Spatial Planning and Structural Safety in Offshore Engineering

through Probabilistic Multivariate Modelling with Copulas

Rieke Santjer

Spatial Planning and Structural Safety in Offshore Engineering through Probabilistic Multivariate Modelling with Copulas

Spatial Planning and Structural Safety in Offshore Engineering through Probabilistic Multivariate Modelling with Copulas

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology
by the authority of the Rector Magnificus prof. dr. ir. H. Bijl,
Chair of the Board for Doctorates,
to be defended publicly on
Monday, 22 June 2026 at 10:00

by

Rieke SANTJER

This dissertation has been approved by the promotors.

Composition of the doctoral committee:

Rector Magnificus,	chairperson
Prof. dr. A.V. Metrikine,	Delft University of Technology, promotor
Prof. dr. T. Rossetto,	Delft University of Technology, promotor
Em. prof. dr.ir. A.W. Heemink,	Delft University of Technology, promotor

Independent members:

Prof. dr.ir. P.H.A.J.M. van Gelder	Delft University of Technology, Netherlands
Dr. J.L.A. Dubbeldam	Delft University of Technology, Netherlands
Prof. dr.-ing. M. Beer	Leibniz University Hanover, Germany
Prof. dr. I. Gitman	Maastricht University, Netherlands
Prof. dr.ir. W.S.J. Uijttewaal	Delft University of Technology, Netherlands, reserve member

This work is co-financed by the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement no 862915 and Stichting Deltares, Netherlands.

TU Delft Delft University of Technology

Deltares

Funded by the European Union (H2020 Grant Agreement no 862915). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them

UNITED

Keywords: dependence modelling, joint distribution, copula, Bayesian network, vine copula, offshore engineering, aquaculture, floating structures, mooring line, floating photovoltaic, structural response

Printed by: Gildeprint - Enschede

Front & Back: Cover art done by Sabine Santjer

ISBN 978-94-6518-355-8

An electronic version of this dissertation is available at https://repository.tudelft.nl/.

CONTENTS

Summary **9**

1 Probabilistic multivariate modelling for offshore applications **1**
1.1 Dependence modelling via copula models . . . 3
1.2 The concept of copula modelling . . . 4
1.3 The role of dependence in aquaculture suitability modelling . . . 5
1.4 Gaussian Copula-based Bayesian Networks for dynamic mooring tensions 6
1.5 Vine Copulas for complex floating structures . . . 7
1.6 Objective and scope . . . 8

2 A probabilistic framework for offshore aquaculture suitability assessment using bivariate copulas **11**
2.1 Introduction . . . 11
2.2 Materials and methods . . . 15
2.2.1 3D hydrodynamic and water quality model . . . 16
2.2.2 Variable selection and sampling . . . 16
2.2.3 Spearman rank correlation . . . 19
2.2.4 Copula modeling . . . 20
2.2.5 Joint multivariate probability calculation . . . 24
2.3 Results . . . 26
2.3.1 Correlation analysis of the data . . . 26
2.3.2 Copula selection and fitting . . . 26
2.3.3 Probability calculation and suitability maps . . . 29
2.4 Discussion . . . 31
2.5 Conclusion . . . 33

3 Gaussian copula-based Bayesian networks for dynamic loads in mooring systems **35**
3.1 Introduction . . . 35
3.2 Modelling framework and study setup . . . 39
3.2.1 3D hydrodynamic model . . . 39
3.2.2 Study locations and variable selection . . . 39
3.2.3 Mooring line model . . . 41
3.3 Modelling approach . . . 45
3.3.1 Spearman's rank correlation coefficient . . . 46
3.3.2 Bivariate copulas . . . 47
3.3.3 Gaussian Copula-based Bayesian Network . . . 47
3.3.4 Processing of numerical tension data . . . 48

3.4 Results . . . 49
3.4.1 Hydrodynamic Gaussian Copula-based Bayesian Networks . . . 49
3.4.2 Methodological application on tension data exemplified by the ReFloc 4 . . . 52
3.4.3 Bayesian Networks of hydrodynamic variables and tension rates . . . 55
3.4.4 Evaluation of tension rates at various percentiles . . . 57
3.4.5 Evaluation of tension rates for different conditionalisation scenarios 59
3.5 Discussion . . . 61
3.6 Conclusion . . . 63

