Sensitivity analysis for hydrodynamic model of the North Sea

Considering the correlations and dependencies between parameters

N. Ye

Delft University of Technology

TU Delft

Sensitivity analysis for hydrodynamic model of the North Sea

Considering the correlations and dependencies between parameters

by

N. Ye

to obtain the degree of Master of Science

at the Delft University of Technology,

to be defended publicly on Thursday October 28, 2021 at 2:30 PM.

Student number:	5135532	
Project duration:	March, 2021 – October, 2021	
Thesis committee:	Prof. dr. dr. ir. A. W. Heemink,	TU Delft, supervisor
	Dr. Ir. El Serafy,	TU Delft, Deltares, supervisor
	Dr. D. Kurowicka,	TU Delft
	R. Santjer, MSc,	TU Delft, mentor

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

The cover image is cited from `https://www.nipic.com/show/22934605.html`.

TU Delft Deltares

Preface

At this time last year, I was in Shanghai doing an internship. Sometimes it is hard to connect life back in China with the life here, especially under the epidemic. The strong contrast seems to be an illusion. I always felt life peaceful but boring in the first year, but after I flew back to the Netherlands this January, I gradually felt the I-called 'inner peace'. It also reminds me why I wanted to come to the Netherlands, specifically TUDelft, for study since my first year of undergraduate. It has been a fulfilling journey for the past two years, courses, homework, projects, internships, thesis, etc., and I still remember taking six mathematical courses simultaneously in the first year. I never recognized myself as a 'mathematics' person, and it is a little surprising to find I have been kept going this way when I look back. No matter the grades, and if I continue studying or seek an unrelated job, I am proud of myself to graduate from mathematics.

Hydraulics is a totally unfamiliar area for me. During the last eight months, the from-zero-to-one thesis project allows me to understand new areas such as hydraulics, aquaculture and some stochastic knowledge. It also gives me some confidence in mathematics, coding and academics. Working in the physical engineering area was not planned either, just like I did not expect I would continue on mathematics. However, it is the unexpectation that led me to where I am.

Working in Deltares is an amazing experience. The environment is very friendly and relaxing, and the colleagues are very helpful and kind. Although my internship was almost finished when we were allowed to come back to the office, we still had some delightful meetings in person. Many thanks to Rieke, my daily supervisor. I was worried about being unfamiliar with hydraulics, all the models, software, clusters and related knowledge. You helped me a lot in different aspects and always told me not to worry. All the suggestions are so useful and detailed. When I saw your comments for my report sentence by sentence, I was so grateful and a little bit surprised, honestly, as I did not expect to have that detailed advice. You always encourage me, approving my work. I could not finish my thesis without your help. I would also like to thank Julien and Firmijn from Deltares. You answered many questions, explaining the knowledge I was not familiar with patiently. With your explanations, I can understand better with my results, and some discussions can be made. You made my results closer to reality. Prof. Heemink and Dr El Serafy, thank you for your trust and valuable advice. Moreover, for Arnold, it is not only during the thesis. Since the first year in the data assimilation course, I have known Arnold, and we also had communications for my internship last year. Doing the internship back in China was a little difficult to be approved before, but you supported me kindly considering the situation.

Personally, I am truly grateful to my parents and brother, letting me feel loved and supported no matter what is going on. Thank you, Terry, for the beautiful cover and for accompanying me when I struggled. Thank you for all my friends, wherever you are now, in Netherlands or back in China, and no matter whether we often meet or not, the connections and true friendships are always what I cherish from the bottom of my heart.

N. Ye
Delft, October 2021

Abstract

In this thesis, sensitivity analysis is used to study the influences of parameters on specific outputs in the hydrodynamic model 3D DCSM-FM of the North Sea. The sensitivity analysis is the study of how uncertainty in the outputs of a model can be divided and allocated to different sources of uncertainty in its inputs. The software Delft3D FM is used to simulate the model, which is developed by Deltares. This thesis is supported by German pilot of the UNITED project, which studies the possibilities of combing blue mussel and sugar kelp cultivation with wind energy. The investigation is conducted at FINO3 platform, 80 km off Sylt.

In this thesis project, temperature and current velocity are selected as outputs in sensitivity analysis, which are influential factors of blue mussel and sugar kelp growth. Specific parameters in the hydrodynamic model are selected as inputs respectively. Three sensitivity methods are conducted: Morris, copula-based and variance-based method. Among them, copula-based and variance-based consider the independency information, while parameters are assumed dependent in Morris method. To generate samples, parameters are transformed into a unit hypercube in Morris method and run on the contour of the grid cell, composing paths. Between two steps within the paths, only one parameter changes at a time (OAT method). The input domain is scanned with a better strategy to separate the paths maximizing the dispersion. Each parameter's elementary effects are calculated within paths, evaluating the changes in outputs contributed by the single parameter. Then (absolute) mean and variance of elementary effects are used as sensitivity indices. Copula-based method uses similar sampling and the same measurements, but it gathers parameters into copulas before sampling to include the dependency information. In variance-based method, the variance of the outputs' conditional expectations is used to measure the sensitivity. Only random samples are needed in this method.

Morris and copula-based method prove the temporal and spatial similarities of parameters' sensitivity behavior. Convective and evaporative heat flux are the most influential parameters of temperature, and they also show correlations with other parameters. Air density influences current velocity the most, while Smagorinsky factor is the most correlated parameter of current velocity. Variance-based method gives similar results about rankings of influences. Correlations are proved to exist among parameters. Uniform horizontal eddy viscosity/viscosity in definition files have no impacts, as they are overwritten. Three methods are compared. Except the differences in independency, sampling and measurement, there are some other differences. Much more samples are required in variance-based method and it is used mainly to decide the existence of significant correlations and whether a parameter can be neglected. While Morris and copula-based method ranks the influences and correlations.

Contents

Abstract v

Notations ix

1 Introduction 1
1.1 Background and basic information of the project . . . 1
1.2 Objectives. . . . 2
1.3 Research questions . . . 2
1.4 Innovations and related study. . . . 3
1.5 Thesis structure . . . 3

2 Literature review 5
2.1 Mussel and seaweed aquaculture . . . 5
2.1.1 Techniques adopted in mussels and seaweeds cultivation . . . 5
2.1.2 Influential factors of blue mussels . . . 6
2.1.3 Influential factors of seaweeds . . . 7
2.2 Introduction of sensitivity analysis . . . 7
2.3 Morris method . . . 7
2.3.1 One-at-a-time (OAT) sampling. . . . 8
2.3.2 Elementary effects and sensitivity measurements . . . 9
2.3.3 Sampling strategy . . . 10
2.3.4 Method of optimized trajectories: enhancing the sampling strategy . . . 11
2.4 Copula-based method . . . 12
2.4.1 The consideration of dependencies and correlations between parameters. . . . 12
2.4.2 Copula . . . 12
2.4.3 Sampling strategy . . . 13
2.4.4 Choosing the target block . . . 14
2.4.5 Choosing the starting point . . . 15
2.4.6 Choosing the traversal order . . . 16
2.4.7 Sensitivity measurements . . . 16
2.5 Variance-based method. . . . 16
2.5.1 General idea of using variance . . . 16
2.5.2 The setting of the model and the first-order sensitivity index . . . 17
2.5.3 Internal effects . . . 17
2.5.4 Total effects . . . 18
2.5.5 Computation strategy of the sensitivity indices . . . 19

3 Methodology 21
3.1 The hydrodynamic model and the software used for simulations . . . 21
3.1.1 The hydrodynamic model: 3D DCSM-FM model . . . 21
3.1.2 The simulation software: Delft3D FM . . . 21
3.1.3 The files and operation process of the model in the thesis project . . . 22
3.2 The choices of observation points. . . . 23
3.3 The choices of outputs and inputs . . . 24
3.3.1 Choices of outputs . . . 24
3.3.2 Choices of inputs . . . 24
3.4 Morris method . . . 25
3.4.1 Sample generation . . . 25
3.4.2 Simulations of the models and time series of outputs . . . 27
3.4.3 Sensitivity measurements . . . 31

3.5 Copula-based method . . . 32
3.5.1 Copula decision . . . 32
3.5.2 Sample generation . . . 33
3.5.3 Simulations of the models and time series of outputs . . . 34
3.5.4 Sensitivity measurements . . . 35
3.6 Variance-based method . . . 37
3.6.1 Sample strategy . . . 37
3.6.2 Simulations of the models and choosing timepoints . . . 38
3.6.3 Sensitivtiy measurements . . . 38

4 Results 39
4.1 Morris method . . . 39
4.1.1 Temperature . . . 39
4.1.2 Current velocity . . . 43
4.2 Copula-based method . . . 50
4.2.1 Temperature . . . 50
4.2.2 Current velocity . . . 53
4.3 Variance-based method . . . 58
4.3.1 Temperature . . . 58
4.3.2 Current velocity . . . 58

5 Discussions, conclusions and recommendations 61
5.1 Discussions . . . 61
5.2 Summary and conclusions . . . 63
5.3 Recommendations . . . 66

A Codes: pretreatments 67

B Codes: read the outputs and create time series plots 71

C Codes: Morris method 73

D Codes: copula-based method 75

E Codes: variance-based method 79

Bibliography 83

Notations

Generally used

X_i	Parameter
Y	Outputs
k	Number of parameters

Morris method

n	Number of inputs
p	The level in the hypercube
r	The number of elementary effects being calculated
s	$s \in \{1,2,\ldots,p-1\}$
δ	Morris step
d_j	Elementary effects
μ_j	Elementary mean
μ_j^*	Absolute elementary mean
σ_j	Standard deviation
d_{ml}	The distance between a couple of trajectories m and l

Copula-based method

C/u	Copulas
ρ	Correlations
$R_j[i]$	The rank statistics, representing the order of the samples
$\mathbf{cell}^i[j]$	The vector containing the coordinated of the origin of the target cell

Variance-based method

$f_0, f_i, f_{ij}, \ldots$	Functions of the factors in the index
V_i	The first-order effect of X_i on Y
S_i	The first-order sensitivity index
V_{ij}^c	The joint effect of the pair (X_i, X_j) on Y
V_{ij}	The joint effect of X_i and X_j, V_{ij}^c minus the first order of X_i and X_j
S_{ij}, S_{ijl}	The higher-order sensitivity index
S_{Ti}	The total effect index

Inputs and timepoints

The abbreviations can be found in table 3.2.

The abbreviations of timepoints can be found in table 3.7.

1 Introduction

1.1. Background and basic information of the project

Covering approximately 71 percent of the Earth's surface, sea has abundant resources and potential to explore and use. Humans have a long history of exploring the ocean. Austronesians on Taiwan have begun spreading into maritime Southeast Asia by 3000 BC [20]; the ancient Egyptians and Phoenicians explored the Mediterranean and Red Sea around 2750 BC [5]; the Chinese Ming Dynasty sailed the Indian and Pacific Oceans in the early fifteenth century [47]; in the late fifteenth century, Western European mariners started making longer voyages. Compared to the history of exploring the ocean itself, the scientific study of oceanography started much later, from approximately 1768 with the voyages describing the Pacific [47]. Nowadays, oceanographic research ranges wide, including marine lifeforms, conservation, the marine environment, the chemistry of the ocean, the studying and modelling of climate dynamics, the air-sea boundary, weather patterns, ocean resources, renewable energy, waves and currents, and the design and development of new tools and technologies for investigating the deep [6]. Through the scientific research, humans can understand more about the ocean and conduct various industries to obtain resources from the ocean, including aquaculture, power generation by ocean waves, tides and salinity differences, extractive industries to obtain minerals, freshwater production, etc.

Among all the resources from the ocean, fish and other fishery products are one of the earliest and most widely obtained sources from the ocean. Fish and fishery products are the main sources of protein and other essential nutrients [47]. In 2011, the total world production of fish, including aquaculture, was estimated to be 154 million tonnes, of which most was for human consumption. The harvesting of wild fish accounted for 90.4 million tonnes, while others are from annually increasing aquaculture [29].

According to the Food Agriculture Organization (FAO), aquaculture "is understood to mean the farming of aquatic organisms including fish, molluscs, crustaceans and aquatic plants. Farming implies some form of intervention in the rearing process to enhance production, such as regular stocking, feeding, protection from predators, etc. Farming also implies individual or corporate ownership of the stock being cultivated." [46] Although there are some issues of aquaculture, such as environmental problems, impacts on wild fish, coastal ecosystems [35], freshwater ecosystems [38], etc., the importance of aquaculture cannot be denied. From a report of FAO, in 2014 output from global aquaculture in 2014 supplied over one half of the fish and shellfish that is directly consumed by humans [4]. To improve the operations of aquaculture, different techniques are investigated, including combination of aquaculture and other industries to increase the efficiency, which is applied in the UNITED project.

This thesis project is part of the UNITED project, which is a research project financed by the European Union Horizon 2020 programme. [1] The acronym UNITED stands for "Multi-Use offshore platforms demoNstrators for boostIng cost-effecTive and Eco-friendly production in sustainable marine activities". The UNITED project runs from 2020 to 2023 and provide evidence for the viability of ocean multi-use, which is the international shared use of marine resources in close geographic proximity by two or more maritime activities, besides the joint use of installations. The UNITED activities will enhance the viability and study current chal-

lenges across five key pillars: technical, economic, environmental, society and legal, policy and governance pillars. To study the large-scale multi-use, the UNITED activities will be based in five pilot sites, including German, Dutch, Belgian, Danish and Greek pilots, where different combinations of activities are carried out.

This thesis project focuses on the German pilot of the UNITED project. The German pilot investigates the possibility of combining offshore wind energy and aquaculture. To be more specific, the blue mussels, or Mytilus edulis, and the seaweeds, or specifically Saccharina latissima (which is also called sugar kelp) are grown. In this thesis project, sensitivity analysis will be carried out in a hydrodynamic model called the 3D Dutch Continental Shelf Model developed in D-HYDRO Flexible Mesh (3D DCSM-FM). The model can simulate the hydrodynamic process within the Northwest European Shelf, including the North Sea and adjacent shallow seas and estuaries such as the Wadden Sea and the Eastern and Western Scheldt [57]. The model has been validated with respect to water levels, water temperature, sea surface salinity and residual currents [57]. To simulate the model, the software Delft3D Flexible Mesh Suite (Delft3D FM) is used. The details of German pilot, 3D DCSM-FM model and Delft3D FM software will be introduced in Chapter 3 Methodology.

The aim of sensitivity analysis is to study the influences of parameters to outputs. This thesis will especially take the dependencies and correlations between parameters into consideration and apply it to Delft3D FM. The theoretical details of sensitivity analysis will be introduced in chapter 2.

The simulations at several observation points in North Sea are carried out and two outputs with corresponding inputs in 3D DCSM-FM model are chosen. Three methods of sensitivity analysis are performed and compared to study the influence of inputs to outputs.

1.2. Objectives

As a powerful model, 3D DCSM-FM model has complex relations between parameters and output variables. To investigate the relations between inputs and outputs of the model, sensitivity analysis is adopted in this thesis project. Instead of giving the concrete physical formulas, which is hard to obtain with abundant parameters, sensitivity analysis evaluates the relations through the indicators showing how much influence the inputs have on outputs. Besides, in the physical process, the correlations between parameters can always be assumed, which is a problem needs to be tackled in sensitivity analysis. On the other word, the main goal of this thesis project is to use sensitivity analysis to study the impacts of parameters to outputs in 3D DCSM-FM model, considering the correlations and dependencies between the parameters.

Around the main goal, some specific goals follow:

- Select the target outputs and corresponding inputs in 3D DCSM-FM to apply sensitivity analysis.
- Find the indices to measure the sensitivity of parameters.
- Find the proper way to include the independence and correlations between parameters into sensitivity analysis.
- Investigate different methods to do sensitivity analysis and compare the differences.

1.3. Research questions

Following the objectives of the thesis, some questions need to be answered. They will be studied around the keywords, sensitivity analysis and independence of parameters. Corresponding to the main goal of the thesis in chapter 1.2, the main question will be, how the uncertainty in the output of the hydrodynamic model can be divided and allocated to the uncertainty of the parameters in the model, which is also the definition of sensitivity analysis [42]. Considering the application of sensitivity in the case, the following detailed questions will be answered:

- Sensitivity measurement: which indices can be selected to measure the sensitivity of parameters and how to explain them?
- Research prioritization: which factors are the most deserving of further analysis or measurement?
- Model simplification: can some factors or compartments of the model be fixed or simplified?

- In the physical process, independence is rare to see, and parameters can be imagined to be dependent and interact with each other, which is also the highlights of the thesis. Some questions related to dependencies and correlations need to be answered:
  - Why do we need to consider the dependencies of the parameters?
  - How to include the dependency information into sensitivity analysis?
  - Which parameters are correlated together?
- From the aspects of different methods, what are the differences between the logic and results of the methods?

1.4. Innovations and related study

This project uses three existed sensitivity analysis methods to study the parameters in 3D DCSM-FM model.

There are some previous studies about sensitivity analysis on models simulated by Delft3D. Zhijie Li et al. [27] developed a two-dimensional algae dynamic model in Taihu Lake in China, using Delft3D-BLOOM based on the results from Delft3D-FLOW, and Morris screening was used to analyze the sensitivity of parameters. In a previous master thesis at TUDelft [52], sensitivity analysis was carried out in fine sediment transport in the humber estuary, which was simulated by Delft3D-Delwaq. In a research by Bastidas et al. [7], a model parameter sensitivity analysis was presented, for the simulation of storm surge and wave by Delft3D. In a paper by Ţene et al. [51], a creative sensitivity analysis method, copula method, was carried out on the simulations of Delft3D-WAQ in North Sea.

The existing studies have some common concerns with the thesis project on studying sensitivity analysis based on the simulations by Delft3D. However, there are still some unique innovations in this project. The choices of temperature and current velocities as outputs in sensitivity analysis are barely seen in the previous studies. About the software, the existing studies include some other models of Delft3D, such as BLOOM, Delwaq, WAQ, etc., while Delft3D FM module is used in this project. From the aspects of sensitivity analysis methods, three different methods are carried out in this project, which will be compared to each other, distinguishing whether to consider independence, while previous research lacks application and comparison of different methods at the same time.

Focusing on the German pilot, this thesis contributes to the UNITED project. Promoting the combination of aquaculture and wind energy is one of the goals of the UNITED project. If the operation is proved to be feasible on the experience platform, it can be promoted to other location. Sensitivity analysis can simplify the research of different sites, allowing us focus mainly on influential parameters, which can be distinguished through sensitivity analysis.

From the aspect of 3D DCSM-FM model, the project can also give some clues for the physical process. The real process of the hydrodynamic movements is complicated, affected by a lot of factors. With the results of sensitivity analysis, we can have an impression of which factors are related to the physical process, and how important they are.

1.5. Thesis structure

The first chapter gives the general introduction of the project, including the motivation, questions and objectives, and related study.

In the second chapter, the theoretical explanation of three methods used to carry out sensitivity analysis is given, along with other literature review about the mussels and seaweeds aquaculture, which is related to the choices of outputs in the analysis.

Chapter 3 describes the methodology used in the thesis project, corresponding to the theoretical part in chapter 2. In the details, this chapter introduces Delft3D FM model used to simulate results, the choices of outputs, inputs, and observation points. Furthermore, the operations of different methods and specific problems met in the process will be mentioned.

Chapter 4 presents the results of the analysis, including sampling, sensitivity indices and analysis. It will separate three different methods apart.

In the last chapter, the conclusions of the studies are summarized and the comparison between different observation points and methods will be discussed about.

2

Literature review

In this chapter, techniques and influential factors of mussel and seaweed aquaculture will be introduced first, as German pilot of the UNITED project adopted longline technique in the growth of the blue mussels called Mytilus edulis, and the seaweeds called Saccharina latissimi (which is also called sugar kelp).

Then the theoretical concepts and three different methods of sensitivity analysis will be presented.

2.1. Mussel and seaweed aquaculture

2.1.1. Techniques adopted in mussels and seaweeds cultivation

Cultivation of mussels and seaweeds provides opportunities for food production while removing excess nutrients in eutrophic coastal water [53].

Two main approaches are used in mussel aquaculture, bottom culture and suspended culture, where 'bottom' means the ocean floor. A bottom culture method means the mussels are growing on the ocean bottom, and in suspended culture the mussels grow without touching the bottom [3]. In 2006, approximately 15% of mussel aquaculture takes bottom culture and suspended culture accounts for 85% [31].

In bottom culture, the mussels grow on natural mussel beds, which are similar to wild growth. There are three main approaches performed in suspended culture, bouchots (or poles, driving poles or stakes into the ground in regions with high tidal regime and mussels grow on the poles), raft (suspending mussels attached on ropes or in socks from a moored rafts with cylindrical floats), and longline system [31]. The first two techniques

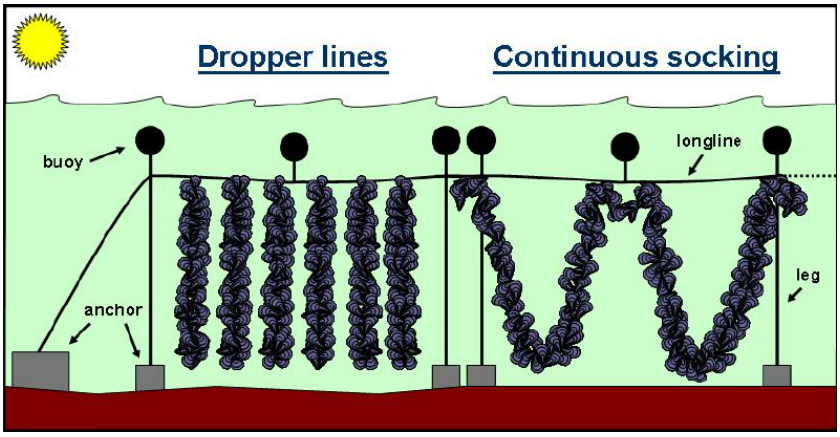
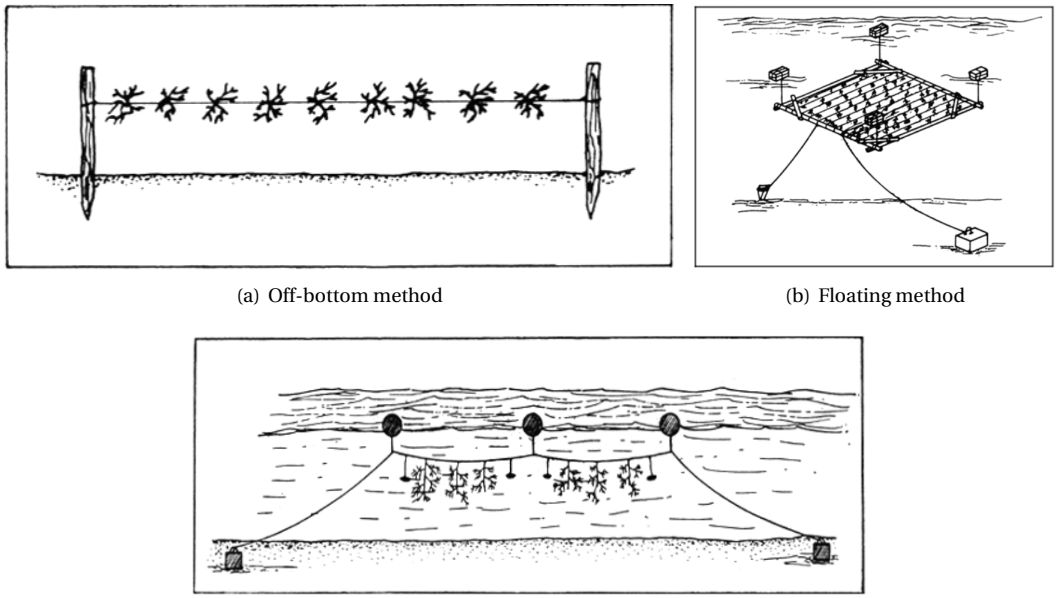


Figure 2.1: Schematic of a basic mussel longline system. The two most common systems used in Canada are shown. Longline systems may also be used for the culture of other species in nets, trays, etc. [31]



(a) Off-bottom method

(b) Floating method

(c) Longline method

Figure 2.2: Three different cultivation methods of seaweeds [16]

are only conducted in a small number of areas, while longline is suitable for many environments and used around the world. As shown in figure 2.1, a series of lines are anchored at ends, and buoys float in the water hanging a series of dropper lines with weights at their ends. Mussels are hung along the lines in socks (or dropper lines) [31].

There are at least three methods applied in seaweed aquaculture, off-bottom (fixed bottom) method, raft or floating method and longline method [16], which are correspondingly similar to the three methods in suspended culture of mussels cultivation.

The off-bottom method is similar with bouchots (poles) techniques in mussel cultivation. Wooden stakes are driven into the sea bottom and seaweeds are planted along the ropes between stakes. In the floating method, seaweed is attached to floating rafts, rising and falling with tidal changes [16], similar to raft technique in mussel cultivation. The longline method is the same as in mussel aquaculture.

Garen et al. [18] compared the growth of mussel on longline, pole and bottom cultures, and observed a clear seasonal growth pattern, similar for all the three cultural type. Also, length and weight growth appeared different in the three type, longline mussels exhibiting the highest performance and bottom-type showed the lowest.

2.1.2. Influential factors of blue mussels

A lot of factors influence the growth of mussels, meat content being the main indicator, such as food availability, seawater velocity, distribution of diarrheic shellfish toxins (DST), temperature, suspended sleeves, etc. Some of the factors have correlation with each other.

Growth of mussels depends largely on food availability [8, 48, 55], which is influenced by seston concentration, composition and transport rate [9, 17, 22].

Seawater velocity and crowdedness will influence transport rate of seston, which in turn affects mussel growth. In the study performed [48], as current velocity and Chlorophyll *a* concentration decreased, the calculated food availability was reduced by more than 80% within the first 30 m of the farm and from 30 m the food availability remained low throughout the farm. It is observed by farmers that meat content along the edges of the farm is higher than the mid-section, which is a probable result of lower seawater velocity, higher crowdedness and friction in the mid-section, and the structure of farms.

A significant negative correlation between meat content and DST is observed [48]. It is suggested that depuration of DST was faster in areas with high food availability.

Temperature is another factor of mussel growth. Higher temperature increased the growth fates of blue mussels, where 16 °C is a proper temperature in experiments [39]. The temperature is also correlated with seston and nutrients in water, which affect the growth directly. Mussels growth was higher during summer at high exposure sites, probably due to a seasonal increase in nutrients [19].

Cultivation methods and the set of tools also influence mussel growth. Longline mussels performed the highest growth among all the techniques [18]. In longline method, the spacing of suspended sleeves is related to seston uptake. Different spacings have no significant effect on the condition index of mussels (the condition index is calculated by the ratio of dry tissue weight and dry shell weight times 100). However, high spacing positively affected shell growth and abundance for small seeds packed within sleeves. High spacing will also increase farm productivity and seston uptake at bay scale [14].

2.1.3. Influential factors of seaweeds

The yield and quality of seaweeds are largely controlled by environmental factors, such as light, salinity, nutrients (such as bioavailable phosphorus), temperature and water velocity.

Light is a main positive-corelated factor for growth of S. latissima in spring [10]. Salinity also showed a positive influence on frond length of seaweeds [36]. Phosphorus availability appeared to control the primary production [10]. Temperature generally shows positive control on the growth and development of juvenile filter-feeders [54].

Water velocity substantially affects the seaweed production, directly increasing the uptake of nutrients and carbon dioxide. Productivity of macrophytes are generally believed to be higher at moderate levels of water velocity compared to slower water speed [40].

2.2. Introduction of sensitivity analysis

According to the definition, sensitivity analysis is the study of how uncertainty in the output of a model (numerical or otherwise) can be apportioned to different sources of uncertainty in the model input [42].

Sensitivity analysis can be categorized into local and global sensitivity analysis. Local sensitivity analysis is the assessment of the local impact of input factors' variation on model response by concentrating on the sensitivity in vicinity of a set of factor values [43]. Gradients or partial derivatives of the output functions are commonly used to evaluate local sensitivity, where the values of other input factors are kept constant while the target input factor is changed. However, local sensitivity has severe restrictions of uncertainty and linearity. In other words, the results are only effective at a point and cannot be proved in the whole space of the input factors. As a solution, global sensitivity is introduced to use large number of data points over the whole input space. The average results will be obtained from different sets of points in the input space, avoiding the local limitations. In this project, all the methods belong to global sensitivity analysis, and 'sensitivity analysis' in the thesis refers to global sensitivity by default.

A lot of methods can be adopted to carry out sensitivity analysis, depends on the requirements of the problems, such as linearity, dependency, calculation time, etc. In this project, three methods will be used to do sensitivity analysis: Morris method, copula-based method and variance-based method. Among them, all the input factors are assumed to be independent in Morris method, while copula-based method and variance-based method take dependency of input factors into consideration.

2.3. Morris method

In this section, Morris method will be introduced, which is also called Elementary Effects (EE) method. The fundamental idea was innovated by Morris in 1991. Two sensitivity measures are used to determine which input factors may have effects which are (a) negligible, (b) linear and additive, or (c) non-linear or involved in interactions with other factors. The sensitivity measures will be introduced in chapter 2.3.2. Morris method

is one of the most fundamental and powerful method, especially when the number of input factors is very large. However, it is worth noting that the input factors are assumed independent with each other in Morris method, which is also a disadvantage of the method. The improvements to this problem will be presented in the next two methods, copula-based method and variance-based method.

2.3.1. One-at-a-time (OAT) sampling

The sampling strategy's main core of Morris method is One-at-a-time (OAT) sampling. Generally, in every two adjacent samples, only one parameter changes at a time, therefore the change in output is only affected by the parameter changing in adjacent samples. In this way, the influences of each parameter will be easier to distinguish.

To explain OAT sampling in details, first an equation will be presented:

$$\begin{aligned} Y &= b_0 + \sum_{r=1}^{k} b_r X_r \\ &= b_0 + b_1 X_1 + b_2 X_2 + \cdots + b_{k-1} X_{k-1} + b_k X_k \end{aligned} \tag{2.1}$$

Usually in the analysis where the relationship between inputs and outputs is unknown, each parameter can be assumed to be linear to the output. In the formula (2.1), Y is the output, X_1 to X_k are the k variables, and b_r are all constants assumed unknown at the start of the sensitivity analysis.

To analysis the sensitivity, the data points are provided. With N simulations, where $N \geq k+1$, the formula (2.1) can be written in the form of matrix, and b_r can be solved with the value of data points:

$$\begin{bmatrix} 1 & x_{11} & \dots & x_{1k} \\ 1 & x_{21} & \dots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{N1} & \dots & x_{Nk} \end{bmatrix} \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_k \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix} \tag{2.2}$$

Using matrix notation, the formula (2.2) can be abbreviated as

$$X_{Nk} B_k = Y_N \tag{2.3}$$

The matrix X_{Nk} has 1 in the first column, and values for the k parameters in N simulations in other columns. B_k contains the $k+1$ unknown coefficient corresponding to the intercept b_0 and the k parameters. Y_N contains outputs from N simulations.

After introducing the linear equation, OAT sampling is easier to understand.

The equation (2.2) can be rewritten as

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \dots & 0 \\ 1 & 1 & 0 & 0 & 0 & \dots & 0 \\ 1 & 1 & 1 & 0 & 0 & \dots & 0 \\ 1 & 1 & 1 & 1 & 0 & \dots & 0 \\ 1 & 1 & 1 & 1 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & 1 & 1 & \dots & 1 \end{bmatrix} \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_k \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{k+1} \end{pmatrix} \tag{2.4}$$

where the variables are assigned the value of 0 and 1, and between two adjacent rows, only one variable changes its value. To make the system more simplified, subtract the entries from the previous row in every

row but the first, and the following can be obtained:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & 1 \end{bmatrix} \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_k \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 - y_1 \\ y_3 - y_2 \\ y_4 - y_3 \\ \vdots \\ y_{k+1} - y_k \end{pmatrix} \tag{2.5}$$

The equation (2.5) shows that if there is any change in value between y_i and y_{i+1}, it can only be attributed to a change in parameter X_i. The change in outputs $\Delta y_i = y_{i+1} - y_i$ is an estimate of the effect on y of changing X_i from 0 to 1.

OAT sampling provides an idea of how to evaluate the effect of one parameter at one time, without the influence of other parameters. However, in real analysis, more specific sampling strategy should be considered, including the distribution of the input samples in each simulation.

Also, there is disadvantage of OAT sampling. OAT sampling is inefficient when the number of parameters k is large and only a few of them influence the outputs. In this situation, most of the simulations are used to find noninfluential parameters, where the outputs keep the same, and only a few of the outputs will include the information of influential parameters. In this project, the number of parameters is limited, allowing us to dismiss the disadvantage.

2.3.2. Elementary effects and sensitivity measurements

Elementary effects are the indicators and measures of factors' sensitivity in Morris method. Therefore, Morris method is also called Elementary Effect method. Consider a model with k independent inputs X_i, $i = 1, \dots, k$, and the ranges of the parameters are divided into p selected levels. The region of experimentation Ω is thus a k-dimensional p-level hypercube. In order to eliminate the influences of different ranges of parameters, evaluating elementary effects independently of parameter ranges, each x_j is first scaled uniformly to $[0, 1]$.

