes in the middle section. The bandwidth is given for parameter uncertainty (width 90% CI, solid red line) and wave climate variability (envelope of the 8 wave climate simulations, dashed red line). The ratio between the two is plotted on a second x-axis (solid back line).

The differences for the shoreline position reveal similar conclusions (Figure 5.13). Only the final time step (October 2012) is shown. As expected, variation among the different wave years is by far strongest in the spit area, similar to parameter uncertainty. In the spit area, variation due to wave climate appears stronger than for parameter uncertainty, but again in the same order of magnitude. Outside the spit area, however, the bandwidth of the wave climate variability is similar to or even smaller than the 90% CI of the GLUE analysis. This is also confirmed by the bathymetry plots in Figure 5.14. High-energy wave years generally lead to a longer spit, as they cause more sediment transport.

From the presented results it follows that the uncertainty sources are in the same order of magnitude over the 14-month study period, with the wave climate inducing, on average twice, as much uncertainty in the first months (August to November). However, it is difficult to draw absolute conclusions, as the GLUE data set is made up of 1024 simulations ($\approx$ 580 of which behavioural), whereas the wave climate variability is based on eight simulations—too few to conduct a full statistical analysis. Yet, the simulated wave years were chosen based on 25 years of wave data, selecting years with different characteristics (see Section 3.5). In that sense, one could even argue that a more statistically rigorous bandwidth of the wave climate variability (e.g. a 90% CI from 1000 simulations based on bootstrapped wave series) might be narrower than the approximation by the eight simulations, as there would be a higher fraction of 'average' years in the bootstrapped set. Furthermore, the 90% CI of the GLUE simulations also depends on the applied behavioural threshold. As shown in Figure 5.11, the width of the prior CI based on all 1024 GLUE simulations (dashed lines) would be larger than the envelope of the wave climate simulations.

What can be concluded from the presented results, is that parameter uncertainty is indeed a significant contributor to the overall prediction uncertainty. The width of the 90% CI is in the same order of magnitude as the approximated bandwidth for wave climate variability, hence not negligible in comparison. That wave climate variability might induce more uncertainty over the simulated 14 months, would also not be a surprise. In fact, this would be in line with the results of Kroon et al. (2020).

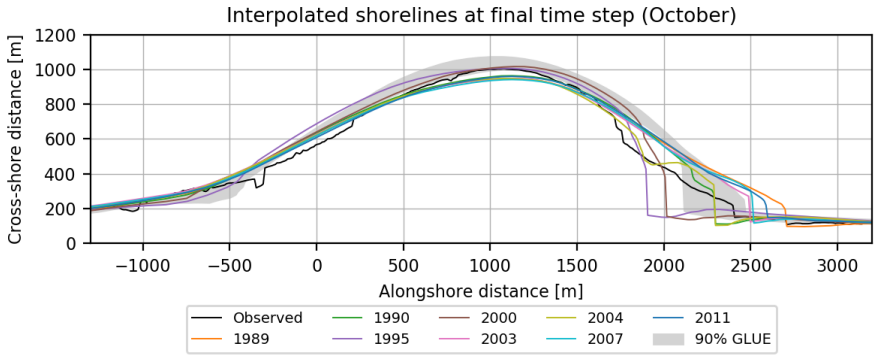

Figure 5.13: Interpolated shoreline position at the final time step (October of the simulated year). Plotted are the observed shoreline in October 2012 (black line), the 90% CI of behavioural GLUE simulations (grey area) and the shorelines of the wave climate simulations (coloured lines, see legend). Note that the year in the legend represents the year at the start of the simulation (i.e. the plotted shorelines are from October of the following year).

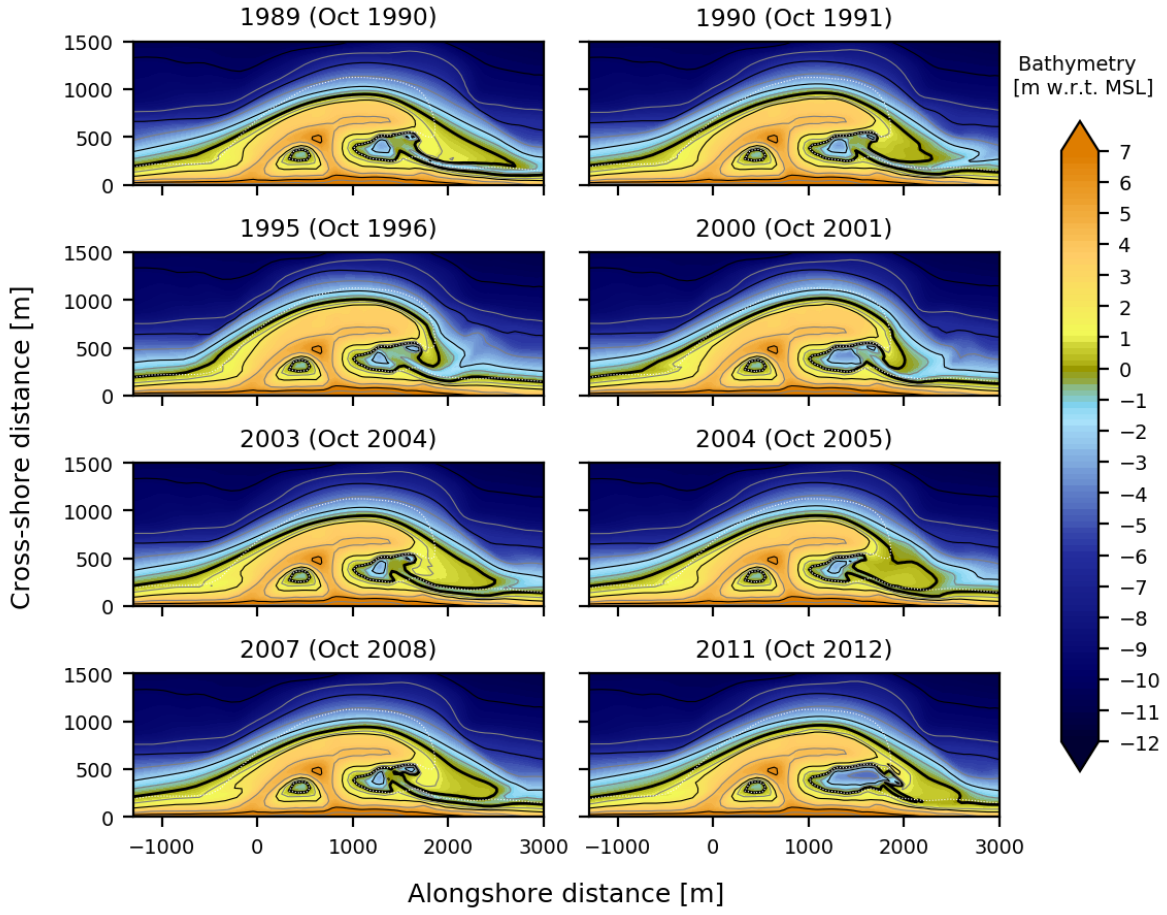

Figure 5.14: Bathymetry plots for the final time step (October) of each of the eight selected wave years. The plots show that the differences between the simulations are mostly confined to the spit area.

6

Discussion

Chapter outline
Chapter 6 provides a further discussion of the results presented in Chapters 4 and 5. Section 6.1 discusses the differences between the SA results and the parameter optimisation of the GLUE analysis and touches on what could be done to improve the SA results. Section 6.2 reflects on the subjectivity of the GLUE method along with some examples how it impacts the results of this study. Then, in Section 6.3, the parameter optimisation is discussed in more detail, specifically the OPS simulations and how the findings relate to equifinality. Next, some remaining model errors are discussed, which lie outside the parameter uncertainty analysed in this study. Finally, Section 6.5 reflects on the implications of the presented results on future mega-nourishment designs and modelling studies.

6.1. Differences in SA and UA results

A surprising observation is that the results of the GLUE analysis are not necessarily in agreement with those of the SA. The aim of the SA is to find the most influential model parameters on the selected output variables. The analysis identifies eight of the initial 16 parameters to be consistently influential, of which five are further examined in the GLUE analysis. The parameter optimisation shows, however, that for α_{rol} and θ_{sd} the posterior likelihood distribution remains similar to the prior uniform distribution (Section 5.4). This suggests that, based on the GLUE analysis, α_{rol} and θ_{sd} do not influence the model skill to the extent that may be expected from the SA results. The distributions for d_{50} and γ—similarly influential to α_{rol} and θ_{sd} according to the SA—are clearly non-uniform.

Several factors might provide a partial explanation. First of all, the SA examines a much larger parameter space (16 dimensions) than the GLUE analysis (five dimensions). Hence, 11 parameters which can vary in the SA, are fixed at constant values for all GLUE simulations. It is difficult to quantify how much this affects the results, but it is likely that it does to some extent. A second consideration is the sampling resolution. The larger parameter space of the SA is screened by a significantly smaller sample size (340 parameter sets against 1024 for GLUE). Moreover, the 340 samples are grouped into 20 trajectories. Hence, only 20 parameter sets are randomly sampled from the parameter space for the SA. While each trajectory provides 17 samples, these only differ in one parameter at a time, hence remain bound to a certain distance from the base vector of the trajectory. As a consequence, the distribution of parameter values among the samples is not uniform (see Figure D.2). The sampling resolution is further affected by the selection of the parameter ranges and the division of the ranges into four levels. The ranges have been chosen based on literature and expert judgement, with the aim to cover the full range of validity for the SE model. But, next to the upper and lower range limits, only two intermediate values are considered for each parameter, due to the division of the ranges into four levels. Consequently, some ranges might be too wide for the applied levels and sample size.

This discussion does not aim to imply that the SA fails to provide any valuable information—influential parameters were identified in f_{sus}, γ, and d_{50}. Still, the above suggests that the SA results appear to be based on an insufficient examination of the 16-dimensional parameter space. This also carries over into the UA (the GLUE results). The parameter optimisation shows that the output uncertainty of the five selected parameters can mainly be attributed to f_{sus}, γ and d_{50}, while α_{rol} and θ_{sd} have less effect. Hence, there might be parameters which were fixed for the GLUE analysis, but may have more influence on the output uncertainty than the examined parameters, specifically α_{rol} and θ_{sd}. For an additional GLUE analysis with new parameters, it would be best to first extend the SA with more simulations (the 20 existing trajectories can still be used), which are better distributed over the parameter space. The results of this supplemented SA can be compared to the existing GLUE results to assess the completeness of the SA. Subsequently, different parameters can be chosen based on the new, more representative SA results, which could be included in an additional GLUE analysis.

The most logical way to improve the SA, would be to increase the sampling resolution. There are several was to achieve this, including a larger sample size (i.e. more trajectories), narrower parameter ranges, the use of more levels to represent the parameter ranges, and a more uniform distribution of the samples over the parameter space. For instance, Campolongo et al. (2007) propose an adjusted sampling strategy for the EE method, which optimises the distribution of the sampled trajectories in the parameter space. While not increasing the total sample size, this approach leads to a more uniform distribution of the sampled parameter sets. The expectation is that this would provide more representative results. Ultimately however, the sample size has the most effect—for example, the use of more levels per parameter range will only improve the effective sampling resolution when combined with a larger sample size. Unfortunately, the author is currently unaware of globally applicable guidelines for combinations of the number of examined parameters, trajectories and levels to be used in the EE method.

6.2. Subjectivity of the GLUE method

The GLUE results presented in this thesis depend on several subjective decisions made during the setup of the GLUE analysis. Such decisions include the applied skill score and behavioural threshold, but also the selected parameters and ranges, and the definition of the performance indicators (i.e. the considered output variables). These choices determine the bounds within which the GLUE analysis operates. For example, the posterior parameter distributions are directly influenced by the behavioural threshold and the parameter ranges. The sampled ranges determine the limits within which the distributions exist, while the behavioural threshold affects the actual width of the distributions—if a low threshold is chosen, the posterior distribution will become wider,

while a high threshold will give a narrower distribution.

Statistically, the ideal scenario would be a posterior distribution which shows a global maximum at the most likely value, local maxima/minima (if any), and a decrease towards zero on both ends of the behavioural range. To achieve this, the modeller would have to select the full range of validity for the considered parameter. Especially for a non-linear model, however, it is difficult to know beforehand what range to sample for an input parameter in order to get the full posterior distribution. While the modeller can make certain assumptions based on the underlying physics of the parameters, there is no way to know for certain if the highest likelihood values represent the global optimum or a local optimum somewhere on the full (unknown) distribution. On the other hand, the wider the applied ranges for the parameters are, the lower the sampling resolution becomes (given a fixed amount of samples), and hence the lower the accuracy of the resulting distribution. One could therefore also argue that it is undesirable to seek the full posterior distribution as this will inevitably lead to many (expensive) model runs with low likelihoods. The difficulty is to find a balance between accuracy and completeness, given a limited sample size.

Similarly, the width of the uncertainty bounds also depends on the behavioural threshold and selected ranges. The uncertainty bounds shown for the volume changes (Figure 5.1) show both the 90% CI based on the likelihoods and behavioural threshold, as well as the prior 90% CI, which considers all 1024 GLUE simulations. Changing the parameter ranges would mainly affect the width of the prior CI. The behavioural threshold essentially determines how much of the prior CI is covered by the constrained uncertainty bounds (posterior CI). Again, a lower threshold would lead to wider uncertainty bounds, while a higher threshold would result in narrower bounds.

Therefore, the results of the GLUE analysis should not be seen as an absolute quantification of the parameter uncertainty in the model. Rather they give estimates of the uncertainty, based on the conditions set by the modeller. As long as this is accepted, the subjectivity of the GLUE method does not need to impede the results, but can actually be a strength, as it allows the modeller to adapt the method to specific models or boundary conditions. In this study, for example, the behavioural thresholds are set at different values for the different model outputs. While this might appear statistically informal, it enables the derivation of combined likelihood values. This would not be possible with a single threshold for all outputs, as that would lead to large differences in the number of behavioural simulations for the different outputs. Moreover, due to the difference in complexity and dimensionality of the outputs (the cumulative volume changes consider an aggregated quantity over a certain area, while the bed levels are considered point-wise in 2D) it would not be reasonable to expect similarly high BSS values for the different outputs. Hence, as long as the subjectivity is acknowledged, the GLUE results can give valuable insights into the parameter induced uncertainty.

6.3. Optimal parameter sets

One of the original aims of the parameter optimisation (Section 5.4) is to identify an optimal parameter set based on the posterior likelihood distributions. However, none of the estimated OPS simulations result in the best skill score of all GLUE simulations. Still, estimating the OPS by taking correlations into account (through the conditional posterior distributions) results in model skills similar to the best GLUE simulations. In fact, the 35 best GLUE simulations, including OPS_1, OPS_2, and OPS_4 all achieve a BSS_{tot} between 0.79 and 0.81. This confirms equifinality, as there are numerous different parameter sets resulting in a similar model skill, and may imply that the maximum achievable model skill—within the bounds of the selected parameter ranges—lies close to $BSS_{tot} = 0.81$, and that further efforts to find an OPS would continue resulting in a model skill near that value.

It is likely that by further refining the posterior distributions (e.g. creating additional conditional distributions and considering each output individually) one might eventually find an OPS with a slightly higher model skill than the best GLUE simulation, but it is debatable how useful such an optimal parameter set is. The posterior distributions depend on all fixed conditions (i.e. fixed parameters, applied forcing conditions, initial bathymetry etc.). Due to equifinality, any optimal parameter set emerging from the analysis, is one of many possible sets, which happens to have the highest likelihood for this particular set of conditions. If one were to change any of the conditions (e.g. adjust a fixed parameter or simulate a different time period), it is highly unlikely that the same exact optimal parameter set would emerge.

Hence, not identifying a single optimal parameter set, does not imply that the parameter optimisation has failed. On the contrary, it has succeeded in identifying parameter sets which lead to reliable model performance over the different outputs, and are likely close to the maximum achievable model skill. Moreover, the best simu-

lations already show an increase of model skill compared to the reference simulation (Luijendijk et al., 2019). Rather than continuing the search for an optimal parameter set for the current conditions, it seems more relevant to find reliable parameter values, which may still achieve a high model skill when some conditions are changed (e.g. different wave conditions). However, this is only possible by analysing the performance of a number of parameter sets under changing conditions, with accompanying observations to asses the model performance. One possibility for this would be to simulate a different time period of the SE with a number of selected parameter sets (i.e. a validation study of the parameter optimisation). If similar optimal parameter sets emerge under different conditions, that would increase the usefulness of such parameter sets for different studies.

6.4. Model imperfection

The uncertainty bounds from the GLUE analysis show that there are certain model imperfections which remain outside the influence of the examined parameter ranges. For example, the model fails to correctly predict the distribution of accretion volumes over the North and South sections (see Figure 5.1). Even for the OPS simulations and the best scoring GLUE simulations, accretion is overestimated in the North and underestimated in the South. Another example is the spit evolution. While practically all behavioural simulations predict the formation of a sand spit—some also correctly predict the deposited volume in the North section—none of the simulations correctly predict the exact location of the spit (see Figure 5.2). The spit is always predicted in a more seaward position than the observed spit. Generally, observations that fall outside the 90% CI of the uncertainty bounds hint to prediction errors which are not governed by the examined parameter ranges.

These errors may be caused by general model inadequacies, such as omitted processes, which are unrelated to the model parameters (e.g. aeolian transport, see Luijendijk et al., 2017). Another possibility is that they are caused by processes which remain unresolved in the applied discretisation (i.e. the spatial or temporal resolutions are too coarse). On the other hand, it is possible that some parameters that have not been considered (or even unexamined values for considered parameters) may partly influence these errors. For example, the applied model assumes a uniform d_{50} over the model domain. In reality, studies have shown that the grain size varies over the SE area, and is also affected by processes such as armouring and sediment sorting (e.g Huisman et al., 2016). Delft3D does allow the implementation of a space-varying d_{50}, which might potentially be able to address the mentioned model errors to some degree. This could be implemented in a GLUE analysis by setting a fixed relation between the d_{50} values of different areas, so the dimensionality of the examined parameter space is not increased.

