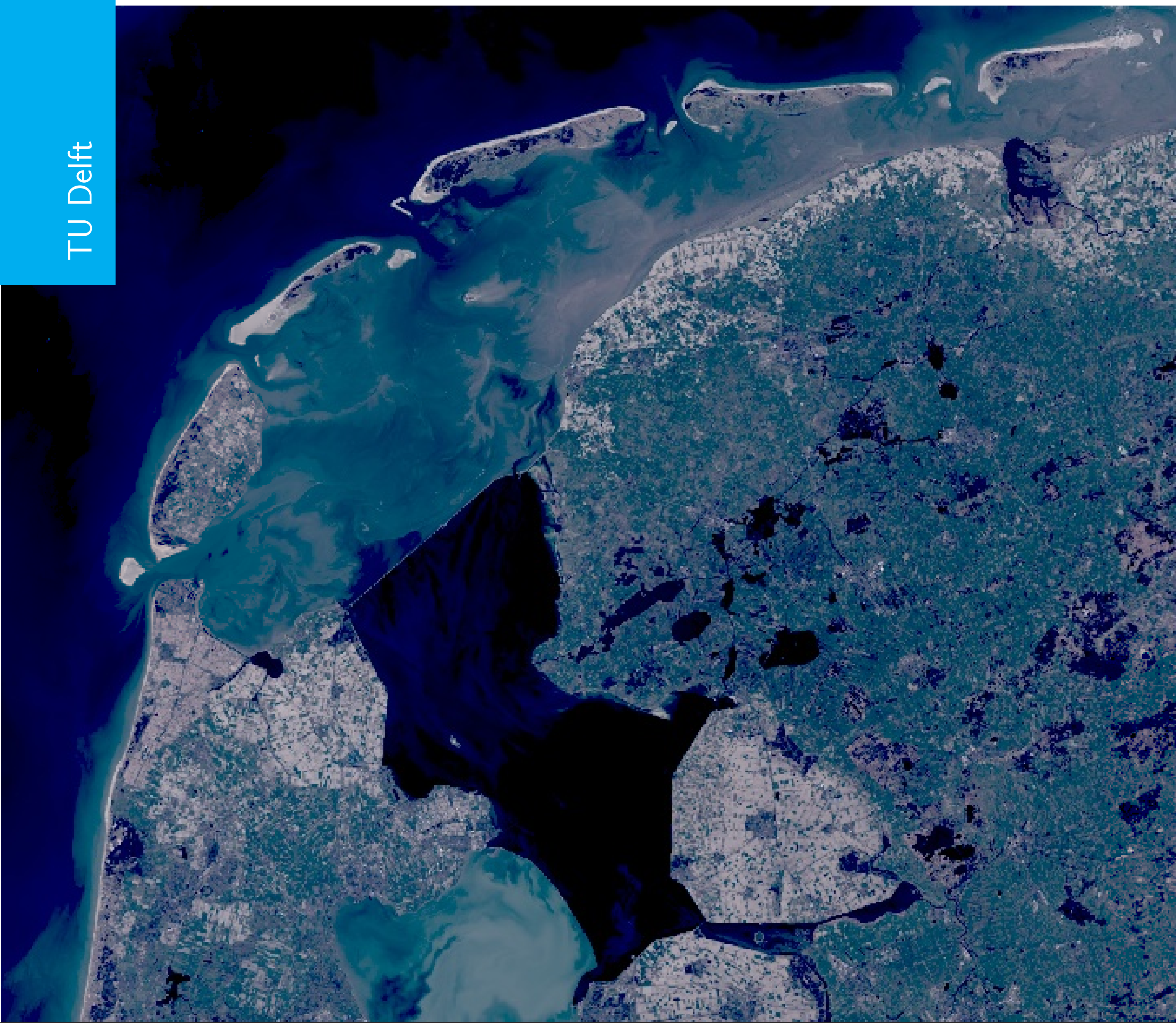
Approach to the use of uncertainty for management purposes

Quantification of the uncertainty in the turbidity for the ecosystem of the Wadden Sea

C.A.M. van de Vries

TU Delft



TU Delft Delft University of Technology

Challenge the future

Deltares
Enabling Delta Life

Cover image: satellite image of the Dutch Wadden Sea area, source: NASA, Landsat 8 mission.

Delft University of Technology

MSc Thesis

Approach to the use of uncertainty for management purposes - study Case: Quantification of the uncertainty in the turbidity for the ecosystem of the Wadden Sea

Cindy van de Vries

January 27, 2017

Student number:	4076141	
Thesis committee:	Prof. Dr. J.D. Pietrzak,	TU Delft
	Dr. Ir. M. Zijlema,	TU Delft
	Dr. Ir. G.Y.H. El Serafy,	Deltares, TU Delft

TU Delft Delft University of Technology

Deltares Enabling Delta Life

Acknowledgements

Writing this thesis was a very challenging and interesting experience, in which I have learned much about water quality, uncertainty and modeling. I made new contacts in the company (Deltares) and I learned to manage my own project. I could not have done it without my assessment committee, which have guided me each in their own way and with their own expertize. I want to thank Ghada el Serafy for offering me this opportunity, guiding me weekly through some difficult times, helping me on the parts of turbidity and uncertainty analysis and guaranteeing that the end product was something to be proud of. To Marcel Zijlema a thank you for your different and unique view on the topic which kept me sharp and keeping me on the right track with writing the report and the expectations from the university. And to Julie Pietrzak for believing in me and pushing me to get the best results out of this topic and sharing her great expertise on the SPM part. Although there were some bumps in the road, they helped me through everything and I am truly grateful for them to have as my committee.

Then a special thanks to the people who made the time writing the thesis enjoyable. To Alex Ziemba, who was always available for answering questions, reviewing, and giving lots of useful suggestions throughout the entire thesis, he did this with great patience and always a warm smile, thank you. Furthermore my fiancée Sherwin, in the most stressful times he always knew how to cheer me up and keep me going. And last but not least, my dear fellow students. Who put up with me every day, listening to my complaints, reviewing my work, but most of all, making my days memorable at Deltares and the previous years of my study. Thank you, Gerben, Melanie, Stef, Rufus, Sjoerd and Nova.

Cindy van de Vries
Delft, January 2017

Summary

Around the world, areas with unique ecosystems are prone to harmful factors deteriorating their environment. Management of these areas are responsible for the protection of the flora and fauna. To fulfill this task as best as possible they need knowledge on the ecological status of these areas. With this knowledge decisions can be made for future problems regarding the area. To do this, they need easily accessible information on indicators describing the conditions of these areas. Obtaining this information can be done using numerical models. The use of models however will introduce uncertainty in the information, which should be taken into account when using model output for the making of decisions. When model output is known together with its uncertainty, it can be efficiently visualized in a toolbox to make it commonly available for users. In this way the use of uncertainty can be incorporated in management purposes.

In this research, an uncertainty analysis is conducted and a visualization of model output together with uncertainty is made for the Wadden Sea area. This area is regarded as a Protected Area (PA) by the ECOPOTENTIAL project. In this area eutrophication and consequently algal bloom causes deterioration of the water quality. Due to Suspended Particulate Matter (SPM) in the water column, the incoming light is reduced, decreasing the amount of algal bloom. With a Delft3D-WAQ model, using the GEM/BLOOM module, the chlorophyll-a concentration can be modelled, which is an indicator of the amount of algae in the water column. The model Delft3D-WAQ Sediments is used to calculate the SPM concentrations, which is used as a forcing for the GEM/BLOOM model. Therefore, they are a main driving force and uncertainty source in the setup for this project. The main research question is formulated as: How can uncertainty from a SPM model as a driving force for a GEM/BLOOM model be identified, quantified and visualized to help decision makers?

The method to identify, quantify and visualize the uncertainty is described shortly. With the use of a literature review the uncertainty sources within the input files are identified. To quantify the uncertainty, first a sensitivity analysis and consequently an uncertainty analysis is used. This sensitivity analysis is done by varying the values of certain input files and assessing the variability it creates in the model output, resulting in the most influential input. The uncertainty analysis is used to obtain the magnitude of the uncertainty coming from this most influential input. This analysis is a Monte Carlo simulation, where different input is assessed by giving these a Probability Density Function (PDF) simulating the uncertainty in the input. From these distributions, samples are drawn to create different experiments to assess the influence of the uncertainty. Using a Latin Hypercube Sampling with Dependence instead of a random sampling, the amount of model runs is reduced from thousands to 188. The dependencies between the input parameters are taken into account. Afterwards, the output is estimated with a PDF from which the PDF characteristics mode and spreading are used to describe the SPM concentration and its uncertainty for each segment in time and space. A toolbox is developed for a 3D visualization of the model output and its uncertainty. A cubic shaped marker is placed in this environment for each segment by its x-, y- and z-coordinates. The SPM concentrations are visualized by coloring the markers and the uncertainty is incorporated using a white hue.

The parameters in the erosion and deposition fluxes in the model equations are identified to be most influential and are therefore assessed on uncertainty. The SPM results are validated by a comparison against measurement data from Rijkswaterstaat and data used in earlier studies for this model. Concluded is that the model gives a good approximation of the SPM concentrations. Some areas indicate a very high uncertainty, mainly where the SPM values are unrealistically high. Overall, the values are within the same order of magnitude as the validation data. The uncertainty in the model output is mainly present in critical areas, where the influence of different factors is significant, such as the effect of a river outflow, stratification and tidal influence. Such as the part of the Wadden Sea area near the discharge simulating the Lake IJssel. This indicates that either the bed load module of the model or the hydrodynamic input is not completely optimized yet. For the quantification of the uncertainty a log-normal distribution is used to estimate the output values at every location and time step. The characteristics of this distribution, the mode and the spreading, are used to quantify the concentrations of SPM and the uncertainty into values that can be used in the visualization.

To answer the research question: the uncertainties are identified with a literature review resulting in the sources of uncertainty, quantified with the uncertainty analysis into a mode and a spreading and visualized in a toolbox with a 3D environment using a marker for each segment and using color for displaying the concentrations and a white hue to visualize the uncertainty. This toolbox than helps decision makers to easily access the data and according uncertainty. The SPM model has some high uncertainties, but is a good estimation to be used as a driving force for the GEM/BLOOM model.

Contents

1 Introduction

1.1 Background information

Earth exists out of many different natural and ecologically valuable areas, with each their own unique consistency and inhabitants and due to their uniqueness should be protected against threats. A well known example of such an ecosystem are the coral reefs [Society MarineBio Conservation, 2015], which can be found amongst other near Australia. Not only do coral reefs have a tremendous biodiversity, their beauty attracts thousands of tourists. Coral reefs are extremely sensitive to changes in light, temperature, overfishing, damaging fishing practices, pollution, and excess sediment from development and erosion [Society MarineBio Conservation, 2015]. Without the help of human management for these areas, these ecosystems could be destroyed. Another example of such an area is the Wadden Sea area located in the North sea [Marencic, 2009], famous for the habitat of seals and birds. Here the ecosystem is heavily influenced by the algal bloom. In order to oppose the problems in these areas, regulations are set up by people monitoring and managing these areas. These regulations are for present and future state of the area, where decision making is needed to preserve the present state. Information regarding the indicators for status and the threats of the ecosystem is needed for knowledge-based decision making. A generic decision making toolbox that incorporates a quantification of ecological indicators and the uncertainty could therefore be very useful to be applied.

One project that has the objective to improve and monitor ecosystems in different Protected Areas (PAs) all over the world, is the ECOPOTENTIAL project [Ecopotential, 2015a]. One of the main goals of ECOPOTENTIAL is to create a framework for ecosystem studies and management of PAs by integrating Earth Observations (EO) and Remote Sensing (RS) into the process of decision making. EO is information about the Earth's physical, chemical and biological systems [El Serafy, 2013]. These observations combined with numerical model output can predictions the ecological status of PAs. These predictions are very useful for monitoring and governing the future state and help in making decisions within a PA. A knowledge-based decision toolbox is needed to support the policy makers in their processes, where they do not have the means to efficiently diagnose and interpret model output or other data regarding PAs.

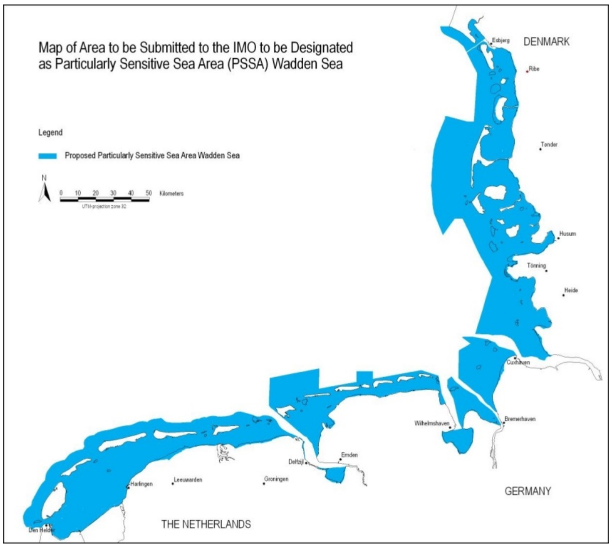


Figure 1.1: The Wadden Sea area, protected area in blue [Waddenzee.nl, 2016].

In this study a method is established to create such a toolbox by means of a study case, the turbidity in the Wadden Sea area. This area is one of the PAs investigated in the ECOPOTENTIAL project and is located in the south-eastern part of the North Sea, see Figure 1.1. The Wadden Sea ecosystem is characterized by tidal flats and a barrier island system with extensive salt marshes; its coastal zone stretches from the Netherlands to Denmark. It is the largest unbroken system of intertidal sand and mud flats in the world [Marencic, 2009]. The Wadden Sea has both the UNESCO World Heritage and Natura 2000 status. This transitional environment between land and sea is characterized by the influence of the tide, fluctuations in salinity, higher temperatures during summer, and occasional ice coverage during winter. Such a diversity creates a wide number of ecological niches. These niches are inhabited by a large variety of species which are adapted to the fluctuating environmental conditions. The area provides a natural habitat for animals such as seals and a various number of birds. The water quality in this area is influenced by pollution such as mainland runoff of pesticides, herbicides and agricultural nutrients.

Figure 1.2: Example of eutrophication *www.quora.nl*.

One of the main problems in the Wadden sea area is eutrophication. This is the phenomenon where excessive amounts of nutrients are present in the water, consequently causing a growth in algal bloom. See Figure 1.2. Eutrophication and algal blooms are serious problems occurring in the Wadden Sea which deteriorate the water quality in many aspects, like oxygen depletion, odor and production of foam, and toxins. Most types of algae are not harmful themselves, however they can still a harmful effect on the ecology in the area. The excessive amount of algae in the water column will block the incoming sunlight, which therefore cannot penetrate into the water and thus preventing light from reaching the algae deeper in the water. When the algae eventually die and decay, the decreased oxygen level in the water can cause mortality among the marine animals living in the area [Li et al., 2013]. Chlorophyll-a is a good indicator of the amount of primary production (algal bloom) in the water column. Turbidity is one of the main sources influencing the phytoplankton growth, due to the blockage of light into the water column [Niu et al., 2015] and is therefore next to chlorophyll-a as important to monitor.

Suspended Particulate Matter (SPM) is the indicator for turbidity and because of the major impact on the chlorophyll-a concentrations, it is important to have a good understanding of this factor. SPM concentrations are in general low in the offshore parts of the North Sea. SPM varies strongly on a short time scale due to tidal action and also on a seasonal scale [Blaas et al., 2012]. SPM is composed of inorganic particles and material of organic origin that is suspended in the water column. Due to the SPM in the water column the light under water is influenced in such a way that it has an impact on the growth of phytoplankton [El Serafy et al., 2011]. Phytoplankton needs enough light and nutrients to grow, therefore the more SPM in the water column the less phytoplankton growth there is. The particles in the SPM can also be a source of nutrients for the phytoplankton.

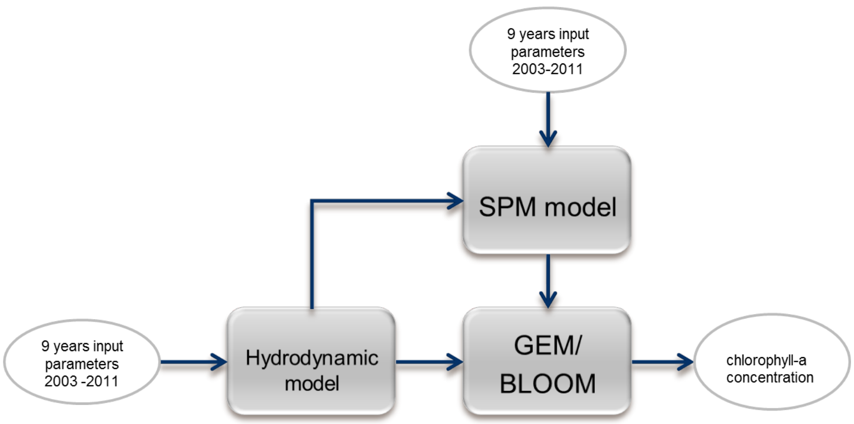


Figure 1.3: Model structure, with the available models together with the years of which sets of input files are available.

The water quality in the Wadden Sea needs to be monitored and modeled to get a good impression on the current status and the future state of the area. The water quality is partly determined by the amount of phytoplankton in the water, in other words the amount of algae. Numerical models are available for the modeling of the water quality in the Wadden Sea area at Deltares. This modeling is done using a combination of three models, see Figure 1.3. Using the Generic Ecological Model (GEM) [Deltares, 2014a] and the phytoplankton module (BLOOM) of this model [Blauw et al., 2009] to calculate the chlorophyll-a concentration in the water column, which is an indicator of the amount of phytoplankton in the water column. The chlorophyll-a concentrations are strongly influenced by forcings, such as the hydrodynamics, which in the GEM/BLOOM model comes from a Delft3D-FLOW model. Another forcing is the SPM in the water column, indicating the turbidity and is calculated by another model, also using Delft3D-WAQ.These models can give a good estimation of the chlorophyll-a concentrations, however, uncertainties will always be present in numerical models. The uncertainties in the GEM/BLOOM model have been topics for previous studies [Jiayuan, 2015]. The SPM model is already calibrated [El Serafy et al., 2013], however in this study the model setup has changed and a numerical model is introduced for the input of SPM. This newly introduced model should be assessed to see if the output gives a good estimation of the SPM concentrations, and what the uncertainty of these values are.

To develop a toolbox for making knowledge-based decisions, not only model output should be taken into account, but also information on the uncertainty in the model output should be given. This information about the uncertainty needs to be quantified first, using an uncertainty analysis. In this study the main focus lies on two aspects, assessing the uncertainty in the model structure coming from the SPM model and the visualization of model output together with uncertainties.

1.2 Problem definition

A problem, concerning the model setup is that at the moment the GEM/BLOOM model uses an SPM forcing based on satellite data. This data, obtained using the MERIS satellite data (a description on how this is done can be found in Eleveld et al. [2007]) is not available anymore since 2011, therefore a new method for implementing the SPM is needed. The latest information on the SPM from different satellite information could be used, such as the Sentinel missions Ḣowever for a more accurate estimation of the SPM, especially in the deeper parts of the water column, this could also be implemented by using a different Delft3D-WAQ model for the calculation of the SPM. Numerical models however are prone to give uncertainty, which needs to be assessed. Therefore an uncertainty analysis of the SPM model is needed. The uncertainty in a numerical model has many different origins, although in this study the focus lies on input uncertainty for the SPM model.

The problem posed by the ECOPOTENTIAL project is that a toolbox is needed to support the decision makers in their processes, by providing them with information about the area and the uncertainty of this information. Thus visualizing these two aspects together is needed to provide decision makers with the information they need for knowledge-based decisions regarding PAs. This toolbox should be a generic tool in which different model output can be displayed. Where in this study the turbidity in the Wadden Sea is addressed, the toolbox can also be used to combine with other studies, such as the study by Meszaros [2016] on the ensemble forecast of the chlorophyll-a concentrations in the Wadden Sea.

1.3 Project goal

For this project the goal is to construct a toolbox in which data in a 3D environment can be visualized together with the uncertainty of this data. The toolbox should be a product that can be universally applied in which different data can be visualized. For testing the toolbox one study case is used. In the Wadden Sea the water quality needs to be monitored and this is done by a water quality model (GEM/BLOOM) that models the chlorophyll-a concentrations. SPM is one of the main parameters influencing this model and is obtained by a numerical model which is prone to uncertainties. Therefore the study case in this research is the data of the SPM model and its uncertainty coming from the model input to see if this model can be used as an alternative input for the GEM/BLOOM model. Thereafter it is tested if this uncertainty analysis can be used as a test case for the developed toolbox. This goal is formulated into the following research question:

How can uncertainty from a SPM model as a driving force for a GEM/BLOOM model be identified, quantified and visualized to help decision makers?

Subquestions are formed from the steps to be taken to answer the main question. Firstly the framework in which the toolbox is generated should be described, to get a clear overview what is needed from the toolbox. The second step is looking at the model setup available from Deltares, understanding the models, the processes, the input and most important investigate what the sources of uncertainty are within these models. In the third step, when these uncertainties are identified the uncertainties need to be assessed, using a uncertainty analysis and to quantify them into a value that can be used in the toolbox. The fourth step is to use the model output and the uncertainties from the previous step to combine these into a visualization. And the last step is to make the toolbox complete and to describe the use of it to users of the toolbox. From this, subquestions are as follows:

1. What information is useful to decision makers?
2. What are the important uncertainties within the regarded model structure and how can they be identified?
3. How can the uncertainties be quantified into values that can be used in the toolbox?
4. How can the model output together with the quantified uncertainties be visualized in a generic toolbox?
5. How can the toolbox be used?

1.4 Thesis outline

In Chapter 2 different literature is consulted to find useful background information. This chapter is divided into five sections, in which information from literature is provided on the earlier uses of GEM/BLOOM, the importance of SPM, why a model is needed for SPM, uncertainty analysis methods and the methods on how to visualize uncertainty.

In Chapter 3 per subquestion the different methods that are used to finding the answer of the subquestion are described. The methods are only described in general, these methods are applied on the case study in the next five chapters.

In Chapter 4 a distinction between what the users of the toolbox, the decision makers, want from the visualization on the one hand and on the other hand what information about this particular PA is important.

In Chapter 5 an overview on general uncertainty in numerical models is given and per model used in this project an overview is given with their processes, input and general uncertainties. Identified are the input files that give uncertainty to the SPM model.

In Chapter 6 these identified sources of uncertainty are analyzed. With a sensitivity analysis the most influential input files are identified. From this result the uncertainty analysis is conducted.

In Chapter 7 the results are then assessed to get a quantification that is usable for the visualization. This chapter focuses on the Python script used to make a toolbox in which a 3D environment is shown of the model results, together with the uncertainty.

In Chapter 8 this toolbox is further elaborated and a description is given on how to use and interpret the toolbox and what additional information should be taken into account to use this tool in decision making processes.

In Chapter 9 the methods used, results and findings are discussed. Points of interest, findings and other observations to be taken into account in follow-up studies are presented here.

In Chapter 10 from the previous chapters and research a conclusion is drawn and the items that were not taken into account in this study are described.

2 | Literature review

In this chapter the literature consulted and useful background information for this project is described. The chapter is divided in different sections, answering questions asked to formulate the background information for this project:

1. Why does SPM need to be modeled numerically?
2. What are studies of the GEM/BLOOM model used in this project and what can we learn from earlier studies
3. What are the methods used to assess uncertainty?
4. What are existing methods of visualizing model output together with the associated uncertainty?

2.1 Modelling SPM concentrations

Why does SPM need to be modeled numerically?

SPM has been monitored with remote sensing techniques, which have limitations when sampling a heterogeneous system. In situ samples are mostly sparse in space and time. The optical remote sensing only measures the surface layer, whereas a large portion of the SPM is found near the sea bed. In these instances, three-dimensional, process based modeling can provide complementary information about the coastal transport system (Baumert et al. [2000], Delhez et al. [2004], El Serafy et al. [2011]). A Delft3D-WAQ model for calculating SPM in the considered area is already developed in previous research [El Serafy et al., 2013]. In which a calibration was made for the model and a sensitivity analysis is conducted, exploring the 71 parameters in the SPM model. From this study 10 parameters were identified to bring the most uncertainty into the model. Still this model gives a good estimation of the SPM in the water column in the North Sea.

The SPM concentrations deteriorate the under water light, which has an influence on the phytoplankton. A numerical model calculation for SPM can help by providing better and more complementary indications of SPM in the water column, therefore reducing the uncertainty for the GEM/BLOOM model.

2.2 Previous studies on GEM/BLOOM

What are studies of the GEM/BLOOM model used in this project and what can we learn from earlier studies?

The GEM/BLOOM model is a module of Delft3D-WAQ and has been calibrated, validated and used in previous studies, which are described in this section. The GEM is regularly applied to assess the ecological quality of Dutch coastal waters and the southern North Sea, the potential effects of e.g. new coastal infrastructure projects and new national and international policies [Los, 2008].

The model has been applied and validated in several projects dealing with estuarine and coastal water systems [Blauw et al., 2009]. An example is the case study of the North Sea, which is a coastal area with relatively shallow waters, in which substantial river discharges result in large fluctuations in salinity, SPM concentrations, nutrient concentrations and algal biomass. Different studies applied GEM on this area, for example in [Los, 2008] where a 3-dimensional GEM/BLOOM model was applied and validated for the (southern) North Sea.

In Jiayuan [2015] a study was done to assess uncertainty for ecological risk mapping, in which the same GEM/BLOOM model for the Wadden Sea was used. This study was part of the ECOSTRESS project, in which the focus lies on better-integrated strategies by improving the risk prevention and disaster management cycle of the coastal zones [Ecostress, 2013]. Jiayuan's thesis describes areas that have the most risk of having large amounts of nutrients and algae bloom occurrences. These risk maps are important to make precautions, relocate fishery resources and interfere recreation activities. The risk is defined by calculating the probability of breaching a certain threshold of clorophyll-a concentrations in the North Sea. Also an uncertainty analysis was conducted using a Monte Carlo sampling and looking into the water quality input parameters. In total there are 422 parameters in the model; of course it is not feasible to investigate them all. In previous research done by Salacinska et al. [2010], a sensitivity analysis was conducted on 71 of these parameters to determine those having the most influence on the output concentrations. These were used by [Jiayuan, 2015] in her uncertainty analysis. These parameters were the Specific Extinction of Inorganic Suspended Matter (ExtV1IM1), the maximum Growth Rate of Diatoms type E (PPMaxMDI_E), the de-nitrification rate (RcDenWat) and the burial rate (VBurDMS1). The uncertainty was represented with confidence bounds around the 90% confidence interval. This thesis concluded that the inlets and coasts regions of the North Sea are the areas that are most at risk of the chlorophyll-a concentration to exceed certain thresholds. The results are shown in Figure 2.1

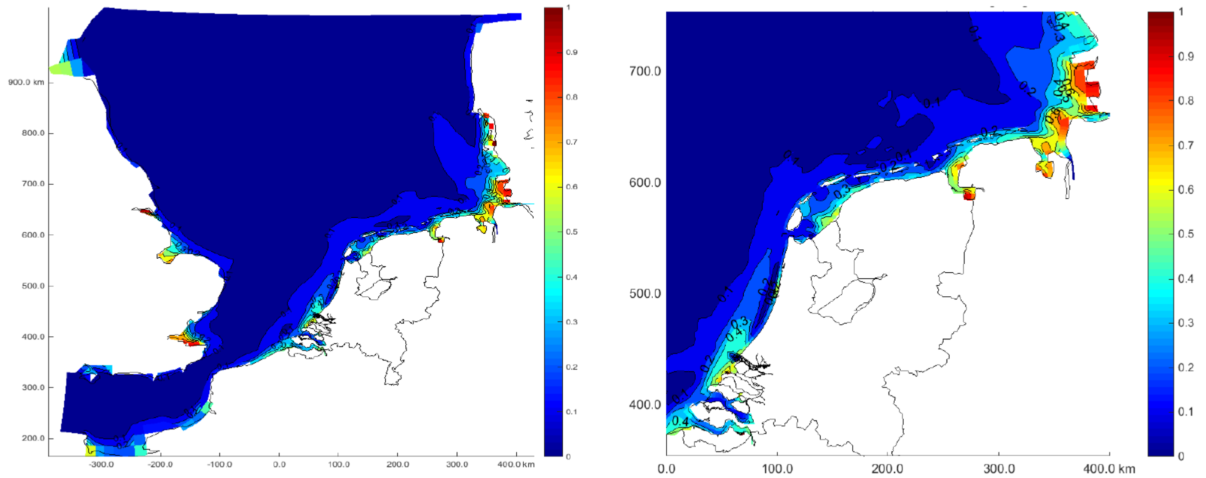


Figure 2.1: Risk maps for chlorophyll-a concentration with a 100 simulations Monte Carlo and a threshold of 7 mg/l [Jiayuan, 2015]. Left: Entire Southern North sea. Right: Zoomed in on the Dutch coast and Wadden Sea area.

Concluded from these literature reviews is that the GEM/BLOOM model has proven to give a good estimation of the chlorophyll-a concentrations in the North Sea and the Wadden Sea area. Different studies for improving the model has been conducted, which increases the accuracy of the model output. Other studies have been conducted to analyze uncertainties in the model. However, it becomes clear that additional research should be performed for looking into the uncertainty caused by the SPM forcing when this forcing is modeled with model output instead of satellite data or a function describing this forcing in the model.

2.3 Assessing uncertainty

What are methods used to assess uncertainty?

For the assessment of uncertainty, literature was consulted on the different ways to make such an uncertainty analysis and on the implementation of different models. First of all the understanding on why uncertainty is needed for decision making was investigated. According to Uusitalo et al. [2015] for decision support models to be useful, the output data should also contain information about the uncertainties related to this output, as the certainty of the desired outcome may be a central criterion on the selection of the management policy. All models are prone to errors coming from different sources within the model. This project looks into the uncertainty in a GEM/BLOOM model, therefore this section focuses on methods to explore these uncertainties.

Li et al. [2013] mentioned four methods for assessing uncertainty: the probability theory method, Monte Carlo analysis, Bayesian method and Generalized Likelihood Uncertainty Estimation (GLUE). These methods are described below in more detail together with their applicability for different models.

Probability theory method - employs probability theory of moments of linear combinations of random variables to find the statistical means and variances of randomly distributed functions. This is applicable for simple linear models, and therefore not feasible for nonlinear systems as GEM/BLOOM.

Monte Carlo analysis - computes output statistics by computing a large number of model simulations, using randomly sampled variables for the input, to comply with probability density functions. This method is generally applicable and is the easiest method for uncertainty analysis [Uusitalo et al., 2015]. The output of all the simulations together is also a probability distribution, which gives statistical information as the mean, the variance, etc. Drawbacks are that thousands of simulations are needed; this sampling technique can result in the clustering of parameter values, as the drawing of samples is done randomly and without correspondence with previous draws [Loucks et al., 2005].

Bayesian method - the key step in this method is to find the prior probability distributions of the model parameters that needs to be investigated. This method assesses and quantifies the uncertainty by calculating probabilistic predictions. It is mainly used on predictive models, in which the parameters can be calibrated in the same time as the uncertainty can be calculated [Li et al., 2013].

GLUE method - is based on the concept of equifinality. This means that different sets of input parameters may result in equally good and acceptable model outputs for a chosen model [Li et al., 2013]. In other words, the parameters from which is randomly chosen are uniformly distributed. This method searches for a set of input parameters that strives to find reliable simulations, rather than looking for the optimal output. Unreasonable combinations of input variables are rejected, while the reasonable and realistic combinations are assigned a posterior probability based upon a likelihood measure that may reflect several dimensions and characteristics of model performance [Loucks et al., 2005].

For the assessment of the uncertainty many model runs need to be performed, in which the input parameters to be assessed are changed for each run. To reduce the computational power needed for these simulations, different sampling methods can be used to reduce to number of model runs. A sampling technique can be applied to draw random values from the given distribution of the input parameters. The Monte Carlo is the most simple sampling and uses random draws from distributions. Another method is the stratified sampling in which a more balanced coverage of the range of input parameters is achieved. This method divides the distribution in different layers and from these layers the same number of parameters are drawn. An example of such a method is the Latin Hypercube Sampling (LHS) [McKay, 1992]. This sampling differs from the Monte Carlo sampling by the number of iterations that is needed to gain a more or less reliable output. LHS forces drawn samples to correspond with the input distribution. The probability distributions of the input parameters are divided into sections of equal probability. After this it draws one observation randomly from each section. In Figure 2.2 an example of a LHS sampling on a probability function is shown and the way the function is divided into different sections.