4 Vine Copula Modelling of Mixed Continuous and Ordinal Data in Offshore Floating Solar Platforms 65
4.1 Introduction . . . 65
4.2 Probabilistic modelling approach . . . 69
4.2.1 (Bivariate) Copula Modelling . . . 69
4.2.2 Vine Copula Modelling for continuous variables . . . 70
4.2.3 Vine Copula Modelling for mixed continuous and ordinal variables 72
4.3 Study setup . . . 72
4.3.1 Location selection and input data . . . 73
4.3.2 Hydroelastic model for FPV . . . 75
4.3.3 Vine fitting approach . . . 76
4.4 Results . . . 77
4.4.1 Selection of extremes and preparation of input data . . . 77
4.4.2 Hydrodynamic and structural data overview . . . 78
4.4.3 Global Vine Model . . . 80
4.4.4 Vine Models per configuration . . . 81
4.4.5 Vine Models per regime . . . 85
4.5 Discussion . . . 92
4.6 Conclusion . . . 95

5 Discussion & Conclusion 97
5.1 Methodological Reflections . . . 97
5.1.1 Variable selection and sampling . . . 97
5.1.2 Significance testing . . . 98
5.1.3 Copula Models . . . 99
5.1.4 Gaussian Copula-based Bayesian Networks . . . 100
5.1.5 Vine Copula Models . . . 100
5.1.6 Sensitivity of Vine structures in low dimensions . . . 101
5.1.7 Performance of GCBN and Vine Copula Models in capturing extremes 103
5.2 Engineering Implications . . . 105
5.2.1 A Roadmap for engineers in practice . . . 106
5.2.2 Methodological considerations for engineering applications . . . 110
5.3 Conclusion . . . 110

Acknowledgements 113

A Appendix **117**
A.1 Timeseries of the studied variables . . . 117
A.2 Tail dependence . . . 118
A.3 Cumulative distribution functions of bivariate parametric copulas . . . 119
A.4 (Inverse) conditional distribution functions . . . 120
A.5 Additional results for the blue mussel case . . . 121
A.5.1 Correlation analysis . . . 121
A.5.2 Angle of rotation for Gumbel and Clayton copula . . . 122
A.5.3 Probability calculation . . . 123
A.6 Additional results for the sugar kelp . . . 124
A.6.1 Example of the sampled data . . . 124
A.6.2 Correlation analysis . . . 125
A.6.3 Copula fitting results . . . 126
A.6.4 Copula parameter and angle of rotation for Gumbel and Clayton copula . . . 126
A.6.5 Probability calculation . . . 128

B Appendix **129**
B.1 Overview of wave characteristics . . . 129
B.2 Example hydrodynamic model data . . . 130
B.3 GCBN structures for each study location . . . 131
B.4 Comparing cumulative distribution functions . . . 132
B.5 Additional results on tension rate percentiles . . . 132

C Appendix **133**
C.1 Selection of extremes . . . 133
C.2 Example of FPV model output . . . 134
C.3 Data per regime . . . 136

Bibliography **137**

Samenvatting **153**

Statement of Novelty **155**

List of Publications **157**

Curriculum Vitae **159**

SUMMARY

Globally, the population is continuously increasing, as is the demand for food and energy. This growing need and limited possibilities for further exploiting on-land resources has led to an increased interest in the exploitation of the marine environment to fill the resource gap. In particular offshore areas are increasingly explored, despite challenges such as harsh environmental conditions characterised by strong variability and complex interactions between physical processes. While offshore sectors such as wind energy are already commercially established, others, such as offshore floating photovoltaic (FPV) systems and aquaculture are still in early stages of development. For both sectors, research has been conducted through small-scale field observations, laboratory experiments, or numerical modelling. However, these approaches often simplify or neglect statistical dependence among the environmental and structural variables, thereby hindering optimal design of offshore structures.