In any case, even if all model parameters could be included, it is likely that there will still remain model imperfections, which are unrelated to parameter choices. An example of a simplified process, which might affect the distribution of volume changes between North and South is the applied BFFC approach. For high energy wave events at the Delfland coasts, the dominant long-term wave direction is from the Southwest sector, leading to an alongshore current towards the Northwest (Wijnberg, 2002). Hence, filtering lower wave heights from the wave series might decrease the fraction of waves coming from the Northwest sector and leading to a Southwest current, which may affect the distribution of sediment between the North and South sections of the SE.

Ultimately, when looking at the uncertainty bounds from the GLUE analysis, it is important to consider that they do not reflect the entire range of outcomes, but merely the possible outcomes that lie within the bounds of the selected parameters. There will always be a possibility that observed changes will lie outside the parameter uncertainty bounds.

6.5. Implications for future projects

The Sand Engine was constructed partly as a nourishment for the Dutch coast and partly as a research experiment, to assess the development of a localised mega-nourishment. Such a project is expensive and can be difficult to implement. While there have been studies to examine the development of potential mega-nourishments at other locations (e.g. Brown et al., 2016), decision-makers around the globe might be hesitant to approve more 'Sand Engines' without better prior knowledge and accurate predictions. In this light, it is interesting to see what this study can contribute to the implementation of future mega-nourishments. A modelling study of a future Sand Engine cannot be based on a GLUE method to quantify parameter uncertainty in predictions, as the method relies on observations, which are non-existent for future projects. Furthermore, running a set of 1024 Delft3D simulations might also not be possible for every design. Yet, this study has shown that predictions might show

significant spread in results depending on the applied parameter settings. So, how can we implement these findings in the design of the next Sand Engine?

Assuming similar beach characteristics as for the Delfland coast (i.e. a sandy, wave dominated coastline), the estimated OPS sets from this study may serve as a starting point for a calibration process of a future modelling study. It is important, however, not to limit such a calibration only to the parameters analysed in the presented GLUE analysis, but possibly include other parameters which were deemed influential in the SA (e.g. $f_{sus,w}$ or Fac_{DSS}) or even other parameters which may be identified as possible sources of parameter uncertainty, based on site specific characteristics. Furthermore, parameter values should never be blindly copied from this study, but always in consideration of site-specific beach characteristics and, possibly, expert judgement. For example, the d_{50} should be primarily based on the local sediment characteristics, not the values used in this thesis. In case a different model is used, the model parameters might differ, but it is likely that there will be parameters with similar effects as for Delft3D, for which the results of this study may be used as inspiration (as long as it is a process based model, which includes similar processes as Delft3D).

Currently, the main limitation to uncertainty analysis with complex models appears to be the resource intensity of these models. The model applied in this thesis uses a relatively strong morphodynamic acceleration approach (filtering $\approx 50\%$ of the wave conditions, see Luijendijk et al., 2019), yet it still took 4–5 days to simulate a time span of 14 months—far shorter than the 20-year design cycle of the Sand Engine. In this light, an interesting finding from this study is the correlation between the initial response phase—in which most morphological changes occur—and the development of parameter uncertainty in the predictions. The existence of an initial response phase is confirmed by observations (de Schipper et al., 2016), and future Sand Engines may experience a similar effect. The different wave year simulations show that this initial response phase may differ in length, however, even for different wave conditions at the same location. Hence, it is likely that the length of the initial response phase might vary for other locations, depending on the sediment characteristics and the forcing conditions (e.g. seasonal variability in the wave climate, magnitude of tidal forces etc.). A new approach to enable longer simulation periods might be to use a complex model to simulate the initial response phase (and potentially high energy wave periods), while switching to a simpler model approach for the remaining time periods. An interesting opportunity might present itself in emulation / surrogate modelling (see also Section 2.4). Specifically, a multi-fidelity approach might provide a promising path. Hereby, the applied surrogate model has the same basic structure as the original complex model, but is applied in a simpler manner. This could be in the form of a lower grid resolution (e.g. Berends et al., 2019), different acceleration techniques, or other simplifications. For example, one could apply BF-based acceleration techniques during the initial response phase, while switching to a computationally cheaper LST approach for the remainder of the simulation period (Luijendijk et al., 2019).

That said, these ideas are based on the results of a 14-month study period, which only includes one storm season, so it should first be researched how larger time-scales affect the presented results. Furthermore, parameter uncertainty only forms one of several uncertainty sources, which should be included in design processes. So while this study presents new information on parameter uncertainty, any design study for a new Sand Engine should include an integral uncertainty analysis which takes into account different uncertainty sources, and how they interact with each other. Section 7.2 gives some recommendations for future research on the ideas discussed in this chapter.

7

Conclusion

Chapter outline
The final chapter of this thesis summarises the conclusions of the research. The original objective—to investigate how, and to what extent, parameter uncertainty influences coastal area predictions of mega-nourishments—was split into four guiding research questions. The conclusions to each question are presented and discussed in Section 7.1, along with an overarching conclusion. Lastly, several recommendations for further research are given in Section 7.2.

7.1. Conclusions

The objective of this thesis was to *investigate how, and to what extent, parameter uncertainty influences coastal area predictions of mega-nourishment*. This was done by creating a large synthetic dataset of morphological Delft3D predictions of the Sand Engine, a localised mega-nourishment along the Dutch coast. First, a sensitivity analysis was conducted to find out which parameters are the most influential and could therefore be considered the largest sources of output uncertainty—sixteen parameters were analysed by the elementary effects method, using 340 Delft3D simulations. The results were used to select five parameters for an uncertainty analysis, examining 1024 Delft3D simulations by means of the *Generalised Likelihood Uncertainty Estimation*: each simulation was assigned a likelihood score based on model performance (BSS), with non-behavioural simulations (BSS below a certain threshold) receiving zero likelihood. The likelihood scores were then used to derive observation based uncertainty bounds for the model outputs and posterior likelihood distributions for the model parameters.

The research objective was split into four research questions. The conclusions to each individual question are discussed below.

1. **Which model parameters contribute most to the total parameter uncertainty?**

   According to the results of the SA (Section 4.2), the suspended sediment transport scaling factor (f_{sus}) is the dominant parameter (1.5–2 times more influential on the considered model outputs than the rest of the parameters). Next follows a group of seven parameters which are consistently influential over all three output categories: $f_{sus,w}$, α_{rol}, β_{rol}, γ, θ_{sd}, d_{50} and Fac_{DSS}. The remaining eight parameters are less influential. Based on absolute influence as well as an effort to minimise potential reciprocity effects, five parameters were selected for the GLUE analysis (Section 4.3): f_{sus}, α_{rol}, γ, θ_{sd}, and d_{50}.

   The posterior likelihood distributions of the GLUE analysis (Section 5.4) do not entirely confirm the SA results, as two parameters (α_{rol} and θ_{sd}) receive near-uniform posterior distributions. For the other three parameters (f_{sus}, γ, and d_{50}), the posterior distributions show clear trends, indicating that they are important uncertainty sources. The near-uniform distributions of α_{rol} and θ_{sd} could be a consequence of equifinality, which is generally caused by parameter interdependence, non-linearity or model insensitivity and is a common issue in GLUE analyses (e.g. Ruessink, 2005; Simmons et al., 2017). However, neither α_{rol} nor θ_{sd} appeared to be strongly correlated to another parameter from the GLUE analysis (Figure 5.8), suggesting it is caused by non-linearity and/or model insensitivity.

   The parameter optimisation shows that parameter correlations should be taken into account when deriving optimal parameter sets from the posterior distributions. The so-estimated parameter sets resulted in similar model skill as the best GLUE simulations ($BSS_{tot} \approx 0.8$), suggesting that this may be close to the maximum achievable model performance within the considered parameter bounds, and confirming the presence of equifinality. Still, the differences in individual outputs for simulations with equal model skill, show that the model is sensitive, even to small changes in the influential model parameters.

   While important sources of parameter uncertainty are identified, the results of the SA and GLUE show clear differences. The GLUE results are more trustworthy as they are based on a more accurate representation of the parameter space. It is likely that the applied sample size for the SA is too small and not sufficiently well distributed over the 16-dimensional parameter space. Combined with potential interaction effects and non-linearities this may cause an incomplete representation of parameter sensitivities. Of the five considered parameters in the GLUE analysis, f_{sus} and γ contribute most to parameter uncertainty. However, the difference between the GLUE and SA results suggests that other parameters, which were fixed for the GLUE analysis, could also form significant contributions to parameter uncertainty.

2. **How does the uncertainty in the selected parameters translate to uncertainty bounds in the predictions?**

   The uncertainty bounds derived from the GLUE results (Section 5.1) provide an uncertainty estimation which is constrained by observations (the CI's are based on the likelihoods). Uncertainty in cumulative erosion volumes of the peninsula after 14 months shows a spread of 1.3 million m^3 $-400.000/+700.000$ m^3. Corresponding accretion volumes in the adjacent South and North sections show spreads of 290.000 m^3 $-100.000/+120.000$ m^3 and 830.000 m^3 $-230.000/+400.000$ m^3, respectively. For all three sections, the width of the 90% CI is over 75% of the median predicted change. Variation in the predicted shoreline position is largest in the spit area, where the 90% CI covered up to 350 m in cross-shore direction. This is associated to the development of the sand spit in the North section, which is a complex process and

difficult to predict for the model. For the remaining shoreline, parameter induced uncertainty is limited ($W_{CI90} \approx 100m$). Finally, uncertainty bounds for predicted bed level changes around the SE show a spread of up to 6 m around the median (which reaches up to 5m of absolute change).

A temporal analysis of the output uncertainty shows that, for all considered outputs, uncertainty mainly forms in the first seven months of the study period (August 2011 to February 2012), although this is most pronounced for the volume changes, and less for the spatially distributed bed levels (Section 5.3). This is also the period in which most of the morphological changes occur, which is in line with the results of de Schipper et al. (2016). During this initial response phase, the uncertainty growth in time correlates strongly to the magnitude of the morphological changes (e.g. $r > 0.95$ for volume changes), which in turn is correlated to the wave energy ($r > 0.9$ for volume changes and $r \approx 0.67$ for the shoreline position). Hence, most uncertainty develops during the storm season from December to January. In the final seven months (March to August 2012), the SE appears to have reached a state where morphological changes are slower, more nuanced and less dependent on the wave energy. In this period the width of the uncertainty bounds for the volume changes only grows by about 10% compared to the first seven months.

The uncertainty estimates are influenced by several subjective choices during the GLUE method (e.g. BSS threshold, parameter ranges). They should therefore not be seen as statistically rigorous uncertainty bounds, but rather as valuable estimates of uncertainty. As such, the results indicate that uncertainty in model parameters translates to significant uncertainty in predicted outputs. This is especially true for complex, three-dimensional features such as the sand spit, making parameter uncertainty particularly important for nature-based designs, which incorporate complex natural processes. This shows that we cannot rely on deterministic predictions, which give no indication of the parameter induced uncertainty, which is in line with similar conclusions for simpler models by Kroon et al. (2020) and Simmons et al. (2017).

3. **How does the output uncertainty vary over the study area?**

   The spatial uncertainty maps presented in Sections 5.2 and 5.3 plot the width of the 90% CI (W_{CI90}) for cumulative bed level changes over the model domain. The maps show the predicted bathymetry with the highest model skill for bed level changes (BSS_z), along with expected bed level changes and accompanying uncertainty. Areas of high uncertainty ($W_{CI90} \approx 5 - 6m$) are found along the head of the peninsula and around the tip of the sand spit. Beyond the $6m$ depth-contour, uncertainty rapidly decreases to negligible levels. Uncertainty correlates with the magnitude of the expected change ($r \approx 0.65$ between the median bed level change and W_{CI90}). There are, however, also locations with high uncertainty ($W_{CI90} \approx 4m$) but low expected changes, most notably in the spit area.

   The spatial uncertainty maps were created with the intention to provide a tool to communicate uncertainty in predictions of designs to the relevant stakeholders and decision-makers. Instead of only showing a deterministic prediction, the maps show the most likely outcome (based on the analysed parameter sets), along with accompanying uncertainty levels. Ideally, such maps should be extended to include other uncertainty sources.

4. **How does the parameter induced uncertainty compare to uncertainty induced by variability in the wave climate?**

   The spread in the results of the eight wave climate simulations and the width of the uncertainty bounds for the GLUE analysis are in the same order of magnitude for both the volume changes and the shoreline position (Section 5.5). The wave climate variability generally exceeds the parameter induced uncertainty bounds by a factor of 1.25 to 1.5. This implies that both uncertainty sources form important contributions to the total prediction uncertainty, and that neither can be neglected.

   This further confirms that parameter uncertainty is an important uncertainty source, and that the common assumption in coastal modelling studies, that intrinsic uncertainty is more important than epistemic uncertainty, is not necessarily true. The fact that the wave climate variability appears slightly more important over the first 14 months is also in line with the findings of Kroon et al. (2020), who observe the relative importance of parameter uncertainty to increase over time. That said, it should be noted that the comparison in this thesis was based on eight simulations representing the wave climate variability against 1024 simulations for parameter uncertainty. Hence, the results should only be seen as a preliminary indication, not as absolute conclusions.

Considering the overarching research objective, an answer was found for all individual research questions. However, the first question is not yet completely answered. The parameter selection for the GLUE analysis was based on the SA results, which proved to give an incomplete picture of the model sensitivity. This may also result in an incomplete representation of the uncertainty bounds, as some of the fixed parameters may yet prove to be significant uncertainty sources. In that sense, the research questions are answered for the considered parameters, but the results can be further extended by including more parameters.

The findings of this thesis have an impact on two key levels. First, they can be used to communicate and address uncertainty in predictions of coastal change. For example, the spatial uncertainty maps can let stakeholders understand the potential range of outcomes for a certain design. Secondly, the results, combined with the created dataset, provide valuable information on how parameter uncertainty affects our predictions of coastal mega-nourishments, which can be used for future morphodynamic studies. Hence, overall, the research objective of this thesis is achieved.

7.2. Future research

Based on the findings of this report, several recommendations for future research are presented.

1. **Extend the presented analysis with additional simulations**

   An extended SA might offer more accurate insights in the importance of the model parameters and reveal if there remain parameters which may form significant contributions to parameter uncertainty. These could then be analysed by a second GLUE analysis, switching the less important parameters (α_{rol} and θ_{sd}) with more influential ones. Also, the ranges of the examined, influential parameters (f_{sus}, γ, and d_{50}) may be adjusted, based on the results of this thesis.

2. **Validation of optimal parameter sets**

   The parameter optimisation, which resulted in optimal parameter sets (or ranges, due to equifinality) is based on a fixed set of conditions applied in this thesis. To examine the reliability of these parameter values, a validation study should be carried out, which assesses their performance under changing conditions. If such a validation is successful, it increases the chance that the found parameter values are applicable to future studies. A validation could be performed by simulating a different time span of the SE evolution, for example.

3. **Continue exploration of the created dataset**

   The synthetic dataset in this thesis provides the model outputs of around 1350 morphological Delft3D predictions (including both the SA and UA simulations). These outputs include numerous variables in several dimensions, providing an enormous amount of data that can be analysed. The presented results have focused on three outputs derived from bed level changes over the SE area, but there are many possibilities to derive other performance indicators (cross-shore profiles, local flow velocity and sediment transport patterns, wave forces etc.). This may also include the derivation of more suitable and sophisticated performance indicators for complex processes such as the evolution of the sand spit. Even the outputs derived for this thesis can be analysed further, exploring additional correlations and patterns. The exploration of the generated dataset is far from exhausted and might offer further insights into the behaviour of the model and the development of uncertainty.

4. **Inclusion of more uncertainty sources**

   A logical next step to this research would be to include additional sources of uncertainty, to work towards an integral quantification of prediction uncertainty. Ultimately, an uncertainty map encompassing all (or most) uncertainty sources could be an extremely useful tool in decision-making and stakeholder communication. This thesis has briefly investigated the relative importance of parameter uncertainty and variability in the wave climate (Section 5.5) but in a very simplified way. This could be improved by carrying out a more thorough analysis, creating a dataset to examine variations in parameter values, as well as wave forcing. Hereby, a bootstrapping approach could be applied, analogous to Kroon et al. (2020).

5. **Analyse correlation between uncertainty sources**

   The comparison between parameter uncertainty and wave climate variability in this thesis was a direct comparison of bandwidths. However, the temporal analysis already shows that the development of uncertainty bounds is closely related to variations in morphological changes and wave climate. In reality,

different uncertainty sources might interact with each other and influence each other's contribution to total prediction uncertainty. Correlation between uncertainty sources has also been observed by Kroon et al. (2020) in their one-line model.

6. **Increase temporal scale**

   The quantification of parameter uncertainty in this thesis is based on a 14-month study period, starting after construction of the SE. However, there can be significant temporal variation in the relative importance of different uncertainty sources. Specifically, Kroon et al. (2020) observe that the relative importance of parameter uncertainty in their one-line model of the SE increases over time. The results of the uncertainty comparison show that different wave years induce different trends in morphological activity, which can be expected to translate to different trends in uncertainty development. Hence it would be interesting to extend the uncertainty analysis to larger time scales. Extending to time scales in the order of several decades, would even allow the inclusion of longer-term uncertainty sources, such as different climate change and SLR scenarios.

7. **Spatial maps and visualisation of uncertainty**

   This thesis has explored a first concept of a spatial uncertainty map which may be used to communicate uncertainty in prediction to stakeholders. However, these maps contain many variables and there might be better ways of presenting them, for example by combining expected changes and corresponding uncertainty in a single map using bivariate colour-scales. An example might be found in Correll et al. (2018) who use *value suppressing uncertainty palettes* for combined value-uncertainty plots.

8. **Surrogate modelling**

   As described in Section 6.5, surrogate modelling might present an interesting opportunity to cutting the resource intensity of future modelling studies. To possibly implement this in a design approach for future Sand Engines, research is needed on the feasibility of such an approach in Delft3D or other process based area models. Is it possible to, and what are the consequences of switching to a different acceleration technique during a simulation? Does this introduce additional uncertainties? Might different parameters become dominant for different approaches? How can we include other uncertainty sources in such a modelling approach, without requiring extreme amounts of model runs to sufficiently explore the response surface? Such questions should be examined before implementing such a new approach in a design process.