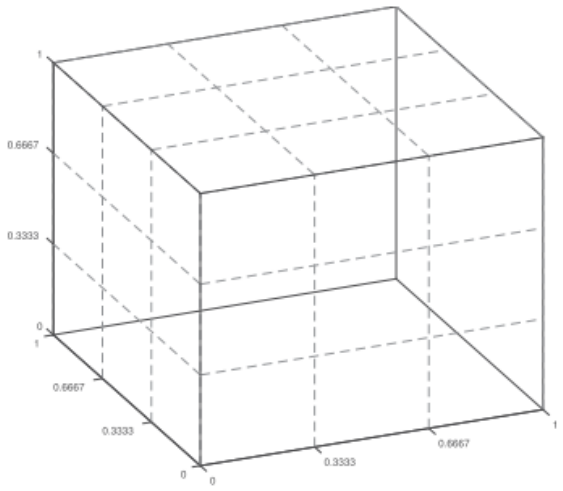


Figure 2.3: Unit hypercube of the parameter space for $n = 3$ parameters and $p = 4$ discretization levels. [51]

Elementary effects d_j are defined as

$$d_j = \frac{M\left(x_1 \dots x_{j-1}, x_j + \delta, x_{j+1} \dots x_n\right) - M\left(x_1 \dots x_n\right)}{\delta} \tag{2.6}$$

The Morris step

$$\delta = \frac{s}{p-1} \quad s \in \{1, \dots, p-1\} \tag{2.7}$$

represents the magnitude of the variation and is chosen as a multiple of the grid cell size, $\frac{1}{p-1}$. A commonly used choice of p and δ is that p is even and delta equal to $\frac{p}{2(p-1)}$. The advantage of this choice is that although the sampling strategy does not guarantee sampling with equal probability from distributions of each parameter, a certain symmetric treatment in inputs can be ensured. For example, when $p = 8$, all the possible values for a parameter ranging within [0, 1] are from $\{0, \frac{1}{7}, \frac{2}{7}, \frac{3}{7}, \frac{4}{7}, \frac{5}{7}, \frac{6}{7}, 1\}$, and the Morris step is $\frac{4}{7}$. Then there will be 4 pairs for inputs: $\{0, \frac{4}{7}\}$, $\{\frac{1}{7}, \frac{5}{7}\}$, $\{\frac{2}{7}, \frac{6}{7}\}$, $\{\frac{3}{7}, 1\}$, as a symmetric treatment in inputs.

To measure the average effect of the parameter variation on the model output, elementary effects are calculated r times for each parameter. Morris proposed two sensitivity measures, the elementary mean and standard deviation:

$$\mu_j = \frac{1}{r}\sum_{i=1}^{r} d_j^{(i)} \tag{2.8}$$

$$\sigma_j = \sqrt{\frac{1}{r-1}\sum_{i=1}^{r}\left(d_j^{(i)} - \mu_j\right)^2} \tag{2.9}$$

The absolute elementary mean is also recommended [11]:

$$\mu_j^* = \frac{1}{r}\sum_{i=1}^{r}\left|d_j^{(i)}\right| \quad j = 1, \ldots, n \tag{2.10}$$

Combing the three measurements, we can conclude the influence of parameter x_j on the model output. The interpretation is as follows.

The mean μ_j evaluates the overall influence of the factor on the output. If μ_j has a high amplitude, it implies the parameter is influential on the output. At the same time, the sign of this effect deos not vary significantly over model simulations.

However, μ_j is not effective to solve type II errors (failing to identify a factor with considerable influence on the model) [42]. Type II errors might occur when both positive and negative elements are included, i.e., when the model is nonmonotonic or has interaction effects. In these cases, some influences may cancel each other when calculating μ_j, thus resulting a low mean value even for an important factor. One solution is to check μ and $sigma$ at the same time.

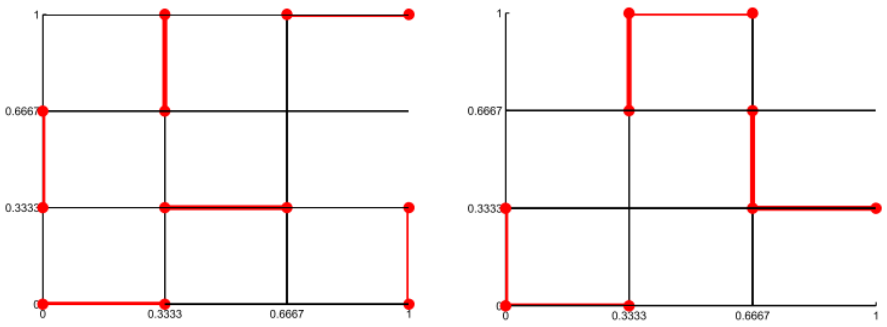
If the parameter is influential but with different signs, it will have a low value of μ but a considerable value of σ. In other words, if σ is high, then the elementary effects relative to this parameter are significantly different from each other [51]. This means that the value of x_j's elementary effects is strongly dependent on the choice of the point in the input space where it is evaluated, i.e., by the choice of the other parameter values. Therefore, it can be generally concluded that this parameter has a high interaction with other parameters. On the other hand, a low value of σ indicates nearly constant values of the elementary effect, therefore implying that the model is almost linearly dependent on x_j.

Another better way to solve the problem is to consider the absolute mean μ_j^*. Compared to μ_j, the absolute mean μ_j^* captures elementary effects of opposing sign, cancelling each other out in the calculation of μ_j. In the case that μ_j and μ_j^* are both high, the factor not only has a large effect on the output, but also the sigh of this effect is always the same. If μ_j is low while μ_j^* is high, the factor has effects of different signs depending on the point of the space at which the effect is computed.

2.3.3. Sampling strategy

In the previous sections, the core idea of sampling strategy, One-at-a-time (OAT) method has been introduced, and the sensitivity measurements has been chosen. In this section, the details of sampling in a $[0,1]^k$ input hypercube will be introduced.

The elementary measurements above require a total of $2kr$ model simulations, for example in 2.4(a) (k: the number of inputs; r: the number of elementary effects being calculated): each site of sensitivity measurements needs two sample points to do the subtraction in d_j, with k input parameters, calculated r times to do the average. By sharing endpoints on the hypercube, for example in 2.4(b), the sampling scheme can be



(a) randomly sampled elementary effects (b) elementary paths with sharing points

Figure 2.4: Efficient sampling in Morris method ($n = 2$, $p = 4$, $r = 3$, $s = 1$): random sampling results in 12 models (left); the number of samples can be reduced if some points are shared (right). [51]

simplified and more efficient, leading to $(k+1)r$ samples. Each factor is moved in a random order at a time. Then the samples connected make up the elementary paths, which can also be called trajectories. In this case, r can also be seen as the number of the trajectories.

The choices for p, r and s have a significant impact on the results of the sensitivity analysis [51]. A high number of levels p may indicate higher accuracy of sampling. A relatively low value of r, the number of sensitivity measurement calculations, will leave many of the levels in the hypercube unexplored. According to Morris [33], a convenient choice is $s = \frac{p}{2}$ (assuming p is even), while Campolongo et al. [11] have demonstrated that $p = 4$ and $r = 10$ produce valuable results in many cases.

2.3.4. Method of optimized trajectories: enhancing the sampling strategy

From the previous section, after simplifying the sampling strategy by sharing endpoints, Morris Method constructed r trajectories in the input space, which is a hypercube. However, this strategy could lead to a non-optimal coverage of the input space, especially for models with a large number of input factors [11]. In this section, an enhanced method of optimized trajectories will be introduced to improve the sampling strategy, aiming at a better scanning of the input domain without increasing the model simulations. The idea is to separate the r trajectories to maximize their dispersion in the input space.

To explain the operations, a high number of different trajectories are selected first, between 500 and 1000, and then a small number of trajectories will be chosen as the highest 'spread' scheme, for example $r = 10$. The concept of 'spread' is based on the following definition of 'distance'. The distance d_{ml} between a couple of trajectories m and l is defined by Campolongo et al [11]. as:

$$d_{ml} = \begin{cases} \sum_{i=1}^{k+1} \sum_{j=1}^{k+1} \sqrt{\sum_{z=1}^{k} \left[X_i^m(z) - X_j^l(z) \right]^2} & \text{for } m \neq l \\ 0 & \text{otherwise} \end{cases} \tag{2.11}$$

where k is the number of input factors and $X_i^m(z)$ indicates the zth coordinate of the ith point of the mth Morris trajectory. As the trajectories used as a starting sample are all different, the distance between two of them is always not zero. Nevertheless, the concept of 'zero' distance is still included to force d to fulfil all the properties of a metric. The space of all possible trajectories with the metric d is thus a metric space [11].

By maximizing the distance d_{ml}, the best r trajectories can be selected out of the starting sample M. To find the best scheme, each possible combination of r trajectories out of M are considered. The quantity D is defined, which is the sum of all the distances d_{ml} between couples of trajectories in this combination. For example, if we consider the combination 2, 4, 5, 8 (i.e. $r = 4$) out of the possible $M = \{1,2,3,4,5,6,7,8\}$, $D_{2,4,5,8}$ is calculated as $D_{2,4,5,8} = \sqrt{(d_{2,4}^2 + d_{2,5}^2 + d_{2,8}^2 + d_{4,5}^2 + d_{4,8}^2 + d_{5,8}^2)}$. After calculating D of all the possible combinations, the combination with the highest value of D will be selected out as the best r trajectories.

The enhanced method of optimized trajectories to improve the sampling strategy is to be preferred to the

original one not considering the 'distance', as it has a better scan of the input space, while keeping the number of simulations the same.

2.4. Copula-based method

2.4.1. The consideration of dependencies and correlations between parameters

Morris method is the fundamental method in sensitivity analysis, under the assumption of parameter independence. Although the correlation information can be explained roughly from sigma (see chapter 2.3.2), the independent sampling of Morris method might break the underlying assumptions in physical systems, where model parameters are usually related to each other. Thus, the explanation of correlations might be inaccurate. The correlation information refers to how related a parameter is with other parameters.

To handle dependency relationships between model parameters, copula-based method and variance-based method are applied. In this section, the concepts and theories of copula-based method will be introduced; variance-based method will be introduced in 2.5.

Copula-based method can be seen as the enhanced version of Morris method, gathering parameters into groups according to the correlations between each other. The main idea of copula-based method is to consider the correlations as the prior factors before sampling. In contrast, variance-based method explains correlations after getting the outcomes of sensitivity measurements and does not need to consider dependencies in sampling. The details will be discussed in 5.1.

2.4.2. Copula

A copula is a joint distribution, defined on an n-dimensional unit hypercube with uniform marginal distributions [34]. It separates the influence of marginal distributions from the influence of parameter dependencies [11].

The joint cumulative distribution function $F(x_1,\dots,x_k)$ of random variables $X = (x_1,\dots,x_k)$ with marginal distributions denoted as $F_i(x_i)$, $i = 1,\dots,k$ can be represented with copula C as follows,

$$F(x_1,\dots,x_k) = C(F_1(x_1),\dots,F_n(x_k)) \tag{2.12}$$

which is unique if $(x_1,\dots,x_k)$ are continuous [34].

There are different kinds of copulas, such as Gaussian, Student-t and copulas from the Archimedean family. Among them, the most popular copulas are Gaussian [11]. Ţene et al. [51] presented an example of a three dimensional Gaussian copula with correlations $\rho(x_1,x_2) = \rho(x_1,x_3) = -0.7$ and $\rho(x_2,x_3) = 0.7$ (see figure 2.5). The larger concentration of points close to the $(1,0,0)$ and $(0,1,1)$ is due to the negative correlation between the first parameter and the remaining two.

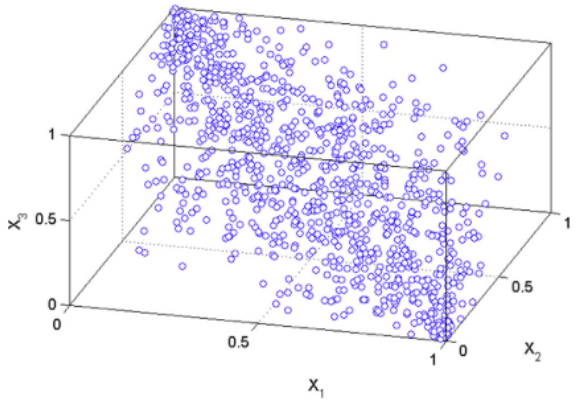
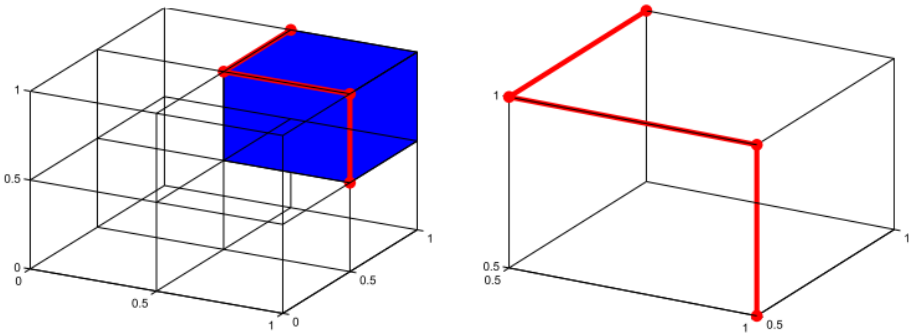
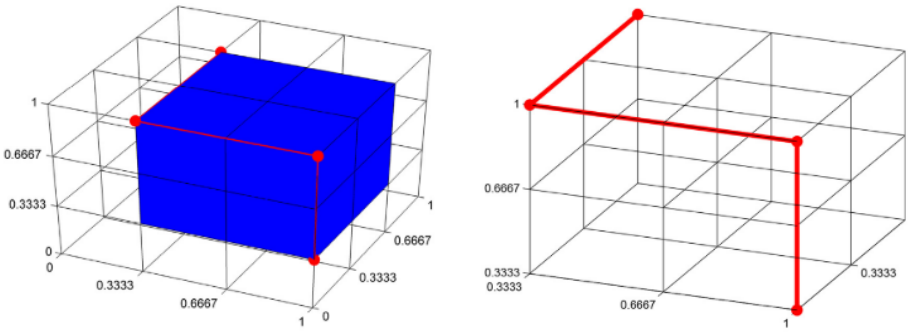


Figure 2.5: Scatter plot of 500 samples from a Gaussian copula with correlations $\rho(x_1,x_2) = \rho(x_1,x_3) = -0.7$ and $\rho(x_2,x_3) = 0.7$. [51]



(a) the whole grid (b) zoom on the cell

Figure 2.6: Geometric reinterpretation of an elementary path ($n = 3$, $p = 3$, $s = 1$). [51]



(a) the whole grid (b) zoom on the cell

Figure 2.7: Geometric reinterpretation of an elementary path ($n = 3$, $p = 4$, $s = 2$). [51]

2.4.3. Sampling strategy

Elementary effects with three sensitivity measurements μ, μ^*, σ are also used to evaluate the sensitivity of parameters in Copula-based method. The main differences between Morris method and Copula-based method are sampling strategies, where Copula-based method considers the dependence. [51] presented a sampling strategy for Copula-based method as follows.

As discussed in chapter 2.3.3, the elementary paths of Morris method are built in the hypercube. Consider the Morris step equal to one cell, i.e., $\delta = \frac{1}{p-1}$, with three parameters and three levels, then each trajectory runs on the contour of a grid cell (see figure 2.6), from one of its corners and ending in the opposite (since all coordinates change with $\pm\delta$). Consider the Morris step higher than one grid cell, trajectories run on a s×s grid cell, starting from one corner and ending at the opposite.

In figure 2.7, the trajectory runs on the contour of the blue gird cell, and the lower-right corner is chosen as the starting point. The same as Morris method, one parameter changes at a time. In this case, x_3 changes first, making the path go up, then x_1 changes, making the path go left, at last x_2 changes, the path going back one step. The order of the traversal is equivalent to choosing a permutation of the set {1,2,3}. There are 3! = 6 different ways to choose the path in the grid cell.

To conclude the example above, the sampling strategy of Copula-based method follows the three steps [51]:

1. Choosing the target grid block.
2. Choosing the starting point as one of the corners of the grid block.
3. Choosing the traversal order of the contour segments, in order to reach the opposite corner.

The total number of possible paths on the unit hypercube can be calculated:

$$N_{\text{cells}} = (p-s)^n, N_{\text{corners}} = 2^n, N_{\text{orders}} = n \tag{2.13}$$

$$N_{\text{paths}} = N_{\text{cells}} \cdot N_{\text{corners}} \cdot N_{\text{orders}} / 2 \tag{2.14}$$

where k is the number of parameters, p is the number of discretization levels, and s is the Morris step size. Through dividing the sampling strategy into three steps, now the dependence constraints can be introduced. The details of the three steps will be introduced in the next sections.

2.4.4. Choosing the target block

As illustrated in figure 2.6 and 2.7, the grid block contains the elementary paths and gives the range of parameters' values. In the grid block, the path runs from one point to the opposite one. This section introduces a method considering dependency constraint.

To choose the grid block, first the copulas should be defined to capture dependencies. As discussed in chapter 2.3.2, each parameter is first scalded to $[0,1]$ to get rid of the influences of different parameters' ranges. A copula between parameters can be extracted from the joint distribution by transforming the margins to be uniform on $[0,1]$. If the parameters are uniformly distributed over the original ranges, it can be done simply by linear scaling. Otherwise, marginal distributions need to be applied. Latin Hypercube Sampling (LHS) [30] is then performed on the copula to ensure a good coverage of the parameter space. However, to choose the grid bock, the points generated by LHS should fall in the hypercube uniformly according to the levels. To achieve this, Latin Hypercube Sampling with Dependence (LHSD) [37] is applied. The difference between LHS and LHSD is that LHSD ranks the copulas to ensure the points fall in the grid cells.

First, figure 2.8 gives a short impression of how LHS operates. By dividing the intervals according to levels in Morris method, the points can also be located in the cells (compared with figure 9.a). However, the problem is that the value will be continuous, thus the general expression of the grid blocks will be difficult to organize.

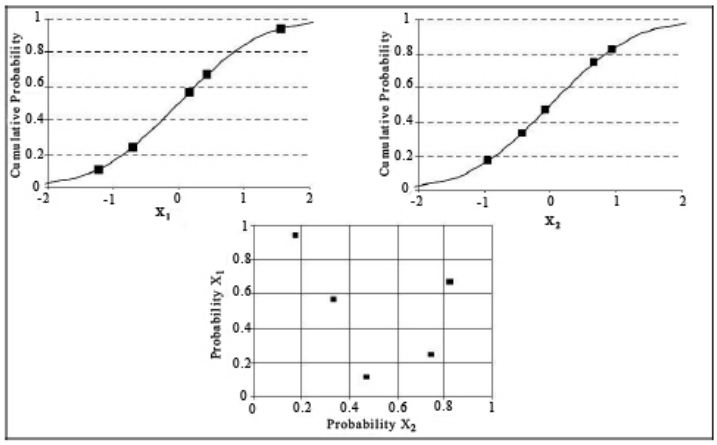


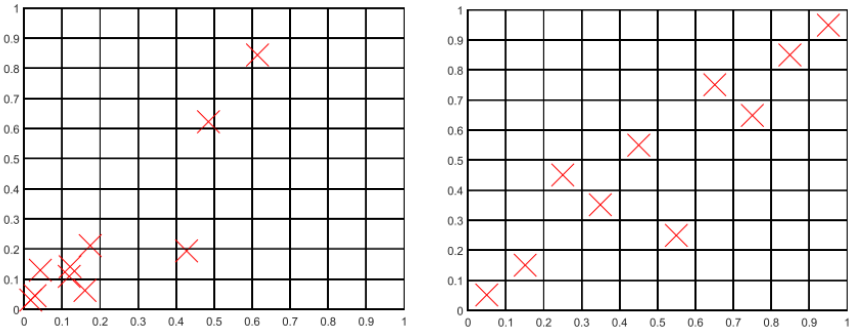
Figure 2.8: Example of LHS: Random stratified sampling of variables x_1 and x_2 at 5 intervals (top) and random painting of sampled x_1 and x_2 forming a Latin hypercube (bottom). [51]

LHSD is operated as follows. Considering a hypercube of l^n grid cells, take l samples from the copula, $u^{(1)}, u^{(2)}, \ldots, u^{(l)} \in R^n$, and arrange them to get one sample in each row and column, while preserving their ranking. The rank statistics of the ith sample of parameter j are computed as

$$R_j[i] = \sum_{k=1}^{l} 1_{\left\{u^{(k)}[j] \le u^{(0)}[j]\right\}} \quad i = 1,\ldots,l \quad j = 1,\ldots,n \tag{2.15}$$

where 1_s denotes the indicator function of set S. $R_j[i]$ effectively represents the order of the sample in $(u^{(1)}[j],\ldots,u^{(l)}[j])$. Finally, the vector containing the coordinates of the origin of the target cell (i.e. its lower-left corner) is determined as:

$$\mathbf{cell}^{(i)}[j] = \frac{R_j[i]-1}{l} \quad i = 1,\ldots,l \quad j = 1,\ldots,n \tag{2.16}$$



(a) Copula samples (b) Latin hypercube reordered samples

Figure 2.9: Using LHSD to ensure an even spread of the copula samples within the parameter space ($n = 2$, $p = 11$, $s = 1$). [51]

The cells are the target grid blocks selected. To relate to Morris method, n is the same as k in Morris method, representing the number of parameters; the number of grid blocks are the times of calculating elementary effects (which is r in Morris method, see chapter 2.3.3). The number of grid blocks can be adjusted, which is explained as follows.

Ţene et al. [51] gave the notification about the number of samples. By the nature of LHSD, the number of samples generated by LHSD, which means the grid blocks here instead of samples to calculate the elementary effects, should be a multiple the size of hypercube. However, sensitivity analysis may require an arbitrary number of samples. To meet the flexible requiments, the LHSD algorithm can be repeated several times, if the number of samples are not enough. If the number exceeds the requiement, part of the samples can be discarded.

To explain it more clearly, figure 2.9 is illustrated. Figure 2.9 gives an example with the parameter space, where $k = 2$, $p = 11$, $s = 1$. Figure 2.9(a) shows 10 copulas selects. In figure 2.9(b), the copulas are located in the cells, ensuring each row and column in each direction has one points. The grids with the red cross are the gird blocks chosen. The coordinates of the points are the lower-left coordinates of the red cross in the figure. There being 10 blocks means the elementary effects are calculated 10 times to do the average, r equal to 10.

2.4.5. Choosing the starting point

After having the grid blocks where the trajectories run on the contour, the starting points should be decided. The ending points will automatically be the opposite point, as each parameter needs to change once in each dimension within the grid block.

For each grid block, $\mathbf{cell}^{(i)}$, the starting corner of the path is randomly sampled. To choose the starting point, the probability of each point being selected will be calculated. The basic idea is that each corner in the grid is a realization of an n-dimensional discrete distribution with p possible values, namely $0, 1, 2, \ldots, p-1$, for each factor.

After marginal transformation or linear scaling, the marginal distributions of the factors are removed. The marginal probability of each factor xi as value $j \in 0, 1, 2, \ldots, p-1$ is $Prob(x_i = j) = \frac{1}{p}$. Using the finite difference formula [34], the probability of each point being selected on the grid can be computed,

$$\text{Prob}\left(X_1 = j_1, \ldots, X_n = j_n\right) = \Delta_{a_1}^{b_1} \Delta_{a_2}^{b_2} \ldots \Delta_{a_n}^{b_n} C \tag{2.17}$$

where $a_i = \frac{j_i}{p}$, $b_i = \frac{j_i+1}{p}$ and

$$\Delta_{a_k}^{b_k} C = C\left(u_1, \ldots, u_{k-1}, b_k, u_{k+1}, \ldots, u_n\right) \tag{2.18}$$

$$- C\left(u_1, \ldots, u_{k-1}, a_k, u_{k+1}, \ldots, u_n\right) \tag{2.19}$$

Hence, when $n = 3$,

$$\begin{aligned}\text{Prob}\left(X_1 = x_1, X_2 = x_2, X_3 = x_3\right) \\ =& C\left(b_1, b_2, b_3\right) - C\left(a_1, b_2, b_3\right) - C\left(b_1, a_2, b_3\right) \\ &+ C\left(a_1, a_2, b_3\right) - C\left(b_1, b_2, a_3\right) + C\left(a_1, b_2, a_3\right) \\ &+ C\left(b_1, a_2, a_3\right) - C\left(a_1, a_2, a_3\right)\end{aligned} \tag{2.20}$$

Take $n = 3$ and $p = 2$ as an example. In this case, the hypercube contains only one cell with eight corners. Each factor can take only two possible values $x_i = 0, 1$. $P_{0,0,0} = Prob(X_1 = 0, X_2 = 0, X_3 = 0) = C(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$, with $a_i = 0$, $b_i = \frac{1}{2}$, $i = 1, 2, 3$, since any copula evaluated at point zero is 0. Using a normal copula with the correlation matrix in Section 2.4.2 ($\rho(x_1, x_2) = \rho(x_1, x_3) = -0.7$ and $\rho(x_2, x_3) = 0.7$). $P_{0,0,0} = 0.0633$. Similarly, $p_{0,1,0} = p_{0,0,1} = p_{1,1,0} = p_{1,0,1} = p_{1,1,1} = 0.0633$. However, the probability of the point $(0, 0, 1)$ and $(1, 1, 0)$ is much higher with $p_{1,0,0} = C(1, \frac{1}{2}, \frac{1}{2})$–$C(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}) = 0.3101$, with $a_i = 0$, $b_i = \frac{1}{2}$, $i = 2, 3$ and $a_1 = \frac{1}{2}$, $b_1 = 1$. After having the distribution as calculated above, the starting point is sampled. In the case above, there is a much larger chance to choose point $(1, 0, 0)$ or $(0, 1, 1)$ as starting points over the other points.

This procedure can be applied to compute the distribution of each corner as starting points for each other grid in the hypercube. However, it needs a lot of calculations when p is large. It is suggested that one can compute the distribution only once and assumed that it applies to all grid cells [51]. The simplified procedure would be sufficient for Gaussian copula; for more complicated copulas such as copulas with asymmetries and tail dependencies, the assumption would not hold.

The ending point is always the opposite corner of the starting point, as each parameter should move one step in the grid block to calculate the elementary effects. The number of all possible paths between the starting and the ending points are 2^n. The choice of the path will be explained in the following section.

2.4.6. Choosing the traversal order

After deciding the starting point, along with the ending point, the differences between paths can be understood as the differences of the traversal orders. The order of traversal is given by a randomly sampled permutation π^i describing the path running from the start point to the end point by changing one factor at one time. For example, the path in figure 2.7 was obtained using the permutation {3, 1, 2}. It means the third parameter changes first, then the first parameter changes, finally the second parameter changes.

$$\begin{pmatrix} x & y & z \\ 1 & 0 & \mathbf{0} \\ \mathbf{1} & 0 & \mathbf{1} \\ \mathbf{0} & \mathbf{0} & 1 \\ 0 & \mathbf{1} & 1 \end{pmatrix} \tag{2.21}$$

2.4.7. Sensitivity measurements

After getting the samples from the hypercube, sensitivity measurements should be defined. The measurements in Copula-based method are the same as in Morris method, using μ, μ^* and σ to indicate the influences of parameters on outputs, the details can be seen in chapter 2.3.2.

2.5. Variance-based method

2.5.1. General idea of using variance

The same as copula-based method, variance-based method is also used to deal with the dependence problems. However, the way of dealing with the dependency problem is totally different between copula-based method and variance-based method. In copula-based method, copulas are used to contain the dependency information, influencing the sample generation. The dependency can be seen as the prior information to evaluate the sensitivity. While in variance-based method, we don't need to consider dependency before calculating the sensitivity measurements. Through a lot of samples and variance, the correlation information

can be interpreted from the sensitivity measurements. Although we can do the same in Morris method, the nature of Morris method, assuming each parameter being independent, will lead to bias in the results.

Why do we choose variance? Speaking of variance, we naturally think of using mean to calculate sensitivity measurements. In fact, mean is used in some cases, for example in risk analysis the model output may happen to be itself a mean [42]. In the case, the outputs are deterministic, and risk level is to be modified by eliminating the uncertainty in the input. However, the elimination is not suggested, and the uncertainty is preferred to be maintained. Then the variance is naturally considered to keep the uncertainty information.

As variance requires a large number of samples, the sensitivity measurement is not constrained by the form of models. In other words, the parameters in the model can be dependent. However, it also means the computational tasks of variance-based method is heavy.

2.5.2. The setting of the model and the first-order sensitivity index

The Russian mathematician I. M. Sobol' provided a concept to compute sensitivity measurements for arbitrary groups of factors. The setting is based on a square integrable function f over Ω^k, the k-dimensional unit hypercube,

$$\Omega^k = \left(X \mid 0 \leq x_i \leq 1; i = 1, \ldots, k\right) \tag{2.22}$$

The function f is expanded into terms of increasing dimensions:

$$f = f_0 + \sum_i f_i + \sum_i \sum_{j>i} f_{ij} + \ldots + f_{12\ldots k} \tag{2.23}$$

in which each individual term is also square integrable over the domain. Each term is a function of the factors in its index, i.e., $f_i = f_i(X_i)$, $f_{i,j} = f_{i,j}(X_i, X_j)$, and so on. The decomposition has a finite number of $2k$ terms. Among them, one is constant (f_0), k are first-order functions (f_i), $C_k^2 = \frac{k!}{(k-2)!2!}$ are second-order functions (f_{ij}), and so on. The expansion is called high-dimensional model representation (HDMR) and is not unique. There can be infinite choices for the terms of a given f.

Sobol' proved that if each term in the expansion has zero mean, i.e., $\int f(x_i)dx_i = 0$, then all the terms of the decomposition are orthogonal in pairs, i.e. $\int f(x_i)f(x_j)dx_i dx_j = 0$. Thus, using the conditional expectations of the model output Y, the terms can be calculated. For example,

$$f_0 = E(Y) \tag{2.24}$$

$$f_i = E\left(Y \mid X_i\right) - E(Y) \tag{2.25}$$

$$f_{ij} = E\left(Y \mid X_i, X_j\right) - f_i - f_i - E(Y) \tag{2.26}$$

The variance of the conditional expectation can be considered to measure the sensitivity. As f_0 is constant, $V(f_i(X_i))$ is $V[E(Y|X_i)]$. $V[E(Y|X_i)]$ can be denoted by V_i, which is called the first-order effect of X_i on Y. By dividing this by the unconditional variance $V(Y)$, the first-order sensitivity index is defined:

$$S_i = \frac{V\left[E\left(Y \mid X_i\right)\right]}{V(Y)} \tag{2.27}$$

The first-order index evaluates the main effect contribution of each input factor to the variance of the output, which is also described as an 'importance measure'.

2.5.3. Internal effects

Two factors are said to interact when their effect on outputs cannot be expressed as a sum of their single effects [42]. Similarly, interactions between more factors can be defined. The extreme values of outputs can be uniquely associated with combinations of inputs, which cannot be described by the first-order effects. In other words, dependency information cannot be interpreted from the first-order index. To investigate the interaction effects between parameters, higher-order effects should be calculated.

From the formulas above, the following equations can be obtained:

$$V_i = V\left(f_i\left(X_i\right)\right) = V\left[E\left(Y \mid X_i\right)\right] \tag{2.28}$$

$$V_{ij} = V\left(f_{ij}\left(X_i, X_j\right)\right) = V\left(E\left(Y \mid X_i, X_j\right)\right) - V\left(E\left(Y \mid X_i\right)\right) - V\left(E\left(Y \mid X_j\right)\right) \tag{2.29}$$

In the equation 2.29, $V\left(E\left(Y \mid X_i, X_j\right)\right)$ describes the joint effect of the pair (X_i, X_j) on Y, which will be denoted by V_{ij}^c. The term V_{ij} is the joint effect of X_i and X_j, V_{ij}^c, minus the first-order effects of X_i and X_j. $V(f_{ij})$ is known as a second-order, or two-way, effect. Similar formulas can be written for higher-order effects.

By simplifying the notations of the variance, i.e., $V(f_i) = V_i$, $V(f_{ij}) = V_{ij}$, etc., and by square integrating each term in the formula 2.23, the ANOVA-HDMR decomposition can be obtained:

$$V(Y) = \sum_i V_i + \sum_i \sum_{j>i} V_{ij} + \ldots + V_{12\ldots k} \tag{2.30}$$

Dividing both sides of the equation by the unconditional variance $V(Y)$, we have

$$\sum_i S_i + \sum_i \sum_{j>i} S_{ij} + \sum_i \sum_{j>i} \sum_{l>i} S_{ijl} + \ldots + S_{123\ldots k} = 1 \tag{2.31}$$

S_{ij}, S_{ijl}, etc., are the higher-order sensitivity index, indicating the interaction between factors.

2.5.4. Total effects

As illustrated in formula 2.31, S123…k is the total effect index, accounting for the total contribution to the output variance due to factor X_i, i.e., its first-order effect plus higher-order effects resulting from interactions with other factors.

Take a three-factor model as an example, the total effect index has the following expression:

$$S_{T1} = S_1 + S_{12} + S_{13} + S_{123} \tag{2.32}$$

where all the terms involving the factor X_1 are considered. Including the first-order sensitivity index S_1, the second-order sensitivity index S_{12}, S_{13}, and the third-order sensitivity index S_{123}.