This dissertation explores the applicability of probabilistic models to explicitly capture such dependencies and to demonstrates how such models can enhance decision-making across different offshore technologies, with a particular focus on aquaculture and floating photovoltaic systems. Two copula-based multivariate modelling approaches with different complexity are applied. On the one hand, the Gaussian copula-based Bayesian Network (GCBN) offers a comparatively accessible and interpretable framework, as its graphical structure can be derived from the underlying physical processes. On the other hand, vine copula models are used to represent complex dependence patterns more flexibly. Unlike GCBNs, they are not restricted to a single copula family, instead, each pair of variables can be modelled using the copula that best captures their dependence. This flexibility, however, comes at the cost of higher model complexity and increased computational and practical challenges.

Subsequently, this dissertation is driven by the overall research question on how multivariate copula-based probabilistic models can support spatial planning and structural safety assessment in offshore engineering. To help answering this, three research avenues are explored in this work.The first examines how variable dependence affects the accuracy of aquaculture site suitability maps. Therefore, the probabilities of meeting optimal and critical survival thresholds are estimated both with and without accounting for dependence among variables. Differences of up to 40% are observed between the two approaches, highlighting the importance of dependence-aware modelling for reliable marine spatial planning.

The second avenue addresses how GCBNs can support the safety assessment of floating offshore structures by modelling mooring tensions under specific hydrodynamic scenarios. Conditioning tension responses on certain hydrodynamic conditions reduces

uncertainties within the safety assessment. The results show that, under harsh conditions, snap loads are counter-intuitively lower due to heavy-tailed marginal distributions.

The third avenue investigates how vine copula models can represent complex dependence among mixed continuous and ordinal variables relevant to the structural response of floating offshore platforms. The results show that vine copula models are able to capture the complex dependence pattern per regime, i.e. per a range of parameters, within which the structure under investigation behaves as a rigid or flexible body. The vine copula models provide detailed insights into how structural configurations influence the dynamic response and enable conditional analyses that help identify configurations that minimise waves-induced stresses in the structure.

Across all chapters, the results demonstrate that explicitly accounting for multivariate dependence leads to more reliable assessments in offshore aquaculture, mooring safety, and structural response modelling. More broadly, the findings emphasise that dependence modelling is not merely a statistical refinement, but a necessary component for understanding and managing offshore systems characterised by interacting environmental and structural processes. By integrating copula-based methods into current practice, this work contributes to developing more robust, transparent, and physically informed frameworks for decision-making in emerging offshore technologies.