Bibliography

Arendt, P. D., Apley, D. W., & Chen, W. (2012). Quantification of Model Uncertainty: Calibration, Model Discrepancy, and Identifiability. *Journal of Mechanical Design*, *134*(10). https://doi.org/10.1115/1.4007390

van Arkel, M. (2016). *Towards an efficient sensitivity analysis of wave forcing in coastal erosion studies (MSc Thesis)*. Delft University of Technology. http://repository.tudelft.nl/islandora/object/uuid%3A3bfe1ad0-426d-4671-810f-2110724c3385

Baart, F., van Gelder, P. H. A. J. M., & van Koningsveld, M. (2011). Confidence in real-time forecasting of morphological storm impacts. *Journal of Coastal Research*, (SI 64), 1835–1839. https://www.jstor.org/stable/26482494

Beiser, V. (2019). Why the world is running out of sand. *BBC Future*. https://www.bbc.com/future/article/20191108-why-the-world-is-running-out-of-sand

Berends, K. D., Scheel, F., Warmink, J. J., de Boer, W. P., Ranasinghe, R., & Hulscher, S. J. (2019). Towards efficient uncertainty quantification with high-resolution morphodynamic models: A multifidelity approach applied to channel sedimentation. *Coastal Engineering*, *152*, 103520. https://doi.org/10.1016/j.coastaleng.2019.103520

Beven, K. (1993). Prophecy, reality and uncertainty in distributed hydrological modelling. *Advances in Water Resources*, *16*(1), 41–51. https://doi.org/10.1016/0309-1708(93)90028-E

Beven, K., & Binley, A. (1992). The future of distributed models: Model calibration and uncertainty prediction. *Hydrological Processes*, *6*(3), 279–298. https://doi.org/10.1002/hyp.3360060305

Beven, K., & Binley, A. (2014). GLUE: 20 years on. *Hydrological Processes*, *28*(24), 5897–5918. https://doi.org/10.1002/hyp.10082

Blasone, R. S., Madsen, H., & Rosbjerg, D. (2008). Uncertainty assessment of integrated distributed hydrological models using GLUE with Markov chain Monte Carlo sampling. *Journal of Hydrology*, *353*(1-2), 18–32. https://doi.org/10.1016/j.jhydrol.2007.12.026

Blasone, R. S., Vrugt, J. A., Madsen, H., Rosbjerg, D., Robinson, B. A., & Zyvoloski, G. A. (2008). Generalized likelihood uncertainty estimation (GLUE) using adaptive Markov Chain Monte Carlo sampling. *Advances in Water Resources*, *31*(4), 630–648. https://doi.org/10.1016/j.advwatres.2007.12.003

Blumberg, A. F., & Georgas, N. (2008). Quantifying Uncertainty in Estuarine and Coastal Ocean Circulation Modeling. *Journal of Hydraulic Engineering*, *134*(4), 403–415. https://doi.org/10.1061/(ASCE)0733-9429(2008)134:4(403)

Booij, N., Ris, R. C., & Holthuijsen, L. H. (1999). A third-generation wave model for coastal regions 1. Model description and validation. *Journal of Geophysical Research: Oceans*, *104*(C4), 7649–7666. https://doi.org/10.1029/98JC02622

Bounford.com, & UNEP/GRID-Arendal. (2006). Coastal population and shoreline degradation. https://www.grida.no/resources/5563

Brière, C., Giardino, A., & van der Werf, J. J. (2011). Morphological modeling of bar dynamics with Delft3D: The quest for optimal free parameter settings using an automatic calibration technique. *Proceedings of the 32nd Coastal Engineering Conference*. https://doi.org/10.9753/icce.v32.sediment.60

Brown, J. M., Phelps, J. J., Barkwith, A., Hurst, M. D., Ellis, M. A., & Plater, A. J. (2016). The effectiveness of beach mega-nourishment, assessed over three management epochs. *Journal of Environmental Management*, *184*, 400–408. https://doi.org/10.1016/j.jenvman.2016.09.090

Bruun, P. (1962). Sea-Level Rise as a Cause of Shore Erosion. *Journal of the Waterways and Harbors Division*. https://doi.org/10.5860/choice.47-1443

Callaghan, D. P., Ranasinghe, R., & Roelvink, D. (2013). Probabilistic estimation of storm erosion using analytical, semi-empirical, and process based storm erosion models. *Coastal Engineering*, *82*, 64–75. https://doi.org/10.1016/j.coastaleng.2013.08.007

Campolongo, F., Cariboni, J., & Saltelli, A. (2007). An effective screening design for sensitivity analysis of large models. *Environmental Modelling and Software*, *22*(10), 1509–1518. https://doi.org/10.1016/j.envsoft.2006.10.004

Castelletti, A., Galelli, S., Ratto, M., Soncini-Sessa, R., & Young, P. (2012). A general framework for Dynamic Emulation Modelling in environmental problems. *Environmental Modelling & Software*, *34*, 5–18. https://doi.org/10.1016/j.envsoft.2012.01.002

Church, J., Clark, P. U., Cazenave, A., Gregory, J., Jevrejeva, S., Levermann, A., Merrifield, M., Milne, G., Nerem, R., Nunn, P., Payne, A., Pfeffer, W., Stammer, D., & Unnikrishnan, A. (2013). Sea level change. In T. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, & P. Midgley (Eds.), *Climate change 2013: The physical science basis. contribution of working group i to the fifth assessment report of the intergovernmental panel on climate change* (pp. 1137–1216). Cambridge University Press. https://www.ipcc.ch/site/assets/uploads/2018/02/WG1AR5_Chapter13_FINAL.pdf

Correll, M., Moritz, D., & Heer, J. (2018). Value-Suppressing Uncertainty Palettes. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*. https://doi.org/10.1145/3173574.3174216

Deltares. (2011a). *Delft3D-FLOW User Manual* (3.15). Deltares.

Deltares. (2011b). *Wave look-up table: Building with Nature* (tech. rep.).

van Gelder, P. H. A. J. M. (2000). *Statistical methods for the risk-based design of civil structures* (Doctoral dissertation). Delft University of Technology. http://resolver.tudelft.nl/uuid:6a62d6fa-cbcc-4c38-af8a-027c3d191a9d

Glasserman, P. (2003). *Monte Carlo Methods in Financial Engineering* (Vol. 53). Springer New York. https://doi.org/10.1007/978-0-387-21617-1

Hall, J. W., Manning, L. J., & Hankin, R. K. S. (2011). Bayesian calibration of a flood inundation model using spatial data. *Water Resources Research*, *47*(5). https://doi.org/10.1029/2009WR008541

Hanson, H., Brampton, A., Capobianco, M., Dette, H., Hamm, L., Laustrup, C., Lechuga, A., & Spanhoff, R. (2002). Beach nourishment projects, practices, and objectives—a European overview. *Coastal Engineering*, *47*(2), 81–111. https://doi.org/10.1016/S0378-3839(02)00122-9

Harrison, P., & Pearce, F. (2001). *AAAS atlas of population & environment*. American Association for the Advancement of Science, University of California Press.

Herman, J., Usher, W., Mutel, C., Trindade, B., Hadka, D., Woodruff, M., Rios, F., & Hyams, D. (2019). SALib. https://salib.readthedocs.io/en/latest/index.html

Holthuijsen, L. H. (2007). *Waves in Oceanic and Coastal Waters*. Cambridge University Press. https://doi.org/10.1017/CBO9780511618536

Huisman, B. J. A., de Schipper, M. A., & Ruessink, B. G. (2016). Sediment sorting at the Sand Motor at storm and annual time scales. *Marine Geology*, *381*, 209–226. https://doi.org/10.1016/j.margeo.2016.09.005

Johnson, M. E., Moore, L. M., & Ylvisaker, D. (1990). Minimax and maximin distance designs. *Journal of Statistical Planning and Inference*, *26*(2), 131–148. https://doi.org/10.1016/0378-3758(90)90122-B

Jonkman, S. N., Vrouwenvelder, A. C. W. M., Steenbergen, R. D. J. M., Morales-Nápoles, O., & Vrijling, J. K. (2017). *Probabilistic Design: Risk and Reliability Analysis in Civil Engineering, Lecture notes CIE4130* (M. H. G. Baas, M. A. van der Lugt, & A. Kosters, Eds.). Delft University of Technology.

Kabat, P., Fresco, L. O., Stive, M. J., Veerman, C. P., Van Alphen, J. S., Parmet, B. W., Hazeleger, W., & Katsman, C. A. (2009). Dutch coasts in transition. *Nature Geoscience*, *2*(7), 450–452. https://doi.org/10.1038/ngeo572

Khare, Y., Martinez, C. J., Muñoz-Carpena, R., Bottcher, A., & James, A. (2019). Effective global sensitivity analysis for high-dimensional hydrologic and water quality models. *Journal of Hydrologic Engineering*, *24*(1), 1–19. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001726

Khu, S. T., & Werner, M. G. (2003). Reduction of Monte-Carlo simulation runs for uncertainty estimation in hydrological modelling. *Hydrology and Earth System Sciences*, *7*(5), 680–692. https://doi.org/10.5194/hess-7-680-2003

Kroon, A., de Schipper, M., den Heijer, K., Aarninkhof, S., & van Gelder, P. (2017). Uncertainty Assessment in Coastal Morphology Prediction With a Bayesian Network. In T. Aagaard, R. Deigaard, & D. Fuhrman (Eds.), *Proceedings of coastal dynamics 2017*.

Kroon, A., de Schipper, M. A., van Gelder, P. H., & Aarninkhof, S. G. (2020). Ranking uncertainty: Wave climate variability versus model uncertainty in probabilistic assessment of coastline change. *Coastal Engineering*, *158*, 1–11. https://doi.org/10.1016/j.coastaleng.2020.103673

Kucherenko, S., Albrecht, D., & Saltelli, A. (2015). *Exploring multi-dimensional spaces: a Comparison of Latin Hypercube and Quasi Monte Carlo Sampling Techniques*. http://arxiv.org/abs/1505.02350

Leatherman, S. P., Zhang, K., & Douglas, B. C. (2000). Sea level rise shown to drive coastal erosion. *Eos*, *81*(6), 55–57. https://doi.org/10.1029/00EO00034

Lesser, G. R., Roelvink, J. A., van Kester, J. A., & Stelling, G. S. (2004). Development and validation of a three-dimensional morphological model. *Coastal Engineering*, *51*(8-9), 883–915. https://doi.org/10.1016/j.coastaleng.2004.07.014

Li, L., Storms, J. E., & Walstra, D. J. (2018). On the upscaling of process-based models in deltaic applications. *Geomorphology*, *304*, 201–213. https://doi.org/10.1016/j.geomorph.2017.10.015

Likhachev, D. V. (2019). Parametric sensitivity analysis as an essential ingredient of spectroscopic ellipsometry data modeling: An application of the Morris screening method. *Journal of Applied Physics*, *126*, 184901. https://doi.org/10.1063/1.5126074

Losada, M. A., Baquerizo, A., Ortega-Sánchez, M., & Ávila, A. (2011). Coastal Evolution, Sea Level, and Assessment of Intrinsic Uncertainty. *Journal of Coastal Research*, *59*(10059), 218–228. https://doi.org/10.2112/SI59-023.1

Luijendijk, A. P. (2019). *Crossing borders in coastal morphodynamic modelling* (Doctoral dissertation). Delft University of Technology. Delft University of Technology. https://doi.org/10.4233/uuid:75ac8d9e-293f-4bff-90b3-467010352032

Luijendijk, A. P., Hagenaars, G., Ranasinghe, R., Baart, F., Donchyts, G., & Aarninkhof, S. (2018). The State of the World's Beaches. *Scientific Reports*, *8*(6641). https://doi.org/10.1038/s41598-018-24630-6

Luijendijk, A. P., Ranasinghe, R., de Schipper, M. A., Huisman, B. A., Swinkels, C. M., Walstra, D. J., & Stive, M. J. (2017). The initial morphological response of the Sand Engine: A process-based modelling study. *Coastal Engineering*, *119*, 1–14. https://doi.org/10.1016/j.coastaleng.2016.09.005

Luijendijk, A. P., de Schipper, M. A., & Ranasinghe, R. (2019). Morphodynamic acceleration techniques for multi-timescale predictions of complex sandy interventions. *Journal of Marine Science and Engineering*, *7*(3), 78. https://doi.org/10.3390/jmse7030078

Mantovan, P., & Todini, E. (2006). Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology. *Journal of Hydrology*, *330*(1-2), 368–381. https://doi.org/10.1016/j.jhydrol.2006.04.046

van Maren, D. S., & Cronin, K. (2016). Uncertainty in complex three-dimensional sediment transport models: equifinality in a model application of the Ems Estuary, the Netherlands. *Ocean Dynamics*, *66*(12), 1665–1679. https://doi.org/10.1007/s10236-016-1000-9

Menéndez, M., Méndez, F. J., Izaguirre, C., Luceño, A., & Losada, I. J. (2009). The influence of seasonality on estimating return values of significant wave height. *Coastal Engineering*, *56*(3), 211–219. https://doi.org/10.1016/j.coastaleng.2008.07.004

Morris, M. D. (1991). Factorial sampling plans for preliminary computational experiments. *Technometrics*, *33*(2), 161–174. https://doi.org/10.1080/00401706.1991.10484804

Mulder, J. P. M., van de Kreeke, J., & van Vessem, P. (1995). Experimental Shoreface Nourishment, Terschelling (NL). *Coastal Engineering 1994*, *3*, 2886–2899. https://doi.org/10.1061/9780784400890.209

Oreskes, N., Shrader-Frechette, K., & Belitz, K. (1994). Verification, validation, and confirmation of numerical models in the earth sciences. *Science*, *263*(5147), 641–646. https://doi.org/10.1126/science.263.5147.641

Plant, N. G., & Holland, K. T. (2011). Prediction and assimilation of surf-zone processes using a Bayesian network. *Coastal Engineering*, *58*(1), 119–130. https://doi.org/10.1016/j.coastaleng.2010.09.003

Pronzato, L. (2017). Minimax and maximin space-filling designs: some properties and methods for construction. *Journal de la Société Française de Statistique*, *158*(1), 7–36.

Radermacher, M., de Schipper, M. A., Price, T. D., Huisman, B. J. A., Aarninkhof, S. G. J., & Reniers, A. J. H. M. (2018). Behaviour of subtidal sandbars in response to nourishments. *Geomorphology*, *313*, 1–12. https://doi.org/10.1016/j.geomorph.2018.04.005

Ranasinghe, R. (2016). Assessing climate change impacts on open sandy coasts: A review. https://doi.org/10.1016/j.earscirev.2016.07.011

Ranasinghe, R. (2020). On the need for a new generation of coastal change models for the 21st century. *Scientific Reports*, *10*(2010). https://doi.org/10.1038/s41598-020-58376-x

Ranasinghe, R., & Stive, M. J. (2009). Rising seas and retreating coastlines. https://doi.org/10.1007/s10584-009-9593-3

Reniers, A. J. H. M., Roelvink, J. A., & Thornton, E. B. (2004). Morphodynamic modeling of an embayed beach under wave group forcing. *Journal of Geophysical Research*, *109*, C01030. https://doi.org/10.1029/2002JC001586

Rijkswaterstaat, & van Houdt, J. (2012). Zandmotor. https://www.flickr.com/photos/zandmotor/albums/with/72157639855997186

Rijn, L. C. v. (1984). Sediment Transport, Part II: Suspended Load Transport. *Journal of Hydraulic Engineering*, *110*(11), 1613–1641. https://doi.org/10.1061/(ASCE)0733-9429(1984)110:11(1613)

van Rijn, L. C. (1993). *Principles of Sediment Transport in Rivers, Estuaries and Coastal Seas*. Aqua Publications.

van Rijn, L. C., Wasltra, D. J., Grasmeijer, B., Sutherland, J., Pan, S., & Sierra, J. P. (2003). The predictability of cross-shore bed evolution of sandy beaches at the time scale of storms and seasons using process-based profile models. *Coastal Engineering*, *47*(3), 295–327. https://doi.org/10.1016/S0378-3839(02)00120-5

van Rijn, L. C. (1997). Sediment transport and budget of the central coastal zone of Holland. *Coastal Engineering*, *32*(1), 61–90. https://doi.org/10.1016/S0378-3839(97)00021-5

Roelvink, J. A. (1993). Dissipation in random wave groups incident on a beach. *Coastal Engineering*, *19*(1-2), 127–150. https://doi.org/10.1016/0378-3839(93)90021-Y

Roelvink, J. A., Huisman, B., Elghandour, A., Ghonim, M., & Reyns, J. (2020). Efficient Modeling of Complex Sandy Coastal Evolution at Monthly to Century Time Scales. *Frontiers in Marine Science*, *7*, 535. https://doi.org/10.3389/fmars.2020.00535

Ruessink, B. G. (2005). Predictive uncertainty of a nearshore bed evolution model. *Continental Shelf Research*, *25*(9), 1053–1069. https://doi.org/10.1016/j.csr.2004.12.007

Ruessink, B. G. (2006). A Bayesian estimation of parameter-induced uncertainty in a nearshore alongshore current model. *Journal of Hydroinformatics*, *8*(1), 37–49. https://doi.org/10.2166/jh.2006.009

Saltelli, A. (2002). Sensitivity Analysis for Importance Assessment. *Risk Analysis*, *22*(3), 579–590. https://doi.org/10.1111/0272-4332.00040

Saltelli, A., Aleksankina, K., Becker, W., Fennell, P., Ferretti, F., Holst, N., Li, S., & Wu, Q. (2019). Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices. *Environmental Modelling & Software*, *114*, 29–39. https://doi.org/10.1016/j.envsoft.2019.01.012

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., & Tarantola, S. (2008). *Global Sensitivity Analysis. The Primer*. John Wiley & Sons, Ltd. http://doi.wiley.com/10.1002/9780470725184

de Schipper, M. A., Ludka, B. C., Raubenheimer, B., Luijendijk, A. P., & Schlacher, T. A. (2020). Beach nourishment has complex implications for the future of sandy shores. *Nature Reviews Earth & Environment*, 1–15. https://doi.org/10.1038/s43017-020-00109-9

de Schipper, M. A., de Vries, S., Ruessink, G., de Zeeuw, R. C., Rutten, J., van Gelder-Maas, C., & Stive, M. J. (2016). Initial spreading of a mega feeder nourishment: Observations of the Sand Engine pilot project. *Coastal Engineering*, *111*, 23–38. https://doi.org/10.1016/j.coastaleng.2015.10.011

Simmons, J. A., Harley, M. D., Marshall, L. A., Turner, I. L., Splinter, K. D., & Cox, R. J. (2017). Calibrating and assessing uncertainty in coastal numerical models. *Coastal Engineering*, *125*, 28–41. https://doi.org/10.1016/j.coastaleng.2017.04.005

Small, C., & Nicholls, R. J. (2003). A Global Analysis of Human Settlement in Coastal Zones. *Journal of Coastal Research*, *19*(3), 584–599. https://www.jstor.org/stable/4299200

Sobol', I. M. (1967). On the distribution of points in a cube and the approximate evaluation of integrals. *Zh. Vychisl. Mat. Mat. Fiz.*, *7*(4), 784–804.