We can calculate the total effect index according to 2.31 by computing the difference between all the terms and 1, but the computation will be a heavy task. Here a technique is introduced to calculate the total indices, demanding only the same cost of first-order indices [21]. First the unconditional variance can be decomposed into main effect and residual:

$$V(Y) = V\left(E\left(Y \mid X_i\right)\right) + E\left(V\left(Y \mid X_i\right)\right) \tag{2.33}$$

Another way to find the total index is to decompose $V(Y)$ conditioning with respect to all the factors but one, i.e. $X_{\sim i}$:

$$V(Y) = V\left(E\left(Y \mid \mathbf{X}_{\sim i}\right)\right) + E\left(V\left(Y \mid \mathbf{X}_{\sim i}\right)\right) \tag{2.34}$$

The measure $V(Y) - V\left(E\left(Y \mid \mathbf{X}_{\sim i}\right)\right) = E\left(V\left(Y \mid \mathbf{X}_{\sim i}\right)\right)$ is the remaining variance of Y that would be left, on average, if the true values of $X_{\sim i}$ can be determined. The average is calculated over all possible combinations of $X_{\sim i}$, since $X_{\sim i}$ are uncertain factors and the 'true values' are unknown. The total effect index for X_i is obtained, dividing by $V(Y)$:

$$S_{T_i} = \frac{E\left[V\left(Y \mid \mathbf{X}_{\sim i}\right)\right]}{V(Y)} = 1 - \frac{V\left[E\left(Y \mid \mathbf{X}_{\sim i}\right)\right]}{V(Y)} \tag{2.35}$$

Total effect index is meaningful to sensitivity analysis. Total effect index gives information on the nonfeatures of the model. According to formula 2.31, for a given factor X_i, a significant difference between S_i and S_{Ti} shows important interaction involving this factor. Total effect index can also give the answer that which factor can be fixed anywhere over the range without affecting the output. $S_{Ti} = 0$ is necessary and sufficient for X_i to be noninfluential on the output. If $S_{Ti} \simeq 0$, then Xi can be fixed at any value within its range, without obviously

affecting the value of the output variance $V(Y)$. Total effect indices are suitable for factor fixing setting (FF), which is used to identify which parameters have no significant contribution to the variance of the outputs. The name of the setting can be understood from the question 'which factor can be fixed anywhere over its range of variability without affecting the output?' [42]

2.5.5. Computation strategy of the sensitivity indices

In the last three sections, the formulas of sensitivity indices are introduced. However, the computation for the conditional variance such as $V(E(Y|X_i))$ and $V(E(Y|X_i,X_j))$ are very heavy. For example, to calculate $V(E(Y|X_i))$, a set of 1000 Monte Carlo points are first used to calculate the conditional mean $E(Y|X_i)$. Then the procedure needs to be repeated 1000 times to estimate the variance. In this section, a simpler strategy introduced by [42] will be described, based on the same formulas above. First, a $(N,2k)$ matrix of random numbers (k is the number of inputs) and two matrices A and B are defined to contain half of the sample each. N is called a base sample, varying from a few hundred to a few thousands. Sequences of quasi-random numbers are suggested by Sobol' [45].

$$A=\begin{bmatrix} x_1^{(1)} & x_2^{(1)} & \dots & x_i^{(1)} & \dots & x_k^{(1)} \\ x_1^{(2)} & x_2^{(2)} & \dots & x_i^{(2)} & \dots & x_k^{(2)} \\ \cdots & \dots & \dots & \dots & & \\ x_1^{(N-1)} & x_2^{(N-1)} & \dots & x_i^{(N-1)} & \dots & x_k^{(N-1)} \\ x_1^{(N)} & x_2^{(N)} & \dots & x_i^{(N)} & \dots & x_k^{(N)} \end{bmatrix} \tag{2.36}$$

$$B=\begin{bmatrix} x_{k+1}^{(1)} & x_{k+2}^{(1)} & \dots & x_{k+i}^{(1)} & \dots & x_{2k}^{(1)} \\ x_{k+1}^{(2)} & x_{k+2}^{(2)} & \dots & x_{k+i}^{(2)} & \dots & x_{2k}^{(2)} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ x_{k+1}^{(N-1)} & x_{k+2}^{(N-1)} & \dots & x_{k+i}^{(N-1)} & \dots & x_{2k}^{(N-1)} \\ x_{k+1}^{(N)} & x_{k+2}^{(N)} & \dots & x_{k+i}^{(N)} & \dots & x_{2k}^{(N)} \end{bmatrix} \tag{2.37}$$

Then define a matrix C_i. In C_i, all the columns are kept the same with matrix B, except the ith column, which is the same with the ith column of matrix A.

$$C_i=\begin{bmatrix} x_{k+1}^{(1)} & x_{k+2}^{(1)} & \dots & x_i^{(1)} & \dots & x_{2k}^{(1)} \\ x_{k+1}^{(2)} & x_{k+2}^{(2)} & \dots & x_i^{(2)} & \dots & x_{2k}^{(2)} \\ \cdots & \dots & \dots & \dots & \dots & \dots \\ x_{k+1}^{(N-1)} & x_{k+2}^{(N-1)} & \dots & x_i^{(N-1)} & \dots & x_{2k}^{(N-1)} \\ x_{k+1}^{(N)} & x_{k+2}^{(N)} & \dots & x_i^{(N)} & \dots & x_{2k}^{(N)} \end{bmatrix} \tag{2.38}$$

The model outputs can be simulated based on the sample matrices A, B and C_i. The output vectors of dimension $N\times 1$ are demoted as

$$y_A=f(A)\quad y_B=f(B)\quad y_{C_i}=f\left(C_i\right) \tag{2.39}$$

Having all the results, the first- and total-effect indices S_i and S_{Ti} for the vector X_i can be obtained. The first-order sensitivity indices are estimated as follows:

$$S_i=\frac{V\left[E\left(Y\mid X_i\right)\right]}{V(Y)}=\frac{y_A\cdot y_{C_i}-f_0^2}{y_A\cdot y_A-f_0^2}=\frac{(1/N)\sum_{j=1}^{N}y_A^{(j)}y_{C_i}^{(j)}-f_0^2}{(1/N)\sum_{i=1}^{N}\left(y_A^{(j)}\right)^2-f_0^2} \tag{2.40}$$

where

$$f_0^2=\left(\frac{1}{N}\sum_{j=1}^{N}y_A^{(j)}\right)^2 \tag{2.41}$$

is the mean, and the symbol $(\cdot)$ denotes the scalar product of two vectors. The total-effect indices are estimated as follows:

$$S_{T_i}=1-\frac{V\left[E\left(Y\mid \mathbf{X}_{\sim i}\right)\right]}{V(Y)}=1-\frac{y_B\cdot y_{C_i}-f_0^2}{y_A\cdot y_A-f_0^2}=1-\frac{(1/N)\sum_{j=1}^{N}y_B^{(j)}y_{C_i}^{(j)}-f_0^2}{(1/N)\sum_{j=1}^{N}\left(y_A^{(j)}\right)^2-f_0^2} \tag{2.42}$$

The total computational cost of this strategy is much lower. For A and B, it cost $N+N$ runs of the models; for C_i, as there are k inputs, it needs $k \times N$ simulations. In total the cost will be $N(k+2)$, while the strategy in the first paragraph needs N^2 simulations.

3

Methodology

3.1. The hydrodynamic model and the software used for simulations

3.1.1. The hydrodynamic model: 3D DCSM-FM model

The sensitivity analysis is conducted in the three-dimensional hydrodynamic model called the 3D Dutch Continental Shelf Model developed in D-HYDRO Flexible Mesh (3D DCSM-FM). The model is originally setup as part of Deltares' strategic research funding and has been used for numerous studies. It is developed in D-HYDRO Flexible Mesh and is based on the horizontal schematization of the 2D DCFM-SM model [57]. 3D DCSM-FM can simulate the hydrodynamic process within the Northwest European Shelf, including the North Sea and adjacent shallow seas and estuaries such as the Wadden Sea and the Eastern and Western Scheldt. Specifically, it covers the area between 15° W to 13° E and 43° N to 64° N. The model divides the area into the computational gird. The network has a resolution that increases with decreasing water depth, and the shapes of the grid can change with the area. The areas of refinement were specified with smooth polygons; areas with different resolution are connected with triangles. Also, the cell size can vary with the square root of the depth. In the case of this thesis project, the grids are uniformly polygons, and the grid size is chosen as 0.5 nm. The model has been validated with respect to water levels, sea surface temperature, sea surface salinity gradients and current velocities in different areas [57].

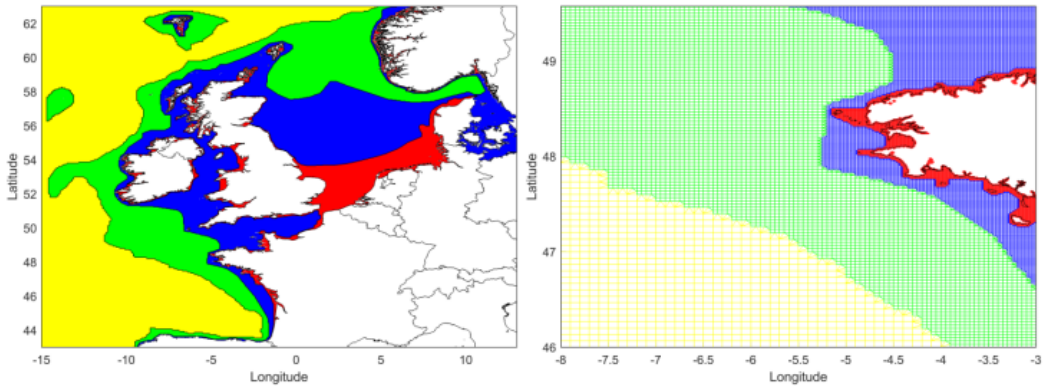


Figure 3.1: Overview (left) and (right) of the DCSM-FM model network with the colours indicating the grid size (yellow: ~ 4 nm; green: ~ 2 nm; blue: ~ 1 nm; red: ~ 0.5 nm). [57]

3.1.2. The simulation software: Delft3D FM

The software Delft3D Flexible Mesh Suite (Delft3D FM) is used in this thesis project to run the model. Specifically, the hydrodynamic module D-Flow Flexible Mesh (D-Flow FM) is used, which is part of Delft3D FM. The

D-Flow FM module is built with use of the Delta Shell framework. There are also other modules compliant with the framework, for example, D-Waves, D-Water Quality, D-Real Time Control and D-Rainfall Runoff.

D-Flow FM can simulate the process of hydrodynamic flow, waves, water quality and ecology. It is a multi-dimensional (1D, 2D and 3D) hydrodynamic (and transport) simulation program which calculated non-steady flow and transport phenomena that results from tidal and meteorological forcing on structured and unstructured, boundary fitted grids. The term Flexible Mesh in the name refers to the flexible combination of unstructured grids consisting of triangles, quadrangles, pentagons, and hexagons [15].

3.1.3. The files and operation process of the model in the thesis project

Input files are needed to set up the hydrodynamic model. All parameters originate from the physical process.

All the input data required in the physical phenomenon and simulation process is collected in the Master Definition Unstructured file, called a mdu-file. Also, some attribute files where relevant data is stored can be defined and referred to in the mdu-file, such as ext-file containing external forcing files, net.nc-files containing grid information, pli-file containing boundary condition location information, etc. The mdu-file mainly has the following sections: [General] containing the name and version, [geometry], [numerics], [physics], [wind], [time], [restart], [external forcing] and [output].

In this thesis project, all the changes of parameters occur in mdu-file, where the parameters are constant values. The attribute files keep the same. Instead of running the models locally by using GUI, the models are uploaded in the cluster in Linux to run, after only changing the parameters in the mdu-files.

The results are exported in the his- and/or map-files. In the his-file, outputs' results of different layers and locations at various timepoints are stored.

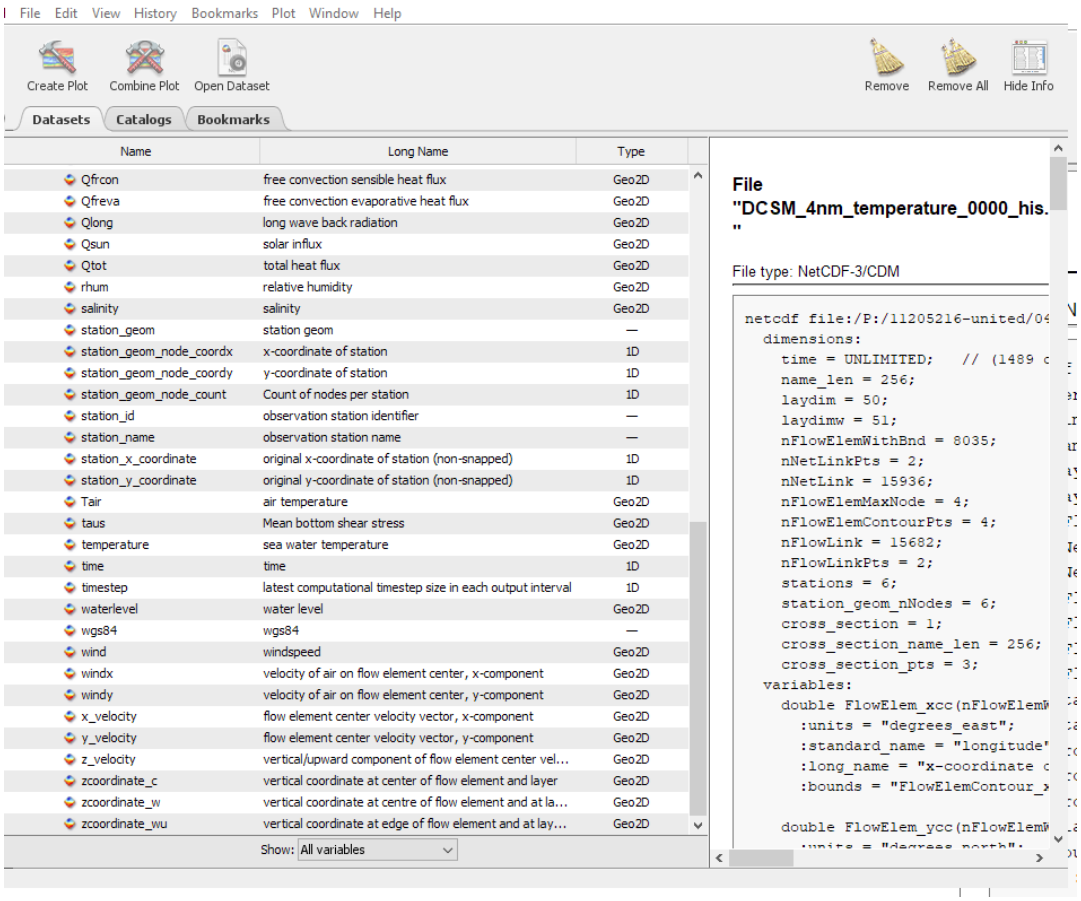


Figure 3.2: The example of his-file opened in Panoply, showing part of the outputs.

3.2. The choices of observation points

3D DCSM-FM can give the simulation results within a large area. However, we would like to know the sensitivity of parameters at some specific points. Thus, some observation points are selected properly to do the sensitivity analysis in this section.

As introduced in chapter 1, the German pilot tests the potential for the cultivation of blue mussels, or Mytilus edulis, and the seaweeds, or specifically Saccharina latissima (which is also called sugar kelp), in combination with wind energy production. The experiment is conducted at a platform called FINO3. In January 2002, three research platforms named FINO were decided to be established in three potential suitable areas in the North and Baltic Seas for the demands of larger offshore wind farms. In 2009, the construction of FINO3 was completed and the platform has been operated since then by the Research & Development (R&D) Centre Kiel University of Applied Science (FuE-Zentrum FH Kiel GmbH) [2].

Before the construction, a variety of meteorological, hydrological and soil investigation were conducted to decide the best location. The FINO3 research platform stands 80 km off Sylt, in the midst of German offshore wind farms Butendiek, DanTysk and Sandbank [2].

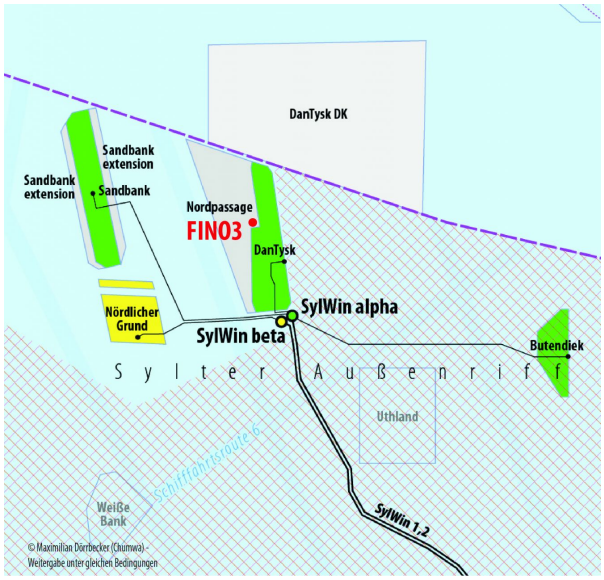
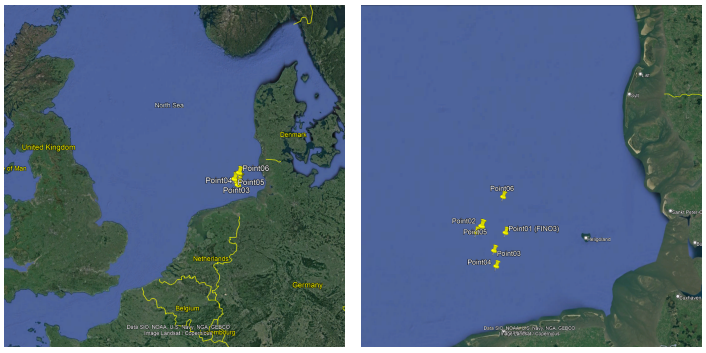


Figure 3.3: The FINO3 research platform stands 80 km off Sylt, in the midst of German offshore wind farms Butendiek, DanTysk and Sandbank. [2]



(a) Locations of observation points (b) Observation points-zoomed in

Figure 3.4: The six observation points are located 80 km off Sylt.

The FINO3 platform is chosen as an observation point. As comparisons of the point, several other observation points around FINO3 are also selected in the model to conduct sensitivity analysis (see figure above). They

are in the different directions of FINO3, also in the different locations towards the wind farms, considering the influences of the wind farms on the water currents. The exact coordinates of the six observation points are displayed in the following table:

Table 3.1: Locations of six observation points.

Observation points	Latitude	Longitude
Point01 (FINO3)	54°11'34.52"N	7° 9'45.66"E
Point02	54°13'28.83"N	6°56'56.73"E
Point03	54° 6'3.41"N	7° 3'27.55"E
Point04	54° 1'12.28"N	7° 4'29.02"E
Point05	54°11'46.73"N	6°54'6.66"E
Point06	54°22'43.67"N	7° 8'50.74"E

3.3. The choices of outputs and inputs

In the simulations of the model are a lot of parameters and outputs. Two outputs relative to mussel and seaweed cultivation are chosen in this thesis project. Based on the output variables, specific inputs related to the outputs are selected, as described below.

3.3.1. Choices of outputs

Influential factors of blue mussels and seaweeds cultivation are discussed in chapter 2.1.2 and 2.1.3. Current velocity and temperature are main factors to both blue mussels and seaweeds. Nutrients are also important to both, but this variable is covered by the water quality model and does not appear in the hydrodynamic model.

As a result, current velocity and temperature are selected in 3D DCSM-FM model as outputs in sensitivity analysis, which largely influence the growth of blue mussels and seaweeds.

3.3.2. Choices of inputs

A large number of parameters exist in the model. In this thesis project, the parameters of the model are considered as 'inputs', and the inputs in sensitivity analysis are all selected from mdu-file (chapter 3.1.3). The inputs corresponding to temperature and current velocity are illustrated as follows, which are confirmed by the scientists from Deltares.

Table 3.2: Ranges of inputs and corresponding outputs.

Parameters	Explanations	Corresponding outputs	Ranges		
			Baseline	Min	Max
Vicouv	Uniform horizontal eddy viscosity (m^2/s)	Temperature, current velocity	0,1	0,1	2
Dicouv	Uniform horizontal eddy diffusivity (m^2/s)	Temperature, current velocity	0,1	0,1	2
Vicoww	Uniform vertical eddy viscosity (m^2/s)	Temperature, current velocity	0,00005	0,000001	0,0001
Dicoww	Uniform vertical eddy diffusivity (m^2/s)	Temperature, current velocity	0,00002	0,000001	0,0001
Smagorinsky	Smagorinsky factor in horizontal turbulence	Temperature, current velocity	0,2	0,05	0,3
Stanton	Convective heat flux	Temperature	0,0013	0,001	0,0016
Dalton	Evaporative heat flux	Temperature	0,0013	0,001	0,0016
Rhoair	Air density (kg/m3)	Current velocity	1,2265	1,1639	1,3669

In the details, some definitions in the explanations are listed below:

- Eddy viscosity is the proportionality factor describing the turbulent transfer of energy as a result of moving eddies, giving rise to tangential stresses [23].
- Eddy diffusivity is the coefficient describing any diffusion process by which substances are mixed in the atmosphere or in any fluid system due to eddy motion [32].
- The Smagorinsky model [44] is considered as the pioneer subgrid-scale model for Large Eddy Simulation (LES), which is a method less computationally expensive than direct flow simulation that can give a high-resolution transient solution for an aerodynamic problem [56]. Smagorinsky Constant is used to calculate the eddy viscosity in the model.
- Convection (or convective heat transfer) is the transfer of heat from one place to another due to the movement of fluid.
- Latent heat flux (He) comprises energy lost (gained) by the stream during evaporation (condensation) as water moves from a higher to lower energy state (or vice versa) [50].

In this table, the baseline value are the default values in the mdu-file; the minimum and the maximum of the ranges are based on it, suggested by the experts from Deltares. And all the parameters are assumed to distribute uniformly within the ranges. According to the table, seven inputs are selected for temperature, and six inputs are for current velocity, and among them five parameters are overlapped.

It is worth noting that the minimum values of the parameters are set as zero at first except Stanton, Dalton and Rhoair. However, some simulations crashed due to the value of zero. The possible explanations can be found in chapter 5.1. After changing zero into the values very close to zero, all simulations can run successfully. In chapter 3.4, an example solution of results with crashed files partly will be introduced.

3.4. Morris method

In the 2, the theories of different methods are introduced, mainly including the concepts, sampling strategies and sensitivity measurements. In the methodology part, the sections of the three methods will be organized in the similar structures, but this part will focus on the real operations in the thesis project, along with the tools used and solutions of problems.

3.4.1. Sample generation

According to chapter 2.3, in a hypercube, several settings decide the sampling strategy of Morris method: number of levels, Morris step size, number of trajectories and number of factors. The number of samples are decided by the number of trajectories and number of factors.

The settings for Morris method are as follows:

Table 3.3: Settings in Morris method.

Number of levels (p)	8
Morris step size ($\delta = p/[2(p-1)]$)	4/7
Number of trajectories (r)	10
Number of factors (k)	Temperature: 7
	Current velocity: 6
Number of samples $(k+1)*r$	Temperature: (7+1)*10=80
	Current velocity: (6+1)*10=70

To generate the samples for Morris method according to the ranges in chapter 3.3.2, the Elementary Effects (EE) Sampling Package is used, which is developed by Drs. Yogesh Khare and Rafael Muñoz-Carpena [24]. The main function of this package, 'Fac_Sampler.m', generates input factor samples into a unit hypercube, and transforms them properly according to the input distributions.

To use this sampling package, the '.fac' file is compulsory in the codes. The file contains the following information: (a) the number of input factors; (b) default distribution truncation values; (c) distribution type and distribution characteristics for each input factor. The '.fac' file can be generated from SimLab v2.2.1 [42], which can be downloaded from the references part in the Sampling Package page [24]. The manual can also be found in this page. Truncated distributions perform better to get accurate results consistent with variance-based method, when parameter distributions have long tails (Normal, LogNormal, Weibull, Gamma, Exponential). Different distributions of parameters can be generated via SimLab, and in this thesis project, all the parameters are assumed as uniformly distributed within the ranges. The detailed examples are in the appendix.

Besides the '.fac' file, the main function 'Fac_Sampler.m' asks for other setting in the table, along with the sampling strategies. Five sampling strategies can be chosen in this package: (a) the method of Optimized Trajectories [OT] [11]; (b) the Modified Optimized Trajectories [MOT] [41]; (c) Sampling for Uniformity [SU] [26]; and (d) Enhanced Sampling for Uniformity [eSU] [12]; and (e) RadialeSU/ReSU [12]. Among them the method of Optimized Trajectories [OT] [11] introduced in chapter 2.3.4 is used in this thesis project. The samples can be generated in the form of 'Excel' or 'Text'.

In the following sections, including the copula-based method and variance-based method, the two systems, temperature and current velocity will be discussed about separately.

temperature

First, part of the samples with ranges starting from zero are shown in the table.

Table 3.4: Samples simulated for temperature in Morris method, ranging from zero.

Vicouv	Dicouv	Vicoww	Dicoww	Smagorinsky	Stanton	Dalton
1,142857	2	4,29E-05	4,29E-05	0,3	0,001171	0,001343
1,142857	2	4,29E-05	4,29E-05	0,3	0,001171	0,001
0	2	4,29E-05	4,29E-05	0,3	0,001171	0,001
0	2	4,29E-05	0,0001	0,3	0,001171	0,001
0	2	4,29E-05	0,0001	0,3	0,001514	0,001
0	0,857143	4,29E-05	0,0001	0,3	0,001514	0,001
0	0,857143	0,0001	0,0001	0,3	0,001514	0,001
0	0,857143	0,0001	0,0001	0,128571	0,001514	0,001
2	1,714286	5,71E-05	2,86E-05	0	0,001514	0,0016
2	1,714286	0	2,86E-05	0	0,001514	0,0016
2	0,571429	0	2,86E-05	0	0,001514	0,0016
2	0,571429	0	8,57E-05	0	0,001514	0,0016
2	0,571429	0	8,57E-05	0	0,001171	0,0016
2	0,571429	0	8,57E-05	0	0,001171	0,001257
0,857143	0,571429	0	8,57E-05	0	0,001171	0,001257
0,857143	0,571429	0	8,57E-05	0,171429	0,001171	0,001257
0	1,142857	4,29E-05	7,14E-05	0,128571	0,001257	0,0016
0	1,142857	0,0001	7,14E-05	0,128571	0,001257	0,0016
0	0	0,0001	7,14E-05	0,128571	0,001257	0,0016
0	0	0,0001	7,14E-05	0,3	0,001257	0,0016
0	0	0,0001	1,43E-05	0,3	0,001257	0,0016
1,142857	0	0,0001	1,43E-05	0,3	0,001257	0,0016
1,142857	0	0,0001	1,43E-05	0,3	0,0016	0,0016
1,142857	0	0,0001	1,43E-05	0,3	0,0016	0,001257

The ranges starting from 0 will result in a problem in simulations. The problem and one possible solution with these ranges will be discussed in chapter 3.4.3. However, to keep the consistency and avoid different possible troubles in following steps, the ranges will be changed to non-zero for all the systems in this thesis project. The new ranges are shown in chapter 3.3.2. The following table gives part of the new samples.

In the table 3.5, each eight lines form a trajectory. Within each eight lines, one parameter only changes once.

Table 3.5: Samples simulated for temperature in Morris method.

Vicouv	Dicouv	Vicoww	Dicoww	Smagorinsky	Stanton	Dalton
0,642857	0,1	8,59E-05	0,0001	0,192857	0,001343	0,001343
0,642857	0,1	8,59E-05	0,0001	0,192857	0,001	0,001343
0,642857	0,1	8,59E-05	0,0001	0,192857	0,001	0,001
1,728571	0,1	8,59E-05	0,0001	0,192857	0,001	0,001
1,728571	0,1	2,93E-05	0,0001	0,192857	0,001	0,001
1,728571	0,1	2,93E-05	0,0001	0,05	0,001	0,001
1,728571	0,1	2,93E-05	4,34E-05	0,05	0,001	0,001
1,728571	1,185714	2,93E-05	4,34E-05	0,05	0,001	0,001
2	2	2,93E-05	5,76E-05	0,157143	0,0016	0,001171
2	2	2,93E-05	0,000001	0,157143	0,0016	0,001171
2	2	2,93E-05	0,000001	0,3	0,0016	0,001171
2	2	2,93E-05	0,000001	0,3	0,0016	0,001514
2	2	8,59E-05	0,000001	0,3	0,0016	0,001514
2	2	8,59E-05	0,000001	0,3	0,001257	0,001514
2	0,914286	8,59E-05	0,000001	0,3	0,001257	0,001514
0,914286	0,914286	8,59E-05	0,000001	0,3	0,001257	0,001514
0,1	1,457143	1,51E-05	0,0001	0,05	0,001257	0,001343
0,1	1,457143	1,51E-05	0,0001	0,05	0,001257	0,001
0,1	1,457143	7,17E-05	0,0001	0,05	0,001257	0,001
0,1	1,457143	7,17E-05	0,0001	0,05	0,0016	0,001
0,1	0,371429	7,17E-05	0,0001	0,05	0,0016	0,001
0,1	0,371429	7,17E-05	4,34E-05	0,05	0,0016	0,001
0,1	0,371429	7,17E-05	4,34E-05	0,192857	0,0016	0,001
1,185714	0,371429	7,17E-05	4,34E-05	0,192857	0,0016	0,001

Between each two line, only one parameter changes, which meets the idea of One-at-a-time (OAT) strategy in a hypercube explained in chapter 2.3. Also, within each eight lines, the endpoints of samples are joint together, making the samples a complete connected path, this is also the idea of reducing the number of samples introduced in chapter 2.3.3. Note that in chapter 2.3.2, the Morris step is corresponding to the ranges of [0, 1]. In the table, the 'steps' of parameters in trajectories, which means the changes of a parameter within one trajectory, are transferred into the real scale of parameters from [0 ,1]. The normalization will be operated automatically in the analysis package. To illustrate the samples related to the ranges and real physical phenomenon, the unnormalized samples are shown.

Current velocity

Current velocity has six inputs, which are introduced in chapter 3.3.2. Table 3.6 shows part of the samples for current velocity in Morris method. The explanations of the table are very similar to the explanations for temperature's samples. Here in table 3.6, seven lines form a trajectory.

3.4.2. Simulations of the models and time series of outputs

After having the samples, the Matlab codes are written to replace the values in mdu-files and generate new ones. Besides the values of parameters, there are some other settings in the models. All the files including the grid and map information, external information, etc. are provided by Deltares. The coordinates of six observation points are written in xyn-files, and the simulation periods are set from 2004-12-25 02:24:00 to 2005-01-25 02:24:00, lasting for one month. The time interval is 30 minutes, and there are 1489 timepoints in total. Specifically, 2004-12-25 02:24:00 is the 1^{st} timepoint, 2004-12-25 02:54:00 is the 2^{nd} timepoint, 2004-12-25 03:24:00 is the 3^{rd}, and so on.

To simplify the simulations, files for each model are saved in a subfolder, and an overall shell file is created to upload the models to the Linux-based cluster together. In the cluster, seven models can run at the same time. Totally there are 150 models needed in Morris method, among them 80 are for temperature and 70 are

Table 3.6: Samples simulated for current velocity in Morris method.

Vicouv	Dicouv	Vicoww	Dicoww	Smagorinsky	Rhoair
0,914286	0,642857	0,000001	4,34E-05	0,157143	1,2509
0,914286	0,642857	0,000001	4,34E-05	0,157143	1,3669
2	0,642857	0,000001	4,34E-05	0,157143	1,3669
2	1,728571	0,000001	4,34E-05	0,157143	1,3669
2	1,728571	0,000001	0,0001	0,157143	1,3669
2	1,728571	0,000001	0,0001	0,3	1,3669
2	1,728571	5,76E-05	0,0001	0,3	1,3669
2	0,914286	0,0001	5,76E-05	0,05	1,2799
2	0,914286	0,0001	5,76E-05	0,05	1,1639
2	0,914286	0,0001	0,000001	0,05	1,1639
0,914286	0,914286	0,0001	0,000001	0,05	1,1639
0,914286	2	0,0001	0,000001	0,05	1,1639
0,914286	2	0,0001	0,000001	0,192857	1,1639
0,914286	2	4,34E-05	0,000001	0,192857	1,1639
1,185714	1,185714	0,000001	7,17E-05	0,3	1,1639
0,1	1,185714	0,000001	7,17E-05	0,3	1,1639
0,1	0,1	0,000001	7,17E-05	0,3	1,1639
0,1	0,1	0,000001	1,51E-05	0,3	1,1639
0,1	0,1	0,000001	1,51E-05	0,157143	1,1639
0,1	0,1	5,76E-05	1,51E-05	0,157143	1,1639
0,1	0,1	5,76E-05	1,51E-05	0,157143	1,2799

for current velocity. All the simulations took about one week to complete in total.

In the outputs file, the vertical directions are divided into different layers. The data of top layer closest to the sea surface, where the blue mussel and seaweed cultivation are conducted, is used in sensitivity analysis. At each observation points, there are 80/70 model results for each output. After doing the average to the models, we can have one time series for each output at each observation points. According to the time series plots, several timepoints were selected to do sensitivity analysis.

Temperature

Figure 3.5 shows the time series plot of temperature simulated in Morris method at six different observation points.

In the plot we can see the trends of the temperature show similar patterns in different observation points. The up-and-down changes indicate daily variations. Also, there is an overall downward trend, meeting the gradual coldness in January. First two timepoints having obvious characteristics are chosen, one crest and one trough. These two will be compared to study if the crests and troughs keep the same patterns in sensitivity analysis. Then to prove if the crests keep the same patterns, and if the trough keeps the same, one more crest and one more trough are selected. They are separated far on the time scale, from which the influences of simulation time can also be understood roughly.

The two crests are marked as red lines in the figure 3.5: the 693^{rd} timepoint at 2005-01-08 12:24:00, the 1159^{th} timepoint at 2005-01-18 05:24:00. The two troughs are marked as blue lines in the figure: the 130^{th} timepoint at 2004-12-27 18:54:00, the 750^{th} timepoint at 2005-01-09 16:54:00.

Current velocity

Figure 3.6, 3.7 and 3.8 show the time series plots of current velocity in x-, y- and z-directions simulated in Morris method at six different points.