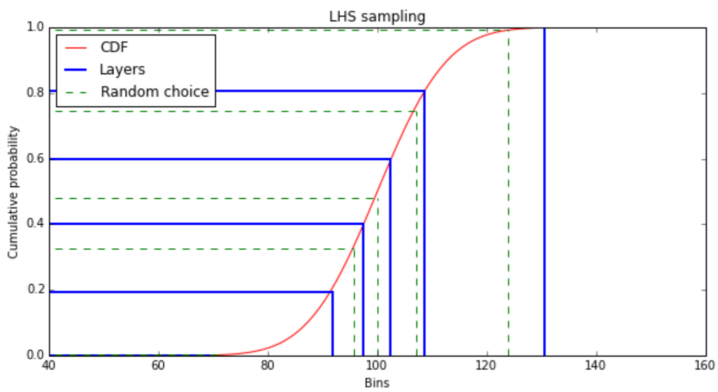


Figure 2.2: LHS example, where the CDF (red) is divided into equally spaced intervals (horizontal blue lines) and a single random sample drawn within each interval (green dashed line), corresponding to a value within the region of the probability function (vertical green dashed line).

In the article of Li et al. [2013], the GLUE method is used together with a LHS to analyze the uncertainties within a GEM/BLOOM model. This ecological model describes such complex processes that a large number of intercorrelated parameters are present. The confidence interval of this uncertainty analysis was obtained by calculating the cumulative distribution functions of model outputs based on the normalized likelihood. The result from this estimation is that the 90% confidence interval that they used on the simulated results failed to enclose the peaks of the observed values in 2009 and 2011. This was due to the inherent uncertainties from inputs, boundaries and the model structure. Also the observed values used for the comparison were space averaged through arithmetic mean other than weighted mean, which could explain any discrepancies. These monthly observation came from four monitoring sites in the Meiliang Bay, including river discharges, water levels, irradiance, temperature, concentrations of ammonia, nitrate, nitrite, phosphate, and biomass concentration of blue- and green algae, and diatom.

In conclusion; for the SPM model it is interesting to use a LHS method together with a method that incorporates the dependencies between input parameters. When including this dependence it reduces the post-processing and increases the efficiency by only generating realistic input combinations. The forcing input is varied and the model is run several times to create the statistical output needed to assess the uncertainty. It should be stated that for assessing the rest of the GEM/BLOOM model, concerning the input parameters, a different approach can be more interesting, such as the GLUE method described above, in which the correlation between the parameters is taken into account, however this falls outside the scope of this project.

2.4 Visualizing uncertainty

What are existing methods of visualizing model output together with the oncoming uncertainty?

The visualization will be a combination of model output and the oncoming uncertainty. What the most efficient method is to visualize this combination was the research in a previous thesis [Vause, 2013]. In this thesis different methods are used to display the water level during a flood in an area, along with the uncertainty of that water level. The tested methods in that research were using color hue, color value, color saturation, shape, size, orientation, texture, transparency and clarity. A survey among the English population was conducted to see what kind of method is best to use for such a visualization. It was concluded that the use of a multi-hue color spectrum is most clear to use in a visualization.

Described in Baart [2013], the way of visualizing values with their uncertainties is described. However here the visualization is based on a one dimensional graph, and not on a two dimensional data. This could be useful for visualizing certain time series on strategic points on the grid. Therefore this study is also taken into account in this chapter. The visualization could be based on different styles, namely: size, color, blur, transparency, jitter, scribble, intervals, and bars. It is shown that the methods of visualizing the uncertainty are similar to the methods used by Vause [2013]. This particular study was on the confidence in coastal forecasts, in which the variance had to be visualized.

From Pfaffelmoser and Westermann [2013] it can be concluded that the normal way to visualize the uncertainty is to visualize by means of color and opacity. However it is also stated that it is difficult to use such methods, because you cannot make clear in what position and structure the specific features in the data are affected by the uncertainties.

In Figure 2.3 different methods of visualizing the uncertainty are shown. These methods are selected from the literature described above as the methods that might be usable in a data visualization. These different methods could be applied in this study to visualize the uncertainty within the toolbox.

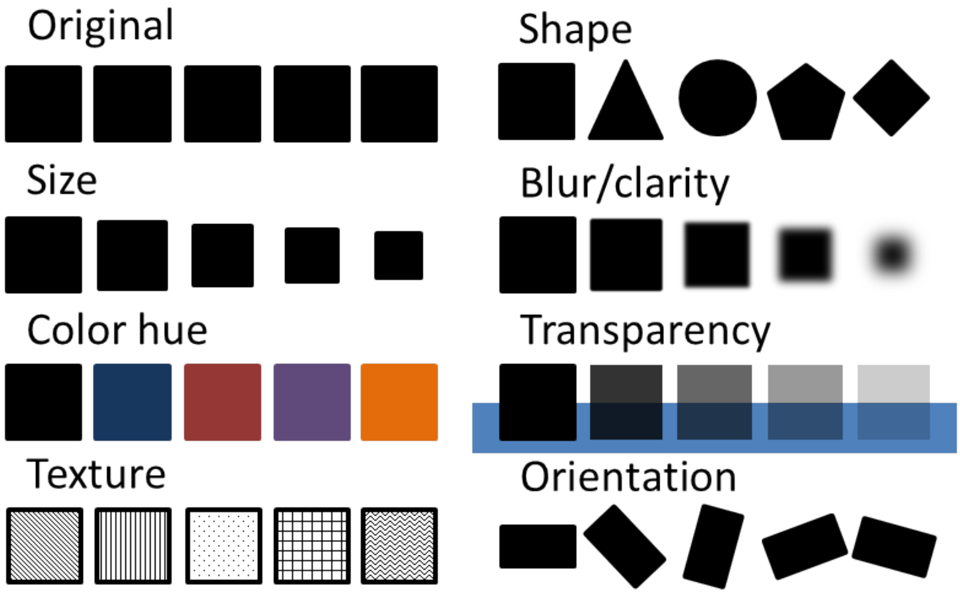


Figure 2.3: Different methods to visualize uncertainty, using characteristics of an object to indicate the amount of uncertainty.

3 | Methodology

In this chapter the methods for answering the subquestions and consequently the main research question are described. This chapter is divided into five sections according to the subquestions stated in the introduction. The first section is about the framework in which the toolbox is created and selection of models to be used. The second about the identifying of the uncertainties in the model structure. In the third section the identified uncertainties will undergo an analysis to quantify the uncertainties. After that the method for the visualization of the uncertainties in the toolbox is described. In the last section a user manual is presented that can help the users of the toolbox. The subquestions are listed below.

1. What information is useful to decision makers?
2. What are the important uncertainties within the regarded model structure and how can they be identified?
3. How can the uncertainties be quantified into values that can be used in the toolbox?
4. How can the model output together with the quantified uncertainties be visualized in a toolbox?
5. How can the toolbox be used?

3.1 Framework for the knowledge-based decision toolbox

One of the two objectives of this research is to provide a toolbox, that will be used to support policy makers and managers by making knowledge-based decisions for PAs. For this project the focus will lie on how to visualize the uncertainty together with information from a model, which will be the base of the eventual toolbox. A toolbox needs to give meaningful and simplified information through which these stakeholders can utilize the data and understand the uncertainties within. This visualization will be made in such a way that it can be used in general by all kinds of models with the associated uncertainty.

To find the answer to the first subquestion, information needs to be gained from different sources (eg. websites, papers) about who the users of the toolbox are and what information about the PA should be represented. First of all one should understand who the decision makers are, for example by research on the ECOPOTENTIAL project website [Ecopotential, 2015a] describes the managers and policy makers of this area. This information is useful to understand the decision makers' needs. On the other hand it is important to understand what information is critical for the Wadden Sea to understand its behavior. Using information about the Wadden Sea, such as the Quality Status Report 2009 [Marencic, 2009] this can be assessed.

3.2 Identifying uncertainties

It is important to have a good overview on the models used and what their general uncertainties, in order to identify the most influential input. All models contain uncertainties, stemming from different sources. Within the scope of this project the uncertainty investigated comes from the input parameters of the SPM model.

To find the answer to this subquestion, research will be done on the background, the setup, input and uncertainties of the models, using literature and calibration reports from previous studies in which these models were used. The in- and output will be assessed and the sources of the uncertainties per model will be listed. Using literature and previous studies also gives insight on the already determined uncertainties and the ones that are yet to be taken into account, narrowing done the input to be assessed.

3.3 Quantifying uncertainties

New uncertainties introduced by using a numerical model for the SPM concentrations needs to be identified. Therefore the uncertainties identified in this SPM model, coming from the input files, are taken into account. First a sensitivity analysis is done on the input files that are identified to contain uncertainty. The most influential input file will consequently undergo an uncertainty analysis. The steps for answering this subquestion are shown in Figure 3.1 and are further elaborated in the section below.

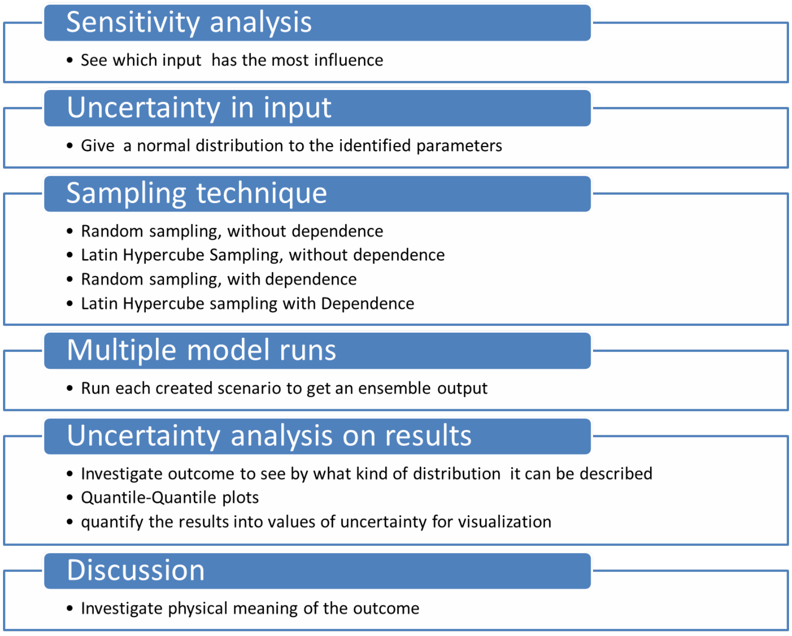


Figure 3.1: Flowchart describing the methodology for conducting the uncertainty analysis.

Sensitivity analysis
Not all of the input files can be assessed within this project as not all of them have a significant influence on the output and it would take up too much computational time. With a sensitivity analysis a selection is made by identifying the most influential input file. Different experiments are performed for the sensitivity considering a single input file. The sensitivity analysis is done in different experiments. In each experiment one input file is considered. The input files are available from the year 2003 to 2011 (each year has a set of files, in total there are 9 sets). In one experiment the model will run 9 times, by changing the specific input file for each year available, and keeping the rest of the input constant. In that way, a rapid assessment method of checking which input has the most influence is made. The actual uncertainty analysis will be focused on the most important input.

Uncertainty in the input
To simulate the uncertainty in the input, a distribution needs to be given to all the input parameters. This is dependent on the base values and the ranges in between which the actual parameter values lie. Finding these distributions is done using previous studies [El Serafy et al., 2013].

Sampling technique
The influence of the uncertainty of the input on model results is investigated by running the model multiple times, using a certain sampling technique. There are many methods to use in an uncertainty analysis, however some disadvantages will limit the amount of methods to be used. In this study a combination of methods is used, which are described below.

First it is generally explained how to obtain multivariate function from a normally distributed parameters and how to get the input for different model runs. It is shown that the steps and the links between the steps can be used the other way around, which is used to include the dependencies (using a Copula). Further on the dependencies and the reducing of the model runs is included in the method.

A Probability Density Function (PDF, notated with $f(x)$) describes the relative likelihood of a value to represent a parameter. This likelihood is on the interval [0, 1], representing the change of being the representation of the actual value. This distribution is used in this study to describe the parameters that have an uncertainty in a range around a base value. A normal distribution is used, which is described by the characteristics average (μ) and a standard deviation (σ). Where the standard deviation is the deviation from this average value [Pfaffelmoser and Westermann, 2013]. The Cumulative Distribution Function (CDF, notated with $F(x)$) is a summation of the likelihood of the values, in total summing up to 1. This CDF is obtained by using the integral of the PDF.

Monte Carlo is an strong method for doing an uncertainty analysis, in which randomly samples are chosen from a probability distribution that describe the input parameters to be assessed. The disadvantages of the Monte Carlo method is that thousands of runs are needed to get a good estimate of the input distributions and the dependencies between input parameters can not be incorporated in this method. Monte Carlo is based on n independent samples, where U are the random variables and X the set of random samples [Gan et al., 2014].

The PDF and CDF of a normal distribution are visualized in Figure 3.2, in the top left and top right. The random sampling of the Monte Carlo method is shown in the bottom left figure, where it can be seen how many samples are taken per interval. To get an uniform distribution, in which the entire PDF is taken into account, many samples are required. The bottom right figure shows the inverse CDF, where the x-axis is taken as the randomly sampled axes. If $f(x)$ is a probability density function, with corresponding CDF $F(x)$, then if U is a vector with random variables with an uniform distribution over $[0, 1]$ (see Figure 3.2 bottom left), it follows that $f(x)$ can be estimated with $F^{-1}(U)$. Where $F^{-1}(U)$ is the inverse CDF, or Percentile Point Function (PPF) (Figure 3.2, bottom right).

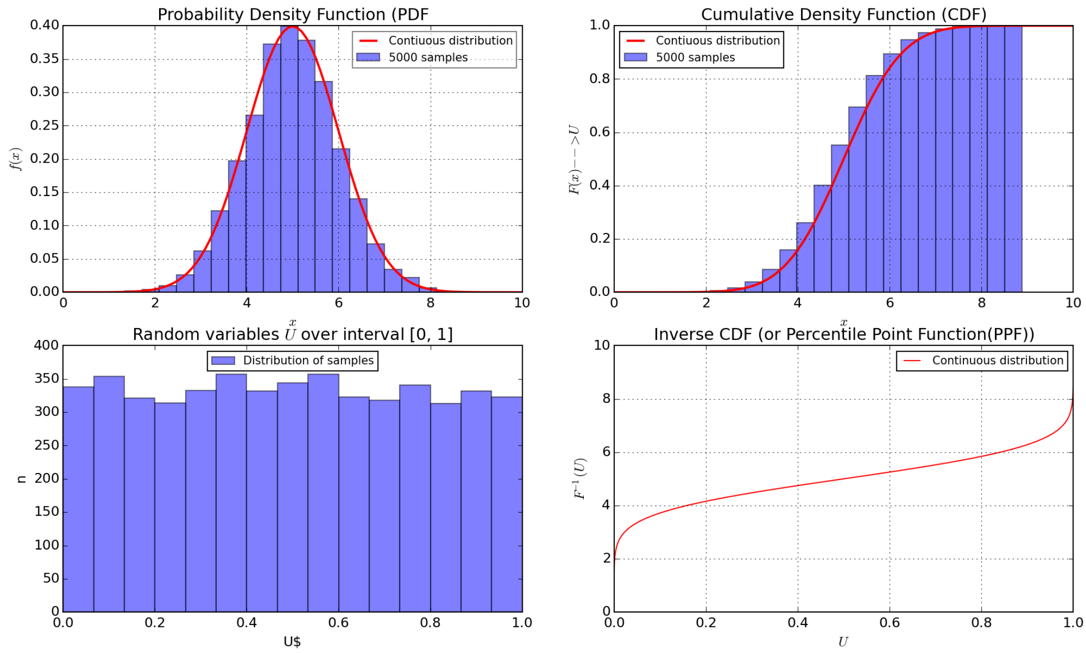


Figure 3.2: Relations between distribution functions (normal distribution), theoretical representation (red) and sampling (blue). Top left: PDF ($f(x)$). Top right: CDF ($F(x)$). Bottom left: Distribution of the samples corresponding with the CDF probability. Bottom right: Inverse CDF ($F^{-1}(x)$).

When combining parameters with a distribution as viewed in Figure 3.2 (normally distributed) a multivariate distribution function can be created. This function is estimated with the scatter plot in the left part of Figure 3.4. Two parameters are given a distribution (one with $\mu = 5$ and $\sigma = 1$ and the other $\mu = 10$ and $\sigma = 3$, which are randomly chosen for this example), and for 5000 experiments, random samples are chosen from both parameter distributions, creating 5000 combined points in the scatter plot. The function that describes the behavior of the scatter plot is the multivariate function.

With the Latin Hypercube Sampling (LHS) the disadvantage of a high number of runs for accurate assessment can be overcome. LHS stratifies the probability distributions of the input parameters into sections of equal probability. After this it draws one observation randomly from each section. In Chapter 2 Literature review more information on the LHS is given and visualized in Figure 2.2.

An example of the LHS method is shown in Figure 3.3, where 5 intervals are chosen and per interval randomly 20 samples. In total 100 samples are visualized in this figure, which is not enough to have a good representation of the distribution, as can be seen in the right panel of this figure. In the left part the distribution of the random variables on the interval [0, 1] are visualized, which is close to a uniform distribution. In the middle panel, the values corresponding to the normal distribution of these variables are shown in the inverse CDF. When transferring this back to a normal distribution, the right figure is obtained and so the normal distribution is estimated.

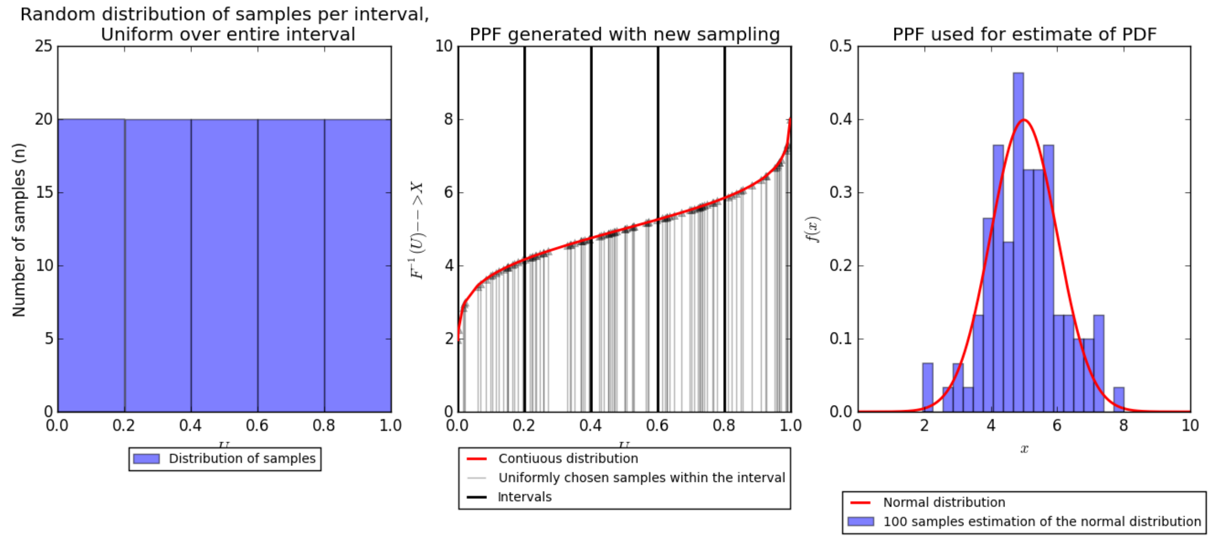


Figure 3.3: Example of LHS, 5 intervals of 20 samples (in total 100 samples), creating a uniform distribution on [0, 1]. Left: chosen samples on interval [0, 1]. Middle: PPF from the random variables U. Right: Transformed PPF to PDF, the 100 samples starts to resemble a normal distribution.

When two parameters are sampled with this method, the two can be converted into a multivariate distribution function as well. This would result in a similar plot as for the Monte Carlo, Figure 3.4 left, however this shape would be obtained with significant lesser sampling points.

The dependencies between the parameters need to be incorporated in the method for this study as well. This can be done using a Copula, which is a multivariate distribution function with uniform marginals [Schmidt, 2006]. The Copula uses uniform random variables from a probability density function. These random variables can be transformed back into the probability function, which is the important last step of the process. In that way the dependencies are integrated into the probability function, and when a sampling method is used these correlations stay intact.

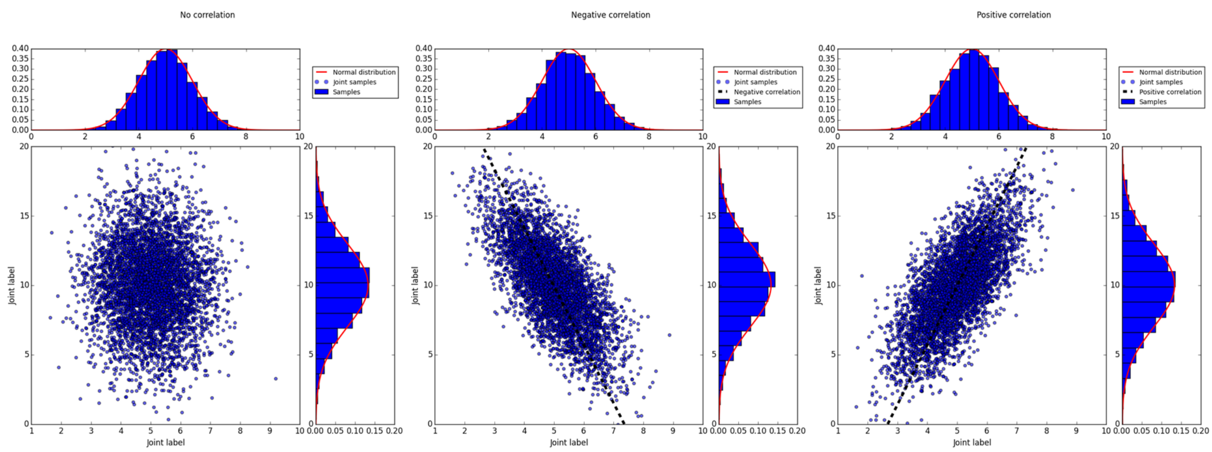


Figure 3.4: Multivariate distributions estimate with a scatter plot from sampling two parameter distributions. The histograms in the top and at the right describe two parameters in this example. Left: No correlation. Middle: Negative correlation. Right: Positive correlation.

The Copula changes the shape of the multivariate distribution function to incorporate the correlation between the parameters. This is visualized in the middle and right figure of Figure 3.4. In these figures a negative correlation ($\rho = -0.7$) and a positive correlation ($\rho = 0.7$) are visualized for this example.

As can be seen, with 5000 experiments a clear and good estimate of the normal distribution is created (see histograms in the figure). Although to reduce the amount of samples this method needs to be incorporated with the Latin Hypercube Sampling technique.

Latin Hypercube Sampling with Dependence (LHSD) is the combination of the Copula method together with the LHS. In that way both the dependencies are incorporated in the method as well as the reduction of sampling needed. This method is adapted from Meszaros [2016] which uses the Copula method from Jiayuan [2015] and the method for LHSD from Tene [2015].

Multiple model runs
In this step the samples taken from the input parameters are implemented into the model and for each sample set the model is run. Each run will produce an outcome for the SPM concentration. The aggregation of the spread of these deterministic runs considering the variations introduced to the input variables of highest importance represent the uncertainty of the input.

Uncertainty analysis of results
For the visualization it is important that the outcome of this analysis will be quantified into one value for the uncertainty. The assessment of the outcome is to see what kind of distribution fits the multiple output of values of SPM. This assessment is done by regarding the PDF, CDF and the Quantile-Quantile (QQ) plot of the outcome. From the histograms of one location in space and time it can be determined what the PDF should be. From the QQ-plot this PDF can be confirmed to be the correct distribution to use to describe the multiple outcomes. From the distribution function a standard deviation is adopted to be a value for the uncertainty. Each location in time and space will then be described in terms of a value for the uncertainty.

Discussion
The last step is to check if the results from the uncertainty analysis on the physical background and if the results are plausible. This is done in a discussion part, where the results are checked with literature on the physical characteristics of the SPM concentrations throughout the year.

3.4 Visualizing uncertainties

The uncertainty found in the previous subquestion should be visualized in such a way that the users of the toolbox can easily have a holistic understanding of the significance of the final outputs from the models being consulted. As deterministic runs can be misleading in the assertion that they represent the truth, the introduction of the uncertainty component is needed to make decisions for future projects in the Wadden Sea.

Using the programming language Python, a 3D environment is set-up to visualize the concentrations SPM in the area of interest, together with the uncertainties. In one location both model output and a value for uncertainty can be presented, using different characteristics of markers. For example, the color values for the model output and the color opacity for the uncertainty. Using Python gives a good opportunity to build in different aspects that can help by making a nice visualization, in which for example a cross section can be made and time series of the entire year can be viewed for each location.

3.5 User manual for the toolbox

Knowledge about the physical character of the area of interest combined with the visualization will give The visualization will provide the model output, in the first case the SPM concentration [mg/l], together with the uncertainties of these values. However to take decisions based on these values is another step, therefore it is important that knowledge of physical aspects in the area of interest is taken into account. Therefore to answer this subquestion a kind of manual is made on how to interpret the visualization and what aspect should be kept in mind when using it.

3.6 Overview

In Figure 3.5 a flowchart summarizing the methodology proposed in this chapter is shown. The squares indicate the chapters and the circles the method and results from these chapters, corresponding with the same color.

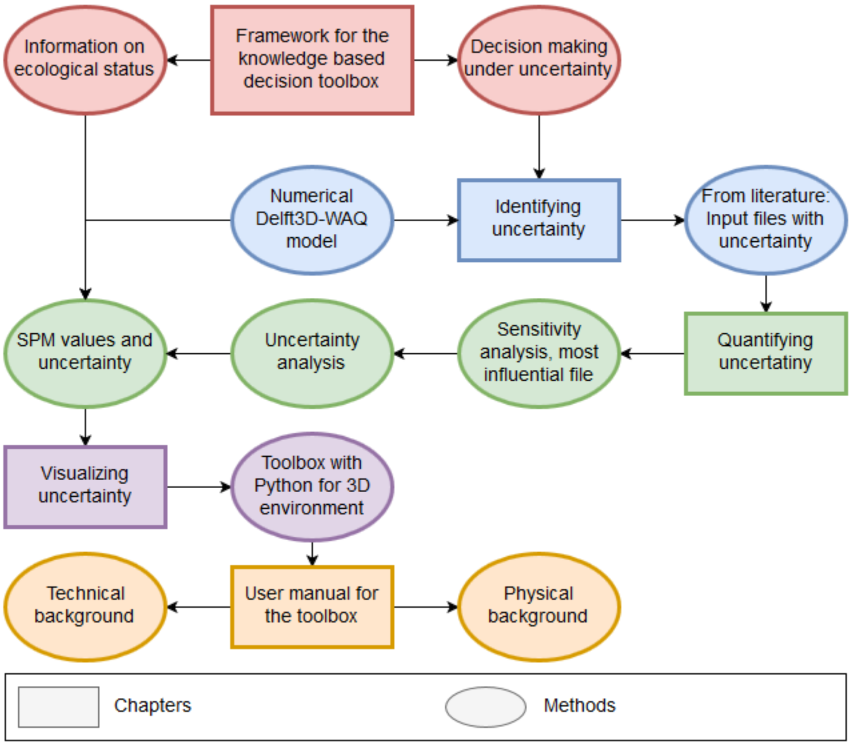


Figure 3.5: Flow chart with chapters and methodology from starting point of describing the framework posing the problem, towards the eventual toolbox. Each chapter is indicated by a color.

4 Framework for the knowledge-based decision toolbox

The toolbox for the Wadden Sea area should visualize information on both physical and biological objectives, to give information needed for the decision makers. Within this project, it is important to make a distinction between what the decision makers want from the visualization and what information about this particular PA is important for them to have. Therefore an understanding of who the decision makers are and characteristics of the Wadden Sea is necessary.

4.1 Decision makers

The decision makers considered are a combination of mainly policy makers, resource managers, scientists and policy makers of the PA Wadden Sea, within the scope of this project. The PAs within ECOPOTENTIAL project are situated in different countries, with each their own regulations, and thus, per PA, the decision makers will differ. Because this area is a stretch of coast bordering three different countries, the Netherlands, Germany and Denmark there is a trilateral cooperation between these countries for the policy making [Ecopotential, 2015b]. The vision of the Trilateral Wadden Sea Cooperation (TWSC) is a unique, natural and dynamic ecosystem with characteristic biodiversity, vast open landscapes and rich cultural heritage, enjoyed by all, and delivering benefits in a sustainable way to present and future generations [Common Wadden Sea Secretariat, 2010].

The Wadden Sea Plan (WSP) lists the objectives of the Trilateral Cooperation and implementations. It is an agreement on how the different countries involved in managing the Wadden Sea oversee the coordination and integration of management of the Wadden Sea area and of the projects and actions that must be carried out to achieve the commonly agreed targets. According to the Convention on BioDiversity (CBD) the ecosystem approach "is a strategy for the integrated management of land, water and living resources that promotes conservation and sustainable use in an equitable way [Common Wadden Sea Secretariat, 2010]. The Habitats Directive state that The Trilateral Monitoring and Assessment Program (TMAP) is a monitoring program for this PA. This program is carried out by the involved countries in the framework of the Trilateral Wadden Sea Cooperation. Its purpose is to provide a scientific assessment of the current state and further development of the Wadden Sea of the trilateral targets of the WSP. The management goals of the Wadden Sea area are primarily at a national level, but agreements are made between all three countries, which brings these goals to a trilateral level. The Trilateral Wadden Sea Plan is an adaption of the WSP from 1997. This plan states that one of the main targets is: A Wadden Sea which can be regarded as a eutrophication non-problem area. [Marencic, 2009].

The natural capital and environmental properties of the Wadden Sea are protected under a variety of regulations. It is designated as a Natura2000 site, RAMSAR site, Water Framework Directive (WFD) transitional water body, and UNESCO world heritage site [Ecopotential, 2015b]. Each protection framework has its own management goals, assessment approaches, and monitoring requirements. To this end, a Wadden Sea Management Council created, with the aim of improving the efficiency and coordination of all different managers of the area.

4.2 Wadden Sea

For this visualization it is important to have information about the process that causes the water quality to deteriorate the water quality in the Wadden Sea area. Fifteen main problems in the Wadden Sea are described in this report (Marencic [2009] and Ecopotential [2015b]), among which climate change, wind farming, and the water quality problems take precedence.

Although all these factors should be assessed, the main focus of this project is on the eutrophication processes calculated with a Delft3D-WAQ model and the parameters influencing this factor. This factor not only has direct consequences on the the species living in the area which as a result, amongst others, also affects the fishery industry. The water quality and the amount of algae bloom caused by excessive nutrients in the water are linked to most of the problems described in the list above. It was found in earlier studies that not in all the months of the year the same amount of chlorophyll-a concentration were present. As found in Jiayuan [2015] are March and April are the months most likely to exceed a chlorophyll-a concentration threshold [Blauw et al., 2009]. In the next section the phenomenon eutrophication is explained in more detail.

4.2.1 Eutrophication

Figure 4.1 shows the processes of eutrophication in a freshwater lake. Eutrophication is the phenomonon when an excessive amount of nutrients are present in the water body, stimulating the growth of algal bloom. When the algae die and decay, they are decomposed and the nutrients contained in that organic matter are then changed into an inorganic form through microorganisms. This can be seen in the right part of Figure 4.1. The process of decomposition of algae consumes a lot of oxygen. Due to the decrease in oxygen in the water, fish and shellfish die. The ecosystem in the area gets disrupted and the water changes color due to the growth of phytoplankton and algal blooms. Moreover the water quality decreases drastically and can cause health problems from contact with the water. The nutrient increase can be caused by human activities, such as pollution. Through these activities the nutrients nitrogen and phosphorus will dissolve into the water and enhances the phytoplankton growth over time. In freshwaters the phosphorus increase is the main factor causing excessive phytoplankton growth, whereas in coastal waters the nitrogen is the key limiting nutrient. In ocean waters the atmospheric fixed nitrogen can easily enter the water.

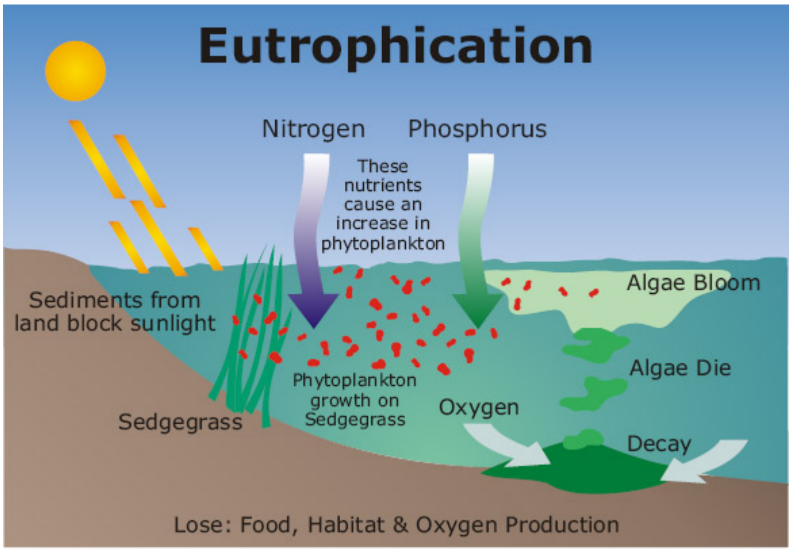


Figure 4.1: Overview of the processes of Eutrophication [Sachin, 2016].