Stedinger, J. R., Vogel, R. M., Lee, S. U., & Batchelder, R. (2008). Appraisal of the generalized likelihood uncertainty estimation (GLUE) method. *Water Resources Research*, *44*(12). https://doi.org/10.1029/2008WR006822

Stive, M. J. F., de Schipper, M. A., Luijendijk, A. P., Aarninkhof, S. G. J., van Gelder-Maas, C., van Thiel de Vries, J. S. M., de Vries, S., Henriquez, M., Marx, S., & Ranasinghe, R. (2013). A New Alternative to Saving Our Beaches from Sea-Level Rise: The Sand Engine. *Journal of Coastal Research*, *29*(5), 1001–1008. https://doi.org/10.2112/JCOASTRES-D-13-00070.1

Sun, N., Hong, B., & Hall, M. (2013). Assessment of the SWMM model uncertainties within the generalized likelihood uncertainty estimation (GLUE) framework for a high-resolution urban sewershed. *Hydrological Processes*, *28*(6). https://doi.org/10.1002/hyp.9869

Sutherland, J., Peet, A. H., & Soulsby, R. L. (2004). Evaluating the performance of morphological models. *Coastal Engineering*, *51*(8-9), 917–939. https://doi.org/10.1016/j.coastaleng.2004.07.015

Teng, J., Jakeman, A., Vaze, J., Croke, B., Dutta, D., & Kim, S. (2017). Flood inundation modelling: A review of methods, recent advances and uncertainty analysis. *Environmental Modelling & Software*, *90*, 201–216. https://doi.org/10.1016/j.envsoft.2017.01.006

The Economist. (2019). The rising oceans: Climate change is a remorseless threat to the world's coasts. *The Economist*. https://www.economist.com/briefing/2019/08/17/climate-change-is-a-remorseless-threat-to-the-worlds-coasts

Tonnon, P. K., Huisman, B. J., Stam, G. N., & van Rijn, L. C. (2018). Numerical modelling of erosion rates, life span and maintenance volumes of mega nourishments. *Coastal Engineering*, *131*, 51–69. https://doi.org/10.1016/j.coastaleng.2017.10.001

Tuffin, B. (1996). On the use of low discrepancy sequences in Monte Carlo methods. *Monte Carlo Methods and Applications*, *2*(4), 295–320. https://doi.org/10.1515/mcma.1996.2.4.295

Uhlenbrook, S., & Sieber, A. (2005). On the value of experimental data to reduce the prediction uncertainty of a process-oriented catchment model. *Environmental Modelling and Software*, *20*(1), 19–32. https://doi.org/10.1016/j.envsoft.2003.12.006

Valverde, H. R., Trembanis, A. C., & Pilkey, O. H. (1999). Summary of beach nourishment episodes on the U.S. East Coast barrier islands. *Journal of Coastal Research*, *15*(4), 1100–1118. https://www.jstor.org/stable/4299028%0A

de Vriend, H. J., van Koningsveld, M., Aarninkhof, S. G., de Vries, M. B., & Baptist, M. J. (2015). Sustainable hydraulic engineering through building with nature. *Journal of Hydro-Environment Research*, *9*(2), 159–171. https://doi.org/10.1016/j.jher.2014.06.004

Vrijling, J. K., & Meijer, G. J. (1992). Probabilistic coastline position computations. *Coastal Engineering*, *17*(1-2), 1–23. https://doi.org/10.1016/0378-3839(92)90011-I

Vrugt, J. A., Gupta, H. V., Bouten, W., & Sorooshian, S. (2003). A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters. *Water Resources Research*, *39*(8). https://doi.org/10.1029/2002WR001642

van Vuren, B. G. (2005). *Stochastic modelling of river morphodynamics* (Doctoral dissertation). Delft University of Technology. http://resolver.tudelft.nl/uuid:d2723fe3-f822-4e87-a7cd-d81fcd50a934

van der Wegen, M., & Jaffe, B. (2013). Towards a probabilistic assessment of process-based, morphodynamic models. *Coastal Engineering*, *75*, 52–63. https://doi.org/10.1016/j.coastaleng.2013.01.009

Wijnberg, K. M. (2002). Environmental controls on decadal morphologic behaviour of the Holland coast. *Marine Geology*, *189*(3-4), 227–247. https://doi.org/10.1016/S0025-3227(02)00480-2

A

Delft3D model & parameters

This Appendix gives a more detailed explanation of Delft3D-FLOW and the initial 16 Delft3D parameters that were examined in this thesis. First a brief overview of the governing equations used in the model is given, after which the parameters are discussed. Hereby, the focus lies on the implementation of the parameters in the model structure. Unless specified otherwise, all information is taken from Lesser et al. (2004) and Deltares (2011a).

A.1. Governing equations

The governing equations of Delft3D-FLOW are the unsteady shallow-water (Navier-Stokes) equations. The Delft3D model used in this thesis operates in two-dimensional horizontal mode (2DH). This means that the equations are only solved in the horizontal dimensions and are averaged over the depth. The system of equations solved by Delft3D consists of: the horizontal momentum equations (Equations A.2 and A.3), the continuity equation (Equation A.5), the transport equation (Equation A.6) and a turbulence closure model. Vertical accelerations are assumed negligible compared to gravitational acceleration, so the vertical momentum equation reduces to the hydrostatic pressure relation. The equations are solved in a *Generalised Lagrangian Mean* (GLM) reference frame. The GLM flow velocities are related to their Eulerian counterparts by Equation A.1.

$$U = u + u_s \quad \text{and} \quad V = v + v_s, \tag{A.1}$$

where U ad V are the horizontal GLM velocity components, u and v the Eulerian velocity components and u_s and v_s the Stokes' drift components. The horizontal momentum equations are then given by:

$$\frac{\partial U}{\partial t} + U\frac{\partial U}{\partial x} + v\frac{\partial U}{\partial y} + \frac{\omega}{h}\frac{\partial U}{\partial \sigma} - fV = -\frac{1}{\rho_0}P_x + F_x + M_x + \frac{1}{h^2}\frac{\partial}{\partial \sigma}\left(\nu_v \frac{\partial u}{\partial \sigma}\right) \tag{A.2}$$

$$\frac{\partial V}{\partial t} + U\frac{\partial V}{\partial x} + V\frac{\partial V}{\partial y} + \frac{\omega}{h}\frac{\partial V}{\partial \sigma} - fU = -\frac{1}{\rho_0}P_y + F_y + M_y + \frac{1}{h^2}\frac{\partial}{\partial \sigma}\left(\nu_v \frac{\partial v}{\partial \sigma}\right) \tag{A.3}$$

The horizontal pressure terms (P_x and P_y) are computed by the Boussinesq approximation (see Lesser et al., 2004). The Reynold's stresses (F_x and F_y) are solved by simplified equations based on the eddy viscosity concept:

$$F_x = \nu_{\mathrm{h}}\left(\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2}\right) \quad F_y = \nu_{\mathrm{h}}\left(\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2}\right) \tag{A.4}$$

The depth-averaged continuity equation is given by:

$$\frac{\partial \zeta}{\partial t} + \frac{\partial [h\bar{U}]}{\partial x} + \frac{\partial [h\bar{V}]}{\partial y} = S, \tag{A.5}$$

Finally, the transport equation is given by the advection-diffusion equation:

$$\frac{\partial[hc]}{\partial t}+\frac{\partial[hUc]}{\partial x}+\frac{\partial[hVc]}{\partial y}+\frac{\partial[\omega c]}{\partial \sigma}=h\left[\frac{\partial}{\partial x}\left(D_{\mathrm{h}}\frac{\partial c}{\partial x}\right)+\frac{\partial}{\partial y}\left(D_{\mathrm{h}}\frac{\partial c}{\partial y}\right)\right]\quad+\frac{1}{h}\frac{\partial}{\partial \sigma}\left[D_v\frac{\partial c}{\partial \sigma}\right]+hS \tag{A.6}$$

In the latter two equations, S represents a source / sink term.

A.2. Hydraulic parameters

- *Bed roughness:* Chézy coefficient, C
  In 2DH mode, the shear-stress at the bed induced by the flow is modelled by a quadratic friction law:

  $$\vec{\tau}_b=\frac{\rho_0 g\vec{U}|\vec{U}|}{C^2} \tag{A.7}$$

  Hence, C influences the shear stress. A larger value for C will lead to a smaller shear stress and vice versa (it is in fact a smoothness coefficient, rather than a roughness coefficient). The bed roughness is used in several formulations in Delft3D, including in the computation of the reference height and bed shear stress for the sediment transport computations.

- *Turbulence:* Horizontal eddy viscosity ,ν_h, and diffusivity, D_h
  ν_h and D_h are required to solve the equations for the Reynolds stresses (Equation A.4) and the transport equation (Equation A.6), respectively. Usually, the model grid is too coarse and the time step too large to resolve the turbulent motion scales. Therefore, for 2DH simulations, ν_h and D_h are implemented in Delft3D as a superposition of two parts: the molecular viscosity and a '2D-turbulence' part. The molecular viscosity of water is constant, while the 2D-turbulence parts need to be specified by the user, in the form of ν_h^{back} and D_h^{back}. These two coefficients are considered for variation in this thesis. They directly influence the values of ν_h and D_h, which are used to solve the governing equations.

A.3. Wave parameters (Roller model)

The roller model is an extension to the FLOW module, which enables the modelling of short-wave groups on long waves (Reniers et al., 2004). This causes long waves to travel along with short-wave groups. The model does not resolve individual long waves, but rather derives a direction field from the dominant frequency and direction of the given wave field. It then propagates wave and roller energy along this directional field. The peak frequency (f_p) is used to determine the group celerity, c_g, at which the energy is transported. The transported energy (E_w) is represented by the short wave-energy balance:

$$\frac{\partial E_w}{\partial t}+\frac{\partial E_n c_g\cos(\theta)}{\partial x}+\frac{\partial E_w c_g\sin(\theta)}{\partial y}=-D_w, \tag{A.8}$$

where D_w is the wave energy dissipation and θ the mean wave incidence angle. The formulation for the wave energy dissipation due to breaking is based on an analogy to a bore (Reniers et al., 2004; Roelvink, 1993):

$$D_w=2\alpha_{rol}f_p E_w\left(1-\exp\left(-\left(\frac{E_w}{\gamma^2 E_{ref}}\right)^{\frac{n_d}{2}}\right)\right), \tag{A.9}$$

where γ is the breaker index, α_{rol} a user-specified coefficient ($O(1)$) and n_d a dissipation parameter. The dissipated energy by wave breaking is then converted to roller energy through the balance for the kinetic roller energy, E_r (Reniers et al., 2004):

$$\frac{\partial E_r}{\partial t}+\frac{\partial 2E_r c\cos(\theta)}{\partial x}+\frac{\partial 2E_r c\sin(\theta)}{\partial y}=-D_r+D_w, \tag{A.10}$$

where c is the wave celerity and D_r represents the roller energy dissipation, given by:

$$D_r=\frac{2g\beta_{rol}E_r}{c} \tag{A.11}$$

β_{rol} is the user-specified mean roller slope ($O(0.1)$). In this study, four parameters of the roller model are investigated, γ, γ_{max}, α_{rol}, and β_{rol}. γ, α_{rol}, and β_{rol} are all directly included in the roller energy balance (Equation A.10). The breaker index γ denotes the ratio between wave height (H) and water depth (h), at which waves start to break. A larger γ implies later wave breaking (i.e. at a lower water depth). As the hydrodynamic time step is smaller than the wave time step, the water depth may change in-between subsequent wave time steps. Therefore, a limiter is used (γ_{max}), which represents the maximum value that γ can reach in-between two wave time steps and enforces wave breaking on the hydrodynamic time step level. α_{rol} and β_{rol} scale the wave energy dissipation and roller dissipation respectively. One can expect reciprocal effects between α_{rol} and β_{rol}, as a change in either D_w or D_r can be compensated by an equal change in the other.

A.4. Sediment parameters

The sediment transport for suspended sediment and bedload sediment is computed following van Rijn (1993). To distinguish the two, a reference height a is defined based on the bed roughness. For suspended sediment transport, the sediment concentration at the reference height (c_a) is computed by a formula adapted from Rijn (1984):

$$c_{\mathrm{a}} = f_{sus}\eta 0.015\rho_{\mathrm{s}}\frac{d_{50}}{a}\frac{T_{\mathrm{a}}^{1.5}}{D_*^{0.3}}, \tag{A.12}$$

where f_{sus} is the user specified scaling coefficient, η is the relative availability of the sediment fraction at the bed, T_a is the dimensionless bed shear stress and D_* is the dimensionless particle diameter (as per van Rijn, 1993).

The bedload transport represents the transport of sediment below the reference height a and is computed as a summation of three components, following van Rijn (1993): (1) a part due to the near-bed current, $S_{b,c}$ (in the same direction as the current), (2) a part due to the waves, $S_{b,w}$ (in the direction of wave propagation), and (3) an approximation of the suspended transport due to asymmetric wave orbital motions, $S_{s,w}$. The latter actually affects the suspended transport, but is included as a component in the bedload transport as it is limited to approximately 0.5m of the bed. See Lesser et al. (2004) for a detailed elaboration of each component. The total bedload transport is then given by the two directional components $S_{b,u}$ and $S_{b,v}$:

$$\begin{aligned} S_{\mathrm{b,u}} &= f_{\mathrm{bed}}\left(\frac{u_{\mathrm{b}}}{|\vec{u}_{\mathrm{b}}|}\left|S_{\mathrm{b,c}}\right| + \left(f_{\mathrm{bed,w}}S_{\mathrm{b,w}} + f_{\mathrm{sus,w}}S_{\mathrm{s,w}}\right)\cos\phi\right) \\ S_{\mathrm{b,v}} &= f_{\mathrm{bed}}\left(\frac{v_{\mathrm{b}}}{|\vec{u}_{\mathrm{b}}|}\left|S_{\mathrm{b,c}}\right| + \left(f_{\mathrm{bed,w}}S_{\mathrm{b,w}} + f_{\mathrm{sus,w}}S_{\mathrm{s,w}}\right)\sin\phi\right) \end{aligned} \tag{A.13}$$

where f_{bed}, $f_{bed,w}$ and $f_{sus,w}$ are user specified scaling coefficients and ϕ is the local angle between the computational grid and the wave direction.

Six sediment parameters are examined in this thesis, namely the median grain diameter d_{50}, the factor for suspended sediment Fac_{DSS}, f_{sus}, f_{bed}, $f_{bed,w}$ and $f_{sus,w}$. It follows from Equations A.12 and A.13 that the four scaling coefficients ($f_{...}$) all directly affect the computed sediment transport. It should be noted, however, that only f_{sus} is included in the computation of the suspended sediment transport, while $f_{sus,w}$ (corresponding to suspended transport due to wave asymmetry) is included in the bedload computation. Hence, all suspended transport, including due to the wave-induced alongshore current, is affected only by f_{sus}.

d_{50} denotes the median grain diameter of the bed material, whereas the representative diameter for suspended sediment is given by $d_s = Fac_{DSS} \cdot d_{50}$. Hence, d_{50} affects all computed sediment transport magnitudes (it is included in Equation A.12 as well as in the computation of the three components for the bedload transport, see Lesser et al. (2004)), while Fac_{DSS} determines the diameter used in the suspended transport computation.

A.5. Morphology parameters

- *Bed slope Coefficients, α_{bs} & α_{bn}*
  The magnitude and direction of the bedload transport can be affected by potential gradients in the bed level. Hereby, it plays a role in which direction the bed is sloped. A slope in the longitudinal direction

(parallel to the sediment transport) influences the magnitude of the transport, while a slope in transverse direction influences the direction of the sediment transport. The bedload transport is changed by a longitudinal slope by:

$$S_b = \left(1 + \alpha_{bs}\left(\frac{\tan(\phi)}{\cos\left(\tan^{-1}\left(\frac{\partial z}{\partial s}\right)\right)\left(\tan(\phi) - \frac{\partial z}{\partial s}\right)} - 1\right)\right) \cdot S_b \tag{A.14}$$

where $\partial z/\partial s$ is the longitudinal bed level gradient, ϕ is the internal angle of friction and α_{bs} is a user-specified calibration coefficient.

A transverse bed slope affects the direction of the bed load transport:

$$\begin{aligned} S_{\mathrm{b,uu}} &= S_{\mathrm{b,uu}} - \alpha_{bn}\left(\frac{\tau_{\mathrm{b,cr}}}{\tau_{\mathrm{b,cw}}}\right)^{0.5} \frac{\partial z}{\partial n} S_{\mathrm{b,vv}} \\ S_{\mathrm{b,vv}} &= S_{\mathrm{b,vv}} - \alpha_{bn}\left(\frac{\tau_{\mathrm{b,cr}}}{\tau_{\mathrm{b,cw}}}\right)^{0.5} \frac{\partial z}{\partial n} S_{\mathrm{b,uu}} \end{aligned} \tag{A.15}$$

where α_{bn} is a user specified calibration coefficient, $\tau_{b,cr}$ is the critical bed shear stress, $\tau_{b,cw}$ is the bed shear stress due to current and waves, and $\partial z/\partial n$ is the transverse bed level gradient.