1

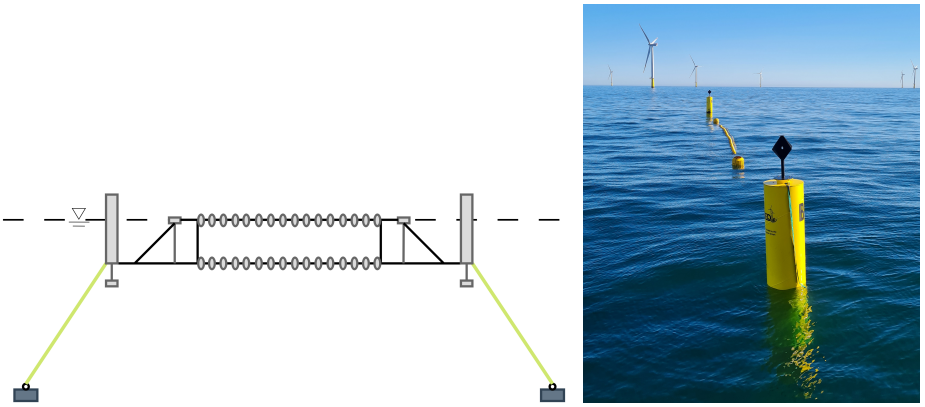
PROBABILISTIC MULTIVARIATE MODELLING FOR OFFSHORE APPLICATIONS

The human world population is continuously increasing and thus the food and energy demand. In addition, given for instance global political situations and the climate change, using local food and energy resources to overcome the challenge of higher demands but avoiding higher dependence on other states around the world becomes more relevant. However, since space at land is limited, the use of the marine environment, near- and offshore, is gaining increasing relevance due to its substantial potential as a source of both food and energy (Yong *et al.*, 2022). Nevertheless, challenges such as limited space (Buck and Langan, 2017; Buck *et al.*, 2018; Pınarbaşı *et al.*, 2017) are hindering to make full use of the potential. One approach to overcome this challenge is the multiple use of marine space (Buck *et al.*, 2004; Wever *et al.*, 2015). Not only space can be shared, e.g. by combining offshore wind farms with aquaculture of mussels or kelp, or floating photovoltaic systems, but also costs can be minimised by shared operation and maintenance. However, main factor hindering the development of such offshore sectors are harsh environmental conditions in offshore regions, consequently leading to higher costs. These conditions are inherently uncertain, often characterised by chaotic and stochastic behaviour. Therefore, across various offshore sectors, both scientific knowledge and practical experience remain limited. The complex and interdependent nature of offshore variables are not yet fully understood, particularly in terms of their interaction with and impact on offshore engineering structures. Given this complexity and uncertainty, probabilistic methods offer a valuable framework to describe these dependencies of different variables.

This dissertation examines the assessment and suitability of offshore structures, particularly aquaculture and floating photovoltaic (FPV) serve as case studies out of two main reasons. First, there is a high potential and support for massive development and significant prospective benefits of offshore aquaculture, such as mussel and kelp cultivations (e.g., Fernand *et al.*, 2017; Kammler *et al.*, 2024; Kerrison *et al.*, 2015; Schiener,

2014) and FPV (e.g., Benjamins *et al.*, 2024; Fan *et al.*, 2025; Nalmpanti *et al.*, 2025; Sahu *et al.*, 2016). Second, despite their potentials, both sectors are still in the early stages of development, with limited knowledge and experience compared to more established fields like offshore wind energy. To date, field research has been predominantly limited to near-shore environments (e.g., Gagnon and Bergeron, 2017) or to fresh water environments (e.g., Dörenkämper *et al.*, 2021), while offshore, both sectors have not been developed beyond a pilot level in the North Sea (UNITED Project, 2022).

Figure 1.1a shows sketch of a potential long-line structure for aquaculture, composed of anchors, mooring lines, buoys at the surface and a vertical net in between. Figure 1.1b shows how this structure is installed in the sea close to an offshore wind farm in the southern North Sea.



(a) Sketch of a potential kelp longline structure (Santjer *et al.*, 2024b). **(b)** Photo of such a structure installed offshore.

Figure 1.1: Example of offshore kelp cultivation in the North Sea from the German Pilot of the UNITED project (Strothotte *et al.*, 2021).

For both, offshore aquaculture and FPV, hydrodynamic impacts such as waves or currents need to be considered for structural safety, but so is the water quality relevant for aquaculture, to enable species to grow properly and to optimise economic outcomes and enable the transition to commercial-scale operations. The relation between these relevant variables is highly complex and interdependent. This needs to be considered in the risk assessment of offshore structures, e.g. in determining the failure probability of a structure given joint behaviour of variables. Montes-Iturrizaga and Heredia-Zavoni (2016) investigated the dependence of wave height and period on the reliability of mooring lines of floating structures. But yet to date, no multivariate analysis was done considering ecological or hydrodynamic variables for offshore structures such as offshore aquaculture or FPV.

The main area of interest for these applications in this dissertation is the southern North Sea, due to strong support from the European Union for advancing sustainable

use of the sea and the numerous studies already conducted in this region on different applications. Despite the extensive availability of high-quality data for various relevant variables in the North Sea, derived from in-situ measurements, remote sensing and high-resolution numerical models such as e.g. CMEMS, the question remains on how to bring this information together for better and safer operation of offshore structures. When adopting a probabilistic approach, independence of the variables is typically assumed. To address the limitations of this assumption, this dissertation employs probabilistic methods that incorporate multivariate dependencies through the use of copula models.

