MSc Thesis

Towards an efficient sensitivity analysis of wave forcing in coastal erosion studies

A case study of Negril Coast, Jamaica

Mieke van Arkel

May 2016

TU Delft

Deltares
Enabling Delta Life

Reference front page figure: http://negril-jamaica-videos.com/town-of-negril/features-of-negril/beachfront.html

Towards an efficient sensitivity analysis of wave forcing in coastal erosion studies

A case study of Negril Coast, Jamaica

Final report by:

ir. Mieke van Arkel

In partial fulfillment of the requirements for the degree of Master of Science in Hydraulic Engineering at the Delft University of Technology.

May, 2016
Delft

Graduation Committee
Prof. Dr. ir. A.J.H.M. Reniers, Delft University of Technology
ir. W.P. de Boer, Deltares
ir. A.P. Luijendijk, Delft University of Technology
Dr. ir. O. Morales Napoles, Delft University of Technology

Preface

This thesis project constitutes the final stage of my MSc degree program Hydraulic Engineering at Delft University of Technology. Before the start of this thesis project, I realized this was the last option to gain knowledge during my time as a student. I wanted to get the most out of it. Therefore, I formulated a personal objective that was to improve my Matlab and Delft3D modeling skills. I am very happy that I can state that I have fulfilled this objective. During the thesis process, that knew of course a lot of ups and downs, I constantly persuaded myself that my personal objective was more important than the actual research's objectives. However, in the end, I am happy with the way the research objectives are fulfilled as well.

First of all, I would like to thank Arjen Luijendijk for giving me the opportunity to write two theses on the same case study and, moreover, to conduct the last research at Deltares. Arjen, thank you for your support throughout the entire process. I wish one day I can work as hard as you do, while being so calm as you are.

Of course thanks to my other supervisors as well: Ad Reniers, thank you for being present at all the meetings and for your approachable attitude. Thank you, Oswaldo Morales Napoles for your help with the probabilistic part. I know supervising students is not your favorite task, so therefore even more thanks.

Wiebe de Boer, thank you for being my daily supervisor. You always made time for me, while you had a lot of stuff to do yourself. Thank you for reading and commenting on all my draft reports. You have really inspired me by your ideas and criticism, but especially by your working and thinking pace.

Lastly, thank you Feico for helping me in the writing process and Yahia for providing me with the bathymetric data.

Mieke van Arkel
arkelmieke@gmail.com

Table of Contents

Summary

Many coastal areas around the globe are threatened by coastal erosion. Numerical models are often used to study the coastal erosion processes and mitigating measures. However, these models inherently entail a degree of uncertainty. *Sensitivity analysis* can be used to appoint uncertainty in the model output to different sources of uncertainty in the model input. This study investigates approaches for conducting an efficient sensitivity analysis of wave forcing in coastal erosion studies.

In the context of coastal erosion studies, an efficient sensitivity analysis is typically characterized by obtaining sufficiently accurate insights in the system dynamics within the project's time and budget constraints. Conducting an extensive sensitivity analysis may be very time-consuming. In order to meet time and budget constraints, engineers can use approaches that speed-up the sensitivity analysis. The following three approaches are considered, the use of: (i) simpler models, (ii) simpler sensitivity methods, and/or (iii) input reduction methods. These approaches may lead to limited or even inaccurate insights in the system dynamics. Hence, engineers face the difficulty of trading accuracy for speed. This inverse relation between speed and accuracy is called the *speed-accuracy trade-off* (SAT).

To investigate approaches for conducting an efficient sensitivity analysis the following sub-objectives need to be fulfilled: Asses the speed-accuracy trade-off between:

- the use of a simple and a complex model to conduct a sensitivity analysis.
- the use of several sensitivity methods to conduct a sensitivity analysis.
- conducting a sensitivity analysis with a reduced and a full input dataset.

Negril Coast, located in Jamaica, is used as a case study to conduct several sensitivity analyses and to evaluate their performance. Negril Coast is a wave-dominated environment that suffers from coastal erosion. The processes driving the erosion are currently unclear. Furthermore, the case is characterized by data poorness on e.g., nearshore wave characteristics and bathymetry. A sensitivity analysis is considered useful to (i) assess the uncertainty in the prediction of the coastline evolution resulting from data uncertainty, and (ii) identify the important parameters for the coastline evolution.

The focus of this study is on bulk alongshore sediment transport rates due to nearshore waves. The following model-sensitivity configurations are considered in order to assess the SATs:

- Models: wave transformation using linear wave theory and a wave energy balance model (Delft3D-WAVE, i.e. SWAN)
- Sensitivity methods: using the non-probabilistic (i) One-at-A-Time (OAT) method, (ii) Morris method, (iii) Full factorial design method, (iv) Fractional factorial design method, and the probabilistic (v) Latin hypercube method and (vi) Crude Monte Carlo method.
- Input reduction methods: a reduced input wave climate using the K-harmonic means method combined with the maximum dissimilarity algorithm and a full input wave climate

The most extensive sensitivity analysis was conducted using (i) the wave energy balance to transform offshore wave characteristics to nearshore, (ii) the Crude Monte Carlo method with 350 samples, and (iii) the full wave input dataset. This analysis generates results as accurate as possible within the time constraints of this research. Moreover, it is considered that a more extensive analysis would only demand more time without a significant increase in accuracy. Therefore, the speed and accuracy of this analysis were established as the *benchmark*. The accuracy is defined by the (i) range of model output, (ii) estimated output distributions, (iii) sensitivities and their level of significance. The speed and accuracy of the other sensitivity analyses were compared to this benchmark.

This research demonstrates that although the achievable efficiency of the sensitivity analysis is very much dependent on the case at hand, a number of generic approaches can be applied that increase the speed while maintaining similar accuracy levels as the benchmark. The research draws the following conclusions concerning these approaches:

- The complexity of the study area determines whether linear wave theory can be applied for wave transformation. In cases the linear wave theory may be applied, it speeds up the sensitivity analysis with a factor up to 180.

- The use of probabilistic methods, compared to non-probabilistic methods, results in significantly more accurate results of the sensitivity analysis.

- A probabilistic sensitivity analysis using the Latin hypercube method is a factor ± 5 faster than the benchmark, while maintaining similar accuracy levels.

- A sensitivity analysis with a reduced input dataset is a factor 6-13 faster than the benchmark. In principal the accuracy of an analysis with a reduced dataset can be similar to the full dataset. However, the accuracy depends on the performance of the input reduction individually and in combination with a sensitivity analysis.

- The use of the Latin hypercube method combined with a reduced wave input dataset is the best approach for an efficient sensitivity analysis with the wave energy balance. This sensitivity analysis is a factor 30-85 faster than the benchmark.

The following recommendations are made to improve approaches for an efficient sensitivity analysis.

- Investigate approaches that speed-up the sensitivity analysis using a complex model with limited computational effort. This is especially relevant for cases, where simple models such as linear wave theory are not valid.

- Conduct the same research with a case study on which sufficient data is available. Then, the accuracy benchmark can consist of actual measurement data.

- Investigate the performance of an efficient sensitivity analysis focusing on other parameters than the bulk alongshore sediment transport rates.

1. Introduction

This first chapter introduces the research topic and the research objective. Furthermore it sets out the methodological steps to fulfill this objective. Section 1.3 provides a brief outline of the report structure.

1.1 Context

1.1.1 Numerical modeling of coastal erosion

Many coastal areas are under threat from coastal erosion (De Vriend and Van Koningsveld, 2012, Filatova et al., 2011, Stive et al., 2013). This is caused by e.g. sea level rise and human-induced coastal changes like the construction of harbor entrances (Bosboom and Stive, 2013). Coastal erosion can have severe effects on human life through an increase in flood risk and/or the loss of beach. Therefore, coastal erosion is an important problem to tackle. Hence, a lot of research on coastal erosion is carried out (Royal HaskoningDHV, 2013, Van Rhee, 2012, Philips and Jones, 2006, Zhang et al., 2004). The development of sophisticated numerical models is an important part of this research. These models simulate coastal erosion processes. With the model results, insights in the important processes of the coastal area under study are obtained (Deltares, 2014b, Deltares, 2010, Roelvink and Walstra, 2004). However, modeling inherently entails a degree of uncertainty. This is due to, inter alia, natural variability and our limited understanding of underlying processes involved in erosion. Full elimination of these uncertainties is impossible. Therefore hydraulic engineers always have to deal with uncertainties. Proper model calibration and validation reduce the degree of uncertainty of model output. This requires data on the coastal area under study, e.g. on nearshore wave characteristics and shoreline evolution.

1.1.2 Data poor environments

In data poor environments not much data on the coastal area under study is available. This complicates model calibration and validation. Consequently, the degree of uncertainty of model output is higher compared to data rich environments. Often time and budget, necessary to obtain more data on the area, are lacking. Therefore, projects in data poor environments have to find a way to deal with the high degree of uncertainties involved in modeling (e.g.: Hamby, 1994, Kurowicka and Cooke, 2006, Van der Klis, 2003).

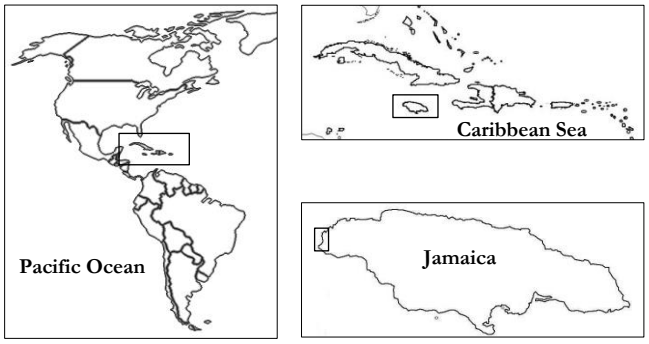

Figure 1.1 – Location case study, Negril - Jamaica

The Negril Coast at Jamaica (Figure 1.1) is an example of a data poor environment. The coastline consists of a strip of hotels, restaurants and resorts (UNEP, 2010, Rhiney, 2012, Edwards, 2009, Robinson et al., 2012). Uncertainty exist on the local shoreline erosion. The inhabitants are convinced that the beach suffers from long-term erosion. This may lead to enormous future income losses for the local tourist sector and the local people. However, studies are not unanimous about the local erosion rates. For instance, Robinson et al. (2012) indicates an erosion rate that is four times lower than rates found by other researchers (Mitchell, 2012, Hendry, 1982, McKenzie, 2012, Mondon and Warner, 2012, Robinson and Hendry, 2012). Actual data on coastline evolution is hardly available.

The data poorness of the Negril Coast is further increased due to limited data on e.g. nearshore wave characteristics and a limited understanding of the local processes. The latter is partly due to the complexity of the area. This complexity is caused by the presence of coral reefs in the bay and coastal inlets in the near area (Mitchell, 2012, Hendry, 1982, McKenzie, 2012, Mondon and Warner, 2012, Robinson and Hendry, 2012).

1.1.3 Sensitivity analysis

A sensitivity analysis can be used to appoint uncertainty in the model output to different sources of uncertainty in the model input (Saltelli, 2004, Confalonieri et al., 2010). This provides insights in the model accuracy and provides a first glance of the system dynamics. For a sensitivity analyis, at least more than one model execution is necessary (Van Gelder, 2000, Kurowicka and Cooke, 2006, Hamby, 1994, Hamby, 1995, Saltelli, 1999). Many methods, called sensitivity methods, for conducting a sensitivity analysis exist (Saltelli, 1999). A sensitivity analysis can be very time-consuming due to: (i) the amount of computation time a numerical process based model may require (Walstra et al., 2013, Deltares, 2010), and (ii) the large number of model executions required by particular sensitivity methods (Saltelli, 1999).

The aim of coastal modeling studies is to obtain *accurate* insights in the system dynamics *within time and budget constraints*. Conducting a sensitivity analysis is a way to obtain these accurate insights when uncertainties are involved (Hamby, 1994, Hamby, 1995). However, sometimes time and budget constraints cannot be met. Therefore, engineers can use approaches that speed-up the sensitivity analysis. The following three approaches are considered:

- Using simpler models in order to avoid excessive computational effort (Deltares, 2010).
- Using a sensitivity method that requires a smaller number of model executions (e.g. Saltelli, 1999).
- Reducing the size of the input dataset (Olij, 2015, Camus et al., 2011, Walstra et al., 2013).

However, the use of the three approaches may lead to limited and even inaccurate insights in the system dynamics (Olij, 2015, Camus et al., 2011, Walstra et al., 2013, Saltelli, 1999). Thus, on the one hand the use of these approaches may meet time and budget constraints, but on the other hand this may be at the cost of accuracy. Hence, engineers face the difficulty of *trading accuracy for speed*. The inverse relation between speed and accuracy is called the *speed-accuracy trade-off* (SAT) (Dambacher and Hübner, 2013, Drury, 1994).

1.2 Research objective

In the context of coastal erosion studies, an *efficient* sensitivity analysis is typically characterized by obtaining accurate insights in the system dynamics within the project's time and budget constraints. Using the three approaches explained this may not be possible. Therefore, the objective of this research is:

Investigate approaches for conducting an efficient sensitivity analysis in coastal erosion studies

For this it is important to obtain insights in the SATs related to the three approaches explained. Therefore the three sub-objectives are to assess the SATs between:

- The use of a simple and a complex model for conducting a sensitivity analysis.
- The use of several sensitivity methods for conducting a sensitivity analysis.
- Conducting a sensitivity analysis with a reduced and a full input dataset.

1.3 Methodology and report structure

In order to fulfill the objectives, several sensitivity analyses needed to be evaluated on their performance based on speed and accuracy. A point of reference by which something can be measured, i.e. a *benchmark,* was necessary for a proper evaluation. The benchmark was established by conducting a sensitivity analysis generating results as accurate as possible within the time constraints of this research. The resulting speed and accuracy functioned as the benchmark against which to compare the speed and accuracy of the other analyses.

A case study was used for conducting several sensitivity analyses and to evaluate their performance. A sensitivity analysis is useful when uncertainties are involved. Hence, the case should involve a certain degree of uncertainty. Therefore Negril Coast, Jamaica (Section 1.1.2) was selected as case in this research. The case-specific findings were, if possible, generalized. In this way the objectives were fulfilled *generically*.

The methodological steps that were undertaken to fulfill the research objectives are explained below. Every chapter of the report discusses one step. This is displayed in the reading guide at the bottom of this page (Figure 1.2).

- Chapter 2. Theory on sensitivity analysis and uncertainties was gained through a *literature review*. Based on the theory, several sensitivity methods for conducting sensitivity analysis were selected. Furthermore, a way to identify and quantify uncertainties in data poor environments was selected.
- Chapter 3. The case study was *investigated* in depth. This included an analysis of the local problem, available data and of the uncertainties involved. Furthermore, the aim of the modeling study was defined.
- Chapter 4. A simple and a complex model were *set-up* to model the wave transformation near Negril Coast.
- Chapter 5. A benchmark was *establishe*d through conducting an extensive sensitivity analysis with the complex model, the most extensive sensitivity method and the full input dataset.
- Chapter 6. A sensitivity analysis was *conducted* using the simple model, the most extensive sensitivity method and the full input dataset. Its speed and accuracy were *compared* with the benchmark.
- Chapter 7. A number of sensitivity analyses were *conducted* using the complex model, the full input dataset and the simpler sensitivity methods selected. Their speed and accuracy were *compared* with the benchmark.
- Chapter 8. Different input reduction techniques were *explored* through a literature review. Consequently, a sensitivity analysis was *conducted* using the complex model and the extensive sensitivity method, but with a reduced input dataset. Its speed and accuracy were *compared* with the benchmark.
- Chapter 9. The case-specific findings were *evaluated* and if possible *generalized*. Furthermore, some recommendations for engineers, who aim to conduct an efficient sensitivity analysis, and for further research were made.

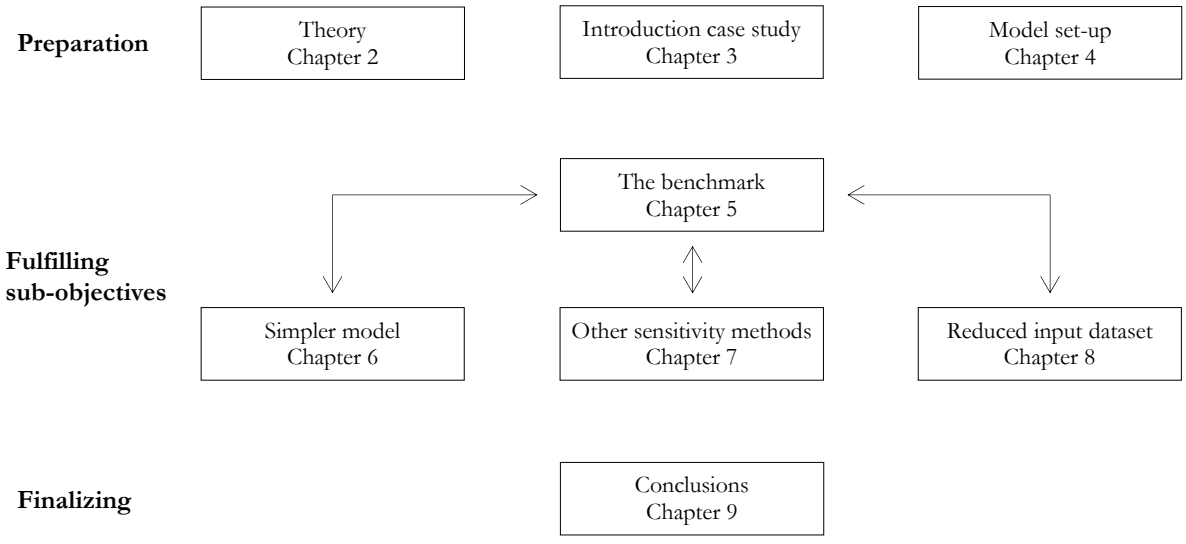

Figure 1.2 - Reading guide

2. Theory on sensitivity analysis, methods and uncertainties

This chapter discusses how a sensitivity analysis deals with uncertainties. The first section discusses what a sensitivity analysis is and what methods can be used to conduct one. Section 2.2 discusses uncertainties. It presents a possible uncertainty classification and it introduces a method for this uncertainty quantification. Section 2.3, the concluding remarks, provides a selection of methods that will be used for conducting the sensitivity analyses and how these methods deal with the uncertainties.

2.1 Sensitivity analysis

Sensitivity analysis is the study of how the uncertainty in the output of a model can be apportioned to different sources of uncertainty in the model input (Saltelli, 2004, Confalonieri et al., 2010). Crucial is to obtain a range of model outputs instead of one fixed model output. Therefore, it is necessary to carry out at least more than one model execution. For every execution the value of some uncertain parameter need to be changed. The *importance* of these uncertain parameters can be calculated on the basis of the range of outputs. Note the difference between sensitive and important parameters. Sensitive parameters have a significant influence on assessment results. Important parameters are those whose *uncertainty* contributes substantially to the uncertainty in assessment results (Hamby, 1994). Sensitive parameters are not necessarily uncertain and thus not necessarily important. This research focusses on *important* parameters. The extent to which a parameter is important depends thus partially on *the degree of uncertainty* of the parameter.

Conducting a sensitivity analysis can serve different purposes (e.g. Saltelli, 2004, Van der Spek, 2013):

- The model output ranges provide insights in the model reliability and accuracy (e.g. De Moel et al., 2012).
- Insights in parameter importance provide a first glance of the system dynamics (De Moel et al., 2012, Den Heijer et al., 2008, Hamby, 1994, Hamby, 1995, Saltelli et al., 1999).
- The results of a sensitivity analysis can steer data collection: it would be efficient to gather proper field data on the most important parameters (De Moel et al., 2012, Den Heijer et al., 2008, Hamby, 1994, Hamby, 1995, Saltelli et al., 1999).
- The range of outputs enables investigation of the occurrence of extreme or undesired events.
- The analysis forces engineers to investigate their own models, which can provide information on model calibration and potential model simplifications (De Moel et al., 2012, Saltelli, 1999, Saltelli, 2004, Campolongo and Saltelli, 1997, Confalonieri et al., 2010).

2.1.1 Categorization of sensitivity methods

Many sensitivity methods for conducting a sensitivity analysis are available. These methods differ in the way they (i) deal with uncertainties, and (ii) compute sensitivities (e.g. Kurowicka and Cooke, 2006). Figure 2.1, page 6, illustrates a possible way to categorize sensitivity methods.

The first and simplest category consists of the One at A Time (OAT) method. The OAT method considers a *base case*, consisting of a model execution using for every parameter its *basic value*. In all other model executions only one parameter value of the base case is changed. The analysis considers how this change affects the model output. Only two model executions, of which one is the base case, are needed to compute the importance of a parameter. The methods belonging to the second category are slightly more complicated. They or (i) changes *more parameters* at a time, or (ii) considers *more than one base case*. They require always more than two runs to compute the importance of a parameter. The probabilistic methods change, contrarily to the other methods, the parameter values in a probabilistic manner. The fourth category consists of the response surface methodology, which is an alternative way for conducting a sensitivity analysis.

The following sections discuss the six sensitivity methods of Figure 2.1 in depth. Note that there are many more sensitivity methods belonging to the categories. Moreover, there are many more categories available and even different ways of categorization possible. Hence, this section does not attempt by any means to be a complete review of all the different methods available in literature.

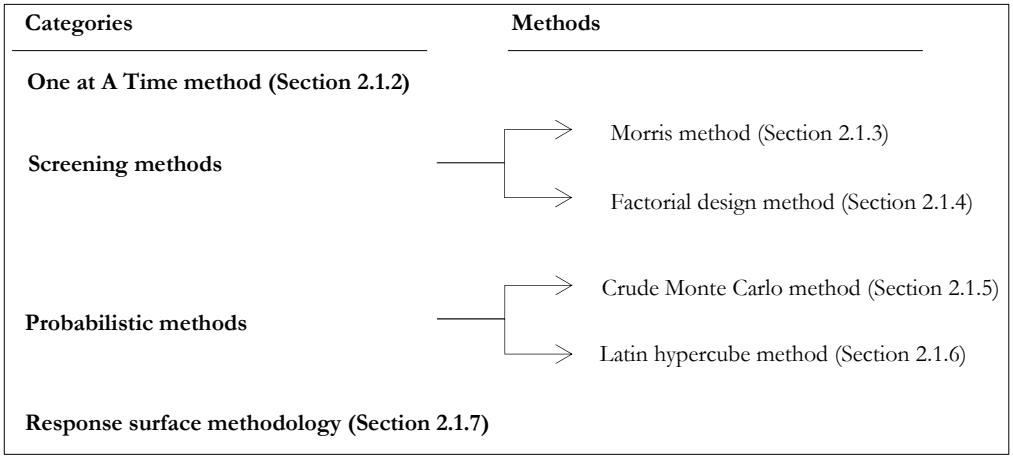

Figure 2.1 - Categorization of sensitivity methods (Based on e.g.: Rose, 1983, Helton, 1993, Kurowicka and Cooke, 2006)

2.1.2 One-At-a-Time method

The OAT method is the common practice in the hydraulic engineering sector. Figure 2.2 provides a visual representation of the method. The method assumes one *base case*, consisting of a model execution with for all the parameters *basic values*. This base case results in the *basic model output*. In every next run one parameter value is changed, while the other parameters keep their basic value (e.g. Saltelli, 1999). The model output of every run is compared with the basic model output. In this way the importance of a parameter is calculated. Of course it is possible to change the input variable twice, for instance to test increasing and decreasing of the parameter value.

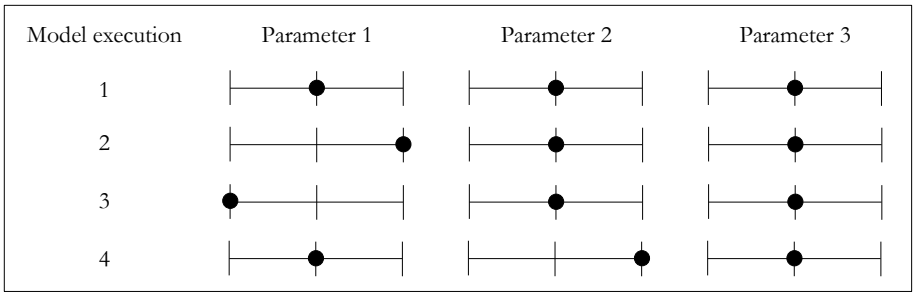

Figure 2.2 - Visual representation of four runs with the OAT method

The total number of model executions that is required to conduct a sensitivity analysis with the OAT method, assuming two changes for every parameter value, can be calculated through the following formulation:

$$Nr\ of\ executions\ required\ for\ OAT\ method = 1 + (n * 2) \quad 2.1$$

With n the number of parameters involved.

The importance of a parameter can be indicated by computing the so-called OAT-effect. The OAT-effect of an increase of the value of parameter one can be computed as follows:

$$OAT\ effect_{\ increase\ value\ parameter\ I} = R_2 - R_1 \quad 2.2$$

With R the model output, with the index indicating the particular model execution.

Since only one parameter is changed in every model execution, the OAT method discounts the existence of parameter interaction effects, i.e. the change in model output due to the change of two parameter

values at the same time. Hence, it ignores the original model complexity. If the model is linear and when the parameter values are independent, this is not a problem. However, when this is not the case, the method provides only a glimpse at the model behavior. Every conclusion drawn is, after all, only valid around the base case. This method is thus easy to implement and computationally cheap, but also very limited. Therefore, many sources in literature *discourage* to use this method (e.g. Saltelli, 1999).

2.1.3 Morris method

The Morris method exists of a randomized experimental plan that varies, like the OAT method, only one parameter at a time. However, the base case *changes* with every model execution. Therefore, the method wanders in the space of the parameter values, rather than just oscillating around the base case as the OAT method (Kurowicka and Cooke, 2006). In this way, the Morris method enables the determination of parameters having negligible, linear, non-linear- and interaction effects (Kurowicka and Cooke, 2006, Morris, 1991, Saltelli, 1999).

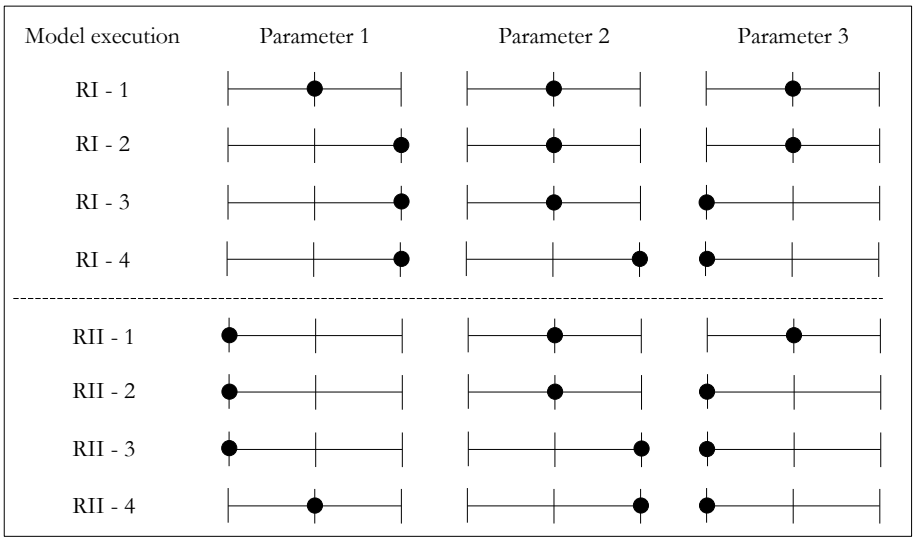

Figure 2.3 - Visual representation of the Morris method, with two orientations and three parameters having three possible values

First the Morris method is explained on the basis of an example. Figure 2.3 provides a visual interpretation of this example. It illustrates that three parameters can take three parameter values. In every model execution one parameter changes its value. *An orientation* consists of the model executions in which every parameter has changed its value once. In case of three parameters, an orientation consists of four model executions. In the first model execution of the new orientation, a *new base case* is formed. Again for every model execution one parameter changes its value. The number of orientations needs to be selected by the user. Values in the range of four-ten orientations giving accurate insights have been reported (Saltelli, 1999).

To compute the importance of a parameter the model output of an execution is compared with the model output of the previous execution. In this case the elementary effect of parameter three in orientation I can be computed as follows:

$$Elementary\ effect\ of\ parameter\ 3\ = \frac{R_{I,3} - R_{I,2}}{\Delta} \qquad 2.3$$

With R the model output, with the index indicating the particular model execution. With Δ the difference of the parameter values between model executions three and two.

In the same way the elementary effect of a parameter can be computed for the other orientations. The sample mean and the variance of all these calculated elementary effects are unbiased estimators of the mean and the variance of the finite distribution of elementary effects. The mean and variance indicate the importance of a parameter. The mean and the variance of the effects are mostly represented as by the markers in Figure 2.4 (page 8). The two lines relate the estimated means and the standard deviation in the following way (Kurowicka and Cooke, 2006):

$$d_i = \pm 2 \frac{S_i}{\sqrt{r}} \quad 2.4$$

With d_i the estimated mean, S_i the standard deviation and r the number of orientations.

When the mean and the variance of the parameter are located outside the triangle (Figure 2.4), it suggests that there is significant evidence that an effect of the parameter on the model output exist (Kurowicka and Cooke, 2006). When the variance of an effect is large, it indicates that the parameter is involved in interactions with other parameters. When the mean is large, it indicates a very important parameter. However, the judgment about the importance of parameters is context dependent (Morris, 1991).

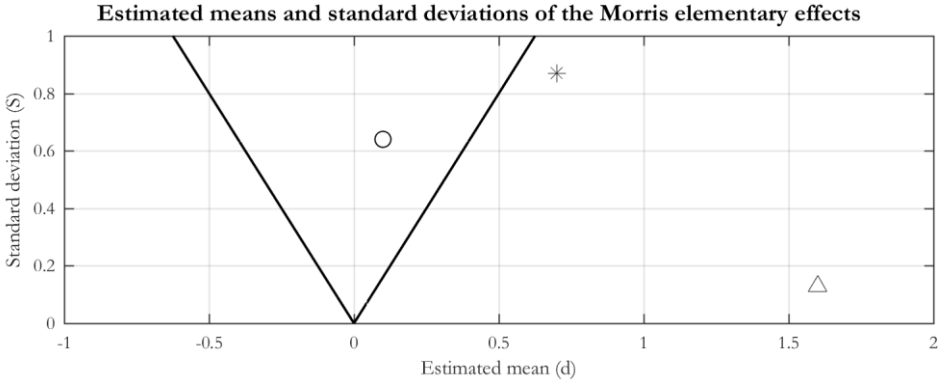

Figure 2.4 - The graphical representation of estimated means and standard deviations of elementary effects (Based on: Kurowicka and Cooke 2006)

The Morris method can be explained in a more generic way. For this a quantity of interest G is considered. G is the function of a number of k parameters x. Mathematically this can be described as: $G(x_1,x_2,\ldots x_k)$. The values of x are not fixed but can adopt different values. For every model execution one value for every parameter needs to be selected. Therefore, Morris (1991) first scales all the parameter x_k in order to be able select values in the interval [0,1]. Hence, the region of interest for the experiment is the k-dimensional unit hypercube. Secondly, he *discretizes* this unit hypercube to a k-dimensional *p-level* grid. For instance in case of p =4, all sides of the hypercube consists of 4 grid lines. These lines represent particular parameter values. When p=4, a parameter can thus take the values [0; 1/3 ; 2/3 ; 1]. The value of p can be freely chosen.

Now for every run a value for every parameter needs to be selected. Therefore Morris (1991) creates a randomized version of the following sampling matrix, based on k=3 and p=2 (so all three parameters can adopt the value 0 or 1):

$$B = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \quad 2.5$$

B is a m*k matrix, with m reflecting the number of model executions (always k+1). The elementary effect of a parameter can be computed as in equation 2.1. A sample of r effects is required to calculate a finite distribution of elementary effects associated with an input. Therefore, r independent random orientations of B need to be used for the analysis. The number of required model executions can be calculated as follows (Saltelli, 1999):

$$Nr\ of\ executions\ required\ for\ Morris\ method = r * (n + 1) \quad 2.6$$

With r the number of random orientations of the matrix, and n the number of parameters involved.

2.1.4 Factorial design method

The Factorial design method is the second method that belongs to the screening methods. The method ensures that all parameters involved adopt a minimum and a maximum value. The model is executed with

all combination of these values (Hamby, 1994, Box and Hunter, 1961, Kurowicka and Cooke, 2006, Edwards et al., 2009). Figure 2.5 illustrates a factorial design with three parameters, requiring eight (3^2) model executions.

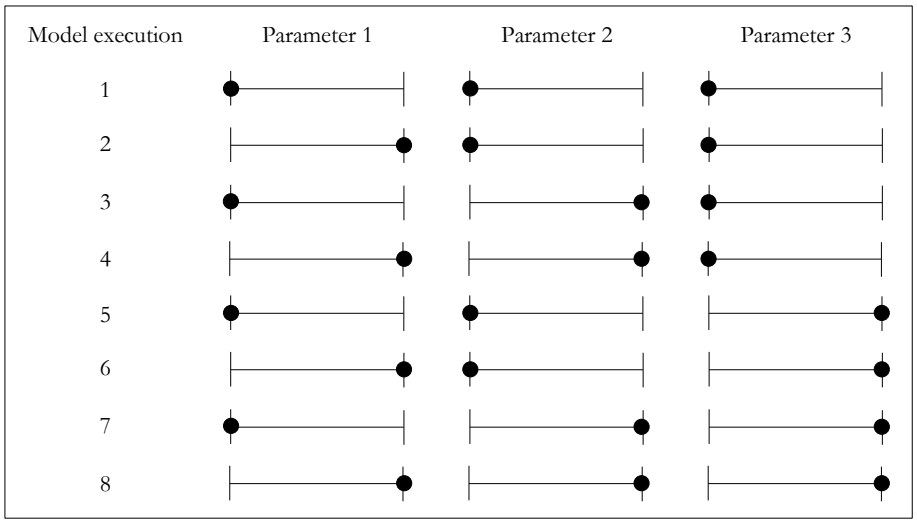

Figure 2.5 - Visual representation of a 2^3 full factorial design (Based on: Box and Hunter, 1961)

The importance of a parameter is investigated by computing the difference between the model outputs resulting from the high and the low level of that parameter. This can be done in the following way (Kurowicka and Cooke, 2006) :

$$Effect\ parameter\ I = \frac{\sum R_{all\ runs\ with\ par\ I\ with\ \max value} - \sum R_{all\ runs\ with\ par\ I\ with \min value}}{m\ /2} \qquad 2.7$$

With R the model output and m the total number of model executions. In the same way the effect of combination of parameters can be calculated. As explained before, these effects are called interaction effects.

Clear is that a large number of uncertain parameters requires a huge number of model runs. A way to decrease this number of model executions required is to use the Fractional factorial design method instead of the just explained Full factorial design method. A fractional factorial design requires half of the number of runs of a full factorial design. This efficiency can be reached because the parameter values are not independently varied between its high and low level, but *based on the change of other parameters*. Figure 2.6 represents a 2^{3-1} fractional factorial design. The change of parameter thee is defined by the product of the change of the first two variables. For instance, considering the first model execution, parameter one and two adopt their minimum values and, since -1 * -1 =+1, parameter three adopts its high value.

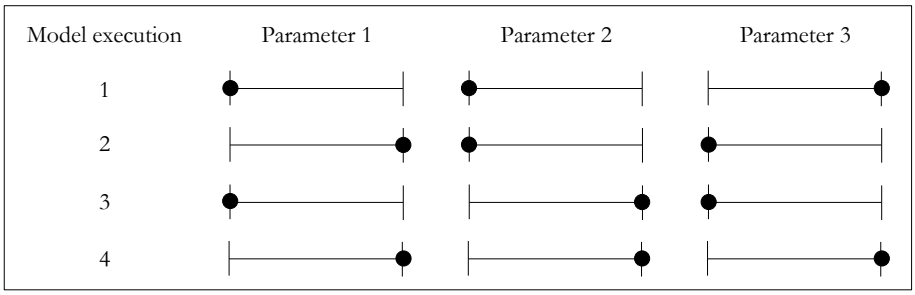

Figure 2.6 - Visual representation of a 2^{3-1} fractional factorial design (Based on: Box and Hunter, 1961)

The fractional design can be useful in case of (i) continuous variables, since the response is expected to vary smoothly, and (ii) negligible multi-factor interaction effects (Hamby, 1994, Box and Hunter, 1961, Kurowicka and Cooke, 2006, Edwards et al., 2009). A drawback of a fractional factorial design is that the main effect of a parameter (equation 2.5), is influenced by the levels of the other parameters. In other words, the main effect is *confounded* with other multi-factor interactions (Box and Hunter, 1961).

Considering the example of Figure 2.6 the combination of all parameters at their low value is not taken into account. When this effect is non-negligible this affects the computed importance of a parameter.