Hence, α_{bs} scales the effect of a longitudinal bed slope on the magnitude of the bed load, while α_{bn} scales the effect of a transverse slope on the direction of the bed load.

- *Dry cell erosion factor, θ_{sd}*
  Delft3D allows the erosion of dry cells (e.g. the emerged beach) by distributing the computed erosion for a wet cell over its adjacent dry cells (evenly). This is governed by the user-defined factor θ_{sd}, which can vary between 0 (all erosion occurs at the wet cell) to 1 (all erosion is distributed over the adjacent dry cells).

B

Wave data

This Appendix provides more detail on the wave data and computations used for the thesis. The first section discusses the computation of wave energy characteristics, after which the long-term wave series used for the uncertainty comparison (Section 3.5) is described.

B.1. Computation of wave energy characteristics

Both for the temporal analysis (Section 5.3) and the uncertainty comparison (Sections 3.5 and 5.5) representative wave energy characteristics have been computed over various intervals. These characteristics are based on the mean wave energy per unit horizontal area (or mean energy density), E (J/m^2). According to linear wave theory, the wave energy density for small-amplitude, periodic waves is given by Holthuijsen (2007):

$$E = \frac{1}{2}\rho g a^2, \tag{B.1}$$

where ρ denotes the density of seawater $(1024 kg/m^3)$, g the gravitational acceleration $(9.81 m/s^2)$ and a the amplitude of the wave. Using the relation $H = 2a$ between the wave height, H, and the amplitude, equation B.1 can be rewritten as:

$$E = \frac{1}{8}\rho g H^2 \tag{B.2}$$

The same energy density can also be described in terms of the variance of the surface elevation as:

$$E = \rho g m_0, \tag{B.3}$$

where m_0 denotes the first order moment of the variance-density spectrum. The latter formulation also holds for random surface waves, described by a variance-density spectrum (Holthuijsen, 2007). Through the definition of the significant spectral wave height (H_{m0}), we can then derive the energy density for a random wave field.

$$H_{m0} = 4\sqrt{m_0} \tag{B.4}$$

$$E = \frac{1}{16}\rho g H_{m0}^2 \tag{B.5}$$

A time averaged value of H_{m0} can be used to compute the time averaged wave energy density over a certain interval. This is applied here in the computation of the wave energy characteristics for the temporal analysis and the uncertainty comparison. The corresponding wave data H_{m0} is averaged over hourly intervals (three-hourly for the uncertainty comparison). The representative value for the wave energy over a certain period is then obtained by taking the sum of all (three-)hourly wave energy densities over the considered time period.

B.2. Long-term wave time series

The wave series used to compute the energy characteristics of the various wave years between 1989 and 2014 gives the near-shore wave conditions. The near-shore conditions are derived from a transformation of offshore waves (Europlatform) to the $10m$-depth contour, using a SWAN model developed by Deltares (2011b). Figure B.1 shows the significant wave height (H_{m0}) and peak wave direction in three-hour intervals, while Figure B.2 presents the energy characteristics of the 25 wave years.

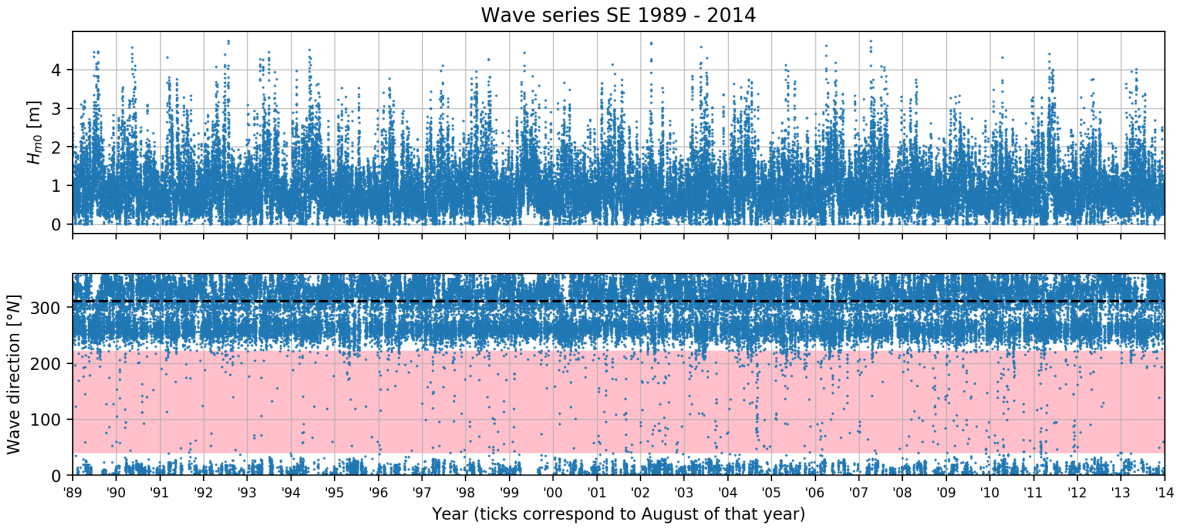

Figure B.1: Time series of significant wave height (H_{m0}, upper panel) and peak wave direction (lower panel) at location $X, Y = (71852.903, 453164.2)$, around the $10m$-depth contour close to the SE. The black dashed line represents the border between the North and West sectors, at $311°N$. The red shaded area in the lower panel denotes wave conditions moving away from the SE, which are ignored in the energy analysis.

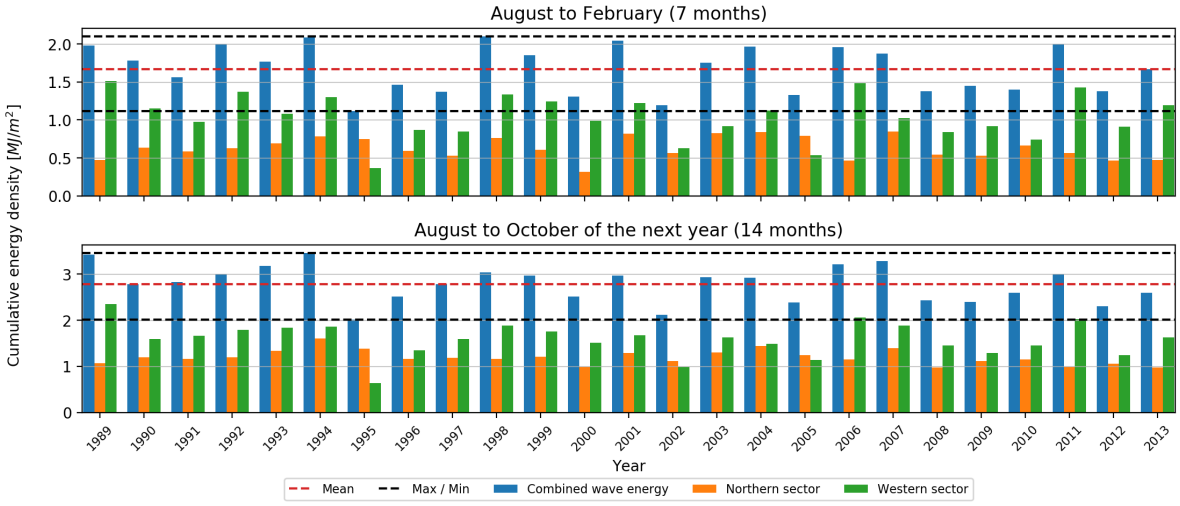

Figure B.2: Bar plots showing an estimate of the cumulative, three-hourly averaged wave energy density (see Appendix B for the computation). The values are computed over the first seven months (Aug–Feb, upper panel) and the total 14-month simulation period (Aug–Oct, lower panel) of the selected wave years. The colours represent the combined directional sectors (blue), the Northern sector (orange), and the Western sector (green). The dashed lines show the maximum, minimum and mean (red) values for the combined conditions. Note that 2011 represents the original wave series (Aug 2011 – Oct 2012).

C

Sobol' sequence

For the advantage of a faster convergence rate over MC sampling to hold, the sample size N needs to be sufficiently large (the LDS convergence rate of $O(N^{-1})$ is an optimum, achieved for $N \to \infty$). The definition of *sufficiently large* is not a generally applicable number, however, and also depends on the dimension of the parameter space, k. For given samples the superiority of the Sobol' sample over an MC counterpart can be tested through three criteria (see also Pronzato, 2017):

1. *Maximin-distance* (Johnson et al., 1990)
   Denotes the absolute distance between the two closest sample points in the parameter space. This distance should be maximised.

2. *Minimax-distance* (Johnson et al., 1990)
   Denotes the distance from the point in the parameter space, which is the most remote from any sample point, to the nearest sample point. The smaller this distance, the better spaced the samples are over the parameter space.

3. *Discrepancy criterion*
   In general, discrepancy describes a sample's deviation from the uniform distribution. Several mathematical definitions of discrepancy exist, but most are very complex to compute. The L_2-discrepancy is commonly applied for practicality, as by Kucherenko et al. (2015).

Within the scope of this thesis, it is assumed that the first two criteria are sufficient to demonstrate the superiority of the Sobol' sample, and the discrepancy test is not considered further. A Sobol' set of 1024 sample points has been compared to ten randomly drawn MC sets of the same size to evaluate if the use of the Sobol' sequence is indeed superior for this study.

The results show that the Sobol' set scores better than all ten MC sets for both criteria (Figure C.1). Therefore, it is assumed that the sample set is large enough for the convergence speed advantage of the Sobol sequence to hold and Sobol sampling is applied for this thesis. Combined with the pre-screening of the parameter space through the SA simulations, this gives a sampling strategy which aims to optimises the sampling resolution of the parameter space, while also preventing too many non-behavioural runs. Figure C.2 gives an overview of the sampled Sobol set in the five-dimensional unit hypercube. The patterns between certain variables are a result of the mathematical algorithm behind the sobol sequence. After drawing the 1024 samples in the unit hypercube, they are converted to the respective parameter ranges.

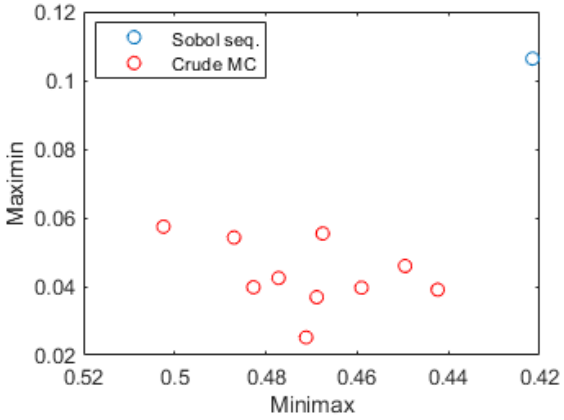

Figure C.1: Results of the computed Minimax and Maximin criteria for the Sobol sample (blue) and 10 equally sized MC samples (red). The Minimax axis is reversed, so that the upper right corner represents the best scoring sample set. The Sobol set scores better than all ten MC sets, in both metrics.

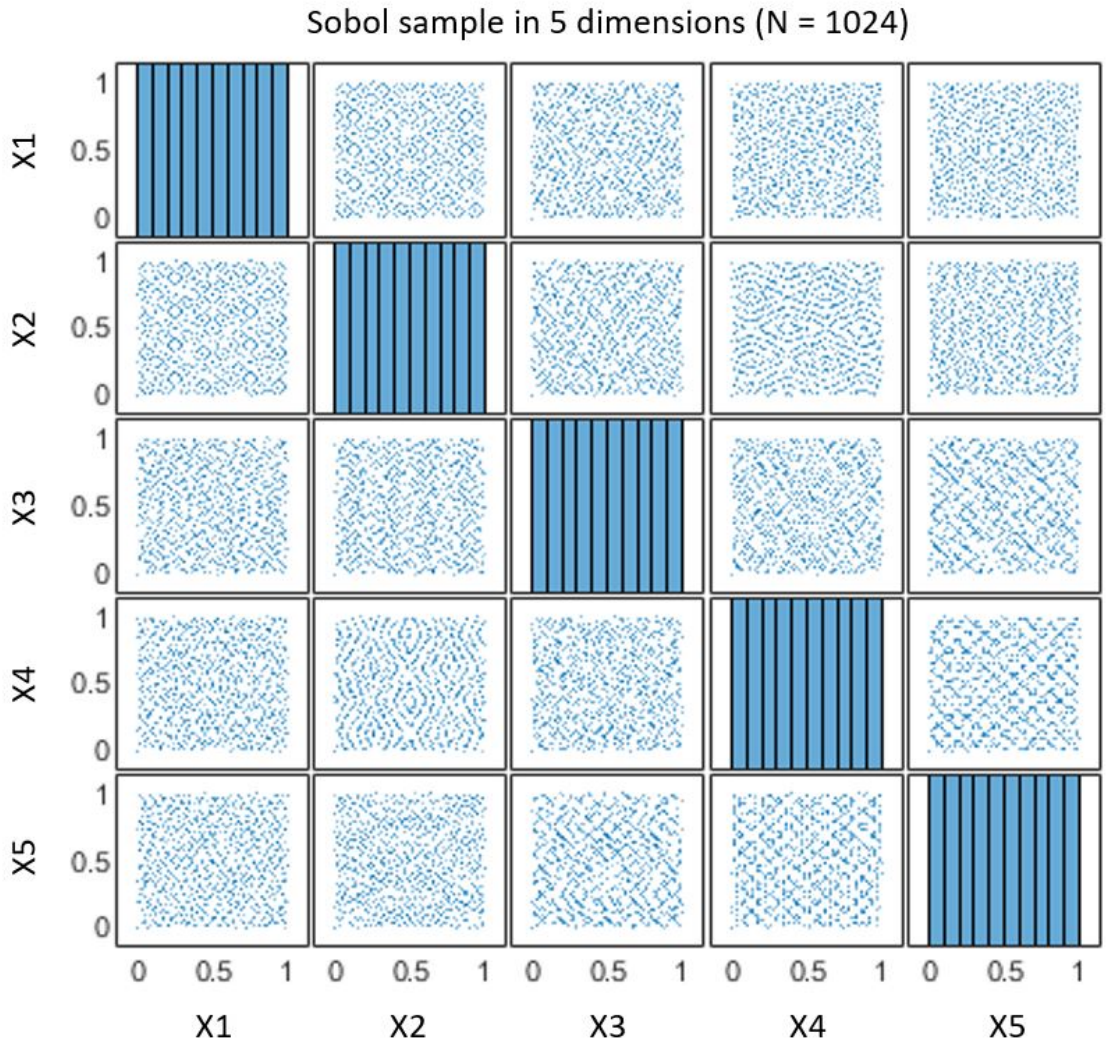

Figure C.2: Matrix of scatter plots showing 1024 samples drawn by the Sobol sequence in 5 dimensions. The scatter plots are for two parameters each. The plots on the diagonal show histograms for the corresponding variables.

D

Sensitivity analysis

D.1. EE analysis

The EE analysis in Section 4.2 only considered 19 of the 20 tracks as one track was deemed to include unstable simulations. An indicator for this was the high standard deviation of the EE's for cumulative volume changes in the entire SE area of $f_{sus,w}$ and α_{rol} (Figure D.1, upper left panel). The high standard deviations were caused by extreme EE's for one trajectory, while the other 19 trajectories showed EE's in the same order. Therefore, this trajectory was deemed unstable.

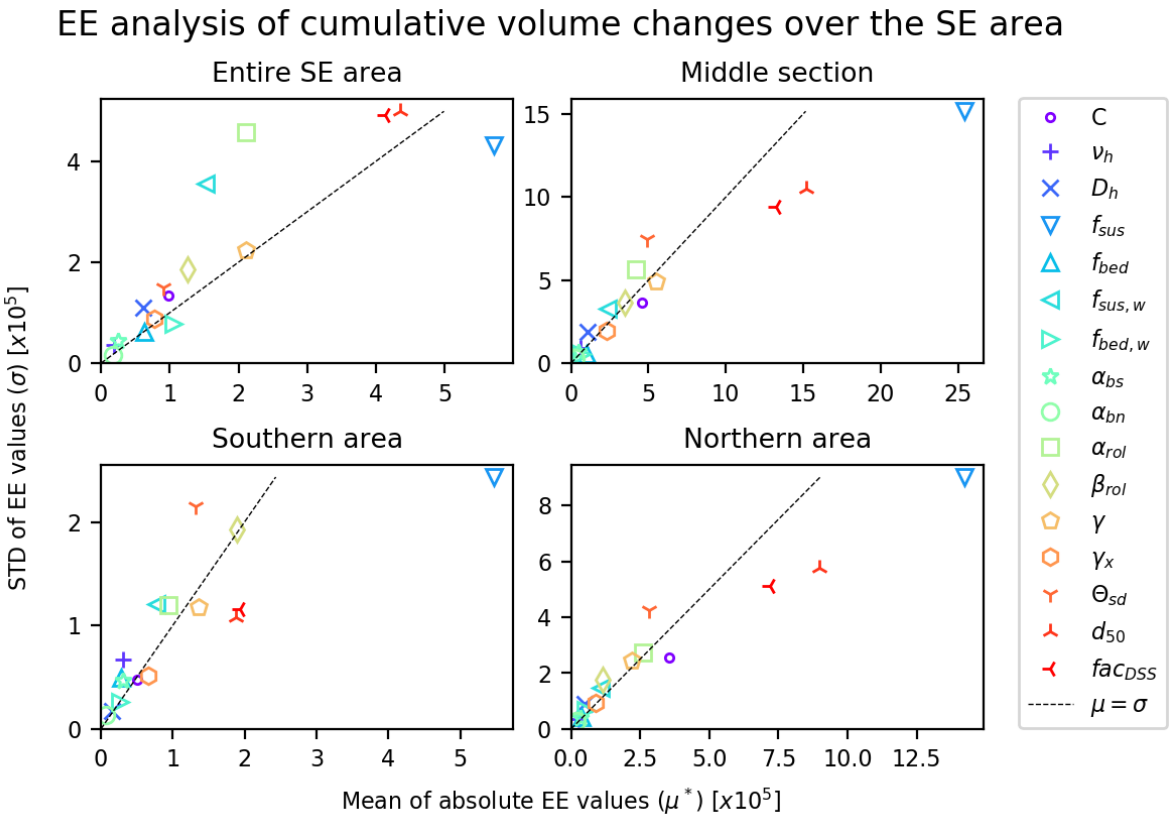

Figure D.1: Results of the EE-analysis of the cumulative volume changes around the SE (as presented in Section 4.2). Shown are the original results, including the trajectory which was later removed from the analysis, due to unrealistic volume changes. This was the cause for the hgih σ values of $f_{sus,w}$ and α_{rol} (upper left panel).