Section 1.1 introduces copula models as the methodological foundation of this dissertation and outlines the motivation for their selection. This is followed by Section 1.2, which describes their concept and provides additional methodological details. Section 1.3 introduces the role of dependence among variables when defining suitable areas for offshore aquaculture installation and operation, while Section 1.4 emphasises the role of multivariate modelling in understanding dynamic mooring tensions under offshore hydrodynamic impacts. How a complex multivariate modelling method can help understanding the interplay of hydrodynamic and structural variables of a floating photovoltaic platform is introduced in Section 1.5. Finally, Section 1.6 describes the overall objective and scope and defines research questions answered throughout this dissertation. [1]

1.1. DEPENDENCE MODELLING VIA COPULA MODELS

Offshore environments are characterised by harsh, highly variable and uncertain conditions, with complex interactions between several processes. In such settings, probabilistic models are a suitable approach for capturing uncertainty and interdependencies among variables.

Copula models are chosen as the central methodological approach, as they offer a flexible way to describe the joint distribution of random variables. In particular, copula models allow to model the dependence structure between variables independently of their marginal distributions (e.g., Genest and Favre, 2007; Joe, 2015; Nelsen, 2006). Different to other popular stochastic modelling approaches, they are able to capture complex and non-linear dependence patterns commonly observed in environmental data, including asymmetries and tail dependence, which may be particularly relevant under extreme offshore conditions and, for example, in the interaction between environmental forcing and structural response.

Copula-based models are a popular approach to model dependence structures among random variables and have been successfully applied across a wide range of disciplines, including finance, health sciences, environmental research and engineering. Building on these advantages, this dissertation implements and compares different copula-based methods tailored to distinct offshore applications, ranging from ecological suitability assessments to structural safety and response modelling. Their applicability, strengths and limitations are explored in detail. Accordingly, the focus of this dissertation lies

[1] This dissertation is primarily structured as a collection of scientific publications. Some information in this chapter overlaps with the content of the chapter-specific introductions of Chapter 2 to 4. Additionally, some definitions may be repeated with potentially different notations across these chapters.

in analysing and improving the understanding of dependence among environmental, hydrodynamic and structural variables in offshore systems.

1.2. THE CONCEPT OF COPULA MODELLING

The concept of copulas is based on Sklar's theorem (Sklar, 1959), which states that any d-dimensional joint distribution can be described in terms of a set of uniform one-dimensional marginal distributions in the interval $[0,1]$ and a d-dimensional copula that models the dependence between random variables. For the 2-dimensional case, let $H(x,y)$ for $(x,y) \in \mathbb{R}^2$ be a joint distribution with univariate marginals $F(x)$ and $G(y)$. Then, there exists a copula C in the unit square $I^2 = ([0,1] \times [0,1])$:

$$H(x,y) = C(F(x), G(y)) \tag{1.1}$$

Equation (1.1) is satisfied for all $(x,y) \in \mathbb{R}^2$ (Genest and Favre, 2007; Joe, 2015; Nelsen, 2006). Following that, the density of the multivariate distribution is:

$$h(x,y) = c(F(x), G(y)) \cdot f(x) \cdot g(y) \tag{1.2}$$

Where c is the density of the copula and $f(x)$ and $g(y)$ are the univariate densities of variables x and y.

The dominantly used parametric copulas considered in the scope of this dissertation are: the Gaussian, Student-t, and three Archimedean copulas, namely Frank, Clayton, and Gumbel (see Figure 1.2). Each of them has different characteristics, with some capturing asymmetries in joint distributions. A description of each copula model is given below.