The so-called *design resolution* indicates to what extent the effects calculated by a Fractional factorial design are confounded with eachother (Box and Hunter, 1961). Considering the example of Figure 2.6, both the maximum and the minimum value of a parameter are combined with the maximum and the minimum value of the other parameters. However, the maximum and the minimum value of a parameter are *not* combined with all the *possible combinations* of the maximum and the minimum values of the other parameters. Hence, the main effect of a parameter is *not* confounded with any other main effect, but *is* confounded with the two-factor interactions of the other parameters. This is called a design of resolution III. Box and Hunter (1961) can be consulted for more information of design resolutions.

2.1.5 Crude Monte Carlo method

The Crude Monte Carlo method is the first probabilistic method that is discussed here. Under Crude Monte Carlo sampling the computer draws different input vector values (x_1, x_2, … x_k) pseudo-randomly from the joint distribution (X_1, X_2, … X_k). The computer is not able to draw random samples from the distributions, but can only simulate random numbers. Therefore it is called *pseudo-random sampling*. This sampling method is repeated m times, resulting in a probabilistic distribution of the model output. Figure 2.7 visually represents the Crude Monte Carlo method with sample size of m is three.

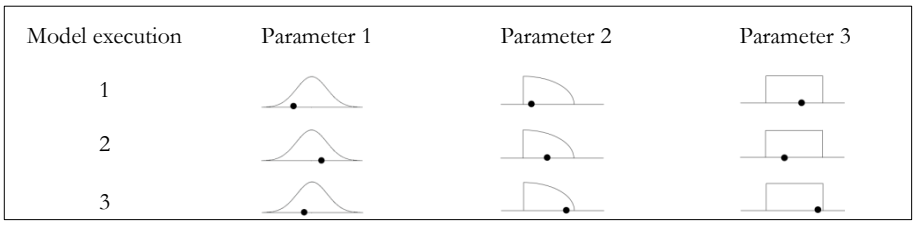

Figure 2.7 - Visual representation of the Crude Monte Carlo method with a sample size of m =3

The accuracy of the Crude Monte Carlo method is dependent on the sample size m (Muthén and Muthén, 2002). Despite of some attempts (Van der Klis, 2003, Morgan et al., 1992), it is not possible to compute the exact sample size that results in accurate insights *before* conducting a sensitivity analysis. The sample size resulting in accurate insights, depend on (Helton et al., 2005, Kurowicka and Cooke, 2006): (i) the amount of uncertain parameters, (ii) the degree of uncertainty of these uncertain parameters, (iii) the complexity of the model used, and (iv) the probabilities related to the insights desired. The sample size of a Crude Monte Carlo analysis can be easily increased, simply through sampling more samples (Sallaberry et al., 2008). Sample sizes in the ranges of hundreds to thousands are common.

Dependence modeling

The Crude Monte Carlo method allows including the dependence, i.e. the correlation, between the uncertain input parameters. Traditionally, the pairwise dependence between parameters has been described using bivariate distributions. The main limitation of this approach is that the individual behavior of the two random parameters, now called variables, must be characterized by the same parametric family of distributions. Of course this is not always the case. Copula models avoid this problem. A copula is a distribution on the unit square with uniform marginal distributions (Kurowicka and Cooke, 2006). A copula is a parametric model and many *copula families* are available to the modeler (Kurowicka and Cooke, 2006). Basically a copula joins or couples bivariate distribution functions to their marginal distribution functions. The copula approach states that the joint cumulative distribution function $H(x_1, x_2, \ldots x_k)$ of any pair of x_k random parameters can be written in the form (Genest and Favre, 2007, Rose and Smith, 2012, Schweizer, 2007):

$$H_{x_1 \ldots x_k}(x_1, x_2 \ldots, x_k) = C\{D(x_1), E(x_2), \ldots, F(x_3)\} \qquad with\ x_1 \ldots x_k \in \mathbb{R} \qquad 2.8$$

With D, E and F marginal distributions and C the copula.

Every continuous bivariate distribution can be represented in terms of a copula. Moreover, it is possible to find a unique copula that corresponds to a given continuous joint distribution (Kurowicka and Cooke, 2006, Genest et al., 2009a).

Results of a probabilistic sensitivity analysis: output ranges

The model output range is an important result of a probabilistic sensitivity analysis. These output ranges allow the following two important aspects:

- Estimating the model output distribution
- Estimating a copula that describes the dependence between an uncertain input parameter and the uncertain model output.

To estimate the output distribution of the model output, an empirical cumulative distribution function ($F_n(R)$) can be used. $F_n(R)$ is defined at a point R_i as the proportion of elements in the dataset that are less than or equal to R_i (Dekking et al., 2005):

$$F_n(R) = \frac{number\ of\ elements\ in\ the\ data\ set\ \leq R}{n} \quad 2.9$$

With n the total number of elements in the dataset.
Moreover, Matlab functions are available that estimate the best fitting probability density function to the data.

The estimation of a copula can be done in many ways (Genest et al., 2009a, Genest et al., 2009b, Kallenberg, 2009, Charpentier et al., 2007). Because this research focusses not on finding the best estimation for a copula, only semi-correlations for the estimation are considered. The way semi-correlations can be used to estimate a copula is described in Appendix A.

Results of a probabilistic sensitivity analysis: Sensitivities

The correlation handles on the degree to which two parameters are related. The correlation between an uncertain input parameter and an uncertain model output can be used as a measure for sensitivity. The correlation can be qualitatively determined by a *scatter plot*. However, for a more quantitative approach, the correlation coefficient can be calculated. The *correlation coefficient* is a number between –1 and 1 that determines whether two paired sets of data are related. The closer to 1 the more confidence is present for a positive linear correlation and the closer to –1 the more confidence for a negative linear correlation. When the correlation coefficient is close to zero there is no evidence of any relationship.

Several correlation coefficients exist. The simplest, Pearson, correlation coefficients is a measure for the degree of *linear dependence* between two parameters. When dealing with non-linear relationships and rank transformation need to be used. The input and output values are replaced by their rank, which results in linear relationships. The *rank correlation coefficient* (RCC) can indicate the degree of monotonicity between the input and the output sample. The RCC can be calculated through the same calculation as the simple correlation coefficient but with operating on the rank transformed data (e.g. Hamby, 1994, Kurowicka and Cooke, 2006). When computing sensitivities, it is crucial to consider the correlation between input parameters. This correlation may influence correlation between an uncertain input parameter and the model output. Therefore *partial correlation coefficients* (PCC) can be calculated. Partial correlation measures the degree of association between two random parameters with the effect of a set of controlling random parameters removed (Kurowicka and Cooke, 2006).The rank transformation can also be applied to these PCC, turning the coefficient in the partial rank correlation coefficients (PRCC) (e.g. Hamby 1994; Kurowicka and Cooke). Details on correlation coefficients can be found in Appendix A.

In order to draw accurate conclusions from the correlation coefficients, it is important to investigate the ir *significanc*e (Fenton and Neil, 2013, Dekking et al., 2005). The standard method that statisticians use to measure the significance is the t-test resulting in a *p-value* (Fenton and Neil, 2013, Dekking et al., 2005). The t-test uses the null hypothesis H0 and the alternative hypothesis H1. The null hypothesis is that the two parameters are uncorrelated. The t-test finds how much evidence one can find against H0, i.e. find the probability that expresses how likely it is to obtain a value of the test statistic z0 at least as extreme as the value Z observed from the data. The z_0 can be calculated as (Dekking et al., 2005):

$$z_0 = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \qquad 2.10$$

With r the correlation found and n the number of samples.

The corresponding probability can be found through tables, or the built-in Matlab function. This probability is the p-value. The p-value is a number between 0 and 1 representing the probability that this data would have arisen if the null hypothesis were true. A high p-value, generally higher than 0.01-0.03, indicates a low significance. A low p-value indicates a highly significant correlation.

2.1.6 Latin hypercube method

To reduce the number of model runs required for a probabilistic sensitivity analysis, variance-reducing techniques can be used. The Latin hypercube method is an example of such a technique. The technique divides the range of each variable into N intervals of equal probability (Kurowicka and Cooke, 2006, McKay et al., 2000). This is shown in Figure 2.8. The cumulative probability distribution is used to divide the probability density function in bins of equal probability. Then one value, for each parameter, in every bin is randomly selected. Finally, random combinations of these values are made. Because of this steering of the sampling procedure, the samples are wider spread compared to the Crude Monte Carlo method. This is the reason a smaller number of samples can be used (Kurowicka and Cooke, 2006, McKay et al., 2000).

The characteristics of the results of a sensitivity analysis using the Latin hypercube method are similar as of a sensitivity analysis using the Crude Monte Carlo method.

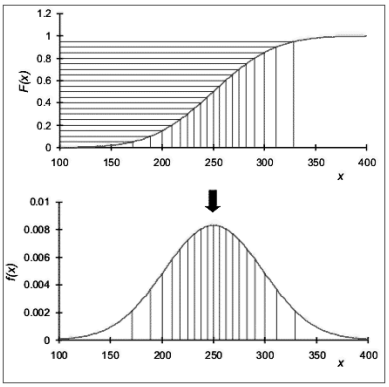

Figure 2.8 - Visual representation of the Latin hypercube method

Sample size

As mentioned before it is impossible to define the sample size of a probabilistic method beforehand. Different theories on the required sample size for the Latin hypercube method can be found (Matala, 2008, Van der Klis, 2003). Literature mentions sample sizes of 25, 30, twice the number of random variables and much larger sample sizes. In most cases the resulting accuracy has not been investigated. Normally the only way to investigate the performance of the method is to repeat the method several times, but this requires again a larger sample size (Van der Klis, 2003).

A drawback to Latin hypercube sampling is that its highly structured form makes it difficult to increase the size of an already generated sample while simultaneously preserving the stratification properties that make Latin hypercube sampling so effective (Sallaberry et al., 2008). However, it is possible through a procedure using rank correlated variables (Sallaberry et al., 2008). Details on this can be found in Sallaberry et al. (2008)

Dependence modeling: Vines

To include the dependence between two or more uncertain input parameters, a special sampling procedure needs to be followed. This procedure requires independent uniform parameters, between which

dependence is induced by transforming the parameters in a way which uses previously sampled values (Kurowicka and Cooke, 2006). The transformations are the conditional cumulative and the inverse cumulative distributions of the copulas in an appropriate vine representation (Kurowicka and Cooke, 2006). Appendix A explains more on (dependence) trees and vines. Here, just the sampling algorithm for the canonical vine on four parameters is presented (for explanation see Appendix A).

This algorithm samples four independent uniform (0,1) parameters, now called variables, $u_1,\ldots, u_4$. The variables $x_1,\ldots,x_4$ are assumed to be uniform (0,1) as well. The conditional correlation between variables (i,j) given k is denoted as $r_{i,j|k}$. In case a variable is correlation with more than one variables, k is a vector. $F^{-1}_{r_{13};u_1}(x_j)$ denotes the inverse cumulative distribution function for x_j given u_i under the conditional copula with correlation $r_{i,j|k}$. The algorithm can be stated as follows (Kurowicka and Cooke, 2006):

$$x_1 = u_i \tag{2.11}$$

$$x_2 = F^{-1}_{r_{12};u_1}(u_2)$$

$$x_3 = F^{-1}_{r_{13};u_1}(F^{-1}_{r_{23|1};u_2}(u_3))$$

$$x_4 = F^{-1}_{r_{14};u_1}\left(F^{-1}_{r_{24|1};u_2}\left(F^{-1}_{r_{34|12};u_3}(u_4)\right)\right)$$

The uniform parameters $u_1,\ldots,u_4$ are sampled independently. The dependent parameters $x_1,\ldots,x_4$, are obtained by applying successive inverse cumulative distributions.

2.1.7 Response surface methodology

The last category of sensitivity methods is the response surface methodology. A response surface is a simple mathematical function which attempt to replace the use of a time-consuming deterministic model. The surface, created by using a moderate number of model runs, is an approximate version of the full model, such that the error of approximation is minimized in the region of interest. Therefore the approximation holds in the small region of interest only. The initial analytical effort to develop a good response surface may be significant, but once it is available the surrogate for the original model can be used to conduct sensitivity and uncertainty analyses fast and easy. The final uncertainty and sensitivity results are naturally not better than the response surface approximation to the original model (Morgan et al., 1992, Helton, 1993).

Several issues concerning the response surface methodology exist (Helton, 1993):

- It is difficult to develop an appropriate design with which the original model can be executed. This is often due to many parameters involved and unknown model behavior.
- The construction of an appropriate response surface approximation may require a considerable amount of statistical sophistication and thus time.
- There are some difficulties in incorporating correlations and restrictions between input parameters. Because of these difficulties the methodology works when there are only a few (e.g. less than ten) input parameters.
- The relationships between the input and output parameters need to be basically linear or quadratic.
- A design that is appropriate for one output parameter may not be appropriate for a different output parameter(Helton, 1993). Therefore the methodology works only for models with a limited number of distinct output parameters.

2.2 Uncertainties

2.2.1 Classification of uncertainties

Following Van der Klis (2003) two main sources of uncertainties exist: variability and limited knowledge. Uncertainties originating from the first source are inherent to particular process for instance the wave

height in the next storm. Therefore is elimination of these uncertainties not possible. Uncertainties originating from limited knowledge can be caused by the level of knowledge of the modeler, the general state of knowledge and by variability. Uncertainties originating from incomplete or uncertain knowledge of processes can be eliminated through obtaining more knowledge.

The ordering based on sources is thus important for the investigation to what extent the uncertainties can be eliminated. However to choose a method to assess the effects of uncertainties on model results, Van der Klis (2003) categorizes uncertainties in terms of their manifestation in a particular model. In this way the following four levels of uncertainties can be distinguished (Van der Klis, 2003):

- Technical uncertainties are uncertainties in input data and model parameters.
- Epistemological uncertainties are uncertainties related to the model completeness and validity.
- Methodological uncertainties are uncertainties due to assumptions underlying the model.
- Model operation uncertainties, are uncertainties caused by implementation and numerical errors. For instance uncertainty introduced due to the truncation error of the numerical scheme.

Table 2.1 provides an overview of some uncertainties related to hydraulic engineering – divided in the categories defined.

Table 2.1- Overview of some uncertainties related to hydraulic engineering

Technical	**Epistemological**	**Methodological**	**Model operational**
Bathymetry	Sea level rise	Linear wave theory	Instabilities
Wave data	Future temperate	Wave energy balance	
Wind data	Future salinity	Grid	
Bottom friction	Flora growth	Morphological factor	
Sediment size		Wave input reduction	
Area characteristics			
Coastline shape			
Water level			

2.2.2 Quantification of uncertainties

Quantification of uncertainties is necessary for the sensitivity analysis to deal with the uncertain parameters. Section 2.1 explains that the way of quantification depends on the sensitivity method used. In case a probabilistic method is used, input distributions are necessary. When a non-probabilistic method is used, this is not necessary. Quantification of distributions is more complex. Therefore this section assumes the use of a probabilistic method. The non-probabilistic quantification can be derived from this.

When some data on the uncertain parameter is available, this can be used for quantification. A best fit probability distribution can be drawn through the data (see Appendix A). When no data is available it becomes more complex. Conducting measurements on the quantities of interest takes a lot of time or can be very expensive. Therefore expert judgment is used to compute the degree of uncertainties. Cooke and Goosens (2008) define the following steps:

- Ask calibration question to experts.
- Score experts' performance as subjective probability assessors based on calibration questions.
- Ask experts to provide predefined percentiles (<5%, 50%, >5%) for the parameters of interest. Note: in non-probabilistic methods these percentiles can function as parameter values.
- Compute resulting distributions and compute the lowest (a) and the highest value (b) of a parameters through (Cooke and Goosens, 2008, Cooke and Slijkhuis, 2003):

$$a = L - k\left(\frac{U-L}{100}\right) \quad 2.12$$

$$b = U + k\left(\frac{U-L}{100}\right) \quad 2.13$$

In which k defines the overshoot rule, which is typically 10% and in which [L,U] is the smallest interval containing all assessed quartiles of all experts.

The range between a and b is divided into five parts – each representing a uniform distribution. Then thousands of numbers are generated from these distributions. Finally a probability density function is drawn through these numbers (Appendix A)

- Evaluating and validate the combination of expert judgments.

The program Excalibur (Lighttwist Software) supports the assessment of expert judgment.

It is very time consuming to follow all the steps defined. The scoring of the experts' performance based on calibration questions can be skipped in order to speed up the process.

Uncertain dependencies

The dependencies between input parameters are often uncertain. This uncertainty can be elicited through expert judgment as well. However Werner et al. (Future), show that there is only limited experience for expert judgment within a copula modeling framework. This is because copulas are distinguished by measures of association, for instance rank correlations, but also by its behavior along the dependence function as indicated by the family. This can be too much detail to query from experts (Werner et al., Future). An option is to ask just the correlation coefficients between several parameters to experts, and to *assume a copula family*. This involves another uncertain parameter, namely the assumed copula family. When it is aimed to asses this uncertainty, different copula families can be assumed and their results can be compared.

2.3 Concluding remarks

This chapter presents theory on sensitivity analysis and uncertainties. The following points are important for conducting the several sensitivity analyses for this research:

- Due to the several issues related to the response surface methodology, this method is not used to conduct a sensitivity analysis. Table 2.2 shows the sensitivity methods, and their most important characteristics related to accuracy, that will be used for conducting sensitivity analyses.

Table 2.2 - Characteristics of several sensitivity methods

	Non-linear / non-monotone model output behavior	Correlated input parameters	Interaction effects between input parameters	Model output distributions	Extra note
OAT	No	No	No	No	Simple
Morris	Limited	No	No	No	Random method
Factorial design	No	No	Yes	No	Focusing on extremes
Crude Monte Carlo	Yes	Yes	Yes	Yes	Easy extension samples
Latin hypercube	Yes	Yes	Yes	Yes	Difficult extension samples

- For probabilistic methods it is not possible to compute the exact sample size that results in accurate insights *before* conducting a sensitivity analysis. The sample size, resulting in accurate insights, depend on: (i) the amount of uncertain parameters, (ii) the degree of uncertainty of these uncertain parameters, (iii) the complexity of the model used, and (iv) the probabilities related to the insights desired.

- For a Crude Monte Carlo method the sample size can be easily increased, simply through sampling more samples. For the Latin hypercube method it is more difficult to increase the sample size while simultaneously preserving the stratification properties that make Latin hypercube sampling so effective.

- The sensitivity analysis that will be conducted focusses on *important parameters*. The extent to which a parameter is important depends thus partially on *the degree of uncertainty* of the parameter.

- The degree of uncertainty is indicated by a *limited expert judgment process*. This process *excludes* the calibration questions to score expert's performance and assumes copula families for the dependencies between input parameters.

3. Introduction to the case: Negril Coast, Jamaica

This chapter introduces the case study. It describes the study site (section 3.1) and analyses the data available to the project (section 3.2). The modeling objective and the aim of the sensitivity analysis are clarified in section 3.3. Section 3.4 discusses the uncertainties involved in this sensitivity analysis. Finally, Section 3.5 presents the chapter's most important findings regarding the further research.

3.1 Introduction

3.1.1 Study site

Figure 3.1 shows that Negril's coastline has been divided into 2 parts: Bloody Bay in the north and Long Bay in the south, of which the latter, with its coastline of 6.4 km, is the focus of this research. The 2 bays are separated by a rocky peninsula and a small island, Booby Cay (Smith Warner International Ltd., 2007, Ten Ham et al., 2006). Negril used to be an isolated and remote fishing village, but since the construction of a highway in 1959 the coastline changed into a strip of development with hotels and resorts (UNEP, 2010, Rhiney, 2012, Edwards, 2009). The Great Morass, through which the Negril River runs, is located just inland of a line of developments (Robinson et al., 2012).

The area in front of the beach is shallow with a mild slope. The inner shelf, with a depth to 20 m, extends out 1-2 km from the shore. A large part of the inner shelf is covered with seagrass beds. Behind the inner shelf the outer shelf start. Along the edge of this outer shelf, at a distance of 2-3 km from the shore, a fringing line of coral reefs is located. Seaward of these reefs the water depths become greater than 200 m (Smith Warner International Ltd., 2007). Besides the coral reefs at the edge of the outer shelf, there are some coral reefs located in the shallow waters of Long Bay. The largest is a reef of around 700 m, from north to south, located 1.5 km offshore (Smith Warner International Ltd., 2007, Ten Ham et al., 2006).

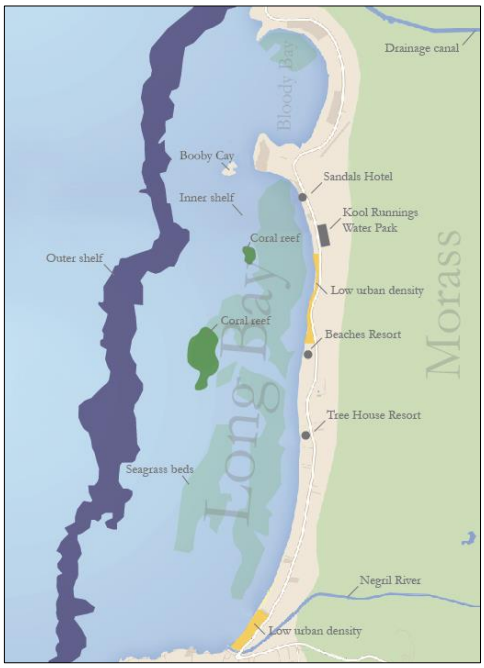

Figure 3.1 - Overview Negril, Jamaica (Van Arkel, 2015)

3.1.2 Coastline evolution

Long Bay is suffering increasingly from erosion since 1950. This is a problem for the local tourist sector (Edwards, 2009, CL Environmental, 2014, Robinson and Hendry, 2012). However, data on shoreline evolution is very poor. Therefore uncertainty on the actual erosion rates exists (Mitchell, 2012, Hendry, 1982, McKenzie, 2012, Mondon and Warner, 2012, Robinson and Hendry, 2012, Robinson et al., 2012). Clear is that the erosion is not an alongshore uniform occurring phenomenon. Besides that there is a big

difference in erosion caused by storm events and time-averaged erosion. Based on literature the following conclusions can be drawn (e.g. Department of Geology and Geography, 2002, Mondon and Warner, 2012):

- The northern and southern parts of the bay experience the highest rate of erosion. Recovery seems to take place during calmer conditions.
- The middle section of the coastline experiences seasonal fluctuation in erosion and appears to be more stable than the northern and the southern parts.
- The part of the beach located behind the coral reef, does not experience much erosion. This section might have even experienced accretion over the last decades.

3.2 Data analysis

3.2.1 Bathymetric data

The following two sources of bathymetric data were available to the project:

- Bathymetric survey by (Smith Warner International Ltd., 2007)between September 23rd and October 5th 2013;
- GEBCO data from 2008 (GEBCO, 2015), supplemented with some details from Google Earth (2015)

Figure 3.2 presents both bathymetries. As can be seen the GEBCO bathymetry is much deeper close to the shoreline compared to the surveyed bathymetry. Consequently, the coral reef is located in much deeper water than in the surveyed bathymetry.

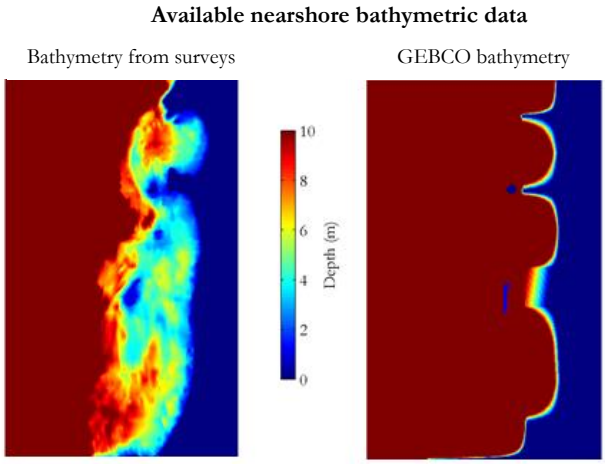

Figure 3.2 – Available nearshore bathymetric data (Smith Warner International Ltd., 2007, GEBCO, 2015)

3.2.2 Remote data sources

The shoreline position of the Negril Coast has not been measured structurally. Therefore Smith Warner International Ltd. (2007)compared historical shoreline data in the form of aerial and satellite images. Aerial photos of 1968, 1980, 1991 were compared with a satellite image from 2003 and a mapped shoreline position of 2006. (Smith Warner International Ltd., 2007) indicates a long term erosional trend. However they note the poor quality of the aerial photographs. Following (McKenzie, 2012) this results probably in incorrect shoreline trends. Therefore the findings of (Smith Warner International Ltd., 2007) are not further taken into account.

For this project more satellite images are compared. Figure 3.3 compares Google Earth satellite images of 2003, 2009, 2013 and 2015. The coastal stretch behind the coral reef is shown. As can be seen the beach width in the pictures of 2009 and 2013 is much smaller than the width of 2003. However, the figure of 2015 shows a beach which is even wider than the beach of 2003. These figures thus suggest a beach that erodes, but also recovers again. A long term erosional trend *cannot* be indicated from the figures.

Figure 3.3 - Comparison between historical satellite images (Google Earth)

Figure 3.4 compares NASA Landsat images, available from 2005 onwards (NASA). The lines show the *average* shoreline position in a period of 180 days. In this way the tidal influence is cancelled out, however uncertainties remain involved. Often the images indicate the wave swash zone as land. Therefore the coastline is most probably located more towards the sea side. The figure indicates very small differences only, which indicates now major changes in the shoreline position. This may imply low sedimentation and erosion rates or high variation in erosion and sedimentation rates at all locations.

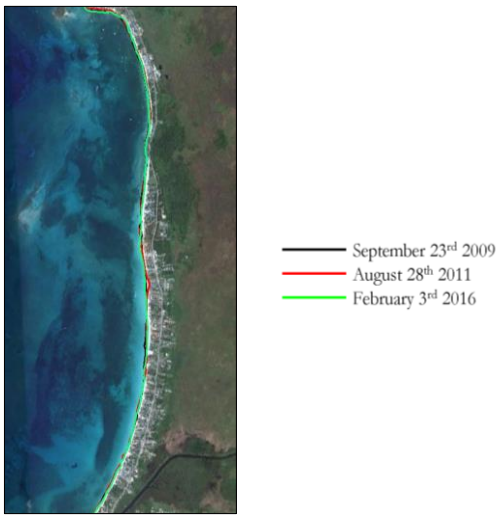

Figure 3.4 - Comparison of Landsat satellite images along the Negril Coast (NASA)

3.2.3 Water levels and currents

The Caribbean Sea is a micro-tidal environment with tidal ranges of a few decimeters only (Torres and Tsimplis, 2011). This is in agreement with the water levels and currents measurements carried out by CL Environmental (2014) in September and October 2013. They measured a maximum tidal range of 0.345 m during spring tide. Furthermore, they measured currents not larger than 0.2 m/s. The currents were generally in North-South direction (CL Environmental, 2014).

3.2.4 Wave and wind climate

No observed offshore wave data are available to this project. Instead only *hindcasted* wave data from, locally non-validated, global wave models, were available. Two sources of hindcasted data are available:

- The European Centre for Medium-Range Weather Forecasts (ECMWF, 2015) provides hindcasted data every 6 hours in the period from 1979 to 2012.
- National Oceanic and Atmospheric Administration (NOAA, 2015) provides hindcasted data every 3 hours in the period from 2005 to 2012.

The models use wind data to determine the local wave climates. The wave and wind data from global wave models contain some error when compared against measured wave data (Van Steijn, 2015, Appendini et al., 2013). This is due to simplification and assumptions in the analytical model used and due to measurement errors in the input data (Brooker et al., 2004). The error can be subdivided into a bias and a random variation. Bias can be eliminated by calibrating hindcasted model results through a measurement in the area. However, the random variation remains. Not much literature on the amount of bias is available (Brooker et al., 2004, Wang et al., 2014, Cardone, 1987, Bitner-Gregersen and Soares, 2007).

Figures 3.5-3.8 presents the wave height and wind speed roses for both datasets. As can be seen the locations of the NOAA and ECMWF data differ from each other. Moreover, the locations of NOAA available wave and wind data differ from each other. Both datasets indicate that: (i) most waves travel in *opposite direction* of the Negril Coast, (ii) most waves are below 1.5 m, but that waves above 4 m do occur.

Figures 3.9-3.12 presents scatter plots from the offshore wave height and period, and offshore wave and wind direction from ECMWF and NOAA respectively. Figure 3.9 shows that a wave period higher than 9s hardly occurs in the ECMWF dataset. This implies that there is hardly any swell. This is in agreement with the clear relationship between the wave and wind directions. Considering the data of NOAA there seems some, but not much, swell present in the area. Figure 3.12 shows a very strange the relation between the offshore wave and wind direction (NOAA). The accuracy of this relation can be considered as doubtful.

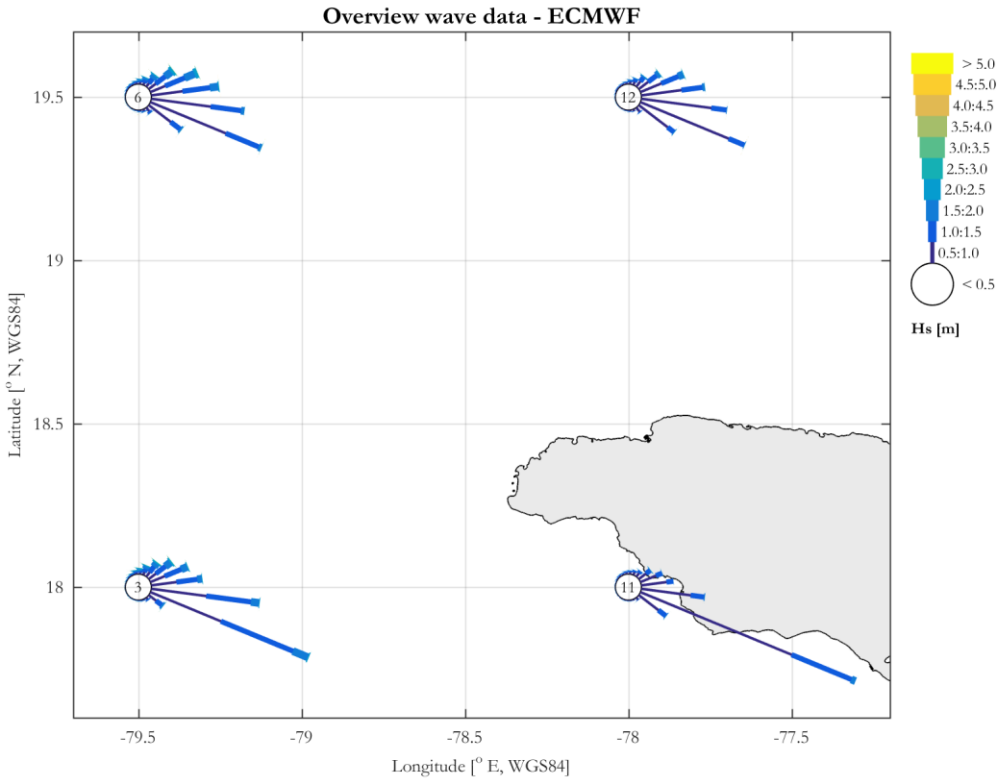

Figure 3.5 - Overview ECMWF wave data (ECMWF, 2015)

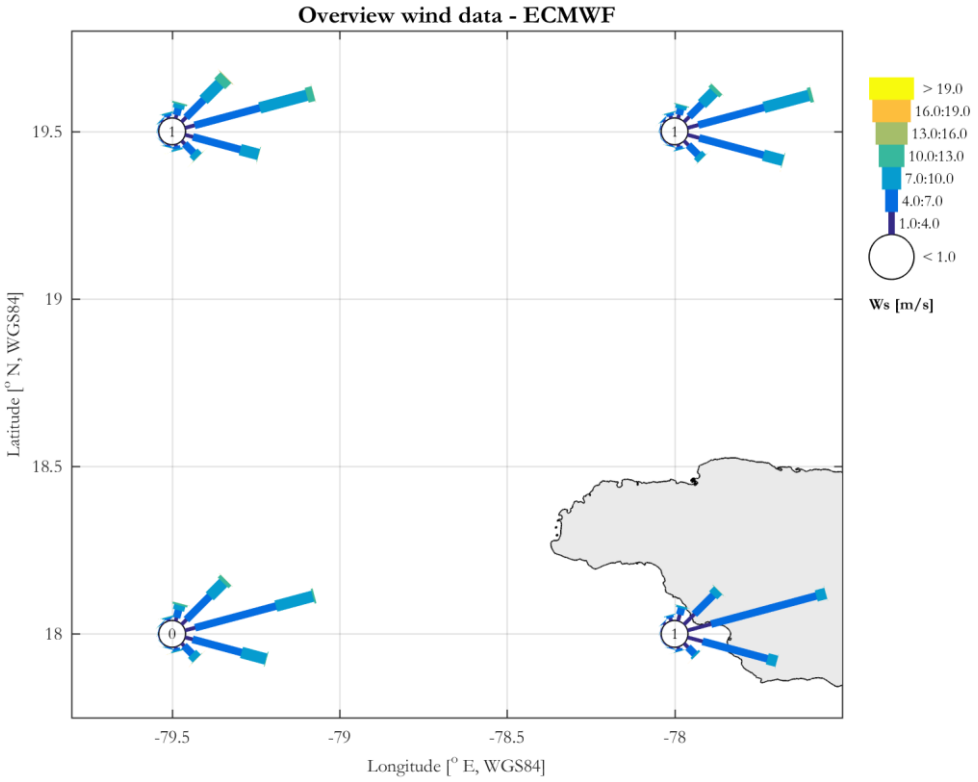

Figure 3.6 - Overview ECMWF wind data (ECMWF, 2015)

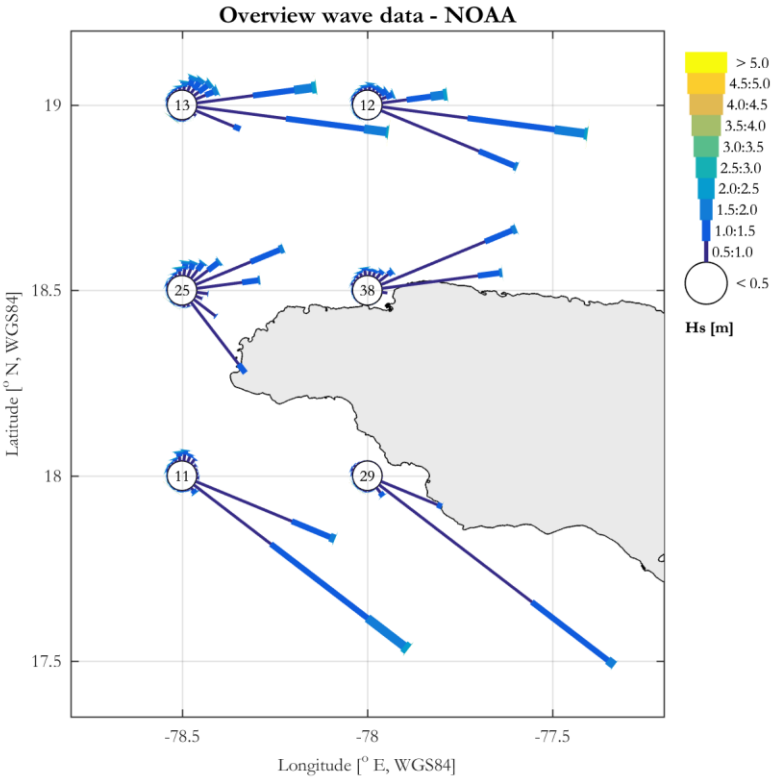

Figure 3.7 - Overview NOAA wave data (NOAA, 2015)