D.2. Parameter distribution

Figure D.2 shows the distribution of the parameter values over the 340 SA simulations. As can be seen the distributions are often non-uniform. Consequently, some parameter values occur more often than others in the SA simulations, possibly leading to bias in the results.

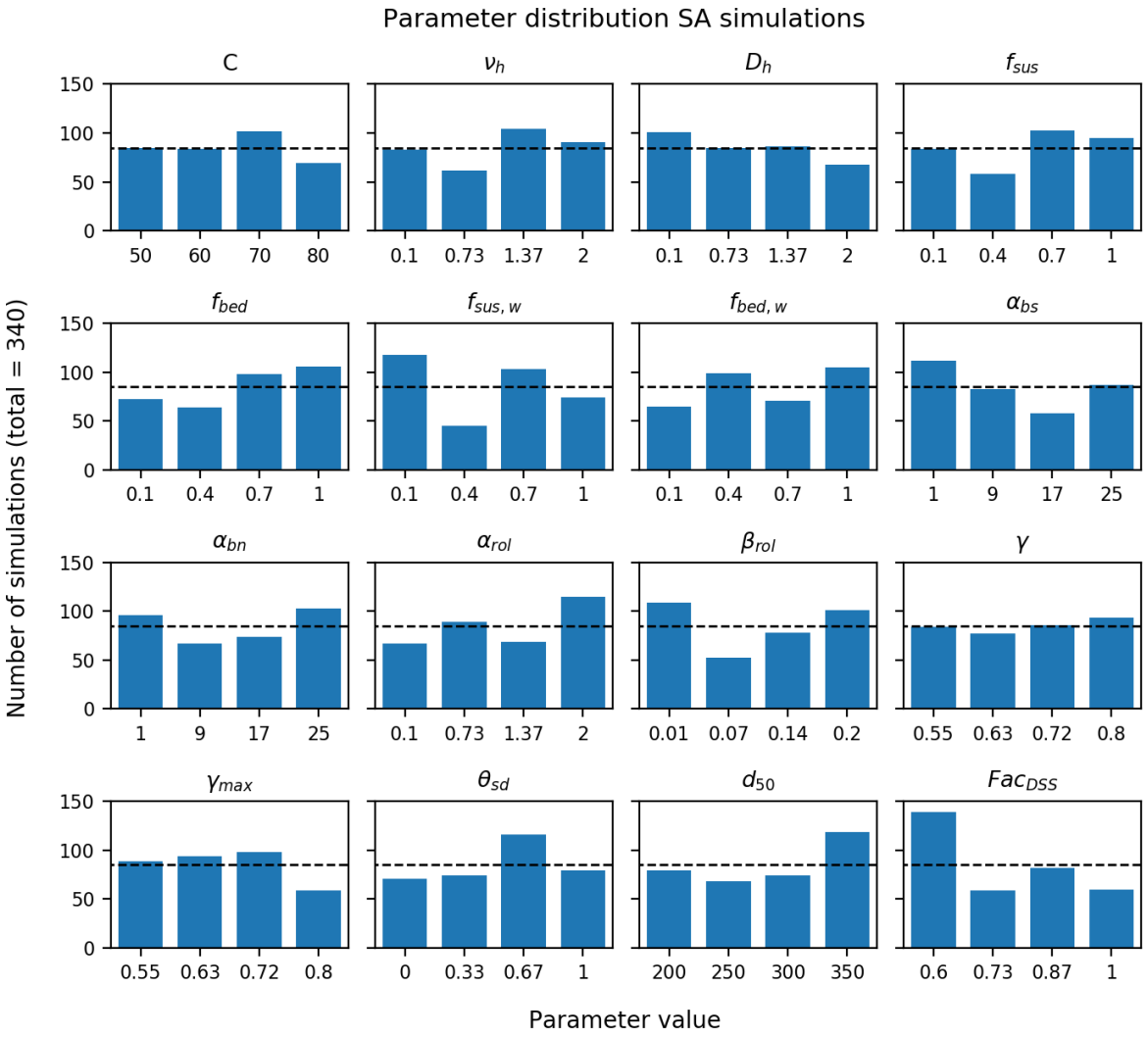

Figure D.2: Distribution of each parameter's values over the SA simulations. The dashed line represents the number of simulations for a uniform distribution (85 per parameter value).

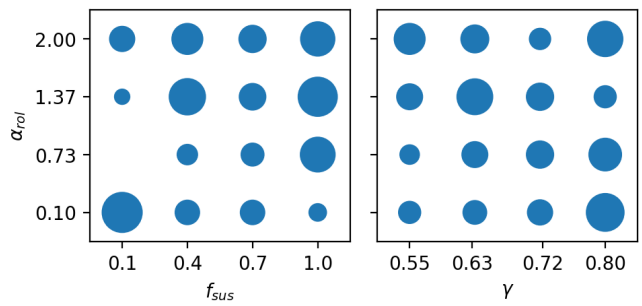

Figure D.3: Bubble plots showing the proportion of SA samples distributed over the values of α_{rol}, f_{sus}, and γ. One can distinguish weak correlations between α_{rol} and f_{sus} (positive), and α_{rol} and γ (negative).

The non-uniform distributions also lead to correlations between sampled parameter values. Table D.1 shows the correlation matrix for the 16 parameters. Correlations range from -0.3 to 0.3. There are two notable correlations between samples of parameters that wer selected for the GLUE: α_{rol} and f_{sus}, and α_{rol} and γ. The sampled values for these two pairs are shown by bubble plots in Figure D.3.

	c	v_h	D_h	f_{sus}	f_{bed}	$f_{sus,w}$	$f_{bed,w}$	α_{bs}	α_{bn}	α_{rol}	β_{rol}	γ	γ_{max}	θ_{sd}	d_{50}	Fac_{DSS}
c		-0.04	0.05	0.02	-0.01	-0.31	0.09	-0.03	-0.09	-0.21	0.10	-0.12	-0.08	-0.22	-0.09	0.00
v_h	-0.04		-0.11	0.13	-0.07	0.09	-0.07	0.20	0.09	0.01	-0.03	0.09	-0.24	0.01	-0.08	-0.08
D_h	0.05	-0.11		-0.03	-0.01	-0.21	-0.30	-0.09	0.08	0.13	0.11	-0.13	-0.01	0.03	-0.18	0.03
f_{sus}	0.02	0.13	-0.03		0.08	0.11	-0.03	0.24	0.27	0.23	-0.09	-0.02	0.20	-0.11	-0.11	-0.18
f_{bed}	-0.01	-0.07	-0.01	0.08		0.12	0.16	-0.14	-0.03	-0.10	-0.22	-0.10	-0.07	0.00	-0.15	-0.26
$f_{sus,w}$	-0.31	0.09	-0.21	0.11	0.12		0.17	0.22	-0.02	-0.04	-0.19	0.02	-0.12	0.23	-0.14	0.05
$f_{bed,w}$	0.09	-0.07	-0.30	-0.03	0.16	0.17		0.06	-0.04	0.03	-0.14	-0.23	0.03	0.05	-0.02	-0.03
α_{bs}	-0.03	0.20	-0.09	0.24	-0.14	0.22	0.06		0.15	0.13	-0.22	0.10	-0.15	0.06	-0.16	0.04
α_{bn}	-0.09	0.09	0.08	0.27	-0.03	-0.02	-0.04	0.15		0.33	0.25	-0.05	0.01	-0.13	-0.17	-0.07
α_{rol}	-0.21	0.01	0.13	0.23	-0.10	-0.04	0.03	0.13	0.33		0.10	-0.17	0.12	0.11	0.08	0.10
β_{rol}	0.10	-0.03	0.11	-0.09	-0.22	-0.19	-0.14	-0.22	0.25	0.10		0.13	0.06	-0.12	0.13	0.12
γ	-0.12	0.09	-0.13	-0.02	-0.10	0.02	-0.23	0.10	-0.05	-0.17	0.13		0.10	-0.03	0.08	-0.02
γ_{max}	-0.08	-0.24	-0.01	0.20	-0.07	-0.12	0.03	-0.15	0.01	0.12	0.06	0.10		-0.14	0.32	-0.18
θ_{sd}	-0.22	0.01	0.03	-0.11	0.00	0.23	0.05	0.06	-0.13	0.11	-0.12	-0.03	-0.14		-0.08	0.17
d_{50}	-0.09	-0.08	-0.18	-0.11	-0.15	-0.14	-0.02	-0.16	-0.17	0.08	0.13	0.08	0.32	-0.08		0.01
Fac_{DSS}	0.00	-0.08	0.03	-0.18	-0.26	0.05	-0.03	0.04	-0.07	0.10	0.12	-0.02	-0.18	0.17	0.01	

Table D.1: Correlation matrix showing Pearson's linear correlation coefficient (r) between all 16 parameters, for the 340 sampled parameter sets. Warmer colours indicate positive correlation, while cold colours indicate negative correlation.

D.3. Performance analysis

In Section 4.3 the model performance of the SA simulations has been analysed, to find out how it relates to the parameter ranges and distributions. For each parameter, the the values can be plotted against the achieve skill scores of the corresponding simulations. Here, this has been done in the form of scatter, box, and bar plots. In Section 4.3 only the scatter and bar plots for overall model performance (BSS_{tot}) are shown. This appendix gives the same figures for the volume changes ($BSS_{\Delta V}$), the shoreline position (BSS_{sl}) and the bed level changes (BSS_z).

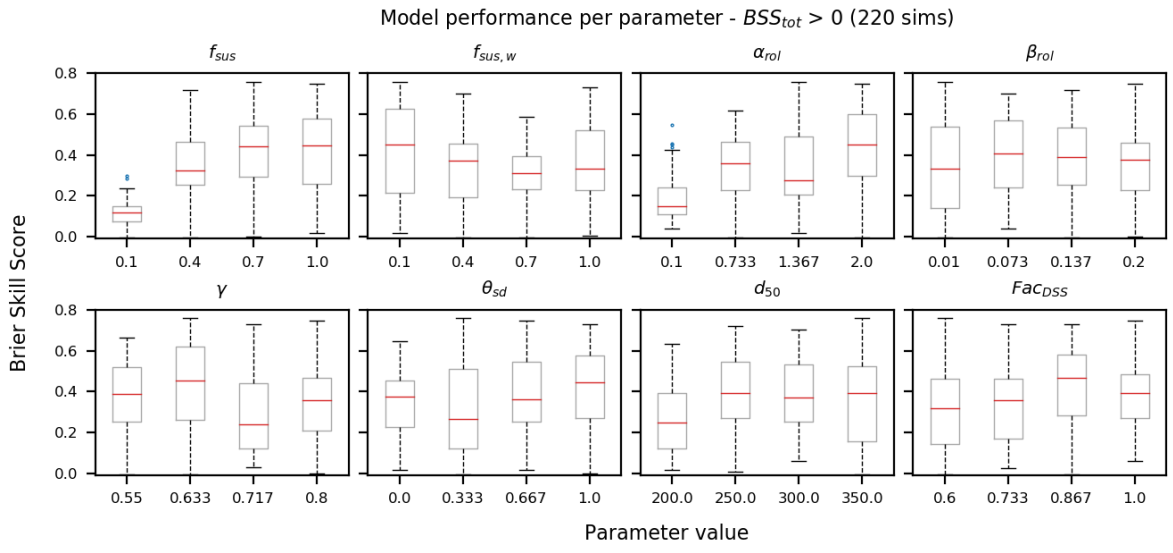

Figure D.4: Box plots showing the distribution of BSS_{tot} for the different parameter values. The box plot shows the median (red line), the interquartile range (grey box), maximum and minimum values (whiskers) and outliers, if present (blue dots).

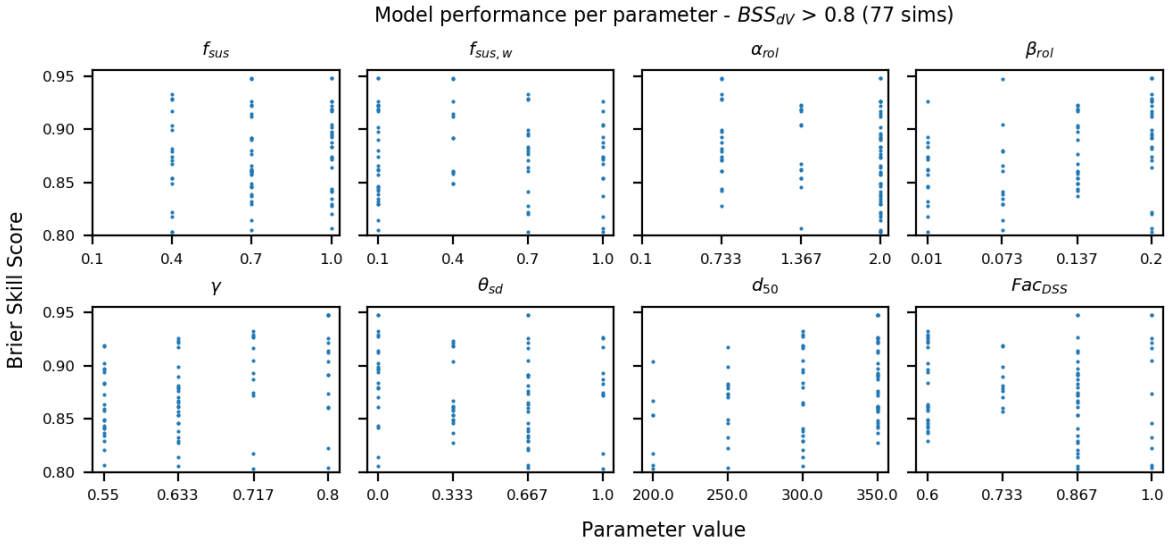

Figure D.5: Scatter plots showing the distribution of all $BSS_{\Delta V} > 0.8$ over the different parameter values.

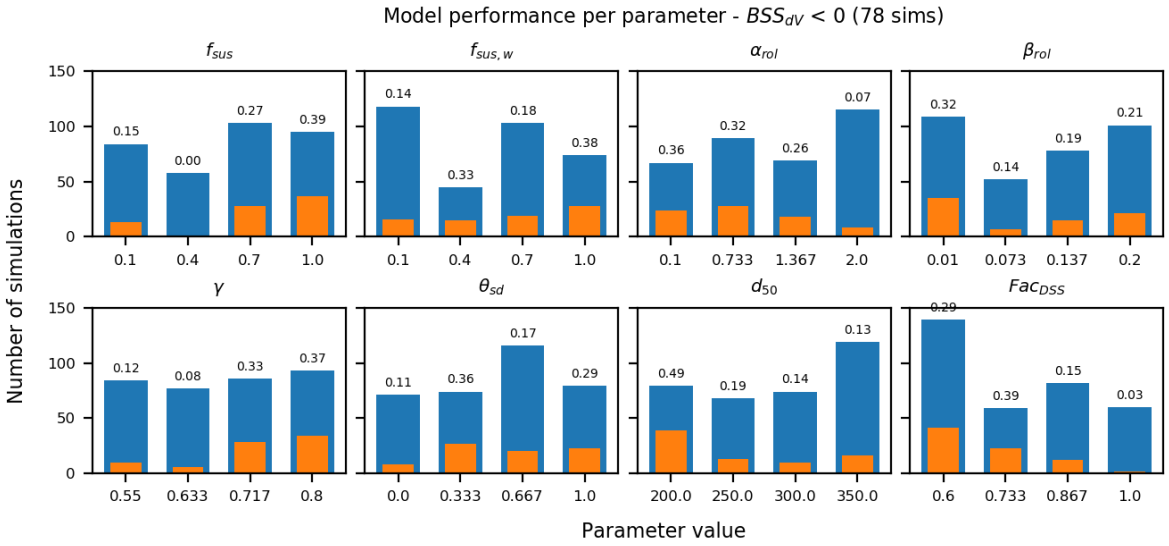

Figure D.6: Bar plots showing the distribution of all $BSS_{\Delta V} < 0$ over the different parameter values.

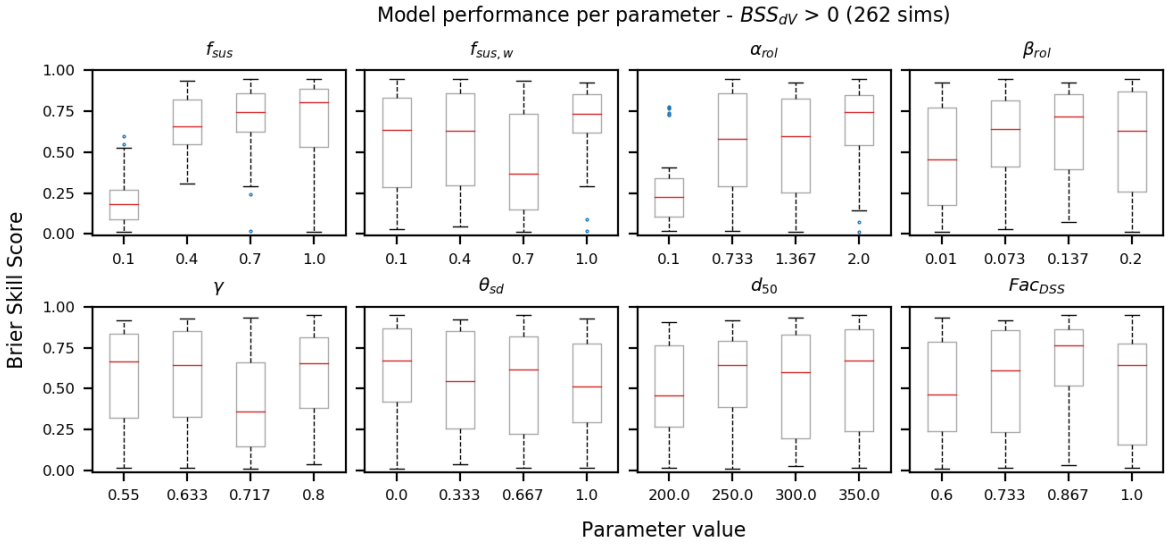

Figure D.7: Box plots showing the distribution of all $BSS_{\Delta V} > 0$ over the different parameter values.