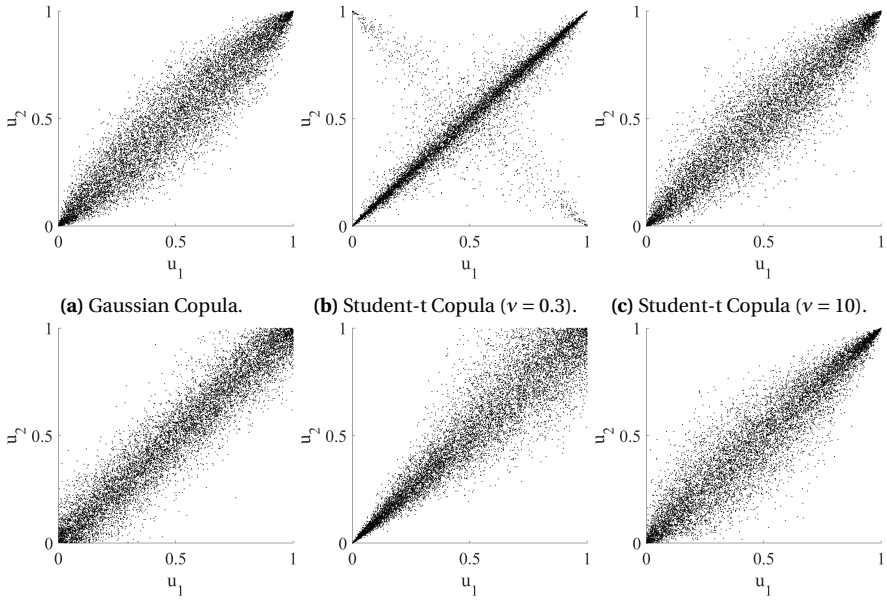
Gaussian copula: This copula, also called the normal copula, is often applied due to its flexibility. Its copula parameter is identical to Spearman's rank correlation coefficient ρ. The density of this copula is ellipse-shaped and therefore symmetric.

Student-t copula: Similar to the Gaussian copula, the density of the student-t copula is also elliptical, however, it is rather star-shaped with a stronger dependence in the tails. This copula is the only copula considered in this dissertation with two parameters, namely the correlation coefficient ρ (identical to the Gaussian copula) and a shape parameter ν, describing the degree of freedom. As ν increases, this copula approaches the Gaussian copula (see Figure 1.2c).

Frank copula: The third symmetric copula is the Frank copula. However, unlike the Gaussian copula, it it is tail independent.

Clayton copula: The Clayton copula is an asymmetric copula, meaning that, with positive correlation, it is characterised by tail dependence in the lower-left quadrant.

Gumbel copula: Similar to the Clayton copula, this copula is also asymmetric, but has an upper tail dependence in the top-right quadrant under positive correlation.



(a) Gaussian Copula. **(b)** Student-t Copula ($\nu = 0.3$). **(c)** Student-t Copula ($\nu = 10$).

(d) Frank Copula. **(e)** Clayton Copula. **(f)** Gumbel Copula.

Figure 1.2: Visualisation of different parametric copula models with $\rho = 0.8$.

Both asymmetric copulas, the Gumbel and the Clayton copula, are characterised by tail dependence. Thus, in the process of copula fitting, particularly of these two copulas, their rotation clockwise by 90°, 180° or 270° is considered.

1.3. THE ROLE OF DEPENDENCE IN AQUACULTURE SUITABILITY MODELLING

As mentioned above, offshore aquaculture has lots of potential for the future but several limitations and challenges hinder their development and implementation. While it is not just a question on how to design such structures, a big concern is where to place them, especially given conflicts with other uses of the sea. Chapter 2 of this dissertation addresses suitability mapping for aquaculture. Despite the abundance of existing literature analysing site suitability and advancing spatial planning of aquaculture at sea, many of them focus on the co-location of wind farms with offshore aquaculture, such as (e.g., Benassai *et al.*, 2014; Di-Tullio *et al.*, 2017; Gimpel *et al.*, 2015; Stockbridge *et al.*, 2025; Weiss *et al.*, 2018). Most of the available studies related to spatial planning often consider suitability or sustainability indexes (e.g., Benassai *et al.*, 2014; Di-Tullio *et al.*, 2017; Medina Madariaga *et al.*, 2025). Only a few approach the topic probabilistically. Jossart *et al.* (2020) for instance makes use of spatial autocorrelation to analyse spatial dependence. Coccoli *et al.* (2018) on the other side employ a discrete Bayesian Belief Network, while Weiss *et al.* (2018) conduct a probabilistic assessment determining potential areas using independent variables.

Chapter 2 of this dissertation will present a framework on how to create a suitability map for mussel or kelp cultivation in the North Sea related to ecological variables. Different to existing studies, dependence among relevant variables is considered. Thus, the focus is on the presented method and the importance of considering dependencies in determining suitable cultivation locations. For mussel and kelp cultivation, each, three relevant variables are chosen, while two variables are conditioned on the water temperature as the most influential variable. For each variable pair, parametric copula models are fitted because the available data is limited in time and could potentially not capture long-term extremes. In addition to that, fitting parametric copula models allow to better understand the relation and potential asymmetries between variables.