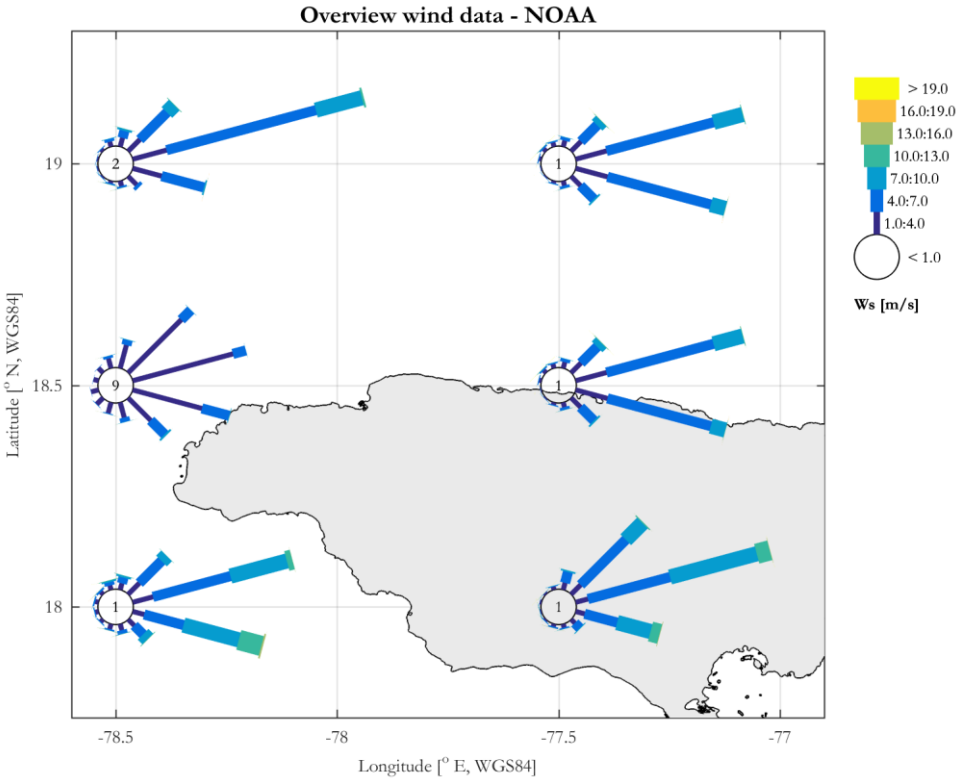

Figure 3.8 - Overview NOAA wind data (NOAA 2015)

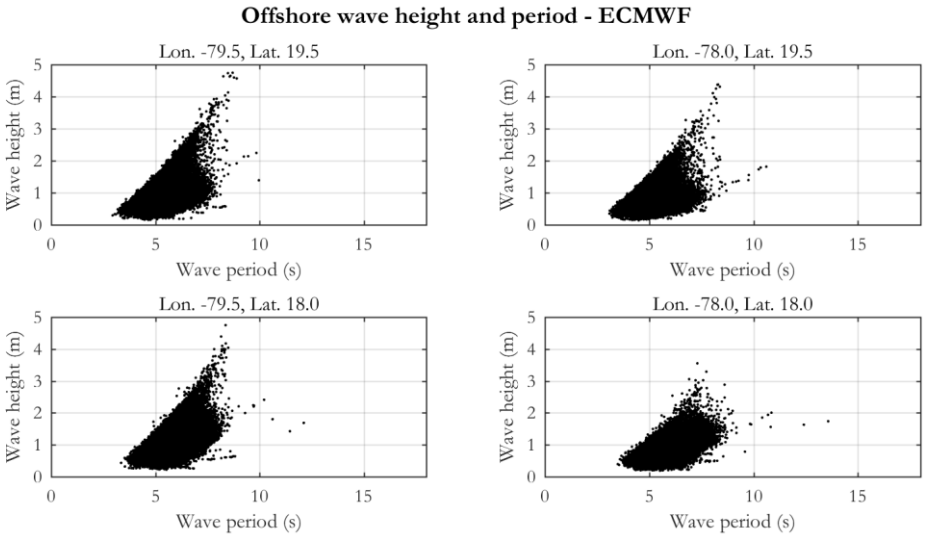

Figure 3.9 - Relation offshore wave height and period – ECMWF (ECMWF, 2015)

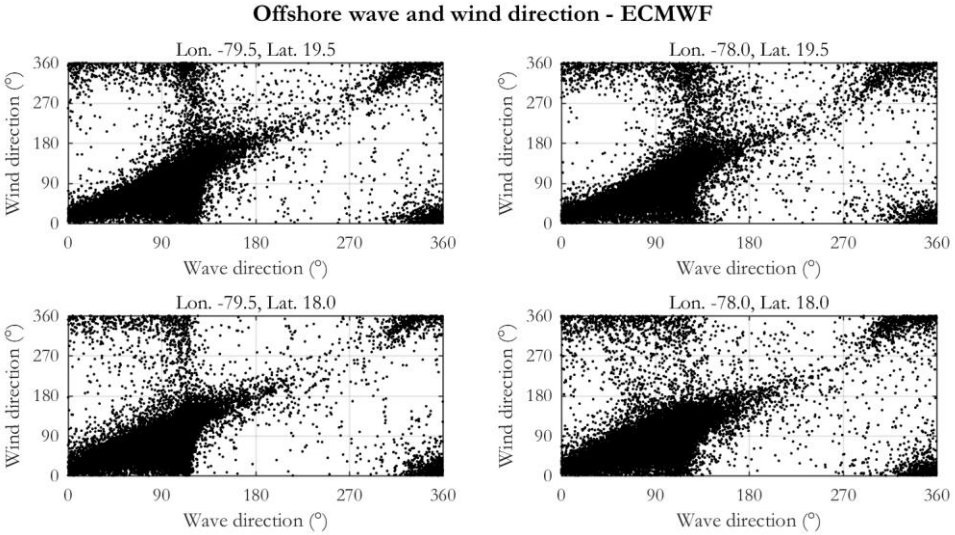

Figure 3.10 - Relation offshore wave and wind direction – ECMWF (ECMWF, 2015)

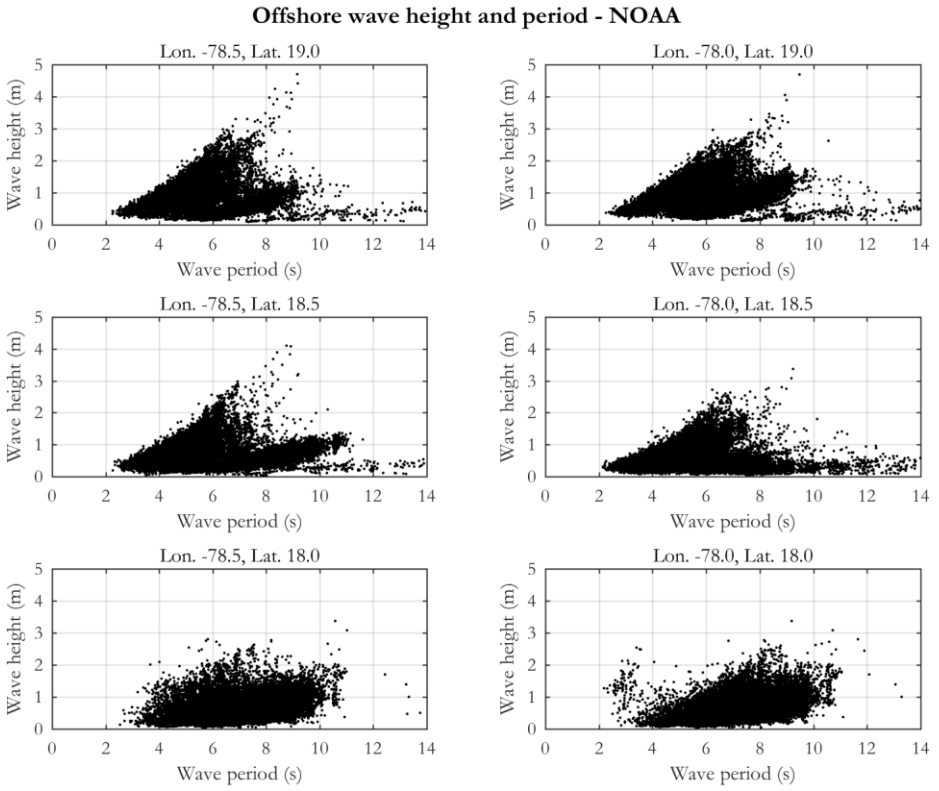

Figure 3.11 - Relation offshore wave height and period – NOAA (NOAA, 2015)

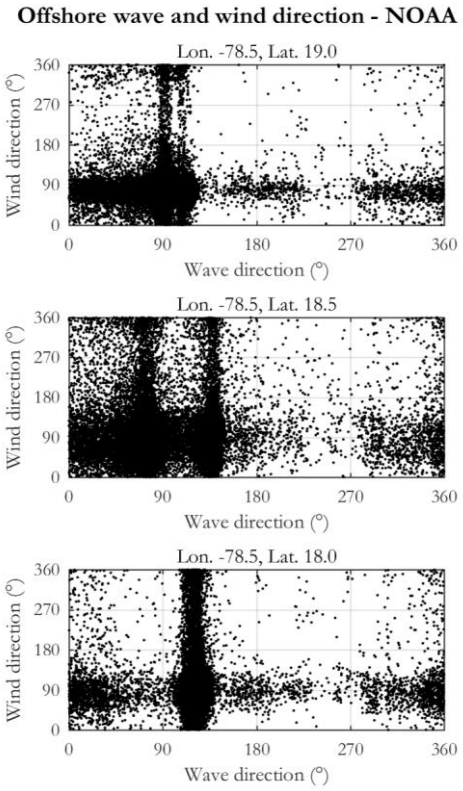

Figure 3.12 - Relation offshore wave and wind direction – NOAA (NOAA, 2015)

In Appendix B the only available wave measurements in the area are included. Smith Warner International Ltd. (2007) carried out wave measurements between October 20th and November 28th 2006 in a water depth of approximately 8 m, that is just offshore the coral reef. The largest wave height observed at this location was 2.91 m, measured at November 21st during a four-day storm event. The figures clearly indicate the storm event by much higher wave heights occurring in that period. The peak period is 8s, which implies that it the wave conditions consist of sea and not of swell. Furthermore, some historical wind data at Montego bay is available in Appendix B (The Weather Company, 2016).

3.2.5 Sediment characteristics

CL Environmental (2014, p. 131) conducted a local sediment analysis and compared the resulting values for the D50 with a preliminary design report by Smith Warner International Ltd. (2007). Both tests agreed on the poor grading of the sand, and resulted in similar values for the grain size for the sand on the beach face. The differences in values for the grain size on the back of the beach can be explained through the wave conditions during the tests. The sediment size found on the beach is smaller during calm periods compared to storm periods. The grain size of the sand and the beach face varies along the bay. The median grain size varies generally between 0.20 and 0.55 mm at most places, but outliers of 0.81 mm are present (Smith Warner International Ltd., 2007).

3.3 Modeling study

No in depth coastal modeling studies on the Negril Coast were carried out at the start of this study. Therefore this study focusses on obtaining an idea of the erosional processes at Negril beach. Since the erosion rates and the system dynamics at the Negril Coast are unknown, this study aims to address the following question:

- Is it plausible that the Negril Coast suffers from long-term erosion?
- What are important factors affecting the long-term shoreline development of the Negril Coast?

Bulk alongshore sediment transport is the net movement of sediment particles, parallel to the shoreline, in the entire active zone (Bosboom and Stive, 2013). Gradients in the bulk alongshore sediment transport rates often result in long-term changes in the coastline (Bosboom and Stive, 2013). Hence, the bulk alongshore sediment transport is an appropriate measure for the examination of the long-term shoreline development. The gradients in bulk alongshore sediment transport can be caused by changing any of the factors influencing the transport rate. Waves are expected to be the main driving factor of these changes at the Negril Coast. Therefore this study focusses on wave characteristics and ignores the minor tides and currents. To compute the bulk alongshore sediment transport rates the local nearshore wave characteristics are essential. Since no nearshore measurements are available to the project, offshore wave characteristics need to be transformed to nearshore.

Section 3.2 discussed that the available data contains many uncertainties. To answer the questions for this coastal modeling study it is necessary to incorporate these uncertainties. Therefore, a sensitivity analysis needs to be conducted. A sensitivity analysis in the case of Negril serves the purposes to obtain:

- Ranges of model output to getting an idea of possible bulk alongshore sediment transport rates.
- Insights in parameter importance to obtain a first idea of the system dynamics.

3.4 Uncertainties

Through expert judgment uncertain parameters that are expected to play a role for the bulk alongshore sediment transport rates at Negril Coast are indicated. The quantification of the uncertainties of these parameters is done through (i) data analysis for the parameters of which data is available, and (ii) a limited expert judgment process for the parameters of which no data is available. Table 3.1 provides an overview of the uncertainties involved, indicated by experts, the reasoning behind the uncertainties and the method of quantification. Appendix C discusses the expert judgment on uncertainties. It presents the questionnaire used, their results and the conclusions drawn. The uncertainty quantification through data analysis is explained below the table.

Table 3.1 - Overview of uncertainties involved

Uncertain parameter	Reason of uncertainty	Method uncertainty quantification
Bias of wave height	Non-validated global model	Expert judgment
Bias of wave period	Non-validated global model	Expert judgment
Bias of wave direction	Non-validated global model	Expert judgment
Bias of wind speed	Non-validated global model	Expert judgment
Bias of wind direction	Non-validated global model	Expert judgment
Uncertainty of coral reef height	Bathymetric survey containing errors Change over years	Expert judgment
D50 (sediment particle size)	Variation along the beach Lack of data	Data analysis
Surfzone slope	Variation along the beach Lack of data	Data analysis

Actual data from the D50 is available. This data is put into a histogram, through which a best probability density function is drawn. Figure 3.13 shows the results. See Appendix A for details on the Generalized Pareto distributions. For sensitivity methods which do not use probabilistic uncertainties the highest and lowest number of the D50 can be used. This will be clearly indicated in later Chapters.

It is very difficult to define the surfzone slope from the detailed bathymetry. The surfzone slope seems to vary roughly between 1/200 and 1/80. Therefore a uniform distribution with a = 1/200 and b = 1/80 is assumed. For non-probabilistic methods 1/200 and 1/80 can be used as extremes and 1/140 as the basic value. See Appendix A for details on probability distributions.

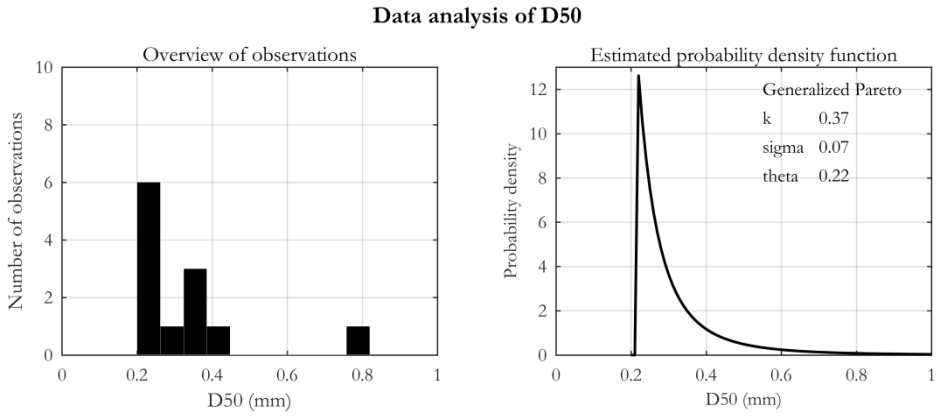

Figure 3.13 - Distribution of the D50

3.5 Concluding remarks

This chapter introduces the case Negril Coast. The following points are important for conducting the several sensitivity analyses for Negril Coast:

- The aim of the modeling study at Negril Coast is to address the following questions: (i) 'Is it plausible that the Negril Coast suffers from long-term erosion?', and (ii) 'What are important factors affecting the long-term shoreline development of the Negril Coast?'.

- Gradients in the bulk alongshore sediment transport rates often result in long-term changes in the coastline. Therefore, bulk alongshore sediment transport rates will be used as a measure for long-term shoreline evolution. For this, a modeling execution that *transforms offshore wave characteristics to nearshore* is needed.

- Because the available data contains many uncertainties a sensitivity analysis needs to be conducted. This sensitivity analysis serves the purposes to obtain: (i) a range of model outputs to get an idea of possible bulk alongshore sediment transport rates, and (ii) insights in parameter importance to obtain a first idea of the system dynamics.

- Experts indicated uncertainties that need to be involved in the sensitivity analysis. The uncertainties are presented in Table 3.1 at page 25. Exact quantification can be found in Appendix C.

- The quantification of most parameters is done through expert judgment. From this, it is concluded that it is very important to construct *clear questionnaires* in order to map the expert judgment correctly. A clear questionnaire is characterized by:

    - The providence of proper information on the parameters.
    - Clear and case-specific examples
    - Eventually, questions referring to percentages of particular parameter values, instead of referring to fixed numbers, in order to indicate the uncertainty of that particular parameter.

4. Introduction and setup of the models for the Negril Coast

Chapter 3 indicates that a model for wave transformation is needed for the modeling study. This chapter introduces two wave transformation models and discusses their set-up. Furthermore, the computation of bulk alongshore sediment transport rates, using the nearshore wave characteristics, is described. Section 4.1 introduces the *complex* model: The wave energy balance. Section 4.2 introduces the *simple* model: The linear wave theory. The concluding remarks are presented in Section 4.3.

4.1 The wave energy balance

The complex model uses the wave energy balance to transform offshore wave characteristics to nearshore. These nearshore wave characteristics are used to compute bulk alongshore sediment transport rates. The following sections present the computational steps, the model set-up and the model verification.

4.1.1 Computational steps

Figure 4.1 shows a schematic representation of the working of the complex model, consisting of the following four steps:

- Transformation of offshore wave characteristics to nearshore using the wave energy balance.
- Computation of bulk alongshore sediment transport rates.
- Computation of coastline evolution.
- Computation of the new shoreline orientation and subsequently the new transport rates.

The last three steps can be repeated as many times as desired, investigating the shoreline evolution in time. The next sections elaborate each on one step.

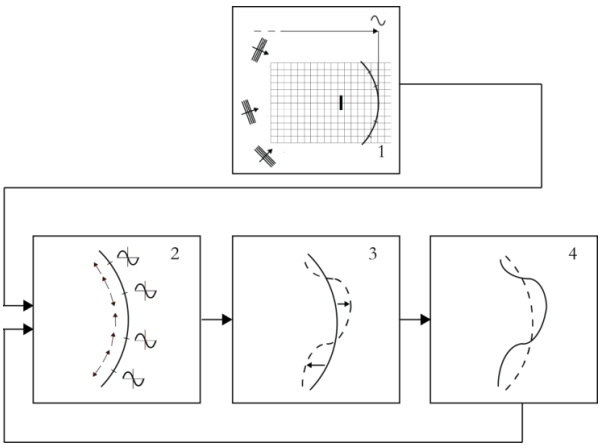

Figure 4.1- Schematic representation of the complex model

Transformation

A *Delft3D-WAVE model* is used to transform offshore wave characteristics to nearshore. The Delft3D-WAVE module is based on the spectral SWAN (Simulating WAves Nearshore) model (Holthuijsen, 2007) This is a commonly used 3rd generation numerical wave model, taking into account all the processes of propagation, generation, wave-wave interactions and dissipation (SWAN, 2015, Holthuijsen, 2007, Deltares, 2014a). The nonlinear character of these processes requires an Eulerian approach. Therefore, a spatial grid, accompanied by a bathymetry, needs to be implemented. Nesting can be used to allow local refinement of the grid at locations of interest (SWAN, 2015). At the boundaries of the grid wave climates can be imposed. Additionally a wind field can be imposed on a second grid. The offshore waves are translated to nearshore wave conditions through solving the wave action balance on the computational grid. In the

absent of an ambient current the action balance equation reduces to the *energy balance equation*. The energy balance equation reads as follows (Holthuijsen, 2007):

$$\frac{\partial E(f,\theta;x,y,t)}{\partial t}+\frac{\partial c_{g,x}(f,\theta;x,y,t)}{\partial x}+\frac{\partial c_{g,y}(f,\theta;x,y,t)}{\partial y}+\frac{\partial c_{\theta}(f,\theta;x,y,t)}{\partial \theta}=S(f,\theta;x,y,t) \quad 4.10$$

With E the wave energy, c_g the group celerity and c_Θ the directional turning rate, and with f, x, y, Θ indicating the frequency, the orientation of the computation and the direction respectively. S is the source term. The source term represents the effects of generation, nonlinear wave-wave interactions and dissipation, through processes as wind, white-capping, bottom friction and depth-induced breaking.

Bulk alongshore sediment transport rates

The model calculates averaged bulk alongshore sediment transport rates using all the nearshore wave characteristics computed. For this linearized (S,φ)-curves are computed along multiple locations at the coast, using the Kamphuis sediment transport formulation. This formulation reads as follows (Bosboom and Stive, 2013):

$$S = 2.27 * H_{s,b}^2 * T_p^{1.5} * (\tan\alpha_b)^{0.75} * D^{-0.25} * (\sin 2\varphi_b)^{0.6} \quad 4.9$$

With $H_{s,b}$ the wave height at the breaker point [m], T_p the peak period [s], α_b surfzone slope [-], D the median sediment grain size d_{50} [m] and φ_b the wave angle at the breaker point [°].

Coastline evolution

The model defines grid cells along the coast. For each of these cells the sediment balance is computed. Subsequently, the appurtenant coastline evolution is computed. Consequently, the sediment transport rates for the new coastline orientation can be calculated on the basis of the already defined (S,φ)-curves. This can be repeated as many times as desired.

4.1.2 Model set-up

Two Delft3D-WAVES models are set-up to transform offshore (i) ECMWF and (ii) NOAA wave characteristics. This section presents the set-up of both models.

Model area and grids

The computational domains used are presented in Figure 4.2. The boundaries of the overall domain correspond with the locations where the wave information is provided. The model for the ECMWF wave applies four computational grids:

- Domain I – resolution 1000 * 1000 m
- Domain II - resolution 500 * 500 m (refinement factor 2)
- Domain III – resolution 100 * 100 m (refinement factor 5)
- Domain IV – resolution 25 * 25 m (refinement factor 4)

The model for the NOAA waves applies three computational grids:

- Domain I - resolution 400 * 400 m
- Domain II – resolution 100 * 100 m (refinement factor 4)
- Domain III – resolution 25 * 25 m (refinement factor 4)

A *spatially varying wind field* is applied in both models. For this, a grid of four and six cells are used for ECMWF and NOAA respectively.

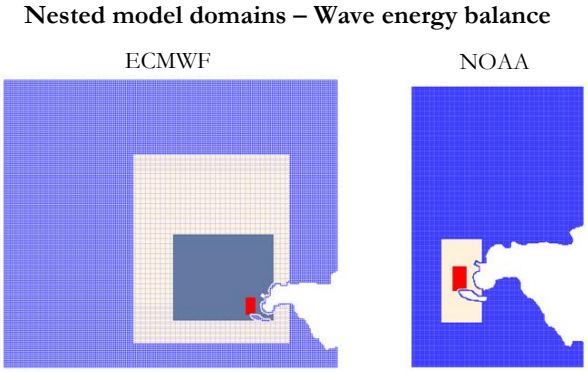

Figure 4.2 - Nested wave model domains – wave energy balance

Bathymetry

The bathymetry as presented in Section 3.2 is used. For both the ECMWF and NOAA data both the GEBCO and the detailed bathymetry are incorporated. This results in four different models. The nearshore wave characteristics are selected from the model results *at a depth of five meters*. This is *behind* the coral reef.

Wave and wind boundary conditions

The wave and wind data described in Section 3.2.4 is used as the basis for the schematization of the normal boundary conditions. The NOAA waves at Lat. -78.0 Lon. 18.5 are located too close to the shore and are therefore not used as a boundary condition. The boundary locations for the wave model are indicated in Figure 4.2. The wave and wind data is *classified* in order to define a set of representative offshore wave and wind conditions, called the *full wave climate*. The *reference location* that used is for ECMWF is Lat. 795. Lon. 18 and for NOAA Lat. 785.5 and Lon. 18.5. These are the locations closest to the Negril Coast. To distinguish sea and swell and since wave, and wind direction do not always degree, the classification is carried out on the basis on the following parameters and class boundaries:

- Wave height 0 – 0.25 – 0.5 – 1 – 1.5 – 2 – 2.5 – inf. (m)
- Wave direction 0 : 30 : 360 (°)
- Wave period 0 – 10 – inf. (s)
- Wind direction 0 : 180 : 360 (°)

This results in a total of 132 classes for the ECMWF data and 160 classes for the NOAA data. For each class, the representative wave and wind condition are determined based on averaging all condition with the class. Finally weights are assigned to the classes. The wave and wind scenario that is applied as boundary conditions for the wave model is constructed by *simultaneous occurrence in the other boundary points*. The spatial variation in wave conditions along the boundaries is accounted for in the model.

Model settings

The following parameter settings are applied in the wave model:

Wave spectrum

- At the model boundary a JONSWAP spectrum with peak enhancement factor of 2.2 is assumed.
- At the model boundary a directional spreading is specified based on the relation with the wave period. The largest spread is used for short period waves (cosine power 2 for wave periods 0-12 s) and the smallest spread for long period waves (cosine power 7 for wave periods of 12 s or higher). This relationship is based on information provided in SWAN (2015).

Physical parameters

- Third-generation mode for wind growth and white-capping, both according to Komen et al. (1984), and quadruplet interactions are incorporated

- Constant depth-induced breaking ($\alpha = 1$ and $\gamma = 0.73$)
- JONSWAP bottom friction with a friction coefficients of 0.038 m^2/s^3
- No non-linear triad interactions and no diffraction

Numerical parameters

- Maximum number of iterations: 50
- Percentage of wet points: 99%
- Relative change H_s-T_{m01} = 0.02
- Relative change H_s with respect to mean = 0.02
- Relative change T_{m01} with respect to mean = 0.02

Other parameters

- D50 = 0.28 mm
- Surfzone slope = 1/140

With these settings it takes around 7-20 hours for one model execution consisting of 132 scenarios. The time is dependent on the characteristics of the boundary conditions used.

4.1.3 Model verification

The model working is verified in two ways: (i) a general verification of modeling results and (ii) a comparison of the model results and the measurements (see Section 3.2.4 and Appendix B).

General

Figure 4.3 indicates the model results of two of the 160 ECMWF scenarios modeled. Comparable plots are made for every ECMWF and NOAA condition in order to verify the wave modeling results. The processes of refraction, wave direction do to bottom friction and breaking and wave focusing over offshore shallow areas can be observed. On the basis of inspection of the model results it is concluded that the processes and boundary conditions are incorporated correctly in the model.

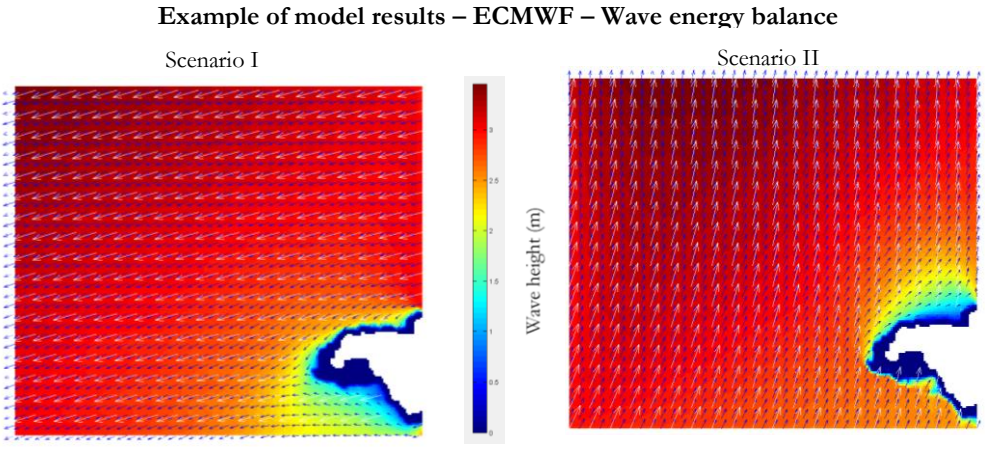

Figure 4.3 - Example of computed wave height and directions for 2 scenarios

Figure 4.4 shows the sediment transport rates resulting from both ECMWF and NOAA wave data combined with both the detailed and a GEBCO bathymetry. The graphs show that the GEBCO bathymetry results in lager transport rates. This is because the nearshore waves are higher and less refracted when the GEBCO bathymetry than is used. This is the case because the GEBO bathymetry is very deep close to the shore. Hence, compared to the detailed bathymetry, the waves are affected less by

the bottom. The detailed bathymetry is considered to be better and therefore worked further with. Furthermore Figure 4.4 shows that NOAA results in slightly larger transport rates. Considering the nearshore wave roses this is mainly because overall higher waves reach the coast when NOAA data, instead of ECMWF data, is used. The difference between NOAA and ECMWF is investigated further in the next section.

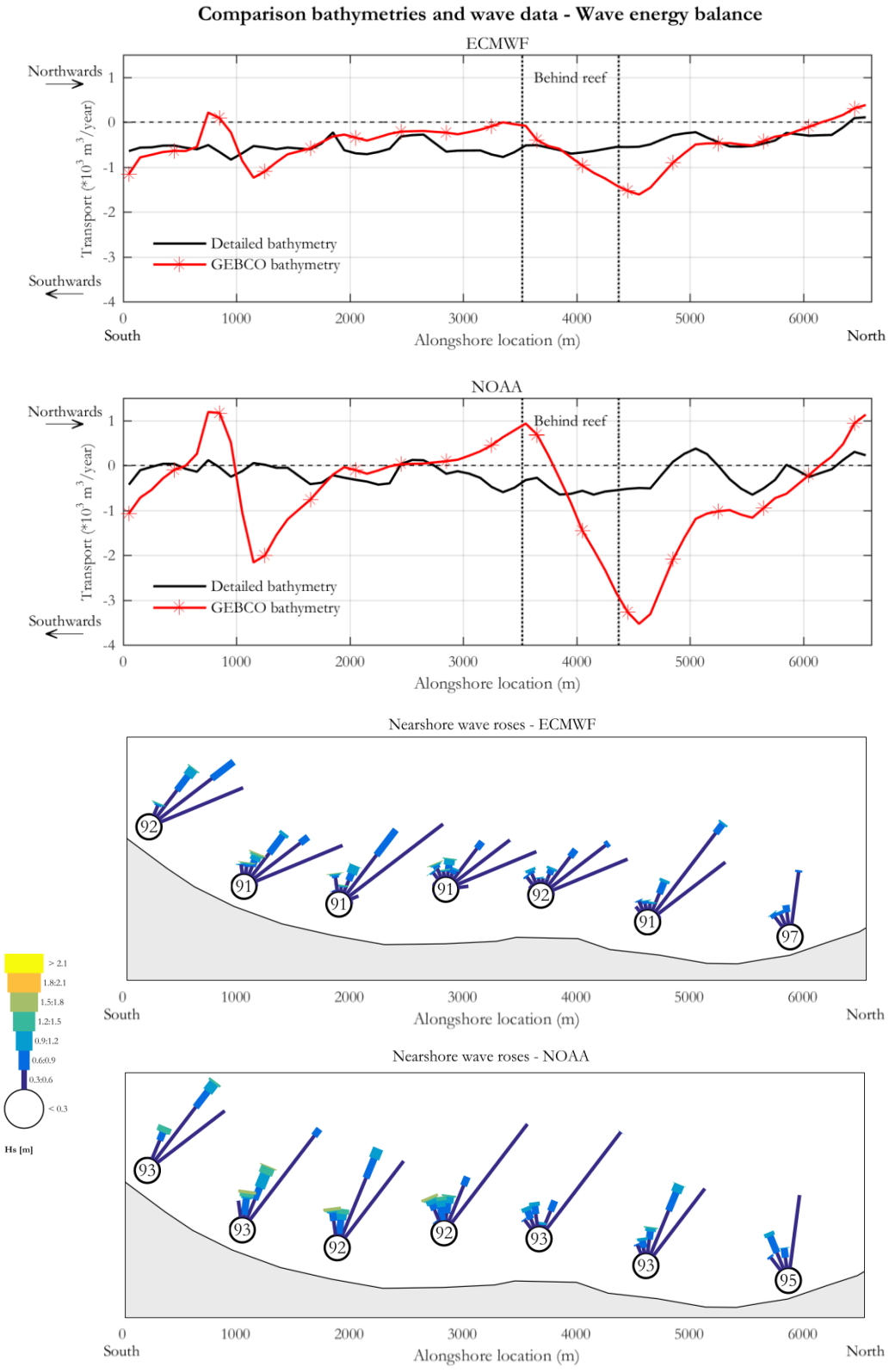

Figure 4.4 – Comparison wave data and bathymetries - the wave energy balance

Comparison with measurements

To assess the quality of the wave transformation, model results are compared to the available nearshore wave measurements. For this the period during which measurements have taken place is modeled using both NOAA and ECMWF time series. Figure 4.6 presents the modeling results, which are compared to the available measurement data Figure 4.5. The peak of the wave heights during the storm event can be clearly indicated by the ECMWF modeling results, contrarily to the NOAA modeling results. This is in line with the research of Appendini et al. (2013), who state that ECMWF wave data is better than the NOAA wave data in the Caribbean sea. Therefore the ECMWF dataset is considered to provide *more accurate results and is selected to work further with.* More modeling results of the time series can be found in Appendix B.

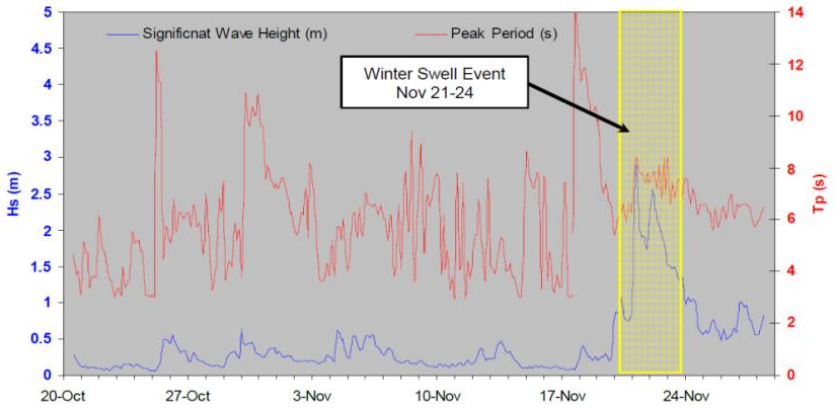

Figure 4.5 – Available wave measurements (Smith Warner International Ltd., 2007)

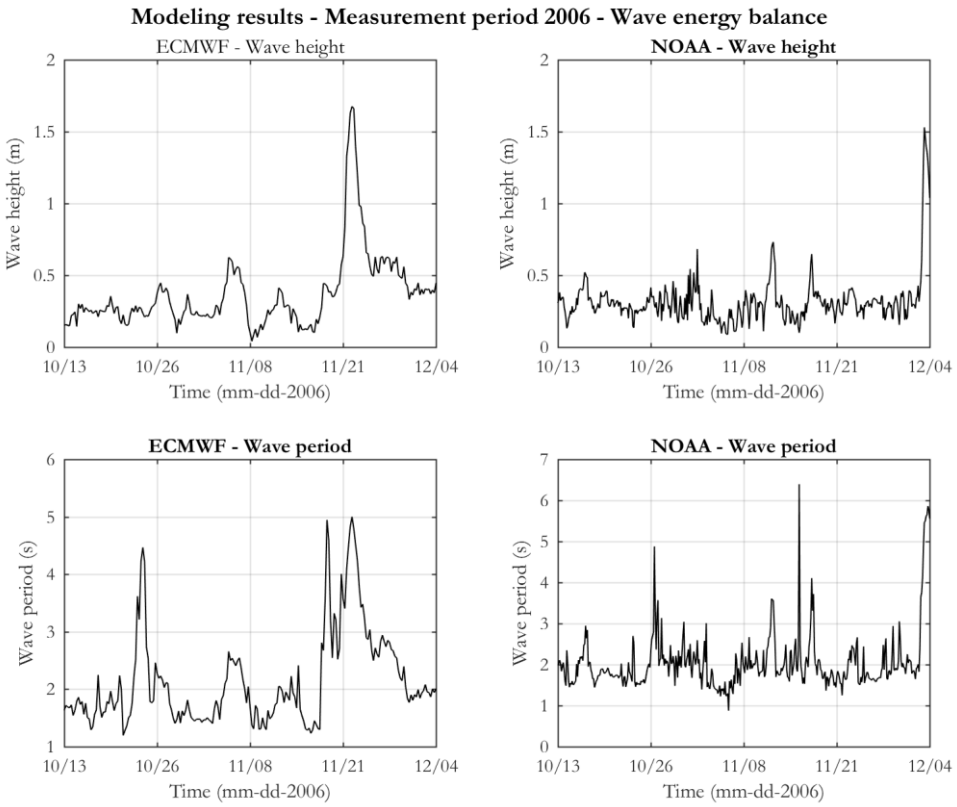

Figure 4.6 - Resulting nearshore wave height and period of modeling the measurement period

4.2 The linear wave theory

This model uses the linear wave theory to transform offshore wave characteristics to nearshore. The following sections present the computational steps, the model set-up and the model verification.