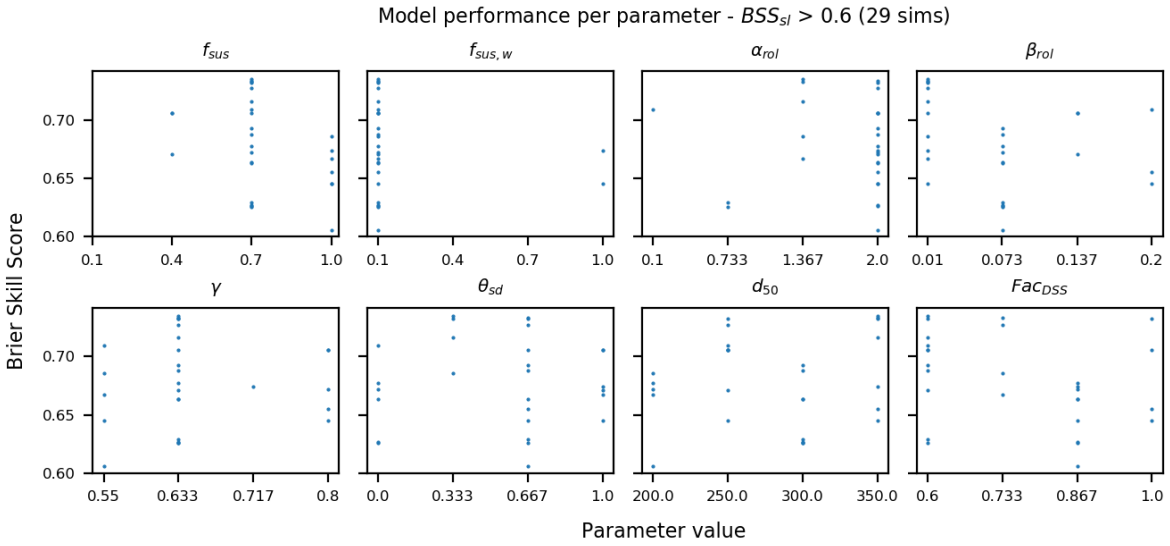

Figure D.8: Scatter plots showing the distribution of all $BSS_{sl} > 0.6$ over the different parameter values.

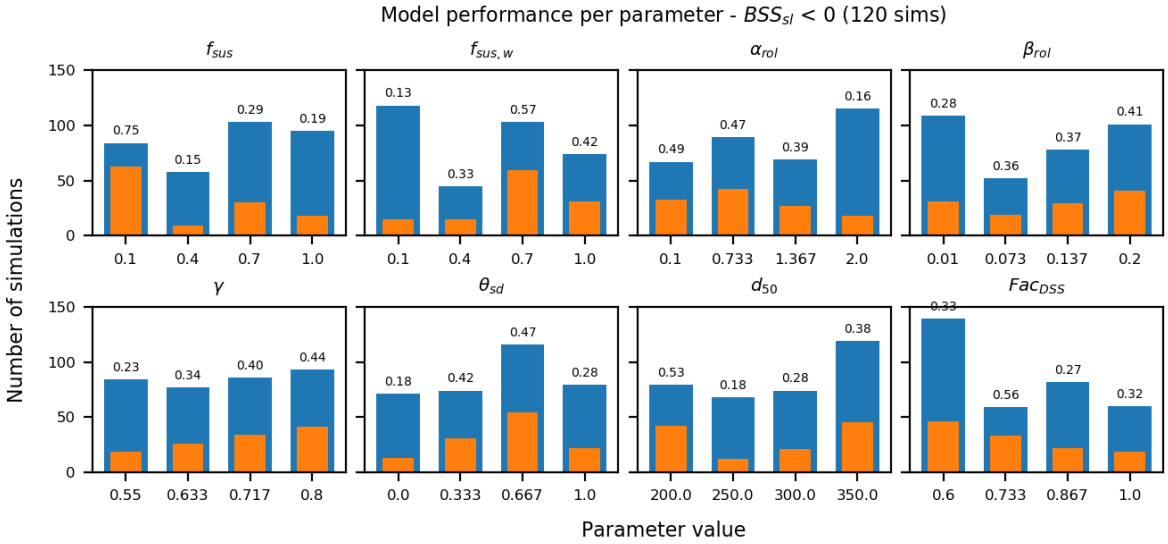

Figure D.9: Bar plots showing the distribution of all $BSS_{sl} < 0$ over the different parameter values.

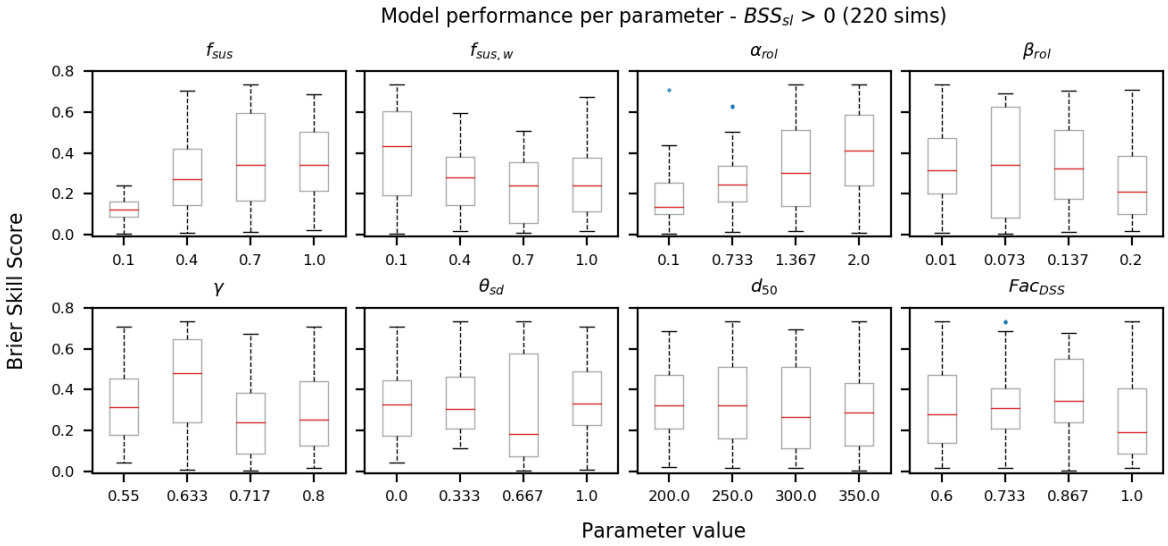

Figure D.10: Box plots showing the distribution of all $BSS_{sl} > 0$ over the different parameter values.

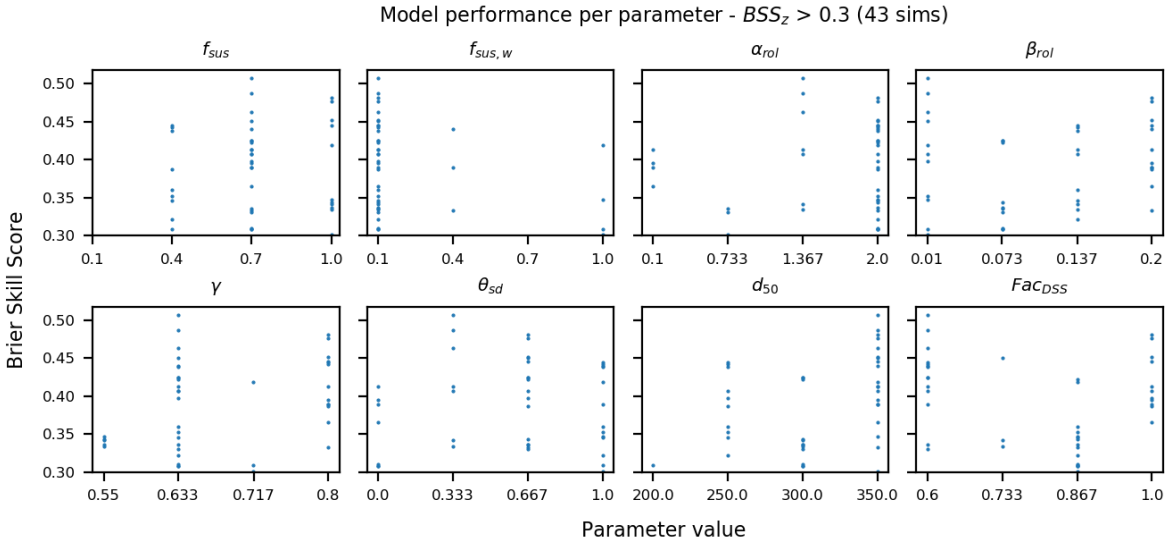

Figure D.11: Scatter plots showing the distribution of all $BSS_{sl} > 0.3$ over the different parameter values.

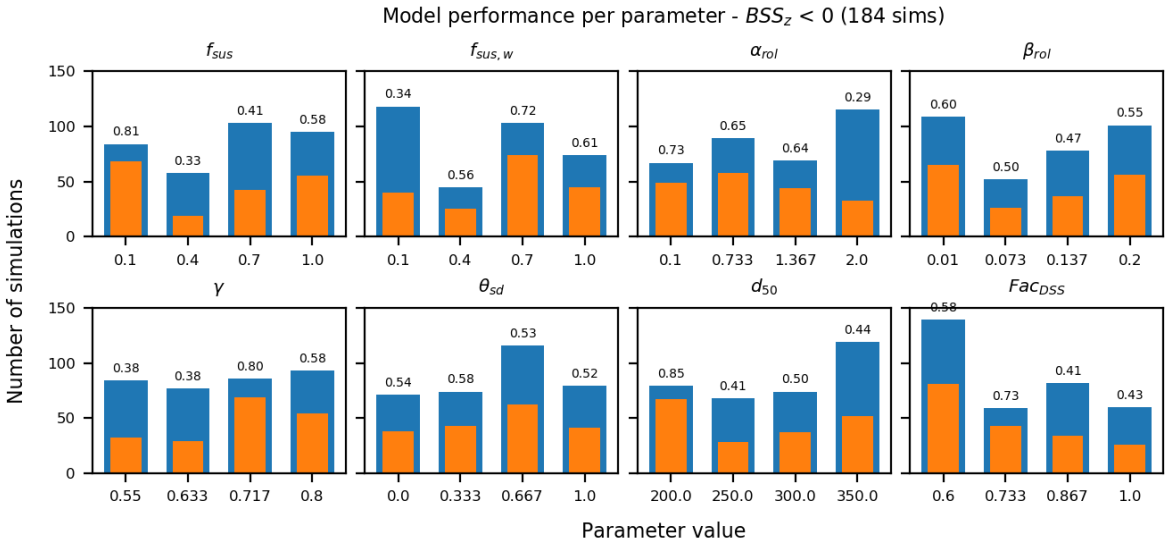

Figure D.12: Bar plots showing the distribution of all $BSS_{sl} < 0$ over the different parameter values.

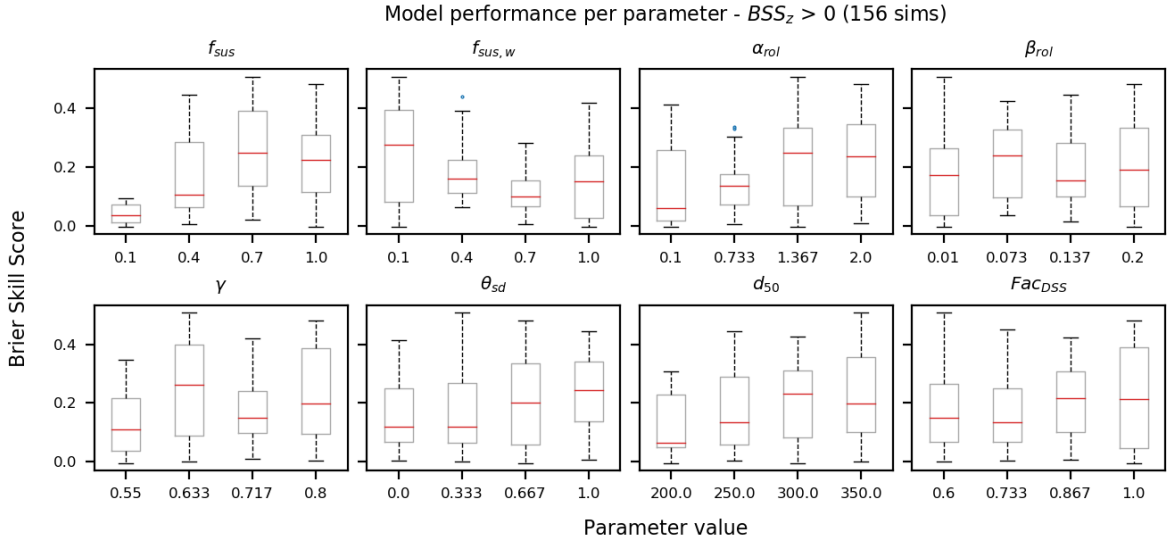

Figure D.13: Box plots showing the distribution of all $BSS_{sl} > 0$ over the different parameter values.

E

Uncertainty analysis

E.1. Uncertainty bounds

The following figure shows the time series of cumulative volume changes with the uncertainty bounds resulting from the GLUE analysis. Here, hey are plotted on equal scales, to visualise the difference in the magnitudes of the predicted volume changes. As can be seen, the morphological activity is highest in the middle section, followed by the North section.

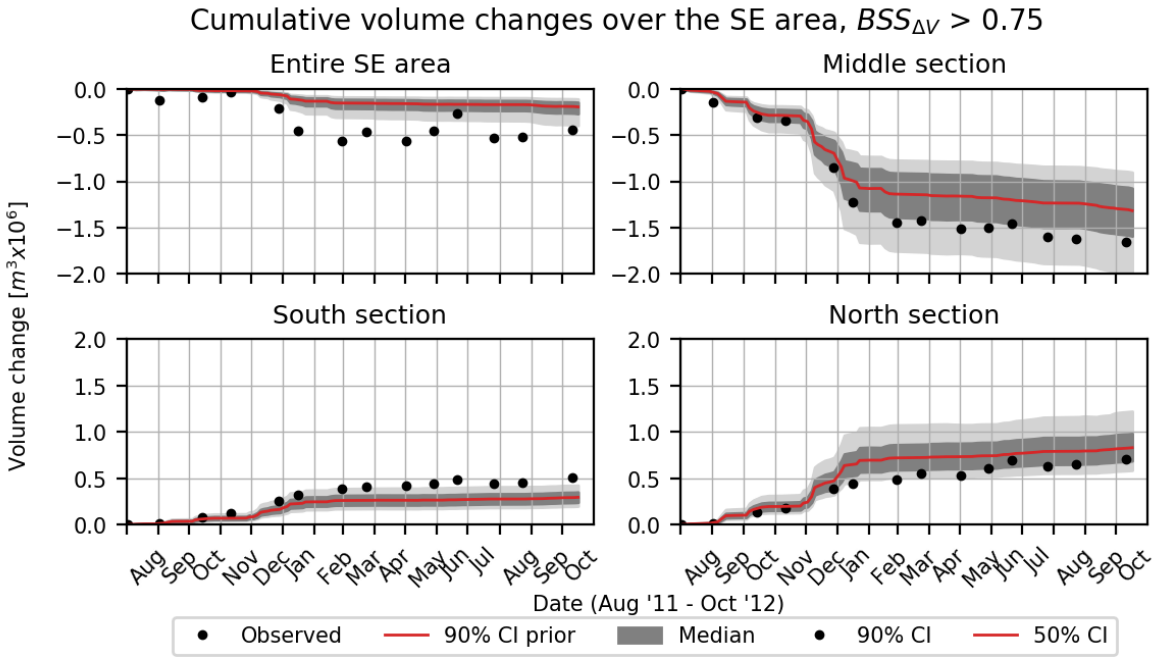

Figure E.1: Subplots showing a time series of cumulative volume changes from the measurements by de Schipper et al. (2016) (black dots) and the combined results of behavioural GLUE simulations (determined by a threshold of $BSS_{dV} \geq 0.75$). The median (red) and confidence intervals (grey araes) are given. All values are in million cubic metres and on an equal scale. The time series is plotted from August 2011 to October 2012, for the four control areas: Entire SE (upper left), middle section (upper right), Southern section (lower left), and Northern section (lower right).

E.2. Temporal analysis

Figure E.2 shows the development between the wave climate, the predicted volume changes and the growth of the uncertainty bounds in time. The correlation is given for all four control areas by means of the Pearson correlation coefficient.