4.2.2 SPM

In the middle left of Figure 4.1 the process of sediments coming from land that block the sunlight is depicted. SPM concentrations have an effect on the plankton growth [El Serafy et al., 2007]. The SPM is composed of fine-grained inorganic particles and material of organic origin that is suspended in the water column [El Serafy et al., 2011] and blocks the light. Which consequently has an effect on the factors needed for the growth of algae that are needed in the process of photosynthesis. Negative effects of this increase in nutrient concentrations are Phaeocystis-blooms, a decline in seagrass, increased bloom of green macroalgae and anoxic sediments [Marencic, 2009]. The influence of the SPM on the algae bloom is therefore significantly important and therefore should be taken into account investigated. Especially when looking at the possibilities regarding the models available at Deltares. The SPM is an external forcing for the model calculating the chlorophyll-a concentrations, as an indicator of the amount of algal bloom. Where the SPM model used is a new asset to the model structure and should be investigated if this forcing is a good indication of the SPM values in the area.

4.3 Discussion

The decision makers for the Wadden Sea area are a wide combination of managers, policy makers, EU directives and more. Because of regulations that are imposed by different parties, there are different boards needed, such as the Wadden Sea Board, to comply for all these different regulations. Therefore the information they need is information about future scenarios and what their impact will be on the area. This information can help decision makers make decisions on projects and impacts that will be within the regulations imposed by the different frameworks.

For this PA information on the eutrophication is important to give an indication of what the current state of the water quality is. From the available models the chlorophyll-a concentration can be obtained, which is a good indicator for the amount of phytoplankton. The SPM has great influence on the chlorophyll-a concentrations and is therefore another important parameter to be assessed.

5 | Identifying uncertainties

This chapter describes the processes, setup, and sources of uncertainty of the numerical models used that gives information about the water quality in the Wadden Sea area. Firstly the general uncertainties you can find in all numerical models are described. After that the uncertainties per model are described in more detail. However it is not feasible to cope with all the uncertainties within these kind of models only the uncertainty coming from the input files are described in more detail.

5.1 Sources of uncertainty in numerical models

Uncertainty comes from different sources within the model, discernible in three categories as found in Uusitalo et al. [2015] and Loucks et al. [2005]. In Figure 5.1 the different sources that can occur are shown. Knowledge uncertainty are the uncertainties that stem from the knowledge gaps when setting up the models initial and boundary conditions, input parameters and forcings. Uncertainties from measurements ranging from wrong measurements due to human flaws, or inaccuracy of the measurement equipment. The second category are the model uncertainties, which occur within the model, such as the numerical errors or uncertain model structures and parameter values. To reduce this category of uncertainties, the model can be calibrated, which has been done in previous research, see Chapter 2 the Literature Study. Therefore these uncertainties fall outside the scope of this project. The last category is the decision uncertainty. These are the uncertainties that occur when interpreting data wrongly. This can be on either end of the model, so model input could have errors because of this source, but also the output can be wrongly interpreted.

Within the scope of this project the uncertainties coming from input files are investigated. In the next sections the three different models which make up the model structure for modeling the water quality in the Wadden Sea will be elaborated. Furthermore, the general uncertainty sources are described within these models and also the uncertainties addressed within this project are discussed. First an overview of the links between the models is given and also a short overview of the grid and the set up of the model will be described.

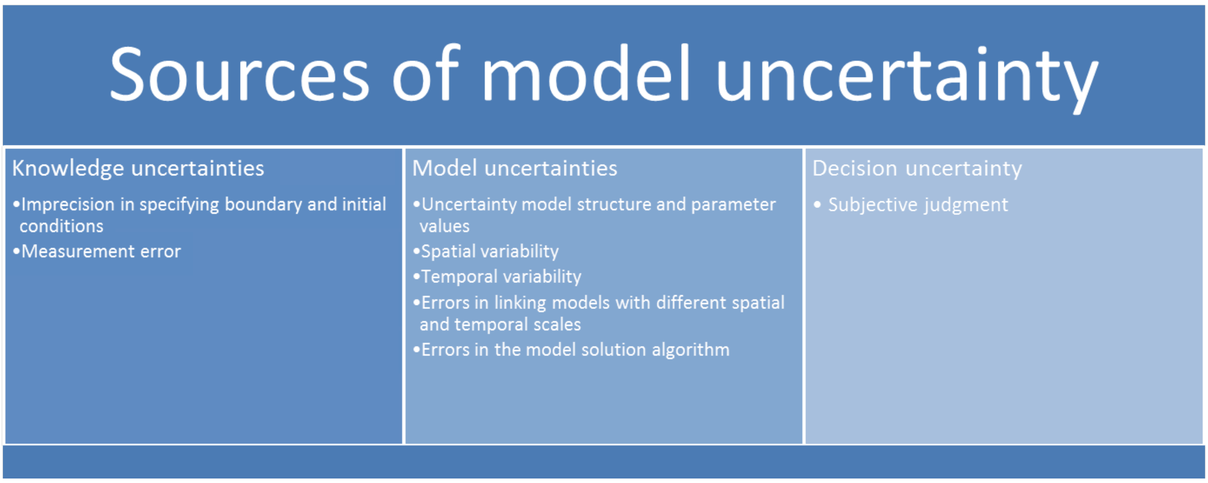


Figure 5.1: Different kinds of uncertainties that occur in numerical models.

5.2 Model structure en setup

In Figure 5.2 below, the model structure of how the models are linked together is shown. A general structure is shown below, for details see Appendix A. For the input of the waves, a Delft3D-WAVE model [Deltares, 2014b] is needed. This model is the SWAN (Simulating WAves Nearshore) model, which is a third generation wave model [SWAN team, 2016] and calculates the wave field for the entire domain.

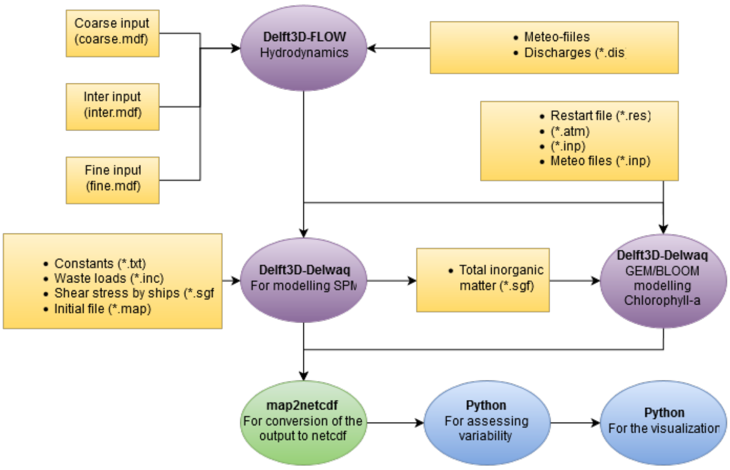


Figure 5.2: Entire model structure for the set up to calculate the water quality in the Wadden Sea area, in purple the numerical models used, in yellow the input uncertainties and in green and blue the post-processing steps towards the visualization.

From Figure 5.2 is it observed that the hydrodynamic model drives both the SPM model and the GEM/BLOOM model. The SPM is thereafter an input forcing for the GEM/BLOOM model. Because the uncertainty of the GEM/BLOOM model needs to be assessed, the first section will explain in more detail how the GEM/BLOOM model is set up and what the different processes are for this model. The following section will explain the SPM model, the processes, the in- and out-put of this model and how the uncertainties are transferred from this model to the GEM/BLOOM model. The same will be done in the last section, in which the hydrodynamic model is further explained.

5.2.1 ZUNO-DD

The models used are for the ZUNO area (Zuidelijke Noordzee, translated into English: Southern North Sea) and are made with Domain Decomposition. This means that there is no uniform grid for the entire area. The grid is split up in three different grids, the coarse, the intermediate, and the fine grid. In the open sea area the grid can be coarser, because this larger mesh width can capture all the larger processes in that area. However, for the areas closer to land and in the areas of interest, the variability is higher because this is where most of the processes takes place. These process need to be determined on a smaller scale and therefore the areas of the Dutch coastal stretch have a fine grid and on some places an intermediate grid. The grid is curvilinear and has 12 non-equidistant sigma layers, which provide a good resolution to cope with the surface mixing layer and the elevated near-bed SPM concentrations [Stuparu, 2012]. Sigma layers divide the profile in different layers, following the bathymetry. In

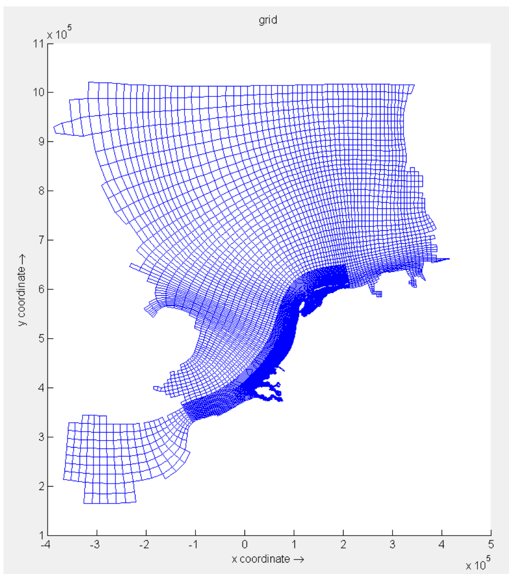


Figure 5.3: ZUNO-DD grid, coupled fine, intermediate and coarse grid.

Table 5.1 the dynamic sigma layers with the relative thickness, based on a percentage of the total water column. Layer 1 is the surface layer and layer 12 is the bottom layer. In total there are approximately 160.000 cells in the grid, with 237 cells in the x-direction and 245 in the y-direction. Figure 5.3 shows the grid and its location in the North Sea. A space varying bathymetry is applied. The local depth is defined relative to NAP (Normaal Amsterdams Peil [m]).

Table 5.1: The 12 different sigma-layers in the grid, with the relative thickness [Blaas et al., 2012]

Layer	1	2	3	4	5	6	7	8	9	10	11	12	**Total**
Relative thickness (%)	4.0	5.6	7.8	10.8	10.9	10.9	10.9	10.9	10.8	7.8	5.6	4.0	**100.0**

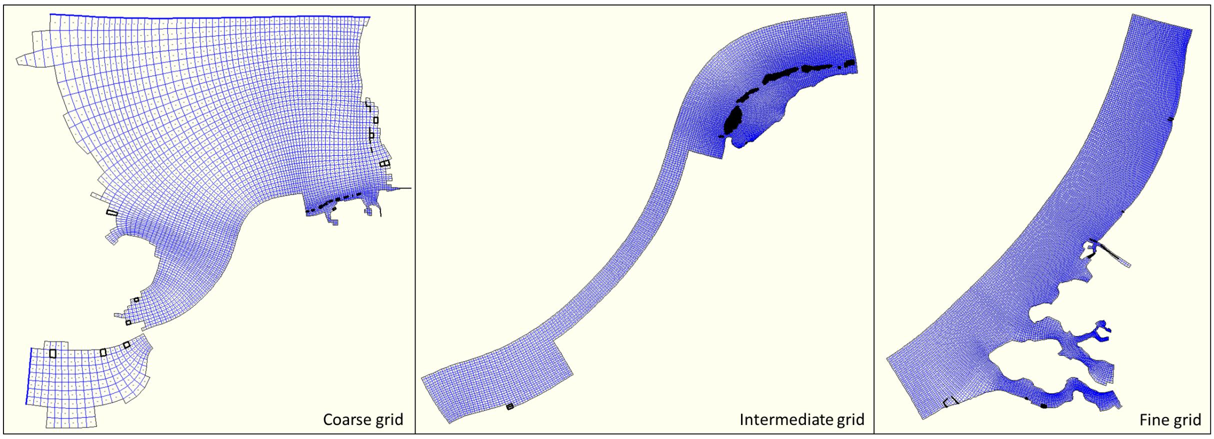


Figure 5.4: Left: Coarse grid. Middle: Intermediate grid. Right: Fine grid.

5.2.2 Coarse grid

As can be seen in the left part Figure 5.4 the coarse grid covers the main part of the North Sea. Bathymetry data in this grid originates from the North West European Shelf Operational Oceanographic System [NOOS, 2016]. The resolution of the coarse grid in the x-direction (Δx) varies between 6,000m and 20,000 m and in the y-direction (Δy) between 5,000m and 30,000m. The coarse grid contains out of 62 cells in the x-direction, 134 in the y-direction and 12 in the z-direction.

An open boundary is implemented on the north part of the domain (see darker blue line) as well as an open boundary in the channel at the south-western part. The water levels at these boundaries are represented bu the astronomical tide.

5.2.3 Intermediate grid

The intermediate grid covers the main part of the Dutch coastal zone and the Wadden Sea. This refinement of the coarse grid was chosen to cover the coast where the most important processes in the water need to be covered. The Bathymetry of this grid originates from the same source as for the coarse grid, the NOOS. The resolution of the coarse grid in the x-direction (Δx) varies between 1,000m and 2,500m and in the y-direction (Δy) between 2,000m and 3,000m, which is a factor 9 smaller than the coarse grid. The coarse grid contains out of 65 cells in the x-direction, 245 in the y-direction and 12 in the z-direction. See the middle figure in Figure 5.4.

5.2.4 Fine grid

The fine grid is given in the right part of Figure 5.4. The coarse grid covered most of the Dutch coastal zone, however there were the influence of the Meuse outflow is greatest a new grid was formed, the fine grid. The bathymetry of the fine grid originates from the most recent Kusstrook Fijn model which contains a compilation of surveys by the Dutch Hydrographic Service and Rijkswaterstaat, the most recent of which was carried out in 2005 [Blaas et al., 2012]. The resolution of the coarse grid in the x-direction (Δx) varies between 500m and 1,000m and in the y-direction (Δy) between 1,000m and 1,500m, which is a factor 36 smaller than the coarse grid. The fine grid contains, out of 110 cells in the x-direction, 214 in the y-direction and 12 in the z-direction.

For the Delft3D-WAQ models (SPM and GEM/BLOOM) another grid is used, however the setup of the ZUNO-DD is still used, the fine part is aggregated 2x2. Blocks of four cells (2x2) are aggregated into one cell, see Figure 5.5. The output of the hydrodynamic model needs to be aggregated as well and after that the model output should be coupled that the input matches the layout of the grid used in the Delft3D-WAQ models. The aggregation was performed by using the DIDO module in the Graphical User Interface (GUI) of Delft3D, creating an aggregation file. With this file, the information of the fine grid and the hydrodynamic components on this grid can be aggregated using Agrhyd.

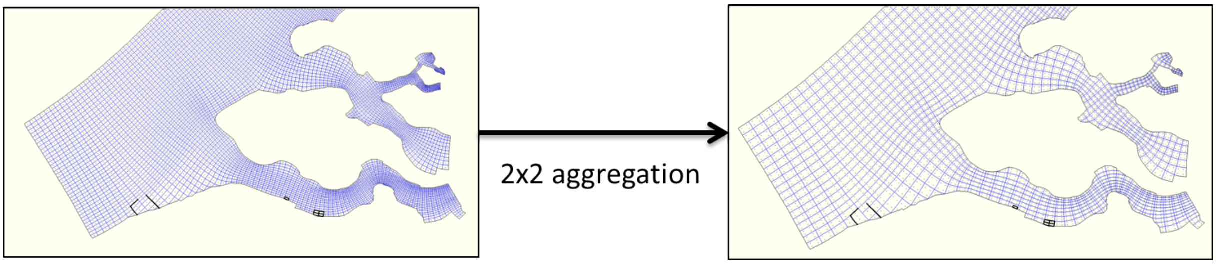


Figure 5.5: 2x2 aggregation of the fine grid.

5.2.5 River discharges

Within the domain many rivers are present, having an influence on the hydrodynamic model, but are also extremely important forcings on the SPM and the nutrient concentrations. Rivers carry sediments from upstream of agricultural run-off and deposit it in the North Sea. In previous research, where these models were implemented, for international monitoring programs, the time series for the discharge for every river was used, coming from measurement stations. Because the rivers are situated in different countries (namely England, Belgium, France, the Netherlands, Germany and Denmark) different sources needed to be consulted, which can be found on page 19 of Blaas et al. [2012]. The locations of the rivers are shown in Figure 5.6. For the study area the most relevant rivers are the ones in the Belgium, the Netherlands and Germany.

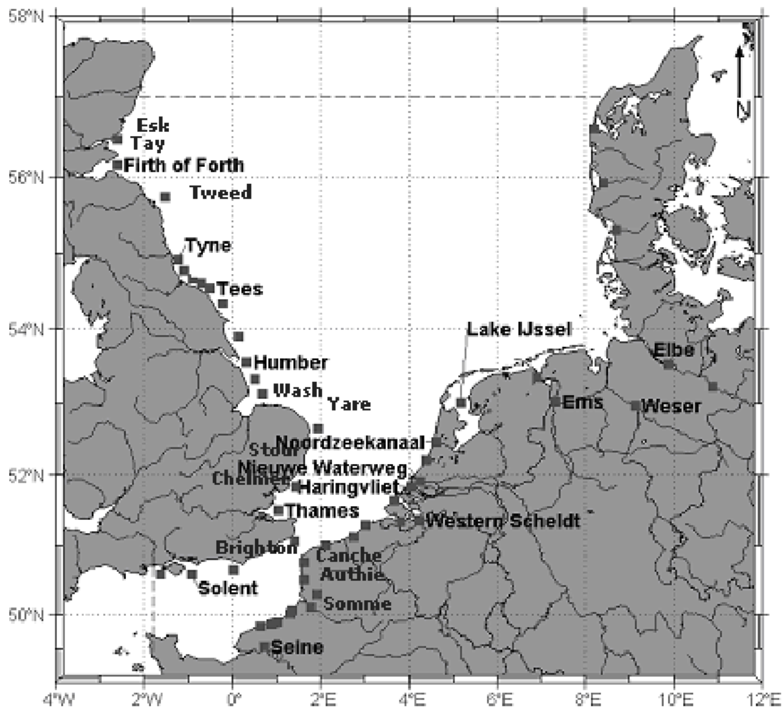


Figure 5.6: Locations of the rivers in the ZUNO-DD grid (Blaas et al. [2012] and Meuwese [2007])

5.2.6 Meteorological forcing

The meteorological forcings are the wind velocity (U_{10} in $[m/s]$ at 10 m height from the water surface), temperature, pressure field, cloud coverage among others. These are boundary conditions for the momentum equations used in Delft3D-FLOW. The data for these forcings come from measurements, supplied by the Royal Dutch Meteorological Institute from the KNMI.

5.2.7 Hydrodynamics

The hydrodynamic input comes from a Delft3D-FLOW model. This model can simulate coastal, riverine and estuarine areas. The flow boundary conditions are coming from a different model, a SWAN (Simulating WAves Nearshore) model, which is a third generation wave model in the Delft3D-WAVE environment [SWAN team, 2016]. This means that wave data measured farther offshore can be recalculated to nearshore data.

Delft3D-FLOW can be applied, as well, on a two as a three dimensional grid. It is a hydrodynamic (and transport) simulation program which calculates non-steady flow and transport phenomena that result from tidal and meteorological forcing on a rectilinear or a curvilinear, boundary fitted grid [Deltares, 2014c]. The numerical model solves the unsteady shallow water equations (Navier-Stokes) in two or three dimensions, using Finite Difference Methods (FDM). These methods solve the differential equations (in this case the shallow water equations) using finite differences to approximate the derivatives of the equation. The system of equations consists of the horizontal equations of motion, the continuity equation and the transport equation for conservative constituents.The model can be applied to predict flows in shallow sea waters, coastal areas, estuaries, lagoons, rivers and/or lakes. It aims to model flow phenomena of which the horizontal length and timescales are significantly larger than the vertical timescales [Blaas et al., 2012]. Different scales can be calculated with this model. The equations are capable of resolving turbulent scales; large eddies. However, it should be taken into account that if these small scales are modeled that the grid should also be adapted to this size to comply with all the processes on that scale, usually the grid is too coarse.

The model calculates the water levels, salinity, temperature and transports. These output are needed for the input of the SPM and the GEM/BLOOM model. The model was validated and calibrated in previous projects, see the Literature Study in Chapter 2.

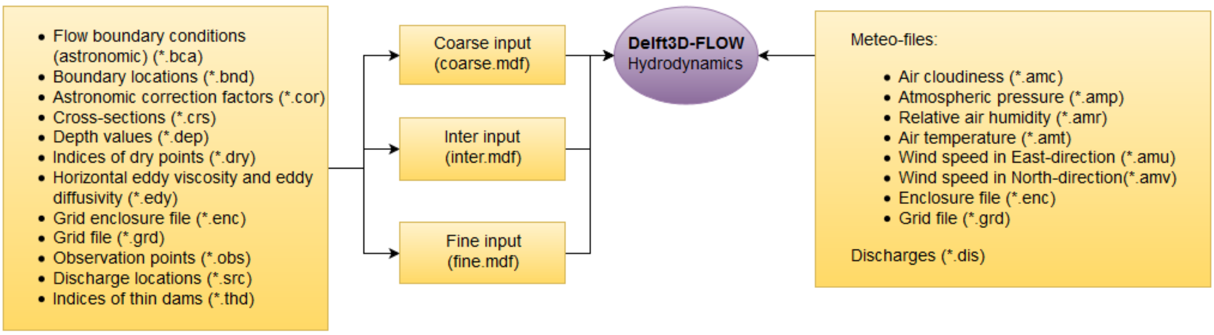


Figure 5.7: Input for the hydrodynamic model, complete structure in Appendix A

Figure 5.7 shows the input of the hydrodynamic model. On the left side the communication files coming from the SWAN model are shown. These files are needed for all three the different grids. Furthermore in the left yellow squares the information on the meteorological data and discharge data are shown, which come from measurement data. Measurements have uncertainties, referred to as knowledge uncertainties in the introduction to this chapter, as knowledge uncertainties. This information comes from measurement stations. The meteorological data is needed for the momentum equation. This data contains information about wind velocities and air pressure at Mean Sea Level (MSL), cloud cover, air temperature (at 2 meters height above MSL) and relative humidity fields. The data comes from the KNMI HIRLAM model (http://hirlam.org and www.knmi.nl).

5.3 GEM/BLOOM model

A water quality model has one or more state variables, pollutants or substances, which enter the modeled area through model boundaries or lateral inflows. They move with the currents through the modeled area. At the same time they may show their own specific behavior in the aquatic environment. This can be a simple decay, but also an interaction of transformation between different state variables. [Deltares, 2014a]. The GEM model is part of the Delft3D-WAQ, which models the water quality and aquatic ecology. With an advection-diffusion equation the sources and sinks of the variables are taken into account [Blauw et al., 2009] and the transport processes are described. The advection and diffusion fluxes between the cells are derived the hydrodynamic model described in 5.2.7. The simplified advection-diffusion equation:

$$\frac{\delta C}{\delta t} = -u\frac{\delta C}{\delta x} - v\frac{\delta C}{\delta y} - w\frac{\delta C}{\delta z} + \frac{\delta}{\delta x}\left(D_x\frac{\delta C}{\delta x}\right) + \frac{\delta}{\delta y}\left(D_y\frac{\delta C}{\delta y}\right) + \frac{\delta}{\delta z}\left(D_z\frac{\delta C}{\delta z}\right) + S + P \tag{5.1}$$

in which:

C	= Concentrations $[g/m^3]$
u, v, w	= Components of the velocity vector $[m/s]$
D_x, D_y, D_z	= Components of the dispersion tensor $[m^2/s]$
x, y, z	= Coordinates in three spatial dimensions $[m]$
S	= Sources and sinks of mass due to loads and boundaries
P	= Sources and sinks of mass due to processes
t	= Time $[s]$

In this section the water quality processes are described. GEM is the Generic Ecological Model GEM simulates the nutrient cycles of nitrogen (N), phosphorus (P) and silicate (Si). In GEM there is a phytoplankton module, BLOOM, which simulates primary production with competition between species, respiration and mortality of phytoplankton [Blauw et al., 2009]. GEM links different physical, chemical and ecological model components into one generic and flexible modeling tool that allows for variable sized, curvilinear grids to accommodate both the requirements for local accuracy and maintaining a relatively short model run-time. Fifteen algae species can be modeled within BLOOM. Further processes, transports and changes that are calculated in GEM/BLOOM are described in more detail in Figure 5.8. The main processes are: extinction of light, de-nitrification and nitrification, particulate organic matter decomposition in water and sediment, settling, burial, filter feeder processes (grazing, respiration, excretion), and re-aeartion processes.

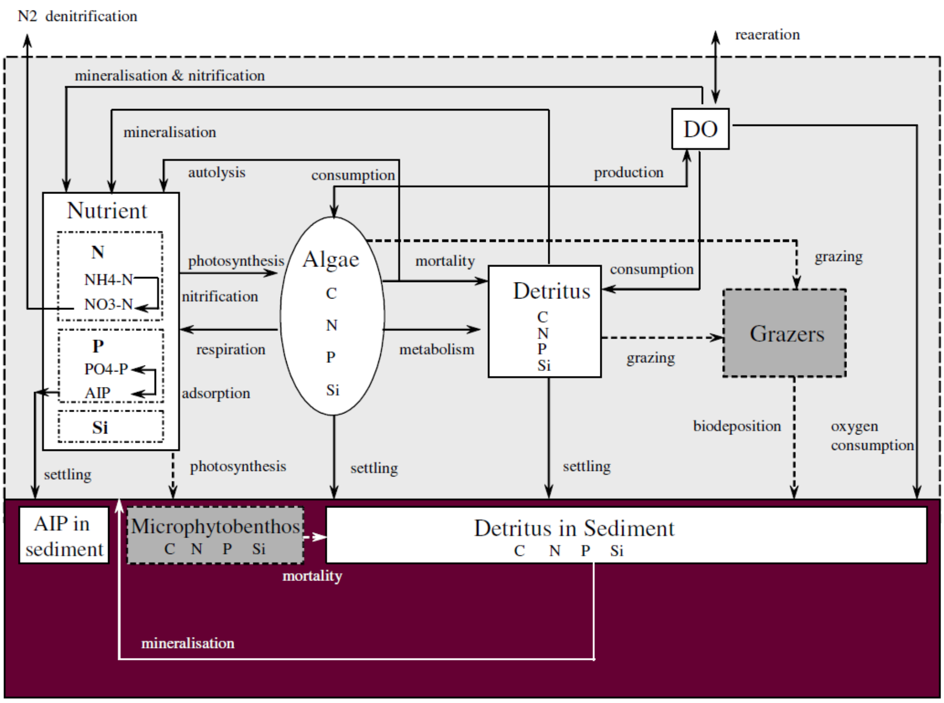


Figure 5.8: The processes on the model variables in the GEM/BLOOM model [Deltares, 2014d]

For eutrophication the state variables are algae, inorganic nutrients (N-NH4, N-NO3, P-PO4, Si) and particulate organic matter. The SPM forcing coming from the SPM model described in the next section. The processes included in GEM are phytoplankton processes, extinction of light, decomposition or particulate organic matter in water and sediment, nitrification and dinitrification, reaeration, settling, burial and filterfeeder processes.

The GEM/BLOOM model is a good choice to model the Wadden Sea, because of the spatial resolution of approximately 1 x 1 km. Other process-oriented ecological models usually perform well for area that they are built for, however when they are applied to other systems, their performance tend to be poor in comparison to the GEM/BLOOM model, even after re-parameterization [Blauw et al., 2009]. The model outputs under consideration include information on chlorophyll-a concentrations, and other parameters as the light climate, primary production and nutrients, which are not looked into in this study.

5.4 SPM model

SPM is composed of fine grained particles of organic and inorganic origin that are suspended in the water column [El Serafy et al., 2011]. Due to the SPM, the water gets turbid and the incoming light in the water column will be blocked, decreasing the phytoplankton growth. However the organic fraction of the SPM is one of the sources of nutrients, which to an extend enhances this growth.

The SPM model is used to recalculate the salinity, which has been calculated in the hydrodynamic model, and the inorganic matter in different layers throughout the entire year. The model was used to simulate mostly surface values of SPM in the Dutch coastal zone, particularly in the fine part of the grid [Cronin, 2014]. It was developed in Delft3D-WAQ module, as was the GEM/BLOOM model, which calculates the advection-diffusion equation for the transport of SPM. More information on the processes in Delft3D-WAQ is given in section 5.3. In the Dutch coastal zone, SPM partly originates from the rivers and erosion, local seabed resuspension and partly from sources outside the North Sea. The SPM concentrations have a strong variability on the short time scale, due to the wave and tidal forcings, but also on a seasonal scale [Stuparu, 2012]. The three important output from the SPM model are the three fractions, namely the coarse (IM1, diameter $40\mu m$), medium (IM2, diameter $15\mu m$) and the fine (IM3, diameter $1\mu m$)sediments (in $[g]$) [Tene, 2015]. Together they form the Total Inorganic Matter (TIM), which is the total of the suspended inorganic matter in the water column, and is a good indicator of the amount of SPM in the water, because the organic fraction is relatively small. Therefore, further in this project, the TIM concentrations will be used when addressing the output of the SPM model, but will be called SPM. The main driving forces of the SPM model are the discharges of the rivers coming out in the North Sea, the meteorological input, and hydrodynamic model output.

The model was developed to simulate mostly surface values of SPM in the Dutch coastal zone, in particular the area covered by the fine grid [Cronin, 2014]. The model calculates the concentration of suspended matter throughout the water column and near the bed on a tidal timescale and seasonal scale over the entire model domain. The boundary conditions are prescribed at the English channel and the Northern model boundaries, connecting the rest of the North sea, based on derived climatology. The grid used for the SPM model differs from the Hydrodynamic grid, by aggregation of the intermediate and the fine grid. Forcings needed for this model is the salinity, the temperature and wind forcings that are calculated by the Hydrodynamic model. The numerical scheme that is used in this model is the iteration solver with backward differences (15) [Deltares, 2014a]. To simulate one year it takes approximately 11 hours.

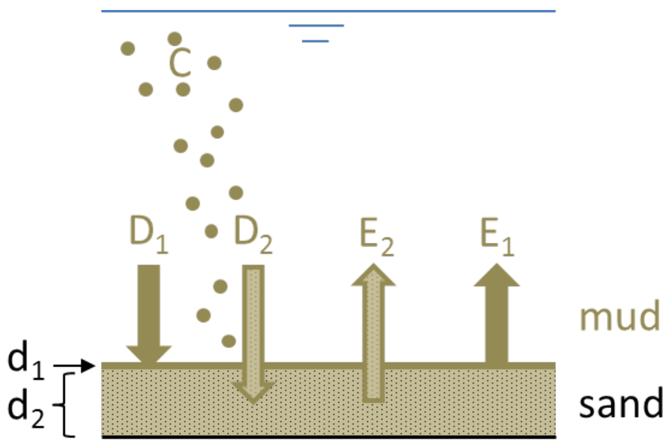


Figure 5.9: Buffer mode as used in the SPM model (Adapted from Van Kessel et al. [2011]). Showing the Erosion (E) and Deposition (D) fluxes between the water column and the two bed layers (d_1 and d_2) and the concentration (C) of suspended matter in the water column.

The SPM model is described in two bed layers that interact with each other, see Figure 5.9. These two layers are a thin fluffy layer (S_1, indicated in Figure 5.9 with the layer thickness d_1 [m]) which is formed during slack tide and accounts for quick resuspension and settling. The total sediment mass in this layer is rather small. The other layer is the sandy buffer layer (S_2, indicated in the Figure, layer thickness d_2 [m]), in which the more finer material is stored and for a longer period. This layer is only disturbed during extreme conditions, such as a storm. (El Serafy et al. [2013] and Van Kessel et al. [2011]). In the Figure, D_1 and D_2 (both [mg/l]) represent the deposition flux to both S_1 and S_2, and E_1 and E_2 (both [mg/l]) indicate the erosion from both layers. These fluxes are described in the following equation, the parameters are described in 5.2:

$$D_{1,IM_i} = (1 - \alpha_{IM_i}) V_{Sed,Im_1} C_{IM_i} \tag{5.2}$$

$$D_{2,IM_i} = \alpha_{IM_i} V_{Sed,Im_1} C_{IM_i} \tag{5.3}$$

$$E_{1,IM_i} = min\left(Z_{Res,IM_i}, V_{Res,IM_i} M_{i,j}\right) \left(\frac{\tau}{\tau_{cr,S_1,IM_i}} - 1\right) \tag{5.4}$$

$$E_{2,IM_i} = F_{ResPup} M_{i,1} \left(\frac{\tau}{\tau_{Sh}} - 1\right) \tag{5.5}$$

$$\tag{5.6}$$

Table 5.2: Parameters in the erosion and deposition equations [El Serafy et al., 2013].