4.2.1 Computational steps

Figure 4.7 shows a schematic interpretation of the functioning of the model. The figure distinguishes five steps of which of which only the first two are different compared with the complex model. The next sections elaborated on these first steps:

- Transformation of offshore wave characteristics to nearshore
- Incorporation the interaction with structures

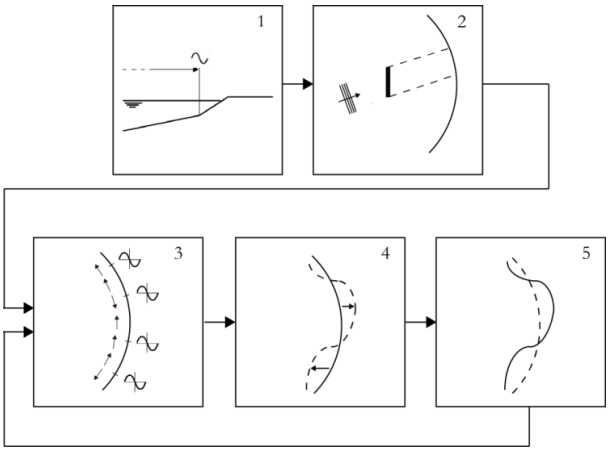

Figure 4.7 - Schematic representation of the simple model

Transformation

The model assumes an alongshore uniform coast, with only one cross-shore profile consisting of a slope. Offshore waves, defined by wave heights, directions, periods and weights, are used as the wave input at the offshore side of the slope. To compute the nearshore wave characteristics, first the offshore and nearshore wave length are calculated through the dispersion relationship (Holthuijsen, 2007):

$$L = \frac{gT^2}{2\pi} \tanh\left(\frac{2\pi d}{L}\right) \quad 4.1$$

With g the gravitational constant [m/s²], T the wave period [s], d the water depth [m] and L the wave length [m]. Since the hyperbolic tangent approaches 1 in deep water, the deep water depth does not have to be defined.

Consequently, the wave celerity is calculated through (Holthuijsen, 2007):

$$c = \frac{L}{T} \quad 4.2$$

With c the wave celerity [m/s], L the wave length [m] and T the wave period [s].

Subsequently the wave angle at shallow water is calculated through Snell's law (Bosboom and Stive, 2013):

$$\frac{\sin \varphi_2}{c_2} = \frac{\sin \varphi_1}{c_1} \quad 4.3$$

With c the wave celerity [m/s], φ the wave angle [°] and where the subscripts 1 and 2 indicate the location at which the parameters are evaluated, at deep and shallow water respectively.

Effects on wave height due to shoaling and refraction are incorporated through the following formulation (Bosboom and Stive, 2013):

$$\frac{H_2}{H_1} = K_{sh}K_r = \sqrt{\frac{c_1\, n_1}{c_2\, n_2}}\sqrt{\frac{\cos\varphi_1}{\cos\varphi_2}} \quad 4.4$$

With H the wave height [m], K_{sh} the shoaling factor [-] and K_r the refraction factor [-], c the wave celerity [m/s], n the ratio of the group celerity and the individual wave celerity [-] and φ the wave angle [°].Again the subscripts 1 and 2 indicate the location at which the parameters are evaluated, at deep and shallow water respectively. The group celerity is defined by (Bosboom and Stive, 2013):

$$c_g = n * c \quad 4.5$$

With n the ratio (Bosboom and Stive, 2013):

$$n = \frac{1}{2} + \frac{kd}{\sinh(2kd)} \quad 4.6$$

With k [m^{-1}] the wave number as follows (Bosboom and Stive, 2013):

$$k = \frac{2\,\pi}{L} \quad 4.7$$

With L the wave length [m].

Finally the depth-induced wave breaking is incorporated. When the wave's height becomes greater than a certain fraction of the depth, the waves break and their heights decrease abruptly. This fraction is described by the breaking index, defined by (Bosboom and Stive, 2013):

$$\gamma = \left[\frac{H}{d}\right]_{max} = \frac{H_b}{d_b} \approx 0.78 - 0.88 \quad 4.8$$

With H the wave height, H_b the wave height at the point of breaking [m], d the still water depth and d_b the still water depth at point of breaking [m].

After the computation of the nearshore wave climate, a selection of relevant offshore waves can be made. Relevant offshore waves are the ones of which it is expected that they actual reach the coastline. The offshore waves traveling in other directions adopt a zero value for all their characteristics.

Interaction with structures

After the nearshore wave characteristics are computed, their interaction with structures is incorporated through the computation of sheltering effects and diffraction. For this the diffraction coefficients of all the waves are calculated following Kamphuis (1992). Consequently the wave characteristics of the diffracted waves are computed. Important to note is that the structures' height cannot be implemented; it is assumed that structures reach above sea-level.

4.2.2 Model set-up

Model settings

Table 4.1 presents the model settings of the simple model for the base case. Since the coast is western orientated and the coast is assumed to be alongshore uniform, the relevant waves have a direction between 225° and 345°. The nearshore water depth is defined at the shallowest depth at which the waves do not break. Based on a breaker index of 0.78-0.88 this is set at 5 m. Only waves above 3.9-4.4 m break at a deeper depth. It is not expect that this nearshore wave height occurs very often. Besides the settings displayed, the number of grid cells defined in the model is important. 68 grid cells, of which the evolution can be computed, are implemented along the coastline. Based on maps from Google Earth (Google Earth, 2015), the coral reef is implemented at the correct location.

Table 4.1 – Model settings base case, linear wave theory

Setting	Values
Selection of relevant waves	225° - 345°
Slope	1/140
Nearshore water depth	5 m
D50	0.28 mm

Wave boundary conditions

The model demands ± 20 seconds to transform a set of 100-1000 waves. Therefore time series are used as boundary conditions. The relevant waves have a direction of 225° - 345°. Because all the other waves get a zero value, only waves with the relevant direction are used as model input. Figure 4.8 and Figure 4.9 show the selection of waves for ECMWF and NOAA respectively. As can be seen only a small fraction of the waves reaches the coast: For ECMWF 1.5% and for NOAA 3.6%.

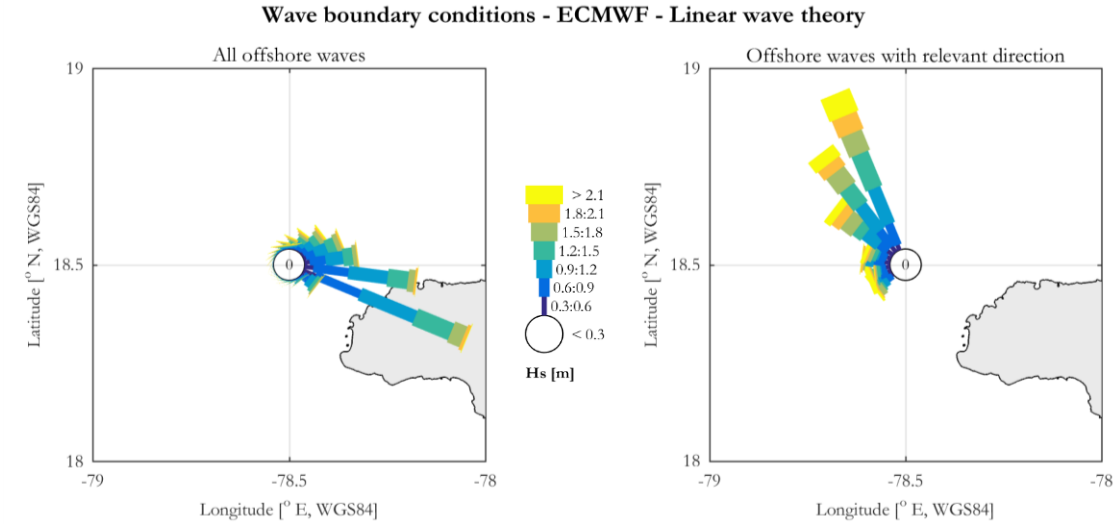

Figure 4.8 - Boundary conditions linear wave theory – ECMWF

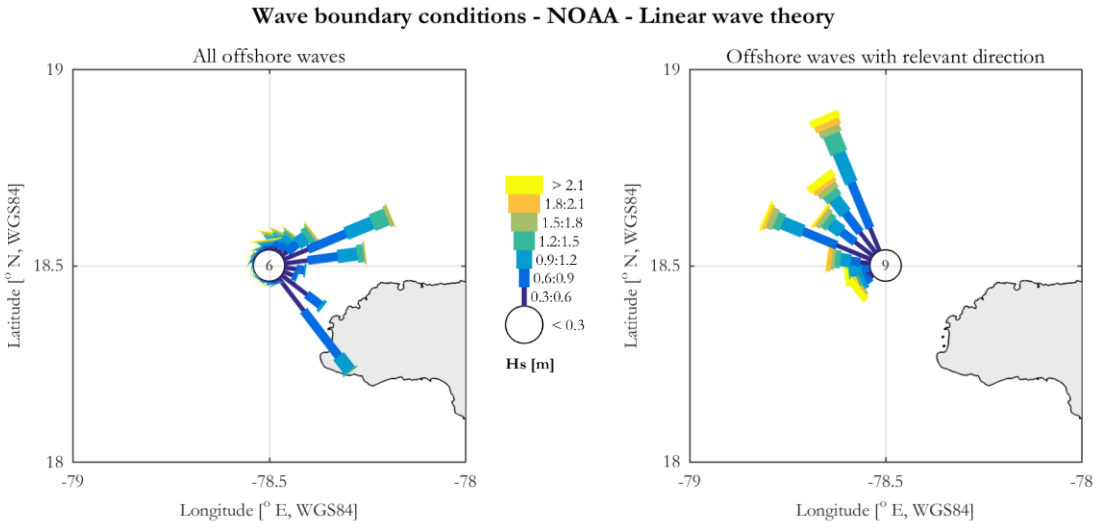

Figure 4.9 - Boundary conditions linear wave theory - NOAA

4.2.3 Model verification

Figure 4.10 shows the nearshore waves roses and bulk alongshore sediment transport rates, resulting from the model executions with the base case using both ECMWF and NOAA wave data. As can be seen the wave roses along the shore are pretty similar. This is because the model assumes one constant cross-shore profile. The influence of the coral reef can be seen from both the transports and the roses.

The wave roses resulting from the two wave datasets do not differ very much: More ECMWF waves approach the coast from the North, however these are mostly small waves. The NOAA waves result in larger transport rates than the ECMWF waves. Based on the wave roses, this is probably not because of the local wave characteristics. This is because only 1.5% of the ECMWF wave reaches the coast, and 3.6% of the NOAA waves. This is a difference of factor two, which is more or less the factor in sediment transport.

It is expected that the wave energy balance provide more accurate results compared to the linear wave theory. Therefore, a comparison with the measurements is not made.

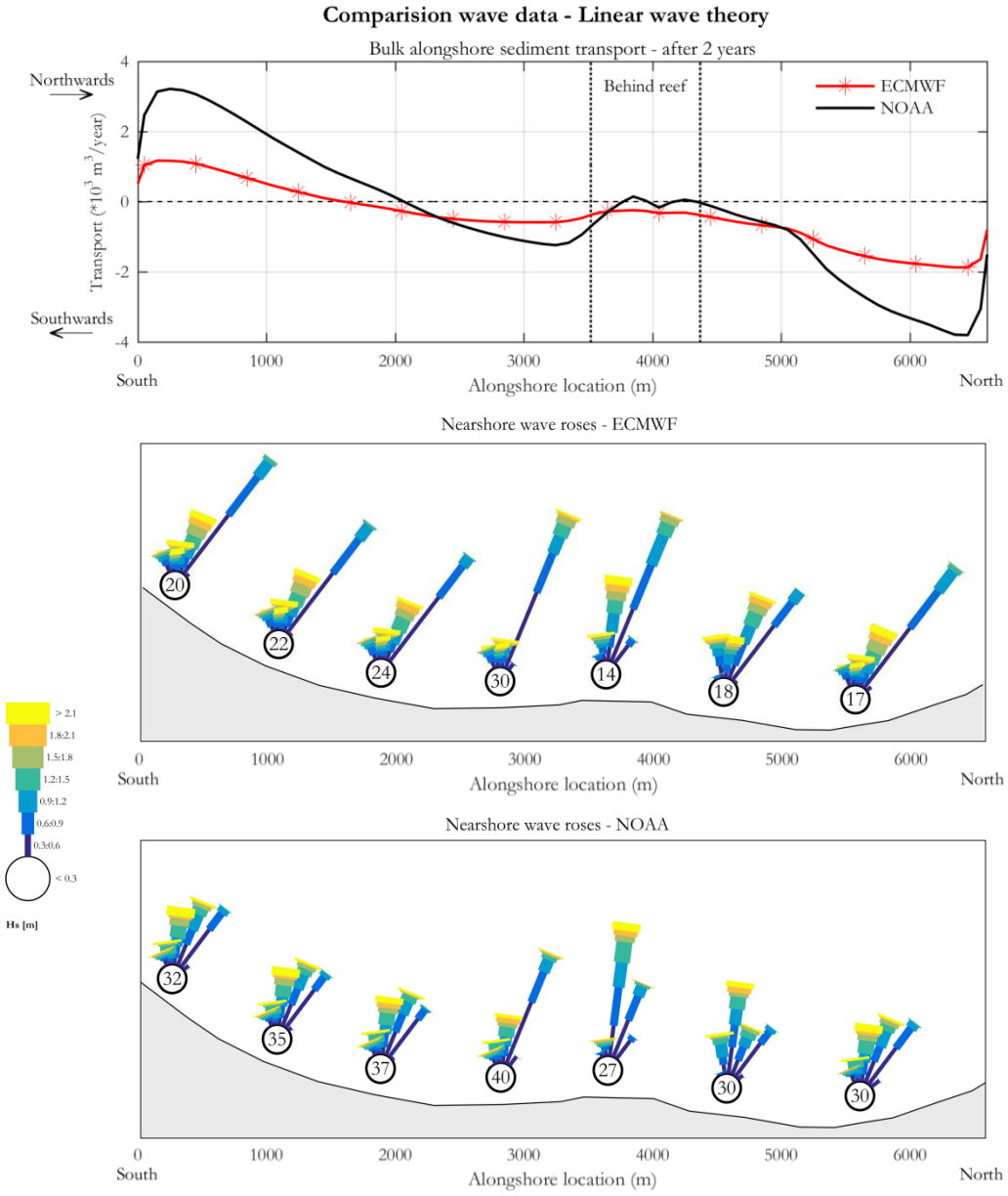

Figure 4.10 - Bulk alongshore sediment transport rates resulting from base cases - Model I

4.3 Concluding remarks

This chapter discusses the set-up of the two models used to (i) transform offshore wave characteristics to nearshore and (ii) compute the bulk alongshore sediment transport rates along Negril Coast. The following points are important for conducting the several sensitivity analyses:

- Table 4.2 provides an overview of the differences between the two models.

Table 4.2 - Origin of the differences between the models

Origin of difference	**Linear wave theory**	**Wave energy balance**
Size of wave input	Typically ~ 1000 – 10000 waves (time series)	Typically ~ 100 waves (wave climates)
Wave input	Waves at 1 location Only onshore directed waves considered	Varying offshore wave boundaries All wave directions considered
Wind input	No wind considered	Varying wind fields
Bathymetry	Alongshore uniform slope, with one constant cross-shore profile	Bathymetry
Coral reef	No reef height considered	Reef height considered
Processes	Only linear processes included	Both (non-)linear processes included
Computational time	Around 20 seconds	Around 7-20 hours

- The detailed bathymetry is expected to provide better results than the GEBCO bathymetry. Therefore the detailed bathymetry is used in the further research.

- ECMWF seems to provide better results than NOAA. This is based on a comparison between measurements and modeled nearshore wave characteristics for both NOAA and ECMWF waves. The difference between the modeling results and the measured data are big. However, they are considerably smaller when the ECMWF wave dataset is used. Therefore the ECMWF dataset is used for conducting the sensitivity analyses.

5. The benchmark - An extensive sensitivity analysis

A benchmark is established and used to evaluate the performance of all the sensitivity analyses conducted. Section 5.1 explains how the benchmark is established. Section 5.2 presents the results of the sensitivity analysis conducted, i.e. the benchmark. Finally, section 5.3 discusses the most important findings.

5.1 The benchmark

A benchmark is established through conducting a sensitivity analysis that generates results as accurate as possible within the time constraints of this research. Moreover, it is considered that a more extensive sensitivity analysis would only demand more time without a significant increase in accuracy. The sensitivity analysis that is used to establish the benchmark is characterized by the use of:

- The wave energy balance to transform offshore wave characteristics to nearshore. The wave energy balance involves more aspects and processes than the linear wave theory (Chapter 4)

- The Crude Monte Carlo method. It is expected that probabilistic methods result in more accurate insights (see Chapter 2). The sample size of a probabilistic method cannot be defined before the analysis. Hence, several sample sizes need to be investigated in order to obtain accurate insights. It is easier to increase the sample size of the Crude Monte Carlo method, compared to the Latin hypercube method. Therefore the Crude Monte Carlo method is used to establish the benchmark. Finally 350 samples are used.

- A full wave input climate (Section 4.2.2).

The uncertainties, and their dependencies, involved in the sensitivity analysis are indicated and quantified through expert judgment (Section 3.4). Table 5.1 and Table 5.2 provide an overview of the uncertainties and their correlations. For the dependencies, Gaussian copulas are assumed. See appendix A for details on distributions.

Table 5.1 – Distributions of uncertainties involved

Parameter	Distribution
Bias of wave height (m)	Generalized Pareto (-0.87, 2.62, -1.5)
Bias of wave period (s)	Generalized Pareto (-1.07, 6.42, -2)
Bias of wave direction (°)	Generalized Pareto (-1.12, 33.66, -15)
Bias of wind speed (m/s)	Generalized Pareto (-0.83, 0.71, -0.80)
Bias of wind direction (°)	Generalized Pareto (-1.12, 33.66, -15)
Bias of coral reef height (m)	Generalized Pareto (-0.95, 1.62, -0.70)
D50 (mm)	Generalized Pareto (0.37, 0.07, 0.22)
Surfzone slope (-)	Uniform (1/200, 1/80)

Table 5.2 – Correlations between the uncertainties involved

	Bias of wave height (m)	Bias of wave period (s)	Bias of wind speed (m/s)	Bias of wave direction (°)	Bias of wind direction (°)
Bias of wave height (m)	1	0.85	0.85	-	-
Bias of wave period (s)	0.85	1	0.83	-	-
Bias of wind speed (m/s)	0.85	0.83	1	-	-
Bias of wave direction (°)	-	-	-	1	0.8
Bias of wind direction (°)	-	-	-	0.8	1

5.2 Results

Conducting the extensive sensitivity analysis takes 2450 hours with one computer with a 3.60 GHz CPU using one core. The following sections present the results of the extensive sensitivity analysis, i.e. the benchmark. The results are divided in three parts. Each part represents an important characteristic of the benchmark: (i) the range of model outputs, (ii) sensitivities (scatter plots and correlation coefficients), and (iii) an estimation of the dependence between an uncertain input parameter and the model output.

5.2.1 Probabilistic model output ranges

Figure 5.1 presents the model output ranges belonging to the benchmark. Additionally, it displays the deterministic model output. From the figure the following phenomena are observed:

- The transport rates never exceed 9,000 m^3/year. The median sediment transport rate never exceeds 1,000 m^3/year. To put these results into context: The transport rates along the central Dutch coast are around 125,000 m^3/year (van Rijn, 1997). Hence, the transport rates at Negril Coast can be considered as very small.
- The median model output and the deterministic model output are very similar.
- The maximum and minimum model outputs differ significantly from each other and from the deterministic model output. This indicates the usefulness of conducting a sensitivity analysis
- The maximum and the minimum model output indicate opposite direction of the sediment transport. This may be very important with regard to the obtained insights in the system dynamics.

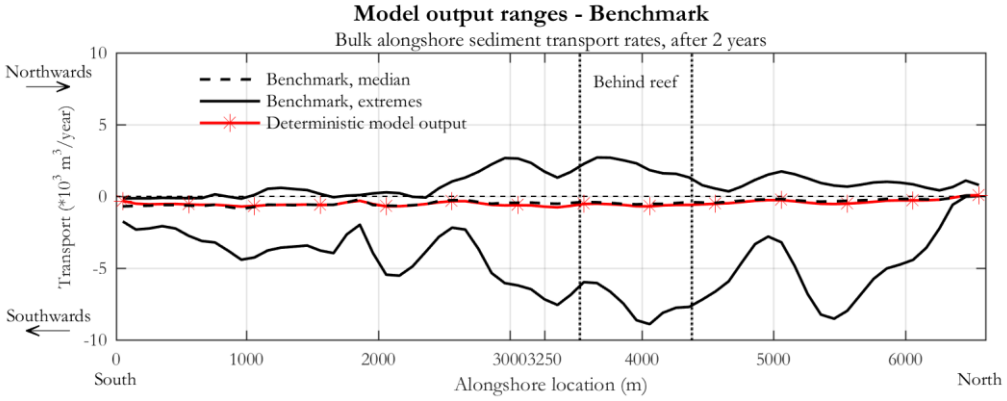

Figure 5.1 – Model output ranges - Benchmark

Important to note is that the median and the extremes of the output ranges do not necessarily represent the results of particular model executions. Hence, drawing conclusions from these lines must be done very carefully. It is possible to find results of the model execution that approaches the median and the extremes best. Figure 5.3 provides the best approximations of the median and maximum sediment transport rates, together with their corresponding nearshore wave roses. Table 5.3 presents the values of the parameters used for the selected model executions.

Table 5.3 – Parameter values of best approximations - Benchmark

Execution	Change of						D50	Surfzone slope
	Wave height	Wave period	Wave direction	Wind speed	Wind direction	Reef height		
Median	0.3 m	-1.6 s	-5.2°	0.5 m/s	-5.1°	0.1 m	0.31 mm	± 1/123
Maximum	1.2 m	0.0	14.4°	0.7 m/s	12.3°	-0.2 m	0.25 mm	± 1/85

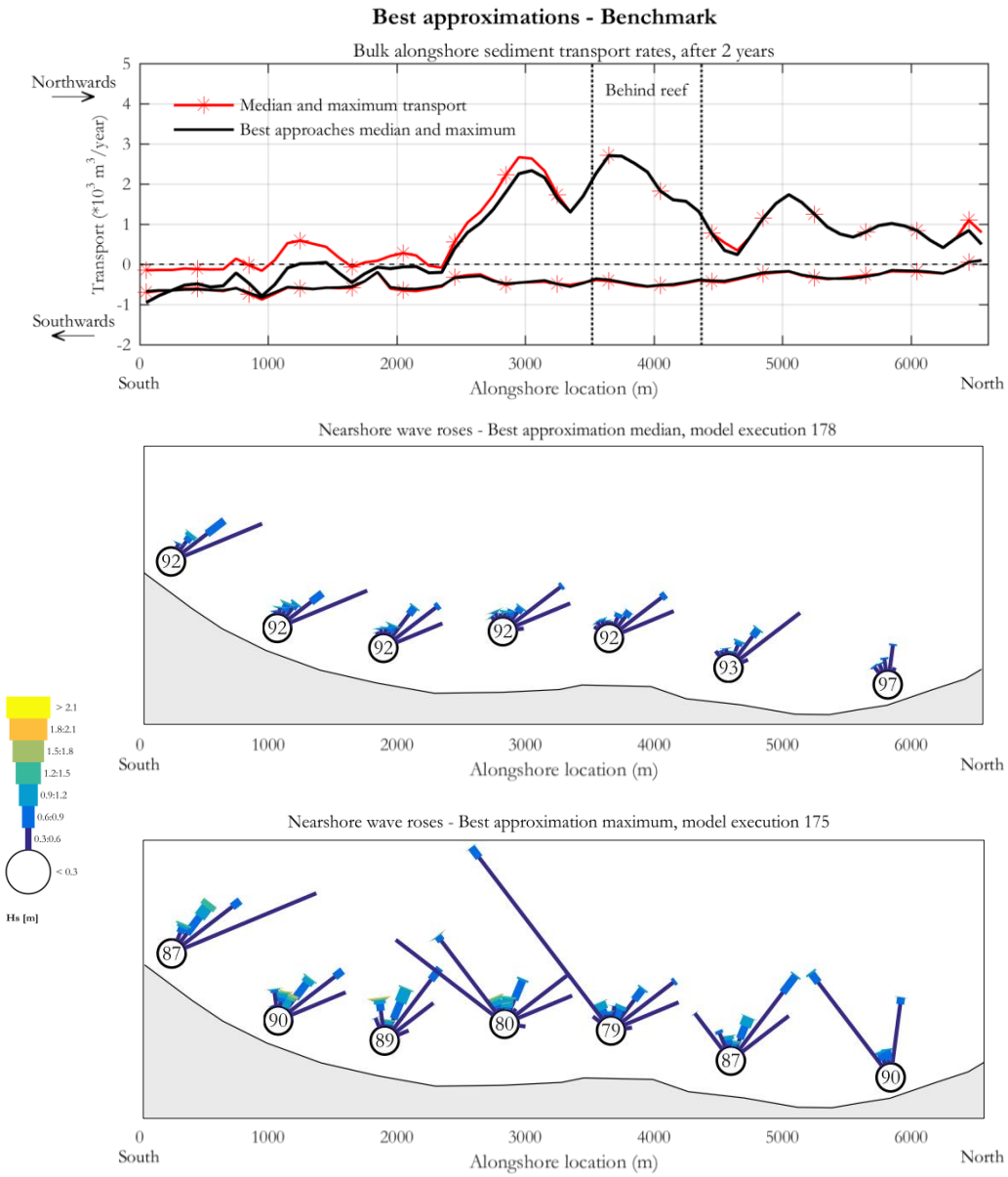

Figure 5.2 – Best approximations – Benchmark

From the figure and the table, the following phenomena are observed:

- The best approximation of the median is very accurate and indicates minor sediment transport rates mostly in southward direction. The alongshore gradients of these transport rates are marginal. Consequently, hardly any erosion or sedimentation is expected.

- The best approximation of the maximum is accurate, only in the South of the bay the line shows some deviation from the maximum model output. This indicates *alongshore varying sensitivities*. The approximation of the maximum indicates southward directed sediment transport in the South and northward directed transport in the North of the bay. It is expected that the gradient in the transport rates cause erosion between, approximately, the alongshore locations 2000 and 3000 m and sedimentation behind the coral reef.

- The difference in sediment transport rates between the two model executions is (partially) caused by the difference in nearshore waves. Compared to model execution 178 (median), model execution 175 (maximum) results in (i) higher, and (ii) more northward directed nearshore waves.

This explains the northwards directed sediment transport resulting from execution 175 (maximum).

- The differences in the nearshore waves between the two model executions are probably caused by the differences in the change of wave direction (Table 5.3). The change in wave direction involves the selection of another subset of waves reaching the coast. The large positive change result in a sub-set of waves that are, on average, more northwards directed. This is agreement with the wave rose plots of Figure 3.5 (page 20)

Considering the best approximation of particular model output, can be useful for model calibration. If some data is available, the best approximation for this data can be investigated. The corresponding input parameter values may support the model calibrations. In case of Negril Coast the only available data on nearshore waves consists of a very small period of wave measurements. Moreover, the exact time series of these measurements is not available. Additionally, the measurements period contained of a storm event, while the uncertainties are focused on average conditions. Attempts to calibrate the model in this way do not make sense. Hence, calibration is not further considered here.

The probability of a particular model output can be investigated by the estimation of probability density function (Appendix A). Figure 5.3 shows that the sediment transport rates at alongshore location 3250 m (see Figure 5.1), can be estimated by an Extreme value distribution with μ = -0.41 and σ = 0.81. Furthermore, the figure shows the high probability of occurrence of the deterministic model output. The estimation of probability density function of the model output makes the results of the sensitivity analysis more *meaningful*.

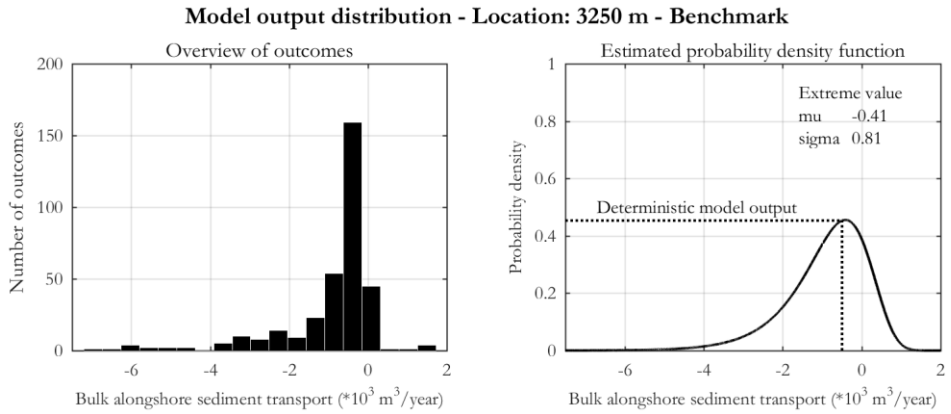

Figure 5.3 – Model output distribution – Benchmark
See Appendix A for details on distributions

5.2.2 Scatter plots

A scatter plot gives quick insights in the relations between an uncertain input parameter and the model output. Figure 5.4 shows the scatter plots, for location 3250 m (see Figure 5.1), that belong to the benchmark. The y-axes represent the bulk alongshore sediment transport rates, the x-axes represent several uncertain input parameters. The circles indicate the samples with a wave height larger than one meter. The following phenomena are observed:

- The changes in wave height and wave period are very dependent of each other. This is indicated by the experts and thus shows the correctness of the sampling procedure.

- The interaction effects of the wave height and/or the wave period, and the wave direction play an important role. As can be seen from the scatter plot of the wave direction, the samples with a large wave height result in large sediment transport rates for that particular wave direction. Additionally, the change in wave direction determines the direction of the transport. This confirms the findings in the previous section.

- Larger sediment transports do occur more often when the coral reef is low. The dependence seems not very big. Furthermore, the dependence between the D50 or the surfzone slope and the sediment transport is observable, but seems very small.

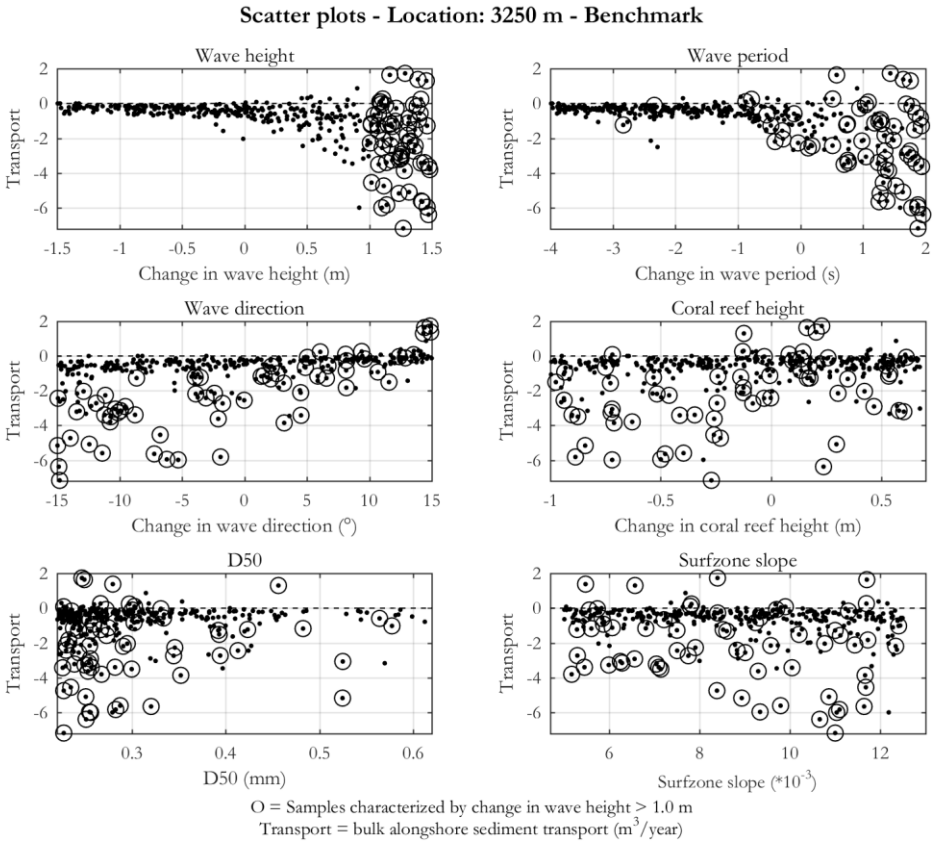

Figure 5.4 – Scatter plots - Benchmark

5.2.3 Correlation coefficients

A disadvantage of a scatterplot is the difficult distinction between the effects of *individual* parameters, which are correlated with other parameters. Moreover, it would be very time-consuming to analyze scatter plots for every alongshore location. The computation of partial (rank) correlation coefficients may be a solution to this. Figure 5.5 presents the partial correlation coefficients and the partial rank correlation coefficients with the absolute bulk alongshore sediment transport rates (see Appendix A). Table 5.4 presents the corresponding P-values, as average values along the coast. Following the P-values most correlations are significant. Therefore the sample size of 350 is considered to be large enough. The following phenomena concerning the correlation coefficients are observed:

- The rank correlation coefficients do not differ much from the correlation coefficients. Hence, the dependence between the input parameters and the model output is close to linear.

- The P-value of the reef height indicates an insignificant correlation. However, the P-values with the transport just at the South of the coral reef are almost zero. The correlation is thus significant at particular locations. Therefore, it can be concluded that a higher coral reef results in smaller sediment transport at the South of the reef. In general fewer waves approach the coast from the south (see Figure 5.2). Therefore, the effect of the reef is smaller at the northern side of the bay.

- The uncertain wave period is the most important parameter at almost all locations. Longer waves, compared to shorter waves with similar wave heights, contain more energy and cause more sediment transport. This explains the difference in importance between the change in wave period and wave height.

- The uncertain wave height is more important in the North and the South of the bay than in the middle of the bay. This is due to the shadowing zone of the headlands. The higher the wave height the more energy travels into these shadowing zones. Therefore, these zones are more sensitive to the uncertain wave height than the rest of the bay. This phenomenon is more important for the North of the bay. This is caused by the larger number of waves coming from the North refracting around the headlands.

- The wind speed has almost no effect on the bulk alongshore sediment transport rates. The wind direction does affect the sediment transport rates.

- The D50 and the surfzone slope are only involved in the sediment transport formulation of Kamphuis (Chapter 3). A larger D50 results in smaller transport rates, whereas a steeper surfzone slope results in larger transport rates. The parameters are more important in the South of the bay, than in the North. This is because in the North of the bay the wave height and period affect the sediment transport rates more (see point four).