The figures all show up-and-down perturbations, which represent the direction changes of tides. Changes in x- and y-direction are more regular around a comparatively fixed value, while the median of the 'waves' in y-direction have a changing trend. It can be understood if the wind is strong in this direction. The data of winds in this period is lost, but this assumption is reasonable. The numerical instability can be seen at first,

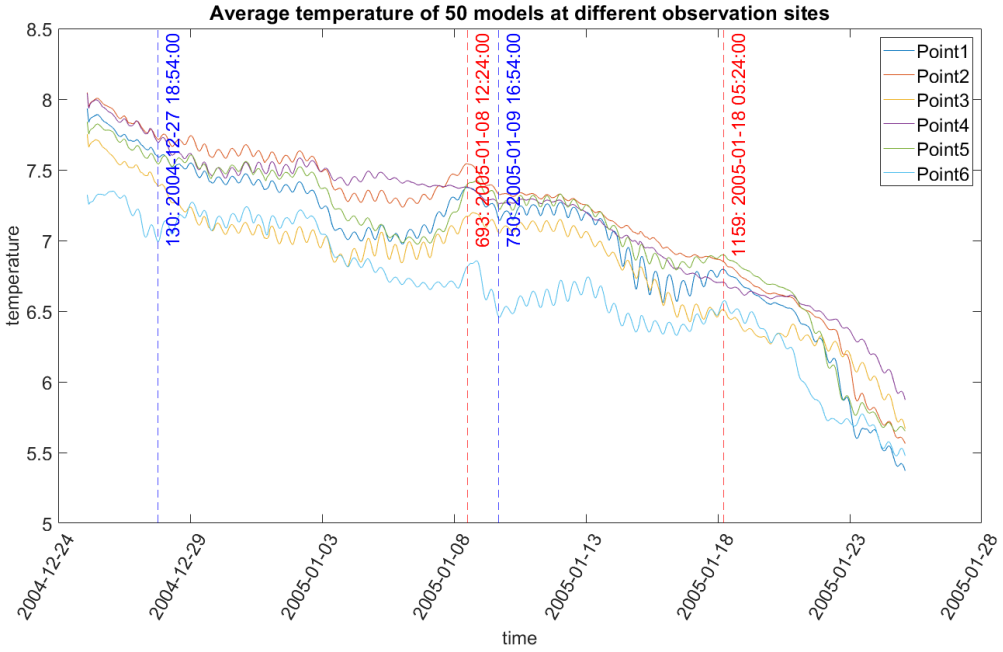


Figure 3.5: Average temperature of 80 models at six observation points in Morris method. The lines from left to the left: the 130th timepoint at 2004-12-27 18:54:00, the 693rd timepoint at 2005-01-08 12:24:00, the 750th timepoint at 2005-01-09 16:54:00, the 1159th timepoint at 2005-01-18 05:24:00. The blue lines are located in troughs, red lines are in crests.

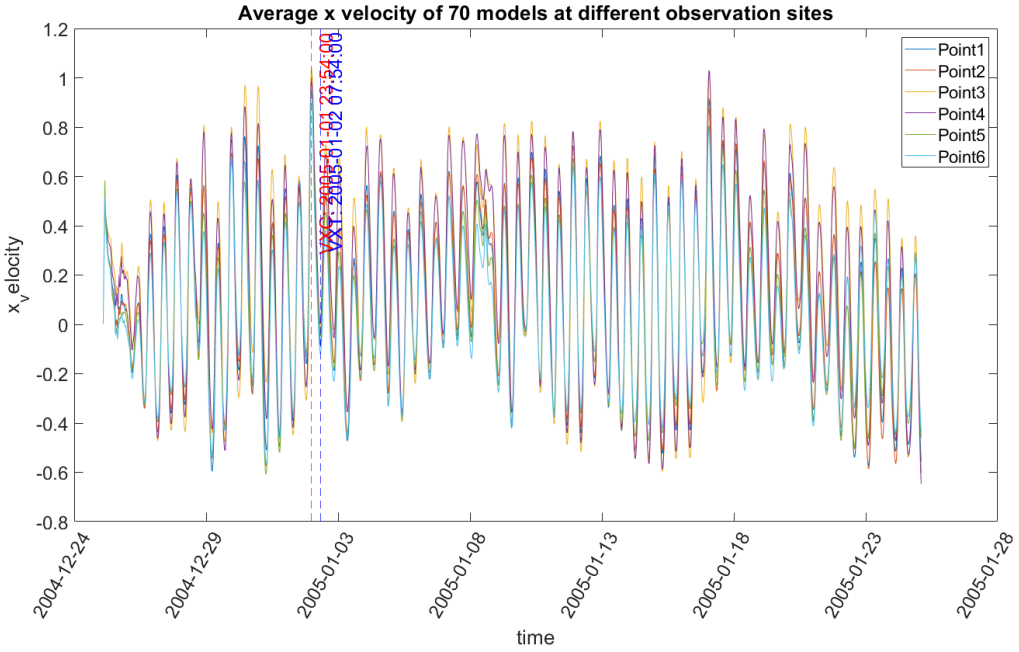


Figure 3.6: Average x-direction velocity of 50 models at six observation points in copula-based method. The left red line is 380^{th} timepoint, 2005-01-01 23:54:00, located in crest; the right blue line is 396^{th} timepoint, 2005-01-02 07:54:00, located in trough.

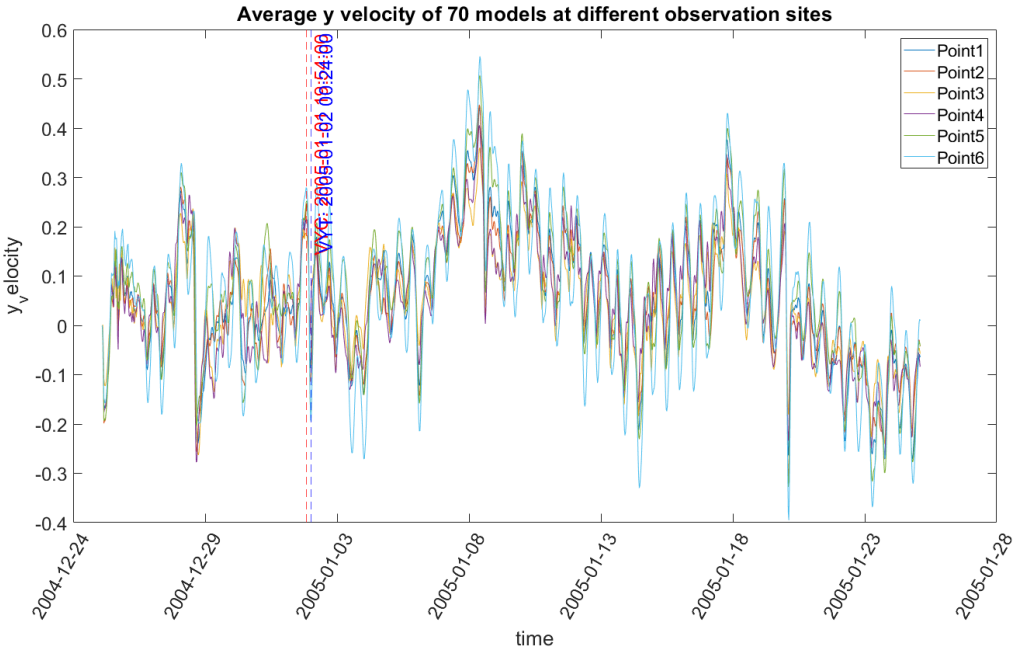


Figure 3.7: Average y-direction velocity of 50 models at six observation points in copula-based method. The left red line is 372^{nd} timepoint, 2005-01-01 19:54:00, located in crest; the right blue line is 381^{st} timepoint, 2005-01-02 00:24:00, located in trough.

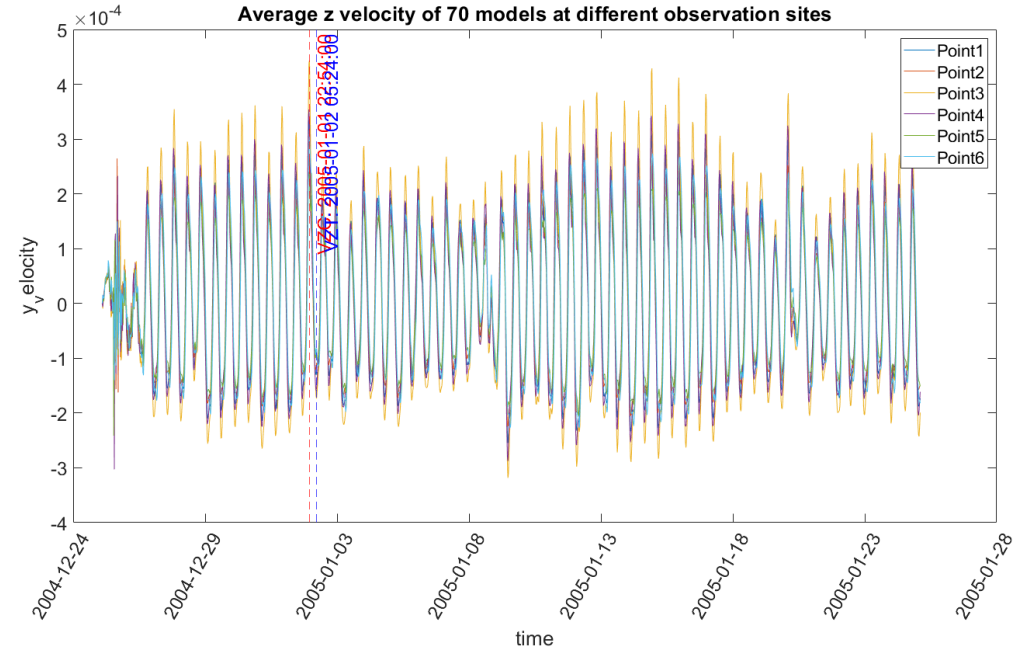


Figure 3.8: Average z-direction velocity of 50 models at six observation points in copula-based method. The left red line is 378^{th} timepoint, 2005-01-01 22:54:00, located in crest; the right blue line is 391^{st} timepoint, 2005-01-02 05:54:00, located in trough.

where the strong fluctuations take place. It can be understood as the boundary conditions are set first, and the locations within the domain take time to be stable.

Similar to temperature, crests and troughs are selected to conduct sensitivity analysis in all three directions. They are marked as red and blue lines separately. In x-direction, the crest is selected at 380^{th} timepoint, 2005-01-01 23:54:00; the trough is selected at 396^{th} timepoint, 2005-01-02 07:54:00. In y-direction, the crest is selected at 372^{nd} timepoint, 2005-01-01 19:54:00; the trough is selected as 381^{st} timepoint, 2005-01-02 00:24:00. In z-direction, the crest is selected at 378^{th} timepoint, 2005-01-01 22:54:00; the trough is selected at 391^{st} timepoint, 2005-01-02 05:54:00.

Notations

To make the timepoints easier to be marked, the timepoints chosen to do sensitivity analysis at will be noted in table 3.7.

Table 3.7: Notations of different timepoints selected to simulate the model.

Outputs	Crests/troughs	Timepoints	Notations
Temperature	Crests	693rd: 2005-01-08 12:24:00	TC1
		1159th: 2005-01-18 05:24:00	TC2
	Troughs	130th :2004-12-27 18:54:00	TT1
		750th :2005-01-09 16:54:00	TT2
Current velocity: x-direction	Crests	380th :2005-01-01 23:54:00	VXC
	Troughs	396th :2005-01-02 07:54:00.	VXT
Current velocity: y-direction	Crests	372nd :2005-01-01 19:54:00	VYC
	Troughs	381st :2005-01-02 00:24:00	VYT
Current velocity: z-direction	Crests	378th :2005-01-01 22:54:00	VZC
	Troughs	391st :2005-01-02 05:54:00	VZT

3.4.3. Sensitivity measurements

As mentioned in chapter 3.3.2, in the simulations of temperature with parameters ranging from zero, some simulations crashed: 10^{th}, 11^{th}, 12^{th}, 14^{th}, 15^{th}, 16^{th} and 77^{th}. Among them, the first six samples are within the 2^{nd} trajectories, and the 77^{th} sample is in the 10^{th} trajectory. Before changing the ranges not from zero, a solution was tried first to solve the problem. ten trajectories are included in the samples, and it is assumed that deleting two trajectories where crashed files existed is tolerated. Thus, the remaining eight trajectories are used to calculate the sensitivity indices. The idea and the calculations are the same, and the results will be showed in chapter.

However, this strategy losses the generality, as we don't know which trajectories will have the failed simulations. It will be complex to distinguish out the incomplete trajectories every time, and it is not clear whether there will be too much incomplete trajectories. Therefore, the solution of deleting part of the paths is not suggested.

To calculate the sensitivity measurements, the Elementary Effects (EE) Measures and Plots Package is used in Matlab, which is developed by Drs. Yogesh Khare and Rafael Muñoz-Carpena [24]. The main function 'EE_SenMea_Calc.m' calculated the elementary mean μ, the absolute mean μ^* and standard deviation σ introduced in chapter 2.3.2. Several files are needed in the function. The '.fac' file is also needed in this function, which is the same as the sampling package (chapter 3.4.1). The sample files and the corresponding outputs in '.xlsx' files are also compulsory. Besides these, the function needs the 'Fac_Sam_Char' file, which is generated along with the samples in the sampling package. It is a '.txt' file, containing the following information: (a) sampling strategy; (b) over sampling size; (c) number of factor levels; (d) number of trajectories; (e) number of factors.

The package will generate several files as outputs. 'EE_SA_Measures.txt' contains the information of μ, μ^* and σ. 'Raw_EE.txt' and 'Raw_EE.mat' save raw elementary effects for all outputs, introduced in chapter 2.3.2. The package also gives two figures for each model output. The first plot is of μ^* and σ while the second consists of μ and σ. Besides the dots in the plots, $\mu^* = \sigma$ line is plotted in the first plot, and $\mu = \pm 2\sigma/\sqrt{r}$ is plotted in the second one (where r is the number of sampling trajectories, and $\sigma/\sqrt{r}$ is noted by SEM, which means the

standard error of the mean). The $\mu^* = \sigma$ line threshold in the first plot is based on Chu-Agor et al. [13] and Khare et al. [25]. The $\mu = \pm 2\sigma/\sqrt{r}$ lines were proposed by Morris [33] to identify factors with dominant non-linear effects. Factors above the lines in both plots are considered to have dominant interactions effects [24]. Also, in the second plot, if the coordinates lie outside of the wedge formed by the two lines, the expectation of the factor's distributions can be seen as non-zero with significant evidence [33]. Note that the lines are not strict standards to evaluate the interaction effects, for example, parameters below them but close to the lines might also include correlation information. Additionally, the monotonicity information of the effects can be found in the plots. The factors are recognized with perfectly monotonic effects when $\mu^* = |\mu|$ holds true. They are indicated by solid red circles otherwise they are indicated by blue asterisk. In addition, bootstrapping based 95% CI for μ^* and μ are indicated on the subplots by horizontal error bars.

Also, to check the results of the package, another codes are recreated according to the logic and formulas in chapter 2.3. The codes can be found in the appendix. The codes find the 'steps' where the differences of each parameter contribute to the changes in the results. Then the elementary effects and sensitivity indices are calculated according to the formulas. The codes only export the values, without the plots produced like above. After testing, the two codes have the same outputs.

3.5. Copula-based method

In chapter 2.4, the theories of copula-based method are introduced. In this chapter, the operations of this method will be presented, and the general steps are similar to Morris method in chapter 3.4.

3.5.1. Copula decision

The copulas including the correlation information will be defined before sample generation, to make the samples more reasonable to the real physical process. The factors related will be gathered within the same copula, and the correlations are within the ranges of [-1, 1]. To simplify the sample generation, in this thesis project, each copula will only have two parameters, and the correlations will be defined as 1 or -1. The copulas containing different number of parameters, with correlations other than 1 and -1 can be studied further (see chapter 5.1). The copulas used in this thesis project are suggested by experts from Deltares.

Table 3.8: Copula defined before simulations.

Copulas	Corresponding outputs	Correlations
Vicouv – Dicouv	Temperature, current velocity	1
Vicoww – Dicoww	Temperature, current velocity	1
Stanton – Dalton	Temperature	1
Smagorinsky	Temperature, current velocity	-
Rhoair	Current velocity	-

The explanations of the parameters can be found in chapter 3.3.2.

As illustrated in chapter 3.1.3, all the parameters changed are background constant values in mdu-file. The horizontal and vertical background eddy viscosity/diffusivity are used to simulate the effects of the mixing of respectively velocity that cannot be resolved by the numerical grid. Therefore, in most cases the rank correlation between Vicouv and Dicouv is positive. The similar situation holds for Vicoww and Dicoww. However, note that these background values in mdu-files and the eventually used viscosity/diffusivity values can be different. Thus, the resulting temperature and current velocity can be different. In vertical direction, the eventually used values are determined by the $K-\epsilon$ turbulence model; in horizontal direction, the horizontal viscosity is the result of background value combined with the contribution of the Smagorinsky (subgrid) model. One might doubt that it is not the 'real' connection between viscosity/diffusivity and temperature/current velocity, but in this thesis project, the parameters in mdu-files are studied, instead of the real values, therefore the differences are tolerated.

Convection and evaporation are different processes, but the values Stanton and Dalton are for the transfer between the same substances. A positive rank correlation between the two parameters can be expected. In practice, the two values are set to the same value in numerical models. If needed, one or the other can be

increased or decreased slightly during calibration, so a negative correlation is also possible. To simplify the operation and keep the same as in practice, the correlation between Stanton and Dalton is set as 1.

As explained above, the eventually used viscosity is contributed to by Smagorinsky horizontal turbulence, and a positive rank correlation can be seen in the computation. Smagorinsky is useful in case of spatially varying grid cell sizes, as the effects of mixing cannot be resolved by the numerical grid which depends on the grid cell size, and a uniform Vicouv value will not be sufficient to capture this correctly. However, there is not a clear relation between the model input keywords Smagorinsky and Vicouv, which serves as the background value.

A similar reason holds for RhoAir and Vicouv/Vicoww: they are background values. However, increasing RhoAir and increasing Vicoww can be understood to have a similar effect on the model results. For example, both will breakdown temperature stratification when the values are increased. Increasing RhoAir (air density) results in more wind stress, generating more vertical shear, and then more vertical mixing and increasing Vicoww will appear.

3.5.2. Sample generation

Based on chapter 2.4, the settings of sample generation in copula-based method are as follows:

Table 3.9: Settings in Copula-based method.

Number of levels (p)	8
Morris step size ($\delta = p/[2(p-1)]$)	4/7
Number of trajectories (r)	10
Number of segments (k)	Temperature: 4
	Current velocity: 4
Number of samples $(k+1)*r$	Temperature: (4+1)*10=50
	Current velocity: (4+1)*10=50

In the table 3.9, the segments include copulas and independent factors. The details of the copulas can be found in table, and the ranges of the factors is in table.

To generate the samples, the Matlab codes developed by Ţene et al. [51] were used. It is part of Deltares OpenEarthTools, and it is specifically used to do copula-based sensitivity analysis. The 'step1_sampling.m' function contains related functions and information, where the settings will be declared, including the information in table above, parameter ranges, correlations and groups. The function assumes each parameter is uniformly distributed within the ranges. In 'step1_sampling.m', 'extended_morris.m' function distinguishes copulas and dependent parameters, and generates the samples based on the three steps in chapter 2.4: first choose the target cells, then compute rank statistics; in each cell, the corner and the order are decided to construct the paths. The samples will be saved in xlsx-file; the settings and paths will be saved respectively in 'setup.mat' and 'paths.mat' files, which will be used later in sensitivity measurements codes.

The following tables give examples of samples in temperature and current velocity.

Temperature

In the table 3.10, there are seven parameters with four segments, each five forming a trajectory. Each parameter changes only once within one trajectory, and only one parameter changes within two adjacent samples. In each trajectory, the first sample and the last sample represents the starting and ending points in a cell respectively, which are on the opposite corner of the chosen cell. The values of Vicouv and Dicouv are always the same as the correlation is 1. It holds the same for Vicoww and Dicoww, Stanton and Dalton.

Current velocity

Table 3.11 gives the first three trajectories as an example. For current velocity, there are six parameters with four segments, each five forming a trajectory. The characteristics are similar with table above. In current velocity, the values of Vicouv and Dicouv, Vicoww and Dicoww are always the same as the correlations are 1.

Table 3.10: Samples simulated for temperature in copula-based method.

Vicouv	Dicouv	Vicoww	Dicoww	Smagorinsky	Stanton	Dalton
2	2	1,51429E-05	1,51429E-05	0,15714	0,00126	0,00126
1,72857	1,72857	1,51429E-05	1,51429E-05	0,15714	0,00126	0,00126
1,72857	1,72857	0,000001	0,000001	0,15714	0,00126	0,00126
1,72857	1,72857	0,000001	0,000001	0,15714	0,00117	0,00117
1,72857	1,72857	0,000001	0,000001	0,19286	0,00117	0,00117
1,45714	1,45714	2,92857E-05	2,92857E-05	0,26429	0,001	0,001
1,45714	1,45714	2,92857E-05	2,92857E-05	0,26429	0,001086	0,001086
1,45714	1,45714	1,51429E-05	1,51429E-05	0,26429	0,001086	0,001086
1,18571	1,18571	1,51429E-05	1,51429E-05	0,26429	0,001086	0,001086
1,18571	1,18571	1,51429E-05	1,51429E-05	0,3	0,001086	0,001086
1,18571	1,18571	7,17143E-05	7,17143E-05	0,22857	0,001257	0,001257
1,18571	1,18571	7,17143E-05	7,17143E-05	0,19286	0,001257	0,001257
1,18571	1,18571	7,17143E-05	7,17143E-05	0,19286	0,001343	0,001343
0,91429	0,91429	7,17143E-05	7,17143E-05	0,19286	0,001343	0,001343
0,91429	0,91429	5,75714E-05	5,75714E-05	0,19286	0,001343	0,001343

Table 3.11: Samples simulated for current velocity in copula-based method.

Vicouv	Dicouv	Vicoww	Dicoww	Smagorinsky	Rhoair
0,37143	0,37143	5,75714E-05	5,75714E-05	0,19286	1,2219
0,37143	0,37143	5,75714E-05	5,75714E-05	0,22857	1,2219
0,37143	0,37143	4,34286E-05	4,34286E-05	0,22857	1,2219
0,1	0,1	4,34286E-05	4,34286E-05	0,22857	1,2219
0,1	0,1	4,34286E-05	4,34286E-05	0,22857	1,2509
1,18571	1,18571	0,000001	0,000001	0,3	1,3669
1,18571	1,18571	0,000001	0,000001	0,26429	1,3669
0,91429	0,91429	0,000001	0,000001	0,26429	1,3669
0,91429	0,91429	0,000001	0,000001	0,26429	1,3379
0,91429	0,91429	1,51429E-05	1,51429E-05	0,26429	1,3379
2	2	1,51429E-05	1,51429E-05	0,05	1,2799
1,72857	1,72857	1,51429E-05	1,51429E-05	0,05	1,2799
1,72857	1,72857	1,51429E-05	1,51429E-05	0,08571	1,2799
1,72857	1,72857	2,92857E-05	2,92857E-05	0,08571	1,2799
1,72857	1,72857	2,92857E-05	2,92857E-05	0,08571	1,3089

3.5.3. Simulations of the models and time series of outputs

The same as chapter 3.5.2, values of parameters in mdu-files are changed through Matlab codes, and each model with different parameters is located in subfolders. Other files including the grid and map information, external information, etc. keep the same for all the models. The same observation points, simulation periods and time intervals are used (see chapter 3.2 and chapter 3.5.2). In the total, there are 100 models needed in copula-based method, where 50 are for temperature and 50 are for current velocity. They are uploaded in the cluster to run, taking approximate four days to finish.

The top layer closest to the sea surface is chosen to be studied. After doing the average on 50 models at each observation points, we can have the time series plots for the outputs, temperature and current velocity.

Temperature

Table 3.9 shows the time series plot of temperature simulated in Morris method at six different observation points.

The plot shows similar features to the time series in Morris method. The trends of the curves are similar at different observation points; the up-and-down changes were from the changes of day and night; the overall

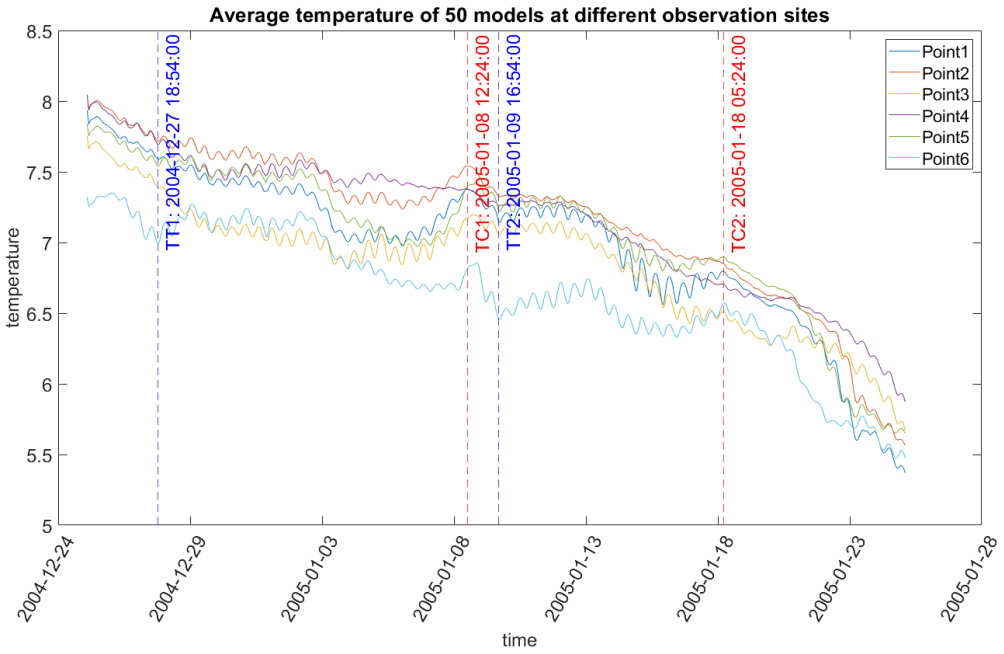


Figure 3.9: Average temperature of 50 models at six observation points in copula-based method. The lines from left to the left: the 130th timepoint at 2004-12-27 18:54:00, the 693rd timepoint at 2005-01-08 12:24:00, the 750th timepoint at 2005-01-09 16:54:00, the 1159th timepoint at 2005-01-18 05:24:00. The blue lines are located in troughs, red lines are in crests.

downward trend could be understood in January.

The two crest timepoints chosen in Morris method, TC1 and TC2, also meets the crests in table 3.10. Thus, these two timepoints are chosen for sensitivity analysis in copula-based method. It holds the same for TT1 and TT2, which are at troughs.

Current velocity

Table 3.10, 3.11 and 3.12 show the time series plots of current velocity in x-, y- and z-directions simulated in Morris method at six different points.

The features of the plots 3.10, 3.11 and 3.12 are similar to the simulation results in Morris method. The up-and-down perturbations indicate the change of tides; the changes in x- and z-directions are comparatively around a fixed value, while the changes in y-direction have different trends, which can be understood as the influences of wind directions. As the boundary conditions are set first and it takes time for the locations within the domain to be stable, the numerical instability can be seen at first.

All the points chosen in Morris method are all located exactly in the crests and troughs. Thus, the same timepoints are chosen for copula-based methods: VXC, VXT; VYC, VYT; VZC, VZT. The explanations of the notations can be found in chapter 3.5.2.

3.5.4. Sensitivity measurements

Ţene et al. developed the Matlab package to calculate sensitivity indices in copula-based method [51], as a follow-up step of sample generation codes introduced in chapter 3.5.2. The main function is 'step2_sensitivity.m'. After loading 'setup_copula.mat' and 'paths_copula.mat' created by 'step1_sampling.m' introduced in chapter 3.5.2, where the information of settings, copulas and paths are saved, combining the outputs of simulations, the elementary effects and three sensitivity indices are calculated according to the definition. The copulas and independent parameters will be distinguished before calculation. The codes for plotting in chapter 3.4.3 will be adjusted, and similar plots will be created.

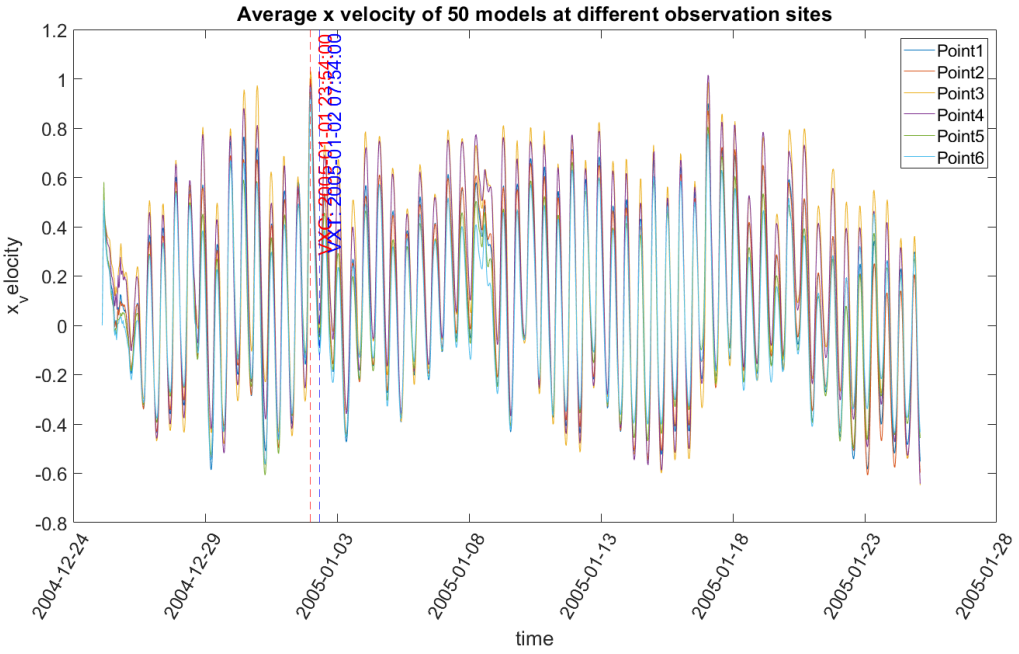


Figure 3.10: Average x-direction velocity of 50 models at six observation points in copula-based method. The left red line is 380^{th} timepoint, 2005-01-01 23:54:00, located in crest; the right blue line is 396^{th} timepoint, 2005-01-02 07:54:00, located in trough.

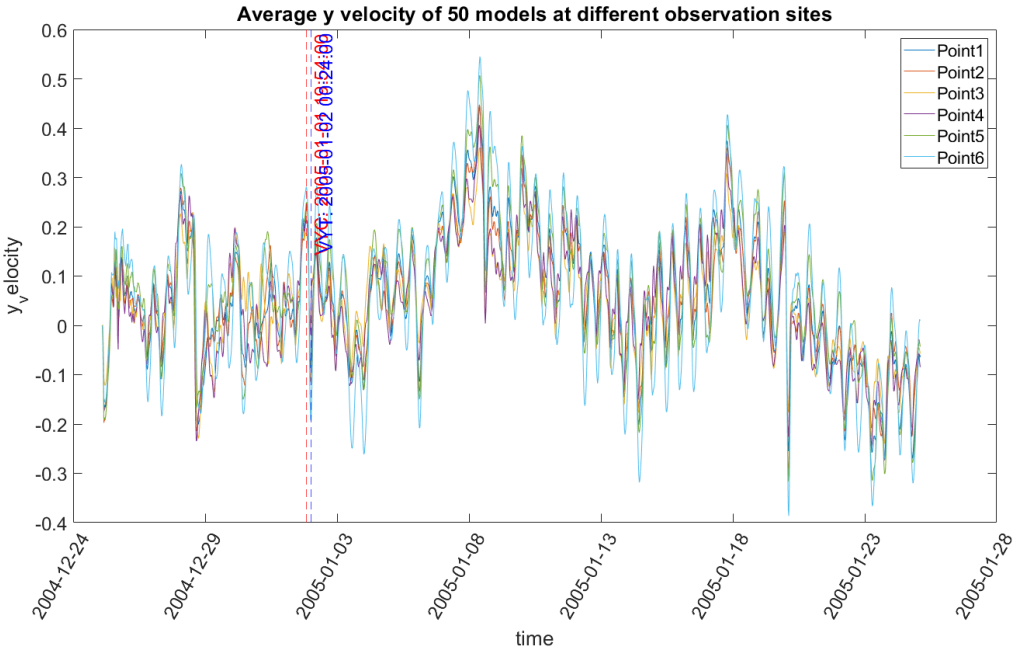


Figure 3.11: Average y-direction velocity of 50 models at six observation points in copula-based method. The left red line is 372^{nd} timepoint, 2005-01-01 19:54:00, located in crest; the right blue line is 381^{st} timepoint, 2005-01-02 00:24:00, located in trough.