Parameter	**Meaning**	**Unit**
C_{IM_i}	Concentration of fraction Inorganic Matter (IM_i) of the fraction class i	$[mg/l]$
V_{Sed,IM_i}	Settling velocity of the fraction class i	$[m/s]$
α_{IM_i}	Proportion of the deposited silt that is stored directly in the sandy layer (S_2)	$[-]$
D_{j,IM_i}	Deposition flux of SPM fraction IM_i from layer S_j	$[g/m^2d]$
E_{j,IM_i}	Resuspension flux of SPM fraction IM_i from layer S_j	$[g/m^2d]$
τ	Bottom shear stress	$[Pa]$
τ_{cr,S_1,IM_i}	Critical shear stress for silt resuspension fraction i from fluff layer (S_1)	$[Pa]$
τ_{Sh}	Critical Shields stress for sand mobilization in buffer layer S_2	$[Pa]$
Z_{Res,IM_i}	Zero-order resuspension rate from layer S_1	$[g/m^2d\ ]$
V_{Res,IM_i}	First order resuspension rate from layer S_1	$[1/d]$
$M_{i,j}$	Mass of sediment fraction i in layer j per surface area	$[g/m^2]$
F_{ResPup}	Van Rijn (1993) pickup factor from buffer layer	[-]

5.5 Uncertainties in regarded models

In this section the uncertainties in the GEM/BLOOM model and consequently the SPM model are identified and discussed.

5.5.1 Uncertainties in GEM/BLOOM

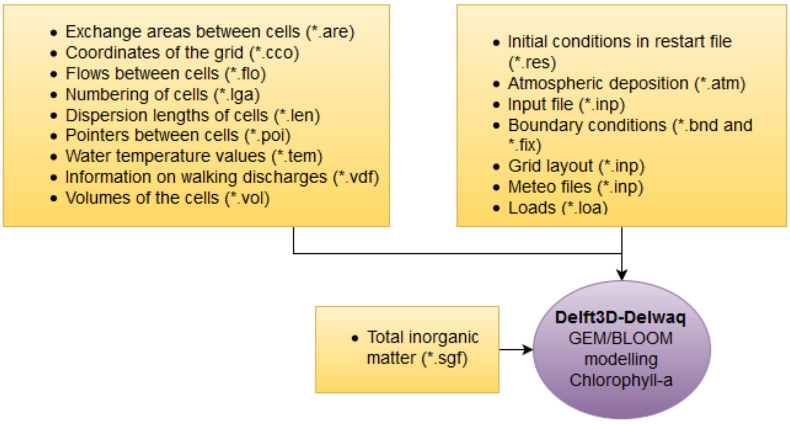


Figure 5.10: Input for the GEM/BLOOM model, complete structure in Appendix A.

As mentioned before, within the scope of this research only uncertainties due to input forcings are taken into account. In Figure 5.10 the input of the GEM/BLOOM model is shown. It can be seen that many input files can be a source for uncertainty in this model. The hydrodynamic model serves for the input of many different files, as described in the top left yellow square, which are all correlated. The SPM input is described in the bottom left yellow square, and is just one file with the TIM concentrations coming from the SPM model. The topmost right yellow square describes the other files. The grid layout and the boundary conditions are all the same, and are not a source of uncertainty that is addressed here.

From Chapter 2 it was concluded that the uncertainty of the model was already assessed in earlier studies. The uncertainty caused by the water quality parameters are the subject of an ongoing study (Meszaros, 2016). The SPM forcing through another method, namely using an SPM transport model is yet to be determined. The uncertainties in this model that are interesting for this project comes from the SPM and hydrodynamic forcing. Therefore the oncoming uncertainties from these models will result in the uncertainty of this model. The inherent uncertainties coming from the processes within the models etc are not addressed in this project.

5.5.2 Uncertainties in a SPM model

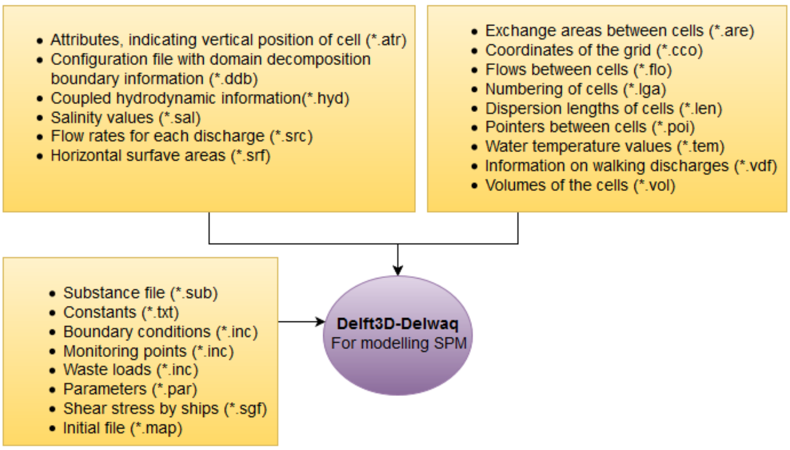


Figure 5.11: Input for the SPM model, complete structure in Appendix A

In Figure 5.11 the input files for the SPM model are shown, which is the focus area of the uncertainty analysis in this project. The two top most yellow squares describe the input coming from the hydrodynamic model. The left yellow square describes the other inputs of the SPM model. In which some are also sources for uncertainty, such as the parameters, constants, waste loads, shear stresses and the initial file. The different input files of the SPM model are described in more detail below [Deltares, 2014e].

Hydrodynamic input The hydrodynamic input is described in the previous section. It contains information about the water temperature, the flows, salinity, processes, exchanges between cells, etc. Because these input files come from a numerical model, there will be sources of uncertainty present. However, assessment of these uncertainties fall outside the scope of this project, the hydrodynamic would need to be run multiple times, which is technically and numerically too demanding (storage of 300 GB and a computational time of 3 days).

Substance file The substance file describes the substances to be calculated in the SPM model, therefore contains no knowledge uncertainty.

Constants This file contains the parameters that are constant for each cell, but differ in time. The file consists out of 71 parameters. These are a major source of uncertainty. Some parameters, such as water density are the same globally in the model. However the model covers such a large area, in which many different conditions are taken into account, one value per parameter can not be sufficient to represent the reality [El Serafy et al., 2013].

Boundary conditions (BC) This file describes the locations of the boundary conditions and the imposed BC at those points, such as a water level BC in the northern boundary. These are the same for each year.

Monitoring points These are the locations of the monitoring points, which are not a source of uncertainty.

Waste loads These are the loads of SPM in the associated riverine discharges taken into account in the model.

Parameters This file contains the information on the surface and bottom depth. This parameter is different for each cell, but unchanged in time, in other words the bathymetry is unchanged throughout the entire model. This however introduces an extra uncertainty in the model, because in reality the bathymetry would differ due to transport of sediments, this however falls outside the scope of this study.

Shear stresses by ships These are the shear stresses induced by ships. Gathered from wave buoys and recalculated by means of the SWAN model output, the yearly averaged wave fields [Cronin, 2014]. This file contains information about the τ_{ship} for each segment in space and time. This input is correlated to the hydrodynamics.

Initial file This file contains the SPM concentrations on the initial time step for each cell. Which can come from the restart file of the a previous model run, or measurements. This can also be a source of uncertainty in the model.

The uncertainties of the SPM forcing stems from the propagation of uncertainties in the hydrodynamic forcing, initial conditions and boundary conditions; in addition to uncertainties in parameterization of processes such as water-bed exchanges of sand-mud mixtures [El Serafy et al., 2011]. Key parameters are Critical Shields stress for sand mobilization in the buffer layer (TauShields), Van Rijn Pickup factor from the buffer layer (FactResPup), Critical bed shear stress (per fraction) in the fluff layer (TcrS1, IMi) the first order resuspension rate from the fluff layer (VResIMi) and the setting velocity (per fraction) [Cronin, 2014]. The model was already calibrated for these parameters. Validation was also done with a set of surface in-situ validation data from MWTL (Rijkswaterstaat), MUMM and CEFAS were used. The model results were also plotted against MERIS data that was interpolated onto the model grid. The bias and root mean square errors of the measurements against the model were assessed. From this it was concluded that nearshore concentrations had a lower bias than offshore concentrations.

5.6 Summary

In this chapter the setup and characteristics per model were described and from this uncertainties to be assessed were identified. Because in the model setup the SPM model is a new asset, the uncertainties in this model are addressed. The model has been calibrated and validated in earlies studies, therefore the sources of uncertainty to address are narrowed down to the uncertainties coming from input files. Regarding the setup and looking at what input is needed for this model, four files are identified of containing uncertainty, the constants, waste loads, shear stresses induced by ships and the initial file. However input as bathymetry and hydrodynamics also introduce uncertainty, computational wise it is not feasible to take these uncertainties into account in this study.

6 Quantifying uncertainties

In this chapter, the quantification of uncertainties in the SPM model is elaborated. In the previous chapter the sources of the uncertainty were identified; the quantification of the uncertainty in the model output will be done using a sensitivity and uncertainty analysis as described in Chapter 3 Methodology. The area of interest in this study is the Wadden Sea, therefor the grid is cropped to this area and some surroundings. However as the PA includes the dutch part of the Wadden Sea, the area is extended from the Rhine influence along the Dutch coast all the way to Denmark, to have a better overview of the surrounding processes. The part of the grid that is used in this study is shown in red in Figure 6.1.

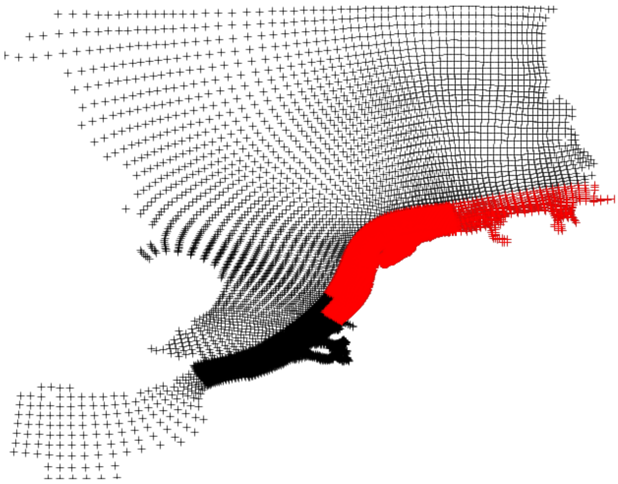


Figure 6.1: In black all the cells in the domain, in red the cells that are used in this study.

6.1 Sensitivity analysis

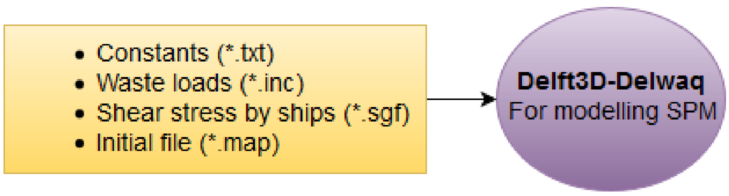


Figure 6.2: The identified input files for the SPM model containing uncertainty.

To see which of these input files cause the main uncertainty in the SPM model a short sensitivity analysis is done. From the previous chapter on identifying uncertainty, the sources of uncertainty in the SPM model are found, see the files in Figure 6.2. From previous projects, there is input data available from the years 2003 to 2011 for the SPM model. This input are all the files listed in the bottom left yellow square in the figure. To see which input has the most influence on the output of SPM concentrations, the model is run multiple times with the same input that only differs for the one file to be assessed. For example, nine model runs are done using for each run the same input of the year 2009 (hydrodynamics, substances, constants, boundary conditions, monitoring points, parameters, shear stresses by ships and initial file) and varying the waste load file for the years 2003 to 2011. This method can not be applied on the constants file, because these constants consist out of one value per parameter (in total 71 parameters), which were the same for all the years between 2003 and 2011. Moreover, were these parameters to differ then it would not be a good assessment to see which individual parameter could have the most influence. However in El Serafy et al. [2013] the sensitivity was already investigated and therefore the information obtained from that report is used for the constants file. Therefore the information from that report is used for the constants file. The parameters are discussed in the next section.

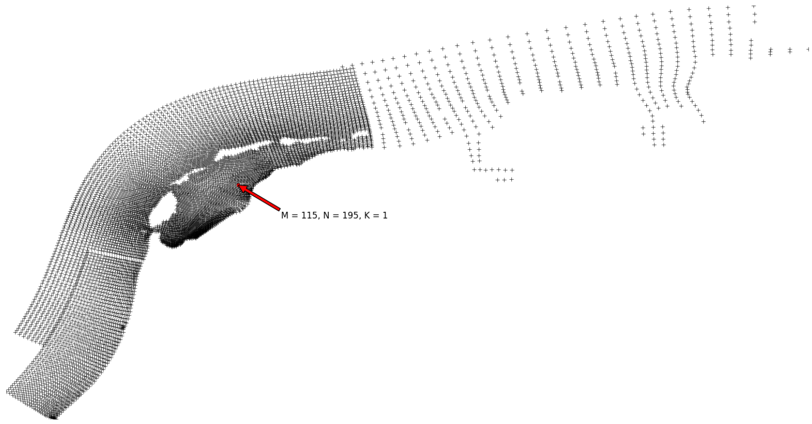


Figure 6.3: Segment 9687, at M = 115, N = 195 and K = 1, the location of the cell that is used as an example in this chapter.

In this chapter the method conducted is shown in figures by time series or data at a certain time step. These figures are posed as an example and for reference the location is always taken on segment 9687, see Figure 6.3. For the sensitivity analysis it is chosen to visualize the time series of the concentrations with a representation of once a week, this is done to better support the explanation of the figures in the sections. Later for the uncertainty analysis it is chosen to use a time step of a day, so there is more detailed information describing the SPM concentrations.

There are two pathways to indicate the variability of the model output, by using the standard deviation and by using the difference from the average to the reference scenario. The figures in this chapter all show in the top image the time series of the different experiments (the grey lines) and the reference case in which all the input files are from the same year, 2009 (the dashed red line). In the bottom image the average of these time series is given (the dashed black line), together with its standard deviation (the grey area) and the reference case (the dashed red line). The average of these time series is calculated by:

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{6.1}$$

The standard deviation is calculated with:

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2} \tag{6.2}$$

x = The SPM value at one location in both space and time for the i-th experiment
μ = Average value of x
σ = Standard deviation of x
N = Number of elements in x

With the standard deviation, as calculated with equation (6.2), the spreading of the output time series is calculated. In the figures this is visualized with the gray area surrounding the average values. Another method is to look at the difference between the mean of the output files (dashed black line) and the reference line (red dashed line), indicating the deviation of the average. These two methods are discussed per case in the sensitivity analysis.

6.1.1 Initial mapfile

The initialization conditions of the SPM concentrations in the model are inserted using a mapfile. A mapfile is s a static set of conditions used to initialize the model runs, this fixes all variables to a singular value at the start of each run. Information of all the calculated substances in every segment is available in this file. In Figure 6.4 the sensitivity analysis on the initial file is shown. It can be seen that the initial and last value differ extremely. The variability of the files is also quite high, the standard deviation is continually present in order of 10 mg/l, however the deviation from the reference case is only in the spring and early summer somewhat different from the reference scenario. Therefore it is chosen to re-use the last values from the reference scenario (where all the input stems from 2009) and rerun the model until a stable SPM field is obtained. This is done in Figure 6.5, where the model is run 6 times in a row, reusing the Restart mapfile from the previous run, in which the last values of the runs are stored. The black line in the figure indicates the stable situation which will be further used in this study.

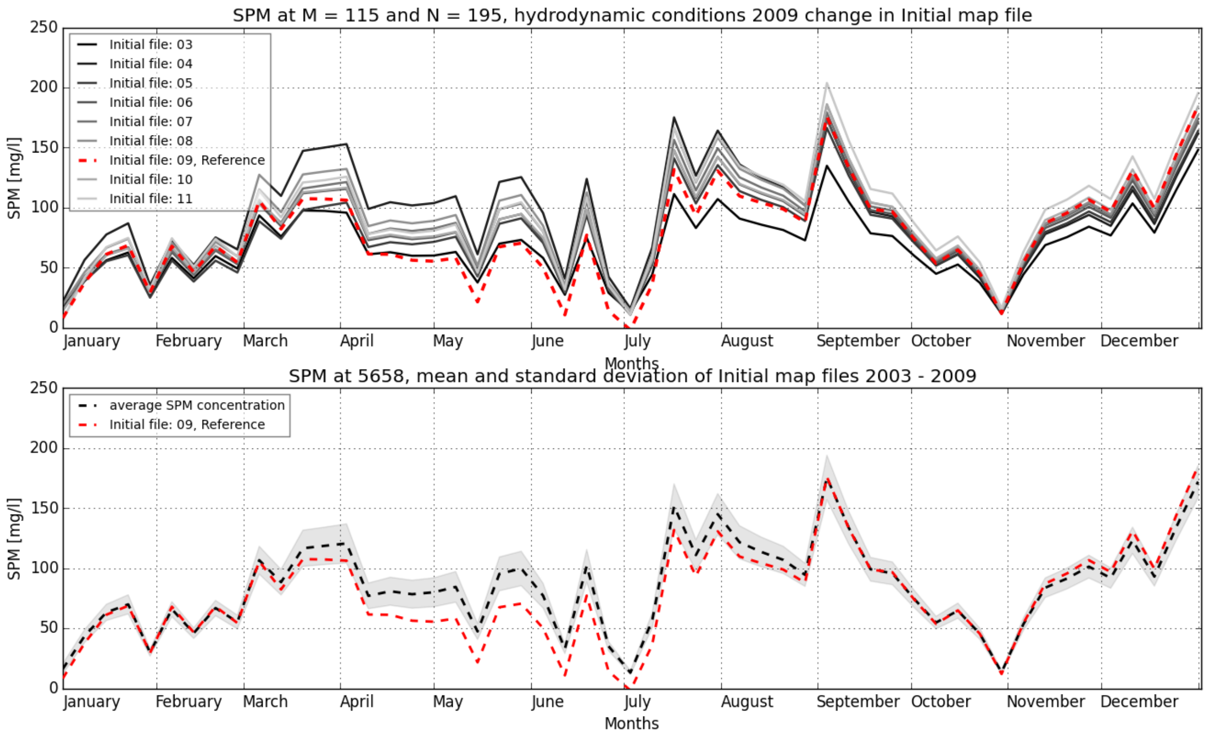


Figure 6.4: Time series of the year 2009 with different Initial map files, the reference situation with initial file from 2009 (red dashed line). Top figure: time series of the 9 representations. Bottom: Time series averaged (black dashed line) and given an uncertainty band by a standard deviation (grey).

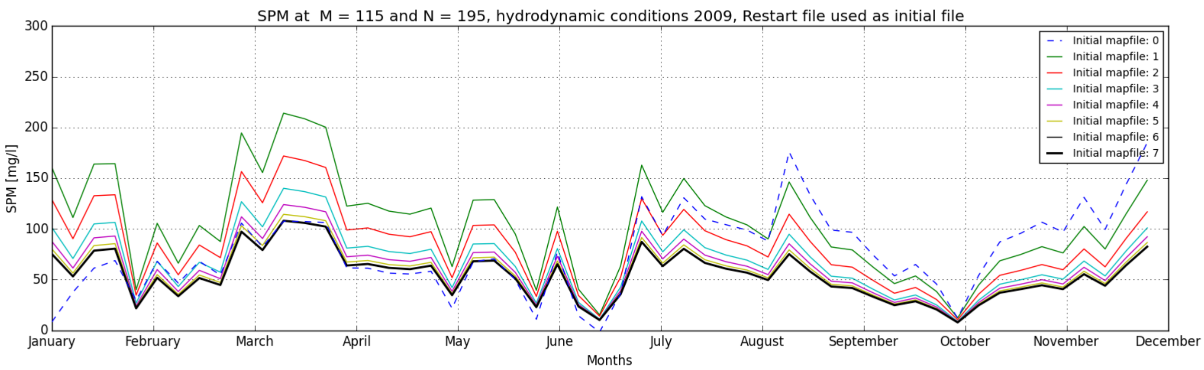


Figure 6.5: Time series re-using the restart file (different colored lines) of the previous model run until an initial SPM value is obtained which is a realistic representation for the entire temporal domain.

6.1.2 Waste loads

Using the initial file discussed in the previous section, the next file to be assessed is the Waste load file, in which the SPM loads from the 91 discharges in the domain are represented. Waste loads are an external forcing to the system in the form of added (or withdrawn) masses per unit of time [Deltares, 2014e]. The loads of suspended sediments from the riverine systems in conjunction from that which is re-suspended from the sea-bed are the sources of inorganic matter in the water column. In Figure 6.6 the sensitivity of the Waste loads are shown. The variability between the reference and the average is within natural variational limits. Furthermore is shown the range of the standard deviation is small, order 1. From these figures it can be concluded that the variability caused by the waste load input is negligible.

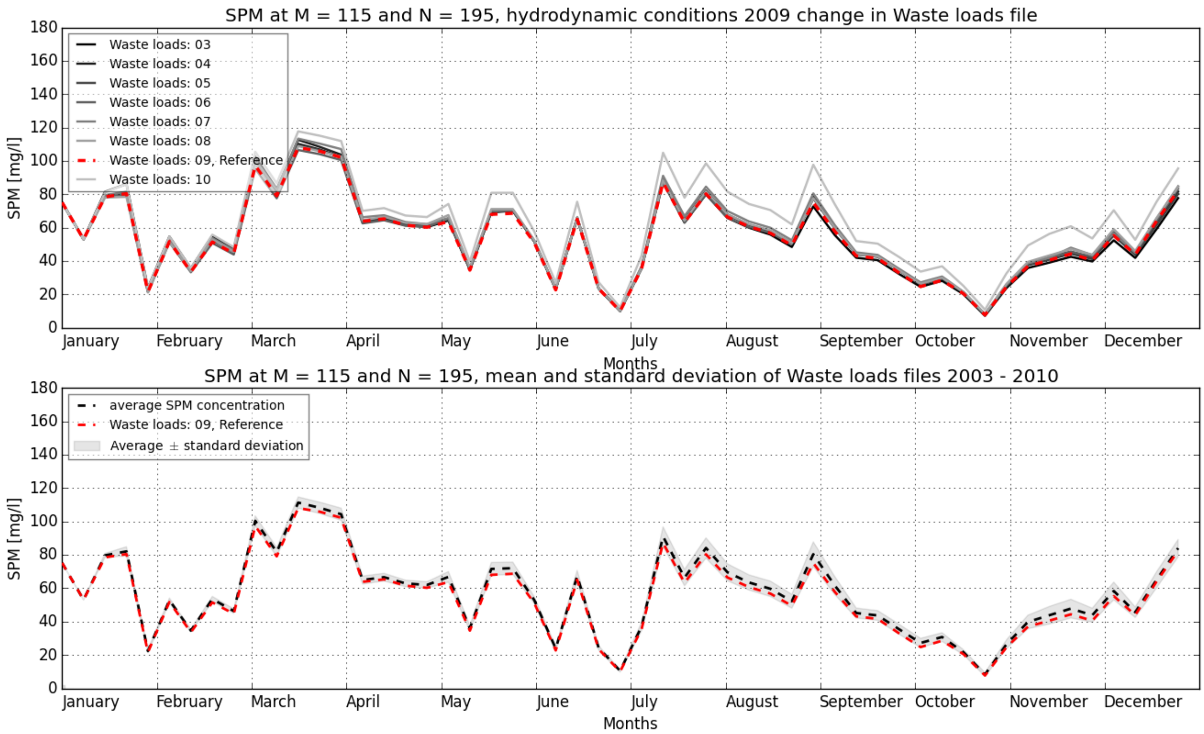


Figure 6.6: Time series of the year 2009 with different Waste load files, the reference situation with waste load file from 2009 (red dashed line). Top figure: time series of the 8 representations. Bottom: Time series averaged (black dashed line) and given an uncertainty band by a standard deviation (grey).

6.1.3 Shear stresses induced by ships

The shear stresses induced by ships, which are calculated using the wave field from SWAN [SWAN team, 2016]. In SWAN, the data for the significant wave height (H_s) and averaged which come from data of wave buoys and are recalculated with SWAN for the entire domain are analyzed for sensitivity. Ships are not singular sources for these stresses, in spite what the name might suggest. Hydrodynamical conditions such as waves, storms and currents all have effect in some extend to the magnitude of these shear stresses. The outcome of this analysis is shown in Figure 6.7. Because the shear stresses are recalculated with SWAN, a link is established with the hydrodynamic forcing in the model. Because of this dependency it is to be expected that the variability of this file will be quite large, when using the data for the shear stresses of different years, where different hydrodynamic conditions are applicable. Taking the figures into account, it can be seen that even with the dependency problem kept in mind that the variability generated by the change of this file is not that significant. Therefore this input file is also not taken into account in the uncertainty analysis.

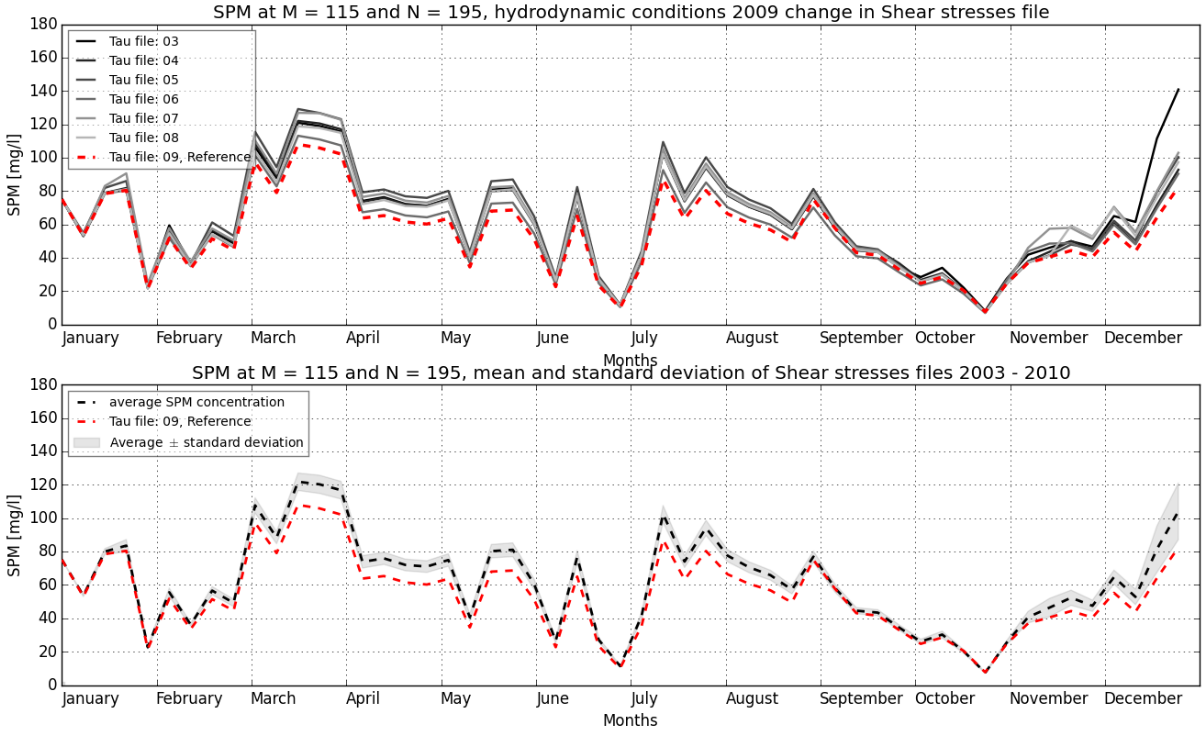


Figure 6.7: Time series of the year 2009 with different shear stresses induced by ships, the reference situation from 2009 (red dashed line). Top figure: time series of the 7 representations. Bottom: Time series averaged (black dashed line) and given an uncertainty band by a standard deviation (grey).

6.2 Uncertainty analysis

The constants file could not be evaluated the same as the waste load file or the shear stresses induced by ships, therefore the results from a previous sensitivity analysis are used. There are 71 parameters in the <constants.txt> file, which are the same for every location throughout the domain. In El Serafy et al. [2013] a sensitivity analysis is already conducted and from this, 10 parameters which have the most influence on the model results are the result of this analysis, listed in Table 6.1. These parameters are used in the equations describing the processes in the model, see section 5.4 of the Literature review on the SPM model. The parameters between brackets are corresponding to the equations 5.2. In this section the correlation between the parameters become clear when looking at the equations given to calculate the deposition and erosion fluxes of the particles and a further explanation is given in section 5.5.2.

Table 6.1: The 10 most influential parameters in SPM model (Adapted from El Serafy et al. [2013])

Parameter	**Explanation**
1. TauShields (τ_{Sh})	Critical shear stress for resuspension
2. FactResPup (F_{ResPup})	Overall factor for resuspension pickup from the sandy layer (S2)
3. VSedIM1 (V_{Sed,IM_1})	Sedimentation velocity of IM1
4. FrIM1SedS2 (α_{IM_1})	Fraction (α) of total settling flux IM1 directly into the sandy layer (S2)
5. VResIM1 (V_{Res,IM_1})	First-order resuspension rate
6. TauRS1IM1 (τ_{cr,S_1,IM_1})	Critical resuspension stress for fluff layer (S1)
7. VSedIM2 (V_{Sed,IM_2})	Sedimentation velocity of IM2
8. FrIM2SedS2 (α_{IM_2})	Fraction (α) of total settling flux IM2 directly into the sandy layer (S2)
9. VSedIM3 (V_{Sed,IM_3})	Sedimentation velocity of IM3
10. FrIM3SedS2 (α_{IM_3})	Fraction (α) of total settling flux IM3 directly into the sandy layer (S2)

The uncertainty analysis on the influence of the parameters on SPM with a Monte Carlo simulation is combined with a Latin Hypercube Sampling technique which incorporates the dependencies between the parameters. This method is explained in detail in chapter 3. The parameters to be assessed are to be given a probability distribution, to simulate the variability of the input. With a Copula the dependencies of the parameters are caught and with the sampling technique the amount of model runs is reduced. In this section the different steps together with the results are elaborated.

6.2.1 Distributions of input parameters

The first step of the process is to give a probability distribution to the input parameters, in order to simulate the uncertainty of the parameter. The influence of a constant parameter is different than varying forcings, therefore, any error therein will introduce a permanent error continuously throughout the model run. It is for this reason that the impact can not be neglected. From the study of El Serafy et al. [2013] the ranges between which the different parameters can vary are given in Table 6.2.

Table 6.2: Ranges and base values of the 10 parameters (Adapted from El Serafy et al. [2013])

Parameter	**Base value**	**Range from**	**Range until**
1. TauShields	0.8	0.4	1.2
2. FactResPup	3.00E-08	8.00E-09	8.00E-08
3. VSedIM1	10.8	5.04	43.2
4. FrIM1SedS2	0.15	0.05	0.4
5. VresIM1	0.2	0.05	0.5
6. TaucRS1IM1	0.1	0.01	0.35
7. VSedIM2	86.4	43.2	172.8
8. FrIM2SedS2	0.15	0.05	0.4
9. VSedIM3	0.1	0.1	5.04
10. FrIM3SedS2	0.15	0.05	0.4

From the study in El Serafy et al. [2013] it was found that the parameters described in Table 6.1 are correlated in pairs. This correlation is taken into account for the uncertainty analysis. The pairs are described in Table 6.3 with a Spearman's rho correlation factor, this concept is explained in section 6.2.4 where the dependencies are taken into account in the sampling method. This factor is taken into account during the sampling for the uncertainty analysis. For the Latin Hypercube Sampling with Dependence this factor is crucial, see next section. The parameters can be described with a multivariate normal distributions, which is used in this study. Together with the information on the ranges and the base value the distributions of the parameters was made, this is visualized in Figure 6.8.