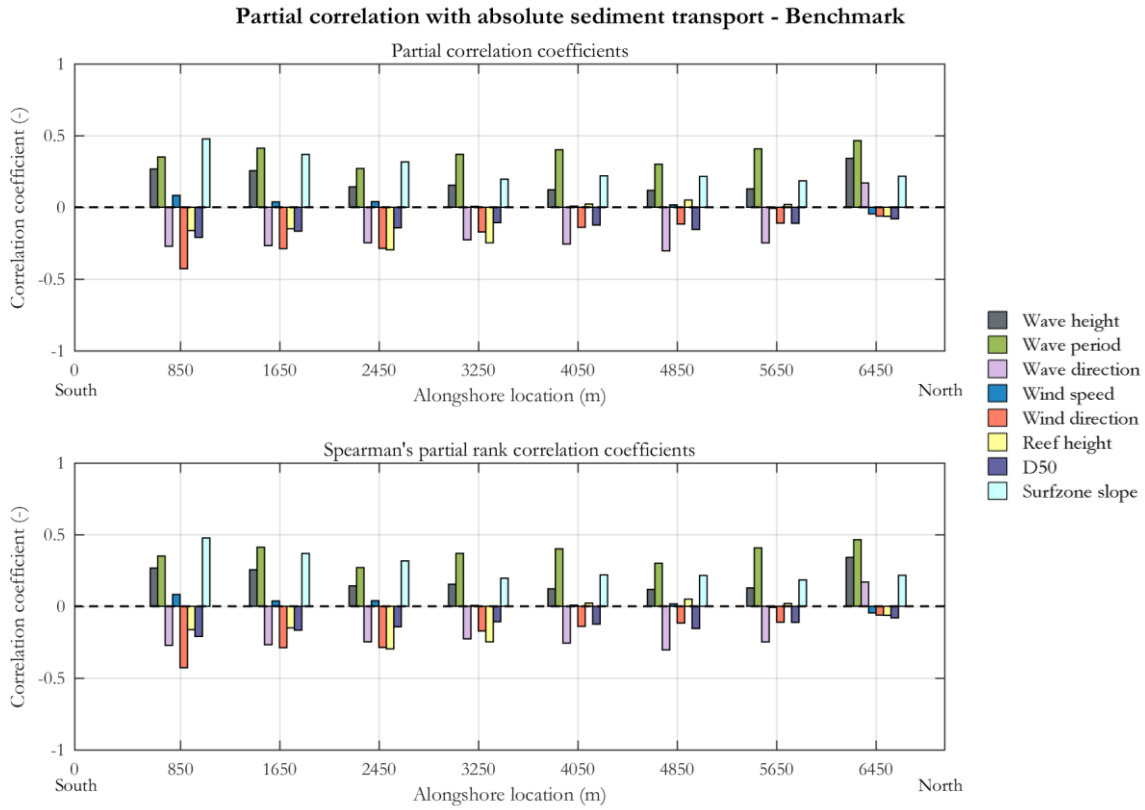

Figure 5.5 –Partial correlation coefficients – Benchmark

Table 5.4 – P-values partial correlation coefficients

	P-values							
	Wave height	**Wave period**	**Wave direction**	**Wind speed**	**Wind direction**	**Reef height**	**D50**	**Surfzone slope**
Partial correlation	0.000	0.003	0.003	0.628	0.021	0.227	0.021	0.000
Partial rank correlation	0.000	0.008	0.016	0.167	0.000	0.061	0.017	0.000

5.2.4 Copula estimations

The uncertain input parameters can be linked to the uncertain model output through an estimated copula. The steps are explained in Appendix A, and are demonstrate here by estimating a copula that relates the change of wave direction and the bulk alongshore sediment transport rates. Figure 5.6 shows a scatter histogram of the change in wave direction and the bulk alongshore sediment transport rate (location 3250 m), transformed to a standard normal distribution. Four families of copulas are fit through the data, and samples are generated from these approximations (Figure 5.7). Consequently, the total correlation and the correlations of the points in the four quadrants are calculated for the copula samples and for the data in Figure 5.6. Table 5.5 below shows the results. The Gumbel and the Gaussian copulas seem to estimate the original data best. Note: As discussed in Chapter 2, many methods to estimate a copulate and to test the best estimate exist. Since this research does not focus on finding the best copula estimation the simplest method is used here

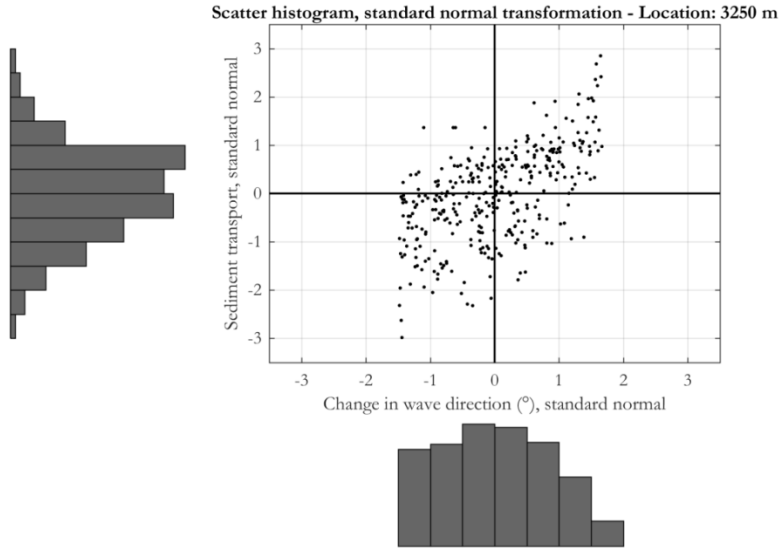

Figure 5.6 – Scatter histogram, standard normal transformation

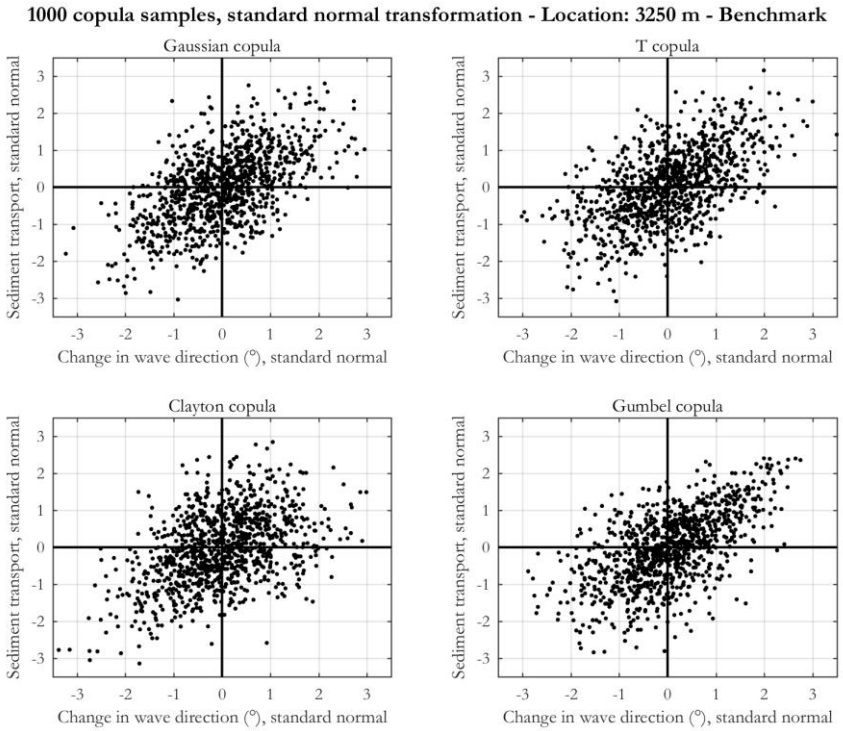

Figure 5.7 – Scatter plots of copula samples, standard normal transformation

Table 5.5 – Overview semi-correlation coefficients, with in bold the most similar correlations

	ρ_{total}	$\rho_{north\ west}$	$\rho_{north\ east}$	$\rho_{south\ east}$	$\rho_{south\ west}$
Original data, transformed	0.57	0.14	0.55	-0.01	0.08
Samples from Gaussian copula	0.57	0.13	0.33	0.07	0.36
Samples from T copula	0.57	0.17	0.43	0.19	0.30
Samples from Clayton copula	0.45	0.02	0.04	0.03	0.41
Samples from Gumbel copula	0.60	0.05	0.56	0.06	0.27

5.3 Concluding remarks

This chapter established the benchmark through conducting a sensitivity analysis that generates results as accurate as possible within the time constraints of this research. Moreover, it is considered that a more extensive sensitivity analysis would only demand more time without a significant increase in accuracy. The following aspects are important findings of this chapter:

- The benchmark is established by conducting an extensive sensitivity analysis using: (i) the wave energy balance, (ii) the Crude Monte Carlo method with 350 samples, and (iii) the full wave climate.

- The extensive sensitivity analysis provide the following insights regarding Negril Coast:
  - The bulk alongshore sediment transport rates, and their alongshore gradients, are rather small. Hence, large erosion rates are *not* expected.
  - The coral reef height affects the sediment transport rates at locations southern from the coral reef. Hence, a degrading coral reef, decreasing in height, may cause erosion in the southern part of the bay.
  - Larger wave periods result in larger bulk alongshore sediment transport rates. Swell events may thus cause coastline erosion.

- The benchmark is characterized by a speed of one sensitivity analysis in 2450 hours, using one computer with a 3.60 GHz CPU using one core.

- Due to practical reasons the scatter plots and copula estimations are not involved in the further assessment. Therefore, the accuracy of the benchmark is characterized by:
  - The model output ranges as presented in Figure 5.1.
  - The output distributions as presented in Figure 5.3.
  - The partial correlation coefficients and their corresponding P-values (Figure 5.5 and Table 5.4 respectively). The rank correlation coefficients are very similar to the partial correlation coefficients and are therefore not considered further.

- The speed and accuracy of the other sensitivity analyses conducted are compared to this benchmark.

6. Comparison of models – Results

The first approach to speed up the sensitivity analysis is the use of a simpler and faster model. In this case, using linear wave theory, instead of the wave energy balance, for wave transformation. This chapter assesses the speed-accuracy trade-off between the use of these two models. Section 6.1 discusses the differences between the models, through comparing single, fixed model outputs and the results of a sensitivity analysis. Section 6.2 aims to explore the application range of the linear wave theory. For this, the case of the Negril Coast is *simplified* in order to make the two models as similar as possible. Single, fixed model outputs and the results of a sensitivity analysis are compared. Section 6.3 discusses the most important findings of this chapter.

6.1 Differences between the models

6.1.1 Comparison fixed model output

Figure 6.1, on the next page, compares the bulk alongshore sediment transport rates and the nearshore wave climates resulting from the base case executed with both models. The settings for the base case are presented in Chapter 4. From Figure 6.1 the following phenomena are observed:

- The orders of magnitude of the transport rates resulting from both models are comparable.
- The linear wave theory incorporates *2d refraction effects only*. This is clearly indicated by the middle figure, which shows very similar nearshore wave roses along the shore. The alongshore differences between the roses are caused by the shadowing zone of the coral reef. The transport rates are computed based on S-φ curves. Consequently, the transport rates follow the coastline orientation. This causes the larger transport rates at both ends of the bay. Furthermore it causes, together with the presence of the coral reef, the gradients in the transport rates. These gradients cause sedimentation (at around 1,600 m) and erosion at the both ends of the bay.
- The wave energy balance incorporates *3d refraction effects*. The nearshore wave rose plots in the lower figure clearly indicates that the waves have been refracted towards the shore. This results in marginal alongshore gradients of the transport rates. Consequently, hardly any erosion or sedimentation is expected.
- The linear wave theory, compared with the wave energy balance, result in higher nearshore waves (see the wave roses). This is because the wave energy balance considers more processes, like white-capping and bottom friction.

6.1.2 Comparison results sensitivity analysis

Both models are used to conduct a sensitivity analysis using the Crude Monte Carlo method. The same uncertainties as in Chapter 0 are used (page 39). Moreover, the same 350 samples are used for both models, in order to avoid the randomness affecting the results. Figure 6.2 and Figure 6.3 present the results of the sensitivity analysis. Table 5.1 presents the P-values for the partial correlation coefficients resulting from both models. From the figures the following phenomena are observed:

- The median and extreme sediment transport rates resulting from both models are very different. The causes underlying this phenomenon are explained in previous section. Despite of these differences, the orders of magnitude of the transport rates are similar.
- The P-values are very similar. Following these P-values, most correlations are significant.
- The sign of the correlation coefficient of most parameters, resulting from both models, are similar at the most locations.
- The alongshore difference of the importance of the wave height resulting from the wave energy balance, is not indicated by the linear wave theory. This is because the 3d refraction effects are not included by the linear wave theory. This causes the differences between the resulting importance of the surfzone slope in the South of the bay of the both models.

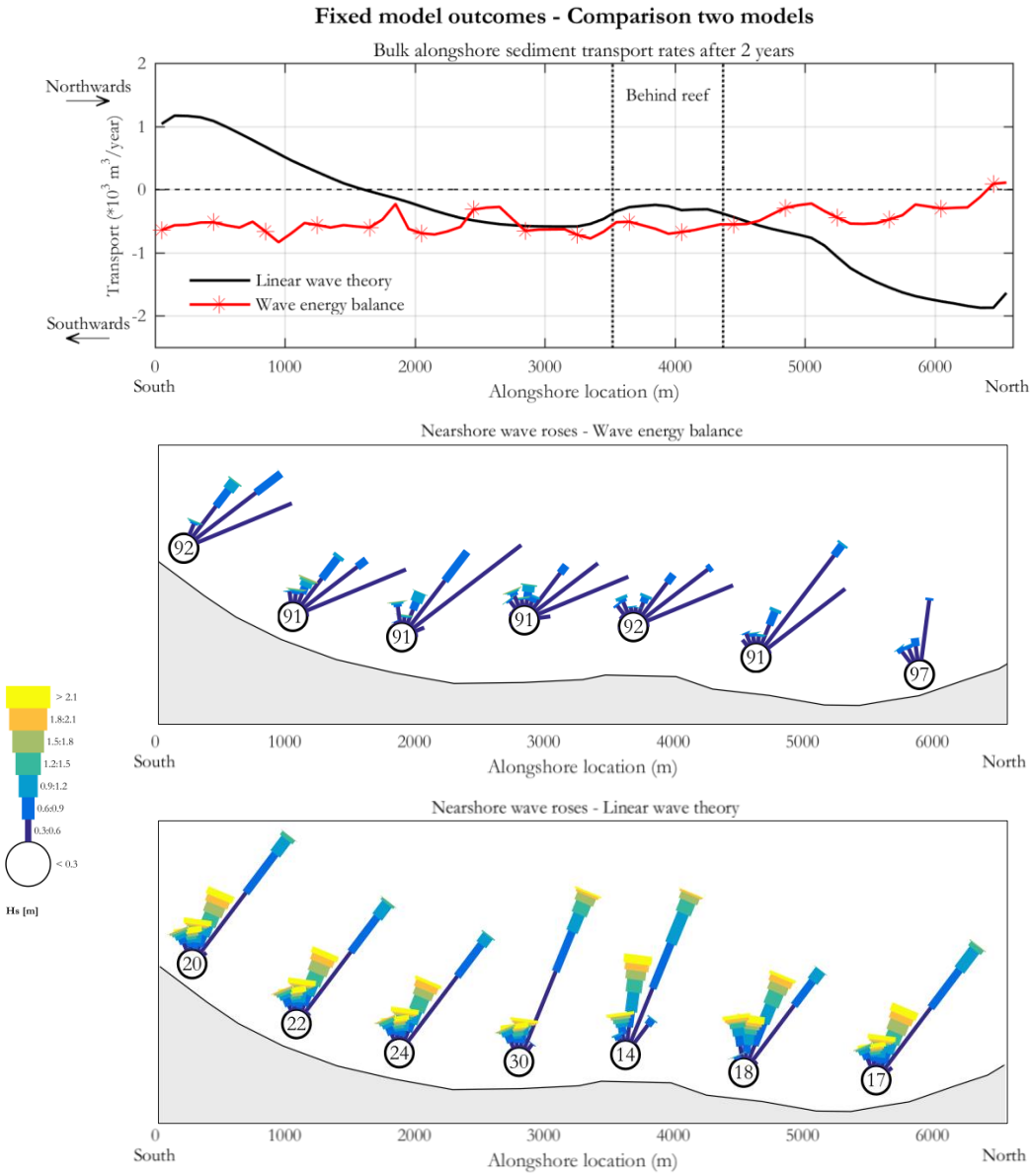

Figure 6.1- Comparison fixed model output, model I and II

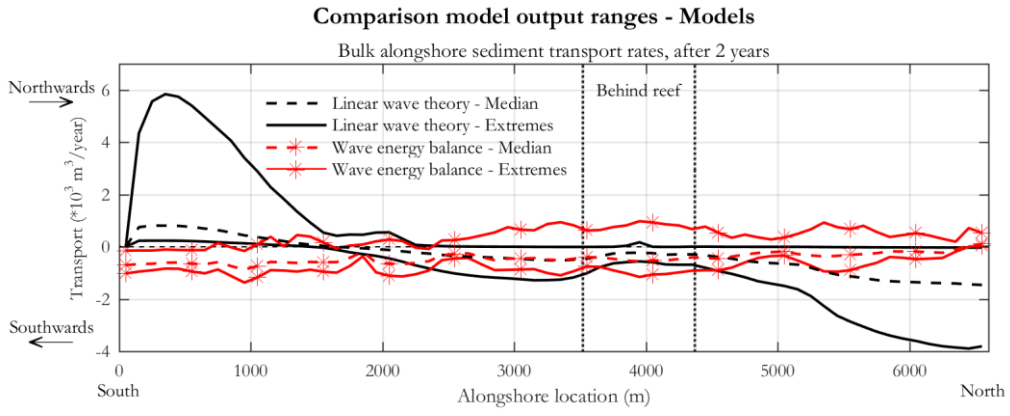

Figure 6.2 - Comparison model output ranges resulting from the two models

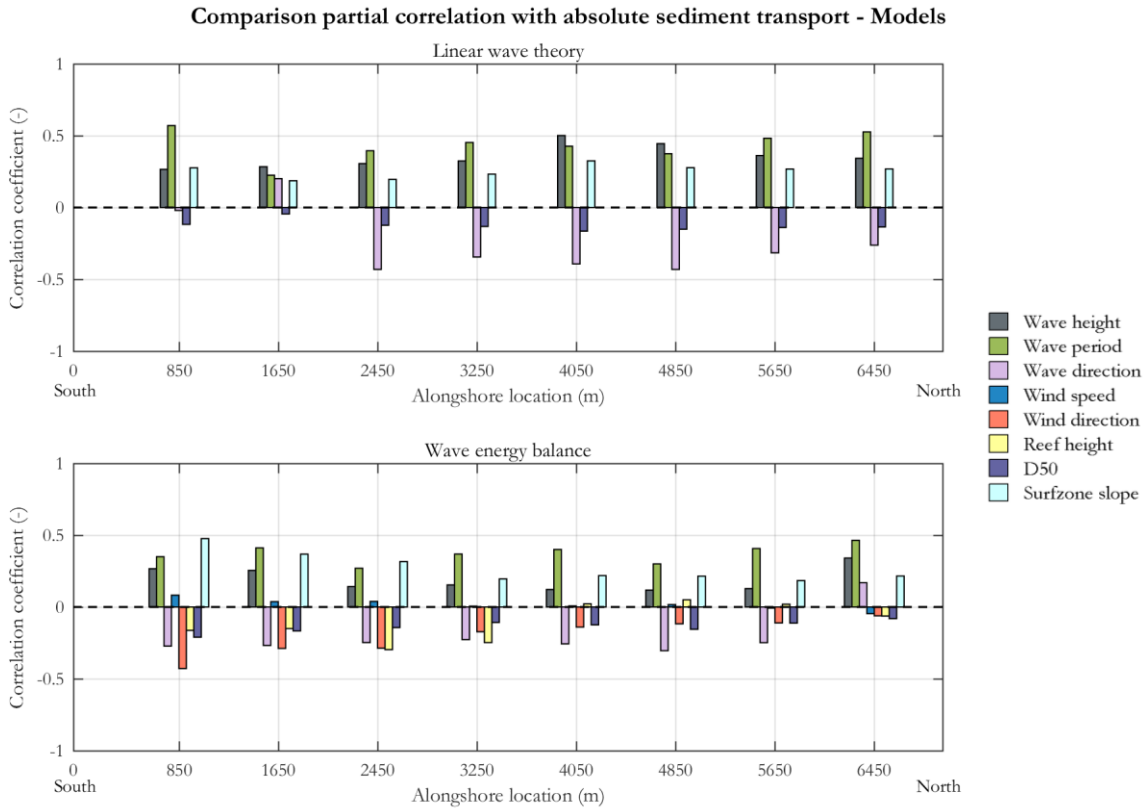

Figure 6.3 - Comparison partial correlation coefficients resulting from the two models

Table 6.1 – Comparison P-values of partial correlation coefficients resulting from the two models

Model	**P-values**							
	Wave height	Wave period	Wave direction	Wind speed	Wind direction	Reef height	D50	Surfzone slope
Linear wave theory	0.00	0.006	0.042	n/a	n/a	n/a	0.035	0.000
Wave energy balance	0.009	0.003	0.003	0.628	0.021	0.227	0.021	0.000

6.2 Exploration of the application range of the linear wave theory

Previous section shows that a sensitivity analysis using the linear wave theory, instead of the wave energy balance, results in other insights in the coastal system dynamics at Negril Coast. The linear wave theory seems only to be useful to indicate *the order of magnitude of the sediment transport rates*. The aim of this section is to explore the application range of the linear wave theory in a more *generic sense*. Therefore, the situation at Negril Coast is *simplified*. In this way the Delft3D-WAVE model, that uses the wave energy balance, can be made very similar to the linear wave theory. Fixed model outputs and results from the sensitivity analysis resulting from both models are compared. The sections below discuss how the situation at Negril Coast is simplified (Section 6.2.1) and the results of the comparisons (Sections 6.2.2 and 6.2.3).

6.2.1 Simplification of the situation at Negril Coast

To make the two models as similar as possible, the situation at Negril Coast is simplified. The simplification is based on the factors causing the differences between the models (Table 4.2, page 37). The simplification is done in steps, in order to investigate what differences are most important. Table 6.2 shows the four steps undertaken to simplify the situation at Negril Coast. The differences that originate from processes involved are not equalized, since this is the foundation of the models. The four steps

results in four different models for the wave energy balance. Figure 6.4 shows the adapted bathymetries at Negril Coast used for the Delft3D-WAVE model in the last two steps.

Table 6.2 – Simplifications the situation at Negril Coast in steps

Step	Factor	Linear wave theory	Wave energy balance
1	Size of wave input	Wave climate	Waves only based on 1 location Onshore directed waves only
2	Wind input	-	Additionally: ignoring wind
3	Bathymetry & coral reef	Additionally: No coral reef	Additionally: constant cross-shore profile and no coral reef
4	Bathymetry	Additionally: No coral reef	Additionally: uniform depth contours

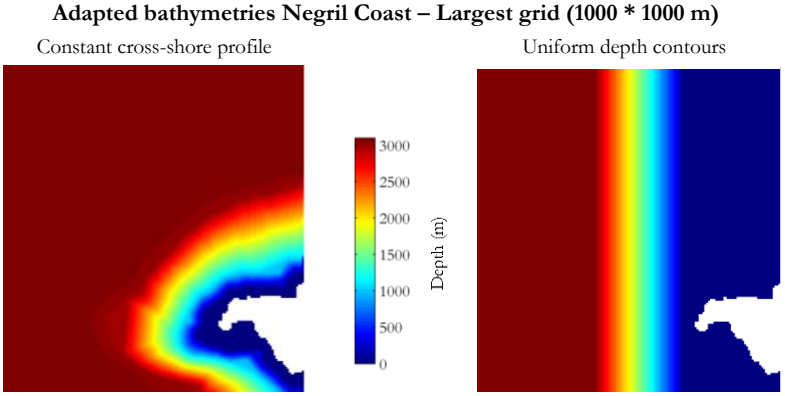

Figure 6.4 – Adapted bathymetries Negril Coast

6.2.2 Comparison fixed model outputs – Simplified cases

Based on the simplified situations, models are made and executed. The fourth step results in two models: the model is executed once with all the waves, and once with the onshore directed waves only. Figure 6.5 presents the fixed model outputs resulting from the model executions. The following phenomena are observed:

- Ignoring the wind results in slightly smaller sediment transport rates. However, the differences are minor. This is in line with the minor importance of the wind speed (Figure 6.3).
- The transport rates resulting from the constant cross-shore profile (wave energy balance) are very different than the transport rates resulting from the linear wave theory. Considering the nearshore wave roses this is due to the focusing and defocusing of waves caused by the varying bathymetry.
- The transport rates resulting from the uniform depth contours (wave energy balance) are very similar to the transport rates resulting from the linear wave theory. This is because the (de-) focusing of waves does not occur. This can be clearly seen from the nearshore wave roses that are very similar alongshore.
- Regarding the simplest situation – Including all the waves in the Delft3d-WAVE model, results in slightly larger transport rates than transport rates resulting from the linear wave theory. This is because of waves from the North and South reach the coast in case of the Delft3d-WAVE model.
- Regarding the simplest situation – Considering the onshore directed waves only in the Delft3d-WAVE model, results in slightly smaller transport rates than transport rates resulting from the linear wave theory. This is because some non-linear processes like bottom friction and white-capping are not included in the linear wave theory.

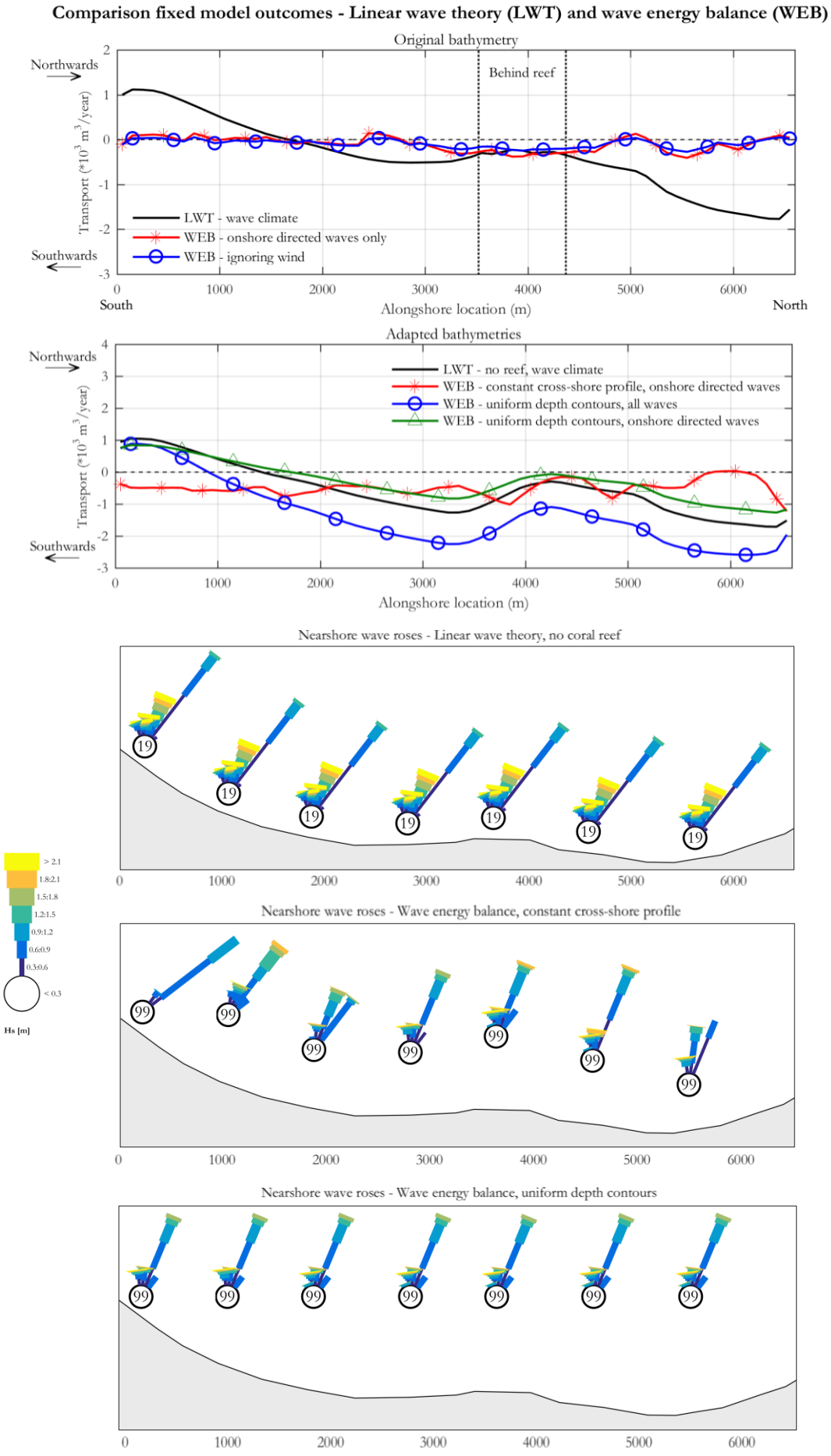

Figure 6.5 – Comparison fixed model outputs both models for the simplified cases

6.2.3 Comparison sensitivity analyses – The simple case

For the simplest situation a sensitivity analysis is conducted with both models. Note that the wave energy balance considers *all waves*. For the sensitivity analyses the Crude Monte Carlo method is used, involving the uncertainties as presented in Chapter 0 (page 39). The same 350 samples are used for both sensitivity analyses, in order to avoid the randomness affecting the comparison. Regarding the accuracy the following comparisons are made: (i) model output ranges (Figure 6.6), (ii) model output ranges (Figure 6.7) and (iii) partial correlation coefficients (Figure 6.8). The following aspects are important regarding the speed and accuracy of the results:

- For conducting the sensitivity analysis, using the same computer as for the benchmark, the wave energy balance demands 350 hours and the linear wave theory demands not even two hours. The linear wave theory speeds-up the sensitivity analysis thus with a factor of around 180.
- The model output ranges, as well as the output distributions, resulting from both models are very similar.
- The comparison between the partial correlation coefficients shows significant differences. The wave energy balance indicates the wave height as relatively unimportant parameter. The wave height for the wave energy balance, however, is a very important parameter.
- Only the partial correlation coefficients of the D50 and the surfzone slope resulting from the linear wave theory are not significant. The importance of the surfzone slope even has an incorrect sign. The insignificance is caused by the wave height, the period and the direction affecting the transport rates that result from the linear wave theory to a large extent.

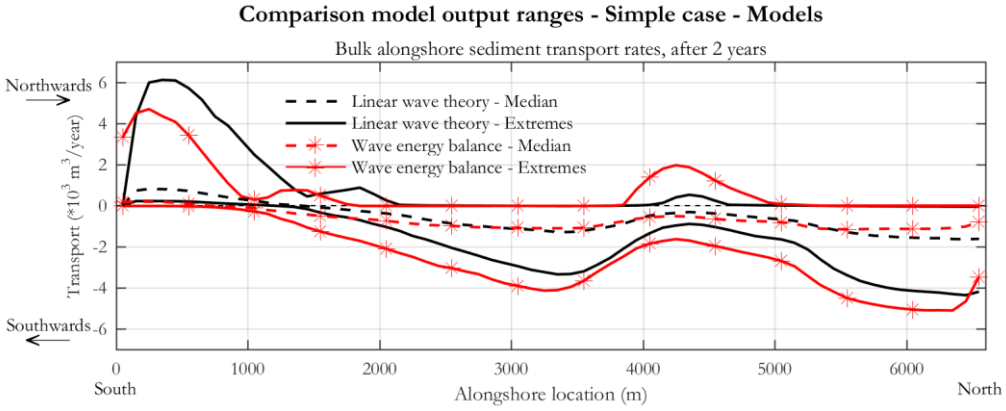

Figure 6.6 – Comparison model output ranges resulting from the two models, the simple case

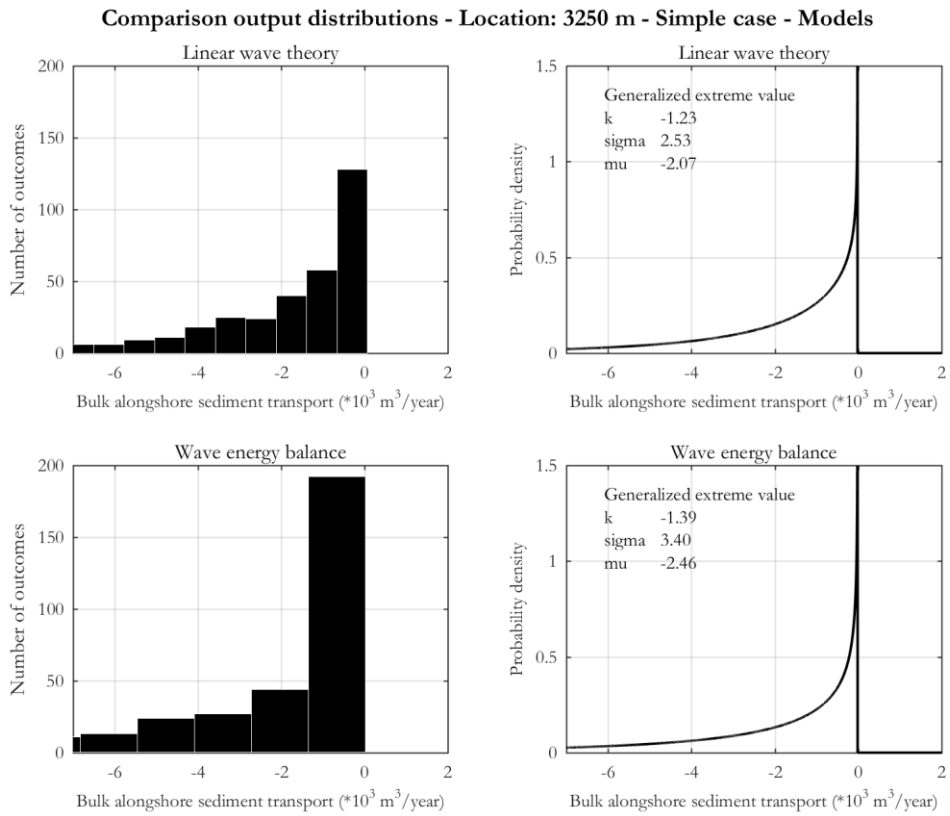

Figure 6.7 - Comparison output distributions, resulting from the two models
See Appendix A for details on distributions

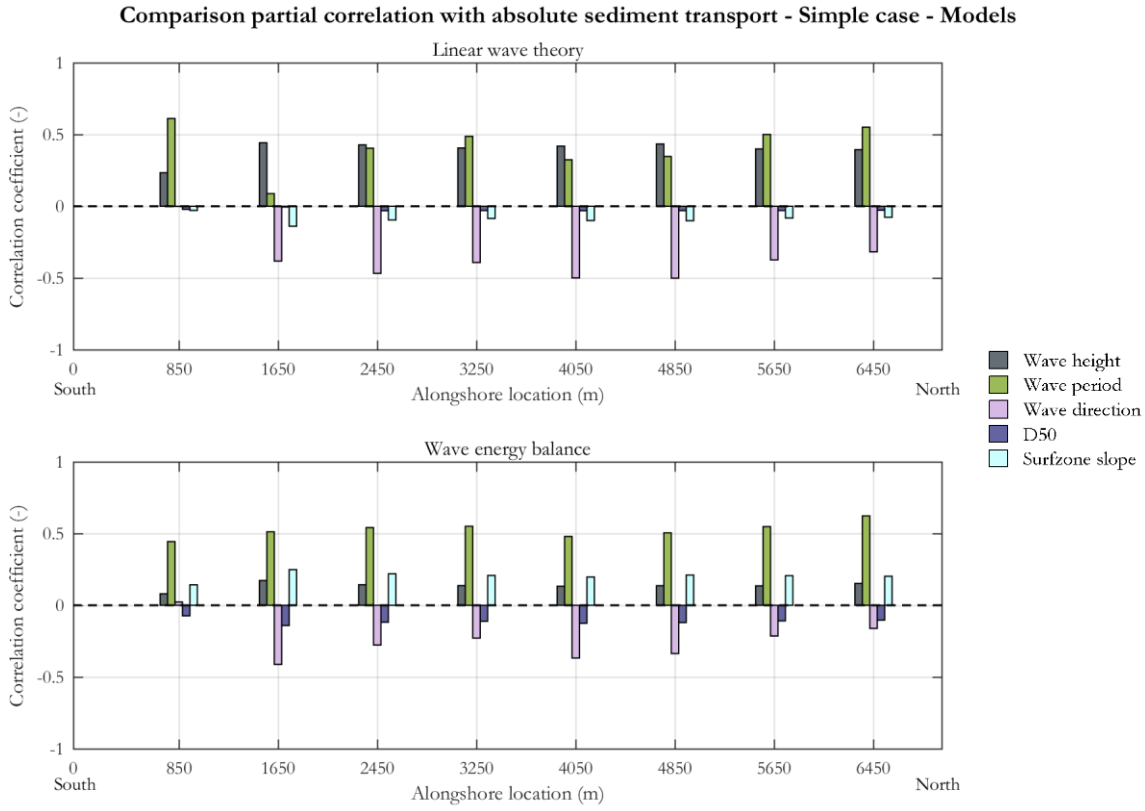

Figure 6.8 – Comparison partial correlations resulting from the two models, the simple case

6.3 Concluding remarks

This chapter assesses the speed-accuracy trade-off between the use of linear wave theory and the wave energy balance for conducting a sensitivity analysis. First, this is done for the case of Negril Coast. After, this case is simplified in order to investigate the application range of linear wave theory. The following aspects are important findings of this chapter:

- Concerning the case of Negril Coast, the insights in the system dynamics obtained by the linear wave theory differ considerably from the insights obtained by the wave energy balance. The insights obtained with the linear wave theory can be considered as inaccurate. The (de)-focusing of waves, due to the variable bathymetry, is the main cause of the differences.

- The linear wave theory can only be used to transform offshore wave characteristics to nearshore in very simple cases. The following factors characterize a very simple case:
  - Minor influence of (de-)focusing of waves due to a variable bathymetry.
  - Minor influence of the wind and other non-linear processes.
  - Minor influence of refracting offshore waves.

- Considering a very simple case, the linear wave theory demands twenty seconds, instead of one hour demanded by the wave energy balance, to transform 130 offshore waves. Therefore, the use of the linear wave theory *speeds-up* the sensitivity analysis with a factor 180.

- Considering a very simple case, the ranges of model outputs resulting from sensitivity analyses using both models are very comparable. Moreover, the resulting estimated output distributions are very similar.

- The partial correlation coefficients resulting from a sensitivity analysis with both model are slightly different. Firstly, the wave is height is relatively unimportant when the wave energy balance is used, but is very important when the linear wave theory is used. Secondly, the D50 and the surfzone slope resulting from the linear wave theory are not significant. The insignificance is caused by the uncertain wave height, period and direction affecting the transport rates that result from the linear wave theory to a large extent.