Using two fitted parametric copula models, a 3-dimensional sample cloud is generated, representing the dependence structure of the chosen variables, in which two variables are conditioned on the water temperature. Given a high number of samples, the probability can be determined, that certain growth or survival thresholds are met. Results indicate that for just three variables, accounting for their interdependence rather than treating them independently can lead to a difference of up to 40% in the probability of meeting the thresholds. This shows the importance of considering dependence among variables in site selection of offshore aquaculture.

1.4. GAUSSIAN COPULA-BASED BAYESIAN NETWORKS FOR DYNAMIC MOORING TENSIONS

After providing suitability maps for offshore aquaculture based on ecological variables and emphasising the importance of accounting for the dependence among variables in Chapter 2, Chapter 3 considers a method for multivariate dependence modelling based on copula models. This is done for hydrodynamic variables relevant for floating offshore structures, such as aquaculture or FPV installations. Specifically, the focus of this application is on the dynamic tension loads experienced by the mooring systems of any floating offshore structure. As previously highlighted, offshore structures are subjected to harsh environmental conditions, including extreme events. In such scenarios, the safety of these structures is no longer assured, and the mooring systems responsible for maintaining their stability may be at risk of failure. Most critical in these cases are snap forces, which are sudden increases in tensions in the mooring line, after they become temporarily slack. These dynamic snap tensions can exceed static tensions by a multiple (Governo *et al.*, 2023; Guo *et al.*, 2017; Hsu *et al.*, 2017; Miškov *et al.*, 2023). Several studies on mooring line tensions exist, related to field measurements (Nguyen *et al.*, 2019), numerical investigations (Desiré *et al.*, 2023; Feng *et al.*, 2021; Hermawan and Furukawa, 2020; Nasyrlayev *et al.*, 2023) and to experiments (Hsu *et al.*, 2017; Hsu *et al.*, 2019; Somoano *et al.*, 2024), however most of them do not consider snap loads. Therefore, snap loads, characterised by peak tension rates, are analysed in Chapter 3. Different locations in the North Sea are selected and up to 13 variables are considered relevant. Among some existing probabilistic analyses related to tensions in mooring lines (Li *et al.*, 2016; Montes-Iturrizaga and Heredia-Zavoni, 2016; Wu *et al.*, 2022; Zhao and Dong, 2023),

Montes-Iturrizaga and Heredia-Zavoni (2016) determined multivariate modelling as a potential next step. Therefore, a probabilistic graphical model representing the joint distribution of the variables is chosen, called Gaussian copula-based Bayesian Network (GCBN) (Hanea *et al.*, 2015; Hanea *et al.*, 2006).

GCBN's are chosen because they are computationally efficient and have no limitation on the number of variables. Additionally they help understanding the underlying physics. GCBN's are directed acyclic graphs (DAGs) consisting of nodes, which represent random variables, and arcs, representing the probabilistic dependence between the nodes. This probabilistic dependence is described via bivariate copula models (see Section 1.2). Nodes without predecessors are described by marginal distributions while nodes with predecessors are described by conditional distributions. The absence of arcs between nodes indicate conditional independence between two variables.

To keep the joint distribution flexible and computationally efficient, one-parameter copulas, which are parameterised by Spearman's rank correlation coefficient ρ, are used for each edge. Commonly, the Gaussian copula is chosen, as in this dissertation. The method of the Gaussian copula-based Bayesian network (GCBN) allows for updating distributions given observations (i.e. inference) and sampling for different conditioning scenarios. Therefore, the second goal of Chapter 3 is to assess tension rates of mooring lines conditioned on certain hydrodynamic conditions.

The joint probability density function (pdf) of a GCBN with m variables can be described as:

$$f_{1,\dots,m}(x_1,\dots,x_m) = f_1(x_1) \prod_{v=2}^{m} f_{v|Pa(v)}(x_v|x_{Pa(v)}) \tag{1.3}$$

where the joint density of m variables is denoted as $f_{1,\dots,m}$, the marginal densities are denoted by f_v and the conditional densities are denoted by $f_{v|}$.. Each node v represents a variable X_v, while the previous nodes of v form the set $Pa(v)$ (Hanea *et al.*, 2015).