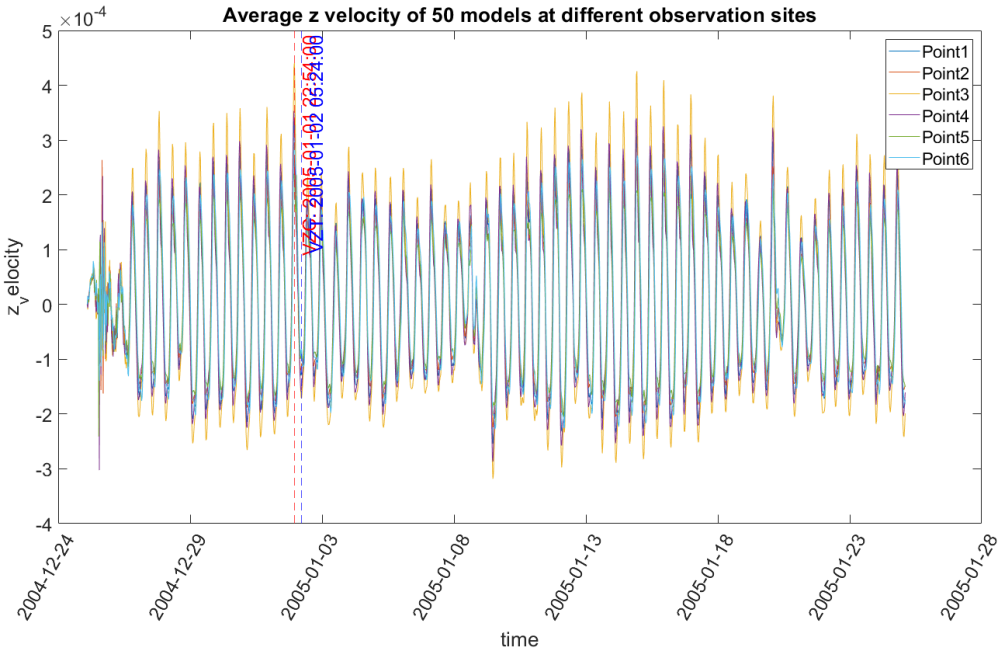


Figure 3.12: Average z-direction velocity of 50 models at six observation points in copula-based method. The left red line is 378^{th} timepoint, 2005-01-01 22:54:00, located in crest; the right blue line is 391^{st} timepoint, 2005-01-02 05:54:00, located in trough.

3.6. Variance-based method

In chapter 2.5, the general idea of using variance in variance-based method is introduced, along with the sensitivity indices and the computation strategy. In this methodology part, the operations of variance-based method will refer to the strategy in chapter 2.5.5.

3.6.1. Sample strategy

Two matrices A and B are compulsory for both temperature and current velocity. The size of A and B is $(N, 2k)$, where k is the number of inputs, and N is called a base sample, which is suggested to vary from a few hundred to a few thousands. Besides, the number of matrices C_i are decided by k, where all the columns are kept the same with matrix B, except the ith column being the same with the ith column of matrix A. Therefore, the number of samples in total is $N * (k+2)$. To reduce the simulations of the model, N is chosen as small as possible. From the previous two methods, the background value Vicouv and Dicouv have no influence on the outputs (it will be discussed in chapter 5.1), thus, to simplify the models, these two values are abandoned in variance-based method. For temperature, Vicoww and Dicoww are also skipped. The following table shows the parameters used in variance-based method:

Table 3.12: Inputs and corresponding outputs in variance-based method

Outputs	Inputs
temperature	Smagorinsky
	Stanton
	Dalton
Current velocity	Vicoww
	Dicoww
	Smagorinsky
	Rhoair

Temperature

In temperature, the matrix samples and results are provided by colleagues from Deltares. N is chosen as 100. Thus, there are $100 * (2+3) = 500$ samples in temperature.

Current velocity

In current velocity, SimLab v2.2.1 [42] introduced in chapter 3.4 is used to generate the random matrices A and B. Through SimLab v2.2.1, quasi-random sampling is used to generate A; random sampling is used to generate B. As SimLab v2.2.1 has default numbers of samples, N is decided as 128 in current velocity, leading to totally $128 * (2+4) = 768$ models.

Under certain conditions, largely governed by the method of compiling the sampling frame or list, a systematic sample of every nth entry from a list will be equivalent for most practical purposes to a random sample. This method of sampling is sometimes referred to as quasi-random sampling [28]. In random sampling, each sample has an equal probability of being chosen.

3.6.2. Simulations of the models and choosing timepoints

The results of temperature are provided by colleagues from Deltares. Thus, only models for current velocity were needed to simulate.

The 768 models are divided into six matrices, and the following procedures are the same: changing the values in mdu-files, creating subfolders to save each model, creating overall shell file to submit the models together to the cluster.

As similarities appear between different timepoints, which will be discussed in the details in chapter 4, to simplify the analysis and avoid duplication, only one timepoint was chosen for each system in variance-based method. Timepoint VXC, VYC and ZYC were picked for current velocity in x-, y- and z-directions respectively.

3.6.3. Sensitivtiy measurements

The first-order sensitivity index and total effect index are calculated according to the formulas in chapter 2.5.5, using Matlab. The codes can be found in the appendix.

4

Results

In this chapter the results of sensitivity analysis are discussed. The results are presented per methods: Morris method, copula-based method and variance-based method. In each method, two different outputs are selected: temperature and current velocity, and current velocity data are available in three directions, which will be discussed separately. In Morris method, four timepoints are selected for temperature; two timepoints are selected for each direction in current velocity. In copula-based method, the same timepoints are selected. In variance-based method, one timepoint is selected for temperature, and one timepoint is selected for each direction in current velocity. Besides, six observation points are studied, including FINO3 platform.

4.1. Morris method

In this section, the results of temperature and three directions of current velocities are discussed respectively. The values of elementary mean μ, absolute mean μ^* and standard deviation σ are listed in tables. To present more directly and clearly, the plots of μ^* vs σ, and μ vs σ are created. Only part of the results will be illustrated in the report, the full information can be found in appendix.

4.1.1. Temperature

As introduced in chapter 3.5, first, the ranges from zero are used in temperature's simulations in Morris method. Afterwards, the analysis will be based on the samples not ranging from zero. As a solution to crashed files, incomplete trajectories are abandoned. This operation is only adapted as a possible solution on temperature in Morris method. In details, the 2^{nd} and 10^{th} trajectories were deleted among the total ten trajectories. To simplify the process, time series plots was not made in this strategy, and TC1 was chosen directly from the analysis in chapter 3.5.2.

Table 4.1 shows the sensitivity analysis results at timepoint TC1. The interpretation of the table will be discussed later, combined with the results without ranging from zero.

Table 4.1: Morris method: sensitivity indices for temperature at timepoint TC1, parameters ranging from zero

Point1				Point2			
facname	mu*	mu	sigma	facname	mu*	mu	sigma
Dalton	0,20190	-0,20190	0,00822	Dalton	0,19798	-0,19798	0,00780
Smagorinsky	0,03066	-0,03066	0,01985	Stanton	0,02108	-0,02108	0,00493
Stanton	0,02347	-0,02347	0,00563	Smagorinsky	0,00193	-0,00118	0,00209
Vicoww	0,00092	-0,00090	0,00049	Vicoww	0,00063	-0,00063	0,00032
Dicoww	0,00028	0,00018	0,00034	Dicoww	0,00020	0,00012	0,00028
Vicouv	0	0	0	Vicouv	0	0	0
Dicouv	0	0	0	Dicouv	0	0	0

Point3				Point4			
facname	mu*	mu	sigma	facname	mu*	mu	sigma
Dalton	0,20694	-0,20694	0,01235	Dalton	0,19684	-0,19684	0,00830
Smagorinsky	0,16594	-0,16594	0,03530	Smagorinsky	0,01883	-0,00503	0,02052
Stanton	0,01183	-0,01024	0,00995	Stanton	0,00529	-0,00245	0,00546
Vicoww	0,00182	0,00132	0,00185	Vicoww	0,00131	-0,00083	0,00118
Dicoww	0,00042	-0,00009	0,00062	Dicoww	0,00031	0,00020	0,00038
Vicouv	0	0	0	Vicouv	0	0	0
Dicouv	0	0	0	Dicouv	0	0	0
Point5				Point6			
facname	mu*	mu	sigma	facname	mu*	mu	sigma
Dalton	0,19720	-0,19720	0,00825	Dalton	0,17375	-0,17375	0,00745
Stanton	0,02663	-0,02663	0,00532	Smagorinsky	0,09701	-0,09701	0,02336
Smagorinsky	0,01001	-0,01001	0,00667	Stanton	0,03009	-0,03009	0,00600
Vicoww	0,00081	-0,00081	0,00025	Vicoww	0,00093	-0,00077	0,00068
Dicoww	0,00012	0,00008	0,00013	Dicoww	0,00037	0,00029	0,00030
Vicouv	0	0	0	Vicouv	0	0	0
Dicouv	0	0	0	Dicouv	0	0	0

In the following parts, all the results are based on the ranges in table 3.2. Table 4.2 shows the three sensitivity indices at timepoint TC1: 2005-01-08 12:24:00. To compare the differences between observation points, the table presents the complete results at TC1. Table ranks the parameters according to the values of μ.

Table 4.2: Morris method: sensitivity indices for temperature at timepoint TC1

Point1				Point2			
facname	mu*	mu	sigma	facname	mu*	mu	sigma
Dalton	0,20232	-0,20232	0,00673	Dalton	0,19813	-0,19813	0,00620
Smagorinsky	0,02255	-0,02255	0,01616	Stanton	0,01904	-0,01904	0,00442
Stanton	0,02114	-0,02114	0,00478	Smagorinsky	0,00160	-0,00054	0,00213
Vicoww	0,00050	-0,00008	0,00061	Vicoww	0,00024	-0,00024	0,00015
Dicoww	0,00026	0,00017	0,00025	Dicoww	0,00018	0,00009	0,00019
Vicouv	0	0	0	Vicouv	0	0	0
Dicouv	0	0	0	Dicouv	0	0	0
Point3				Point4			
facname	mu*	mu	sigma	facname	mu*	mu	sigma
Dalton	0,20648	-0,20648	0,00945	Dalton	0,19713	-0,19713	0,00695
Smagorinsky	0,13774	-0,13774	0,01996	Smagorinsky	0,01663	-0,00038	0,01951
Stanton	0,01057	-0,00831	0,01085	Stanton	0,00330	-0,00002	0,00395
Vicoww	0,00294	0,00261	0,00285	Vicoww	0,00173	0,00094	0,00199
Dicoww	0,00033	-0,00013	0,00046	Dicoww	0,00019	0,00009	0,00030
Vicouv	0	0	0	Vicouv	0	0	0
Dicouv	0	0	0	Dicouv	0	0	0
Point5				Point6			
facname	mu*	mu	sigma	facname	mu*	mu	sigma
Dalton	0,19771	-0,19771	0,00668	Dalton	0,17507	-0,17507	0,00652
Stanton	0,02452	-0,02452	0,00472	Smagorinsky	0,07982	-0,07982	0,01367
Smagorinsky	0,00726	-0,00726	0,00535	Stanton	0,02874	-0,02874	0,00644
Vicoww	0,00069	-0,00069	0,00018	Vicoww	0,00117	-0,00117	0,00061
Dicoww	0,00016	0,00010	0,00019	Dicoww	0,00046	0,00046	0,00018
Vicouv	0	0	0	Vicouv	0	0	0
Dicouv	0	0	0	Dicouv	0	0	0

Generally, Dalton is the most influential parameter to temperature, showing highest μ^*, while Vicoww and Dicoww have the least influences except Vicouv and Dicouv. Smagorinsky has medium influences among parameters. The rankings of Stanton and Smagorinsky may exchange the orders at different observation sites.

Although each parameter is assumed as independent in Morris method, leading to the possible irrationality in simulation results and correlations, we can still have a rough interpretation of correlations from the values of sigma. Except Vicouv and Dicouv having no influences, Vicoww and Dicoww has smallest values of sigma, indicating the least correlations with other parameters. Generally, Smagorinsky has the biggest value of sigma, representing the most correlations with other parameters. Slight differences at one or two observation points can be tolerated.

Three sensitivity indices of Vicouv and Dicouv all show as zero. The parameters are selected from mdu-file, however, Vicouv and Dicouv are also specified in the ext-file. The values in mdu-file are set as a background value, but they are overwritten and adjusted by the values in external files. Therefore, the values we changed have no impact on the computation.

Different observation points show similarities in the results, as interpretated above. Also, in the first table, the solution of deleting incomplete trajectories have similar results with table below, such as the orders, the values, the characteristics of the signs, and so on. It indicates the solution of deleting part of the trajectories is also feasible.

Besides the table containing the specific values, the sensitivity indices are presented in coordinates as illustrated in chapter 3.5.3. Figure 4.1 shows the results at timepoint TC1 at observation point 1 (FINO3).

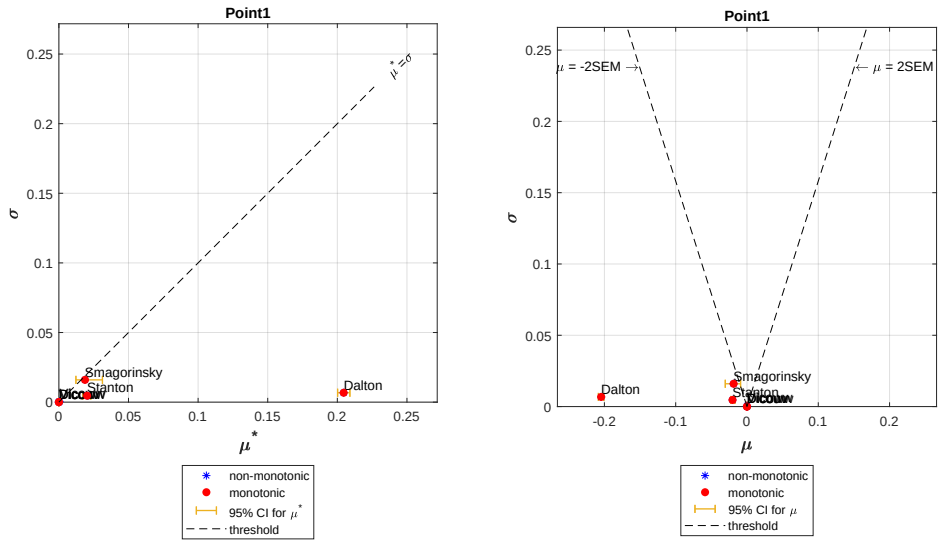


Figure 4.1: Morris method: temperature's sensitivity indices of FINO3 at TC1

The left plot shows the value of μ^* vs σ. Dicouv and Vicouv are exactly located at the origin (0, 0); Vicoww and Dicoww are very close to the origin, the distance being almost zero; Dalton is far away from other parameters; Smagorinsky and Staton are very close to each other. The sigma value of all the parameters are close to zero, and the differences are not obvious. Besides, all the parameters are below the $\mu^* = \sigma$ line, which means the evidence is not strong enough to consider the parameters having dominant interactions effects.

μ and σ are showed in the right. Similar to the left, Dicouv and Vicouv are located at the origin; Dalton is far away from other parameters; others are very close to the origin. Except Dicouv and Vicouv being zero, other parameters all lie out of the wedge formed by the two lines. Two conclusions can be made roughly made from this evidence: first, their expectations of distributions are non-zero; second, they don't have strong correlation effects with other parameters. The first conclusion is what is expected according to the ranges of parameters.

The types of the markers show the monotonicity of the parameters. The solid red points indicate the parameters are all monotonic to temperature. The factors are recognized with perfectly monotonic effects when

$\mu^* = |\mu|$ holds true.

The results of the table and figures above are at the same timepoint. To compare the properties at different timepoints, results at TC2, TT1, TT2 are also necessary to be illustrated. In order to avoid repeat and redundancy, only the observation point 1 (FINO3) will be stated in the results.

Table 4.3: Morris method: temperature's sensitivity indices of FINO3 at TC2

facname	mu*	mu	sigma
Dalton	0,27572	-0,27572	0,01065
Smagorinsky	0,04705	-0,04705	0,02651
Stanton	0,00710	0,00394	0,00769
Vicoww	0,00136	0,00136	0,00121
Dicoww	0,00043	0,00027	0,00057
Vicouv	0	0	0
Dicouv	0	0	0

Table 4.4: Morris method: temperature's sensitivity indices of FINO3 at TT1

facname	mu*	mu	sigma
Dalton	0,04600	-0,04600	0,00016
Stanton	0,02368	-0,02368	0,00015
Smagorinsky	0,00440	-0,00440	0,00141
Vicoww	0,00014	0,00005	0,00016
Dicoww	0,00006	0,00005	0,00005
Vicouv	0	0	0
Dicouv	0	0	0

Table 4.5: Morris method: temperature's sensitivity indices of FINO3 at TT2

facname	mu*	mu	sigma
Dalton	0,22732	-0,22732	0,00976
Smagorinsky	0,12205	-0,12205	0,02409
Stanton	0,01598	-0,01598	0,00952
Vicoww	0,00098	0,00041	0,00133
Dicoww	0,00044	0,00034	0,00038
Vicouv	0	0	0
Dicouv	0	0	0

The μ^* vs σ and μ vs σ plots of FINO3 at TC2, TT1 and TT2 are shown is figure 4.2 4.3 and 4.4.

At each timepoint, the analysis on six observation points is conducted, and they show similar properties. The details are illustrated as follows.

Comparing the results of TC2 with TC1, the order of parameters' μ is the same, and Smagorinsky still has the highest sigma, indicating the most correlations with other parameters. Dalton and Smagorinsky are below the lines $\mu^* = \sigma$ and $\mu = \pm 2SEM$ both in the left and right, showing no dominant interactions in the simulations, and the expectations of Dalton and Smagorinsky are significantly non-zero. Divoww and Vicoww are non-monotonic at TC2, while all parameters are monotonic at TC1. The differences can be tolerated, considering the value being very close to 0, and the bias of simulations.

Looking at TT1 and TT2, the orders of μ^* are similar, and the orders of Smagorinsky and Stanton may exchange. Smagorinsky shows highest sigma. Dalton and Smagorinsky are below the lines and the distances between the lines and these two parameters are much more obvious than other parameters, which can be interpreted as significant evidence of being non-zero. Stanton, Vicoww and Dicoww are close to the lines and the origin. Except Dicouv and Vicouv at the origin, other parameters don't show strong correlation information. All the parameters are monotonic to the output. However, note that the value of sensitivity indices at TT1 are much smaller than the values at other 3 timepoints. It can be understood as caused by being close to the starting point when the system is not stable enough as the timepoints further away.

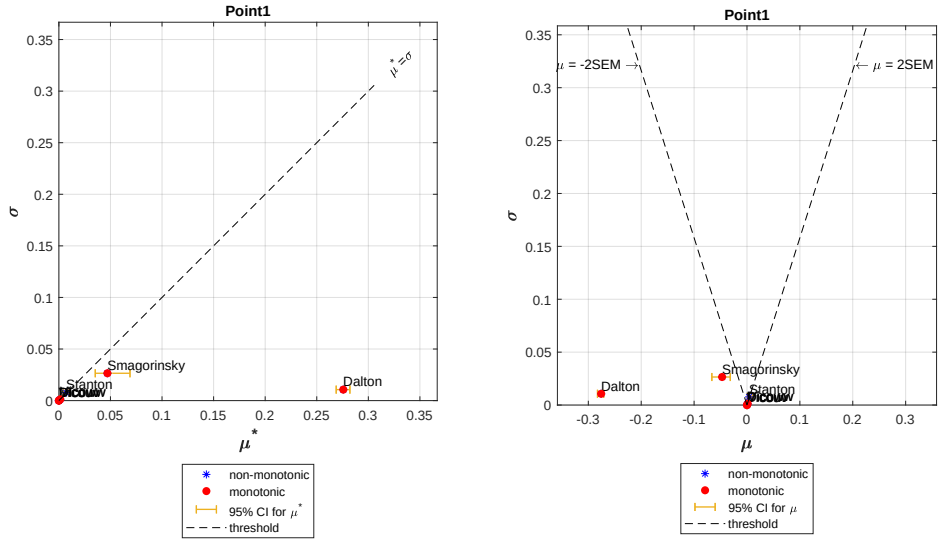


Figure 4.2: Morris method: temperature's sensitivity indices of FINO3 at TC2

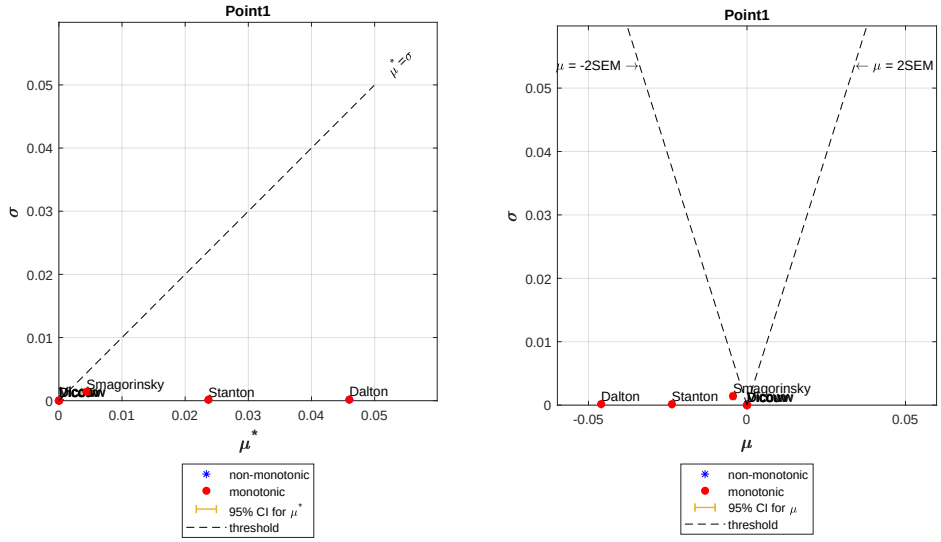


Figure 4.3: Morris method: temperature's sensitivity indices of FINO3 at TT1

After comparing the different timepoints, it can be concluded that different timepoints at a fixed observation point have similarities.

4.1.2. Current velocity

The results of current velocity will be divided into three directions.

X-direction

Table 4.6 shows the sensitivity indices at VXC (2005-01-01 23:54:00) in x-direction.

Generally, Rhoair is the most influential parameter. Vicoww might have bigger influence than Dicoww and Smagorinsky, but the differences are very limited at point 5 and 6. The μ^* values of Dicoww and Smagorinsky are very close to each other except point 3 and 4. Considering the values of μ^* and the differences are small, the three parameters, Vicoww, Dicoww and Smagorinsky can be considered to have the similar influences on current velocity in x-direction.

From σ, we can have a rough impression of the dependency information. Rhoair shows the highest σ at

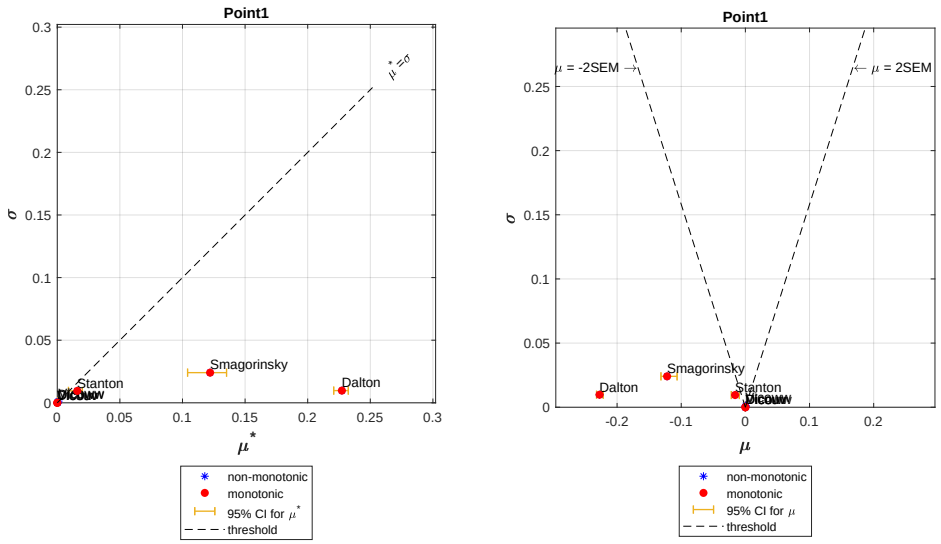


Figure 4.4: Morris method: temperature's sensitivity indices of FINO3 at TT2

Table 4.6: Copula-based method: x-direction velocity's sensitivity indices at VXC

Point1				Point2			
facname	mu*	mu	sigma	facname	mu*	mu	sigma
Rhoair	0,09341	-0,09341	0,03582	Vicoww	0,04083	0,04083	0,03679
Vicoww	0,02710	0,02652	0,02742	Rhoair	0,03107	-0,03069	0,03504
Dicoww	0,01443	-0,01068	0,01793	Dicoww	0,01790	-0,00205	0,02529
Smagorinsky	0,01294	0,00394	0,01696	Smagorinsky	0,01121	0,00359	0,02036
Vicouv	0	0	0	Vicouv	0	0	0
Dicouv	0	0	0	Dicouv	0	0	0
Point3				Point4			
facname	mu*	mu	sigma	facname	mu*	mu	sigma
Rhoair	0,02066	0,02066	0,00231	Rhoair	0,02558	0,02558	0,00297
Vicoww	0,00371	0,00371	0,00192	Vicoww	0,00420	0,00408	0,00353
Smagorinsky	0,00213	0,00167	0,00204	Dicoww	0,00152	0,00148	0,00171
Dicoww	0,00058	0,00054	0,00042	Smagorinsky	0,00078	-0,00057	0,00083
Vicouv	0	0	0	Vicouv	0	0	0
Dicouv	0	0	0	Dicouv	0	0	0
Point5				Point6			
facname	mu*	mu	sigma	facname	mu*	mu	sigma
Rhoair	0,02342	0,00906	0,02690	Rhoair	0,01511	0,00421	0,01810
Vicoww	0,01522	0,01279	0,01892	Dicoww	0,01496	-0,01248	0,01405
Smagorinsky	0,01310	0,00792	0,01828	Vicoww	0,01348	0,00623	0,01560
Dicoww	0,01189	-0,01063	0,01523	Smagorinsky	0,01157	-0,00095	0,01661
Vicouv	0	0	0	Vicouv	0	0	0
Dicouv	0	0	0	Dicouv	0	0	0

four observation points, and σ of Rhoair are very close to the highest at the other two observation points. It indicates that Rhoair is correlated most with other parameters. In fact, the differences of parameters' σ are limited, showing no significant differences in correlation information.

The sensitivity indices of Vicouv and Dicouv are all zero at six observation points. As explained in chapter 4.1.1, The values changed in mdu-file will be overwritten by external files, having no impact in simulations.

It can be concluded from the table that different observation points have similar properties.

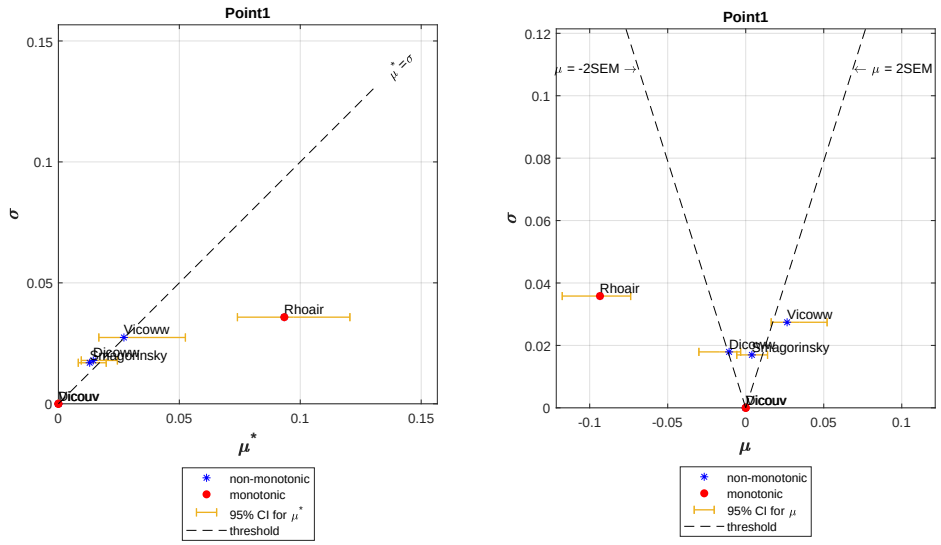


Figure 4.5: Morris method: x-direction velocity's sensitivity indices of FINO3 at VXC

Figure 4.5 shows the μ^* vs σ and μ vs σ plots of current velocity in x-direction at timepoint VXC.

The left shows the values of μ^* vs σ. Dicouv and Vicouv are exactly located at the origin. Rhoair is far away from other parameters, showing highest μ^*. Dicoww, Vicoww and Smagorinsky are close to each other.

$\mu^* = \sigma$ is used to evaluate the correlations involving the parameters. Dicoww, Vicoww and Smagorinsky are all above but very close to the line, while Rhoair is below it and the distance is more significant. The evidence is not strong enough for significant correlations within parameters.

μ vs σ is presented in the right. Dicouv and Vicouv are at the origin. Rhoair is far away from other parameters. Considering the two lines, Rhoair lies outside of the wedge; Vicoww, Dicoww and Smagorinsky are close to the lines and the last two are above the lines. It can be interpreted that the parameters are not significantly correlated, and the expectation of Rhoair and Vicoww's distributions are significantly non-zero.

The blue asterisks of Dicoww, Vicoww and Smagorinsky show non-monotonicity, while Rhoair is monotonic to the current velocity in x-direction.

In order to verify whether the properties keep the same at other timepoints, timepoint VXT (2005-01-02 07:54:00) is selected. The table shows the sensitivity indices of FINO3 at timepoint VXT.

Table 4.7: Morris method: x-direction velocity's sensitivity indices at VXT

facname	mu*	mu	sigma
Rhoair	0,02458	0,02458	0,00393
Vicoww	0,00441	0,00389	0,00358
Smagorinsky	0,00280	-0,00207	0,00285
Dicoww	0,00258	-0,00104	0,00390
Vicouv	0	0	0
Dicouv	0	0	0

As presented, Rhoair has the highest μ^*, being the most influential parameter. The differences between μ^* of

Vicoww, Dicoww and Smagorinsky's are not significant. It is the same in sigma, where the values of parameters are very close. Vicouv and Dicouv still have all the indices as zero.

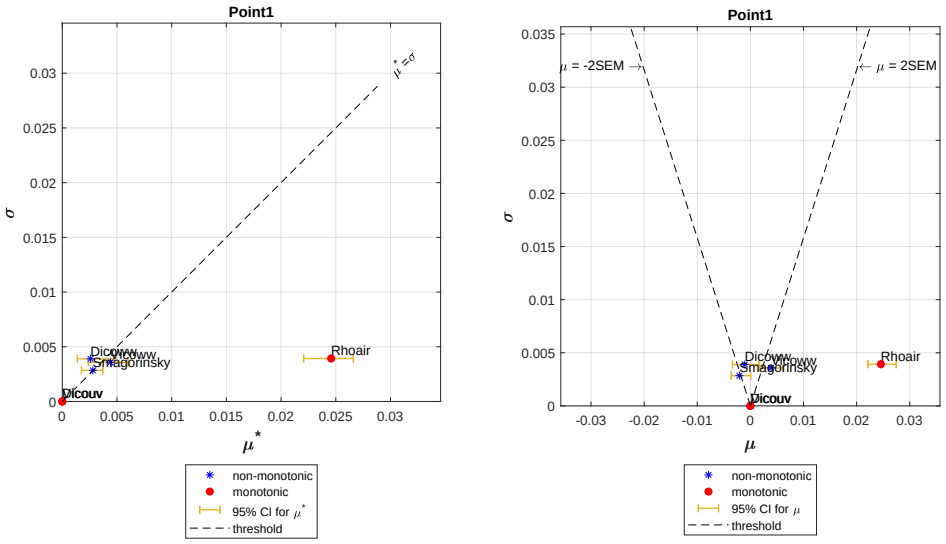


Figure 4.6: Morris method: x-direction velocity's sensitivity indices of FINO3 at VXT

The sensitivity indices are presented directly in figure 4.6. The properties keep the same at timepoint VXC. As seen in the left, parameters are in a cluster expect Rhoair, which is located at bottom right; Dicoww, Vicoww and Smagorinsky can be seen as in a cluster and are all close to the line $\mu^* = \sigma$; Dicouv and Vicouv are at the origin. In the right plot, Rhoair is outside the wedge, keeping the distance with other parameters; Dicouww, Vicoww and Smagorinsky are in a cluster and approximately have the property $\mu = \pm 2SEM$; Dicouv and Vicouv are located at the origin. Two conclusions can be generated: first, the evidence is not significant to prove the parameters have strong correlations; second, Rhoair is significantly non-zero. Except Dicouv and Vicouv, Rhoair is monotonic, while other parameters are non-monotonic.

Generally, different timepoints show similar results.

Y-direction

The same analysis is applied in y- and z-directions. In each direction, two timepoints are selected to conduct sensitivity analysis at six observation points. To check the differences between observation points, the results of four observation points at one timepoint will be presented; to compare between timepoints, the result of FINO3 at another timepoint will be presented. To avoid redundancy and repetition, only part of the results will be presented in this chapter. The complete results can be found in the appendix.

Table 4.8 shows the sensitivity indices of four observation points at VYC.

The results are similar at different observation points. The μ^* value of Rhoair is much higher than other parameters, while the μ^* values of the Vicoww, Dicww and Smagorinsky are close to each other. It shows Rhoair is the most influential parameter to current velocity in y-direction, Vicoww, Dicoww and Smagorinsky can be seen to have the similar influence on the output. Smagorinsky has the highest sigma, indicating the most correlation with others.

The figure 4.7 shows the sensitivity indices of current velocity in y-direction. In both the plots, Rhoair is not in the cluster and located at bottom right; Divcoww, Vicoww and Smagorinsky can be seen as a cluster, and all of the three parameters approximately meet $\mu = \pm 2SEM$; Dicouv and Vicouv are located at the origin. It can be concluded from the evidence that none of the parameters are significantly correlated with others, and Rhoair is significantly non-zero. From the type of the dots, Rhoair is monotonic, Dicoww, Vicoww and Smagorinsky are non-monotonic.

The properties above keep the same with that in x-direction. Then to prove if it is the same at different timepoints, the results of FINO3 at VYT are listed in table 4.9. The rankings of μ^* show similar patterns as in y-direction.