7. Comparison of sensitivity methods - Results

The second approach to speed up the sensitivity analysis is the use of other sensitivity methods than the Crude Monte Carlo method. This chapter assesses the speed-accuracy trade-off between the use of the sensitivity methods selected in Chapter 2. Section 7.1 and 7.2 compares the probabilistic and non-probabilistic method(s) to the benchmark. Section 7.3 discusses the chapter's most important findings.

7.1 Comparison probabilistic methods

This section investigates the speed-accuracy trade-off between the use of the Crude Monte Carlo method or the Latin hypercube method for conducting a sensitivity analysis. The sample size of probabilistic methods is hard to define beforehand (see Section 2.1). Therefore, the sensitivity analysis is conducted with 75 and 50 Latin hypercube samples, using the same uncertainties and correlations as for the benchmark (Table 5.1, page 39). Figure 7.1-7.3 and Table 7.1 present the results of the two sensitivity analyses and the benchmark. The findings are discussed at page 56.

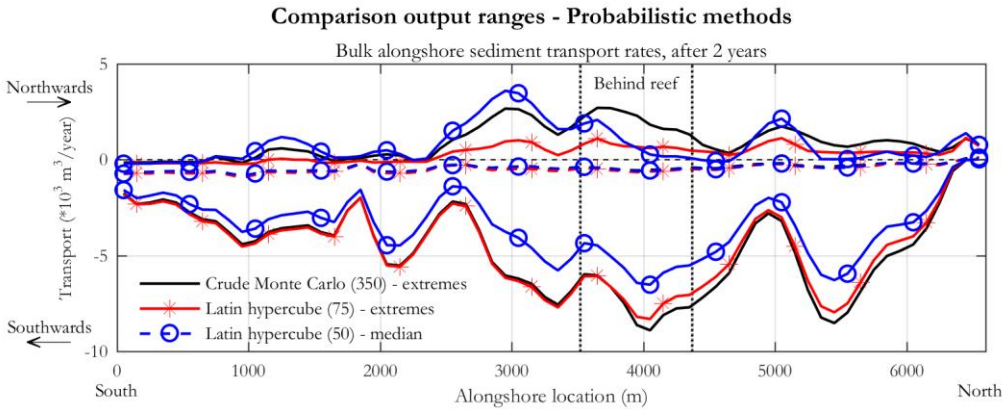

Figure 7.1 - Comparison output ranges resulting from probabilistic methods

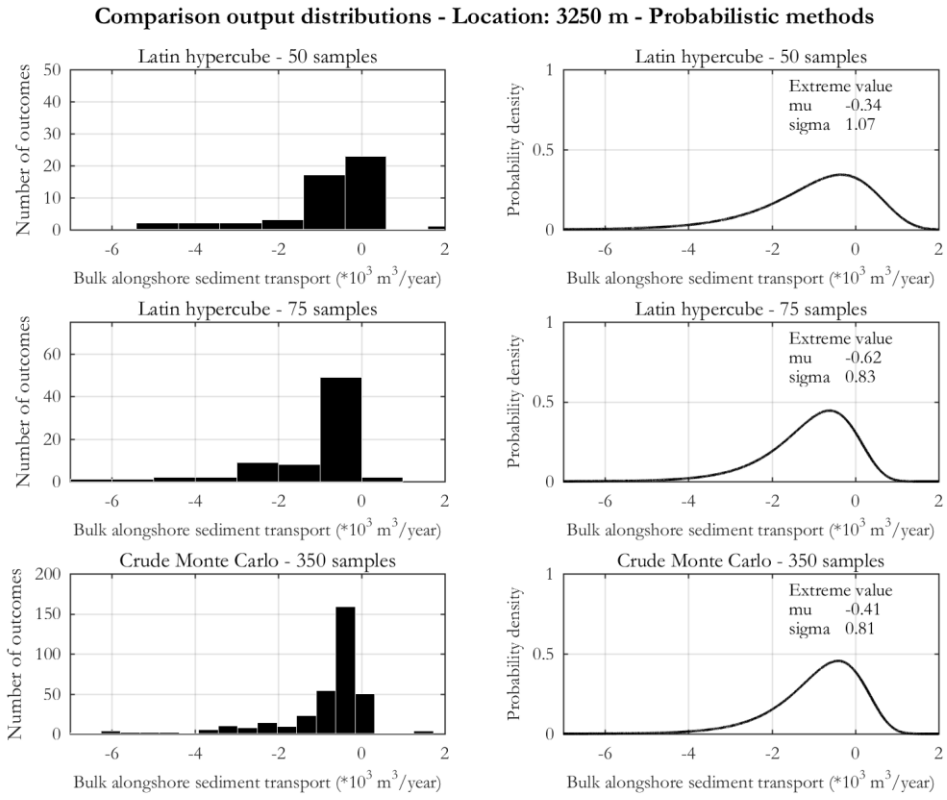

Figure 7.2 – Comparison output distributions resulting from probabilistic methods. See Appendix A for details on distributions

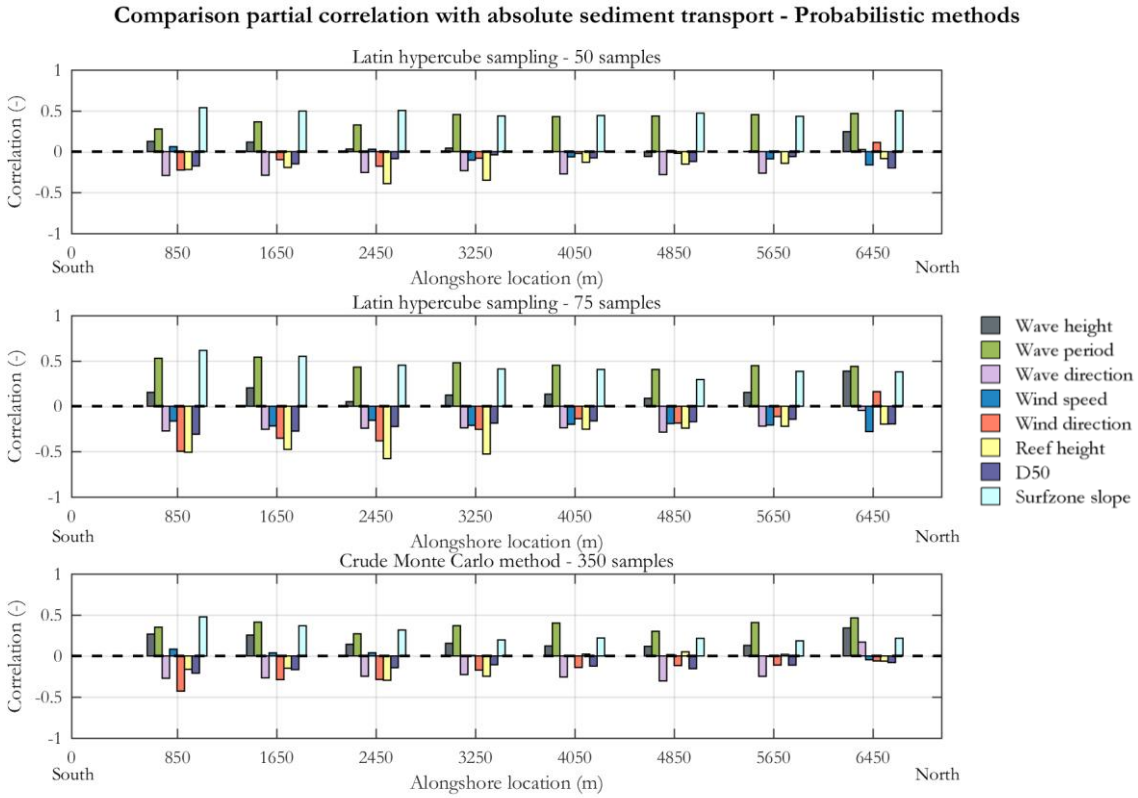

Figure 7.3 – Comparison partial correlation coefficients resulting from probabilistic methods

Table 7.1 – Comparison P-values resulting from probabilistic methods

Method & sample size	**P-values**							
	Wave height	Wave period	Wave direction	Wind speed	Wind direction	Reef height	D50	Surfzone slope
Latin hypercube 50	0.668	0.056	0.135	0.646	0.553	0.234	0.480	0.003
Latin hypercube 75	0.291	0.007	0.061	0.149	0.131	0.023	0.102	0.020
Crude Monte Carlo 350	0.010	0.003	0.003	0.628	0.021	0.227	0.021	0.000

From the figures the following phenomena, regarding the relative accuracy of the Latin hypercube method, are observed:

- Model output ranges: The medians resulting from all the probabilistic methods are very similar. The extremes resulting from the Latin hypercube method with 50 samples are smaller than the extremes resulting from the other methods.
- The output distributions from the Latin hypercube method 75 and the Crude Monte Carlo method 350 are very similar. The output distribution resulting from the Latin hypercube method with 50 samples is significantly smoother. This proofs that not only the extremes are different (previous point), but that it is a more structural difference over the range of the model output.
- The P-values belonging to the partial correlations resulting from the Latin hypercube method with 50 samples are considerably higher than the other two sets of P-values. In case of the 50 samples, only the partial correlation between the surfzone slope and the transport is significant.
- The P-values belonging to the partial correlations resulting from the Latin hypercube method with 75 samples are, except from the wind speed, all higher than the values resulting from the benchmark. Especially they P-vale of the wave height increased significantly. This can be due to different behavior of the model output in case of extreme values for the change in wave height.

- The partial correlation coefficients resulting from the Latin hypercube method with 75 samples are larger than the coefficients resulting from the Crude Monte Carlo method. This is because through fewer samples select the same extreme input parameters are selected. However, the partial correlation coefficients do follow the same alongshore pattern, i.e. both methods result in comparable orders of parameter importance. The correlation with the wind speed differs, but the P-values indicate that this correlation is not significant in case of both methods.

- The partial correlations resulting from the Latin hypercube method with 50 samples are considerably different than the correlations resulting from the two other sensitivity analyses.

7.2 Comparison Crude Monte Carlo method & non-probabilistic methods

This section assesses the speed-accuracy trade-off between the use of the Crude Monte Carlo method and non-probabilistic sensitivity methods. Table 7.2 presents the values for the uncertain parameters for the non-probabilistic methods and the number of model executions required. The Fractional factorial design is a 2^{8-2} design of resolution V. This means that two factor interactions are confounded with three factor interactions (Chapter 2) (Box and Hunter, 1961). Chapter 2 explains that non-monotone and non-linear behavior of the model output can affect the accuracy of a non-probabilistic method. Section 7.2.1 investigates this model output behavior. Section 7.2.2 assesses the resulting output ranges and Section 7.2.3 deepens the OAT method. Estimated probability distributions can, logically, not be compared.

Table 7.2 - Overview of uncertainties for all non-probabilistic sensitivity methods

Variable of interest	**Sensitivity method**		
	OAT	**Morris**	**Factorial design**
Bias of wave height (m)	-1.3 : 0 : 1.2	-1.3 : -0.5 : 0.3 : 1.2	-1.3 : 1.2
Bias of wave period (s)	-1.8 : 0 : 3.9	-1.8 : 0.1 : 2.0 : 3.9	-1.8 : 3.9
Bias of wave direction (°)	-14.7 : 0 : 14.7	-14.7 : -4.9 : 4.9 : 14.7	-14.7 : 14.7
Bias of wind speed (m/s)	-0.75 : 0 : 0	-0.75 : -0.5 : -0.25 : 0	-0.75 : 0
Bias of wind direction (°)	-14.7 : 0 : 14.7	-14.7 : -4.9 : 4.9 : 14.7	-14.7 : 14.7
Bias of coral reef height (m)	-0.9 : 0 : 0.6	-0.9 : -0.4 : -0.1 : 0.6	-0.9 : 0.6
D50 (mm)	0.22 : 0.28 : 0.60	0.22 : 0.35 : 0.47 : 0.60	0.22 : 0.60
Surfzone slope (-)	1/(200 : 140 : 80)	1/(200 : 160 : 120 : 80)	1/(200 : 80)
Number of model executions	17	36	Full: 256 & Fract: 64

7.2.1 Monotonicity and linearity

Figure 7.4 presents the model output behavior resulting from the change of some parameter values.

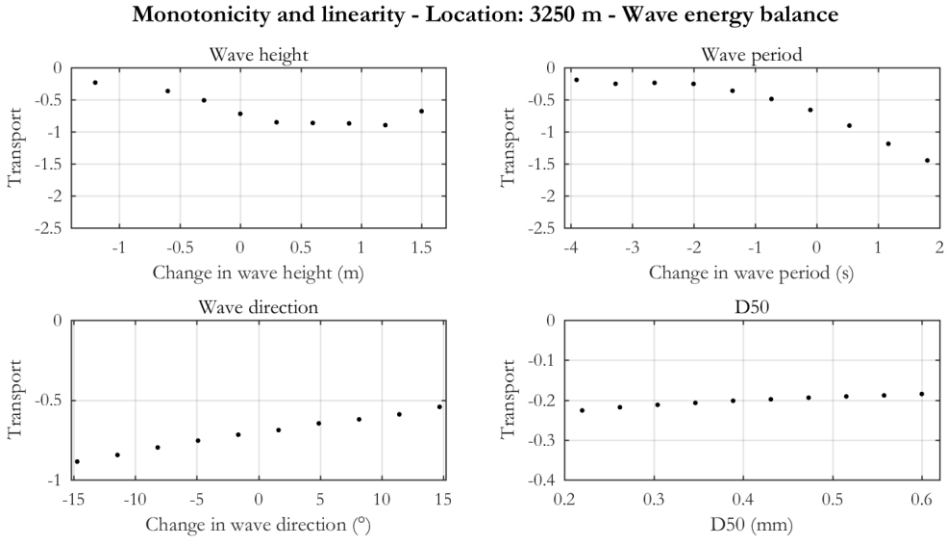

Figure 7.4 – Monotonicity and linearity wave energy balance

The figure indicates that the model output behaves monotonically increasing (absolutely decreasing) to a change in wave direction or a change in the sediment size. The relationships seem close to linear. The behaviour as a reaction on a change in wave height and wave period is clearly not linear. The reaction of the model output on a change in wave height is even non-monotone. The model outputs for a wave height increase of 0 m and 1.5 m are almost similar. The implications of this behaviour are discussed in the next sections.

7.2.2 Output ranges

Figure 7.5 presents the resulting output ranges. Table 7.3 provides parameter values for the Full factorial design method resulting in the maximum model output. The following phenomena are observed:

- The OAT method results in a relatively small range of model outputs. From this it is concluded that interaction effects cannot be neglected in the case of Negril Coast.
- The Full and Fractional factorial design methods result in very similar ranges of model output. The small difference at some locations is caused by two-factor interaction effects involving the coral reef height. These interaction effects are obviously minor.
- The ranges resulting from both the factorial designs are very large. Table 7.3 shows that the maximum model output is caused by, inter alia, an extreme change of the wave direction and wind direction in opposite direction. Considering the correlations indicated by experts this is not realistic. Hence, the large ranges of model output can be considered as *unrealistic* as well.
- The Morris method changes more parameters at the same time, without taking into account dependencies between these parameters. Moreover, the Morris method is a random method. The Morris method results thus *accidentally* in a realistic range of model outputs (Figure 7.5). It is impossible to know beforehand what parameters the Morris method selects and thus whether the resulting range of model output will be realistic.

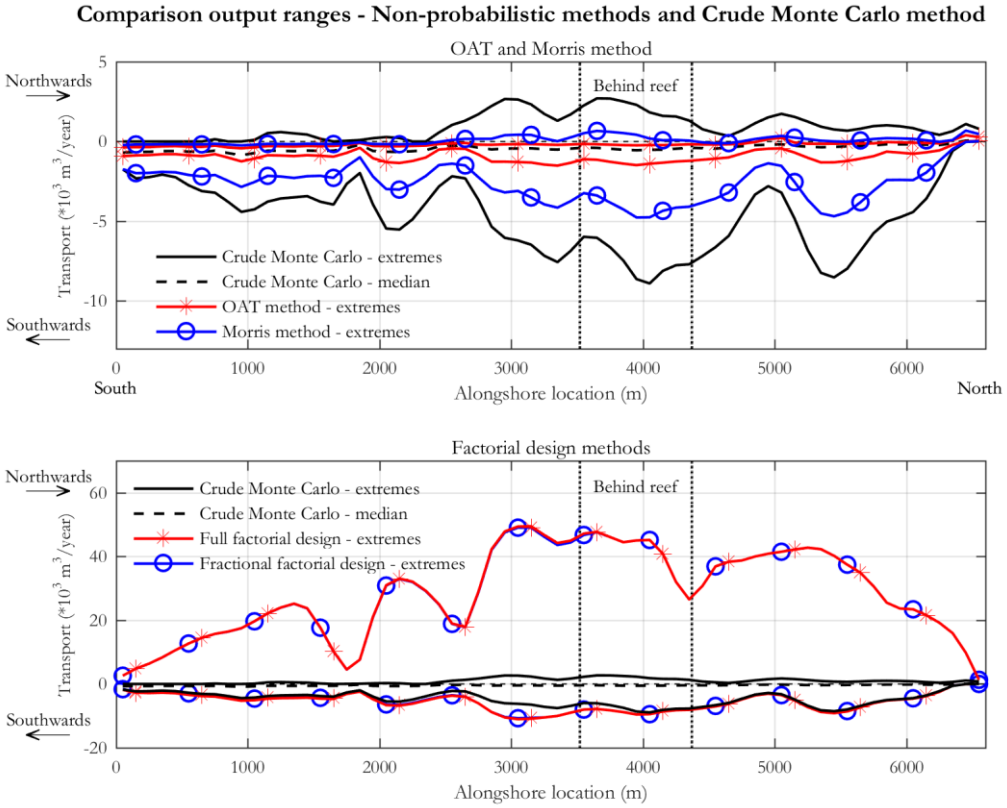

Figure 7.5 - Comparison output ranges non-probabilistic methods with Crude Monte Carlo

Table 7.3 – Values of uncertainties resulting in particular output, Full factorial design and Crude Monte Carlo

Method	Change of						D50	Surfzone slope
	Wave height	Wave period	Wave direction	Wind speed	Wind direction	Reef height		
Full factorial des. (max)	1.3 m	-3.9 s	14.7°	0 m/s	-14.7°	0.9 m	0.22 mm	1/80
CMC (max)	1.2 m	0.0 s	14.4°	0.7 m/s	12.3°	-0.16 m	0.25 mm	± 1/85

7.2.3 OAT method

Previous section shows that methods that change more parameters at the same time, without taking into account dependencies between these parameters, may result in unrealistic model outputs. Therefore, this section focusses on the sensitivities of the OAT method, presented in Figure 7.6. A negative effect means a *decrease* in *absolute* sediment transport.

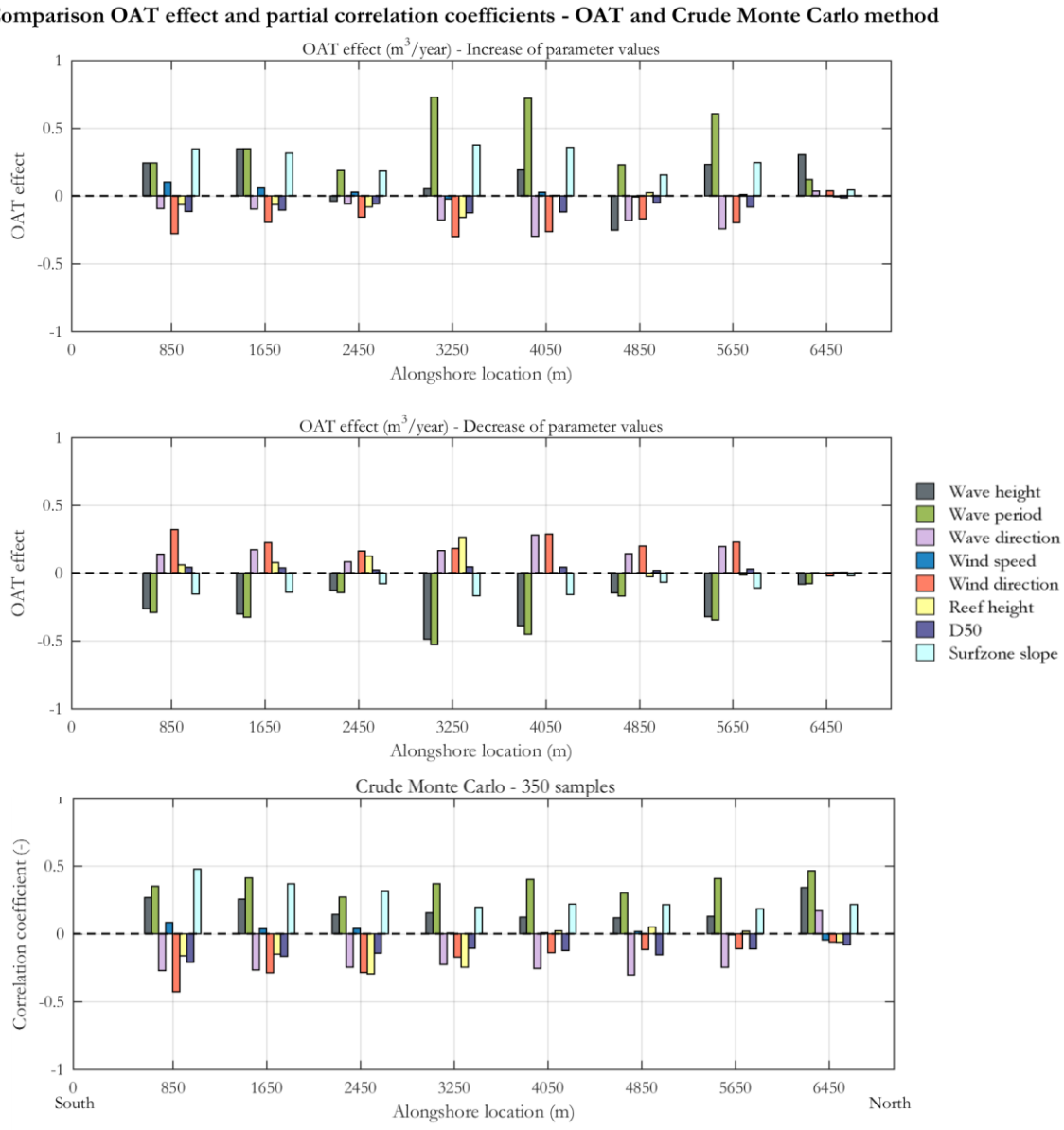

Figure 7.6 – Comparison sensitivities OAT and Crude Monte Carlo method

From Figure 7.6 the following phenomena are observed:

- The OAT effects resulting from an increase and decrease of the same parameter often differ from each other. This indicates that the model output behaviour is *non-linear*.

- The OAT effects, compares to the partial correlation coefficients, indicate the correct sign of a correlation.

- Regarding location 3250 m: An increase of the wave height seem to hardly affect the transport rates at this location. Considering Figure 7.4, this is because of the non-monotone behaviour of the model output due to a change of the wave height. In case the increase in wave height would have been the only model execution conducted in order to check the importance of the uncertain parameters, wrong conclusions on parameter importance would have been drawn.

7.3 Concluding remarks

This chapter assesses the speed-accuracy trade-off between the use of several sensitivity methods for conducting a sensitivity analysis. The following points are important concerning this speed-accuracy tradeoff:

- The use of probabilistic sensitivity methods, compared to non-probabilistic sensitivity methods, results in considerably more accurate insights in the coastal system dynamics. This is because probabilistic methods include: (i) the entire range of possible values of every input parameter, (ii) interaction effects of input parameters, (iii) the dependencies between the input parameters, and (iv) probabilistic distributions that make the results more meaningful.

- A probabilistic sensitivity analysis using the Latin hypercube method is a factor ± 5 faster than the benchmark, while maintaining similar accuracy levels. This holds for other case studies as well.

- When parameter interactions effects *and* dependencies between these parameters play a role, the Factorial design methods and the Morris method may result in unrealistic ranges of model output. This is caused by the selection of unrealistic combinations of parameter values functioning as model input.

- The Morris method results in realistic model output ranges when dependencies between parameters do not play a role. Then the method considers, although in a limited way, non-monotone and non-linear behavior of the model output. The method's randomness is a disadvantage. The randomness can be decreased by a large number of orientations. However, this decreases the speed of the sensitivity analysis.

- The OAT method does not consider non-monotone and non-linear behavior of the model output, neither the parameter interaction effects nor the dependence between the input parameters. However, since the method changes just one parameter value at a time, more or less realistic model output can always be expected. Additionally, the OAT-method is very fast, straightforward and its results are very simple to interpret.

- The accuracy of a non-probabilistic sensitivity analysis can be increased manually. For instance, by changing more than one parameter at the same time, while incorporating the dependencies between these parameters to the best possible extent.

8. Wave input reduction – Theory and results

The third approach to speed up the sensitivity analysis is the use of a reduced input dataset: in this case, the use of a reduced, instead of a full, wave climate. This chapter assesses the speed-accuracy trade-off between the use of a reduced and a full wave climate. Section 8.1 introduces theory on wave input reduction and especially on combining wave input reduction with a sensitivity analysis. Section 8.2 presents the wave input reduction method used in case of the Negril Coast. Section 8.3 compares the results of a sensitivity analysis conducted with the full and the reduced wave climate. Finally, Section 8.4 presents the most important findings of this chapter.

8.1 Theory on wave input reduction in combination with a sensitivity analysis

8.1.1 Introduction to wave input reduction

A *wave climate* is a collection of observations that describe the sea state over time. For this description several variables can be used, for instance wave height, period and direction. The collection of these variables at a certain time is called a *wave condition* (Olij, 2015). Input reduction aims at selecting a *limited number of conditions* with which a specific target, for instance sediment transport rates, obtained with the original wave climate is accurately reproduced (Walstra et al., 2013). Two fundamental options are available to derive this limited number of conditions (Walstra et al., 2013):

- Reconstructing offshore time series to *nearshore time series*.
- Constructing of a *reduced wave climate* consisting of a limited number of conditions in a particular or arbitrary order.

Figure 8.1 presents a scheme that summarizes the steps belonging to the first option. From the offshore wave database a number of waves is selected that represents the wave climate variability at deep water. This is called the *reduced wave climate*. This selection of offshore waves is transformed to nearshore wave conditions by a wave transformation model. Consequently, the original time series of the offshore database can be reconstructed to a nearshore time series. For this, statistical methods can be used (Camus et al., 2011). Finally, the nearshore time series need to be validated. This can be done through e.g. a comparison with measurements (Camus et al., 2011). The resulting nearshore time series maintain the same pattern of wave chronology as the offshore wave database. This is crucial in cases when wave chronology affects the model output (Walstra et al., 2013). This is for instance the case when coastline dynamics is directly linked to individual storm events (Walstra et al., 2013).

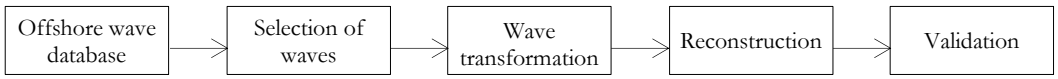

Figure 8.1 - Scheme providing overview of methodology of wave input reduction (Based on: Walstra et al., 2013, Camus et al., 2011)

Figure 8.2 presents a scheme that summarizes the steps belonging to the second option. The offshore wave database is reduced to a reduced wave climate. The sequence of the wave conditions in this climate potentially affects the model output. In these cases it is essential to investigate the correct chronology of the wave conditions. The durations of the individual wave conditions also affects the model output and hence need to be investigated. Validation can take place through e.g. comparing the representation of a specific target, for instance transport rates, by the reduced and the full wave climate (Walstra et al., 2013).

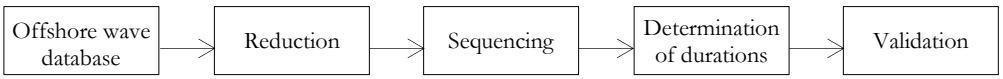

Figure 8.2 - Scheme providing overview of methodology of wave input reduction (Based on: Walstra et al., 2013, Olij, 2015)

Concerning the second option, many reduction methods are available (Olij, 2015). Following Olij (2015) a distinction can be made between: (i) *clustering methods*, and (ii) *dissimilarity methods*. Clustering methods *bundle* similar observed wave conditions in one cluster (Olij, 2015). Every wave condition of the reduced wave climate represents one cluster. The similarity of wave conditions can be based on many aspects, e.g., their contribution to the sediment transport. The dissimilarity methods select wave conditions that, by scaling up their probability of occurrence*, represent the diversity* of the full dataset the best (Olij, 2015).

8.1.2 Single vs. repeated reduction in a sensitivity analysis

Basically there are the following two options, presented in Figure 8.3 and Figure 8.4 respectively, to deal with wave input reduction in a sensitivity analysis:

- *Repeated reduction* changes, for every model execution again, the uncertain parameter values of the full offshore wave data base, where after the offshore wave data is reduced to a number of representing wave conditions. The computational time of the sensitivity analysis is affected by the time the wave input reduction method requires.

- *Single reduction* reduces the offshore wave dataset only once, where after the parameters values of the reduced wave climate are changed or every model execution. The computational time of the sensitivity analysis is not very much affected by the time the wave input reduction method requires.

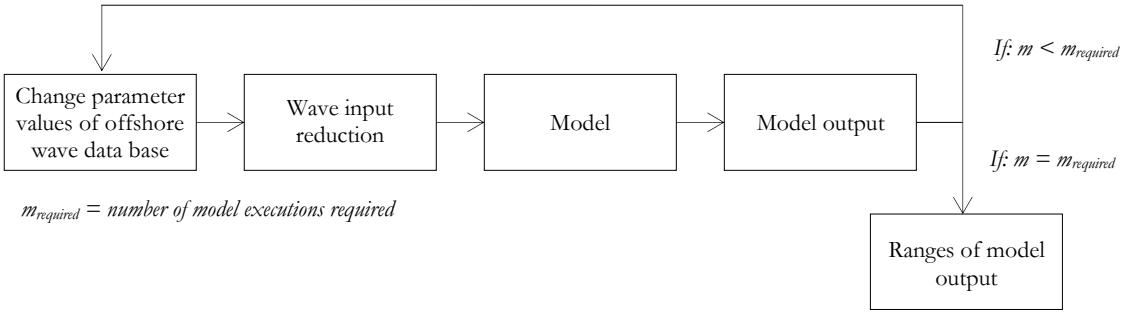

Figure 8.3 – Flow chart of repeated reduction

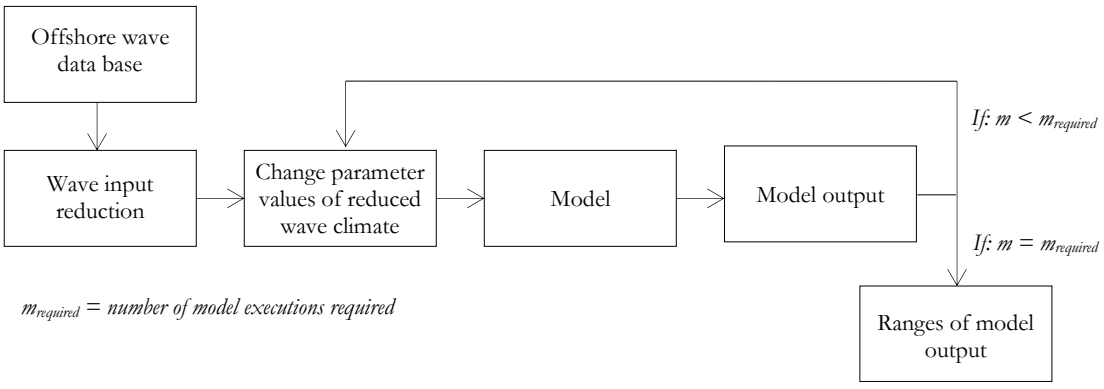

Figure 8.4 - Flow chart of single reduction

The choice between the repeated and single reduction can be crucial regarding the accuracy of wave input reduction in combination with a sensitivity analysis. This is because of the following two factors:

- *Consistency of the input reduction method,* i.e. the extent to which the method results in the same reduced wave climate and hence the same model output for the same input parameter values. The consistency is crucial in case of repeated reduction. An inconsistent method results in a model output that is sensitivity for the reduction method *itself*. Consequently, it would be very

complicated to compute the importance of a parameter. Single reduction, contrarily, allows the use of an inconsistent method.

- *Omitting of waves*. Some methods omit waves that do not contribute to the modeling target. Repeated reduction allows this omitting of waves. Contrarily, the omitting of waves should be done very carefully in case of single reduction: Waves can be omitted *only* when they do not play a role for *all the changes of parameter values* necessary for conducting the sensitivity analysis.

8.2 K-harmonic means combined with the maximum dissimilarity algorithm

The following considerations steer the selection of a wave input reduction method in case of Negril Coast:

- The specific target for the wave input reduction is the bulk alongshore sediment transport.
- The models compute the average conditions, hence the wave chronology does *not* affect the resulting bulk alongshore sediment transport rates. Therefore, a reduced offshore wave climate can be constructed, without taking into account sequencing.
- Olij (2015) investigated algorithms to reduce an offshore wave climate. He recommends using the *K-harmonic means method*. To reproduce the bulk alongshore transport rates with the Kamphuis formulation. Therefore, this method is the starting point of the wave input reduction.

The K-harmonic means method is a clustering method (Olij, 2015). The clustering procedure is explained in Appendix D. Important to know is that the distribution can take place based on wave characteristics only, or on both wave and wind characteristics. Both options are used to reduce the offshore wave data base to 20 representative conditions - 15 times each. Figure 8.5 shows that the reduced wave climate represent the transport rates rather well. Furthermore, the figures show that the excluding of wind parameters may represent the full wave climate best. This is because the wind has only minor effects on the bulk alongshore sediment transport rates (Chapter 0). Finally, the figure shows that the reduction method is *inconsistent*. This is caused by the random selection of the first centroid (see Appendix D)

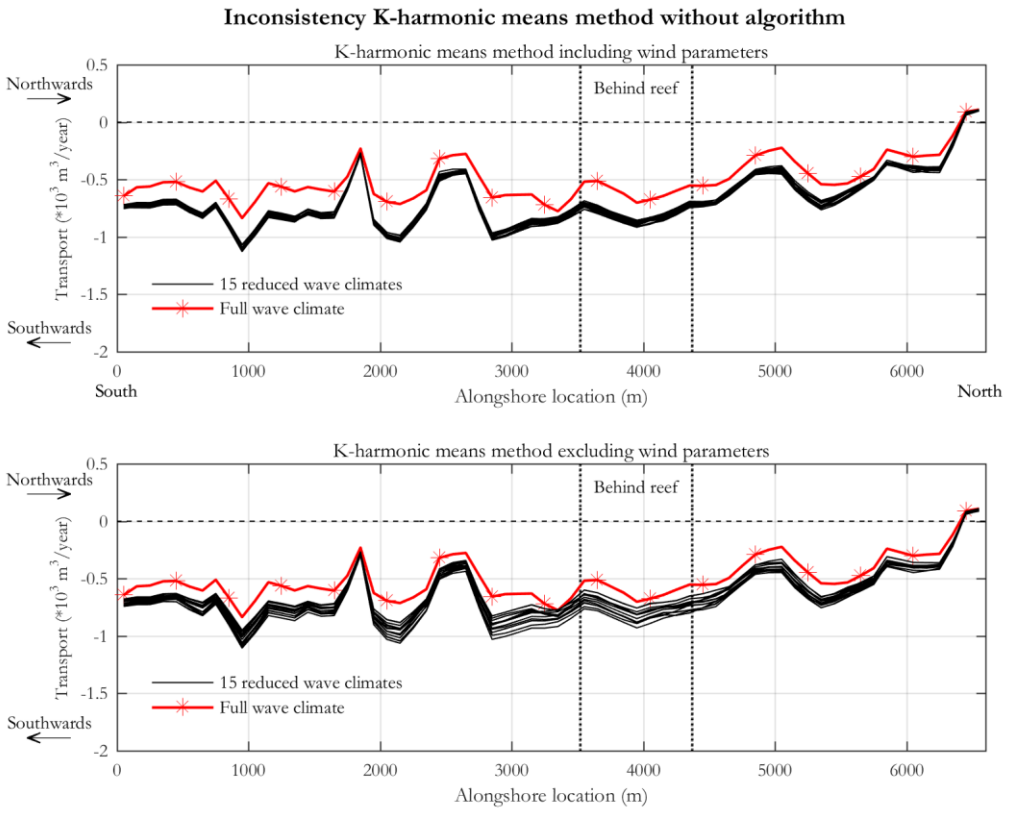

Figure 8.5- Inconsistentness of the K-harmonic means method

In order to increase the performance of the wave input reduction method and to make it consistent, the K-harmonic means is combined with the *maximum dissimilarity algorithm*. The maximum dissimilarity algorithm is a dissimilarity method that selects a sub-set of centroids that represents the full diversity of the full offshore wave dataset (Olij, 2015, Camus et al., 2011). These centroids function as the initial centroids of the K-harmonic means method. The algorithm is explained in detail in Appendix D. The computational time this reduction method requires is considerably - the repeated reduction method needs around 14 hours to create 200 reduced wave climates. Therefore *single reduction* is used.