Table 6.3: Correlation pairs and their correlation factor

Pair	**Correlation**	**Factor of correlation (rho)**	**Index of importance**
1. TauShields - 2. FactResPup	positive	0.7	1
3. VSedIM1 - 4. FrIM1Sed2	negative	-0.7	2
5. VResIM1 - 6. TaucRS1IM1	positive	0.7	5
7. VSedIM2 - 8. FrIM2Sed2	negative	-0.7	4
9. VSedIM1 - 9. FrIM1Sed2	negative	-0.7	3

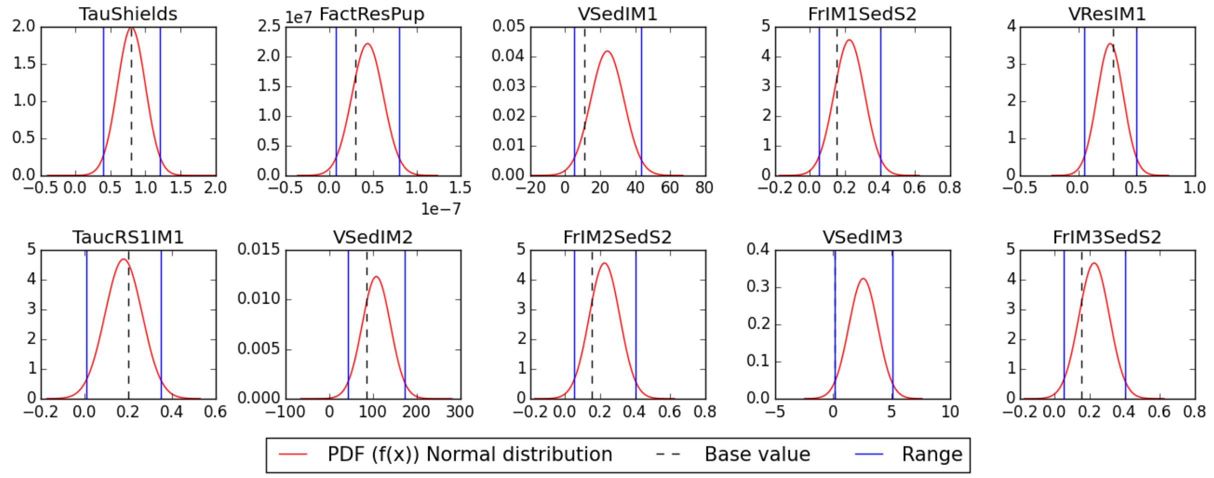


Figure 6.8: Representation of the PDFs of the input parameters, with on the x-axis the values of the parameters (x) and on the y-axis the PDF (f(x)).

6.2.2 Random sampling without dependence

The Monte Carlo method, as described in Chapter 3, for assessing uncertainty uses a random sampling, without incorporating the dependencies between parameters. When using this method, it would take thousands of samples to come close to a represion of the probability distribution of the parameter. This is shown in figure 6.9, where the parameter TauShields is taken as an example. With the histogram showing the samples taken from the distribution and the red line the distribution. This example demonstrate that Monte Carlo would become a good representation of the system after using a few thousand samples.

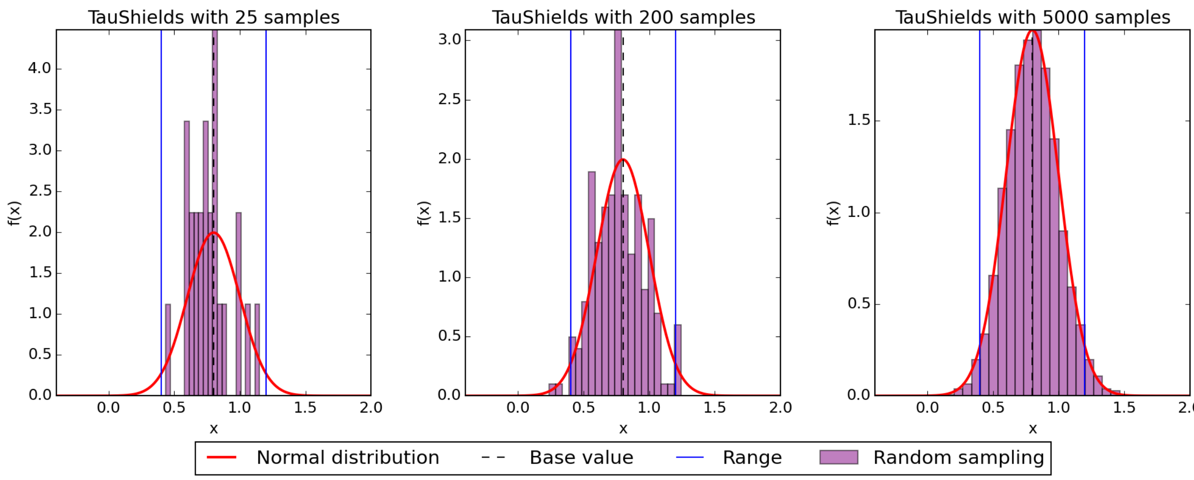


Figure 6.9: Distribution of TauShields with Random sampling, for 25, 200 and 5000 samples.

This method has two disadvantages. First, the amount of model runs needed, secondly it is not able to incorporate dependencies between the parameters. This can be seen in Figure 6.10. The amounts of samples needed is shown in the amount of blue dots in the joint scatter plot. It can be seen that there is no dependency whatsoever in the shape of the cloud, this problem needs be addressed as well. To reduce the amount of samples, an other sampling technique is used, the Latin Hypercube Sampling. To cope with the dependencies between the parameters a Copula is applied.

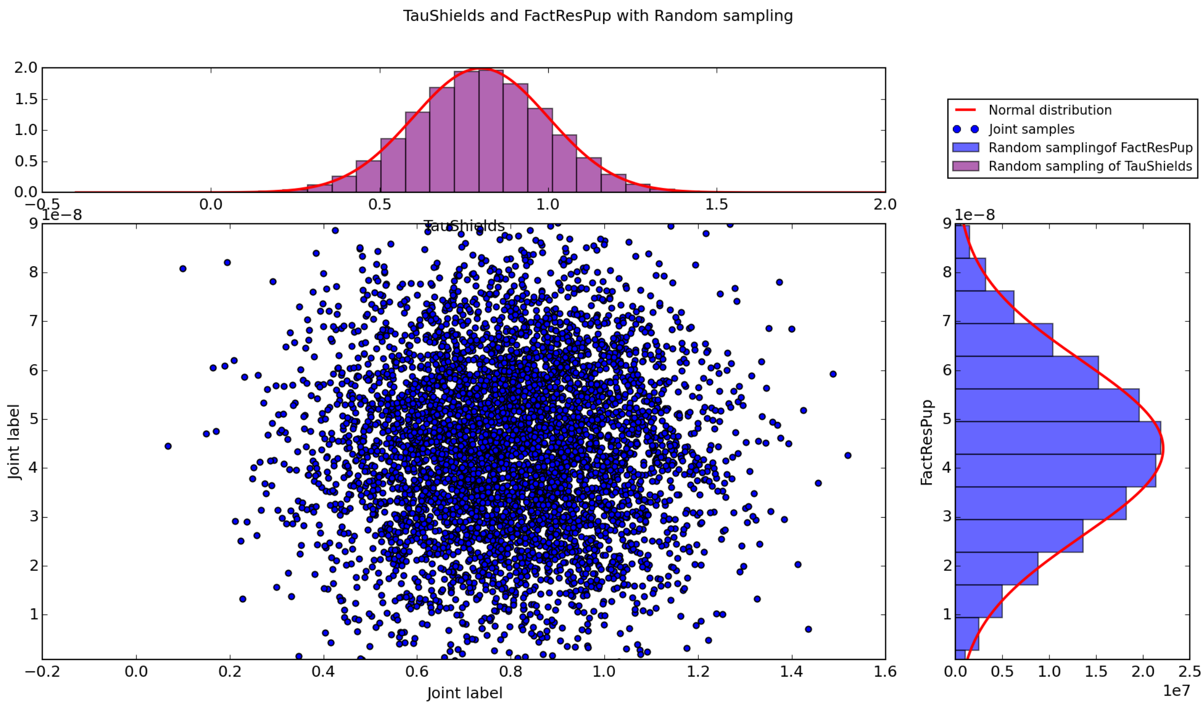


Figure 6.10: The joint scatter plot of the first dependency pair, TauShields and FactResPup, with 5000 samples and no dependency.

6.2.3 Latin Hypercube Sampling without dependence

The same experiment as for the Monte Carlo was done with the LHS technique, to reduce the amount of samples needed for a proper representation of the normal distribution. This can be seen in Figure 6.11. In contrast to the Monte Carlo, where a few thousand samples were needed is the LHS method already quite accurate with a few hundred samples. From the study of Jiayuan [2015] it was proven that between 150 and 200 samples give a good estimation. Therefore the amount of samples at the beginning of the uncertainty analysis in this study is set to 200. Which will be decreased in a later stadium, when the sampling actually takes place, to filter out the non-existing values (negative values for one of the parameters).

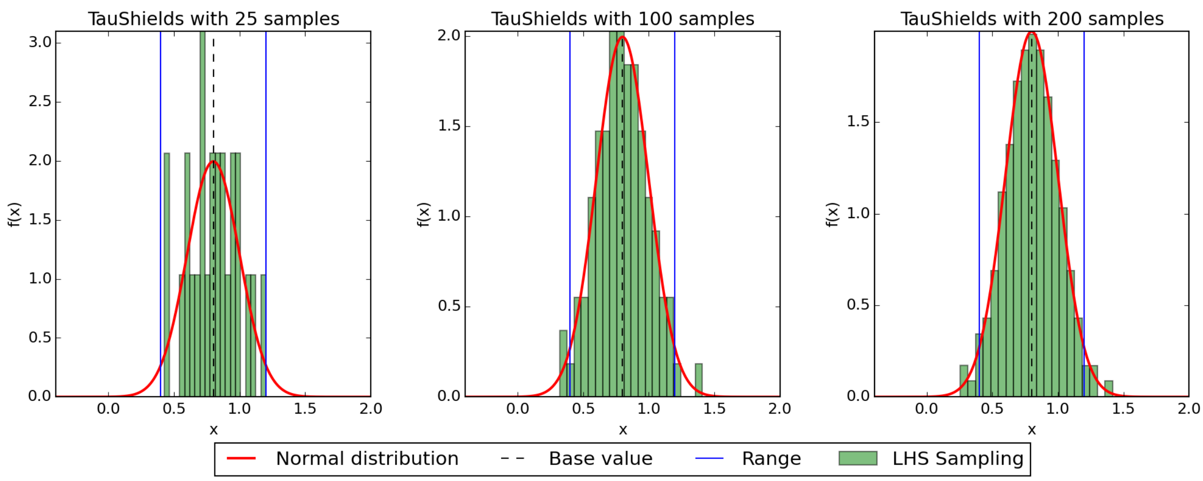


Figure 6.11: Distribution of TauShields with Latin Hypercube Sampling, for 25, 100 and 200 samples.

6.2.4 Random sampling with dependence

Using a Copula the dependencies between the parameters, as described in Table 6.3,are taken into account. The Copula uses a dependency matrix to include the dependencies, see Table 6.4. In this matrix the correlation coefficients are shown, the pairs all had a correlation magnitude of $\rho = \pm 0.7$, as well negatively or positively correlated. This correlation is the Spearmans's rho, which is a ranking for the dependence between two parameters on the interval [-1, 1]. Where -1 is completely negatively correlated, 1 completely postive correlated and 0 is no correlation [Davis and Sampson, 1986]. The correlation of the pairs are known [El Serafy et al., 2013], the other relations between parameters is assumed to be 0. In this way the dependency pairs are taken into account. This matrix is included into the creation of a multivariate probability distribution of the Copula, in which the distributions for each parameter is generated.

Table 6.4: Dependency matrix for the 10 parameters.

	TauShields	FactResPup	VSedIM1	FrIM1SedS2	VresIM1	TaucRS1IM1	VSedIM2	FrIM2SedS2	VSedIM3	FrIM3SedS2
TauShields	1	0.7	0	0	0	0	0	0	0	0
FactResPup	0.7	1	0	0	0	0	0	0	0	0
VSedIM1	0	0	1	-0.7	0	0	0	0	0	0
FrIM1SedS2	0	0	-0.7	1	0	0	0	0	0	0
VresIM1	0	0	0	0	1	0.7	0	0	0	0
TaucRS1IM1	0	0	0	0	0.7	1	0	0	0	0
VSedIM2	0	0	0	0	0	0	1	-0.7	0	0
FrIM2SedS2	0	0	0	0	0	0	-0.7	1	0	0
VSedIM3	0	0	0	0	0	0	0	0	1	-0.7
FrIM3SedS2	0	0	0	0	0	0	0	0	-0.7	1

A Gaussian Copula is used in this case, because the parameters are normally distributed. This Copula uses the correlation matrix as given in the table above, which is the same as the covariance matrix of the multivariate distribution for these parameters. The Copula is made with Copula vectors, which are taken randomly from a uniformly distribution on [1, 0]. For the first correlation pair this method is visualized in Figure 6.12. It can be seen that the distributions of both parameters (blue and purple histograms) are not a good representation of the normal distribution (red lines), due to the low amount of samples. However the joint scatter plot shows a distinct pattern of a positive correlation. As has been mentioned before the amount of samples has been reduced from 200 to 188, discarding of a few non-existing values.

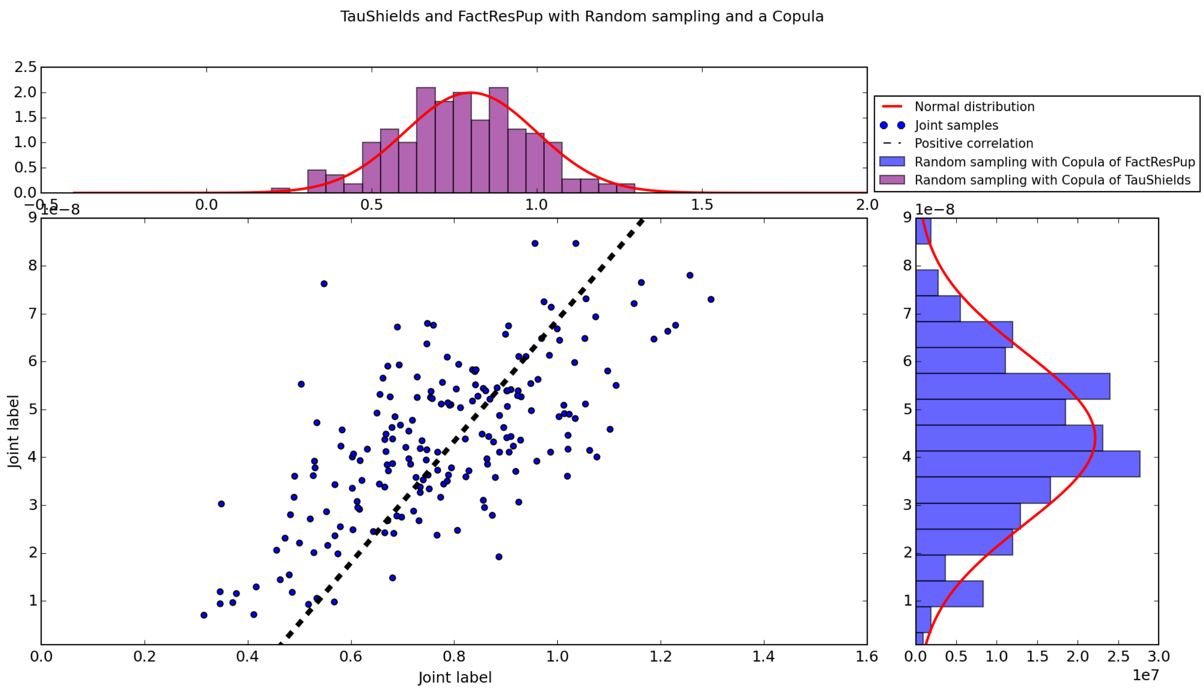


Figure 6.12: The joint scatter plot of the first dependency pair, TauShields and FactResPup, with 188 samples and dependencies incorporated with a Copula.

6.2.5 Latin Hypercube Sampling with Dependence

For this uncertainty analysis a combination of the Copula and the LHS is needed to cope with both the dependencies and to reduce the amount of model runs needed. LHSD transforms samples, such that they are not chosen randomly, but that the marginals U are spread uniformly over [0, 1] and at the same time preserves the Copula made for the dependency links between the parameters. This uniform spreading of the marginals after using the LHSD method is shown in Figure 6.13 for each input parameter.

These marginals are used to create the inverse CDF and from this making a histogram to estimate the PDF for each parameter. This is shown in Figure 6.14, which illustrates the density distribution function of actual samples used in the uncertainty analysis in this project. The red line in the figure indicates the continuous function of the normal distribution. The histogram with the chosen samples of the parameters gives a good estimation of this function. In total, 188 samples are taken from the different parameter input. At the start of the sampling there were 200 samples, but the sample sets with impossible values (negative values) were filtered out.

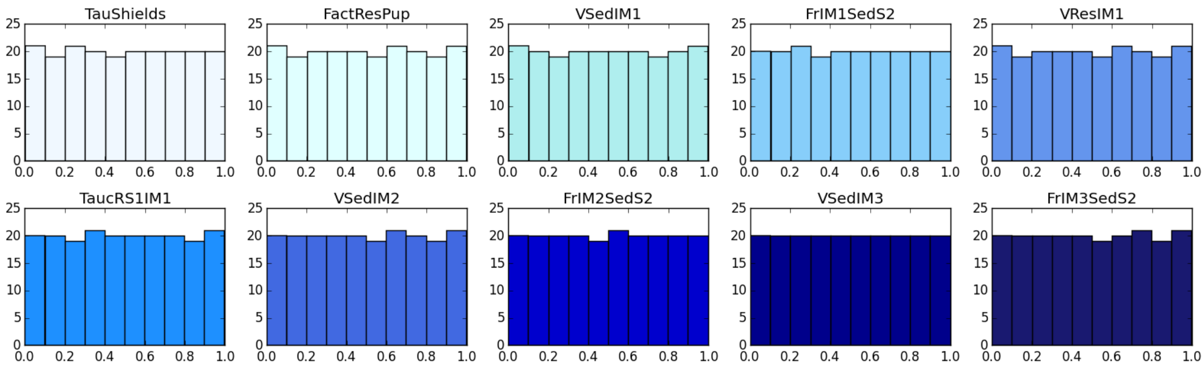


Figure 6.13: Copula vectors chosen with LHSD, forming an almost uniform distribution. On the y-axis the amount of samples within the interval and on the x-axis the intervals that divide the domain [0, 1].

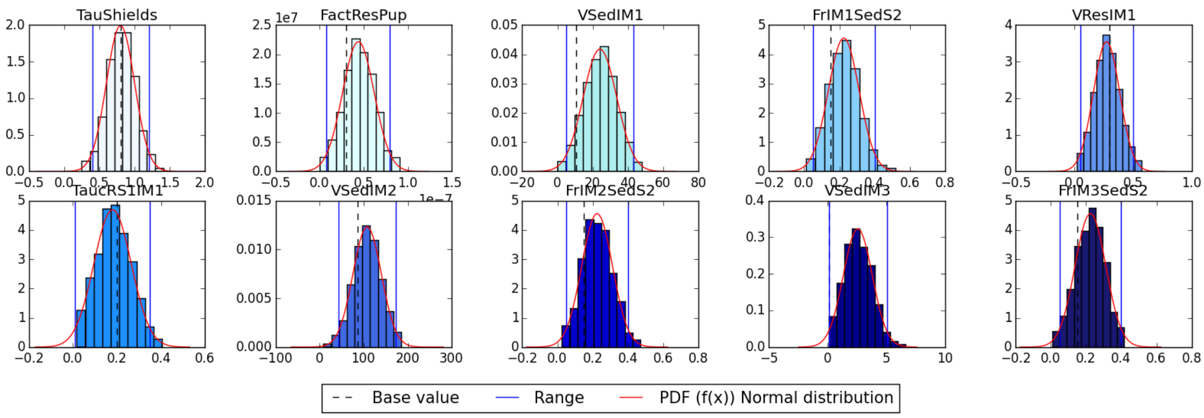


Figure 6.14: Using the vectors made with LHSD to create an inverse CDF into a PDF. On the x-axis the values of the parameters (x) and on the y-axis the PDF (f(x))

From Figures 6.13 and 6.14 above it can be seen that the method reduces the amount of samples, however it has yet to be analyzed whether the correlation is still in tact. In Figure 6.15 the five pairs are set out to each other, showing a histogram for each parameter and their joint scatter plot resembling the multivariate probability function. It can be seen that the correlation is still in tact. The scatter plots that show a downward slope of the correlation fit is negatively correlated, where a upward slope is a positively correlation. The figure in the bottom right shows a joint scatter of the parameters TauShields and FrIM2SedS2, which are not correlated.

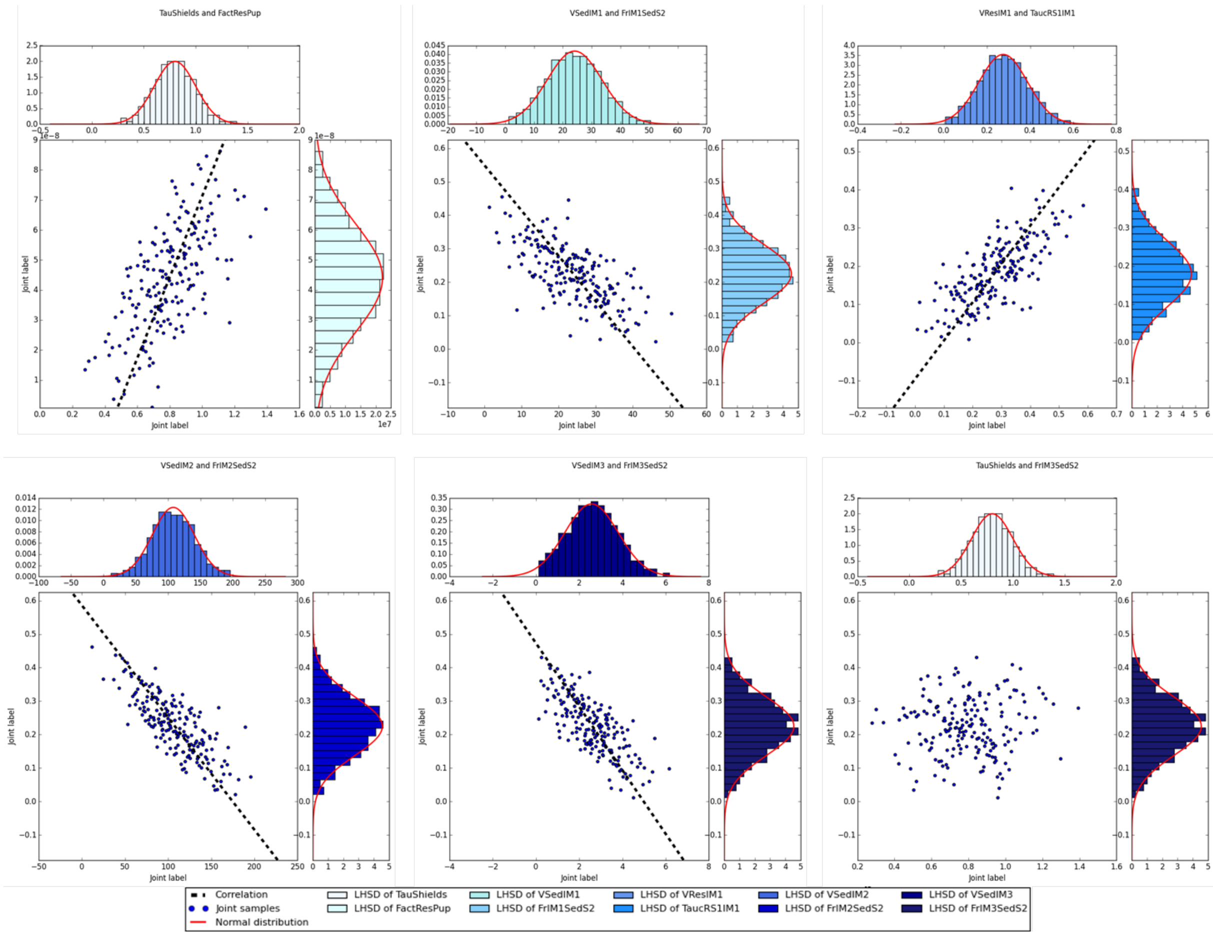


Figure 6.15: The joint scatter plots of the 5 dependent pairs, corresponding to Table 6.3 and one plot of two not correlated parameters (Bottom right), between TauShields and FrIM3SedS2.

6.3 Results

After the LHSD method was applied 188 sets of parameter input was found, which were used to run the SPM model and obtain as much outputs of SPM concentrations for the year 2009 at every location in time and space. The result of the uncertainty analysis are the 188 generated SPM output. In Figure 6.16 in the top figure all time series of SPM output is given at the same location as has been used throughout this entire chapter. It should be noted that this SPM time series does not look similar to the ones in the figures of the sensitivity analysis, as the used time step is 1 day for the visualization of the year. The red line indicates the reference scenario where the base values are taken. In the bottom figure the variability as has been used in the previous section on sensitivity is visualized. In this figure the spreading of the output data, the mean and the reference scenario is shown. In the next section 6.4 Discussion it will be further elaborated what the physical meaning of these graphs is at what this means on the realistic representation of the values for SPM concentrations.

In this section the output will be assessed in such a way that the uncertainty is quantified into a value which can be used in the toolbox in the next chapter. The time series in Figure 6.16 are divided into 12 transects to show as an example in this section what happens on every location at every timestep for the SPM concentrations. Figure 6.17 shows the histograms at these 12 transects, together with the average value (black dashed line), the spreading from the average with 1 standard deviation (blue vertical lines), the median value of the output (yellow dashed line) and the reference scenario (red dashed line).

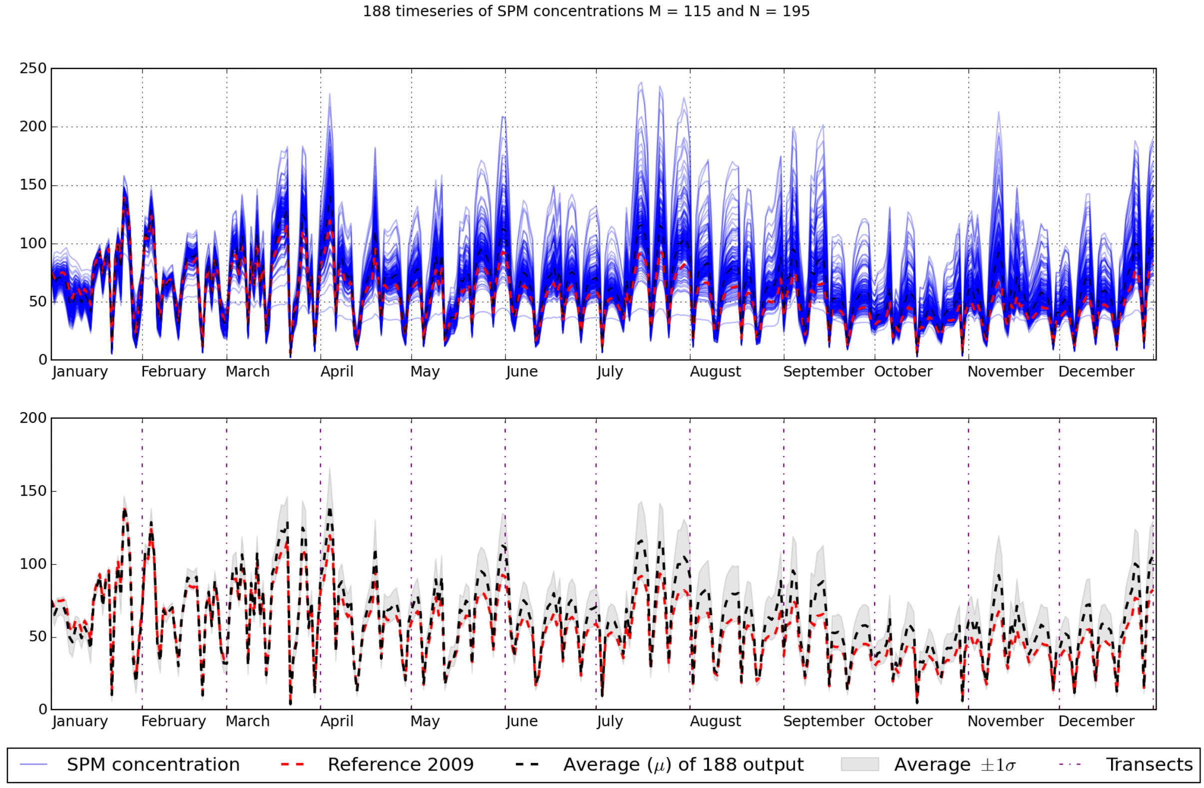


Figure 6.16: Time series of the SPM concentrations of the 188 runs, on the y-label SPM [mg/l] against the reference scenario with the base values as input (red dashed line). Top: 188 realizations (blue lines). Bottom: Average value of these runs (black dashed), with a confidence band (grey area) and the transects chosen at the first day in each month (purple dash-dotted line).

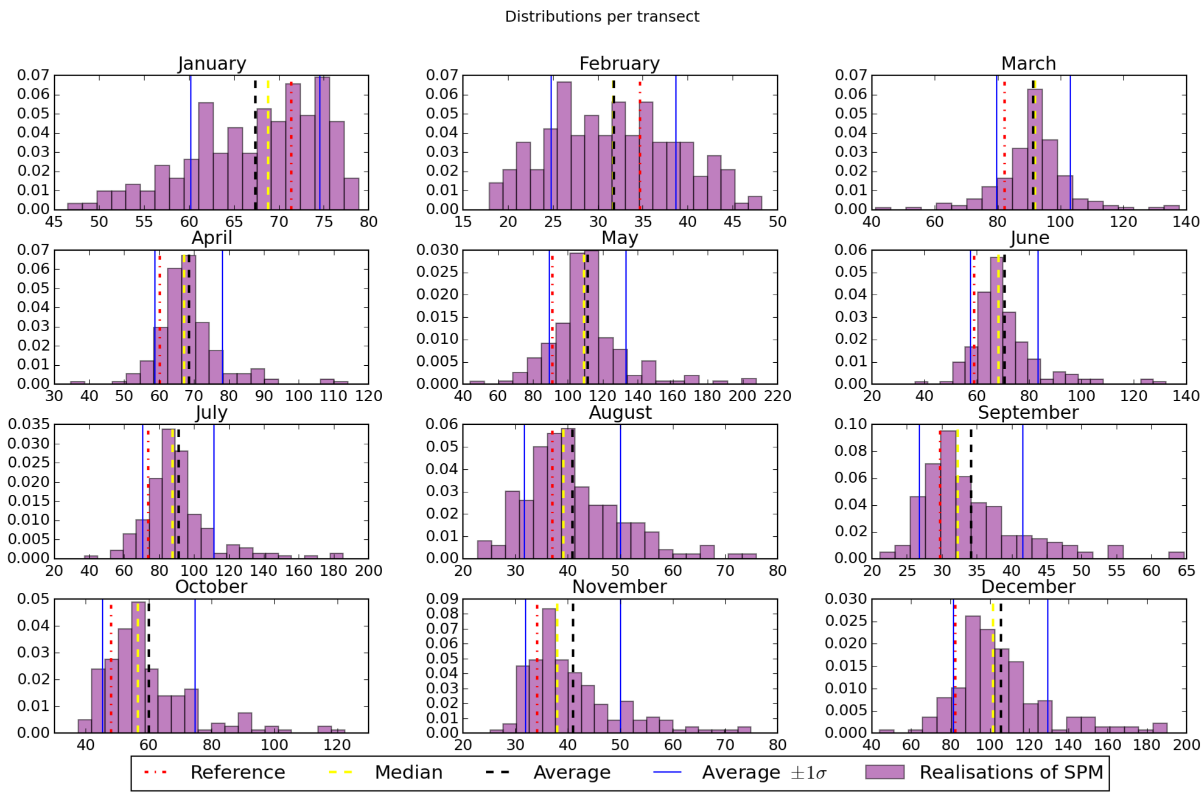


Figure 6.17: Histograms at transects as defined in Figure 6.16 of SPM results.

Using the information from the histograms, the distribution function that best describes the output is found using Quantile-Quantile plots (QQ-plots). A quantile is the percent points of a distribution. A QQ-plot is a probability plot in which the quantiles of the theoretical probability density function are plotted against the quantiles of the SPM output [Scott, 2007]. The QQ-plot calculates the coefficient of determination R^2. This value indicates proportion of the variance in the dependent variable. In this case, when the R^2 goes to 1 it means that the distribution is well fitted. For the example only 1 histogram is chosen to show the method of finding the correct distribution to describe the SPM concentrations, the histogram of December. Later in the section this method is applied on all the locations in the domain. On this histogram the normal and log normal distributions are fitted. In the middle and right figure the QQ-plots for the normal and log normal distributions are shown, with on the x-axis the quantiles of the functions and the y-axis the quantiles of the calculated values. For this example it can be seen that the R^2 of the log normal QQ-plot is higher than of the normal QQ-plot, meaning that the fit of the log normal distribution is a better estimate of describing the SPM output.