Table 4.8: Morris method: y-direction velocity's sensitivity indices at VYC

Point1				Point2			
facname	mu*	mu	sigma	facname	mu*	mu	sigma
Rhoair	0,03099	0,03099	0,00178	Rhoair	0,04735	0,04735	0,00730
Vicoww	0,00262	-0,00255	0,00208	Vicoww	0,01049	-0,00895	0,01033
Dicoww	0,00243	0,00159	0,00265	Dicoww	0,00371	-0,00062	0,00553
Smagorinsky	0,00218	0,00092	0,00280	Smagorinsky	0,00311	-0,00139	0,00480
Vicouv	0	0	0	Vicouv	0	0	0
Dicouv	0	0	0	Dicouv	0	0	0
Point3				**Point4**			
facname	mu*	mu	sigma	facname	mu*	mu	sigma
Rhoair	0,01515	0,01515	0,00077	Rhoair	0,01688	0,01688	0,00161
Smagorinsky	0,00171	-0,00152	0,00136	Vicoww	0,00202	0,00195	0,00251
Vicoww	0,00056	0,00034	0,00078	Dicoww	0,00153	-0,00075	0,00249
Dicoww	0,00037	0,00022	0,00048	Smagorinsky	0,00110	0,00070	0,00117
Vicouv	0	0	0	Vicouv	0	0	0
Dicouv	0	0	0	Dicouv	0	0	0

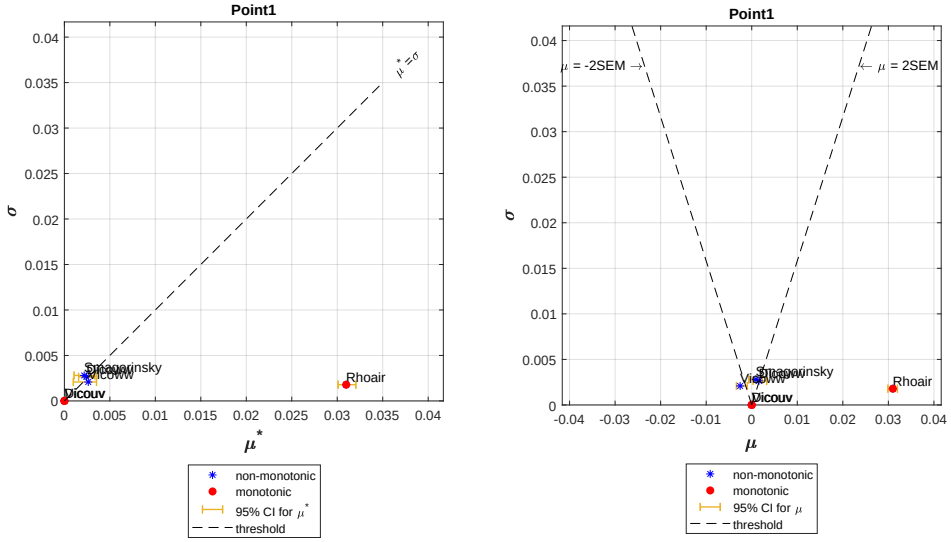


Figure 4.7: Morris method: y-direction velocity's sensitivity indices of FINO3 at VYC

As can be seen above in figure 4.8, Rhoair is far away from other parameters; Vicoww, Dicoww and Smagorinsky are above but also very close to the lines $\mu^* = \sigma$ and $\mu = \pm 2SEM$. The same conclusions can be obtained: the correlations are not significant, and Rhoair is non-zero; except Dicouv and Vicouv, only Rhoair is monotonic, and others are non-monotonic.

The similarity between different timepoints in y-direction is proved.

Table 4.9: Morris method: y-direction velocity's sensitivity indices of FINO3 at VYT

facname	mu*	mu	sigma
Rhoair	0,04127	0,04127	0,01179
Smagorinsky	0,00756	-0,00530	0,00998
Dicoww	0,00608	0,00333	0,00851
Vicoww	0,00573	-0,00418	0,00973
Vicouv	0	0	0
Dicouv	0	0	0

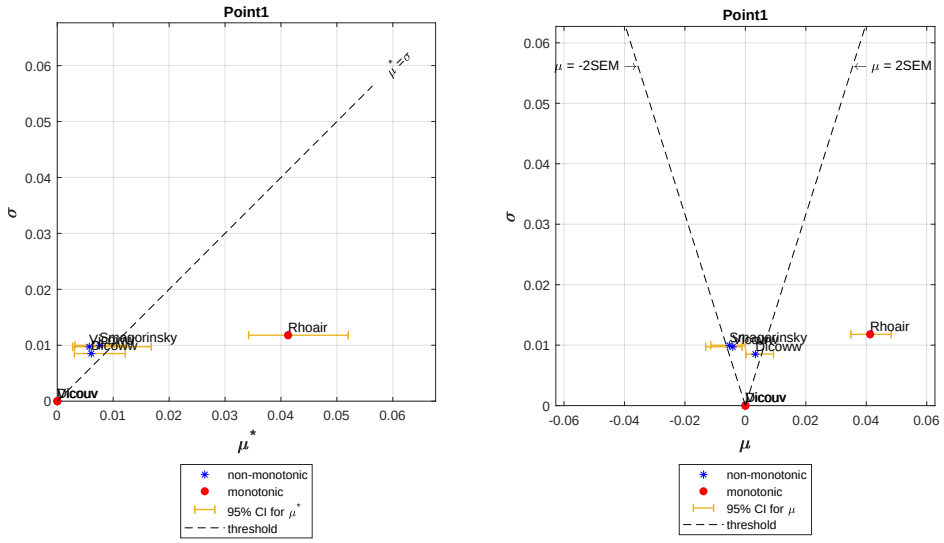


Figure 4.8: Morris method: y-direction velocity's sensitivity indices of FINO3 at VYT

Z-direction

Similarly, the results of four observation points at VZC are presented first, to test the consistency between observation points. Then the results of FINO3 at VZT are shown to prove the similarity between timepoints.

From table 4.10, the ranking of μ^* values are similar with in x- and y-directions: Rhoair is the most influential parameter, Vicoww, Dicoww and Smagorinsky can be seen in a cluster; Vicouv and Dicouv keep zero. Smagorinsky has the highest value of σ, indicating most correlations with others. However, the values of sensitivity indices in z-direction are much smaller than in x- and y-directions, while the values in x- and y-directions are within the same scale. It will be discussed in chapter 5.1.

Note that the scale in the figure is much smaller than in x- and y-directions. Similarly, Rhoair is not in the cluster and located underneath the lines; the μ^* and μ value of Smagorinsky is larger than σ and $2SEM$ more significantly than in x- and y-directions; Vicoww and Dicoww can be seen as in the same cluster; Dicouv and Vicouv are at the origin. Except Dicouv and Vicouv, only Rhoair is monotonic, and others are non-monotonic.

The properties keep the same as the explanations of table 4.10.

In the figure, the properties of parameters are similar to the explanations of the figure 4.9.

It can be concluded that different timepoints show similarities in sensitivity indices.

Table 4.10: Morris method: z-direction velocity's sensitivity indices at VZC

Point1				Point2			
facname	mu*	mu	sigma	facname	mu*	mu	sigma
Rhoair	1,09E-05	1,09E-05	3,21E-06	Rhoair	1,07E-05	1,07E-05	2,55E-06
Vicoww	2,39E-06	-1,3E-06	3,14E-06	Vicoww	3,41E-06	-3,3E-06	3,61E-06
Smagorinsky	1,68E-06	-7,8E-07	1,92E-06	Smagorinsky	1,59E-06	-7,2E-07	1,87E-06
Dicoww	1,2E-06	-3,2E-07	1,95E-06	Dicoww	5,11E-07	5,82E-08	6,1E-07
Vicouv	0	0	0	Vicouv	0	0	0
Dicouv	0	0	0	Dicouv	0	0	0
Point3				Point4			
facname	mu*	mu	sigma	facname	mu*	mu	sigma
Rhoair	9,9E-06	9,9E-06	2,76E-06	Rhoair	9,04E-06	9,04E-06	3,12E-06
Vicoww	3,65E-06	1,48E-06	3,98E-06	Vicoww	2,4E-06	-8,2E-07	3,3E-06
Smagorinsky	2,65E-06	9,93E-07	3,51E-06	Smagorinsky	2,38E-06	-1,3E-06	2,59E-06
Dicoww	1,5E-06	-4,4E-07	2,14E-06	Dicoww	1,06E-06	-3,6E-07	1,29E-06
Vicouv	0	0	0	Vicouv	0	0	0
Dicouv	0	0	0	Dicouv	0	0	0

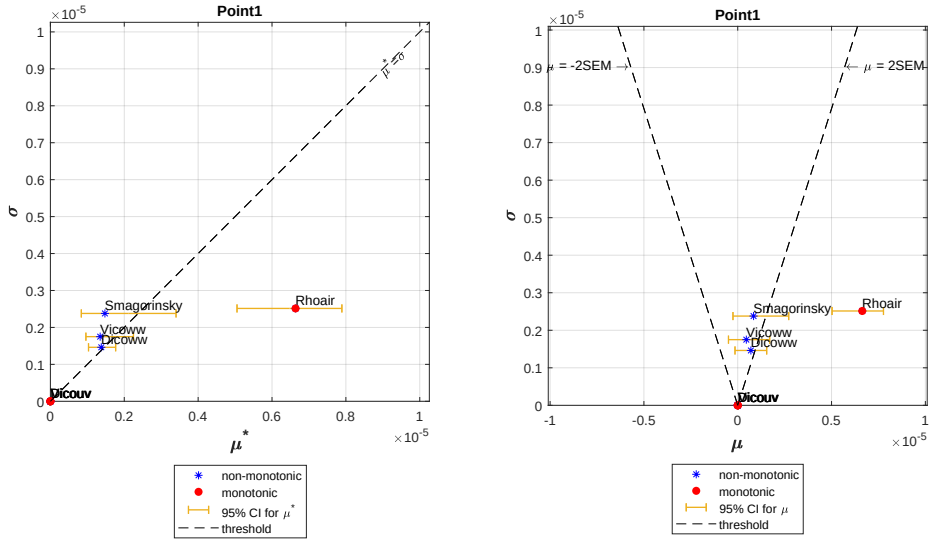


Figure 4.9: Morris method: z-direction velocity's sensitivity indices of FINO3 at VZC

Table 4.11: Morris method: z-direction velocity's sensitivity indices of FINOE3 at VZT

facname	mu*	mu	sigma
Rhoair	1,09E-05	1,09E-05	3,21E-06
Vicoww	2,39E-06	-1,3E-06	3,14E-06
Smagorinsky	1,68E-06	-7,8E-07	1,92E-06
Dicoww	1,2E-06	-3,2E-07	1,95E-06
Vicouv	0	0	0
Dicouv	0	0	0

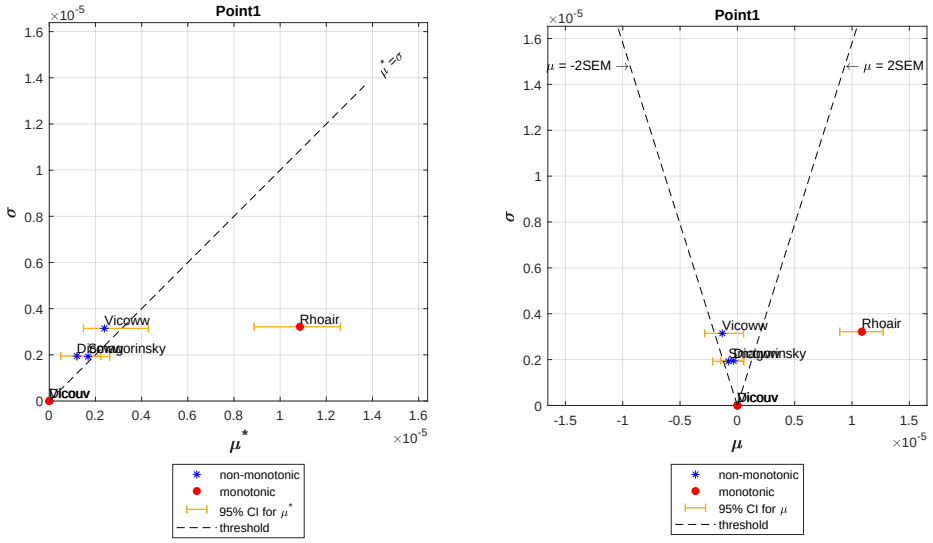


Figure 4.10: Morris method: z-direction velocity's sensitivity indices of FINO3 at VZT

4.2. Copula-based method

In this section, the results of sensitivity analysis will be illustrated in the same structure as Morris method. The results will be separated into temperature and current velocity and the current velocity will be divided into three directions. In each part, the results of different observation points at one timepoint will be presented first, to check if the properties keep the same at different observation points. Then the results of FINO3 at different timepoints will be shown to prove if the results are similar at different time.

4.2.1. Temperature

Table show the results of μ^*, μ and σ of six observation points at the timepoints TC1.

The six observation points all show the same ranking of μ^*: the copula of Stanton and Dalton is the most influential copula, having the highest value of μ^*; Smagorinsky and the copula of Vicoww and Dicoww are the second and third influential copula (parameter); the copula of Vicouv and Dicouv has zero in all the indices, which can be explained from the same reason of being overwritten by ext-file. These results can be seen in every system. To avoid duplication, the analysis of Vicouv and Dicouv will be skipped in the following parts. From σ, the copula of Stanton and Dalton has the most correlation with other parameters.

In the μ^* vs σ plot, the distance between Stanton-Dalton and Smagorinsky are significant, while Dicoww-Vicoww is very close to the origin. Considering the line $\mu^* = \sigma$, the μ^* value of Stanton-Dalton and Smagorinsky are slightly larger than the σ value. In the right plot, Stanton-Dalton and Smagorinsky are within the edge. Two conclusions can be obtained: first, Stanton-Dalton and Smagorinsky have comparatively significant evidence of having correlations with other parameters; second, Stanton-Dalton and Smagorinsky do not have strong evidence of being non-zero. Besides this, all the copulas except Vicouv-Dicouv are non-monotonic.

To prove the similarity of different timepoints, the tables show the results of FINO3 at different timepoints.

The rankings of μ^* at TC2, TT1 and TT2 keep the same as TC1. Stanton-Dalton also shows the most correlation information with the highest σ. Note that the values at TT1 are much smaller than the other three timepoints. It might be caused by the instability when it is too close to the starting points.

The μ^* vs σ and μ vs σ plots of FINO3 at TC2, TT1 and TT2 are as follow:

The figures 4.12 4.13 AND 4.14 all show similar patterns with 4.11. As shown on the left, Stanton-Dalton and Smagorinsky are far away from each other, and the μ^* values are approximately equal to σ values. In the right plots, Stanton-Dalton and Smagorinsky are within the edge. Vicoww-Dicoww is very close to the origin in every plot. We can conclude that Stanton-Dalton and Smagorinsky is correlated with other parameters, but

Table 4.12: Copula-based method: temperature's sensitivity indices at TC1

Point1				Point2			
facname	mu*	mu	sigma	facname	mu*	mu	sigma
Stanton-Dalton	0,22660	0,04470	0,23458	Stanton-Dalton	0,21960	0,04370	0,22724
Smagorinsky	0,02570	-0,00730	0,03468	Smagorinsky	0,00220	-0,00044	0,00400
Vicoww-Dicoww	0,00110	-0,00083	0,00155	Vicoww-Dicoww	0,00061	-0,00018	0,00075
Vicouv-Dicouv	0	0	0	Vicouv-Dicouv	0	0	0
Point3				Point4			
facname	mu*	mu	sigma	facname	mu*	mu	sigma
Stanton-Dalton	0,22030	0,04750	0,22748	Stanton-Dalton	0,20020	0,03820	0,20766
Smagorinsky	0,14180	-0,06190	0,13987	Smagorinsky	0,02720	0,00800	0,02871
Vicoww-Dicoww	0,00290	-0,00180	0,00363	Vicoww-Dicoww	0,00140	-0,00094	0,00196
Vicouv-Dicouv	0	0	0	Vicouv-Dicouv	0	0	0
Point5				Point6			
facname	mu*	mu	sigma	facname	mu*	mu	sigma
Stanton-Dalton	0,22490	0,04410	0,23292	Stanton-Dalton	0,20760	0,04230	0,21462
Smagorinsky	0,00860	-0,00270	0,01223	Smagorinsky	0,08370	-0,03510	0,08349
Vicoww-Dicoww	0,00072	0,00004	0,00087	Vicoww-Dicoww	0,00120	-0,00003	0,00201
Vicouv-Dicouv	0	0	0	Vicouv-Dicouv	0	0	0

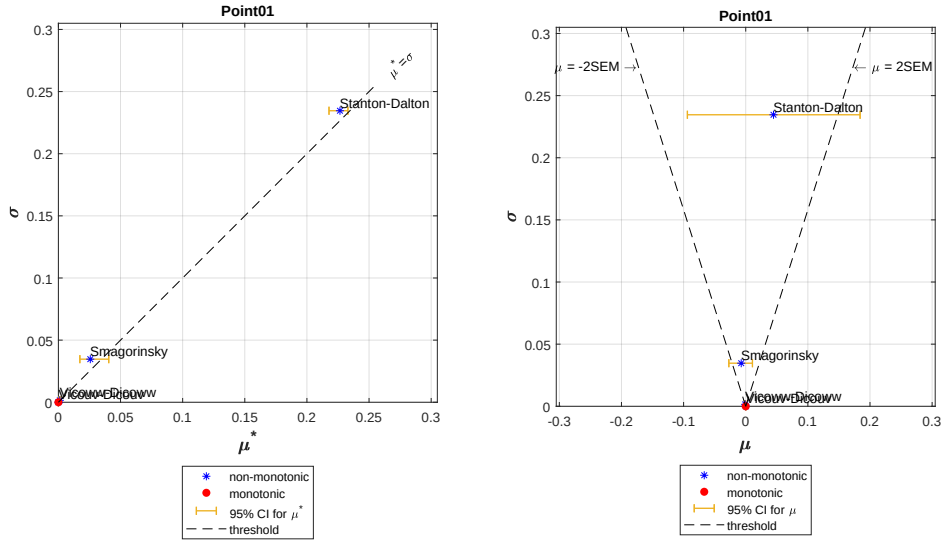


Figure 4.11: Copula-based method: temperature's sensitivity indices of FINO3 at TC1

Table 4.13: Copula-based method: temperature's sensitivity indices of FINO3 at TC2

facname	mu*	mu	sigma
Stanton-Dalton	0,275603	0,052237	0,286039
Smagorinsky	0,049533	-0,01531	0,061593
Vicoww-Dicoww	0,001838	-0,00088	0,003204
Vicouv-Dicouv	0	0	0

Table 4.14: Copula-based method: temperature's sensitivity indices of FINO3 at TT1

facname	mu*	mu	sigma
Stanton-Dalton	0,069726	0,013886	0,072026
Smagorinsky	0,004609	-0,00171	0,004902
Vicoww-Dicoww	0,000184	3,67E-05	0,000257
Vicouv-Dicouv	0	0	0

Table 4.15: Copula-based method: temperature's sensitivity indices of FINO3 at TT2

facname	mu*	mu	sigma
Stanton-Dalton	0,247049	0,051891	0,25522
Smagorinsky	0,126597	-0,05187	0,127966
Vicoww-Dicoww	0,001969	-0,00175	0,003057
Vicouv-Dicouv	0	0	0

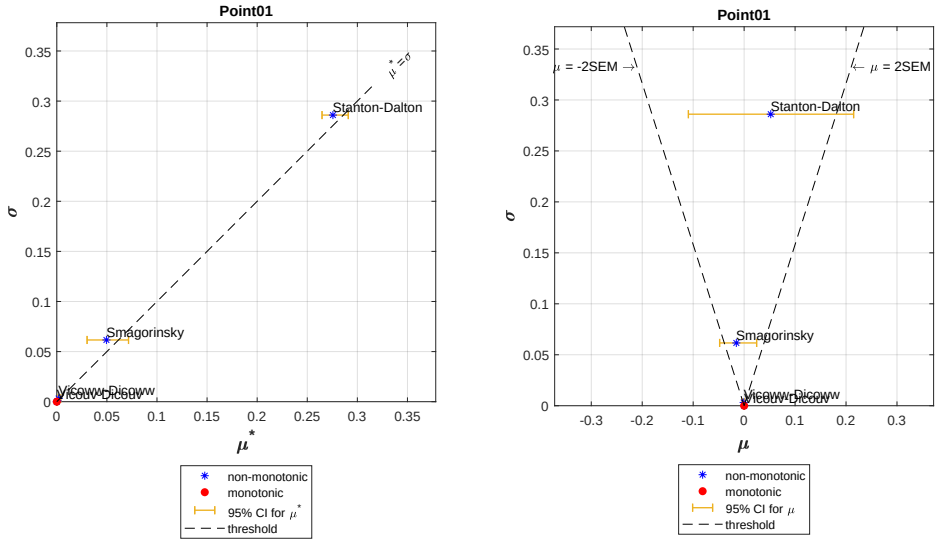


Figure 4.12: Copula-based method: temperature's sensitivity indices of FINO3 at TC2

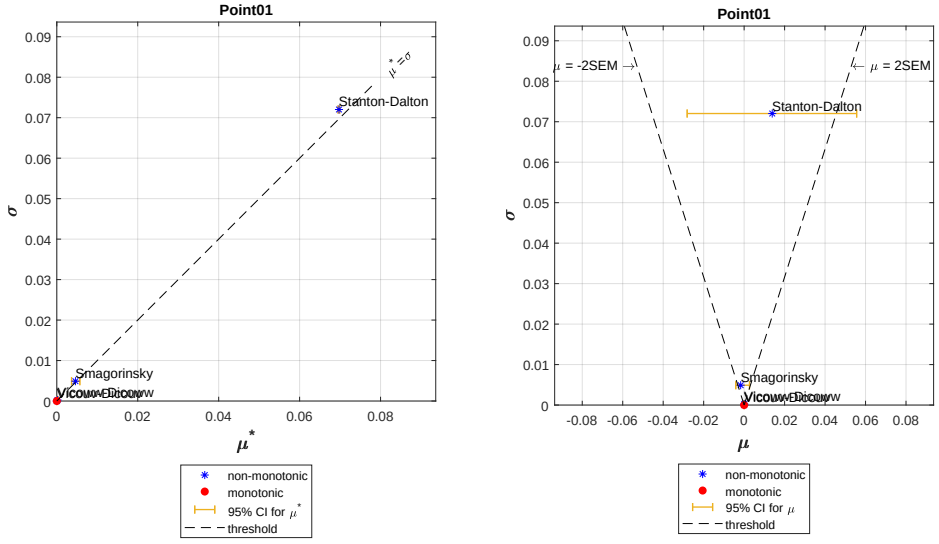


Figure 4.13: Copula-based method: temperature's sensitivity indices of FINO3 at TT1

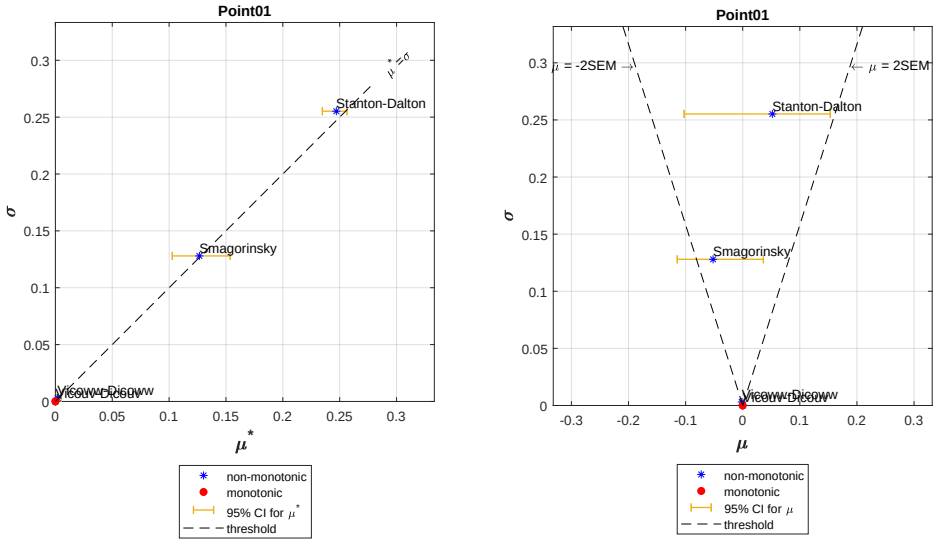


Figure 4.14: Copula-based method: temperature's sensitivity indices of FINO3 at TT2

it is not very significant. Also, the evidence is not strong enough to prove Stanton-Dalton and Smagorinsky non-zero. Besides, except Vicouv-Dicouv, all the copulas (parameters) are non-monotonic.

It can be concluded that the results are similar at different timepoints.

4.2.2. Current velocity

In this section, three directions will be discussed separately.

X-direction

In order to avoid redundancy, only four observation points will be listed in the table.

Table 4.16: Copula-based method: x-direction velocity's sensitivity indices of FINO3 at VXC

Point1				Point2			
facname	mu*	mu	sigma	facname	mu*	mu	sigma
Rhoair	0,089006	-0,01655	0,130262	Rhoair	0,031911	0,003403	0,061886
Vicoww-Dicoww	0,071166	0,008456	0,133734	Vicoww-Dicoww	0,01483	0,002484	0,028163
Smagorinsky	0,036161	-0,00616	0,070105	Smagorinsky	0,008138	-0,0061	0,01341
Vicouv-Dicouv	0	0	0	Vicouv-Dicouv	0	0	0
Point3				Point4			
facname	mu*	mu	sigma	facname	mu*	mu	sigma
Rhoair	0,023324	-2,52E-05	0,02504	Rhoair	0,031515	0,001889	0,03358
Vicoww-Dicoww	0,004208	0,000853	0,006418	Vicoww-Dicoww	0,005652	0,002675	0,009602
Smagorinsky	0,001749	-0,00048	0,002378	Smagorinsky	0,002478	-0,00091	0,003459
Vicouv-Dicouv	0	0	0	Vicouv-Dicouv	0	0	0

The ranking of μ^* values are the same at each observation point. Rhoair is the most influential parameter, having the largest μ^*, Vicoww-Dicoww and Smagorinsky are the second and third influential copula (parameter). Generally, Rhoair has the highest value of σ, indicating that it contains the most correlation information with others.

The sensitivity indices are coordinated in the figure 4.15. In both the left and right plot, Rhoair and Vicoww-Dicoww are close to each other, compared with Smagorinsky. The three copulas/parameters are all located above the lines. It can be concluded that all the three segments have significant correlations with other parameters, and they don't have strong evidence to be non-zero. Besides, the three segments are non-monotonic.

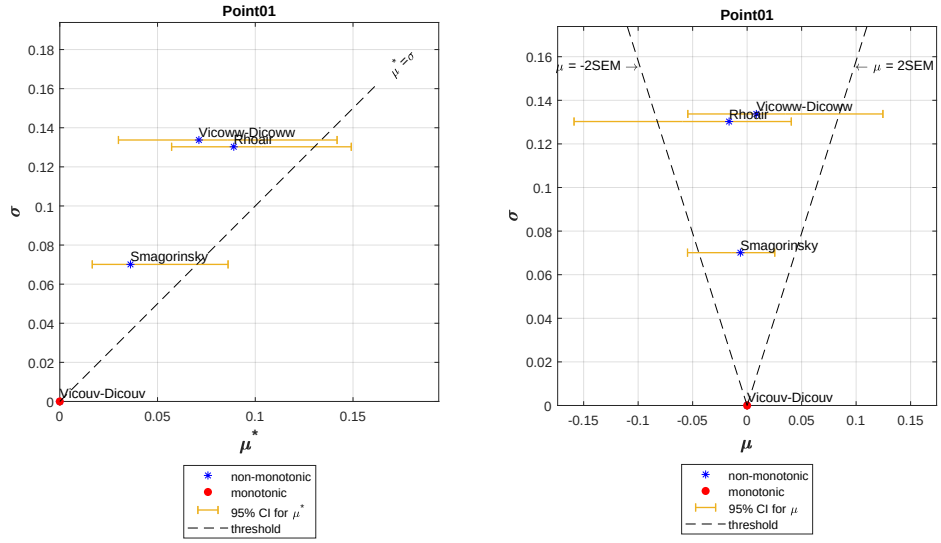


Figure 4.15: Copula-based method: x-direction velocity's sensitivity indices of FINO3 at VXC

The results of FINO3 at timepoint VXT are as follow.

Table 4.17: Copula-based method: x-direction velocity's sensitivity indices of FINO3 at VXT

facname	mu*	mu	sigma
Rhoair	0,027455	-0,00153	0,029899
Smagorinsky	0,002746	0,000163	0,004433
Vicoww-Dicoww	0,002581	-0,00092	0,004029
Vicouv-Dicouv	0	0	0

The ranking of μ^* is the same: Rhoair > Smagorinsky > Vicoww-Dicoww. Also, Rhoair has the largest σ, indicating the most correlation information. The details are the same as at VXC.

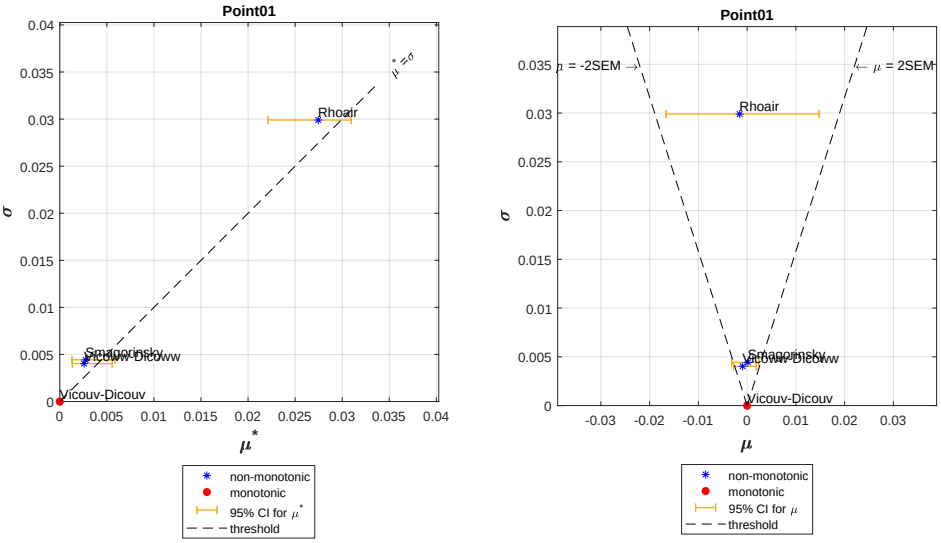


Figure 4.16: Copula-based method: x-direction velocity's sensitivity indices of FINO3 at VXT

In figure 4.16, the distance between Rhoair and the origin is more significant than the distance between Vicoww-Dicoww/Smagorinsky and the origin. Rhoair, Smagorinsky and Dicoww-Vicoww are all above the

lines. The same conclusions can be generated: all the three segments have significant correlations with other parameters, and they don't have strong evidence to be non-zero. Besides, the three segments are non-monotonic.

The similarity between different timepoints in x-direction can be proved.

Y-direction

The same procedures as in x-direction will be applied to y- and z-directions. The table show the results of 4 observation points at VYC.

Table 4.18: Copula-based method: y-direction velocity's sensitivity indices at VYC

Point1				Point2			
facname	mu*	mu	sigma	facname	mu*	mu	sigma
Rhoair	0,028365	-0,00029	0,030891	Rhoair	0,04836	-0,00158	0,055153
Vicoww-Dicoww	0,005691	-0,0012	0,009712	Vicoww-Dicoww	0,011492	0,003091	0,01894
Smagorinsky	0,002366	0,000526	0,004609	Smagorinsky	0,003784	0,002019	0,007049
Vicouv-Dicouv	0	0	0	Vicouv-Dicouv	0	0	0
Point3				Point4			
facname	mu*	mu	sigma	facname	mu*	mu	sigma
Rhoair	0,013604	-7,23E-05	0,014411	Rhoair	0,017925	0,000991	0,018991
Smagorinsky	0,001977	0,000859	0,002406	Smagorinsky	0,001799	0,000417	0,002079
Vicoww-Dicoww	0,001703	0,000809	0,001815	Vicoww-Dicoww	0,001425	-0,00092	0,0016
Vicouv-Dicouv	0	0	0	Vicouv-Dicouv	0	0	0

The ranking of μ^* values is a little different on the orders of Vicoww-Dicoww and Smagorinsky. At observation point 1 and 2, Vicoww-Dicoww is more influential than Smagorinsky; at observation point 3 and 4, the orders are exchanged. The differences can be tolerated as the values are close. Rhoair is the most influential parameter, and also the parameter having the most correlation information, which is the same as in x-direction.