Figure 8.6 shows that the reduced wave climates (20 waves), resulting from the K-harmonic means method in combination with the maximum dissimilarity algorithm, *overestimates* the transport with an average of 35%. This is because there are fewer northward directed nearshore waves resulting from the reduced wave climate than resulting from the full wave climate (see nearshore wave ross). Furthermore, the nearshore wave roses show that the reduced wave climate, compared to the full wave climate results, in smaller nearshore waves.

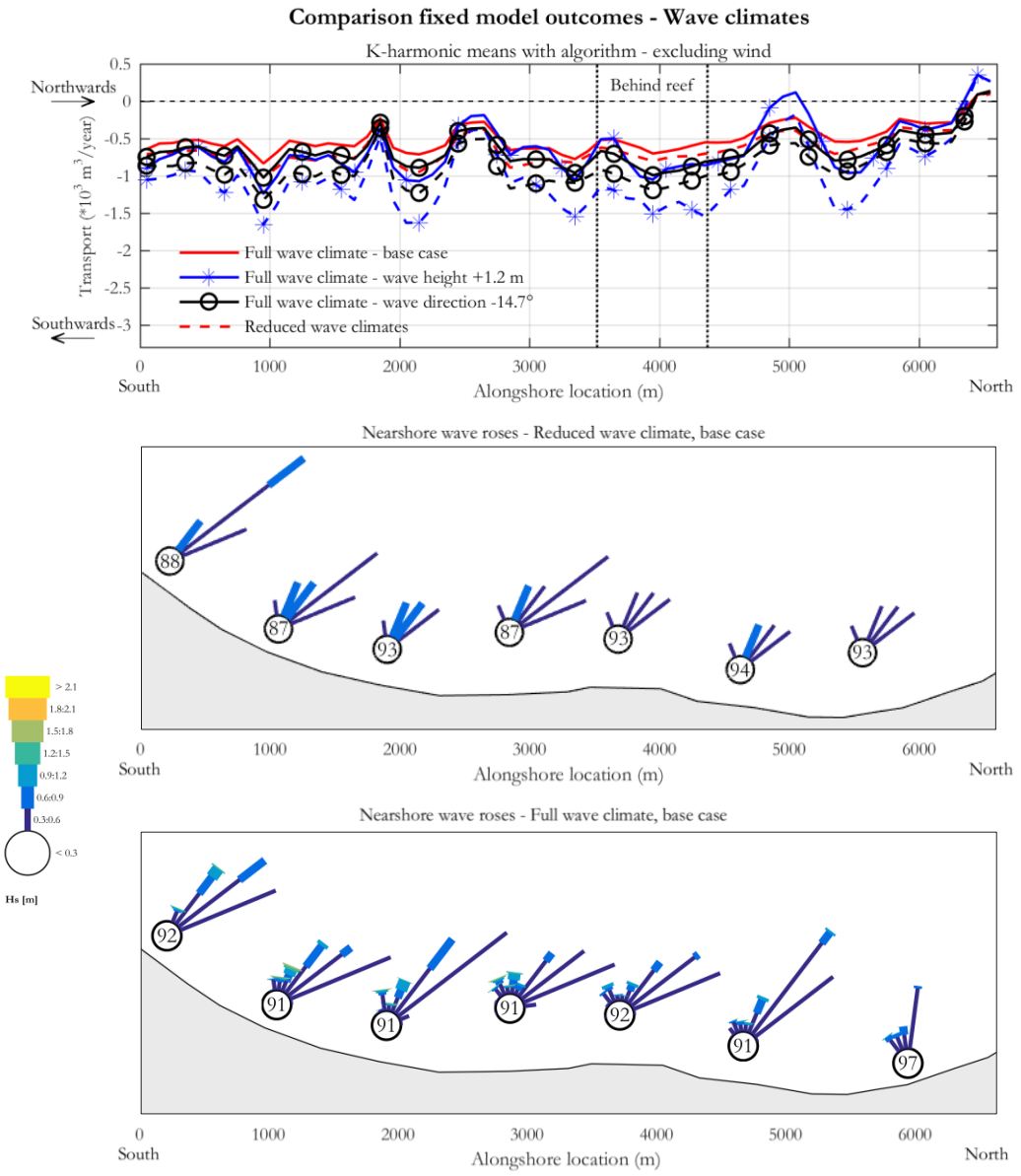

Figure 8.6 - Results of K-harmonic means method combined with maximum dissimilarity algorithm

8.3 Comparison sensitivity analysis

The full and the reduced wave climate are both used for conducting a sensitivity analysis. The Figures 8.7-8.9 present the results. The following phenomena are observed:

- The use of a wave input reduction method reduces the computationally time by a factor ± 7.
- Figure 8.7: The reduced wave climate results in median sediment transport rates within, average, a 10% range of median resulting from the full wave climate. The extremes resulting of the reduced wave climate are larger than of the full wave climate. This overestimation is on around 35%.
- Figure 8.8: The estimated output distributions resulting from both analyses are very similar. The distribution resulting from the reduced wave climate is slightly smoother, which is due to the overestimation of the transport rates that increases when the transport rates increases (Figure 8.7)
- Figure 8.9: The resulting P-values are very comparable. Furthermore, the resulting partial correlation coefficients are very similar.

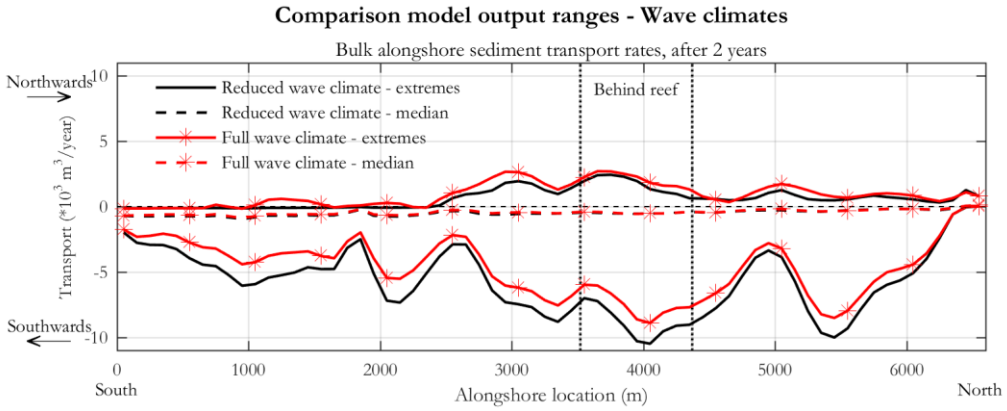

Figure 8.7 – Comparison model output ranges resulting from the full and the reduced wave climate

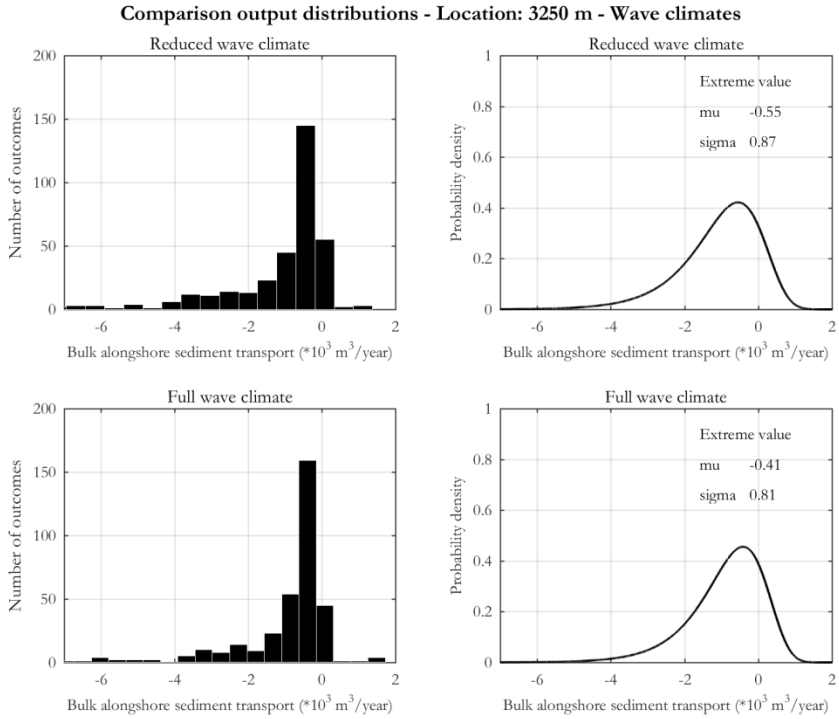

Figure 8.8 - Comparison output distributions resulting from the two wave climate. See Appendix A for details on distributions

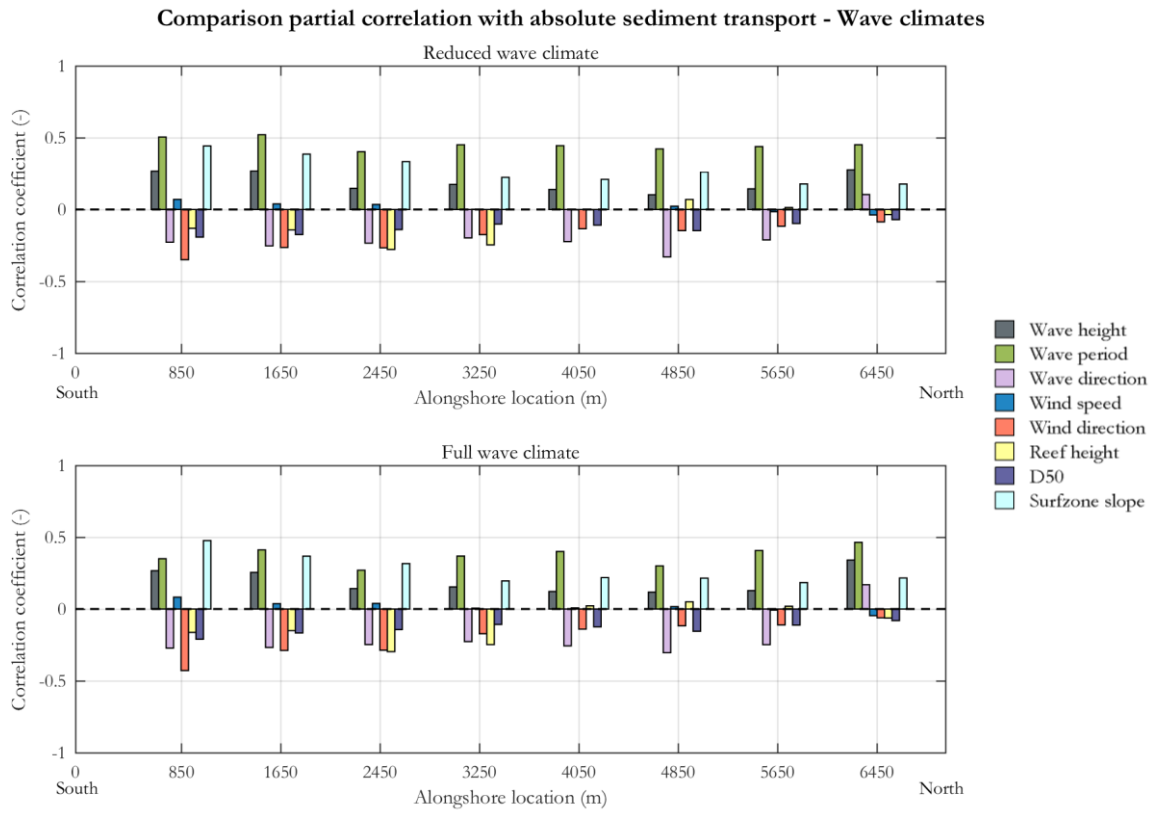

Figure 8.9 – Comparison partial correlation coefficients resulting from the full and the reduced wave climate

8.4 Concluding remarks

This chapter assesses the speed-accuracy trade-off between the use of a reduced and a full wave climate for conducting a sensitivity analysis. The following points are important concerning this speed-accuracy tradeoff:

- A sensitivity analysis with a reduced input dataset is a factor 6-13 faster than a sensitivity analysis with the full input data set. The exact factor depends on the number of waves in the full and the reduced wave climate.
- In principal the accuracy of an analysis with a reduced dataset can be similar to the full dataset. However, the accuracy depends on the performance of the input reduction individually and in combination with a sensitivity analysis.
- Input reduction methods that omit waves are likely to introduce inaccuracies in the sensitivity analysis in case of single reduction.
- Inconsistent input reduction methods are likely to introduce inaccuracies in the sensitivity analysis in case of repeated reduction.
- The K-harmonic means method is an inconsistent input reduction method.
- The K-harmonic means method in combination with the maximum dissimilarity algorithm is a consistent input reduction method.

9. Conclusions & recommendations

This last chapter is divided into three parts. Section 9.1 presents the most important conclusions related to the objectives. Section 9.2 presents findings useful for application, especially for engineers who aim to conduct an efficient sensitivity analysis. Finally, Section 9.3 presents recommendations for further research.

9.1 Conclusions

This thesis investigated approaches to conduct an efficient sensitivity analysis of wave forcing in coastal erosion studies. For that the following three speed-accuracy trade-offs have been assessed:

- Between the use of a simple and a complex model to conduct a sensitivity analysis.
- Between the use of several methods to conduct a sensitivity analysis.
- Between conducting a sensitivity analysis with a reduced and a full input dataset.

For these assessments several sensitivity analyses were conducted. The most extensive sensitivity analysis was conducted using (i) the wave energy balance to transform offshore wave characteristics to nearshore, (ii) the Crude Monte Carlo method with 350 samples, and (iii) the full wave input dataset. This analysis generates results as accurate as possible within the time constraints of this research. Moreover, it is considered that a more extensive analysis would only demand more time without a significant increase in accuracy. Therefore, the speed and accuracy of this analysis were established as the benchmark. The accuracy is defined by the (i) range of model output, (ii) estimated output distributions, (iii) sensitivities and their level of significance. The speed and accuracy of the other sensitivity analyses were compared to this benchmark.

This research demonstrated that although the achievable efficiency of the sensitivity analysis is very much dependent on the case at hand, a number of generic approaches can be applied that increase the speed while maintaining similar accuracy levels as the benchmark. The research draws the following conclusions concerning these approaches:

- The complexity of the study area determines whether linear wave theory can be applied for wave transformation. Generally, if a case is characterized by (de-)focusing of waves due to a variable bathymetry and/or governed by wind or other non-linear processes, a wave energy balance (or similar) model is needed for accurate results. In these cases linear wave theory would result in inaccurate insights. In other cases, the linear wave theory may be applied, speeding up the sensitivity analysis with a factor up to 180.

- The use of probabilistic sensitivity methods, compared to non-probabilistic sensitivity methods, results in considerably more accurate insights in the coastal system dynamics. This is because probabilistic methods include: (i) the entire range of possible values of every input parameter, (ii) interaction effects of input parameters, (iii) the dependencies between the input parameters, and (iv) probabilistic distributions that make the results more meaningful.

- A probabilistic sensitivity analysis using the Latin hypercube method is a factor ± 5 faster than the benchmark, while maintaining similar accuracy levels.

- A sensitivity analysis with a reduced input dataset is a factor 6-13 faster than the benchmark. In principal the accuracy of an analysis with a reduced dataset can be similar to the full dataset. However, the accuracy depends on the performance of the input reduction individually and in combination with a sensitivity analysis. Input reduction methods which omit waves or which are inconsistent likely introduce inaccuracies in the sensitivity analysis.

- The use of the Latin hypercube method combined with a reduced wave input dataset is the best approach for an efficient sensitivity analysis with the wave energy balance. This sensitivity analysis is a factor 30-85 faster than the benchmark.

9.2 Findings useful for application

For engineers who aim to conduct an efficient sensitivity analysis in wave-driven coastal erosion studies it is recommended to take the following steps:

1. Identify and quantify the uncertainties through a limited expert judgment process. For this, it is advised to use questionnaires with: (i) proper information on the parameters, (ii) clear and case-specific examples, and (iii) eventually, questions referring to percentages of parameter values, instead of referring to fixed numbers, in order to indicate the uncertainty of that parameter.

2. Investigate whether linear wave theory can be used to transform waves. For this it is necessary to investigate the extent to which the model output is affected by: (i) (de-)focusing of waves, (ii) wind and other non-linear processes, and (iii) waves other than strictly onshore directed. In doubtful cases, this can be investigated through a model execution using the wave energy balance.

3. Use a reduced instead of a full wave climate. Consider the following aspects carefully:
    - Check the performance of the input reduction method through a check with the results of a model execution with the full wave climate.
    - Make a careful choice between single and repeated reduction, take hereby into account whether the input reduction method is consistent and to what extent it omits waves.
    - The K-harmonic means method is an inconsistent method, hence not suitable to use for repeated reduction. The K-harmonic means method combined with the maximum dissimilarity algorithm is a consistent method, hence suitable to use for repeated reduction.

4. Estimate the sample size required by the probabilistic sensitivity methods. The 350 samples used for the Crude Monte Carlo method and the 75 samples used for the Latin hypercube method, to obtain accurate insights in the average sediment transport rates in this research, can be used as reference point. Take into account that a larger, and when opposite a smaller, sample size is needed in case of:
    - A larger number of uncertain parameters.
    - A larger degree of uncertainty of the uncertain parameters.
    - A more complex model.
    - The aim is to obtain insights in more extreme events.

5. Use a probabilistic sensitivity method in case the project's time and budget constraints allow this. Select a probabilistic sensitivity method, while taking into account that the size of an already generated sample can be increased easily when the Crude Monte Carlo method is used, but that this more difficult when the Latin hypercube method is used.

6. Use a non-probabilistic method in case the project's time and budget constraints demand this. Take into account the potential drawbacks in accuracy by assessing the following factors:
    - The dependencies between the uncertain input parameters indicated by experts.
    - The extent to which parameter interaction effects play a role. This can be indicated by experts or checked through a small number of model executions.
    - The extent to which the model output behaves non-monotonic and/or non-linear. This can be indicated by experts or checked through a small number of model executions.

7. When using a non-probabilistic method, consider to increase the accuracy of the sensitivity analysis manually. For instance, by changing more than one parameter at the same time, while incorporating the dependencies between these parameters to the best possible extent.

9.3 Recommendations for further research

The following recommendations could help to improve approaches for an efficient sensitivity analysis:

- Investigate approaches that speed-up the sensitivity analysis using a complex model with limited computational effort. This is especially relevant for cases, where simple models such as linear wave theory are not valid. A possible way to reduce the computational effort for complex models is decreasing the grid resolution. In this way the complex model physics are accounted for, while coarser grids help to decrease the computational effort. Further research should focus on the extent to which the resolution can be decreased while maintaining a sufficient level of accuracy in the model physics.

- Conduct the same research with a case study on which sufficient data is available. Then, the accuracy benchmark can consist of actual measurement data. This allows a good comparison between deterministic and probabilistic modeling, and moreover the possible improvement of the latter.

- Investigate the performance of efficient sensitivity analysis focusing on other parameters than the bulk alongshore sediment transport rates. In this way, it can be investigated how to deal with aspects in a sensitivity analysis that are neglected for this research. For instance, the dealing with extreme events or input reduction methods that require sequencing or the reconstruction of time series.

References

APPENDINI, C. M., TORRES-FREYERMUTH, A., OROPEZA, F., SALLES, P., LÓPEZ, J. & MENDOZA, E. T. 2013. Wave modeling performance in the Gulf of Mexico and Western Caribbean: Wind reanalyses assessment. *Applied Ocean Research,* 39**,** 20-30.

BITNER-GREGERSEN, E. M. & SOARES, C. G. 2007. Uncertainty of average wave steepness prediction from global wave databases. *Advancement in marine structures*. London, UK.

BOSBOOM, J. & STIVE, M. J. F. 2013. *Coastal Dynamics I - Lecture notes,* Delft, VSSD.

BOX, G. E. P. & HUNTER, J. S. 1961. The 2k-p Fractional Factorial Designs Part I. *Technometrics,* 3**,** 311-351.

BROOKER, D. C., COLE, G. K. & MCCONOCHIE, J. D. The influence of hindcast modeling uncertainty on the predictionof high return period wave conditions. 23rd International Conference on Offshore Mechanics and Arctic Engineering 2004 Vancouver, British Columbia, Canada. 7.

CAMPOLONGO, F. & SALTELLI, A. 1997. Sensitivity analysis of an environmental model: an application of different analysis methods *Reliability Engineering & System Safety,* 57**,** 49-69.

CAMUS, P., MENDEZ, F. J. & MEDINA, R. 2011. A hybrid efficient method to downscale wave climate to coastal areas. *Coastal Engineering,* 58**,** 851-862.

CARDONE, V. J. 1987. The presenst status of operational wave forecasting. *John Hopkin APL Technical Digest,* 8**,** 24-32.

CHARPENTIER, A., FERMANIAN, J. & SCAILLET, O. 2007. The estimatino of copulas: theory and practice. *Copulas: from theory to application in finance*. London: Risk Books.

CL ENVIRONMENTAL 2014. Environmental Impact Assessment - Construction of two breakwaters.

CONFALONIERI, R., BELLOCCHI, G., BREGAGLIO, S., DONATELLI, M. & ACUTIS, M. 2010. Comparison of sensitivity analysis techniques: A case study

with the rice model WARM. *Ecological Modelling,* 221**,** 1897-1906.

COOKE, R. M. & GOOSENS, L. L. H. J. 2008. TU Delft expert judgment data base. *Reliability Engineering & System Safety,* 93**,** 657-674.

COOKE, R. M. & SLIJKHUIS, K. A. 2003. Expert Judgment in the uncertainty analysis of dike ring failure frequency. *Case Studies in Reliability and Maintenance*. John Wiley & Sons, Inc.

DAMBACHER, M. & HÜBNER, R. 2013. Investigating the speed–accuracy trade-off: Better use deadlines or response signals? *Behavior Research Methods,* 45**,** 702-717.

DE MOEL, H., ASSELMAN, N. E. M. & AERTS, J. C. J. H. 2012. Uncertainty and sensitivity analysis of coastal flood damage estimates in the west of the Netherlands. *Natural Hazards and Earth System Sciences,* 12**,** 1045-1058.

DE VRIEND, H. J. & VAN KONINGSVELD, M. 2012. *Building with Nature: Thinking, acting and interacting differently. ,* Dordrecht, The Netherlands, EcoShape, Building with Nature.

DEKKING, F. M., KRAAIKAMP, C., LOPUHAÄ, H. P. & MEESTER, L. E. 2005. *A modern introduction to probability and statistics*, Springer.

DELTARES 2010. Approach for eco-morphological modelling of mega-nourishments along the Holland coast. Draft. B. Huisman and A. Luijendijk, Projectnr. 1200893-000.

DELTARES 2014a. Delft3D-WAVE User Manual.

DELTARES 2014b. Modelling coastline evolution. Practices for the setup of coastline models.: B.J.A. Huisman, Projectnr. 1209381-005.

DEN HEIJER, C., VAN DE GRAAFF, J. & VAN GELDER, P. H. A. J. M. 2008. Probabilistic sensitivity analysis of dune erosion calculations. *4th International Symposium on Flood Defence*. Toronto, Canada.

DEPARTMENT OF GEOLOGY AND GEOGRAPHY 2002. Beach sand resource assessment Negril, Jamaica: Final report on Phase 1. University of the West Indies, Mona, Jamaica

DRURY, C. G. 1994. The speed-accuracy trade-off in industry. *Ergonomics,* 37**,** 747-763.

ECMWF. 2015. *European Centre for Medium-Range Weather Forecasts* [Online]. Available: http://www.ecmwf.int/ [Accessed October, 15th 2015].

EDWARDS, K. L., VEERAMONY, J., WANG, D., HOLLAND, K. T. & HSU, Y. L. Sensitivity of Delft3D to input conditions. OCEANS 2009 - Marine Technology for Our Future: Global and Local Challenges, 26-29 Oct. 2009 2009 Biloxi, MS. IEEE, 1-8.

EDWARDS, P. E. T. 2009. Sustainable financing for ocean and coastal management in Jamaica: The potential for revenues from tourist user fees. *Marine Policy,* 33**,** 376-385.

FENTON, N. & NEIL, M. 2013. *Risk assessment and decision analysis with bayesian networks*, CRC Press.

FILATOVA, T., MULDER, J. P. M. & VAN DER VEEN, A. 2011. Coastal risk management: How to motivate individual economic decisions to lower flood risk? *Ocean & Coastal Management,* 54**,** 164-172.

GEBCO. 2015. *The General Bathymetric Chart of the Oceans* [Online]. Available: http://www.gebco.net/ December 2015].

GENEST, C. & FAVRE, A. 2007. Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of Hydrologic Engineering,* 12**,** 347-368.

GENEST, C., MASIELLO, E. & TRIBOULEY, K. 2009a. Estimating copula densities through wavelets. *Insurance: Mathematics and Economics,* 44**,** 170-181.

GENEST, C., RÉMILLARD, B. & BEAUDOIN, D. 2009b. Goodness-of-fit tests for copulas: A review and a power study. *Insurance: Mathematics and Economics,* 44**,** 199-213.
GOOGLE EARTH.
GOOGLE EARTH. 2015. *Google Earth,* [Online]. [Accessed October, 15th 2015].
GRUBER, L. F. & CZADO, C. 2014. Bayesian model selection of regular vine copulas. *In:* LANZARONE, E. & IEVA, F. (eds.) *The Contribution of Young Researchers to Bayesian Statistics: Proceedings of BAYSM2013*. Cham: Springer International Publishing.
HAMBY, D. M. 1994. A review of techniques for parameter sensitivity analysis of environmental models. *Environmental Monitoring and Assessment,* 32**,** 135-154.
HAMBY, D. M. 1995. A comparison of sensitivity analysis techniques. *Health Physics,* 3**,** 195-204.
HELTON, J. C. 1993. Uncertainty and sensitivity analysis techniques for use in performance assessment for radioactive waste disposal. *Reliability Engineering and System Safety,* 42**,** 327-367.
HELTON, J. C., DAVIS, F. J. & JOHNSON, J. D. 2005. A comparison of uncertainty and sensitivity analysis results obtained with random and Latin hypercube sampling. *Reliability Engineering & System Safety,* 89**,** 305-330.
HENDRY, M. D. 1982. *The structure, evolution and sedimentology of the reef, beach and morass complex at Negril, Western Jamaica*, Petroleum Corporation of Jamaica.
HOLTHUIJSEN, L. H. 2007. Waves in oceanic and coastal waters.
KALLENBERG, W. C. M. 2009. Estimating copula densities, using model selection techniques. *Insurance: Mathematics and Economics,* 45**,** 209-223.
KAMPHUIS, J. W. Proceedings of the short course on design and reliability of coastal structures. 23rd ICCE, 1992 Venice, Italy.
KOMEN, G. J., HASSELMANN, K. & HASSELMANN, K. 1984. On the existence of a fully developed wind-sea spectrum. *Journal of Physical Oceanography***,** 1271-1285.
KUROWICKA, D. & COOKE, R. 2006. *Uncertainty analysis with high dimensional dependence modelling*, John Wiley & Sons, Ltd.
LIGHTTWIST SOFTWARE. Available: http://www.lighttwist.net/.
MATALA, A. 2008. Sample size requirment for Monte Carlo simulations - using Latin Hypercube sampling.
MCKAY, M. D., BECKMAN, R. J. & CONOVER, W. J. 2000. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics,* 42**,** 55-61.
MCKENZIE, A. 2012. Beach responses to hurricane impacts: A case study of Long Bay Beach, Negril, Jamaica. *Caribbean Journal of Earth Science,* 43**,** 51-58.
MITCHELL, S. F. 2012. Editorial - Beach erosion at Negril. *Caribbean Journal of Earth Science,* 43**,** 1-2.
MONDON, E. M. & WARNER, P. S. 2012. Evaluation of a proposed solution to beach erosion at Negril. *Caribbean Journal of Earth Science,* 43**,** 11-23.
MORALES-NÁPOLES, O. 2015a. Bivariate dependences (copulas). A powerpoint presentation. .
MORALES-NÁPOLES, O. 2015b. Structured expert judgment. A powerpoint presentation. .
MORGAN, M. G., HENRION, M. & SMALL, M. 1992. *Uncertainty. A guide to dealing with uncertainty in quantitative risk and policy analysis,* Cambridge, Cambridge University Press.
MORRIS, M. D. 1991. Factorial sampling plans for preliminary computational experiments. *Technometrics,* 33**,** 161-174.
MUTHÉN, L. K. & MUTHÉN, B. O. 2002. How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling: A Multidisciplinary Journal,* 9**,** 599-620.
NASA. Available: http://landsat.gsfc.nasa.gov/.
NOAA. 2015. *National Oceanic and Atmospheric Administration* [Online]. Available: http://www.noaa.gov/ [Accessed November 20th 2015].
OLIJ, D. J. C. 2015. *Wave climate reduction for medium term process based morphodynamic simulations.* MSc, Delft University of Technology.
PHILIPS, M. R. & JONES, A. L. 2006. Erosion and tourism infrastructure in the coastal zone: Problems, consequences and management. *Tourism Management,* 27**,** 517-524.
RHINEY, K. C. 2012. The Negril tourism industry: growth, challenges and future prospects. *Caribbean Journal of Earth Science of The Total Environment,* 43**,** 25-34.
ROBINSON, E. & HENDRY, M. 2012. Coastal change and evolution at Negril, Jamaica: A geological perspective. *Caribbean Journal of Earth Science,* 43**,** 3-9.
ROBINSON, E., KHAN, S., COUTOU, R. & JOHNSON, M. 2012. Shoreline changes and sea-level rise at Long Bay, Negril, western Jamaica. *Caribbean Journal of Earth Science,* 43**,** 35-49.
ROELVINK, J. A. & WALSTRA, D. J. R. 2004. Keeping it simple by using complex models. *Int. Conference on Hydroscience and Engineering (IHCE-2004).* Brisbane, Australia.
ROSE, C. & SMITH, M. D. 2012. Multivariate Distributions, Ch. 6. *Mathematical statistics with mathematica.* New York: Sprinter-Verlag.
ROSE, K. A. 1983. A Simulation comparison and evaluation of parameter sensitivity methods applicable to large models. *In:* WILLIAM K. LAUENROTH, G. V. S. & MARSHALL, F. (eds.) *Developments in Environmental Modelling.* Elsevier.
ROYAL HASKONINGDHV 2013. Quick scan consideration and cases for sandy strategies. Amersfoort: Fiselier, J.; Henrotte, J. and Stuij, S.

SALLABERRY, C. J., HELTON, J. C. & HORA, S. C. 2008. Extension of Latin hypercube samples with correlated variables. *Reliability Engineering & System Safety,* 93**,** 1047-1059.

SALTELLI, A. 1999. Sensitivity analysis: Could better methods be used? *Journal of Geophysical Research,* 104**,** 3789-3793.

SALTELLI, A. 2004. Global sensitivity analysis: An introduction. *Los Alamos National Laboratory.*

SALTELLI, A., TARANTOLA, S. & CHAN, K. P.-S. 1999. A quantitative model-independent method for global sensitivity analysis of model output. *Technometrics,* 41**,** 39-56.

SCHWEIZER, B. 2007. Introduction to copulas. *Journal of Hydrologic Engineering,* 12**,** 346.

SMITH WARNER INTERNATIONAL LTD. 2007. Preliminary engineering report for beach restoration works at Negril.

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**,** 1001-1008.

SWAN 2015. SWAN user manual.

TEN HAM, D., HENROTTE, J., KRAAIJEVELD, R., MILOSEVIC, M. & SMIT, P. 2006. Erosion Negril Beach.

THE WEATHER COMPANY. 2016. *Weather underground* [Online]. Available: https://www.wunderground.com.

TORRES, R. R. & TSIMPLIS, M. N. 2011. Tides and long-term modulations in the Caribbean Sea. *Journal of Geophysical Research,* 116.

UNEP 2010. Linking ecosystems to risk and vulnerability reduction: The case of Jamaica.

VAN ARKEL, M. 2015. *Capturing value from sandy strategies.* MSc. thesis - Wageningen University.

VAN DER KLIS, H. 2003. *Uncertainty analysis applied to numerical models of river bed morphology.* PhD dissertation - Delft University of Technology.

VAN DER SPEK, B. J. T. 2013. *Data poor environments. Uncertainty propogation in hydrodynamic modelling.* Master of Science, MSc. thesis - Delft University of Technology.

VAN GELDER, P. H. A. J. M. 2000. *Statistical methods for the risk-based design of civil structures.* PhD, PhD dissertation - Delft University of Technology.

VAN RHEE 2012. Handreiking adaptief deltamanagement - definitief concept. Leiden: Stratelligence.

VAN RIJN, L. C. 1997. Sediment transport and budget of the central coastal zone of Holland. *Coastal Engineering,* 32**,** 61-90.

VAN STEIJN, P. 2015. *Global assessment on the lifetime of mega nourishments.* Msc. , Delft University of Technology.

WALSTRA, D. J. R., HOEKSTRA, R., TONNON, P. K. & RUESSINK, B. G. 2013. Input reduction for long-term morphodynamic simulations in wave-dominated coastal settings. *Coastal Engineering,* 77**,** 57-70.

WANG, H., ZHU, J. & YANG, J. 2014. Error analysis on ESA's envisat ASAR wave mode significant wave height retrievals using triple collocation model. *Remote Sensing,* 6**,** 12217-12233.

WERNER, C., HANEA, A. M., MORALES-NÁPOLES, O., BEDFORD, T. & COOKE, R. M. Future. Expert judgment for dependence modelling: A systematic literature review and future research directions.

ZHANG, K., DOUGLAS, B. & S., L. 2004. Global warming and coastal erosion. *Climatic Change,* 64**,** 41-58.

A. Statistics

A.1 Probabilistic distributions

A.1.1 Generalized Pareto distribution

The probability density function for the generalized Pareto distribution with shape parameter k ≠ 0, scale parameter σ and threshold parameter θ is (Dekking et al., 2005):

$$y = f(x|k,\sigma,\theta) = \frac{1}{\sigma}\left(1 + k\,\frac{(x-\theta)}{\sigma}\right)^{-1-\frac{1}{k}} \quad \text{A.1}$$

A.1.2 Extreme value distribution

The probability density function for the extreme value distribution with location parameter μ and scale parameter σ is (Dekking et al., 2005):

$$y = f(x|\mu,\sigma) = \sigma^{-1}\exp\left(\frac{x-\mu}{\sigma}\right)\exp\left(-\exp\left(\frac{x-\mu}{\sigma}\right)\right) \quad \text{A.2}$$

A.1.3 Generalized extreme value distribution

The probability density function for the generalized extreme value distribution with location parameter μ, scale parameter σ and shape parameter k ≠ 0 is (Dekking et al., 2005):

$$y = f(x|\mu,\sigma) = \frac{1}{\sigma}\exp\left(-\left(1 + k\frac{x-\mu}{\sigma}\right)^{\frac{-1}{k}}\right)\left(1 + k\frac{x-\mu}{\sigma}\right)^{-1-\frac{1}{k}} \quad \text{A.3}$$

A.1.4 Uniform distribution

The probability density function for the uniform distribution with a the lower value and b the upper value is (Dekking et al., 2005):

$$y = f(x|a,b) = \frac{1}{b-a} \quad \text{A.4}$$

A.2 Correlations

A.2.1 Expectation

The expectation of a continuous random variable X with probability density function f is the number (Dekking et al., 2005):

$$E[X] = \int_{-\infty}^{\infty} x\,f(x)dx \quad \text{A.5}$$

E[X] is also called the expected value or mean of X.

A.2.2 Variance and standard deviation

The spread around the mean of a random variable is measured by the expected squared deviations from the mean. This is called the variance. The variance Var(X) of a random variable X is the number (Dekking et al., 2005):

$$Var(X) = E[(X - E[X])^2] = E[X^2] - (E[X])^2 \quad \text{A.6}$$

The variance of a random variable is always positive or zero. Furthermore is valid (Dekking et al., 2005):

$$\sqrt{Var(X)} = \sigma_X \quad \text{A.7}$$

Where σ is the standard deviation of X.