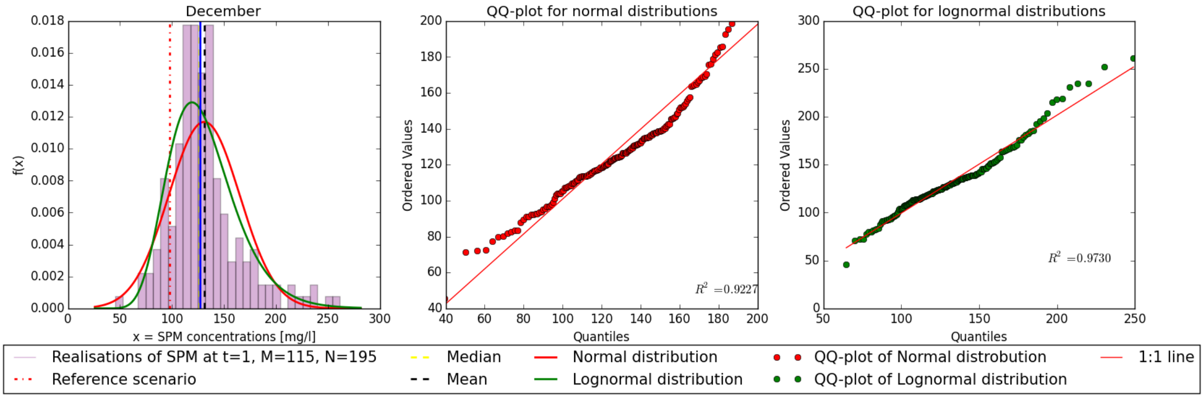


Figure 6.18: Left: histogram of December transect, with a fitted normal distribution (red) and a log-normal distribution (green), bins = 30. Middle: QQ-plot for a normal distribution, $R^2 = 0.9130$. Right: QQ-plot for log normal distribution $R^2 = 0.9642$.

This method is applied over the entire domain, on each segment in each layer and over the entire year with a time step of 1 week, to see which distribution fits in most locations. Conclusion from this test was that the log normal function is the best distribution to use to describe the SPM concentrations, because in almost all locations and time steps this distribution had the best R^2 value.

Concluded is that the log normal distribution the best fit is for the concentration of SPM in the segments. This distribution has an μ which indicates the location parameter and a σ for the scale parameter. These parameters are the average and standard deviation (equations 6.1 and 6.2) of the logarithm of the distribution. Figure 6.19 shows the mean and the median of the normal distribution, where the mode is the value that is most often represented by the data. When the logarithm of the data is used this value would represent the mean. Therefore the mode is used in this study for the indication of the SPM values, where the uncertainty is represented by the scale parameter. The difference with the normal distribution is that the log normal distribution will not take any negative values into account. Which is also realistic for this outcome, because the SPM concentrations can not be below 0. For a normal distribution the average value would be sufficient to use as the concentration value in the visualization to represent the most likely value, for the log normal function this would be the mode.

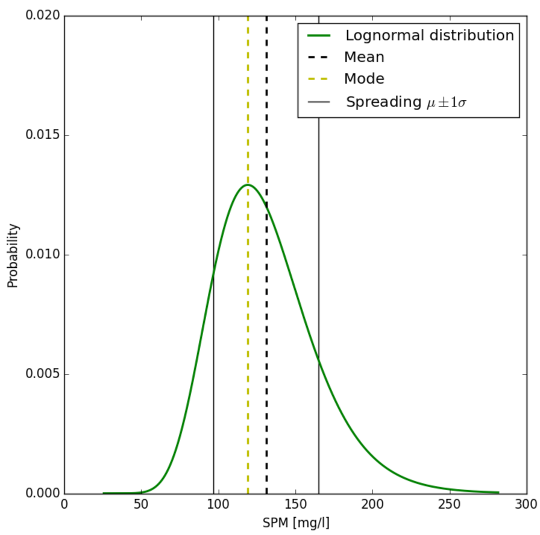


Figure 6.19: Histogram of the December transect, with the mode, median and mean values, showing that the mode represents the data best.

6.4 Discussion

In this section the physical meaning and background of the results will be discussed and a validation is made comparing the obtained model output with three types of data. The concentrations of SPM coming from the SPM model are plotted against the originally used SPM concentrations, data from measurement stations and data processed from the MERIS missions. These data sources are first explained in the next sections and in the last sections the observations of these validation steps are discussed. Furthermore the areas of high uncertainty in the model are pointed out and elaborated. From Figure 6.16 it is noted that the reference scenario and the average outcome are not in line. The average scenario overestimates the output in almost all locations. This is explained when looking at table 6.2, where the base value of the 10 parameters is given. These values differ from the mean value of the ranges, and in all cases it is a lower value. When looking at the equations that are solved in the SPM model (equation 5.2 to 5.6) it can be seen that when these parameters decrease it will have a decreasing effect on the equation outcome. This is the case for all the parameters, except the TaucRS1IM1. Furthermore from the last section the SPM concentrations is estimated with a log-normal distribution and the mean value is changed to the mode value of this distribution, decreasing the differences between the reference and the mode. This can be seen in figure 6.21 where the mode is shown in a black dashed line and the reference as a red dashed line for four different locations.

6.4.1 Validation of SPM model output

The originally used SPM concentrations are given by a segment function, which means that the SPM concentrations are needed for each segment and for each time step [Deltares, 2014e]. In earlier models only the IM1 fraction could be used and therefore the original data only contains this IM1 fraction. Because this is only part of the total SPM values, these IM1 values will always be lower. This file has a time step of one day in which the IM1 values are represented. The data is plotted by the blue dashed line in Figure 6.21. This IM1 data comes from two models, the PACE (DELWAQ-SPM_Waddenzee-PACE_j09_v01) model for the slib concentrations in the Wadden Sea area and DELWAQ-SPM_Noordzee-ZUNO-DD_j03-11_v02 for the rest of the North Sea [Arentz et al., 2012]. The input for GEM/BLOOM is an combination of these models together with satellite data obtained from the MERIS missions [Perez, 2015].

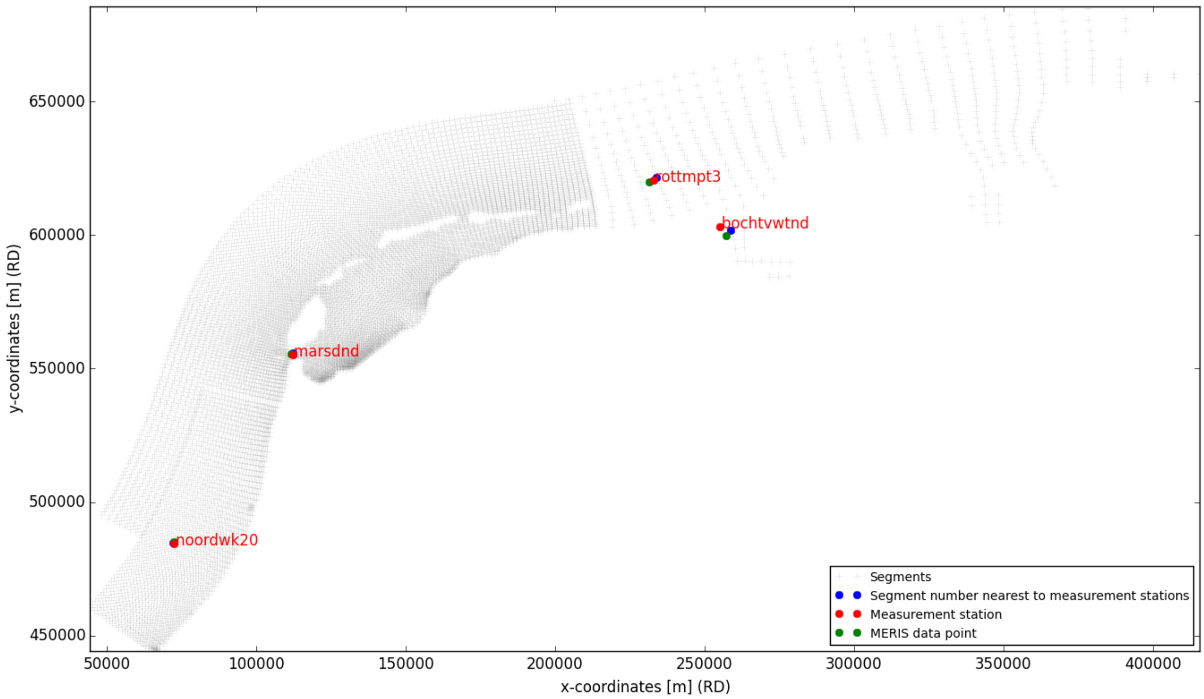


Figure 6.20: Locations of 4 stations from Rijkswaterstaat (red) and the nearest found locations of the MERIS data (green) and the nearest segment within the SPM model (blue).

Using the ocean color retrieved by MERIS (MEdium-spectral Resolution, Imaging Spectrometer) provided SPM data in previous projects [El Serafy et al., 2007]. Where IVM's HYDROPT algorithm (van der Woerd and Pasterkamp [2008] and for SPM explained in [Eleveld et al., 2007]) to estimate SPM concentrations from MERIS reflectance. This data has already been calculated in previous study and could therefore easily be used as a validation data [El Serafy et al., 2007]. It should be noted that from satellite images only one value for SPM is calculated for the entire water column, because the image is in a 2D space and therefore per location only one SPM value will be available. This means that the processes in the vertical water column are merged into one value. However more realistic is that in the lower parts of the water column more SPM is available (due to gravity and settling forces) and less in the upper parts, however this and other processes can not be taken into account.

The measured data comes from the measurement stations of the waterbase from Rijkswaterstaat (Rijkswaterstaat [2016], Deltares [2013]). These measurements are from the Monitoring Waterstaatkundige Toestand des Lands (MWTL). In Figure 6.20 the locations of the measurement stations used in this section to validate the SPM concentrations are shown. There are many more measurement stations, however the choice was made to use only these four stations, due to two reasons. Firstly, it was chosen to have a station on all three parts of the domain: the rottmpt3 (Rottumerplaat 3 km from the coast) and bochtvwtnd (bocht van Wattum north) both are situated on the coarse grid, while marsdnd (Marsdiep north) is on the inter grid and the noordwk20 (Noordwijk 20 km from the coast) is on the aggregated fine grid. The second reason for this choice was that only a few stations contained data from the year 2009. Further more they all lie on interesting points in the area, where marsdnd is in an inlet where ebb and tide are of great importance and will cause a continuous disturbance in water column regarding the SPM concentrations. The noordwk20 station lies 20 km outside the coast, which, regarding the other stations lies in a more stable offshore position. bochtvwtnd lies in the estuary, which has a great influence from the incoming tide and is therefore also highly dynamic. In Figure 6.21 the cyan colored squares indicate the measured data on the dates available for that year. The data represents the concentration of suspended matter in the water, this means all the organic and inorganic fractions present in the water column, measured at the surface in mg/l [Deltares, 2013]. The comparison is made with the top most layer of the SPM model.

In Figure 6.16 the SPM results of the uncertainty analysis are plotted against multiple validation data. The black line indicates the mode of the 188 model outputs together, plotted with an uncertainty band indicating the 95% confidence interval of this data. The validation data is used to check is this black line and its confidence interval are approximating this data. The red line is the reference scenario where the base values of that year were implemented to see if this reference case as has been used in previous studies is within the confidence bands, which it is. This was already expected because the base values never represented an extreme value for the input parameters. The blue dashed line indicates the original data from the sgf file. This data does not follow the calculated data, underestimating the values at most times. Concerning that the IM1 if only a fraction of the SPM data this is also as expected. Furthermore the data does show a somewhat similar pattern, having peaks at the same locations. The data from Rijkswaterstaat and the MERIS data are not continuous throughout the year and are therefore only good for an indication of what the order of magnitude of the data should be. It should be noted that this data also contains uncertainty (in the measurements) and is not taken into account in this validation, because it is used for a general validation of the correctness. The MERIS data is missing for the Bocht van Wattum. Conclusions that can be drawn on this validation is that in general the order of magnitude in all locations are correct. Most data points lie within the confidence interval, indicating that the estimate of the SPM is correct. It should be noted that the model output has a time step of one day (starting from 01-01-09 at 00:00). Due to this the data points from either Rijkswaterstaat or MERIS will never be in line with the model data and therefore a really accurate comparison can not be done. The influence of factors changing on a smaller scale than diurnally are not completely taken into account. For further validation of the model, the same method could be redone on a smaller period of time with a much smaller time step.

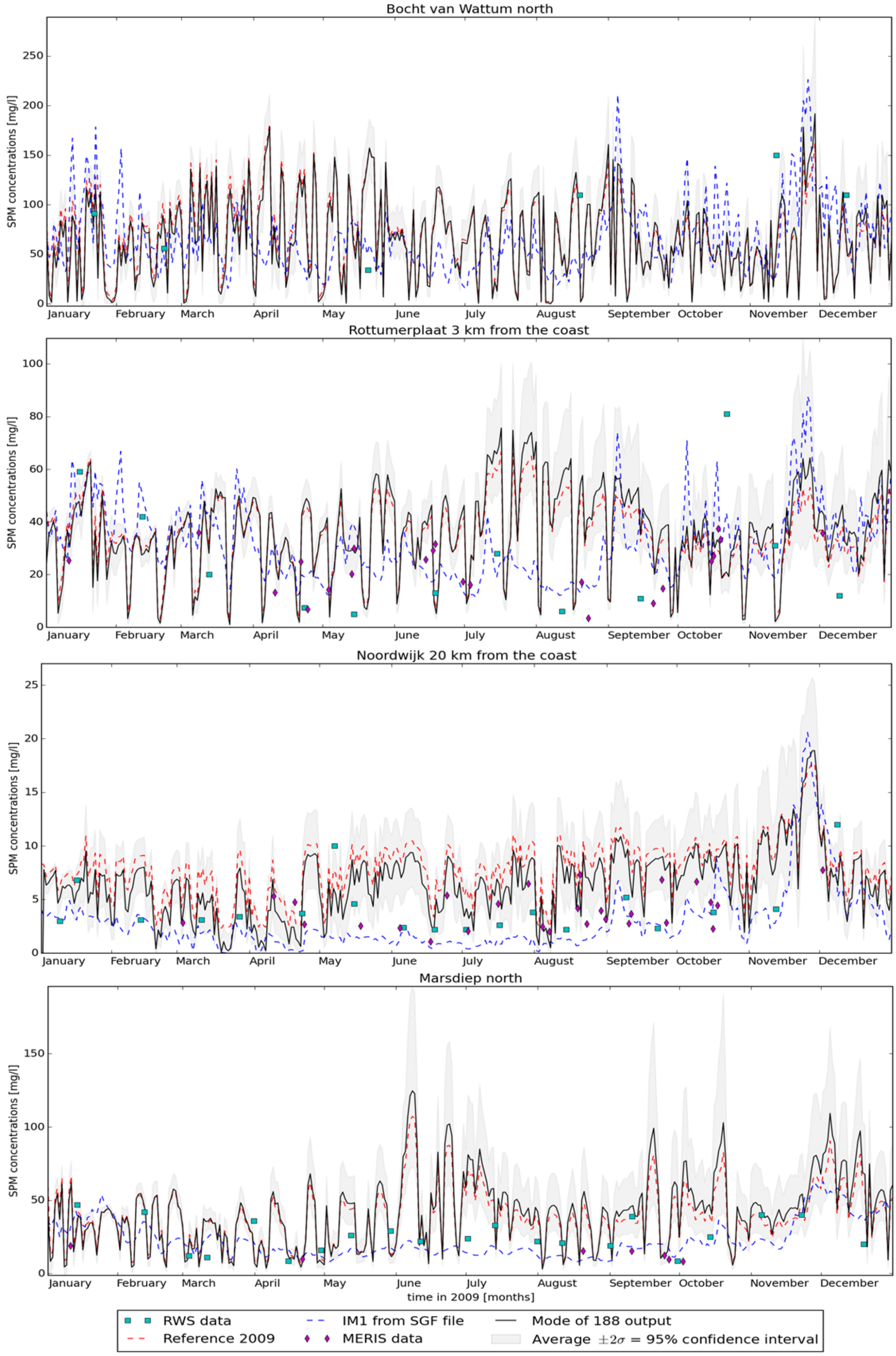


Figure 6.21: Validation for the SPM concentrations from the model, using the mode (black) and the spreading (grey area), compared with reference from 2009 (red dashed), original data (blue dashed), MERIS data (magenta squares) and Rijkswaterstaat data (cyan squares) at four locations within the domain.

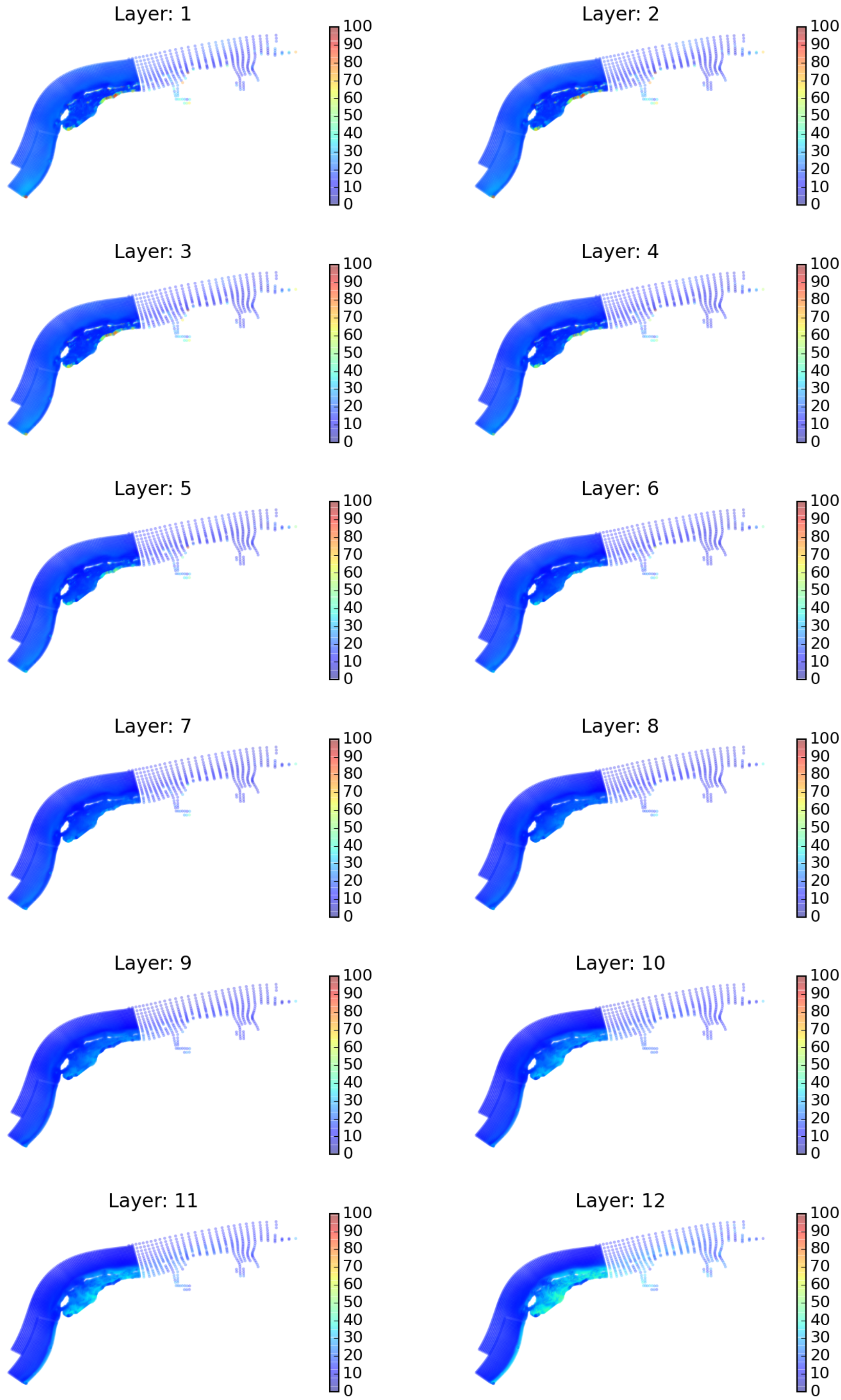


Figure 6.22: Risk map visualizing high uncertainty regions.

6.4.2 Locations of uncertainty

Using the quantified uncertainties, a risk map can be made, by averaging the uncertainty of the entire year and visualizing this on a 2D map. This is done for each layer and is shown in Figure 6.22. From this information it can be concluded that uncertainty is very high in areas where the concentrations are very high and where the processes are highly dynamic and dependent on multiple factors. For example in the Wadden Sea area, where the tidal influence and local processes are very dynamic in time. In Chapter 8 the physical background of the area is discussed in more detail, and here the link between the high uncertainty and the physical processes are further described. However from Figure 6.22 it can be concluded that in general the model has a low uncertainty in the offshore areas. But in the complex areas the model has some exceptions and these areas might need to be assessed again. One reason why these areas contain more uncertainty might be because the bed load module (as described in Section 5.4) is not entirely complete. It is possible that the many factors influencing the SPM mechanisms at a critical area are not correctly or completely taken into account. The dutch coastal region is influenced by storm impact, eroding the coast. The waves influencing the hydrodynamic aspects could be another source for the high uncertainty in these critical areas. From this study the critical areas can be pointed out, however extensive research is needed to further validate and optimize the model.

To summarize this chapter: The SPM model gives a good estimation for the SPM concentrations in most locations within the domain. However in critical areas where the uncertainty is very high, such as outflows of rivers and parts were the SPM processes are very dynamic. For the quantification of the outcome of the uncertainty analysis a log-normal distribution is used to estimate the 188 outputs. The characteristics of this distribution, the mode and the spreading is used to quantify the concentrations of SPM and the uncertainty into a value that can be used in the visualization in the next chapter. The SPM values with this range of uncertainty is validated against measurements from Rijkswaterstaat, which is in all cases in the same order of magnitude and mostly follows the pattern throughout the year. Another validation is against the previously used IM1 data for the GEM/BLOOM model input. The same pattern throughout the year is shown, however the SPM overestimates these values. This is explained to an extend by the fact that IM1 is only a fraction of the SPM, however a big fraction and therefore it is also an indication that the SPM concentrations are plausible and realistic. A last validation was done using SPM values converted from satellite data. This gave a similar result as the measurements. Where a comparison is very tricky, because the time instances do not overlap. For the general assessment of the model it can be said that the model gives a reasonable estimation for the SPM values. The regions of high uncertainty can be identified using a top view projection per layer. From this map it is observed that the highest uncertainty occurs in regions of highly dynamic and influenced by many factors, such as the outflow from a river or the tidal flats of the Wadden Sea.

7 Visualizing uncertainties

The toolbox developed in this study is used to visualize both data and uncertainty in a 3D environment. In this chapter this visualization is presented and the different functionalities are described. To explain the method used for creating the toolbox, a case study is used: SPM concentrations in the Wadden Sea with uncertainties coming from the parameters, as quantified in the previous chapter. Section one describes the aspects of visualizing SPM (or other model output) in a 2D and 3D environment. The second section describes the visualization in the developed toolbox. The different aspects are: visualizing uncertainty, the cutplane, selection of a marker, time implementation and other features are described.

7.1 Visualization of SPM in the Wadden Sea

The SPM concentrations are calculated with the Delft-3D WAQ model for SPM (see Chapter 5 for more details of this model) and are represented in Figure 7.1. The Total Inorganic Matter (TIM) $[mg/l]$ in the water column is calculated with the SPM model and gives a indication of the SPM in the different layers. The TIM is the total suspended inorganic matter in the water column, whereas SPM also includes organic matter. However this organic fraction is very small, therefore the TIM is used as the indicator of the amount of SPM in the water column. In this chapter TIM and SPM can therefore be interchanged.

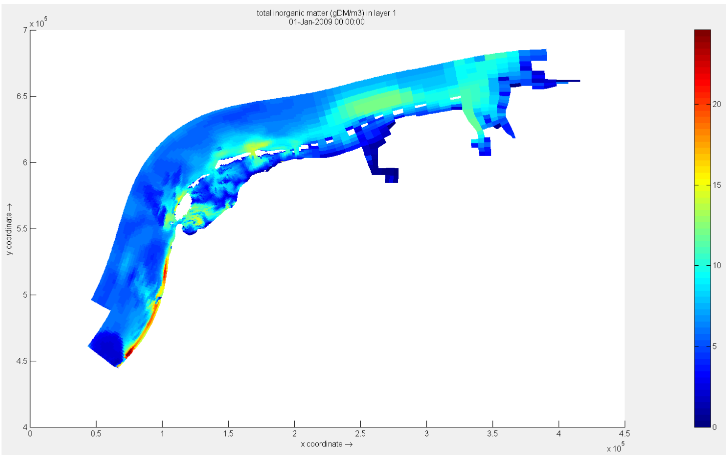


Figure 7.1: Plan view of TIM $[mg/l]$ in the Wadden Sea area at January 1, 2009 in the topmost layer visualized, X- and Y-coordinates in RD $[m]$.

The entire domain in the Delft3D-WAQ model is 3 dimensional and therefore, the data is best visualized in a 3D environment, making the data for all segments available in one interface. The concentrations of TIM are calculated in all 12 layers at every time step.

It was chosen to develop the toolbox using the programming language Python. With Mayavi ([Ramachandran and Varoquaux, 2011b]) an easy interaction and visualization of 3D data is made. This data visualizer can be used within the Python environment. This visualization creates an interface in which the 3D data is plotted. Figure 7.2 shows the visualization of the same data represented as in Figure 7.1. Per segment in the model output a marker on that location represents the segment, where the color is the indication of the concentration of SPM at that point in $[mg/l]$. Further details on the technical requirements to use this toolbox are described in Chapter 8.

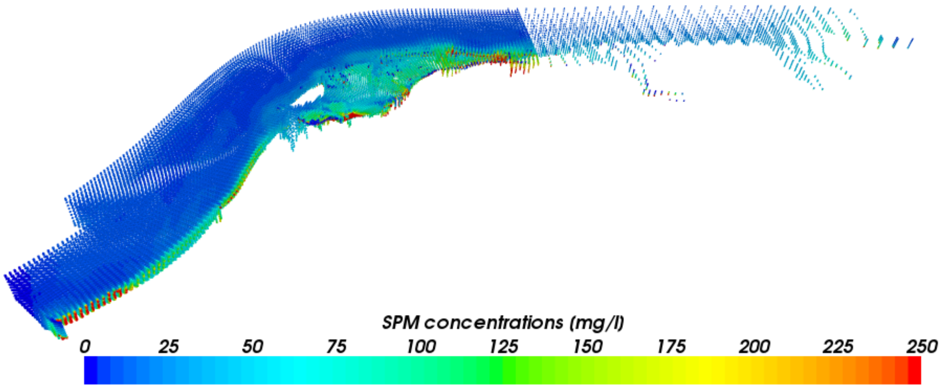


Figure 7.2: 3D environment of the TIM in the Wadden Sea area at a time step in February 2009, in all segments visualized with mayavi.

7.2 Toolbox aspects

A still of the developed toolbox is presented in Figure 7.3, visualizing the SPM concentrations in the Wadden Sea area. In this figure the different functionalities of the toolbox are shown, which are explained in more detail in the next sections. The SPM concentrations are represented in the 3D environment (pane A) by markers with colors indicating the amount of SPM. The colors indicate the amount of SPM and corresponds with the colorbar. The white tint of the markers indicates the uncertainty of the calculated concentrations. The red outline is the cutplane which makes it possible to make a cross section of the area. Bar B shows the different buttons, by which the animation can be started and stopped and the time step is indicated. The last button draws the cross section at the location of the cutplane. When selecting a marker in the 3D environment (by means of a mouse click), the bottom image changes and gives the time series of the SPM concentrations at the location corresponding to the selected marker.

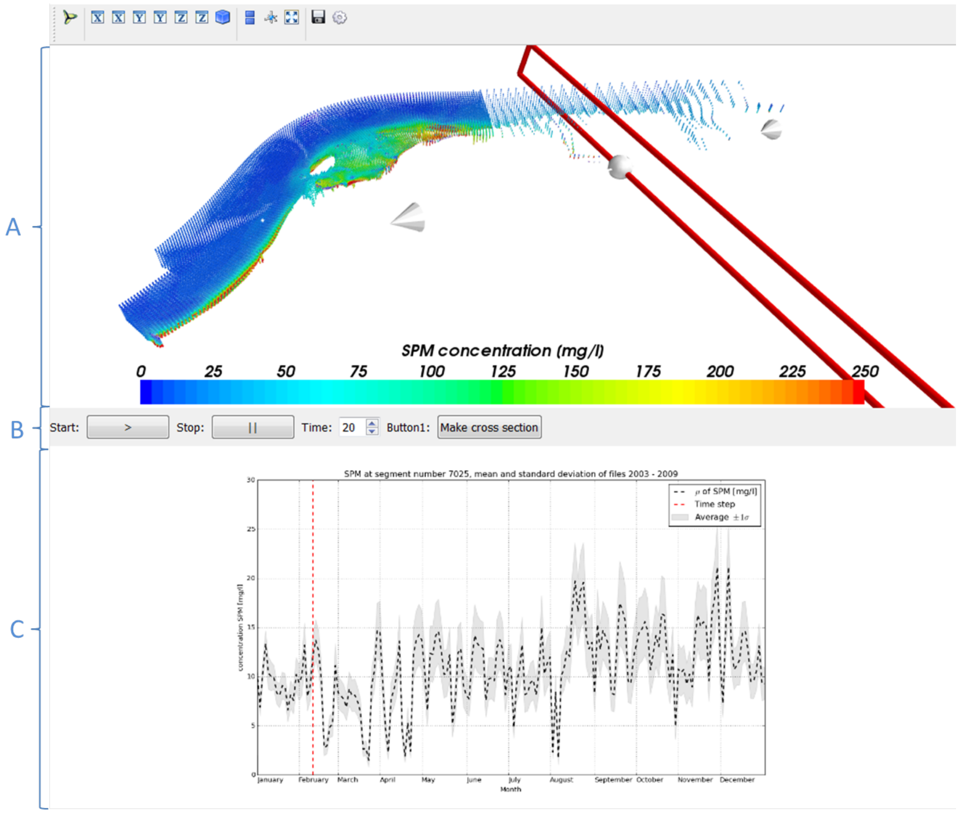


Figure 7.3: A still of the toolbox visualizing SPM concentrations in the Wadden Sea area. Top pane (A) the 3D environment. Bar (B) containing buttons. Bottom pane (C) Time series of one selected location (segment 6951).

7.2.1 Visualization of uncertainty

As the representation of the concentrations without implementing uncertainty (Figure 7.1) can give a biased representation of the system, it is also important to present data along with its uncertainty to make knowledge-based decisions. In the previous section model output was visualized in a 3D environment. From the literature study it was found that a combination of displaying both data and uncertainty is possible by using different properties of the markers indicating the data in a location. Therefore, several methods such as: changing the size, shape, opacity, color hue, transparency, orientation or texture are tested. Because of the limitations of the programs used, not all these options could be applied to visualize the uncertainty and eventually it was chosen to use a combination of the characteristic of opacity and hue to visualize uncertainty. Objects made in Mayavi are limited to only apply one scalar per object, therefore, in the visualization one object with markers is used for the color according to the concentrations. Another object to indicate uncertainty is also an object with markers, which are placed on the exact same location as the other markers, is used to apply opacity over the markers. This object contains completely white markers that are slightly larger than the markers of the first object, surrounding these markers entirely. When applying an opacity to these second markers, the marker within (containing the concentrations color) is visible with a white hue. The whiter the marker, the more uncertain the concentration value in that location at that time step. This concept is illustrated in Figure 7.4, where only one color is used for 5 markers (blue), and the opacity of the surrounding marker is changed. The opacity is indicated by an α, where 0 is complete opaque (only the blue marker is visible) and 1 is no opaqueness (only the white marker is visible).

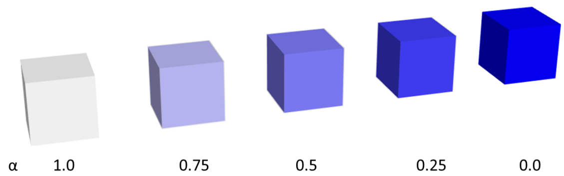


Figure 7.4: Example of the implementation of uncertainty in the toolbox. The color indicating a value for the marker and the white hue indicates the uncertainty. From left (completely white) the uncertainty is 100% to right (completely blue) where the uncertainty is 0%.

7.2.2 Cutplane

Within the visualization a cutplane is available, which is a plane that can be dragged through the visualization. The cutplane clips the data, so only the data on one side of the plane is still visible. This cutplane makes it possible to get more information on the inside of the 3D area. For projects that need information on really specific locations it can be a very useful tool. The cross section is made by moving the cutplane and clicking the button to create a plane corresponding to the cutplane. With this vertical plane the layers from surface to bottom are displayed. On a wanted location it can make a cut of the 3D environment and after a cross-section of the area is displayed. The coordinates of the normal and the origin of the cutplane object are used to calculate which markers are on which side. On one side the markers are than cropped out from the entire domain, leaving the markers on the other side of the plane.