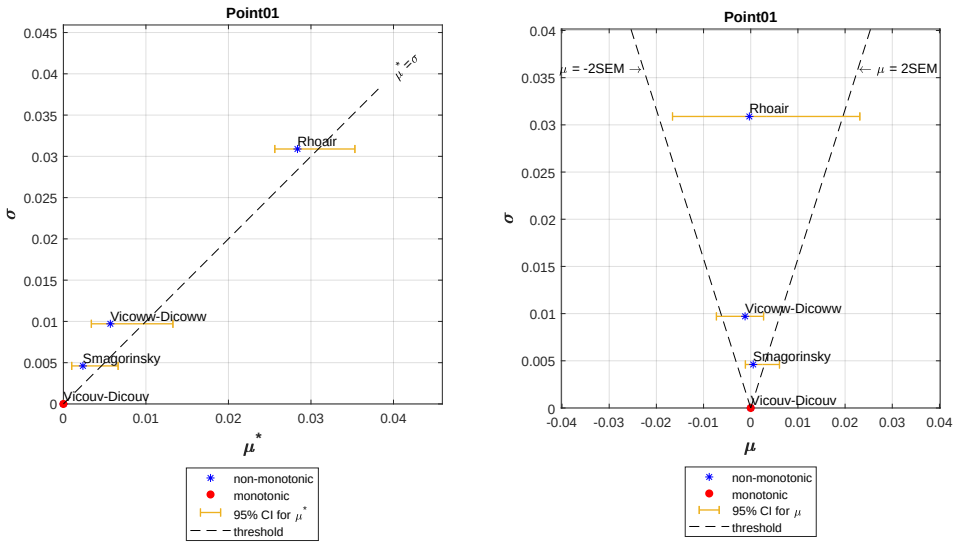


Figure 4.17: Copula-based method: y-direction velocity's sensitivity indices of FINO3 at VYC

The figure 4.17 has the similar patterns as figure 4.16, but Rhoair, Vicoww-Dicoww and Smagorinsky are more far away from each other compared to figure 4.16. The conclusions can be generated similarly: all the three segments have significant correlations with other parameters, and they don't have strong evidence to be non-zero. Besides, the three segments are non-monotonic.

It can be concluded that the results at different observation points are similar.

The results of FINO3 at timepoint VYT are as follow.

Table 4.19: Copula-based method: y-direction velocity's sensitivity indices of FINO3 at VYT

facname	mu*	mu	sigma
Rhoair	0,039005	0,001252	0,05227
Vicoww-Dicoww	0,023641	-0,00351	0,047239
Smagorinsky	0,013843	0,002952	0,02965
Vicouv-Dicouv	0	0	0

The ranking of μ^* is the same as FINO3 at VYC. The sigma value of Rhoair is the highest.

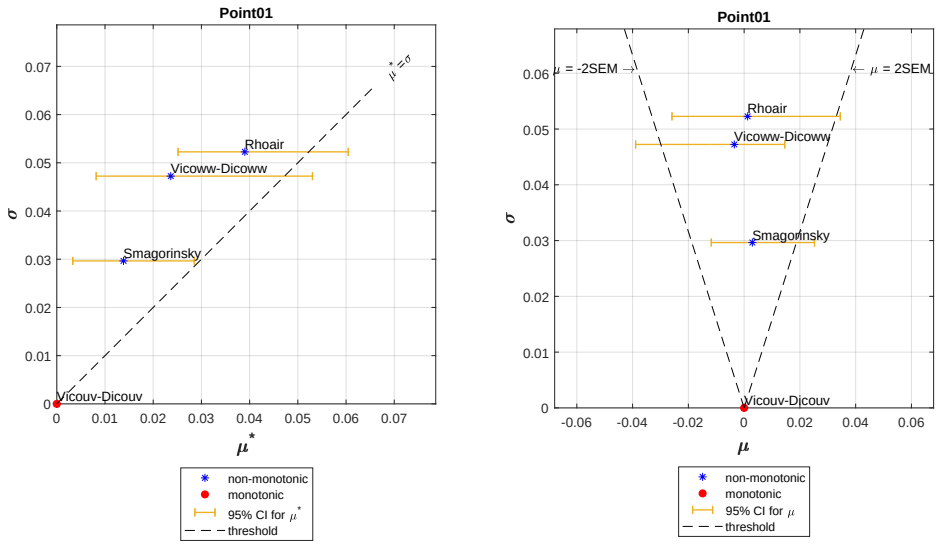


Figure 4.18: Copula-based method: y-direction velocity's sensitivity indices of FINO3 at VYT

Figure 4.18 shows the results of FINO3 at VYT, similar to figure VYC.

The similarity between different timepoints in x direction can be proved.

Z-direction

Table 4.20 shows the results of four observation points at VZC.

Table 4.20: Copula-based method: z-direction velocity's sensitivity indices at VZC

Point1				Point2			
facname	mu*	mu	sigma	facname	mu*	mu	sigma
Rhoair	7,86E-06	1,37E-06	9,12E-06	Rhoair	4,09E-06	-2,3E-06	4,6E-06
Vicoww-Dicoww	2,19E-06	-2E-07	3,33E-06	Smagorinsky	3,46E-06	3,46E-06	2,64E-06
Smagorinsky	1,81E-06	9,18E-07	2,16E-06	Vicoww-Dicoww	2,04E-06	-1,1E-06	2,46E-06
Vicouv-Dicouv	0	0	0	Vicouv-Dicouv	0	0	0
Point3				Point4			
facname	mu*	mu	sigma	facname	mu*	mu	sigma
Rhoair	6,22E-06	-2E-06	7,32E-06	Rhoair	8,84E-06	-6,1E-06	8,11E-06
Vicoww-Dicoww	2,57E-06	7,75E-07	3,47E-06	Vicoww-Dicoww	3,49E-06	1,06E-06	5,16E-06
Smagorinsky	1,69E-06	-3,7E-07	2,22E-06	Smagorinsky	2,6E-06	2,5E-06	2,86E-06
Vicouv-Dicouv	0	0	0	Vicouv-Dicouv	0	0	0

The ranking and the highest sigma generally keep the same as in x and y directions. However, the values are much smaller than in x and y directions, which also appears in Morris method.

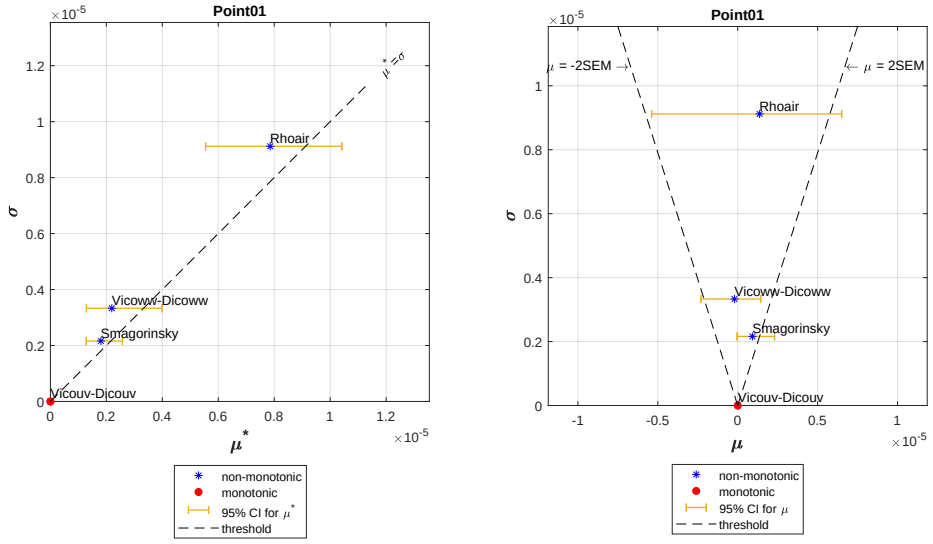


Figure 4.19: Copula-based method: z-direction velocity's sensitivity indices of FINO3 at VZC

The properties of figure 4.19 are similar to VYC, while the scale is much smaller.

Table 4.21: Copula-based method: z-direction velocity's sensitivity indices of FINO3 at VZT

facname	mu*	mu	sigma
Rhoair	8,45E-06	1,38E-07	9,54E-06
Vicoww-Dicoww	5,3E-06	-3,5E-06	6,4E-06
Smagorinsky	2,27E-06	1,12E-06	3,19E-06
Vicouv-Dicouv	0	0	0

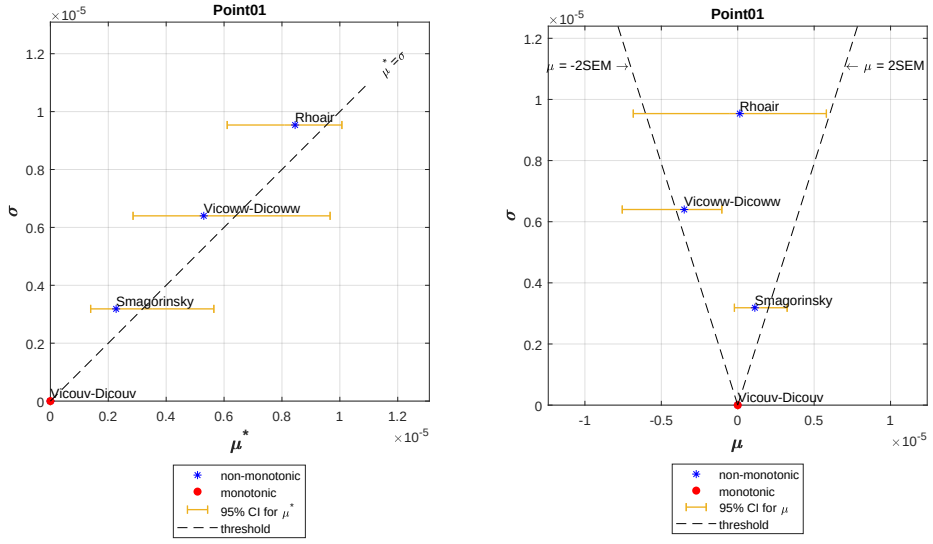


Figure 4.20: Copula-based method: z-direction velocity's sensitivity indices of FINO3 at VZT

The table and the figure are all similar to VZC.

Thus, in z-direction, the similarities between different observation points and timepoints are also proved.

Table 4.22: Variance-based method: first-order and total-order sensitivity indices for temperature

Factor	First-order indices	Total-order indices	Differences (total - first)
Dalton	0.9215	0.5157	-0,4058
Smagorinsky	0.4734	-0.2091	-0,6825
Stanton	0.0926	-0.2159	-0,3085

4.3. Variance-based method

In this section, according to chapter 3.6.2, only one timepoint will be selected to do sensitivity analysis for temperature and three directions in current velocity. First-order indices and total-order indices of selected parameters are presented as results. As illustrated in chapter 2.5.2, the first-order sensitivity indices indicate the main effect contribution of each input, where dependency information are not included. Two conclusions can be obtained from total-order sensitivity indices. First, important interaction involving the factors can be concluded if the differences between first-order and total-order sensitivity indices are significant. Second, the factors can be considered to have no significant influences on the outputs if the total-order indices $S_{Ti} \simeq 0$. However, note that the standards of 'significant differences' and 'approximately equal to zero' are vague, and the conclusions should be considered in real cases.

4.3.1. Temperature

Dalton, Stanton and Smagorinsky are selected for temperature.

From the first-order indices, we can get the information of influence degrees. Dalton is the most influential factor and Stanton is the least influential factor. However, the first-order indices do not include the influences of combinations involving the corresponding factors.

The correlation information can be generated from the total-order indices. To evaluate if the differences between total-order and first-order indices are significant, the values of first-order indices are used as references. As the differences and the first-order indices are in the same order of magnitude, the differences are evaluated as significant. Thus, important interactions are involved in the three factors.

The total-order sensitivity indices of the three parameters are not approximately equal to zero. Therefore, all the three factors cannot be neglected.

In Morris and copula-based method, the rankings of correlation involved in parameters can be concluded by comparing the values of sigma. However, in variance-based method, the values of differences are only used to evaluate if the interactions are important. The rankings of correlations are not considered in the reference [42]. It will be discussed in chapter 5.1.

4.3.2. Current velocity

The results of first-order and total-order indices in three directions are listed below (x, y, z). The timepoints chosen for x, y and z are respectively VXC, VYC, VZC.

Table 4.23: Variance-based method: first-order and total-order sensitivity indices for x-direction velocity

Factor	First-order indices	Total-order indices	Differences (total - first)
Vicoww	-1.0000	2.1257	3,1257
Dicoww	-1.0954	2.2195	3,3149
Smagorinsky	-1.2016	2.3364	3,5380
Rhoair	1.4164	1.6710	0,2546

The value smaller than zero means the effects of the parameters to the outputs are negative. To evaluate the influences, we will consider the absolute value, no matter the influences are positive or negative.

Considering the first-order indices, in x- and y-direction, Rhoair is the most influential parameter. Smagorinsky is the most influential parameter in z-direction.

Table 4.24: Variance-based method: first-order and total-order sensitivity indices for y-direction velocity

Factor	First-order indices	Total-order indices	Differences (total - first)
Vicoww	-0.6196	1.3177	1,9373
Dicoww	-0.6781	1.3761	2,0542
Smagorinsky	-0.7157	1.4150	2,1307
Rhoair	0.8404	1.8118	0,9714

Table 4.25: Variance-based method: first-order and total-order sensitivity indices for z-direction velocity

Factor	First-order indices	Total-order indices	Differences (total - first)
Vicoww	14.8289	-14.1139	-28,9428
Dicoww	-5.4503	5.9601	11,4104
Smagorinsky	17.6466	-17.0058	-34,6524
Rhoair	-5.6954	8.1310	13,8264

Considering the differences between the total-order and the first-order indices, the values can not be considered as unsignificant compared to the values of first-order indices. Therefore, the parameters cannot be considered unrelated to others. As the total-order is significantly not equal to zero, all the parameters cannot be neglected.

Note that the values in z direction are 'unnormal', while the values in x- and y-directions are in the same order of magnitude. It can probably be explained by the simulation results. In x- and y-directions, the values of simulation outputs are between 0 and 1, while the current velocity in z- direction (in vertical direction) is very small, around $0,0003$. The calculation formulas in chapter 2.5.5 may lead to the big differences using the small values in variance.

5

Discussions, conclusions and recommendations

5.1. Discussions

In this section, some assumptions and settings are mentioned first, which may vary in other possible cases. Then the results are discussed. Last the limitations of different methods and the applications are suggested.

Assumptions

In this thesis project, all the data are generated from Delft 3D. To keep consistency in comparisons, the simulation periods are set the same from 2004-12-25 02:24:00 to 2005-01-25 02:24:00, lasting for one month, with the same time intervals. All other parameters, influential factors and external files, besides the target parameters studied, are set the same. The same observation points are chosen (see chapter 3.2). The simulations are assumed to have no errors.

The data of temperature in Variance-based method (chapter 4.3.1) is offered by the colleagues from Deltares. The timepoint of the simulation is lost, but the results is assumed to apply to all timepoints, as in the first two methods, Morris and copula-based method, different timepoints show similarities in sensitivity results.

Also, in chapter 4, four timepoints are selected for temperature and two are selected for each direction in current velocity, then the temporal consistency is concluded from the similarities of chosen timepoints. Avoiding the repeat and reducing calculation costs, the conclusion from the limited timepoints is assumed to be convincible.

Discussions of results

In the results of Morris and copula-based method, the parameters Vicouv and Dicouv (horizontal viscosity and diffusivity) show no impacts on both temperature and current velocity, while Vicoww and Dicoww (vertical viscosity and diffusivity) have influences on outputs. The reasons are briefly explained in chapter 4.1.1, that the values in horizontal directions in mdu-files are overwritten by the values in the external files. In the details, if spatially varying values of horizontal viscosity/diffusivity are prescribed in the external files, the uniform values specified in mdu-files are not used. The spatial values are prescribed to be 0.1 m^2/s everywhere, except a small band along the open boundaries. In the horizontal direction, the total viscosity/diffusivity is a combination of the specific background values and the values computed with Smagorinsky.

In the vertical direction, the viscosity/diffusivity is computed with a k-eps model. A background value can also be specified in mdu-file. However, the background values will not have any contributions if the background value is lower than what is computed with the k-eps model. In the tidal systems, the vertical viscosity/diffusivity values computed with the k-eps model are expected to be usually higher than the prescribed background values in mdu-files. But it is not always the case, which can be seen from the non-zero results of Vicoww/Dicoww. The resulting vertical viscosity/diffusivity in the simulations can be checked and compared with the specific background values to prove the assumption.

In chapter 3.5.3, some simulations crashed when the ranges of the parameters started from zero. It might result from viscosity. In the model, viscosity tends to stabilize the numerical solution. When the unrealistically low value of zero was used, the problems might arise from instability. The zero value of viscosity might result from Smagorinsky being zero.

In the results of Morris and copula-based method, the values of sensitivity indices in z-direction are much smaller than in x- and y-directions. For example, in Morris method, mu* value of Rhoair at FINO3 in x-direction is about 0.09, while the value is about 0.00001. Flows on these scales are strongly anisotropic. Velocities and viscosities in vertical (z-) direction are much lower than in horizontal (x- and y-) directions. It can also cause the 'unnormal' variance-based method results in z-direction.

Limitations and application of different methods

In the thesis projects, three different methods are used to conduct sensitivity analysis: Morris, copula-based and variance-based methods. They can be applied in different cases according to the properties of parameters. In other words, limitations of the methods should be considered.

First, dependency information is not included in Morris method. Correlation can be interpreted from the sensitivity index sigma. However, bias and unreasonableness may occur in the results, as Morris method assume every parameter independent before sampling, and the samples may not conform to the physical process. While copula-based and variance-based include the independency information within parameters.

Second, in copula-based method, the correlations need to be described roughly before sampling. As mentioned in the first point, both copula-based and variance-based method take independency into consideration, but the way of including the information is different. In variance-based method, the dependency is concluded from the variance, and it does not need to be considered in sampling. While in copula-based method, the correlations need to be decided before generating the samples (the details can be found in chapter 2.4). The relations between parameters are not clear in many cases, especially when there are various parameters in a complex physical process. In addition, the correlations might be difficult to measure. For example, in the case of this thesis project, the parameters are defined roughly in the copulas of two parameters, and the correlations are only decided as 1 or -1, indicating the positive or negative relations (see chapter 3.6.1). As illustrated in chapter 3.6.1, the vague decisions are not the only possibilities. While this limitation increases the complexity of operating sensitivity analysis, it might also cause uncertainty in the results.

Third, the calculation costs of variance-based method are very high. In Morris method and copula-based method, the number of samples is $(k+1)*r$, where k is the number of segments (including copulas and parameters) and r is the number of paths. In variance-based method, the number of samples is $(k+2)*N$, where k is the number of parameters and N is the number of random samples in each matrix, varying from hundreds to thousands. Take temperature in this thesis project as an output example: there are 80 samples in Morris method with seven parameters, 50 samples in copula-based method with four copulas and single parameters, while variance-based method requests at least 500 samples with only three parameters.

Fourth, variance-based method does not compare the correlation information between different parameters, but only makes the judgements whether the parameter is strongly correlated with other parameters. In other words, variance-based method is suitable for factor fixting setting (FF), to identify uninfluential factors. For example, through Morris method, Smagorinsky contains the most correlation information in temperature, while the parameters are only concluded to have important interactions with other factors in variance-based method. It may be able to be explained from the strategies of different methods. In Morris and copula-based method, other parameters are fixed when the elementary effects of one parameter are calculated. They correlations between the parameter and all other parameters are considered together, which means the correlations are not distinguished between specific parameters. In variance-based method, the interaction effects are divided into different combinations. Take a three-factor model as an example, $S_{T1} - S_1 = S_{12} + S_{13} + S_{123} = 1 + 2 + 1 = 4$, indicating all the interaction effects of X1. Similarly, we can have $S_{T2} - S_2 = S_{12} + S_{23} + S_{123} = 1 + (-3) + 1 = -1$. The values are chosen as a simple example. As different combinations may have different signs, it cannot be simply concluded that X_1 is more correlated than X_2 according to values of the differences.

In addition, the standards in variance-based method are vague. Significant differences between the total-order and first-order sensitivity indices are evidence for important interaction involving the factors. The total-order sensitivity indices approximately equal to zero indicates the factor having no significant influences on

the outputs. However, it is hard to define 'significant differences' and 'approximately equal to zero'. In this thesis project, first-order indices are used as a reference to evaluate the magnitude.

After comparing the limitation of the three methods, some suggestions are given to decide which method to use in different cases.

If the dependency correlations are not significant in the case, Morris method is suggested to use. If correlations cannot be ignored, copula-based and variance-based method is preferred. Considering the heavy simulation and calculation costs, with the vague standards in variance-based method, copula-based method is better to choose, if the relations between the parameters do not need to be defined strictly. Otherwise variance-based method can be used.

5.2. Summary and conclusions

In this section, first the summary of the results will be presented. Then answers to the questions in chapter 1.3 will be given.

Summary

Three methods are adopted to conduct sensitivity analysis on the hydrodynamic model in this thesis project: Morris, copula-based and variance-based method. Temperature and current velocity in three directions are selected as outputs, and Delft3D is used to simulated the models. In Morris method, 80 and 70 samples are generated respectively for temperature and current velocity; 50 samples are generated respectively in copula-based method; 500 and 768 models are generated respectively in variance-based method. 1518 models were simulated in total, and the simulations took about one month to run. In the simulations, parameters are changed, subfolders and shell files are created to submit the models to the Linux cluster. Besides codes from packages are used for sensitivity analysis, some codes are created to check the results. As sensitivity results, 120 tables and 120 plots are created, including 6 observation points, each one corresponding to one point.As similarities show in chapter 4, the data are not listed in the report.

In Morris and copula-based method, different observation points and timepoints are selected. The parameters show temporal and spatial similarities in sensitivities. In other words, the sensitivity indices show similar properties, such as values and rankings, at different observation points and timepoints. But the values of sensitivity indices are smaller when the timepoints are close to the start point, as the system is not stable enough at the beginning.

As the sensitivity indices are the same in Morris and copula-based method, the evaluations using the two methods can be compared as follows, divided into temperature and current velocity.

Table 5.1: Comparisons between Morris and copula-based method on temperature results

Temperature	Morris method	Copula-based method
Rankings of influences	Dalton >Stanton, Smagorinsky >Vicoww, Dicoww >Vicouv, Dicouv	Stanton-Dalton >Smagorinsky >Vicoww-Dicoww >Vicouv-Dicouv
Correlation information	Smagorinsky contains the most correlation information with other parameters	Stanton-Dalton is the most correlated copula
	The correlations are not significant in all parameters	Stanton-Dalton is significantly correlated, while other don't have significant information
Expectations of parameters' distributions	Dalton and Smagorinsky are non-zero	Stanton-Dalton and Smagorinsky can't be proved non-zero significantly
Monotonicity	All parameters are monotonic	All parameters are non-monotonic

From the table, the rankings are similar, while some differences appear in other aspects. As the difference between the values of sigma is small, the most correlated parameter being different can be tolerated. The significance of correlations and expectations of parameters' distributions are concluded by comparing mu, mu* and sigma (the details can be seen in chapter 3.5.3), but the standards are not absolute, and the conclusions

can be vague if the results are similar to the equations in chapter 3.5.3. Thus, the differences of correlations and expectations in table 5.1 can be tolerated. The bias can occur in simulations, which might lead to the differences in monotonicity. As this thesis project focuses on sensitivity, rankings and correlations are paid more attention on.

In current velocity, the properties in three directions are similar, thus they are combined in the following table. However, note that the values in z-direction are much smaller than in x- and y-direction. It is discussed above in chapter 5.1.

Table 5.2: Comparisons between Morris and copula-based method on current velocity results

Temperature	Morris method	Copula-based method
Rankings of influences	Rhoair >Vicoww, Dicoww, Smagorinsky >Vicouv, Dicouv	Rhoair >Vicoww-Dicoww >Smagorinsky >Vicouv-Dicouv
Correlation information	Rhoair contains the most correlation information with other parameters	Rhoair is the most correlated parameter
	The correlations are not significant in all parameters	The correlations are significant in all parameters
Expectations of parameters' distributions	Rhoair is significantly non-zero	All parameters can't be proved non-zero significantly
Monotonicity	Rhoair is monotonic; Vicoww, Dicoww, Smagorinsky non-monotonic	All parameters are non-monotonic

The rankings and the most correlated parameter are similar, while there are some differences in other aspects. The significance of correlations and expectations are also concluded by comparing different sensitivity indices. As discussed in temperature, the standards are not absolute, possibly leading to vague conclusions. Thus, the differences between Morris and copula-based method in correlation information and expectations of parameters' distributions can be tolerated. The differences in monotonicity can be caused by the bias in simulations.

Variance-based method uses different sensitivity indices. Different higher-order sensitivity indices can be possible evaluations in this method, but in this thesis project, only first-order and total-order sensitivity indices are calculated. The main effect contribution of parameters can be interpreted from the first-order sensitivity indices. Two conclusions can be obtained from the total-order indices. First, important interaction involving in the factor can be concluded from significant differences between the total-order and first-order sensitivities. Second, if the total-order index is approximately equal to zero, the factor can be considered to have no influences on the outputs and neglected, which is factor fixing setting (FF) (see chapter 2.5.4).

Table 5.3: Comparisons of variance-based method on temperature and current velocity results

	Temperature	x-/y-direction velocity	z-direction velocity
Ranking of influences	Dalton >Smagorinsky >Stanton	Rhoair >Smagorinsky >Vicoww, Dicoww	Smagorinsky >Vicoww >Dicoww, Rhoair
Correlation information	Correlations exist in each parameter	Correlations exist in each parameter	Correlations exist in each parameter
	No parameter can be neglected	No parameter can be neglected	No parameter can be neglected

In the ranking of influences, the most influential factors of temperature and x-/y-direction current velocity keep the same as Morris and copula-based method. As the results in z-direction is biased due to the small values of outputs, the differences in z-direction between variance-based method and other two methods can be tolerated.

The correlation information generated from variance-based method is different from Morris and copula-based method. Considering the limitation of vague standards in variance-based method and, the correlation

results of variance-based method are not suggested compared to copula-based method.

Conclusions

Some questions are listed in chapter 1.3. To conclude the report, the answers corresponding to the questions will be presented.

- Sensitivity measurement: which indices can be selected to measure the sensitivity of parameters and how to explain them?

  In Morris and copula-based method, the elementary effects are used to calculate elementary mean, absolute elementary mean and standard deviation. The absolute mean gives the information of influence degree and standard deviation gives the correlation information. In variance-based method, the first-order and total-order effect indices are used. First-order effect indices gives the information of influence degree; the differences between first-order and total-order effect indices show if the significant correlation exists; the total-order effect indices not being zero indicated the parameter's influences cannot be neglected.

- Research prioritization: which factors are the most deserving of further analysis or measurement?

  It can be answered from the tables above.

- Model simplification: can some factors or compartments of the model be fixed or simplified?

  Dicouv and Vicouv can be abandoned in sensitivity analysis, as the parameters changed will be overwritten, thus have no influence on outputs.

- In the physical process, independence is rare to see, and parameters can be imagined to be dependent and interact with each other, which is also the highlights of the thesis. Some questions related to dependencies and correlations need to be answered:

  * Why are the dependencies of the parameters needed to be considered?

    Because independence is rare to see in physical process. In this thesis project, the parameters have correlations between each other. Before creating copulas in chapter 3.6.1, some examples are explained.

  * How to include the dependency information into sensitivity analysis?

    In this thesis project, copula-based and variance-based method are used to include the dependency information. In copula-based method, the dependency is included in the copulas to generate the sample, making the samples more reasonable. In variance-based method, the dependency is included in the variance, specifically embodied in first-order and higher-order indices, including the total-order indices.

  * Which parameters are correlated together?

    It can be answered in copula-based method. For temperature, Stanton and Dalton, Vicoww and Dicoww, Vicouv and Dicouv are correlated together; for current velocity, Vicoww and Dicoww, Vicouv and Dicouv are correlated together. However, note that the correlations are not interpreted from the results, but are from the experience and physical processes. It is difficult for indicators to specifically indicate which variables are relevant. It can be obtained generally from higher-order indices in variance-based method, for example second-order indices for specific parameters can give us a cognition of how big the correlation is. The computation tasks are heavy for this, which can be done in further study.

- From the aspects of different methods, what are the differences between the logic and results of the methods?

  The answers to the questions can also be found in the limitations in the discussions. First, Morris method assume every parameter independent before sampling, which may cause the bias and unreasonableness in results, while copula-based and variance-based method include the independency information within parameters.

Second, Morris method and variance-based method use the same sensitivity indices: elementary mean, absolute mean, which indicating the influence degree, and standard deviation indicating the correlation degree. Variance-based method uses first-order indices to indicate the influential degree, and higher orders including the total-order indices to show the correlations.

Third, the dependency information is included differently in copula-based and variance-based methods. In copula-based method, the correlations need to be decided roughly before sampling, which may be unclear in many cases. In variance-based method, the correlations do not need to be considered manually in operations. Also, the correlations are hard to define in copula-based method. In this thesis project, the correlation is only considered as 1 or -1, which is very rough.

Fourth, the calculation time is different between the methods. In Morris method and copula-based method, the number of samples is (k+1)*r, where k is the number of segments (including copulas and parameters) and r is the number of paths. In variance-based method, the number of samples is (k+2)*N, where k is the number of parameters and N is the number of random samples in each matrix, varying from hundreds to thousands. The calculation costs of variance-based method is much higher than Morris method and copula-based method, which is a disadvantage of variance-based method.

Last, Morris and copula-based method can rank the correlation information of parameters, in other words, which parameters are correlated more with others. However, variance-based method is mainly used to decide if the parameters are involved in significant interactions.

5.3. Recommendations

In this section, some recommendations for future work will be given.

First, the codes of Morris method use the files created by SimLab. However, the software is not updated any more. The codes not relying on SimLab are suggested to investigate.

Second, the settings of copulas in copula-based method are just one rough possibility. In this thesis project, only two parameters are included in one copula. But in other cases, it is possible that more than 2 parameters are concluded together. The logic and steps of the general idea are the same, but it will be a difficult task to adjust the codes to extend the copulas. Besides the number of parameters in the copulas, more possible correlations can be tried. The correlation of 1 or -1 is very rough are unrealistic in most physical processes. The values between -1 and 1 can be used in further study.

Third, in variance-based method, more higher-order indices can be tried. In this thesis project, only the first-order and the total-order indices are calculated. The first-order indices are used to show the influence degree, and the differences of the first-order and the total-order indices are used to indicate the correlation degree of the parameter. As mentioned in the last point of the answers to the fourth questions in conclusions, higher-order indices can show the correlations between specific parameters. Saltelli [42] didn't give the calculation strategy for the higher-order indeices, it is worth investigating the calculations and codes based on the same settings for the first-order and the total-order indices.

Fourth, in variance-based method, the approximation error are produced in calculations. A method is provided by Sobol' et al. to estimate the approximation error when fixing unessential factors in global sensitivity analysis [49]. It is suggested to apply the method in the model.

Different time periods, input parameters and outputs in the model are suggested to try.

What's more, more methods of sensitivity analysis can be tried, besides the three methods used in this thesis project.