A.2.3 Covariance

The covariance expresses to some extent the way two random variables, X and Y, influence each other. The covariance between X and Y is defined by (Dekking et al., 2005):

$$Cov(X,Y) = E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y] \quad \text{A.8}$$

A.2.4 Pearson's linear correlation coefficient

The Covariance may not be always suitable to express the dependence between C and Y. For this reason there is a standardized version of the covariance called the correlation coefficient of X and Y (Dekking et al., 2005). The product moment correlation, also called linear or Pearson correlation, of X and Y, with finite expectations E(X),E(Y) and finite variances σ_X^2 σ_Y^2 is defined as (Kurowicka and Cooke, 2006):

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} \quad \text{A.9}$$

A.2.5 Spearman's rank correlation coefficient

The (Pearson) correlation gives insight in the strength of linear relationship between two variables. However for non-linear relationships it may not give a correct measure for statistical dependence. A method to reduce the effect of non-linear data is the use of rank transformation (Hamby, 1994). Rank order correlation or the Spearman correlation coefficient is a measure for the strength of (non-linear) monotonic relations (Hamby, 1994). The observations of the two parameters, (x; y) are given a specified rank (rank transformation) and from this the correlation is computed. The rank correlation of random variables X and Y with cumulative distributions F_X and F_Y is given by (Kurowicka and Cooke, 2006):

$$\rho_r(X,Y) = \rho(F_X(X), F_Y(Y)) \quad \text{A.10}$$

A.2.6 Partial correlation coefficients

Partial correlation measures the degree of association between two random parameters with the effect of a set of controlling random parameters removed. Partial correlations can be computed from correlation with the following recursive formula (Kurowicka and Cooke, 2006):

$$\rho_{12;3,\ldots,n} = \frac{\rho_{12;3,\ldots,n-1} - \rho_{1n;3,\ldots,n-1} * \rho_{2n;3,\ldots,n-1}}{\sqrt{1 - \rho^2_{1n;3,\ldots,n-1}}\sqrt{1 - \rho^2_{2n;3,\ldots,n-1}}} \quad \text{A.11}$$

Where n is the number of uncertain parameters involved.

A.3 Kernel density estimation

The Kernel density estimation is a way to estimate the probability density function of a random parameters. For this it is necessary to choose a Kernel K and a tuning parameters h, called the bandwidth. A kernel K typically satisfies the following conditions:

- K is a probability density
- K is symmetry around zero
- For parameters larger, in absolute sense, than one K is zero.

Suppose the construction of a kernel density estimate $f_{n,h}$ for a dataset $x_1,x_2,\ldots,x_n$. First, the kernel K need to be scaled into the function. Next, a scaled kernel is put around each elemant x_i in the dataset. The graphs of the shifted kernels overlap whenever x_i and x_j are close to each other. Now, the kernel density estimate is $f_{n,h}$ is constructed by summing the scaled kernels and dividing them by n (Dekking et al., 2005):

$$f_{n,h}(t) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{t-x_i}{h}\right) \quad \text{A.12}$$

With h the tuning parameter, i.e. the bandwidth and t the scaling parameter.

The bandwidth h needs to be chosen by trial and error: A too small bandwidth produces a curve with many insolated peaks, a too large bandwidth produces a very smooth curve but may smoothing away important features of the data. A formula resulting from recent research, however, may effectively be used is $h=1.06*s*n^{-1/5}$, where s denotes the sample standard deviation. For this study built-in functions of Matlab has been used to determine the kernel density estimation.

A.4 Dependence modeling – Vines

A tree is an *acyclic* undirected graph. Dependence trees are used to describe dependence structures in high dimensional distributions. A tree has nodes N and edges E. Ad edge is a subset of unordered pairs of N with no cycle. The degree of a note denotes the number of edges attached on the node.

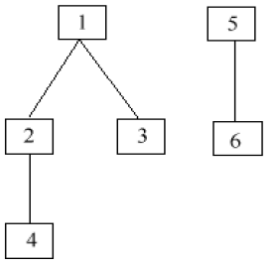

Figure A.1- A tree with six nodes (Kurowicka and Cooke, 2006)

Edges of trees can be used to specify bivariate dependencies. Copula trees are special cases of bivariate trees. The dependence between two parameters is presented by a rank correlation, the correlation of the copula, together with a family of copula indexed by correlation. We specify a copula tree by assigning one-dimensional distribution functions to the nodes, and rank correlations in [−1, 1] to the edges. Figure A.2 shows a tree with six nodes with rank correlations assigned to the edges. The values of parameter one and five are selected independently of all other parameters. The values of parameters two, three and six, however, are dependent on the values of parameter one and five respectively. The value of parameter four is dependent on the value of parameter two, and hence indirectly of one (Gruber and Czado, 2014).

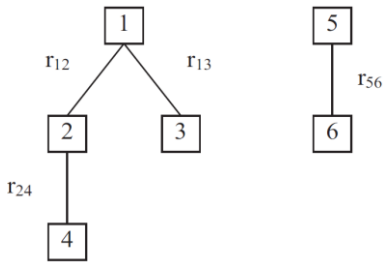

Figure A.2 - A tree with six nodes with rank correlations assigned to the edges (Kurowicka and Cooke, 2006)

A vine of n parameters is a nested set of trees, where the edges of tree j are the nodes of tree j+1, and each tree has the maximum number of edges. A regular vine of n variables is a vine in which two edges in tree j are joined by an edge in tree j+1, but only if these edges share a common node. There are two families of regular vines. Here the focus is on the Canonical vine. The vine is canonical if each tree T_i has a unique node of degree n-i

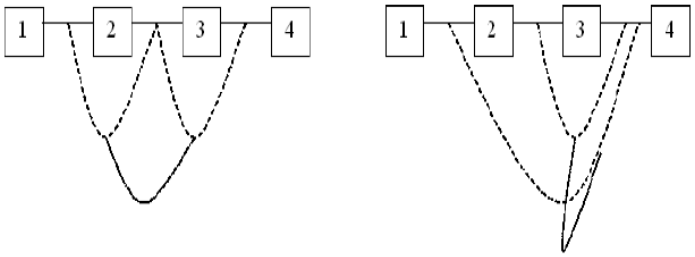

Figure A.3 - A regular vine (left) a non-regular vine (right) (Kurowicka and Cooke, 2006)

The basic idea for copula vines is to enrich the tree in Figure A.2 with additional information on the conditional dependence between parameter two and three, given parameter one. This is shown in Figure A.4.

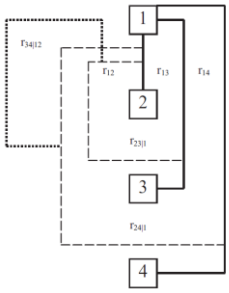

Figure A.4 - The canonical vine on 4 variables with (conditional) rank correlations assigned to the edges (Kurowicka and Cooke, 2006)

To sample dependent parameters a special sampling procedure need to be followed. This procedure requires independent uniform parameters, between which dependence is induced by transforming the parameters in a way which uses previously samples values (Kurowicka and Cooke, 2006). The transformations are the conditional cumulative and the inverse cumulative distributions of the copulas in an appropriate vine representation (Kurowicka and Cooke, 2006).

Here, the sampling algorithm for the canonical vine on four parameters is presented (to sample from a vine as presented in Figure A.4. This algorithm samples four independent uniform (0,1) parameters $u_1,\ldots,u_4$. The parameters $x_1,\ldots,x_4$ are assumed to be uniform as well. The conditional correlation between parameters (i,j) given k is denoted as $r_{i,j|k}$. $F^{-1}_{r_{13};\,u_1}(x_j)$ denotes the inverse cumulative distribution function for x_j given u_i under the conditional copula with correlation $r_{i,j|k}$. The algorithm can be stated as follows (Kurowicka and Cooke, 2006):

$$x_1 = u_i \qquad \text{A.13}$$

$$x_2 = F^{-1}_{r_{12};\,u_1}(u_2)$$

$$x_3 = F^{-1}_{r_{13};\,u_1}(F^{-1}_{r_{23|1};\,u_2}(u_3))$$

$$x_4 = F^{-1}_{r_{14};\,u_1}\left(F^{-1}_{r_{24|1};\,u_2}\left(F^{-1}_{r_{34|12};\,u_3}(u_4)\right)\right)$$

The uniform parameters $u_1,\ldots,u_4$ are sampled independently. The dependent parameters $x_1,\ldots,x_4$, are obtained by applying successive inverse cumulative distributions.

A.5 Estimation of copula through semi-correlations

Through the following steps a copula can be estimated through semi-correlations (Morales-Nápoles, 2015a):

- Transform the data to the copula scale (unit square) using a kernel estimator of the cumulative distribution function.
- Transform this data to standard normal
- Compute the total correlation of this data and the correlation of data distinguished in four quadrants (North-East, North-West, South-West, South-East)
- Fit several parametric copula to the transformed standard normal data.
- Draw a sufficiently big amount of samples from these fitted copula
- Compute the total correlation of this data, and the correlation of the copula samples distinguished in the defined quadrants.
- Compare these correlations to the correlations of the original data and consider what copula to use best.

B. Case specific data

B.1 Data of measurements

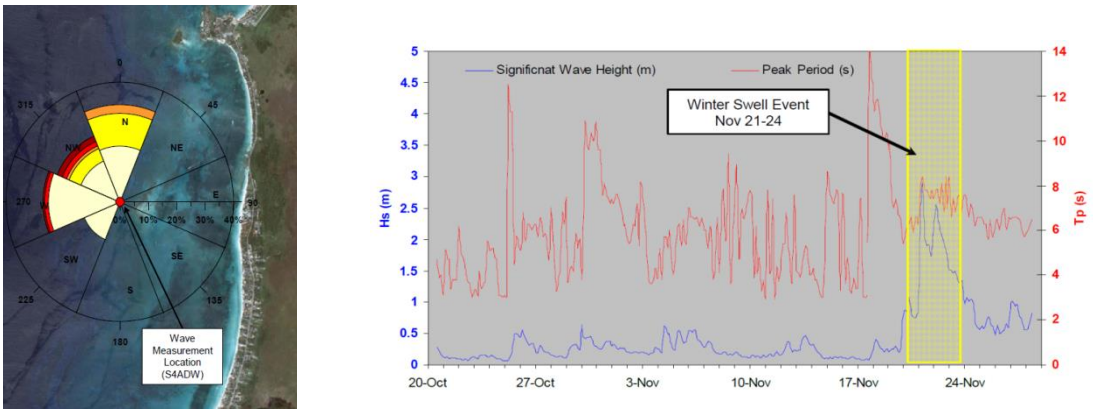

Figure B.1 - Wave measurements between October 20th and November 28th 2006 (Smith Warner International Ltd., 2007)

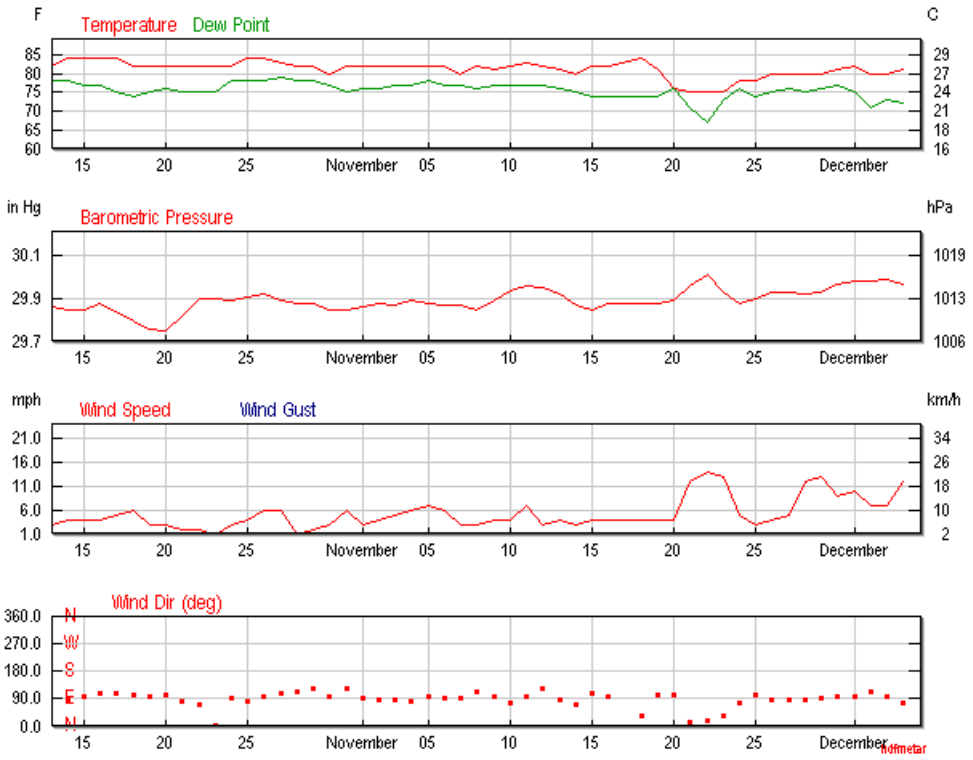

Figure B.2 – Wind measurements at Montego Bay (The Weather Company, 2016)

B.2 Results of modeling measurements

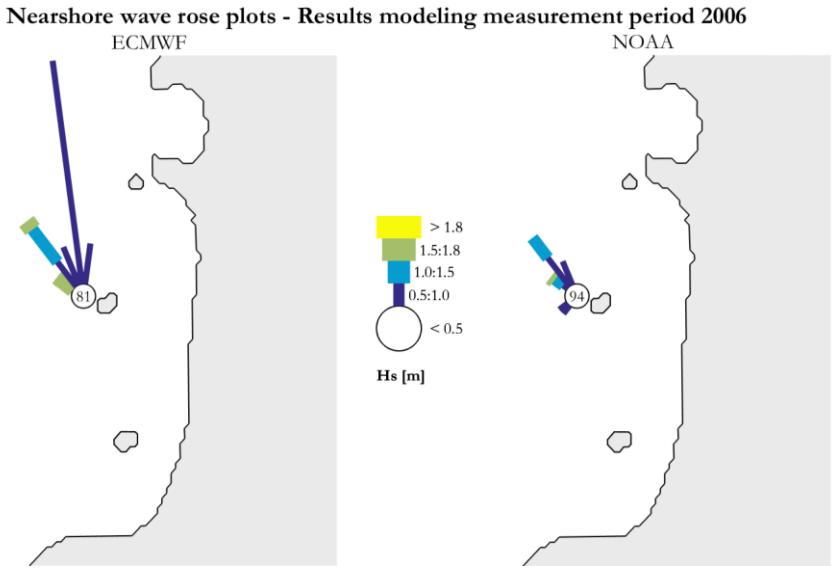

Figure B.3 - Nearshore wave rose plots of results modeling the measurement period

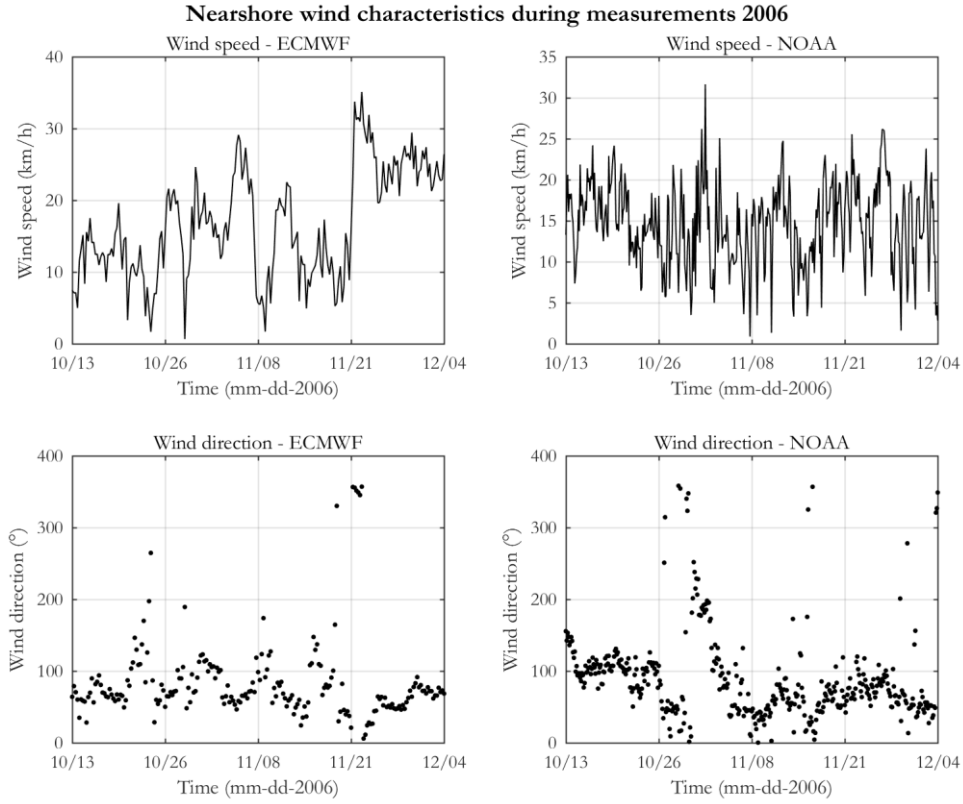

Figure B.4 - Resulting nearshore wind speed and direction of modeling the measurement period

C . Expert judgment on uncertainties

This Appendix presents the expert judgment on uncertainties for the case Negril Coast. Section C.1 presents the final questionnaire used. Section C.2 presents the results of the questionnaires. Section C.3 draws the conclusions from the questionnaire.

C.1 Questionnaire

Section 1. Introduction

In my research on sensitivity analyses uncertainties of several variables are taken into account. Since the lack of data on these variables, expert judgment is necessary to include these uncertainties correctly. The table in section 4 maps these expert judgments. The first column indicates the variable of interest, with the adequate units. An explanation of these variables can be found in section 2 below. In the next three columns the following values have to be filled:

- A value of which it is unlikely that the value of the variable would be lower. Unlikely indicates a probability of 5%.
- A value of which the probability is 50% that the value of the variable is lower, and of which the probability is 50% that the value of the variable is higher.
- A value of which it is unlikely that the value of the variable would be higher. Unlikely indicates a probability of 5%.

Section 3 provides two examples

Section 2. Explanation variables of interest

Variable 1 – 4 : Wave & wind data

Hindcasting of metocean data using meteorological charts is a widely used alternative for direct measurements of wind or waves. Data obtained from hindcasting studies contain some error when compared against measured data, due to simplification and assumptions in the analytical model used and due to measurement errors in the input data (Brooker et al., 2004). The error can be subdivided into a bias and a random variation. The bias can be eliminated by calibrating hindcast model results through a measurement in the area, but the random variation remains. Therefore the focus is on the bias of the metocean data.

Some typical characteristics of the metocean data:

- Almost all the waves are sea waves, with a mean wave height around 1 m.
- The average wave period is 7.3 s.
- The average wind speed is around 5 m/s.

Note i: The bias can be both positive and negative.
Note ii: In case the global wave model would overestimate the wave height, the bias is positive. In case the global wave model would underestimate the wave height, the bias is negative.

Variable 5. Uncertainty in coral reef height (m)

The depth used in the models is obtained from a boat-based bathymetric survey, using an echo-sounder and a GPS devise. **The coral reef is located in a water depth of around 5 m.** The question reads as follows:

5. What is the uncertainty in the height of the coral reef?

Note: When the height of the coral reef is overestimated, a positive value has to be filled. This would mean that the height of the coral reef is lower than the data indicates.

Variable 6. Bottom friction (m^2/s^3)

The bottom friction is used in the SWAN model. The question maps the **actual value** of the bottom friction:

6. What do you expect to be the value of the bottom friction?

Section 3. Examples

Example 1 - What is the bias of the hindcasts of the offshore wave height?

Variable of interest	Unlikely that the value is lower than	Expected value	Unlikely that the value is higher than
Bias hindcasts offshore wave height (m)	-1.5	0.5	2

From this one expert the following conclusions would be drawn:

- 5% probability that the bias is lower than 1.5 m.
- The expected value of the bias is 0.5 m.
- 5% probability that the bias is higher than 2 m

This expert expects the bias to be mostly positive. This would mean that it is most probable that the actual wave height is lower than the global wave model indicates.

Example 2 - What is the uncertainty of the coral reef height?

Variable of interest	Unlikely that the value is lower than	Expected value	Unlikely that the value is higher than
Uncertainty of the coral reef height (m)	-1	0	1

From this one expert the following conclusions would be drawn:

- 5% probability that the uncertainty is lower than -1 m
- The expected value of the uncertainty is 0 m
- 5% probability that the uncertainty is higher than 1 m

This expert expects the uncertainty to be 50% positive and 50% negative. When the uncertainty of the coral reef height is +1 m, it means that the actual coral reef height is 1 m lower than the survey indicates.

Section 4 – Fill in tables

Variable of interest	Unlikely that the value is lower than	Expected value	Unlikely that the value is higher than
1. Bias hindcasts offshore wave height (m)			
2. Bias hindcasts offshore wave and wind direction (°)			
3. Bias hindcasts offshore wave period (s)			
4. Bias hindcasts offshore wind speed (m/s)			
5. Uncertainty coral reef height – from a survey (m)			
6. Bottom friction (m^2/s^3)			

Variable 1	Variable 2	Correlation
Bias of wave height	Bias of wave period	
Bias of wave height	Bias of wind speed	
Bias of wave period	Bias of wind speed	
Bias of wave direction	Bias of wind direction	

C.2 Questionnaire replies by experts

Table C.1 and Table C.1 present the results of the questionnaires

Table C.1 – Results of questionnaire, uncertainties

Variable of interest	Expert	Unlikely that the value is lower than	Expected value	Unlikely that the value is higher than
1. Bias hindcasts offshore wave height (m)	1	-0.5	0	+0.5
	2	-1	0.25	1
	3	-20%	-10%	0
	4	-1.5	-0.5	1.5
2. Bias hindcasts offshore wave and wind direction (°)	1	-15	0	15
	2	-10	0	10
	3	-15	0	15
	4	-15	0	15
3. Bias hindcasts offshore wave period (s)	1	-2	0	1
	2	-2	2	4
	3	0%	10%	15%
	4	0	2	4
4. Bias hindcasts offshore wind speed (m/s)	1	-2.5	0	2.5
	2	-3	3	6
	3	-15%	-10%	0
	4	-3	0	+3
5. Uncertainty in coral reef height (m)	1	-0.5	0	0.5
	2	-0.5	0.3	1
	3	-	-	-
	4	-0.7	0	0.7
6. Bottom friction (m^2/s^3)	1	0.02	0.038	0.067
	2	0.01	0.038	0.05
	3	-	-	-
	4	0.01	0.065	0.1

Table C.2 – Results of questionnaire, correlations

Variable 1	Variable 2	Expert	Correlation
Bias hindcasts offshore wave height (m)	Bias hindcasts offshore wave period (s)	1	Positive in case of sea
		2	Positive
		3	0.9
		4	0.8
Bias hindcasts offshore wave height (m)	Bias hindcasts offshore wind speed (m/s)	1	Positive
		2	Positive
		3	0.9
		4	0.8
Bias hindcasts offshore wave period (s)	Bias hindcasts offshore wind speed (m/s)	1	Positive in case of sea
		2	Negative
		3	0.85
		4	0.8
Bias hindcasts offshore wave direction (°)	Bias hindcasts offshore wind direction (°)	1	Positive in case of sea
		2	No correlation in case of swell
		3	0.7
		4	0.9

C.3 Conclusions

The following two tables present the results of the questionnaire. These results are computed with the help of Excalibur. The lowest (a) and the highest value (b) of a parameters are computed using (Morales-Nápoles, 2015b):

$$a = L - k\left(\frac{U-L}{100}\right) \quad \text{C.1}$$

$$b = U + k\left(\frac{U-L}{100}\right) \quad \text{C.2}$$

In which k defines the overshoot rule, which is typically 10% and in which [L,U] is the smallest interval containing all assessed quartiles of all experts.

The figures at page 87 and 88 present the estimated probability density functions of the uncertain parameters.

Table C.3 – Uncertainties indicated by experts

Variable of interest	a	< 5%	5 - 50%	50 - 95%	b
1. Bias hindcasts wave height (m)	-1.5	-1.3	-0.1	1.2	1.5
2. Bias hindcasts wave & wind direction (°)	-15	-14.7	0	14.7	15
3. Bias hindcasts wave period (s)	-2	-1.8	0.8	3.9	4
4. Bias hindcasts wind speed (m/s)	-0.8	-0.75	-0.5	0	0.05
5. Uncertainty in coral reef height (m)	-0.7	-0.6	0.1	0.9	1

Table C.4 – Correlations indicated by experts

Variable 1	Variable 2	Correlation
Bias hindcasts offshore wave height (m)	Bias hindcasts offshore wave period (s)	0.85
Bias hindcasts offshore wave height (m)	Bias hindcasts offshore wind speed (m/s)	0.85
Bias hindcasts offshore wave period (s)	Bias hindcasts offshore wind speed (m/s)	0.83
Bias hindcasts offshore wave direction (°)	Bias hindcasts offshore wind direction (°)	0.8

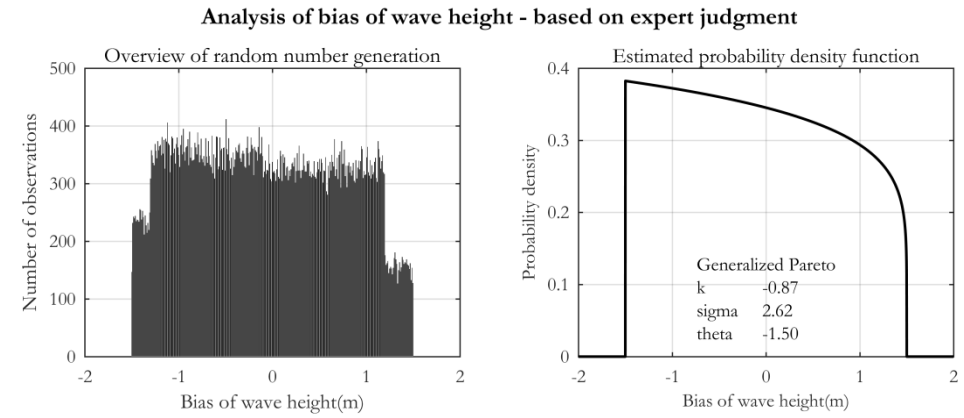

Figure C.1 – Analysis of bias of wave height

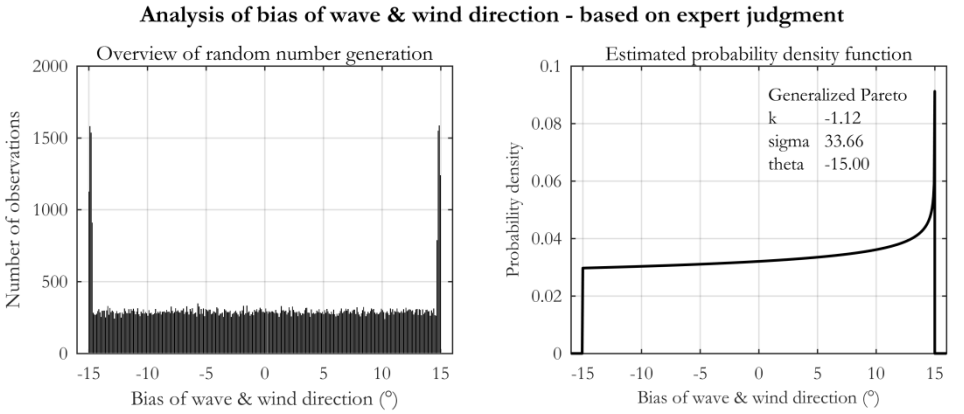

Figure C.2 – Analysis of bias of wave and wind direction

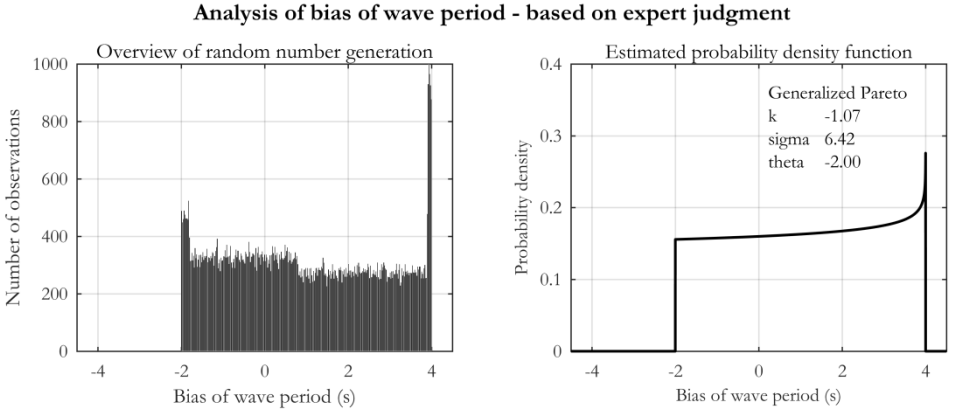

Figure C.3 – Analysis of bias of wave period

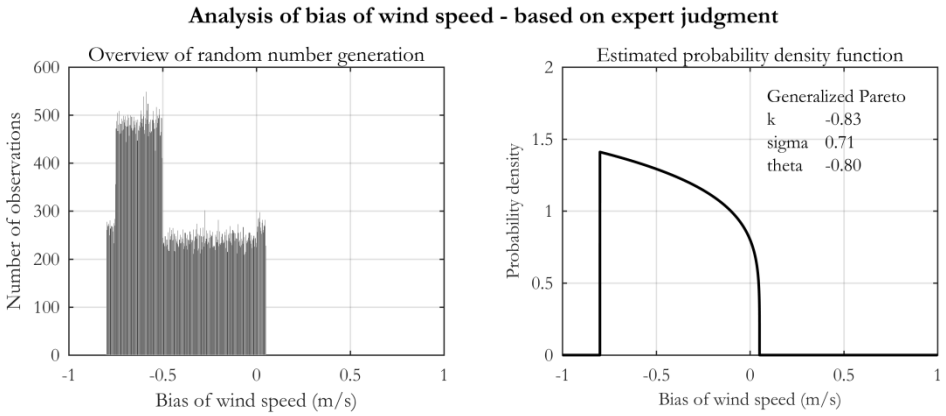

Figure C.4 - Analysis of the bias of the wind speed

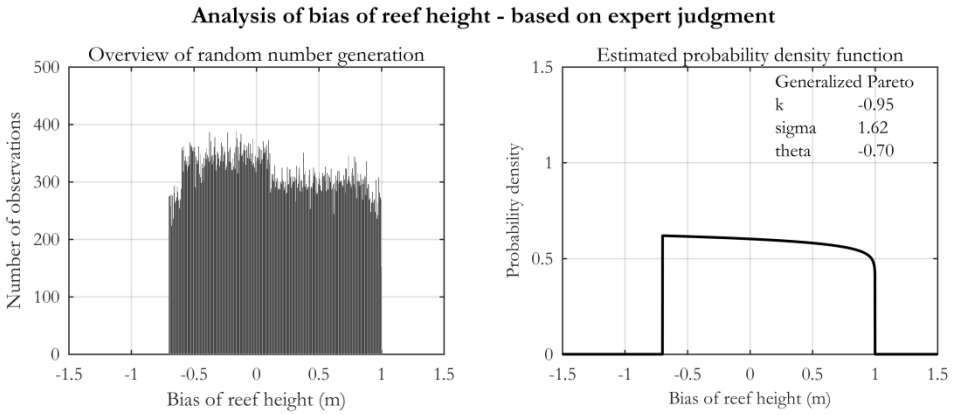

Figure C.5 – Analysis of the bias of the reef height

D. Wave input reduction method

D.1 K-harmonic means method

In the description the following standard notations are used:

- A wave conditions x_i describes a wave condition with multiple variables. This is notated as: $x_i = \{ x_{i,1}, , x_{i,2} \ldots x_{i,z} \}$, where z is the number of variables that is used to describe a wave condition, for instance wave height, period and direction.
- All the waves in the full offshore wave dataset are collected in the database $X = \{x_1, x_2 \ldots x_N \}$, where N is the total number of waves in the full dataset.
- The reduced wave climate consists of k wave conditions. These are called centroids. A centroid I, like a wave condition, described by multiple variables. This is notated as: $v_i = \{ v_{i,1}, , v_{i,2} \ldots v_{i,z} \}$, where z is again the number of variables that is used to describe a condition or centroid.
- The centroids are collected in the database $X = \{v_1, v_2 \ldots v_k \}$, with k the number of centroids.

The following steps describe the K-harmonic means method in detail (Olij, 2015):

1. The first centroid is one wave of the offshore wave data base. This wave is *randomly* selected.
2. The distances from every wave to the first centroid are computed. For this a distance measure is used. A distance measure is a technique to determine the distance between two points in a multivariate space. The distance determines how similar observations are to each other: the smaller the distance between two observations the more similar they are. The K-harmonic means method uses *Euclidean distances D*. The Euclidean distance is generally given by:

$$D_{i,j} = \|x_i, v_j\| = \sqrt{(w_1 R_1 (x_{i,1}, v_{j,1}))^2 + (w_2 R_2 (x_{i,2}, v_{j,2}))^2 + (w_z R_z (x_{i,z}, v_{j,z}))^2} \quad \text{D.1}$$

   Where the two vertical bars indicate a distance measure, where R is the distance measure of a variable that is used to describe a wave conditions, and where w is a user defined weight. This research uses a weight of 1 for all distance measures. Two examples of a distance measure are R_H, the distance measure for the wave height and R_φ the distance measure for the direction:

$$R_H(x_i, v_j) = H_i' - H_j^{rep\prime} \quad \text{D.2}$$

$$R_\varphi(x_i, v_j) = \min(\|\varphi_i' - \varphi_j'\|, 2 - \|\varphi_i' - \varphi_j'\|) \quad \text{D.3}$$

   Where the accents indicate that the variables are normalized as follows:

$$H' = \frac{H - H^{min}}{H^{max} - H^{min}} \quad \text{D.4}$$

$$\varphi^i = \frac{\varphi}{\pi} \quad \text{D.5}$$

3. The remaining *k-1* centroids are selected with a probability based on *D*. For instance for the second centroid:

$$P(x_i = v_2) = \frac{{D_{i,1}}^2}{\sum_{j=1}^{j=(N-1)} {D_{j,1}}^2}, i = 1 \ldots . N \quad \text{D.6}$$

4. The distances between all observations and all k centroids are computed.

5. Determine how much each observation belongs to every cluster through the *membership function*. The membership reads as follows:

$$M_{i,g} = \frac{{D_{i,j}}^{-2}}{\sum_{j=1}^{j=(N-1)} {D_{1,j}}^{-2}} \, , i = 1 \ldots . . N \qquad \text{D.7}$$

The weight of an observation is computed based on the membership. Some observations have more influence on the value determination of the centroids. This is called boosting. Boosting means that observations that are far from a representative wave condition in the observation space get more weight during the determination of the next location of the representative wave condition.

6. The new representative wave conditions of a cluster are calculated through averaging, using the weights, all wave conditions in a cluster. Note that for averaging wave direction the circular mean needs to be used:

$$\varphi_{mean} = \arg\left(\sum_{i=1}^{N} \exp(\mathrm{i}\, \varphi_i\right) = atan2(\sum_{i=1}^{N} \sin \varphi_i \, , \sum_{i=1}^{N} \cos \varphi_i \,) \qquad \text{D.8}$$

With i is the imaginary number.

7. The steps 4-6 are repeated a certain amount of times. In this research 500 iterations are used, since (Olij, 2015) indicates that with 500 iterations good results are obtained.

D.2 K-harmonic means method combined with the maximum dissimilarity algorithm

The maximum dissimilarity algorithm is a dissimilarity method and selects a sub-set of centroids that represents the full diversity of the full offshore wave dataset. This selection can replace the first four steps of the K-harmonic means method (previous section). This combination results in a *consistent* reduction method.

The maximum dissimilarity algorithm consists of the following steps (Olij, 2015, Camus et al., 2011):

1. Determine the distances of all waves to each other. The distance can be interpreted as the dissimilarity between the waves. The first initial centroid is the wave with the largest total distance to the other waves. The distance of all waves to each other is given as:

$$D_{i,j} = \|x_i, v_m\|^2 \quad with\ i = 1 \ldots N, m = 1 \ldots N \qquad \text{D.9}$$

2. Determine the distances of all the remaining waves to the first initial centroid. The observation with the largest dissimilarity is choses as the second initial centroid.

3. Determine the distance of all the remaining waves to the first and second centroid. The representative distance of a wave is the smallest distance to a centroid. The observation with the largest dissimilarity, so the largest smallest distance to a centroid, is selected as the third centroid. The largest smallest distance can be described as follows:

$$D_{i,Vj}^{min} = min\left[D_{i,v_j}, D_{i,V_{j-1}}\right] \qquad \text{D.10}$$

Where v indicates one centroid, and where V indicates the subset of centroids

4. The last step is repeated until k centroids are selected. These k centroids represent the full diversity of the full offshore wave dataset.