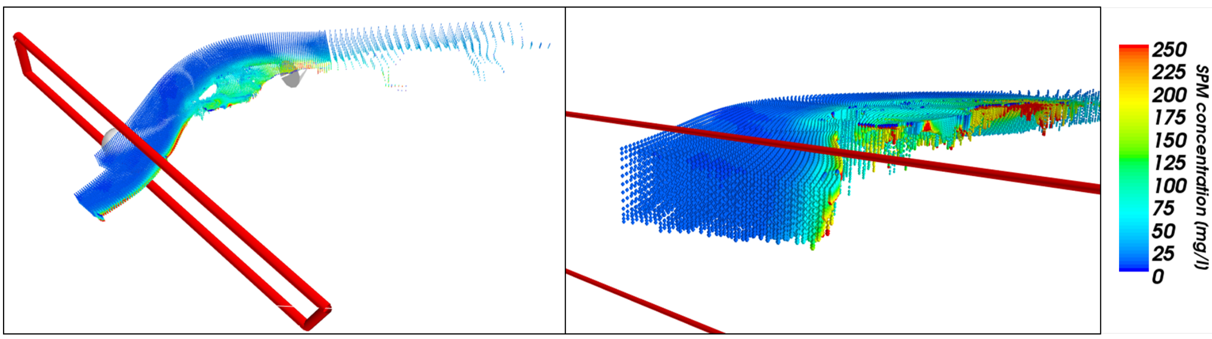


Figure 7.5: Left: Placing the cutplane over the visualization by the user. Right: Close-up of the cross section made.

7.2.3 Marker selection for time series

Another aspect is the marker selection in the 3D environment. When clicking on one of the locations in the domain, information on that location is displayed in the window in the bottom pane. This clicking shows a time series of that location, together with the mean and the uncertainty through time, displayed in a 2D graph. An example is shown in Figure 7.6, in which the selection of the marker and the time series is shown. In the quantification of the uncertainty just a value for the average and the uncertainty are saved to the input file for the visualization. Therefore the time series is a graph based on these values. It is not feasible to store all the time series of the different model runs, the toolbox would get very slow if all these data needs to be processed every time a selection is made. Therefore just visualizing the concentrations with the uncertainty in the time series is chosen.

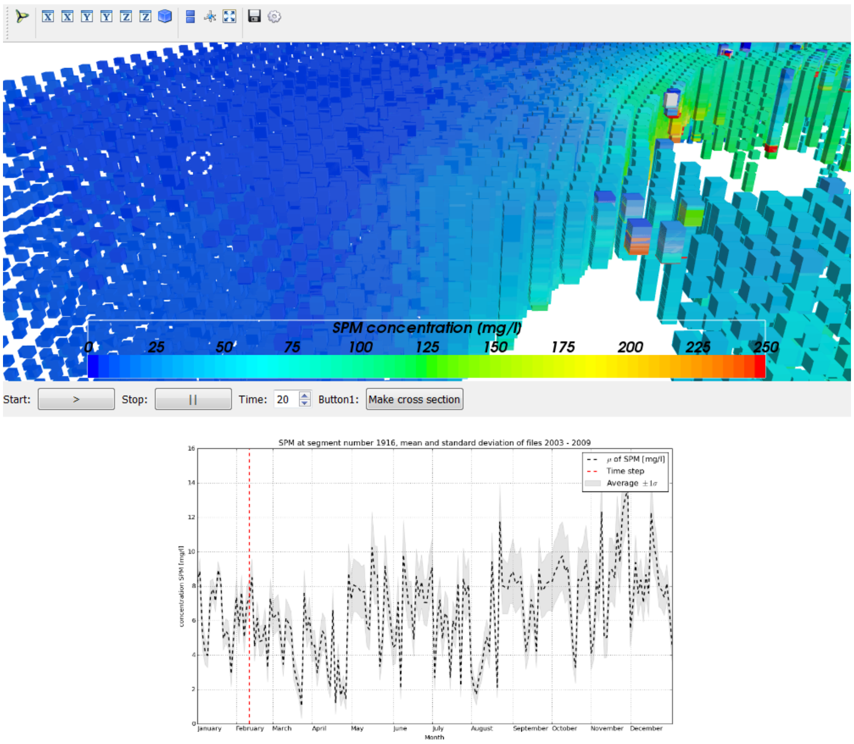


Figure 7.6: The selection of a marker in the 3D environment in the top part of the figure and the time series belonging to that marker in the bottom part. At time step 20 (one time step is 2 days), in the time series indicated by the blue vertical line.

7.2.4 Time implementation

An animation function is implemented to view the sequential time steps. In that way the visualization can show the movement of SPM in the water column. With this time lapse the user can view the information on SPM movement and see the connection with the surrounding cells. Normally a 2D visualization is made for model output, such as the plan view shown in Figure 7.1 in the Quickplot tool of the Delft3D GUI (Graphical User Interface) it is also possible to make a time lapse of the 2D plot. However in that way the connection with different layers is lost. Shortly, previous methods were limited in its representation of a 3D environment in a 2D representation, which is incorporated in this method.

Every update of the time step changes the color and opacity according to that time step. In the time series in the bottom pane a blue vertical line is used to also indicate the time step in this plot. There is a start button to start the animation, where the time step increases 1 at the time. Consequently with the stop button the animation can be stopped at all times. In a text box the time step is tracked. A time step can be displayed manually by inserting a time step of choice in this textbox. This time animation is visualized in Figure 7.7, where four timesteps are displayed, changing the markers accordingly.

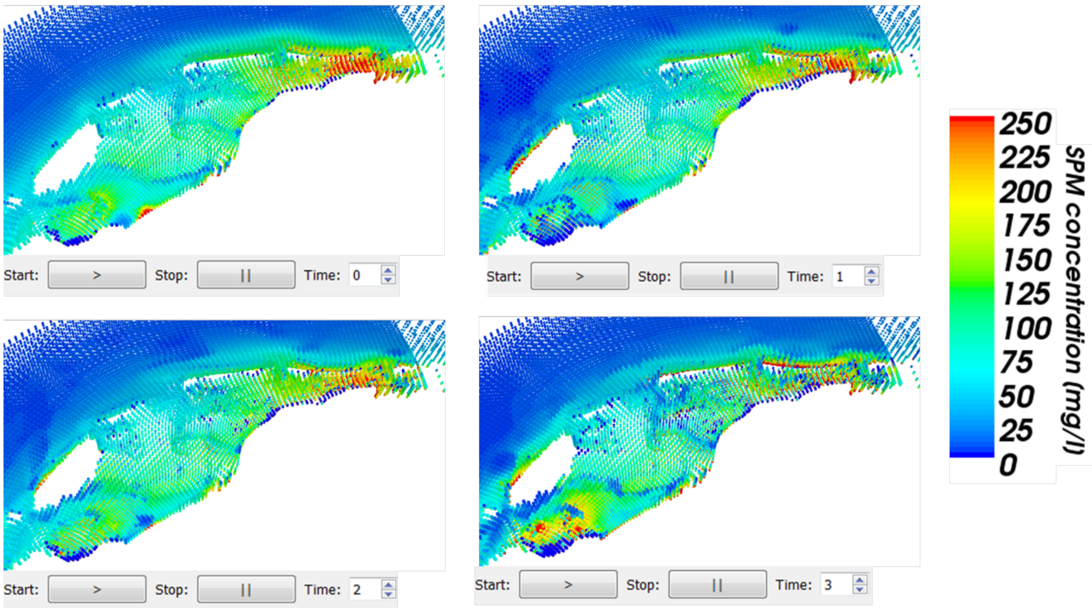


Figure 7.7: Part of the animation of the toolbox, four follow-up time steps of the 3D environment.

7.2.5 Other features

In the toolbox it is possible to change the observation angle in multiple directions. It is possible to pan and zoom through the interface. Furthermore it is possible to translate, rotate, enlarge and move the created objects (such as the markers, which are stored in one object, or the cutplane). This makes the visualization very adjustable and easy to interact with. The area of interest can be looked from different angles which makes it possible to get much more information from this visualization than from the normal 2D visualization. Another feature of the toolbox is the colorbar, representing the amount of SPM per cell in $[mg/l]$. It should be noted that this colorbar is set to have a range of [0, 250], and as a consequence it does not scale to the values in time.

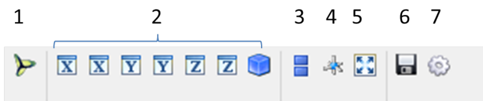


Figure 7.8: Built-in toolbar of the mayavi scene.

The Mayavi scene in which the 3D environment is created has an built-in tool bar, where more functions are available to use for the toolbox. This tool bar is shown in Figure 7.8, where the numbers are an indication for the buttons. Button 1 is an option to open the more advanced Mayavi pipeline, in which the layout of the interface can be adjusted. More objects, filters and other options are available from there. For creating and editing the toolbox this is a useful tool, however this is not the case for users of the toolbox. The buttons indicated by 2 are the ones to set the camera to a certain view along a certain axis. Button 3 toggles the parallel projection. Button number 4 displays the axes on the environment. Button 5 toggles the window to full screen and reverse this when clicked again. With button 6 a still can be made and saved of the view and number 7 opens an extra setting menu.

7.3 Discussion

It is possible to create an interface where both model output in space and time can be combined together with uncertainty, using color for displaying the concentrations and a white hue to visualize the uncertainty. In this toolbox the entire 3D domain is visualized in one interface and using an animation the fourth dimension of time is also taken into account. Different tools are needed to help make the toolbox more intuitive and easy to use, such as the cutplane for more in depth information on a certain location and the selection of a marker to display time series per location.

8 Physical background

Some physical background information on the behavior of SPM concentrations is needed to interpret the toolbox. This is the main focus of this chapter to discuss the physical interpretation and to discuss how this toolbox can be interpreted, using the study case of SPM in the Wadden Sea area. The interpretation of the changing in the SPM concentrations in space and time can only be explained by the physical background. The equations solved in the SPM model that give the outcome of the magnitude of the different fractions of inorganic matter in the domain are dependent on many different variables, such as the bathymetry, waves, flows and discharges.

8.0.1 General behavior of SPM

SPM are the small solid particles in the water column that remain in suspension [El Serafy et al., 2007]. A distinction can be made between fine sediments and coarse sediments, which correspond to the particle fractions within the model. They are stirred up from the bottom by waves, which creates shear stresses onto the sea bed. Oscillatory motion and turbulence created near the bed bring the particles into the water column and oppose the settling velocity, resulting in the particles to remain in suspension [Bosboom and Stive, 2015]. Due to advection the particles are transported with currents. When the particles are lifted from the sea bed into the water column it is called resuspension, while the sedimentation occurs when the particles settle back to the seabed from the water column. When the water movement is slow enough the gravity forces (settling velocity) will take over and give the matter chance to sink back to the bottom of the water column. The top layer of the seabed is very dynamic and sedimentation and resuspension occur often, creating an interaction between water column and seabed. Higher waves cause stronger forces on the bed, resulting to cause more erosion, while bioturbation (the reworking of the particles by animals and plants on the bed) let the particles stay permanently at a location [Gayer, 2009]. With this basic behavior of the SPM the general behavior in the research domain can be explained.

In Figure 5.9 the processes of erosion (E) and deposition (D) are shown as they are integrated in the numerical model. These processes are simulated with the equations 5.2 and 5.6.

8.0.2 Location specific behavior of SPM

From the previous section it became clear that all the transport mechanisms of SPM are induced by water movement. When looking in the water column the concentrations of SPM will be highest near the bed, because of gravity and thus the settling of the particles. Due to the water movement towards and along the coast, the influence of salinity gradients (stratification) and the tide, the SPM concentrations increase drastically, coming closer to the coast [Arentz et al., 2012]. Looking at the toolbox, this behavior is clearly visible, where offshore the concentrations are very low (blue) while the two areas in the circles have very high values, these are the Dutch coast and the Wadden Sea area. In Figure 8.1 a still of the toolbox is shown with SPM concentrations. It can be noted that further offshore the amount of SPM is low, and going almost to zero at most locations. As mentioned is SPM transported by processes in the water column. Further offshore the waves do not affect the bed significantly and the particles are not suspended into the water column. The behavior in the two areas are explained shortly below, combining the behavior of SPM to the specific location.

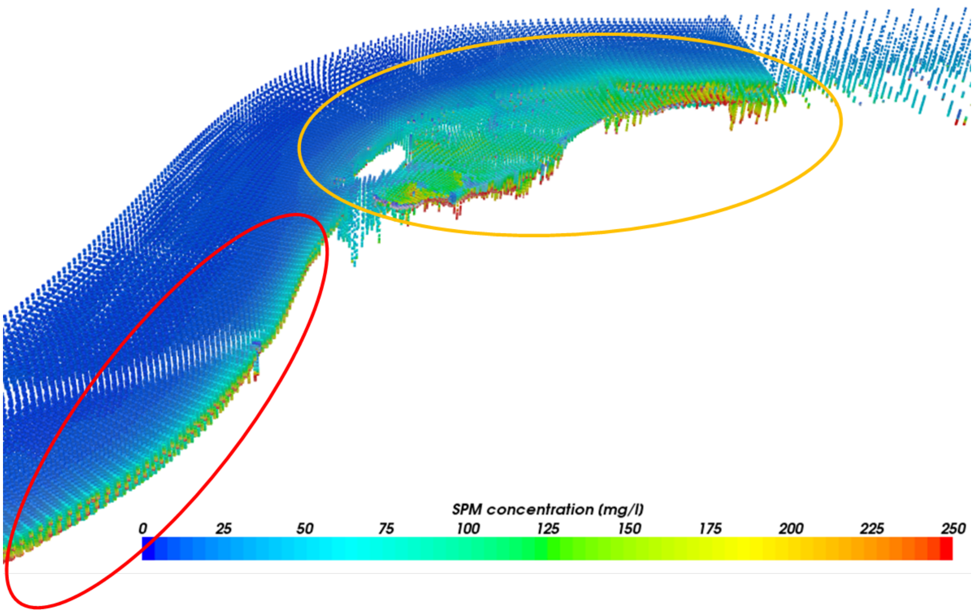


Figure 8.1: Highlighted are the areas with high SPM concentrations. The area in the red circle is the dutch coast. The area in the orange circle is the Wadden Sea area (Dutch and German part).

The Dutch coast is the area within the red circle, see Figure 8.1. Due to an alongshore current near the Dutch coast keeps the SPM in suspension, and transports it along the coast. However a source is needed to keep enough SPM in the system. The main source of SPM comes from the Rhine ROFI (Region Of Fresh Water Influence). SPM can also come into the water column due to erosion of the shoreline and transport mechanisms created by waves or winds. SPM disappears from this part of the system towards the Wadden Sea, when transported with the longshore current northwards. Furthermore, SPM can disappears from the water column and be deposited at the shore or at other coastal structures, or when it is transported with an undertow offshore. Sediments are furthermore extracted to be used in nourishments elsewhere, and therefore human activities could also be a sink for the SPM [Bosboom and Stive, 2015]. From the plume of suspended matter created by the Rhine ROFI into the North Sea, the SPM is transported along the Dutch coast by a long shore current, which is created due to the Coriolis effect. This effect deflects the outflow in a anti-clock wise direction (in the Northern hemisphere), resulting in the longshore current along the Dutch coast in the Northwards direction [van der Hout et al., 2015]. This creates high concentrations of SPM near the coast. When looking at the vertical water profile, SPM concentration tend to be higher in the lower parts, near the bed, due to gravity . The movements stirring up the sediments are greatly influenced by the stratification of the fresh river water from the Rhine and the ocean water. A pycnocline is established when the more dense river water flows underneath the less dense ocean water. Due to this stratification the turbulent motion near the bed is trapped underneath this pycnocline and therefore the suspended matter is shut down to the lower regions [Pietrzak et al., 2011].

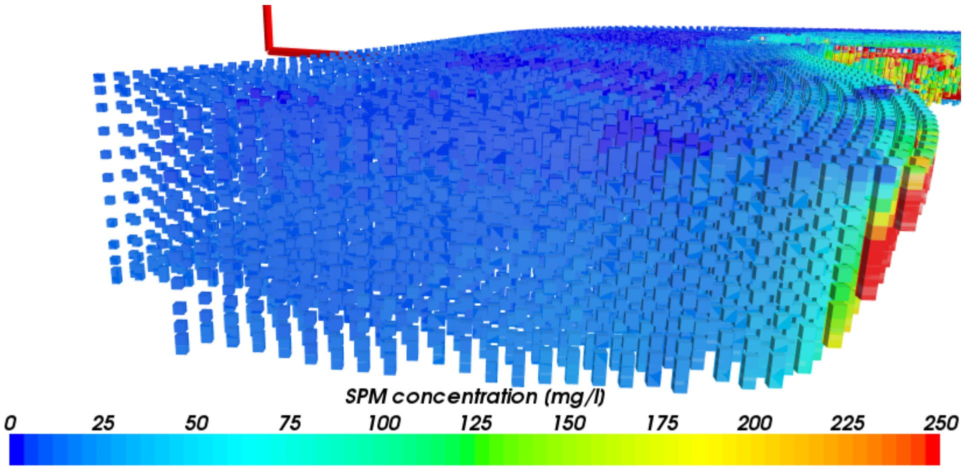


Figure 8.2: A cross section near the Dutch coast, to the left (offshore) very low concentrations of SPM are visible, near the shore the concentrations are high, in that part a very clear distinction is seen in the surface layers (low concentrations) and the bottom layers (high concentrations).

The hydrodynamic forces are not constant through time, but change in direction and magnitude, influencing the behavior of the SPM through time. For example the spring and neap tide cycle changes the influence of the SPM plume from the river on the long shore current. The extend to which the underlying fresh water flow is present is larger in the neap tide cycle, than in the spring tide. In this toolbox it is difficult to see this influence, because of the time steps chosen (1 day) however when looking at the figure of the locations of uncertainty, it can be seen that a large plume near the outflow of the river has a larger uncertainty, indicating that the processes occurring there are more difficult to calculate. From this it can be concluded that the processes of the spring and neap tide are taken into account in the model, however other data sets are needed to investigate these and other processes further.

The Wadden Sea area is a large tidal flat in the northern parts of the Netherlands. The Wadden islands create multiple tidal inlets in the system. This area is highly influenced by the tide, during high tide the water floods the area and during ebb the water flows out again. These constant flows transport the sediments around in the Wadden Sea and in the inlets. Therefore, the SPM in the water column is constantly relatively high (from Figure 8.1, the entire area has a concentration of 75 mg/l or higher, also around the outsides of the Wadden islands). In the time between ebb and flood, called slack tide, the water is more or less stationary, giving the sediments suspended time to settle. Sources of SPM are the SPM coming from the long shore current along the Dutch Coast, and the river outflows near the Wadden Sea (see locations of rivers in Figure 5.6). The in- and outflowing currents create a distinct bathymetry in the area, creating gullies with strong currents. In Figure 8.3 the area is shown on January 1st 2009. In this figure it can be seen that the gullies indeed visible and have a low concentration of SPM, due to the strong current.

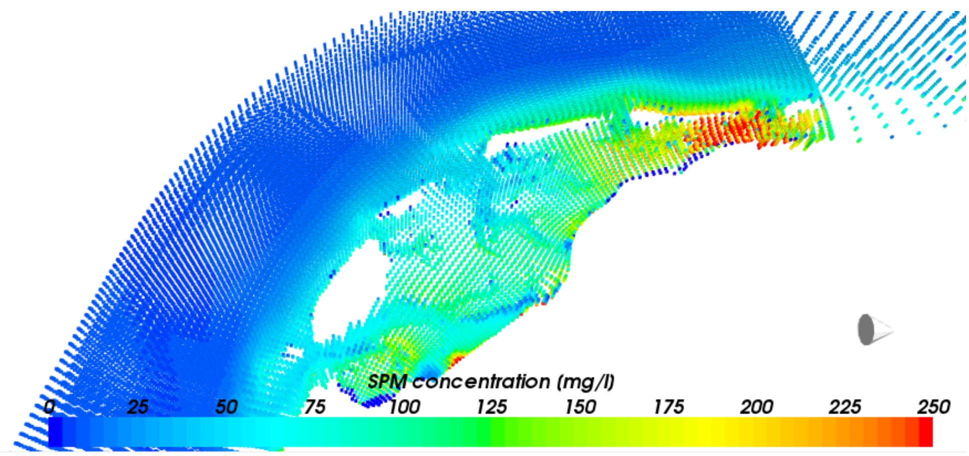


Figure 8.3: Wadden Sea area on January 1st 2009.

8.1 Summary

In this chapter the basic processes of SPM are discussed, in order to validate if the processes known to happen are captured by the model. From this it can be seen that the expected processes are indeed visible. Processes that are time dependent, such as spring and neap tide differences are harder to differentiate and could be interesting as a topic for further studies. The regions where the model create the most uncertain values can be related back to the physical background. Areas with low concentrations (offshore) have a very low uncertainty, where areas near the coast have a much higher uncertainty. Especially near the outflow of a river or on the tidal flats of the Wadden Sea. The relation between the uncertainty and these critical areas, where the SPM concentrations are influenced by many different factors can be related back to possible sources of inaccuracy. First of all it could be possible that the equations 5.2 and 5.6 for the deposition and erosion fluxes, describing the bed load module [Van Kessel et al., 2011] do not completely incorporate all processes. Secondly it is possible that the main driving force, the hydrodynamics are not completely correct. Where SWAN is used to recalculate the wave field for the entire domain, using data from buoys, it can very well be that due to measurement uncertainty or gaps, not all processes are incorporated. Where the erosion of the Dutch Coast is heavily influenced by the storm surges, and wind driven waves, it can not be guaranteed that these processes are incorporated when using non-continuous data from wave buoys.

9 Discussion

In this study it has been shown that by quantifying the output and uncertainty from a numerical model, this information can be combined into a generic toolbox to be easily accessed for management purposes concerning ecosystems. Model output can give information on specific indicators for an area of interest, such as ecological valuable areas. Where the uncertainty gives extra information on how to interpret these result and to what extent a numerical model is useful.

The Wadden Sea is used as a case study in this research to develop this generic toolbox and to assess the uncertainty of a numerical model describing the turbidity in the area. The turbidity is modeled in the concentrations of Suspended Particulate Matter (SPM), using an existing Delft3D-WAQ model [El Serafy et al., 2013]. The SPM blocks the underwater light and thus influences the algal bloom, which is one of the determining factors for the water quality in the area [Marencic, 2009]. A GEM/BLOOM model exists for calculating the indicator for algal bloom, chlorophyll-a concentrations, which has been calibrated and validated (Los [2008] and Blauw et al. [2009]). The SPM model is already calibrated [El Serafy et al., 2013], however in this study the model setup has changed and a numerical model is introduced for the input of SPM. To assess what the new uncertainties are that this model brings into the system, a sensitivity analysis showed that the initial file (file including the initial conditions of the substances in the model for each segment), the waste loads (SPM discharges from the rivers for each time step) and the bottom shear stresses induced by ships did not have a significant influence on the SPM values. From another sensitivity analysis [El Serafy et al., 2013], it was found that the parameters did have a major influence. In total there are 71 parameters present within the model, however from this study 10 parameters were identified to have the most influence. These 10 parameters are present in the erosion and deposition fluxes describing the behavior of SPM. It was chosen to use these parameters in an analysis to quantify the uncertainty in the model.The SPM is calculated with equations for the erosion and deposition fluxes from the bed load module, as described by [Van Kessel et al., 2011]. The parameters in these equations introduce uncertainty within the model.

The uncertainty analysis is done using a Latin Hypercube Sampling with Dependence [Meszaros, 2016], to include the dependencies between parameters and to reduce the amount of model runs needed. It should be kept in mind that in this way the amount of runs is reduced drastically and that the outcome will be an estimation of the reality. From this analysis the results were validated and conclusions could be drawn. The model gives a reasonable estimation for the SPM concentrations in most locations within the domain. However in critical areas where the uncertainty is very high, such as outflows of rivers and parts were the SPM processes are very dynamic. For the quantification of the outcome of the uncertainty analysis a log-normal distribution is used to estimate the 188 outputs. The characteristics of this distribution, the mode and the spreading is used to quantify the concentrations of SPM and the uncertainty into a value that can be used for the toolbox. The SPM values with this range of uncertainty are validated against measurements from Rijkswaterstaat, which is in all cases in the same order of magnitude and mostly follows the pattern throughout the year. Another validation is against the previously used IM1 data for the GEM/BLOOM model input [Blauw et al., 2009]. The same pattern throughout the year is shown, however the SPM overestimates these values. This is explained to an extent by the fact that IM1 is only a fraction of the SPM. Though it is a major fraction of the SPM and therefore it is also an indication that the SPM concentrations are plausible and realistic. A second validation is done with a comparison of SPM values obtained from satellite images from the MERIS missions. Conclusions that can be drawn on this validation is that in general the order of magnitude in all locations are correct. Most data points lie within the confidence interval, indicating that the estimate of the SPM is correct. It should be noted that the model output has a time step of one day (starting from 01-01-09 at 00:00). Due to this the data points from either Rijkswaterstaat or MERIS will never be in line with the model data and therefore a really accurate comparison can not be done. The influence of factors changing on a smaller scale than diurnally are not completely taken into account. A last validation was done using previously. The regions of high uncertainty can be identified using a top view projection per layer. From this map it is observed that the highest uncertainty occurs in regions of highly dynamic and influenced by many factors, such as the outflow from a river or the tidal flats of the Wadden Sea.

Data for concentrations and for uncertainty need to be merged together in the toolbox. This means that there are two distinct characteristics needed to cope with both data sets. The method used to visualize uncertainty in this study is using the color values for the concentrations and white hue values to display uncertainty. From the literature study it was found that there are more ways to display the uncertainty, such as opacity, shape, size, texture, orientation or a blur [Baart, 2013]. By using the programming language Python and in this study markers to visualize the concentrations in the water column, some restrictions are encountered. A blur can not be coped with in the Mayavi objects. Furthermore, things as shape or texture are discrete, and not every distinct value of uncertainty can be shown in the toolbox, due to the limitations of the amount of shapes and textures. The orientation could show the continuously distributed uncertainties, however this gave a very chaotic effect on the visualization. The size of the markers was another option, however the markers are not evenly spread over the area, due to the different grids, and in the z-directions there is already overlap with the next marker. When changing the size the visualization got really messy and could not provide a clear overview of the uncertainty. The opacity was tested with discrete values, which made it clear that using this option that if the markers got too opaque, the markers behind it would be seen through them and the real value of the markers got distorted. However this method was by far the most useful, but the restrictions in Python for the markers do not allow to give the markers separate values for both color and for opacity, because these two parameters are linked to each other in a Look-Up Table (LUT) where Red, Green, Blue, Alpha (RGBA) values are defined in a specif range. Therefore the marker can only cope with one scalar. A solution was to make a new LUT, where for every marker the concentration and according uncertainty are saved for a value for color and opacity. However this was so devious and made the toolbox very slow. Therefore a similar solution was found using color for the concentrations and the opacity for uncertainty. The original markers are given a color for their concentrations, without opacity. Using a second marker, slightly larger than the original marker and completely white, and giving an opacity value accordingly to the uncertainty. In that way the two problems are solved, the devious solution is bypassed and furthermore the marker itself is not opaque, blocking the colors of the marker behind it.

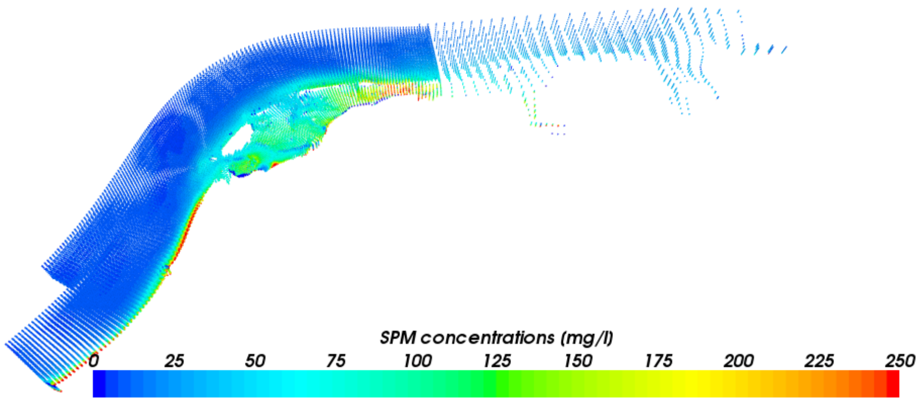


Figure 9.1: Still from the 3D environment in the toolbox, visualizing SPM concentrations [mg/l] and uncertainty.

Using the methods described a toolbox is developed in which the model output and uncertainty is visualized together. The turbidity and its uncertainty are visualized in one interface, see Figure 9.1. When combining the knowledge gained from the uncertainty analysis with the information on the SPM values different results are visible. Expected hydrodynamic processes influencing the SPM concentrations are visible in the toolbox. Processes that are time dependent, such as spring and neap tide differences are harder to differentiate and could be interesting as a topic for further studies. The regions where the model create the most uncertain values can be related back to the physical background. Areas with low concentrations (offshore) have a very low uncertainty, where areas near the coast have a much higher uncertainty. Especially near the outflow of a river or on the tidal flats of the Wadden Sea. The relation between the uncertainty and these critical areas, where the SPM concentrations are influenced by many different factors can be related back to possible sources of inaccuracy. First of all it could be possible that the equation describing the bed load module [Van Kessel et al., 2011] do not completely incorporate all processes. Secondly it is possible that the main driving force, the hydrodynamics are not completely correct. Where SWAN is used to recalculate the wave field for the entire domain, using data from buoys, it can very well be that due to measurement uncertainty or gaps, not all processes are incorporated. Where the erosion of the Dutch Coast is heavily influenced by the storm surges, and wind driven waves, it can not be guaranteed that these processes are incorporated when using non-continuous data from wave buoys.

10 Conclusions and recommendations

10.1 Conclusions

In this section the conclusions of this research are summarized and the main question will be answered. This will be done by summarizing the answers to the subquestions which were found throughout the chapters and connecting them with the discussion points described in the previous chapter.

10.1.1 Subquestions

What information is useful to decision makers?
A combination of two aspects are important and useful to decision makers and for creating the toolbox. Firstly the information on what a decision maker wants from such a toolbox is needed. This is the information on the water quality in the Wadden Sea through time. Secondly, the information on the problems within an PA is important. The water quality (regarding eutrophication) in the Wadden Sea is described by the indicator of chlorophyll-a concentrations in the water column. One of the main factors influencing the chlorophyll-a is the Suspended Particulate Matter (SPM), which influences the light intrusion in the water column and therefore reduces eutrophication. Chlorophyll-a and SPM are therefore important concentrations, which are calculated using numerical models. For this study the focus lies on SPM, because the in the model structure used to calculate the water quality th

What are the important uncertainties within the regarded model structure and how can they be identified?
The chlorophyll-a concentrations are calculated with a Delft3D-WAQ GEM/BLOOM model, where the SPM input is calculated with an Delft3D-WAQ SPM model and the hydrodynamical input with a Delft3D-FLOW model. These models are built and calibrated in previous studies, therefore many sources of uncertainty are already addressed. In previous studies the SPM input came from satellite images, in this study it is investigated to see if this SPM model can be used for the input of the GEM/BLOOM model. Therefore the uncertainty of this SPM model is yet to be addressed. Throughout the entire model structure many different uncertainty sources can be found, coming for example from: model input and inherent uncertainties. A selection is made to narrow the uncertainty sources to the ones that have yet to be assessed. The used SPM model has been the topic of many different studies, and is already calibrated to reduce inherent uncertainty, thus the remaining uncertainty sources are the ones originating from the input. This includes: hydrodynamical input, the parameters in the constants file, waste loads (substances in the river discharges), shear stresses induced by ships and the initial map file (initial value of the substances).

How can the uncertainties be quantified into values that can be used in the toolbox?
To do an uncertainty analysis identified sources need to be further reduced to only cope with one source, this is done using a sensitivity analysis on the identified input files. Not all files can be coped with, otherwise the amount of computational time needed would be too much for only one project. The hydrodynamical input was not taken into account, because the hydrodynamic model has a computational time of a few days. The sensitivity analysis is used to investigate the influence of the Waste loads, the stresses induced by ships and the initial map file. Neither the Waste loads nor the shear stresses induced by ships introduced many uncertainties. Even if the stresses induced by ships did have a great influence, it would not be realistic to use this parameter for the uncertainty analysis, due to its correlation with the hydrodynamics. The initial map file introduced a large variability in the output. To cope with this variability a stable SPM field is needed. By rerunning the model using the restart file of the model output as the initial file. The remaining file is the constants file, on which a sensitivity analysis was already

done in a previous study [El Serafy et al., 2013], in which the separate parameters were assessed. From this study 10 most influential parameters were found: TauShields, FactResPup, VSedIM1, FrIM1Sed2, VResIM1, TaucRS1IM1, VSedIM2, FrIM2Sed2, VSedIM1 and FrIM1Sed2.