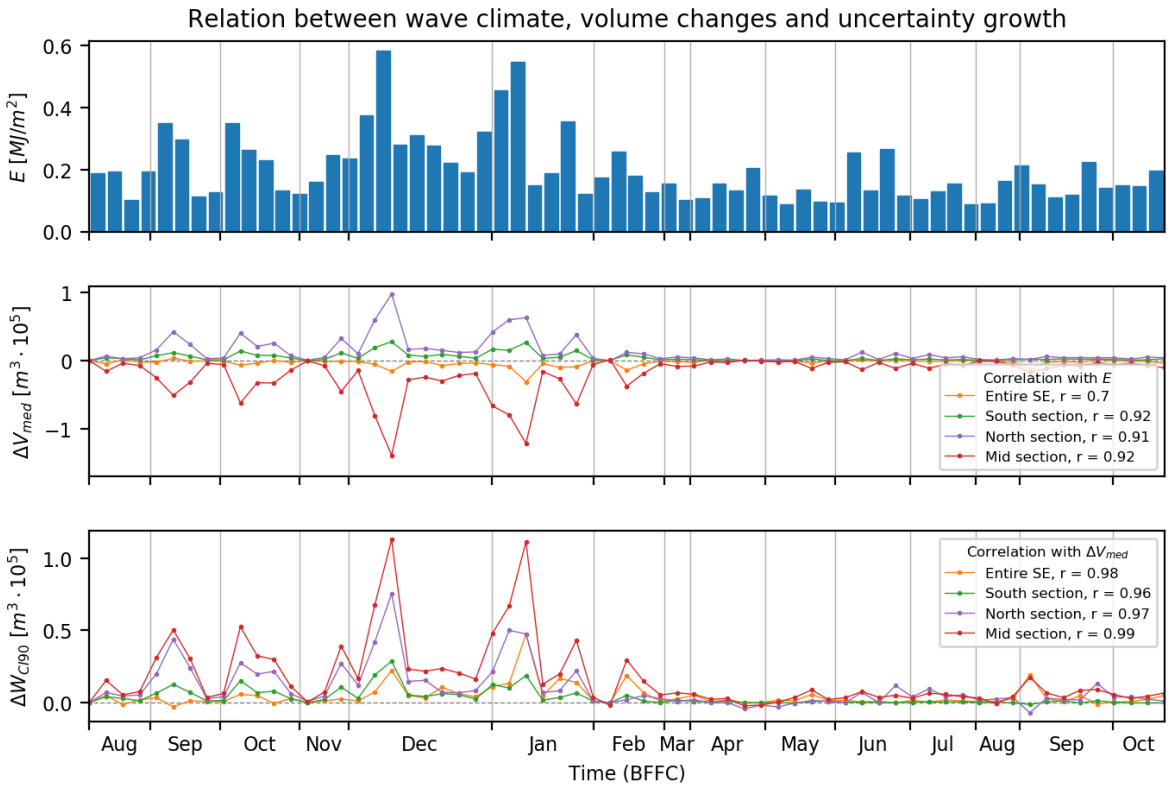

Figure E.2: From top to bottom: 1) Cumulative mean wave energy density per model output interval (E); 2) Median of predicted behavioural volume changes per model output interval (ΔV_{med}); 3) Change in width of the 90% CI for cumulative volume changes per model output interval (ΔW_{CI90}). The correlation coefficients (r) are based on the absolute value of the volume changes. Due to the BFFC approach the output intervals span varying amounts of 'real' time, resulting in the distorted time axis.

E.3. Parameter optimisation

E.3.1. Posterior distributions individual output variables

In Section 5.4, only the posterior parameter distributions based on the combined likelihood are presented. Here, the posterior parameter distributions based on the individual output variables are shown.

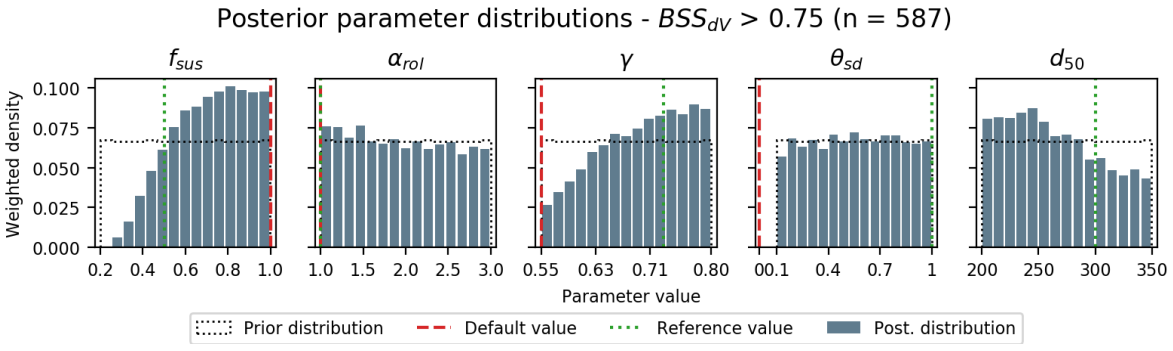

Figure E.3: Posterior distribution of the parameter values (histogram bars) for the cumulative volume changes around the SE.

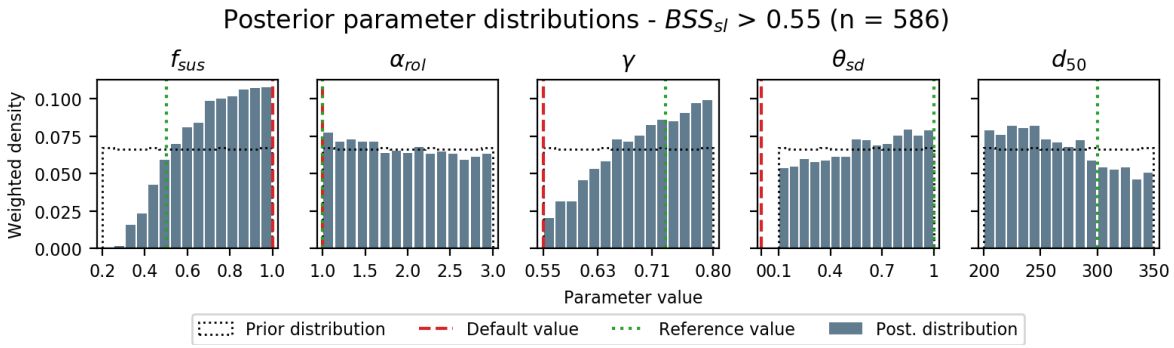

Figure E.4: Posterior distribution of the parameter values (histogram bars) for the shoreline position around the SE.

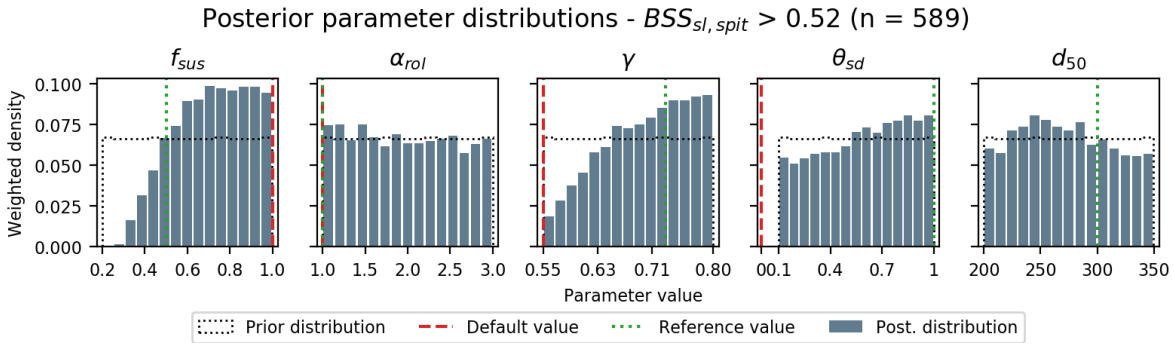

Figure E.5: Posterior distribution of the parameter values (histogram bars) for the shoreline position in the spit area.

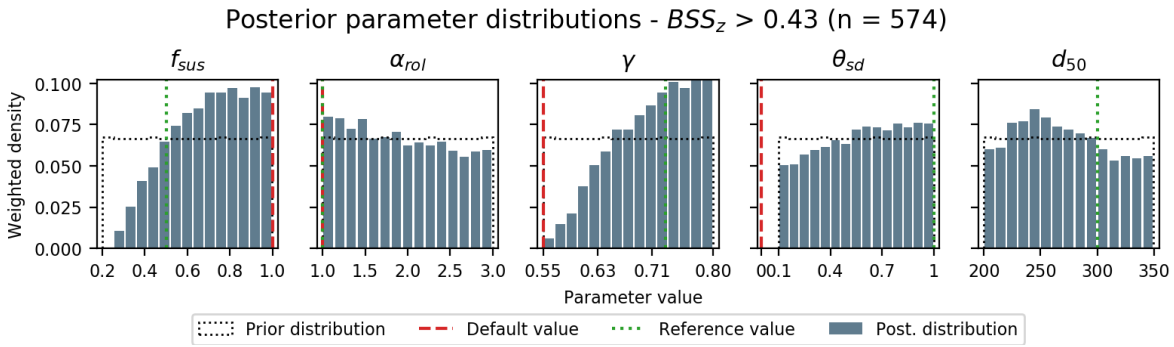

Figure E.6: Posterior distribution of the parameter values (histogram bars) for the bed level changes around the SE.

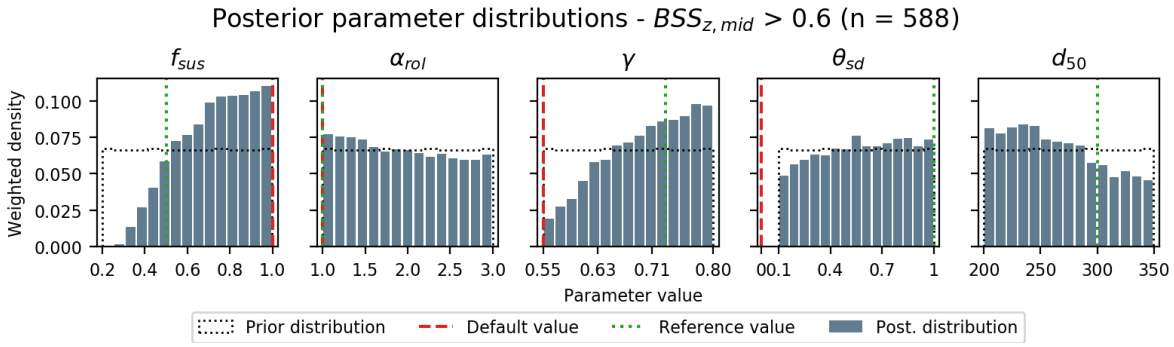

Figure E.7: Posterior distribution of the parameter values (histogram bars) for the bed level changes in the head section.

E.3.2. Conditional posterior distributions

Three of the OPS discussed in Section 5.4 are based on conditional posterior distributions, which are presented below. Each OPS is based on one fixed parameter. The conditional distribution is then derived from the behavioural parameter sets for which the fixed parameter is within 5% of the selected value (relative to the applied parameter range). For instance, $f_{sus} = 0.75$ is fixed for OPS_1. Of the 507 behavioural runs (based on combined likelihood) there remain around 70 simulations for which f_{sus} is within 5% of this value (i.e. for which $0.71 < f_{sus} < 0.79$). However, a simple weighted histogram might show a distorted distribution in case there are more simulations in certain bins than in others. Therefore, for each bin of the conditional distribution, the average likelihood per simulation in this bin is computed.

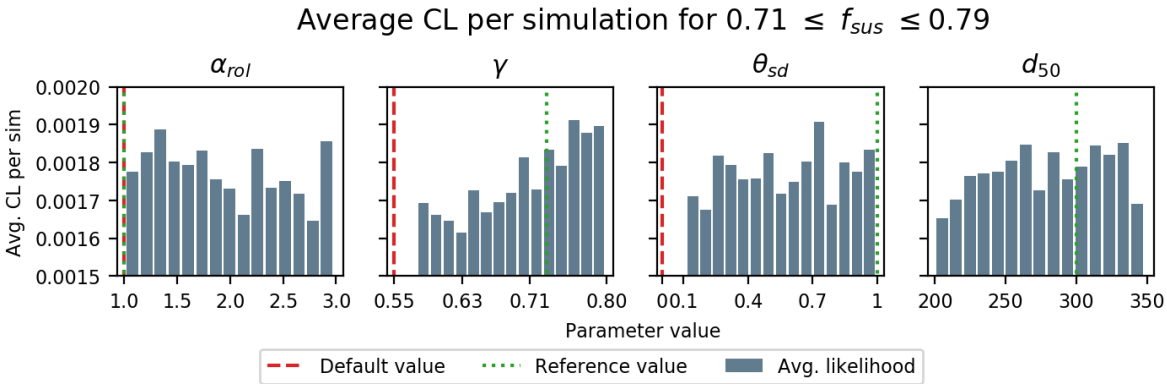

Figure E.8: Conditional posterior parameter distributions for $0.71 < f_{sus} < 0.79$.

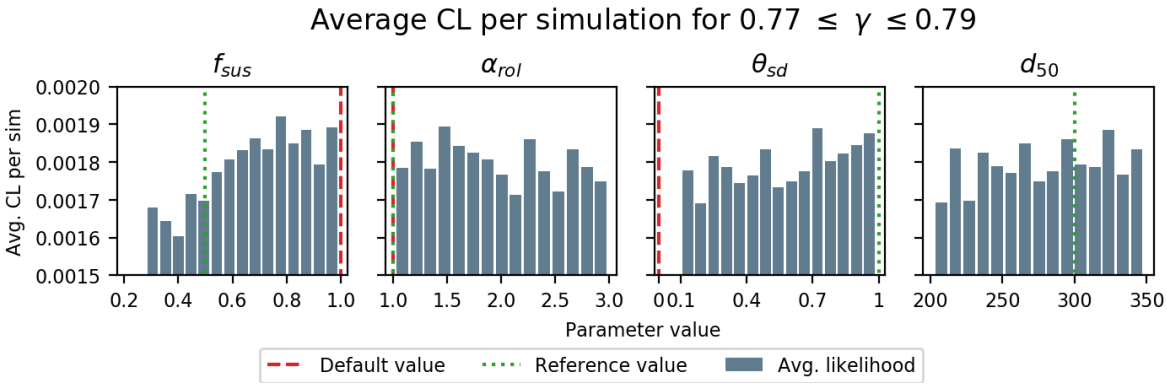

Figure E.9: Conditional posterior parameter distributions for $0.77 < \gamma < 0.79$.

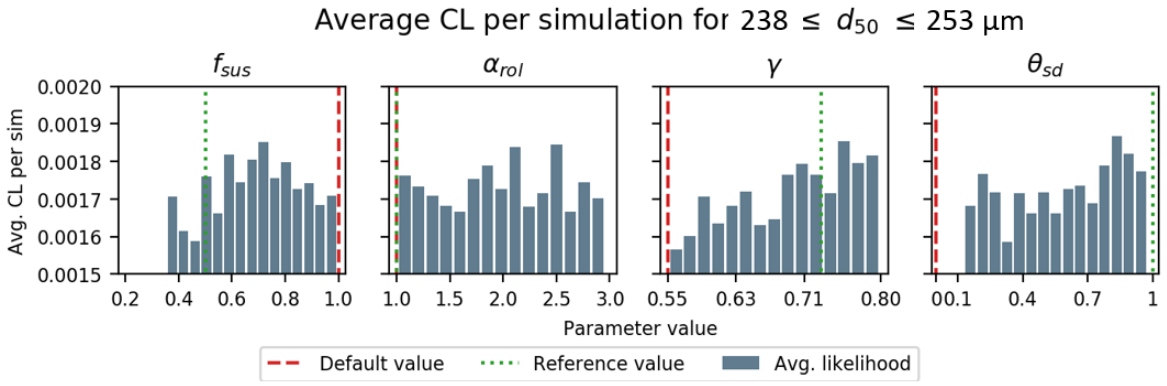

Figure E.10: Conditional posterior parameter distributions for $237.5\mu m < d_{50} < 252.5\mu m$.

E.4. OPS evaluation

Table E.1 shows the rank and percentile in which the OPS and reference simulations rank among the 1024 GLUE simulations, based on the various model outputs. Figure E.11 shows the predicted shoreline position for each of the five OPS simulations.

Sim		BSS_{tot}	$BSS_{\Delta V}$	BSS_{sl}	$BSS_{sl,spit}$	BSS_z	$BSS_{z,head}$
OPS_0	Rank	298	432	241	404	153	146
	Percentile	70.9	57.9	76.5	60.6	85.1	85.8
OPS_1	Rank	32	106	63	42	16	41
	Percentile	96.9	89.7	93.9	95.9	98.4	96.0
OPS_2	Rank	6	27	25	20	5	11
	Percentile	99.4	97.4	97.6	98	99.5	98.9
OPS_3	Rank	63	61	30	55	111	107
	Percentile	93.9	94.0	97.1	94.6	89.2	89.6
OPS_4	Rank	30	154	27	48	16	36
	Percentile	97.1	85.0	97.4	95.3	98.4	96.5
Ref	Rank	279	43	509	579	274	179
	Percentile	72.8	95.8	50.3	43.5	73.3	82.5

Table E.1: Overview of the rank and corresponding percentile of the five OPS simulations and the reference simulation (Luijendijk et al., 2019) among the 1024 GLUE simulations. As can be seen OPS_2 achieves the highest rank for all outputs.

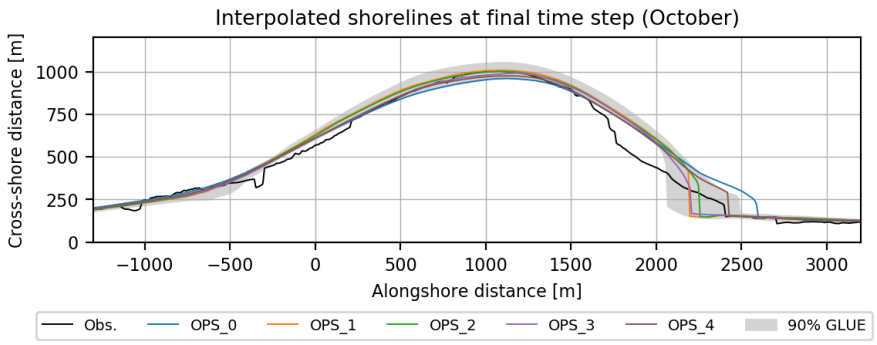

Figure E.11: Overview of the predicted shoreline position in October 2012 for the five OPS simulations.