A

Codes: pretreatments

Change files containing observation point coordinate

```
% change the files in mdu-file containing the information of observation points.
% capable to other files need to be changed in mdu-files
new_obs = 'ObsFile                            = /p/11205216-united/04-MSc_students/Ni_Ye
    /DCSM_4nm_current/3D-DCSM-FM_4nm_20191129_obs.xyn  # Points file *.xyn with
    observation stations with rows x, y, station name';
for k =  1:70;
    old_file = ['DCSM_4nm_current_' num2str(k.','%04d') '/DCSM_4nm_current.mdu'];
    new_file = ['DCSM_4nm_current_' num2str(k.','%04d') '/DCSM_4nm_current.mdu'];
    text_modify_wholeline(old_file,new_file,new_obs,259);
    % replace line 259 with 'new_obs' in old_file, saved as 'new_file'
end
```

```
function text_modify(filename_ini,filename_new,new_contents,line)
% read the file 'filename_ini', replace the whole 'line' with 'new_contents', output
    file is 'filename_new'

%% read the file and adjust
 fileID = fopen(filename_ini,'r+');
 % open the file in readable and writable way
 i=0;
 while ~feof(fileID)
        tline=fgetl(fileID);
        % read the original file line by line
        i=i+1;
        newline{i} = tline;
        % create new lines to accpect original file's contents in each line
        if i==line
            % determine whether the row to be modified is reached
            newline{i} = strrep(tline,tline,new_contents);
            % strrep: replace 'old' with 'new_contents'
        end
    end
    fclose(fileID);
    % close the file

%% export the file
    fileID = fopen(filename_new,'w+');
    % open the output file in readable and writable way.
    % if the file exists, clear the file contents and write from the beginning;
    % if not, create the new file with the name and open only writable
    for k=1:i
        fprintf(fileID,'%s\t\n',newline{k});
        % write newline line by line
    end
    fclose(fileID);
    % close the file
```

```
end
```

Change parameters in mdu-file

```
% after having the samples, generate the mdu-files in different sub-folders
% to run the models (example: copula-based method: temperature)
samples = readmatrix('parameters_temperature_simulations.xlsx','Range','A1:G50');
% rows: number of samples; columns: number of parameters;
line = [157,158,159,160,163,190,191];          % the lines of different parameters
for m = 1:50                                   % m: the number of samples
    output_filename = ['temperature_copula_' num2str(m.','%04d') '/
        DCSM_temperature_copula.mdu'];
    % the mdu-files are located in sub-folders
    text_modify('DCSM-FM_4nm_adjusted.mdu',output_filename,num2str(samples(m,1)),line
        (1,1));
    % replace the parameters' values of line m in 'output_filename' with
    % 'line(1,1)', saved as the new file 'DCSM-FM_4nm_adjusted.mdu'
    for n = 2:7                                % n: roop for parameters
        new_value = num2str(samples(m,n));     % new values for parameters
        text_modify(output_filename,output_filename,new_value,line(1,n));
        % the new file will cover the old one with the same name
    end
    fprintf('%d\n',m);
end
```

```
function text_modify(filename_ini,filename_new,new_contents,line)
% read the file 'filename_ini', replace contents in 'line' with 'new_contents', output
    file is 'filename_new'

%% read the file and adjust
 fileID = fopen(filename_ini,'r+');
 % open the file in readable and writable way
 i=0;
 while ~feof(fileID)
        tline=fgetl(fileID);
        % read the original file line by line
        i=i+1;
        newline{i} = tline;
        % create new lines to accpect original file's contents in each line
        if i==line
            % determine whether the row to be modified is reached
            value_par = str2num(cell2mat(regexp(tline,'\d(\.\d)?','match')));
            % extract the values in this row, using regular expression in Matlab
            old = num2str(value_par);
            % transform values into string
            newline{i} = strrep(tline,old,new_contents);
            % strrep: replace 'old' with 'new_contents'
        end
    end
    fclose(fileID);
    % close the file

%% export the file
    fileID = fopen(filename_new,'w+');
    % open the output file in readable and writable way.
    % If the file exists, clear the file contents and write from the beginning; if not,
        create the new file with the name and open only writable
    for k=1:i
        fprintf(fileID,'%s\t\n',newline{k});
        % write newline line by line
    end
    fclose(fileID);
    % close the file

end
```

Create subfolders to save the models

```
% create subfolders to run models together
for k = 1:50
    folderName{k} = ['temperature_copula_',num2str(k.','%04d')];
    mkdir(folderName{k});
end
```

Copy files into subfolders

```
% copy xml-file and sh-file to subfolders in order to run the models
for k =  1:50
    copyfile('dimr.xml',['temperature_copula_' num2str(k.','%04d')]);
    copyfile('run_dimr_h6-c7.sh',['temperature_copula_' num2str(k.','%04d')]);
    fprintf('%d\n',k);
end
```

B

Codes: read the outputs and create time series plots

Read the outputs

Take temperature in copula-based method as an example.

```
%% read outputs from the simulations and save as mat-file
output_temperature = zeros(50,6,1489);
% 1st dimension: 50 models
% 2nd dimension: 6 observation points
% 3rd dimension: 1489 timepoints
for k = 1:50
    ncFilePath = ['temperature_copula_' num2str(k.','%04d') '/
        DFM_OUTPUT_DCSM_temperature_copula/DCSM_temperature_copula_0000_his.nc'];
    % get the file containing target information in each subfolder
    temperature = ncread(ncFilePath,'temperature');
    % size: 50*6*1489. for one model, 50 layers, 6 points, 1489 timepoints
    temperature_top = reshape(temperature(50,:,:),6,1489);
    % choose 50th layer, after reshape each row is a time series for a site
    output_temperature(k,:,:) = temperature_top;
    % save all the models together
    fprintf('%d\n',k);
end
save('temperature.mat','output_temperature','k');

%% write the outputs into excel: 693
% the same to other timepoints
% find the timepoint in the time series plot
output_temperature_693 = reshape(output_temperature(:,:,693),50,6);
writematrix(output_temperature_693,'output_temperature_simulations_693.xlsx','Sheet',1,
    'Range','A1:F1489');
```

Create time series plots

Take temperature in Morris method as an example.

```
%% first plot with x_label as timepoint, to decide the timepoints used in sensitivtiy
    analysis.
load('temperature.mat');
output_avr = mean(output_temperature);
% 3D data
temperature_avr = reshape(output_avr(1,:,:),6,1489);
% transform into 2D data
plot(temperature_avr');
title('Average temperature of 80 models at different observation sites');
xlabel('time');
```

```
ylabel('temperature');
legend('Point1','Point2','Point3','Point4','Point5','Point6');

%% x_label: date
load('temperature.mat');

num_st=datenum('2004-12-25 02:24:00');
num_et=datenum('2005-01-25 02:24:00');
x=linspace(num_st,num_et,1489);
% change x label into date
timepoint693=datenum('2005-01-08 12:24:00');
timepoint1159=datenum('2005-01-18 05:24:00');
timepoint750=datenum('2005-01-09 16:54:00');
timepoint130=datenum('2004-12-27 18:54:00');
% decide the timepoints from above
output_avr = mean(output_temperature);
% calculate the average values of 6 observation points
temperature_avr = reshape(output_avr(1,:,:),6,1489);
% transform into 2D data

set(gcf,'unit','normalized','position',[0.1,0.1,0.8,0.8]);
plot(x,temperature_avr');
xline(timepoint693,'--r',{'2005-01-08 12:24:00'},'fontsize',18);
xline(timepoint1159,'--r',{'2005-01-18 05:24:00'},'fontsize',18);
xline(timepoint750,'--b',{'2005-01-09 16:54:00'},'fontsize',18);
xline(timepoint130,'--b',{'2004-12-27 18:54:00'},'fontsize',18);

datetick('x','yyyy-mm-dd','keepticks')
xtickangle(60)
title('Average temperature of 50 models at different observation sites','fontsize',18);
set(gca,'fontsize',18);
xlabel('time','fontsize',18);
ylabel('temperature','fontsize',18);
legend('Point1','Point2','Point3','Point4','Point5','Point6','fontsize',18);
```

C

Codes: Morris method

Codes used to double check

```
%% read inputs and outputs: 80 samples, 6 observation points
normlized = readmatrix('normlized_inputs.xlsx');
% in this codes, the inputs should be normalized before
output = readmatrix('temperature_time697');
% the way of read outputs has different possibilities

%% calculate elementary elements
diff_between_samples = zeros(79,6);
% calculate the differences between samples
for i = 1:79;
    diff_between_samples(i,:) = output(i+1,:) - output(i,:);
end
EE = zeros(10,7,6);
% 10 trajectories, 7 parameters, 6 observation points
% in each trajectory (10 totally), there is a EE for each parameter
for j = 1:7
    % 7 parameters
    k = 1;
    for ii = 1:79
        % in each trajectory, find the two sample points where the difference is
        old = normlized(ii,j);
        new = normlized(ii+1,j);
        if old ~= new && mod(ii,8) ~= 0
            % the first statement: EE lies on the step where the parameter
            % changes, which occurs only once for one parameter within one
            % trajectory
            % the second statement indicates: the difference is not included
            % if 'old' and 'new' show over the gap of two trajectories
            for num_obs = 1:6
                % save the EE in the matrix
                EE(k,j,num_obs) = diff_between_samples(ii,num_obs) / (new - old);
            end
            k = k+1;
        end
    end
end

%% mean, absolute mean, standard deviation
mu = reshape(mean(EE),7,6);
mu_star = reshape(mean(abs(EE)),7,6);
sigma = reshape(std(EE),7,6);

%% save the measurements in txt-file
par_name = ["Vicouv", "Dicouv","Vicoww","Dicoww","Smagorinsky","Stanton","Dalton"];
fileID = fopen('EE_measurement.txt','w');
for num_obs = 1:6
```

```
47      fprintf(fileID,'Point%d\n',num_obs);
48      fprintf(fileID,'parameter        mu*      mu      sigma\n');
49      for j = 1:7
50          fprintf(fileID,'%s       %.4e        %.4e        %.4e\n',par_name(j),mu_star(j,
                num_obs),mu(j,num_obs),sigma(j,num_obs));
51          if j == 7
52              fprintf(fileID,'-----------------------------------------------\n');
53          end
54      end
55  end
56  fclose(fileID);
```

Morris method package[24]

The codes from the package introduced in chapter 3.4.1 and 3.4.3 can be downloaded from [24].

D

Codes: copula-based method

The codes for copula-based method are created by Ţene et al. Some adjustments are made according to the cases in the project, thus the main function are listed. The library of the full codes can be found from [51]

Sampling

```
% take the example of temperature
% Use Morris method to sample model runs

clear all;
close all;
% add library folder, containing the necessary functions
addpath(fullfile(pwd,'lib'));

% model parameter definition (index and name)
Vicouv = 1;          var_names{Vicouv} = 'Vicouv';
Dicouv = 2;          var_names{Dicouv} = 'Dicouv';
Vicoww = 3;          var_names{Vicoww} = 'Vicoww';
Dicoww = 4;          var_names{Dicoww} = 'Dicoww';

Smagorinsky = 5;     var_names{Smagorinsky} = 'Smagorinsky';

Stanton = 6;         var_names{Stanton} = 'Stanton';
Dalton = 7;          var_names{Dalton} = 'Dalton';

% parameter ranges (in the order of the indices defined above)
low =     [ 0.1,   0.1,   0.000001,   0.000001,   0.05,    0.001 ,   0.001  ];
high =    [ 2   ,   2   ,   0.0001   ,   0.0001   ,   0.3 ,    0.0016,   0.0016 ];

% Sensitivity analysis settings (Morris method)
p = 8;                           % Morris grid points
k = 1;                                                % Morris step size (in # ofgrid
     cells)
r = 10;                          % number of Morris paths
n = 7;                           % number of variables
corr = diag(ones(1,n));                  % n x n correlation matrix
c = 1;
% Define the rank correlations between parameters: -1 (fully inversely correlated) <=
    corr(i,j) <= 1 (fully positively correlated)
% The correlations may be different and will define the copula.
corr(Vicouv,Dicouv) = c; corr(Dicouv,Vicouv) = c;
corr(Vicoww,Dicoww) = c; corr(Dicoww,Vicoww) = c;
corr(Stanton,Dalton) = c; corr(Dalton,Stanton) = c;

index_freevar = [5];
% Define the index matrix of free var, used in extended_morris.m to
% replace the samples of free var by end of the cube

```

```
% Define groups
groups(1).indices = [Vicouv,Dicouv];
            % variables in this group
groups(1).name = [var_names{Vicouv} '-' var_names{Dicouv}];              % also their
    names, for plotting
groups(1).prob = [0.5 1; 0 0.5];                % upper triangular probability matrix
                                                % prob(i,j) = P(Xi increases | Xj
                                                    increases)
                                                % prob(i,i) = P(Xi increases)
                                                % lower triangular part will be filled
                                                    in automatically, according to
                                                    Bayes' rule
%  type = 2, to determine sensitivity of the group (as one entity)
%  type = 1, to determine individual parameter sensitivities

groups(1).type = 2;

groups(2).indices = [Vicoww,Dicoww];
groups(2).prob = [0.5 1; 0 0.5];
groups(2).type = 2;
groups(2).name = [var_names{Vicoww} '-' var_names{Dicoww}];

groups(3).indices = [Stanton,Dalton];
groups(3).prob = [0.5 1; 0 0.5];
groups(3).type = 2;
groups(3).name = [var_names{Stanton} '-' var_names{Dalton}];

% use Morris to generate the paths
[paths,vars] = extended_morris(p,k,r,corr,groups,index_freevar);

% extract model evaluations
A = cell2mat(paths(:));

% remove duplicates (highly unlikely to have duplicates)
[A_unique,~,map] = unique(A,'rows');
[nPaths,nPoints] = size(paths);
map = reshape(1:nPaths*nPoints, nPaths, nPoints);

% scale to parameter ranges
A = bsxfun(@times, A, high) + bsxfun(@times, 1-A, low);

% rewrite A in the order of paths
for i = 1:10
    for j = 1:5
        A_order((i-1)*5+j,:) = A((j-1)*10+i,:);
    end
end

% these files save the configuration and the paths, in Matlab format
save('setup_temperature.mat','var_names','low','high','groups','p','k','r','n','corr','
    index_freevar');
save('paths_temperature.mat','paths','vars','map','A','A_unique','A_order');

% write the inputs into excel
writematrix(A,'parameters_temperature_simulations.xlsx','Sheet',1,'Range','A1:G50');
writematrix(A_order,'parameters_temperature_in_order.xlsx','Sheet',1,'Range','A1:G50');
```

Sensitivity measurements

```
% Use model results to do sensitivity analysis

% clear all;
% close all;

% load configuration and paths
load('setup_temperature.mat','var_names','var_names','low','high','groups','p','k','
    index_freevar');
load('paths_temperature.mat','paths','vars','map');
```

```
step = k / (p-1);
[nPaths,nPoints] = size(paths);
nSegments = nPoints-1;

% get model results
M_all = importdata('output_temperature_simulations_130.xlsx');
for j=1:6
M = M_all(:,j);
% compute elementary effects
% iVars = [groups([groups.type] ~= 2).indices];
% iVars = [iVars setdiff(1:length(var_names), [groups.indices])];
    iVars = 1;
    iGroups = find([groups.type] == 2);
    Effects = zeros(nPaths,nSegments);
    for iPath=1:nPaths
        for iPoint=2:nPoints
            iSegment = iPoint-1;
            l1 = map(iPath,iPoint-1);
            l2 = map(iPath,iPoint);

            iVar = vars(iPath,iSegment);
            if iVar > 0
                iEntity = find(iVars == iVar);
            else
                iEntity = length(iVars) + find(iGroups == -iVar);
            end
            Effects(iPath,iEntity) = (M(l2) - M(l1)) / step;
        end
    end
% compute mu, mu_star, sigma
    mu(j,:) = mean(Effects);
    mu_star(j,:) = mean(abs(Effects));
    sigma(j,:) = std(Effects);
    names = [var_names(index_freevar) {groups(iGroups).name}];
% save the effects
    struct_value{j} = transpose(Effects);
end
EERaw = struct('Point01',{struct_value{1}},'Point02',{struct_value{2}},'Point03',{
    struct_value{3}},'Point04',{struct_value{4}},'Point05',{struct_value{5}},'Point06'
    ,{struct_value{6}});

%%
% plot
for j=1:6
    mu_star1 = mu_star(j,:);
    sigma1 = sigma(j,:);
    figure(j);
    % h=figure('units','normalized','outerposition',[0 0 1 1]);
    set(gcf,'unit','normalized','position',[0.1,0.1,0.8,0.8]);
    plot(mu_star1,sigma1,'*','MarkerSize',6);
    xlim([0 5]);
    title(strcat({'Sensitivity indices at Point'},num2str(j)));
    xlabel('Absolute mean');
    ylabel ('Standard deviation');
    lim = xlim();
    text(mu_star1+0.01*(lim(2)-lim(1)), sigma1, names);
    xlim([-0.1 0.5]);
    ylim([-0.1 0.5]);
end

%%
save('EE_copula_temperature_timepoint130_new.mat','names','mu_star','mu','sigma','EERaw
    ');
```

E

Codes: variance-based method

```
%% read matrix, the result is on FINO3, x: 380; y: 372; z: 378
A = readmatrix('current_A.xlsx','Sheet',1,'Range','A2:D129');
B = readmatrix('current_B.xlsx','Sheet',1,'Range','A2:D129');
C1 = readmatrix('current_C1.xlsx','Sheet',1,'Range','A2:D129');
C2 = readmatrix('current_C2.xlsx','Sheet',1,'Range','A2:D129');
C3 = readmatrix('current_C3.xlsx','Sheet',1,'Range','A2:D129');
C4 = readmatrix('current_C4.xlsx','Sheet',1,'Range','A2:D129');
% where the input matrixes are saved

load('current_A.mat');
load('current_B.mat');
load('current_C1.mat');
load('current_C2.mat');
load('current_C3.mat');
load('current_C4.mat');
% where the simulation results of different matrixed are saved

xdir_f_A = x_velocity_A(:,1,380);
xdir_f_B = x_velocity_B(:,1,380);
xdir_f_C1 = x_velocity_C1(:,1,380);
xdir_f_C2 = x_velocity_C2(:,1,380);
xdir_f_C3 = x_velocity_C3(:,1,380);
xdir_f_C4 = x_velocity_C4(:,1,380);

ydir_f_A = y_velocity_A(:,1,372);
ydir_f_B = y_velocity_B(:,1,372);
ydir_f_C1 = y_velocity_C1(:,1,372);
ydir_f_C2 = y_velocity_C2(:,1,372);
ydir_f_C3 = y_velocity_C3(:,1,372);
ydir_f_C4 = y_velocity_C4(:,1,372);

zdir_f_A = z_velocity_A(:,1,378);
zdir_f_B = z_velocity_B(:,1,378);
zdir_f_C1 = z_velocity_C1(:,1,378);
zdir_f_C2 = z_velocity_C2(:,1,378);
zdir_f_C3 = z_velocity_C3(:,1,378);
zdir_f_C4 = z_velocity_C4(:,1,378);
% extract the results of current velocities in different directions

var_name = {'Vicoww','Dicoww','Smagorinsky','Rhoair'};

save('current_variance.mat');

%% x direction
xdir_f0 = mean(xdir_f_A);

xdir_y_AC1 = mean(xdir_f_A .* xdir_f_C1);
xdir_y_AC2 = mean(xdir_f_A .* xdir_f_C2);
xdir_y_AC3 = mean(xdir_f_A .* xdir_f_C3);
```

```
xdir_y_AC4 = mean(xdir_f_A .* xdir_f_C4);

xdir_y_BC1 = mean(xdir_f_B .* xdir_f_C1);
xdir_y_BC2 = mean(xdir_f_B .* xdir_f_C2);
xdir_y_BC3 = mean(xdir_f_B .* xdir_f_C3);
xdir_y_BC4 = mean(xdir_f_B .* xdir_f_C4);

xdir_y_AA = mean(xdir_f_A .* xdir_f_A);

% first-order sensitivity indices
xdir_S(1,1) = (xdir_y_AC1 - xdir_f0^2)/(xdir_y_AA - xdir_f0^2);
xdir_S(2,1) = (xdir_y_AC2 - xdir_f0^2)/(xdir_y_AA - xdir_f0^2);
xdir_S(3,1) = (xdir_y_AC3 - xdir_f0^2)/(xdir_y_AA - xdir_f0^2);
xdir_S(4,1) = (xdir_y_AC4 - xdir_f0^2)/(xdir_y_AA - xdir_f0^2);

% total-effect indices
xdir_ST(1,1) = 1 - (xdir_y_BC1 - xdir_f0^2)/(xdir_y_AA - xdir_f0^2);
xdir_ST(2,1) = 1 - (xdir_y_BC2 - xdir_f0^2)/(xdir_y_AA - xdir_f0^2);
xdir_ST(3,1) = 1 - (xdir_y_BC3 - xdir_f0^2)/(xdir_y_AA - xdir_f0^2);
xdir_ST(4,1) = 1 - (xdir_y_BC4 - xdir_f0^2)/(xdir_y_AA - xdir_f0^2);

save('xdir_sensitivity_indices.mat','xdir_S','xdir_ST','var_name');

%% y direction
ydir_f0 = mean(ydir_f_A);

ydir_y_AC1 = mean(ydir_f_A .* ydir_f_C1);
ydir_y_AC2 = mean(ydir_f_A .* ydir_f_C2);
ydir_y_AC3 = mean(ydir_f_A .* ydir_f_C3);
ydir_y_AC4 = mean(ydir_f_A .* ydir_f_C4);

ydir_y_BC1 = mean(ydir_f_B .* ydir_f_C1);
ydir_y_BC2 = mean(ydir_f_B .* ydir_f_C2);
ydir_y_BC3 = mean(ydir_f_B .* ydir_f_C3);
ydir_y_BC4 = mean(ydir_f_B .* ydir_f_C4);

ydir_y_AA = mean(ydir_f_A .* ydir_f_A);

% first-order sensitivity indices
ydir_S(1,1) = (ydir_y_AC1 - ydir_f0^2)/(ydir_y_AA - ydir_f0^2);
ydir_S(2,1) = (ydir_y_AC2 - ydir_f0^2)/(ydir_y_AA - ydir_f0^2);
ydir_S(3,1) = (ydir_y_AC3 - ydir_f0^2)/(ydir_y_AA - ydir_f0^2);
ydir_S(4,1) = (ydir_y_AC4 - ydir_f0^2)/(ydir_y_AA - ydir_f0^2);

% total-effect indices
ydir_ST(1,1) = 1 - (ydir_y_BC1 - ydir_f0^2)/(ydir_y_AA - ydir_f0^2);
ydir_ST(2,1) = 1 - (ydir_y_BC2 - ydir_f0^2)/(ydir_y_AA - ydir_f0^2);
ydir_ST(3,1) = 1 - (ydir_y_BC3 - ydir_f0^2)/(ydir_y_AA - ydir_f0^2);
ydir_ST(4,1) = 1 - (ydir_y_BC4 - ydir_f0^2)/(ydir_y_AA - ydir_f0^2);

save('ydir_sensitivity_indices.mat','ydir_S','ydir_ST','var_name');

%% z direction
zdir_f0 = mean(zdir_f_A);

zdir_y_AC1 = mean(zdir_f_A .* zdir_f_C1);
zdir_y_AC2 = mean(zdir_f_A .* zdir_f_C2);
zdir_y_AC3 = mean(zdir_f_A .* zdir_f_C3);
zdir_y_AC4 = mean(zdir_f_A .* zdir_f_C4);

zdir_y_BC1 = mean(zdir_f_B .* zdir_f_C1);
zdir_y_BC2 = mean(zdir_f_B .* zdir_f_C2);
zdir_y_BC3 = mean(zdir_f_B .* zdir_f_C3);
zdir_y_BC4 = mean(zdir_f_B .* zdir_f_C4);

zdir_y_AA = mean(zdir_f_A .* zdir_f_A);

% first-order sensitivity indices
zdir_S(1,1) = (zdir_y_AC1 - zdir_f0^2)/(zdir_y_AA - zdir_f0^2);
zdir_S(2,1) = (zdir_y_AC2 - zdir_f0^2)/(zdir_y_AA - zdir_f0^2);
zdir_S(3,1) = (zdir_y_AC3 - zdir_f0^2)/(zdir_y_AA - zdir_f0^2);
```

```
121 zdir_S(4,1) = (zdir_y_AC4 - zdir_f0^2)/(zdir_y_AA - zdir_f0^2);
122
123 % total-effect indices
124 zdir_ST(1,1) = 1 - (zdir_y_BC1 - zdir_f0^2)/(zdir_y_AA - zdir_f0^2);
125 zdir_ST(2,1) = 1 - (zdir_y_BC2 - zdir_f0^2)/(zdir_y_AA - zdir_f0^2);
126 zdir_ST(3,1) = 1 - (zdir_y_BC3 - zdir_f0^2)/(zdir_y_AA - zdir_f0^2);
127 zdir_ST(4,1) = 1 - (zdir_y_BC4 - zdir_f0^2)/(zdir_y_AA - zdir_f0^2);
128
129 save('zdir_sensitivity_indices.mat','zdir_S','zdir_ST','var_name');
```

Bibliography

[1] URL https://www.h2020united.eu/.

[2] Fino3 - research platform in the north sea and the baltic no.3. URL https://www.fino3.de/en/.

[3] The different methods of growing oysters. URL https://www.pangeashellfish.com/blog/the-different-methods-of-growing-oysters.

[4] Half of fish consumed globally is now raised on farms, study finds, Sep 2009. URL https://www.sciencedaily.com/releases/2009/09/090907162320.htm.

[5] The ancient world - egypt, Jul 2010. URL https://web.archive.org/web/20100723003850/http://www.marinersmuseum.org/education/ancient-world-egypt.

[6] Research topics, Sep 2013. URL https://scripps.ucsd.edu/research/topics.

[7] Luis A Bastidas, James Knighton, Shaun W Kline, and Justin Pistininzi. Delft3d sensitivity and uncertainty snalysis for hurricane simulations in the north atlantic. In *2016 Ocean Sciences Meeting*. AGU, 2016.

[8] Joseph A Berg and Roger IE Newell. Temporal and spatial variations in the composition of seston available to the suspension feeder crassostrea virginica. *Estuarine, Coastal and Shelf Science*, 23(3):375–386, 1986.

[9] Juan Blanco, Manuel Zapata, and Ángeles Moroño. Some aspects of the water flow through mussel rafts. 1996.

[10] Annette Bruhn, Ditte Bruunshøj Tørring, Marianne Thomsen, Paula Canal-Vergés, Mette Møller Nielsen, Michael Bo Rasmussen, Karin Loft Eybye, Martin Mørk Larsen, Thorsten Johannes Skovbjerg Balsby, and Jens Kjerulf Petersen. Impact of environmental conditions on biomass yield, quality, and bio-mitigation capacity of saccharina latissima. *Aquaculture environment interactions*, 8:619–636, 2016.

[11] Francesca Campolongo, Jessica Cariboni, and Andrea Saltelli. An effective screening design for sensitivity analysis of large models. *Environmental modelling & software*, 22(10):1509–1518, 2007.

[12] Janhavi Chitale, Yogesh Khare, Rafael Muñoz-Carpena, George S Dulikravich, and Christopher Martinez. An effective parameter screening strategy for high dimensional models. In *ASME International Mechanical Engineering Congress and Exposition*, volume 58424, page V007T09A017. American Society of Mechanical Engineers, 2017.

[13] Maria Librada Chu-Agor, Rafael Muñoz-Carpena, G Kiker, A Emanuelsson, and Igor Linkov. Exploring vulnerability of coastal habitats to sea level rise through global sensitivity and uncertainty analyses. *Environmental Modelling & Software*, 26(5):593–604, 2011.

[14] LA Comeau, A Drapeau, T Landry, and J Davidson. Development of longline mussel farming and the influence of sleeve spacing in prince edward island, canada. *Aquaculture*, 281(1-4):56–62, 2008.

[15] Deltares. Delft3d flexible mesh suite: D-flow flexible mesh, technical reference manual. 2019. URL https://www.deltares.nl/en/software/delft3d-flexible-mesh-suite/.

[16] Roberto Foscarini, Jayant Prakash, et al. *Handbook on Eucheuma seaweed cultivation in Fiji*. Ministry of Primary Industries, Fisheries Division and South Pacific . . . , 1990.

[17] Marcel Frechette, Cheryl Ann Butman, and W Rockwell Geyer. The importance of boundary-layer flows in supplying phytoplankton to the benthic suspension feeder, mytilus edulis l. *Limnology and oceanography*, 34(1):19–36, 1989.

[18] Pierre Garen, Stephane Robert, and Serge Bougrier. Comparison of growth of mussel, mytilus edulis, on longline, pole and bottom culture sites in the pertuis breton, france. *Aquaculture*, 232(1-4):511–524, 2004.

[19] Yvette L Garner and Marian K Litvaitis. Effects of wave exposure, temperature and epibiont fouling on byssal thread production and growth in the blue mussel, mytilus edulis, in the gulf of maine. *Journal of Experimental Marine Biology and Ecology*, 446:52–56, 2013.

[20] Per Hage and Jeff Marck. Matrilineality and the melanesian origin of polynesian y chromosomes. *Current Anthropology*, 44(S5):S121–S127, 2003.

[21] Toshimitsu Homma and Andrea Saltelli. Importance measures in global sensitivity analysis of nonlinear models. *Reliability Engineering & System Safety*, 52(1):1–17, 1996.

[22] LS Incze, RA Lutz, and L Watling. Relationships between effects of environmental temperature and seston on growth and mortality of mytilus edulis in a temperate northern estuary. *Marine Biology*, 57(3): 147–156, 1980.

[23] Peter Jarvis, Bruce Jefferson, JOHN Gregory, and Simon A Parsons. A review of floc strength and breakage. *Water research*, 39(14):3121–3137, 2005.

[24] Yogesh Khare and Rafael Muñoz Carpena. Morris su (sampling uniformity) code, 2014. URL `https://abe.ufl.edu/faculty/carpena/software/SUMorris.shtml`.

[25] Yogesh Khare, Christopher J Martinez, Rafael Muñoz-Carpena, Adelbert "Del" Bottcher, and Andrew James. Effective global sensitivity analysis for high-dimensional hydrologic and water quality models. *Journal of Hydrologic Engineering*, 24(1):04018057, 2019.

[26] Yogesh P Khare, Rafael Muñoz-Carpena, Robert W Rooney, and Christopher J Martinez. A multi-criteria trajectory-based parameter sampling strategy for the screening method of elementary effects. *Environmental Modelling & Software*, 64:230–239, 2015.

[27] Zhijie Li, Qiuwen Chen, and Qiang Xu. Modeling algae dynamics in meiliang bay of taihu lake and parameter sensitivity analysis. *Journal of Hydro-environment Research*, 9(2):216–225, 2015.

[28] Francis Henry Charles Marriott et al. *A dictionary of statistical terms*. Number Ed. 5. Longman Scientific & Technical, 1990.

[29] AM Mathiesen. The state of the world fisheries and aquaculture 2012, 2012.

[30] Michael D McKay, Richard J Beckman, and William J Conover. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 42(1): 55–61, 2000.

[31] Christopher W McKindsey, M Robin Anderson, Penelope Barnes, Simon Courtenay, Thomas Landry, and Marc Skinner. *Effects of shellfish aquaculture on fish habitat*. Fisheries and Oceans, 2006.

[32] Alan D McNaught, Andrew Wilkinson, et al. *Compendium of chemical terminology*, volume 1669. Blackwell Science Oxford, 1997.

[33] Max D Morris. Factorial sampling plans for preliminary computational experiments. *Technometrics*, 33 (2):161–174, 1991.

[34] Roger B Nelsen. *An introduction to copulas*. Springer Science & Business Media, 2007.

[35] Donna J Nickerson. Trade-offs of mangrove area development in the philippines. *Ecological Economics*, 28(2):279–298, 1999.

[36] Mette Møller Nielsen, Dorte Krause-Jensen, Birgit Olesen, Rikke Thinggaard, Peter Bondo Christensen, and Annette Bruhn. Growth dynamics of saccharina latissima (laminariales, phaeophyceae) in aarhus bay, denmark, and along the species' distribution range. *Marine Biology*, 161(9):2011–2022, 2014.

[37] Natalie Packham and Wolfgang M Schmidt. Latin hypercube sampling with dependence and applications in finance. *Available at SSRN 1269633*, 2008.

[38] Michael J Paterson, Cheryl L Podemski, Wilhelmina J Findlay, David L Findlay, and Alex G Salki. The response of zooplankton in a whole-lake experiment on the effects of a cage aquaculture operation for rainbow trout (oncorhynchus mykiss). *Canadian Journal of Fisheries and Aquatic Sciences*, 67(11):1852–1861, 2010.

[39] Jan A Pechenik, Linda S Eyster, John Widdows, and Brian L Bayne. The influence of food concentration and temperature on growth and morphological differentiation of blue mussel mytilus edulis l. larvae. *Journal of Experimental Marine Biology and Ecology*, 136(1):47–64, 1990.

[40] César Peteiro and Óscar Freire. Biomass yield and morphological features of the seaweed saccharina latissima cultivated at two different sites in a coastal bay in the atlantic coast of spain. *Journal of applied phycology*, 25(1):205–213, 2013.

[41] MV Ruano, J Ribes, Aurora Seco, and José Ferrer. An improved sampling strategy based on trajectory design for application of the morris method to systems with many input factors. *Environmental Modelling & Software*, 37:103–109, 2012.

[42] Andrea Saltelli, Marco Ratto, Terry Andres, Francesca Campolongo, Jessica Cariboni, Debora Gatelli, Michaela Saisana, and Stefano Tarantola. *Global sensitivity analysis: the primer*. John Wiley & Sons, 2008.

[43] Shashi Shekhar and Hui Xiong. *Encyclopedia of GIS*. Springer Science & Business Media, 2007.

[44] Joseph Smagorinsky. General circulation experiments with the primitive equations: I. the basic experiment. *Monthly weather review*, 91(3):99–164, 1963.

[45] Il'ya Meerovich Sobol'. On the distribution of points in a cube and the approximate evaluation of integrals. *Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki*, 7(4):784–802, 1967.

[46] FAO Statistics and Information Service (FIPS). Global aquaculture production fishery statistical collections, Oct 2011. URL `http://www.fao.org/fishery/statistics/global-aquaculture-production/en`.

[47] Dorrik Stow. *Encyclopedia of the Oceans*. Oxford University Press, 2004.

[48] T Strohmeier, J Aure, A Duinker, T Castberg, A Svardal, and Ø Strand. Flow reduction, seston depletion, meat content and distribution of diarrhetic shellfish toxins in a long-line blue mussel (mytilus edulis) farm. *Journal of Shellfish Research*, 24(1):15–23, 2005.

[49] Stefano Tarantola, D Gatelli, SS Kucherenko, Wolfgang Mauntz, et al. Estimating the approximation error when fixing unessential factors in global sensitivity analysis. *Reliability engineering & system safety*, 92 (7):957–960, 2007.

[50] Jennifer Telling, Andrew Lyda, Preston Hartzell, and Craig Glennie. Review of earth science research using terrestrial laser scanning. *Earth-Science Reviews*, 169:35–68, 2017.

[51] Matei Ţene, Dana E Stuparu, Dorota Kurowicka, and Ghada Y El Serafy. A copula-based sensitivity analysis method and its application to a north sea sediment transport model. *Environmental Modelling & Software*, 104:1–12, 2018.

[52] TJ Vermeulen. Sensitivity analysis of fine sediment transport in the humber estuary. 2004.

[53] Wouter Visch, Per Bergström, Göran M Nylund, My Peterson, Henrik Pavia, and Mats Lindegarth. Spatial differences in growth rate and nutrient mitigation of two co-cultivated, extractive species: The blue mussel (mytilus edulis) and the kelp (saccharina latissima). *Estuarine, Coastal and Shelf Science*, 246: 107019, 2020.

[54] John Widdows. Physiological ecology of mussel larvae. *Aquaculture*, 94(2-3):147–163, 1991.

[55] Jürgen E Winter. A review on the knowledge of suspension-feeding in lamellibranchiate bivalves, with special reference to artificial aquaculture systems. *Aquaculture*, 13(1):1–33, 1978.

[56] Dan Zhao, Nuomin Han, Ernest Goh, John Cater, and Arne Reinecke. Chapter 8 - aeroacoustics of wind turbines. In Dan Zhao, Nuomin Han, Ernest Goh, John Cater, and Arne Reinecke, editors, *Wind Turbines and Aerodynamics Energy Harvesters*, pages 463–491. Academic Press, 2019. ISBN 978-0-12-817135-6. doi: https://doi.org/10.1016/B978-0-12-817135-6.00008-9. URL https://www.sciencedirect.com/science/article/pii/B9780128171356000089.

[57] Firmijn Zijl, Jelmer Veenstra, and Julien Groenenboom. 3d dutch continental shelf model – flexible mesh (3d dcsm-fm): Setup and validation. 2018. URL https://www.deltares.nl/en/projects/3d-dutch-continental-shelf-model-flexible-mesh-2/.