The uncertainty analysis on these parameters is performed using a Latin Hypercube Sampling with Dependence. This method includes the correlation between the parameters and reduces the amount of model runs needed for an accurate uncertainty assessment. Where a random sampling would need thousands of runs, the LHSD eventually 188 runs are done for calculating the SPM concentrations in the entire domain for an entire year. The results of these concentrations resembled a log normal distribution. From this distribution the location parameter (the mode, the value that is represented most often) is used to indicate the estimate of the SPM concentrations and the scale parameter is used as an indication for the certainty of this concentration. The output is validated with in-situ measurement data from Rijkswaterstaat, SPM data used as input for the GEM/BLOOM model in previous studies [Blauw et al., 2009] and SPM obtained from MERIS satellite images [Eleveld et al., 2007]. From this validation it could be observed that the data resembles the measurements in most cases and were always in the same order of magnitude. The input file for the GEM/BLOOM model consisted of only the IM1 fraction of the particulate matter and therefore is an underestimation for the SPM, however from this file it could also be observed that the order of magnitude was in all cases the same. Concluded is that the SPM model gives a good estimation and can be used as an input for the GEM/BLOOM model in future studies.

How can the model output together with the quantified uncertainties be visualized in a toolbox?
To answer this question, two variables, model output and its the uncertainty need to be displayed together in one toolbox. At every location in time a grid cell changes value for concentration and uncertainty. Chosen was to use a marker per grid cell to visualize the data for every grid cell, which enables it to the incorporation of the 3D character of the model domain. The characteristics for these markers can be used to display the two variables, concentration and uncertainty. In the developed toolbox color is used to display the concentrations and a white hue indicates the uncertainty of this value. In that way all the desired values can be displayed in one interface, at every location in the 3D grid. Using an animation the values can change through time and therefore this fourth dimension is also coped with.

How can the toolbox be used?
The toolbox is a tool that is easy to use with certain background on technical requirements and physical processes. With some understanding of the basic processes in which the deposition and erosion of SPM works the toolbox can be interpreted. Depending on the data used in the toolbox, the physical interpretation changes. In Appendix B B, the User manual more details on the technical background is explained and how the toolbox can be opened, and used.

10.1.2 Main question

How can uncertainty from a SPM model as a driving force for a GEM/BLOOM model be identified, quantified and visualized to help decision makers?
The identification of uncertainties comes from literature studies on calibration reports for the regarded model. The quantification is done by using an uncertainty analysis, in which dependencies are integrated and a sampling technique is used to reduce the amount of model runs needed. This analysis results in an ensemble of multiple values for SPM per location per time step. With a log normal distribution function this output can be estimated. The mode of this function results in a value for SPM concentrations and its standard deviation for uncertainty. The visualization is done by developing an interface in which per segment an object, such as a marker, is placed. The characteristics of this marker can be used to display the different values, color for the SPM value and a white hue for the uncertainty as found in the quantification. The validation of the SPM output shows that this SPM values van further be used as an input for the GEM/BLOOM model.

10.2 Recommendations

This research combines an uncertainty analysis (on a SPM model) and the development of a toolbox. However a method is found and conducted, some aspects of this method could be a good base for further research. These points are addressed in this section.

In this study a generic toolbox is developed that can be used for different data, dependent on the area of interest of a project. For example, when more detailed information of the SPM processes are required. The SPM concentrations are dependent on the ebb and tide cyclus, therefore it would be very interesting to investigate the model on a smaller time step. Now the spring en neap tide are neatly captured within the model output, because this simulation was for an entire year it was not feasible to use a smaller time step, because of computational intensity and storage problems. However when using a smaller time step, one can also reduce the total model time and only assess a part of the year. Another study case could be to use the work done in Meszaros [2016], where a ensemble forecast was done on the GEM/BLOOM model. Calculating the chlorophyll-a concentrations in the Wadden Sea area. Using this data could help the decision makers for the protected area to make knowledge-based decisions.

For further optimizing the SPM calculations in the numerical model, the sources contributing to the extreme uncertainties can be further investigated. From this study it was concluded that either the bed load module or the hydrodynamic input are not yet optimized in critical areas. And further studies could show what is missing in these regions.

In some cases it might be more interesting to visualize the different layers of the model in the z-direction, creating top views. This is done using the cutplane, but enabling the cut in the z-direction. Using this method the visualization creates a 2D top view of the model. This view can also be used to not only display a risk map of the area, where colors are used as an index of the risk in the area for example three colors (green = low risk , yellow = neutral, red = high risk). This risk can be linked to the concentrations of SPM or chlorophyll-a, where a certain threshold is exceeded. This might be a nice extension of the toolbox.

Certainly the output gives a good representation of the spreading of the SPM in the area, however decision makers need to make decisions for future scenarios based on these results. Therefore it could be interested to use this toolbox to visualize information from a forecasting study. Or use the model set up to model a certain human intervention, such as an extra discharge with a large amount of SPM into the water, simulating a deposition of a company into the area. This could help in the process of decision making for specific projects.

Instead of using markers for every segment, another option is to use a grid, where the nodes per cell are connected to each other. In that way every cell will have its own shape, but the total grid would be continuously divided, creating a more intuitive field. However this would require a lot more calculation for finding the nodes and connecting these into a grid for the visualization. This combined with the fact that it was not proven that this representation would improve the toolbox were reasons not to test it in this thesis.

List of Concepts

Concept	Description
Arghyd	Program in Delft-3D for the aggregation of a grid.
CDF	Cumulative Density Function, a summation of the likelihood of a certain value.
Copula	A multivariate distribution function with uniform marginals [Schmidt, 2006], which can incorporate the correlation between the parameters described in the multivariate distribution.
Delft-3D FLOW	3D/2D modeling suite for integral water solutions. Simulation of multi-dimensional hydrodynamic flows and transport phenomena, including sediments [Deltares, 2014c].
Delft-3D WAQ	Water quality and aquatic ecology modeling suite [Deltares, 2014a].
Delft-3D WAVE	See SWAN.
DDcouple	Program in Delft-3D for the domain decomposition of different grids.
GEM/BLOOM	The Generic Ecological Model (GEM) [Deltares, 2014a] and the phytoplankton module (BLOOM) of the Delft-3D WAQ model [Blauw et al., 2009].
IM	Inorganic Matter, which is denoted with a fraction number (IM1, IM2 or IM3) describing a fraction of the Total Inorganic Matter, namely the coarse (IM1, diameter $40\mu m$), medium (IM2, diameter $15\mu m$) and the fine (IM3, diameter $1\mu m$)sediments (in $[g]$).
LHS	Latin Hypercube Sampling, stratification of a CDF to a nearly random sampling of a input parameter.
LHSD	Latin Hypercube Sampling with Dependence [Meszaros, 2016], name of the method used to include the LHS into the Monte Carlo simulation and the dependencies between the input parameters.
Meris	MEdium-spectral Resolution, Imaging Spectrometer, satellite mission used for obtaining SPM values.
Monte Carlo	Method to do an uncertainty analysis, simulating physical possible input multiple times.
PDF	Probability Density Function, a function describing the likelihood of a value describing a parameter.
PPF	Percentile Point Function, is the inverse of a CDF.
Sentinel	Satellite missions of ESA (Europian Space Agency) used to obtain for example SPM data.
Segment	Computational volume in the Delft3D models.
SPM	Suspended Particalute Matter is the total of suspended particles, organic and inorganic in the water column, causing the turbidity.
SWAN	Simulating WAves Nearshore, which is a third generation wave model in the Delft3D-WAVE environment [SWAN team, 2016].
TIM	Total Inorganic Matter present in the water column.
ZUNO-DD	ZUdelijke NOordzee with Domain Decomposition, indicating the total grid, including the coarse, intermediate and fine grid for the Southern Northsea, as is used in the numerical models.

List of Figures

List of Tables

List of Symbols

Symbol	Units	Description
α_{IM_i}	-	Proportion of the deposited silt that is stored directly in the sandy layer (S_2)
C	g/m^3	Concentrations
C_{IM_i}	mg/l	Concentration of fraction Inorganic Matter (IM_i) of the fraction class i
D_{j,IM_i}	g/m^2d	Deposition flux of SPM fraction IM_i from layer S_j
D_x	m^2/s	x-Component of the dispersion tensor
D_y	m^2/s	y-Component of the dispersion tensor
D_z	m^2/s	z-Component of the dispersion tensor
E_{j,IM_i}	g/m^2d	Resuspension flux of SPM fraction IM_i from layer S_j
μ	-	Mean
F_{ResPup}	-	Van Rijn pickup factor from buffer layer
$M_{i,j}$	$[g/m^2]$	Mass of sediment fraction i in layer j per surface area
N	-	Numbers of samples
P	-	Sources and sinks of mass due to processes
S	-	Sources and sinks of mass due to loads and boundaries
σ	-	Standard deviation
t	s	Time
τ	Pa	Bottom shear stress
τ_{cr,S_1,IM_i}	Pa	Critical shear stress for silt resuspension fraction i from fluff layer (S_1)
τ_{Sh}	Pa	Critical Shields stress for sand mobilization in buffer layer S_2
x	-	Random sample
x	m	x-coordinate
X	-	Set of x
u	m/s	Components of the velocity vector in x-direction
v	m/s	Components of the velocity vector in y-direction
V_{Res,IM_i}	$1/d$	First order resuspension rate from layer S_1
V_{Sed,IM_i}	m/s	Settling velocity of the fraction class i
w	m/s	Components of the velocity vector in z-direction
y	m	y-coordinate
z	m	z-coordinate
Z_{Res,IM_i}	g/m^2d	Zero-order resuspension rate from layer S_1

Bibliography

L. Arentz, V. Harezlak, T. V. Kessel, and T. V. D. Kaaij. Kalibratie slibtransport- en GEM-model, 2012.

F. Baart. *Confidence in coastal forecasting (PhD thesis).* Doctoral dissertation, Technical University Delft, 2013.

H. Baumert, G. Chapalain, and H. Smaoui. Modelling and numerical simulation of turbulence , waves and suspended sediments for pre-operational use in coastal seas. pages 2–12, 2000.

M. Blaas, K. Cronin, G. Y. El Serafy, Y. F. Friocourt, S. G. Aguilar, and G. H. Keetels. MoS2 : Model setup, data assimilation and skill assessment. 2012.

A. N. Blauw, H. F. J. Los, M. Bokhorst, and P. L. A. Erftemeijer. GEM: A generic ecological model for estuaries and coastal waters. *Hydrobiologia*, 2009. ISSN 00188158. doi: 10.1007/s10750-008-9575-x.

J. Bosboom and M. Stive. Coastal Dynamics 1. Lecture notes CIE4305., 2015.

Common Wadden Sea Secretariat. Wadden Sea Plan 2010. Eleventh Trilateral Governmental Conference on the Protection of the Wadden Sea., 2010.

K. Cronin. Memo - DELWAQ-SPM_Noordzee-Zuno-DD_j03-11_v01 metadata Waterkwaliteitsmodelschematisatie. Technical report, Deltares, 2014.

J. C. Davis and R. J. Sampson. *Statistics and data analysis in geology*, volume 646. Wiley New York et al., 1986.

J. M. Delhez, P. Damm, E. D. Goede, J. M. D. Kok, F. Dumas, and H. Gerritsen. Variability of shelf-seas hydrodynamic models : lessons from the NOMADS2 Project. 45:39–53, 2004. doi: 10.1016/j.jmarsys.2003.09.003.

Deltares. Concentration of suspended matter in water, 2013. URL `http://opendap.deltares.nl/thredds/catalog/opendap/rijkswaterstaat/waterbase/concentration{_}of{_}suspended{_}matter{_}in{_}water/catalog.html`.

Deltares. D-Water Quality User Manual. Versatile water quality modelling in 1D, 2D or 3D systems including physical, (bio)chemical and biological processes, 2014a.

Deltares. Delft3D-Wave user manual. Simulation of short-crested waves with SWAN, 2014b.

Deltares. Delft3D-FLOW, User Manual. Simulation of multi-dimensional hydrodynamic flows and transport phenomena, including sediments, 2014c.

Deltares. Delft3D Functional specification (User Manual), 2014d.

Deltares. D-WAQ Input File. Documentation of the input file (User Manual), 2014e.

Ecopotential. ECOPOTENTIAL: improving future ecosystem benefits through earth observations, 2015a. URL `http://www.ecopotential-project.eu/`.

Ecopotential. Wadden Sea and Dutch Delta, 2015b. URL `http://ecopotential-project.eu/protected-areas/25-the-wadden-sea-dutch-delta-area`.

Ecostress. ECOSTRESS: Ecological Coastal Strategies and Tools for Resilient European Societies, 2013. URL `http://www.ecostress.eu/`.

G. Y. El Serafy. Introduction to Earth Observation Services. 2013.

G. Y. El Serafy, M. Blaas, M. A. Eleveld, and H. J. V. D. Woerd. Data Assimilation of Satellite Data of Suspended Particulate Matter in Delft3D-Waq for the North Sea. (October 2015):1–8, 2007.

G. Y. El Serafy, M. a. Eleveld, M. Blaas, T. Kessel, S. G. Aguilar, and H. J. Woerd. Improving the description of the suspended particulate matter concentrations in the southern North Sea through assimilating remotely sensed data. *Ocean Science Journal*, 46(3):179–204, 2011. ISSN 1738-5261. doi: 10.1007/s12601-011-0015-x. URL `http://link.springer.com/10.1007/s12601-011-0015-x`.

G. Y. El Serafy, S. Gaytan Aguilar, and M. Blaas. Mos2-II : Forecasting SPM in the North Sea. Parameter Optimisation & Forecasting Capability, 2013.

M. A. Eleveld, H. J. der Woerd, M. Blaas, and G. Y. El Serafy. Using SPM observations derived from MERIS reflectances in a data assimilation scheme for sediment transport in the Dutch coastal zone. In *Proceedings of the Joint*. Citeseer, 2007.

Y. Gan, Q. Duan, W. Gong, C. Tong, Y. Sun, W. Chu, A. Ye, C. Miao, and Z. Di. Environmental Modelling & Software A comprehensive evaluation of various sensitivity analysis methods : A case study with a hydrological model. *Environmental Modelling and Software*, 51:269–285, 2014. ISSN 1364-8152. doi: 10.1016/j.envsoft.2013.09.031.

G. Gayer. Suspended particulate matter distribution in the North Sea., 2009. URL `http://www.coastalwiki.org/wiki/Suspended{_}particulate{_}matter{_}distribution{_}in{_}the{_}North{_}Sea`.

C. Jiayuan. Framework For Assessing Uncertainty in Ecological Risk Mapping A Case Study of North Sea in ECOSTRESS Project Framework For Assessing Uncertainty in Ecological Risk Mapping by. (August), 2015.

Z. Li, Q. Chen, Q. Xu, and K. Blanckaert. Generalized likelihood uncertainty estimation method in uncertainty analysis of numerical Eutrophication models: Take BLOOOM as an example. *Mathematical Problems in Engineering*, 2013. ISSN 1024123X. doi: 10.1155/2013/701923.

F. H. Los. A 3-dimensional primary production model (BLOOM / GEM) and its application to the southern North Sea (coupled physical-chemical-ecological model). J Mar Syst. (November 2015), 2008. doi: 10.1016/j.jmarsys.2008.01.002.

D. P. Loucks, E. van Beek, J. R. Stedinger, J. P. Dijkman, and M. T. Villars. *Water Resources Systems Planning and Management: An Introduction to Methods, Models and Applications*. 2005. ISBN 9231039989. doi: ISBN:92-3-103998-9. URL `https://www.utwente.nl/ctw/wem/education/afstuderen/Loucks{_}VanBeek/09{_}chapter09.pdf`.

H. Marencic. Quality Status Report 2009. Wadden Sea Ecosystem No. 25. *Quality Status Report 2009. Wadden Sea Ecosystem No. 25.*, (25):597, 2009. URL `http://www.waddensea-secretariat.org/sites/default/files/downloads/qsr-2009.pdf`.

M. D. McKay. Latin hypercube sampling as a tool in uncertainty analysis of computer models. In *Proceedings of the 24th conference on Winter simulation*, pages 557–564. ACM, 1992.

L. Meszaros. Ensemble forecasting system for Chlorophyll - a concentration prediction in the southern North Sea. Technical Report August, 2016.

H. Meuwese. *Nutrient Loads on the North Sea Feeding the North Sea (Master thesis)*. PhD thesis, Delft University of Technology, 2007.

L. Niu, P. H. A. J. M. van Gelder, C. Zhang, Y. Guan, and J. K. Vrijling. Statistical analysis of phytoplankton biomass in coastal waters: Case study of the Wadden Sea near Lauwersoog (The Netherlands) from 2000 to 2009. *Ecological Informatics*, 2015. ISSN 15749541. doi: 10.1016/j.ecoinf.2015.08.003.

NOOS. North West Shelf Operational Oceanographic System (NOOS), 2016. URL `http://www.noos.cc`.

N. C. Perez. Validation of ecological models : DELFT3d-GEM. Case study –North Sea and Wadden Sea, 2015.

T. Pfaffelmoser and R. Westermann. CORRELATION VISUALIZATION FOR STRUCTURAL UNCERTAINTY ANALYSIS. 3(2):171–186, 2013.

J. D. Pietrzak, G. J. de Boer, and M. A. Eleveld. Mechanisms controlling the intra-annual mesoscale variability of SST and SPM in the southern North Sea. *Continental Shelf Research*, 31(6):594–610, 2011. ISSN 02784343. doi: 10.1016/j.csr.2010.12.014.

P. Ramachandran and G. Varoquaux. Mayavi: 3D visualization of scientific data, 2011a. ISSN 15219615.

P. Ramachandran and G. Varoquaux. Mayavi: 3D Visualization of Scientific Data. *Computing in Science & Engineering*, 13(2):40–51, 2011b. ISSN 1521-9615.

Rijkswaterstaat. Waterdata, 2016. URL `https://www.rijkswaterstaat.nl/water/waterdata-en-waterberichtgeving/waterdata{#}39327`.

Sachin. The effects of Eutrophication, 2016. URL `http://sachinkbiology11.weebly.com/`.

K. Salacinska, G. Y. El Serafy, F. J. Los, and A. Blauw. Sensitivity analysis of the two dimensional application of the Generic Ecological Model (GEM) to algal bloom prediction in the North Sea. 221: 178–190, 2010. doi: 10.1016/j.ecolmodel.2009.10.001.

T. Schmidt. Coping with Copulas, 2006. ISSN 10978690.

D. Scott. Quantile-Quantile (q-q) Plots, 2007. URL `http://onlinestatbook.com/2/advanced{_}graphs/q-q{_}plots.html`.

Society MarineBio Conservation. Coral Reefs - MarineBio.org, 2015. URL `http://marinebio.org/oceans/coral-reefs/`.

D. E. Stuparu. Improving the description of the SPM concentration in the North Sea using a generic estimation technique for the process controlling parameters. Technical report, 2012.

SWAN team. SWAN - User manual SWAN cycle III version 41.10, 2016.

M. Tene. A novel copula-based sensitivity analysis method and its application to the Delft3D-WAQ sediment transport model, 2015.

L. Uusitalo, A. Lehikoinen, I. Helle, and K. Myrberg. Environmental Modelling & Software An overview of methods to evaluate uncertainty of deterministic models in decision support. *Environmental Modelling and Software*, 63:24–31, 2015. ISSN 1364-8152. doi: 10.1016/j.envsoft.2014.09.017. URL `http://dx.doi.org/10.1016/j.envsoft.2014.09.017`.

C. M. van der Hout, T. Gerkema, J. J. Nauw, and H. Ridderinkhof. Observations of a narrow zone of high suspended particulate matter (SPM) concentrations along the Dutch coast, 2015. ISSN 18736955. URL `http://dx.doi.org/10.1016/j.csr.2015.01.002`.

H. van der Woerd and R. Pasterkamp. HYDROPT: A fast and flexible method to retrieve chlorophyll-a from multispectral satellite observations of optically complex coastal waters., 2008.

T. Van Kessel, H. Winterwerp, B. Van Prooijen, M. Van Ledden, and W. Borst. Modelling the seasonal dynamics of SPM with a simple algorithm for the buffering of fines in a sandy seabed. *Continental Shelf Research*, 31(10 SUPPL.), 2011. ISSN 02784343. doi: 10.1016/j.csr.2010.04.008.

R. Vause. End-user Specific Flood Uncertainty Mapping : Effective Map Design from the Perspective of the English Public End-user Specific Flood Uncertainty Mapping : Effective Map Design from the Perspective of the English Public. (3832280), 2013.

Waddenzee.nl. De Waddenzee als PSSA, 2016. URL `http://www.waddenzee.nl/themas/veiligheid/scheepvaart/waddenzee-als-pssa/`.

Appendices

A Project structure

This appendix gives detailed information about the structure of the model set up. In Figure A.1 the model structure is visualized.

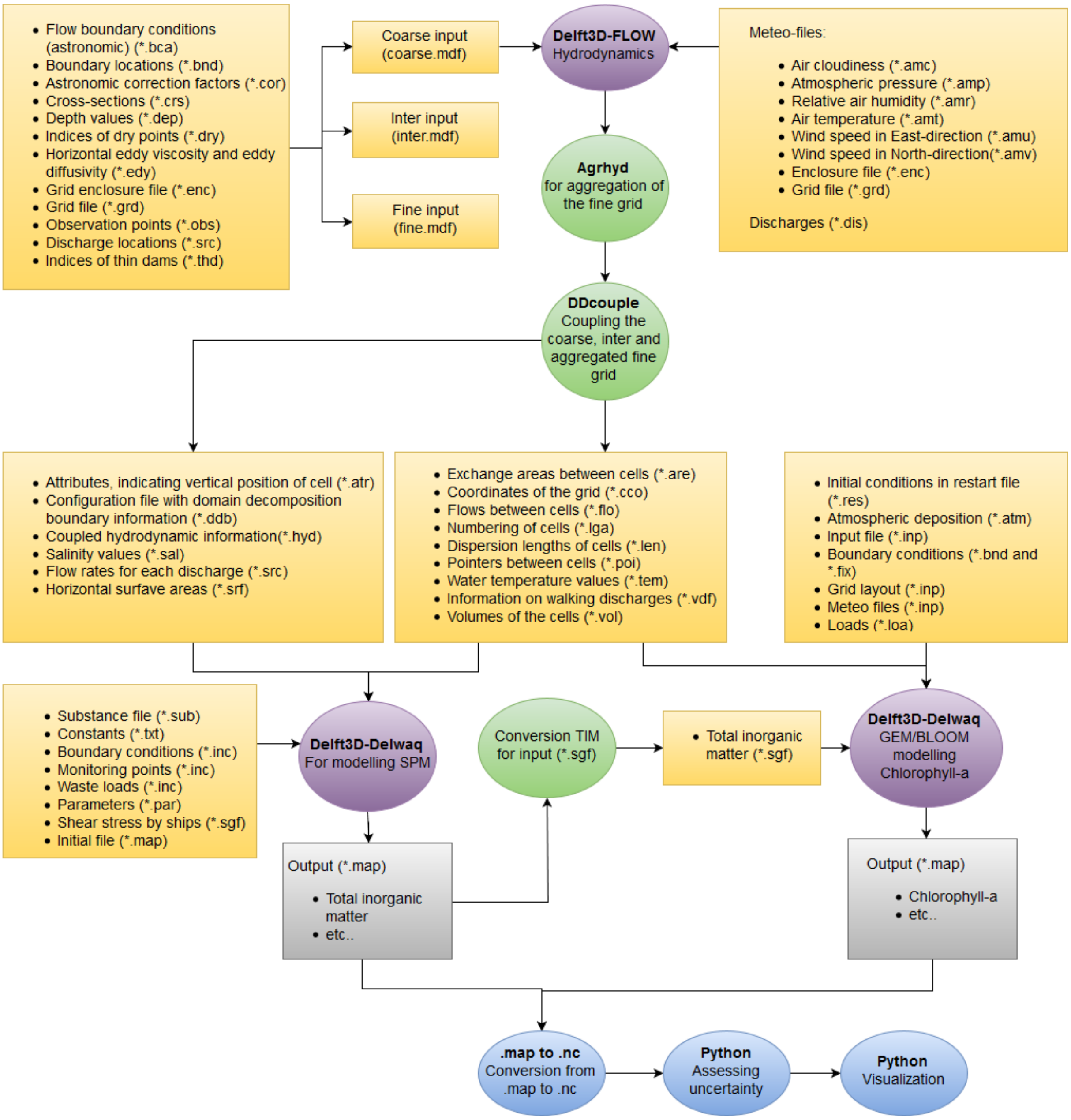


Figure A.1: Structure of the different models, files and processes.

In the flow chart a distinction is made between different steps in the process. The legend in Figure A.2 shows what these different boxes represent. The circles are programs used for the different steps, the squares indicate all the files that are output from these programs and input for next models. For more information on the processes and the output of the different models see Chapter 5.

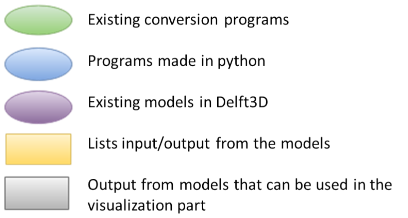


Figure A.2: Legend by the model structure in Figure A.1.

Hydrodynamic model
First of all the hydrodynamic model, which is a component of Delft3D-FLOW, needs to be run to get the information on the hydrodynamic processes, which is a driving force for the SPM and GEM/BLOOM models. This model uses the meteo-input, discharges and input data (such as depths, boundaries, dry points etc) for each of the underlying grids (the coarse, intermediate and the fine grid). From this model the output was data on hydrodynamics.

Agrhyd
To use the data files obtained from the hydrodynamic model for the next step, the SPM model in Delft3D-WAQ, some adjustments needed to be made. The SPM model uses measurement input data (such as the initial conditions file) which is defined on a different grid. This grid consists out of the same coarse and intermediate grid as is used in the hydrodynamic grid. But the fine grid is aggregated 2x2. In other words in the fine grid 4 cells in each layer is combined into 1 cell, see Figure A.3. This aggregation is done by using the program Agrhyd on the fine grid.

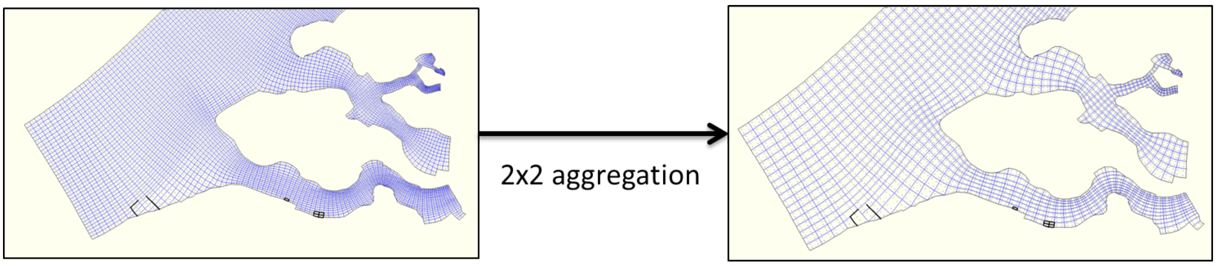


Figure A.3: 2x2 aggregation of the fine grid.

DDcouple
Another step that has to be taken before the SPM model can run with the hydrodynamic output is the coupling part. The hydrodynamic model has a domain decomposition between three different grids and therefore also the output will be in three different groups of files. With the program DDcouple these files are coupled together and only one set of hydrodynamic output becomes available, which corresponds to the grid in the SPM model.

SPM model
This model is set up in the Delft3D-WAQ environment. It requires a different different set of inputs, mainly on the forcings coming from the hydrodynamic processes. Also the initial file on the inorganic matter and the shear stresses by ships is needed. From this model the data on Total Inorganic Matter (TIM) is calculated, which is used as one of data values to test the visualization.

GEM/BLOOM
The model to calculate the other substance that is used for the visualization, is the GEM/BLOOM model which is set up in the Delft3D-WAQ environment. Similar to the SPM model, this model needs input from the hydrodynamic model, but also information on the meteo data, boundary conditions, grid layout and an initial conditions. Special of this model is the input of SPM concentrations, which comes from the SPM model. The output is the chlorophyll-a concentration, which is the second data set used for the visualization in this thesis.

Conversion SPM
Before the output of the SPM model can be used in the GEM/BLOOM model, the data needs to be converted to the correct file format to be accepted by the GEM/BLOOM model. GEM/BLOOM expects an include file (*.inc) to describe the forcing. Therefore a program is needed to convert the SPM data to this format.

Output data to netcdf file
For the visualization the outputted data needs to be easily accessible. However the .map file created by

the Delft3D models is a binary file which does not have a structure that can be easily used without the Delft3D-QUICKPLOT environment. Therefore it is chosen to convert this file to a netcdf (*.nc) format in which the information on longitude, latitude, depth and the required substances is stored. This program was already existing in form of Python code, but is was just a basis program. Much adaption was needed to get all the information on the segments in all layers, in multiple time steps and multiple substances from the map file.

Assessing variability and uncertainty
When the data is gathered and can be used in the Python environment a new program is written in Python to assess the variability and uncertainty. Because many model runs are required to give an indication of the uncertainty and the variability there are lots of output files, and after conversion a lot of netcdf files. However for the visualization only a mean value of the values in these files is needed to give an indication of the SPM and the uncertainty is therefore given by just one value in time and space for every location in th domain. This conversion step is needed to only have one file that will be used for the input of the visualization in the next step.

Visualization
This last step is to actually visualize the data together with the uncertainty in a 3D environment using Python. By means of the previous steps all data is neatly stored and can be accessed in the interactive visualization.

B | User manual

This appendix is a short user manual for the setup of the toolbox. The technical background is elaborated in which the requirements for using this toolbox are explained. The toolbox is created with the programming language Python, combining many different packages and methods, which are described shortly in this section. Technical requirements to use the code are given, as are the requirements for input and the data files. Furthermore explanations are given on the different event handles in the interface and how to use the toolbox.

B.1 Technical requirements

As mentioned, the code for the toolbox is written in a Python file <*.py>. This code can be run in any Python environment, for example Spyder (Scientific PYthon Development EnviRonment). In this code many standard installed libraries are used, as Numpy and Matplotlib. However for the interactive parts, the 3D environment and the graphical objects some extra libraries are needed. In table B.1 a short overview of the needed libraries is given.

Table B.1: Software and libraries used for the visualization together with their functionalities and links to their website for more information (adapted from Ramachandran and Varoquaux [2011a]).

	Function	**Link**
Python	Open-source fast and easy to use programming language.	`https://www.python.org/`
Mayavi	3D scientific data visualization and plotting in Python. Used to create the 3D environment and interface.	`http://docs.enthought.com/mayavi/mayavi/#`
VTK	Visualization ToolKit: general-purpose, open-source, visualization and graphics library for Python. A pipeline architecture used by Mayavi, where the output of one element is the input for the next, where the elements are connected in serie. In this toolbox it visualizes the data.	`http://www.vtk.org/`
Traits	The Traits library extends Python object attributes and give them extra characteristics: initialization, validation, delegation, notification and visualization. With these Traits interaction is enabled in the toolbox, e.g. when a change in a Trait is notified, it enables another element to initiate.	`http://code.enthought.com/projects/traits`
TVTK	Traited TVTK provides a traits enabled version of VTK, which wraps around VTK objects and provide a convenient Pythonic API.	`http://docs.enthought.com/mayavi/tvtk/README.html`

B.2 Input

The toolbox is based on two input files, one for the information on the x-, y-, and z-coordinates of the grid and the other on the information of the concentration and the uncertainty. Both files are stored as NetCDF files (*.nc). The depth fie should contain three variables: 'lon' (for longitudinal information), 'lat' (for latitudinal information) and 'LocalDepth', representing the x-, y-, and z-coordinates of all the segments in the domain. These coordinates do not change through time. The file with the concentrations should contain the variables 'mean' representing the average values of the concentrations from the uncertainty analysis and the 'var', representing the variance of spreading in the output, which will be used as the range of uncertainty in the visualization. This variance is stored as one standard deviation from the average value. The shape of the arrays in which the values are stored should be equal. For the file with concentrations another dimension is implemented for the time.

B.3 User keys

The interface that is created has many functionalities, which can be used with certain proceedings of the mouse. The functionalities are all described in Table B.2. In Chapter **??** more information is given on what these functions mean and where they can be used for.

Table B.2: Important user keys and functions of the toolbox.

Key combination	**Location in toolbox**	**Function**
Hold left mouse button + drag	Scene 1	Rotation of the 3D environment
Hold right mouse button + drag	Scene 1	Zoom in/out of the 3D environment
Hold with mouse wheel button + drag	Scene 1	Pan through the 3D environment
Hold left mouse button + drag	Cutplane in Scene 1	Move Cutplane over 3D environment
Hold left mouse button + drag	Center of Cutplane in Scene 1	Rotate orientation of Cutplane
Click left mouse button	Marker in Scene 1	Selection of marker
Click left mouse button	Start button	Start time animation
Click left mouse button	Stop button	Stop time animation
Click left mouse button + insert time step	Time box	Change time step manually
Click left mouse button	Make cross section button	Cut 3D environment in scene 1 through at location of the cuplane