1.5. VINE COPULAS FOR COMPLEX FLOATING STRUCTURES

Building on the multivariate dependence modelling based on GCBN introduced in Chapter 3, the final chapter addresses a key limitation of GCBNs: their restricted ability to represent asymmetric dependencies. This constraint becomes particularly relevant for offshore structures, where the interaction between environmental loads and structural response is often characterised by asymmetric dependence patterns such as pronounced tail dependence. To overcome this limitation, Chapter 4 introduces vine copula models, a sequence of 'linked' trees whose building blocks are bivariate copulas. This structure enables a full factorisation of the joint density of a d-dimensional distribution into marginal densities and conditional bivariate copula densities, allowing asymmetries and complex dependence features to be captured explicitly (Aas *et al.*, 2009; Czado, 2019).

$$f_{1,\dots,d}(x_1,\dots,x_d) = f_{1|2,\dots,d}(x_1|x_2,\dots,x_d)\ f_{2|3,\dots,d}(x_2|x_3,\dots,x_d)\ ,\dots,\ f_d(x_d) \tag{1.4}$$

The method is applied to floating offshore structures, particularly a FPV structure, for which structural responses under harsh offshore conditions are analysed. While FPV research has so far focused primarily on field deployments in lakes (Dörenkämper *et al.*, 2021; Karatas and Yilmaz, 2021; Kjeldstad *et al.*, 2022; Manoj Kumar *et al.*, 2022), laboratory experiments (Ji *et al.*, 2025; Sree *et al.*, 2022), and numerical studies (C.J. *et al.*, 2024a; Dörenkämper *et al.*, 2021; Ikhennicheu *et al.*, 2021; Ji *et al.*, 2025; Jifaturrohman *et al.*, 2024), probabilistic approaches remain rare. An exception is Zhou *et al.* (2022), who applied neural networks and copulas but also highlighted the suitability of vine copula models for representing the dependence patterns in FPV systems. The research in this dissertation therefore extends the probabilistic modelling of FPV structures by employing vine copulas to examine dependence among hydrodynamic variables, shear forces and longitudinal stresses.

A further methodological contribution of this chapter is the treatment of mixed continuous and ordinal datasets. While vine copulas are traditionally formulated for fully continuous variables, engineering datasets frequently include ordinal variables or variables that are available only in a limited number of possible values. To appropriately handle the discontinuities in their empirical distributions, the vine models in this chapter explicitly incorporate ordinal variables alongside continuous ones, thereby ensuring a consistent representation of mixed data within the vine structure.

1.6. OBJECTIVE AND SCOPE

Floating offshore structures are exposed to severe environmental impacts. In addition, there is a competition about space at sea among different sectors. The promising sectors of offshore aquaculture and FPV hold considerable potential but remain in their early stages of development. This dissertation examines the impact of environmental factors, both ecological and hydrodynamic, on various offshore structures, as well as their interdependencies, with the aim of advancing the scientific understanding of floating offshore systems and mitigating associated risks to facilitate their practical implementation in emerging sectors.

Figure 1.3 gives a graphical overview of the research of this dissertation. First, in Chapter 2, bivariate copula models are used to support spatial planning based on ecological variables for mussel and kelp cultivation. This is done for the whole area of the southeastern North Sea. In Chapter 3 the tension rates in a mooring line of an arbitrary floating structure are assessed under hydrodynamic impacts for 12 locations spread across the southeastern North Sea, whereas the location 4 serves as a reference location. Therefore, a multivariate approach using Gaussian copula-based Bayesian Network (GCBN) is employed at each of the studied locations. However, GCBNs are limited to Gaussian copulas to describe the dependence between variables. This copula might not be the best choice, particularly in extreme value analyses, given potential asymmetries in the dependence patterns of variables. Therefore, vine copula models are introduced as a more flexible alternative in Chapter 4, as this method is not limited to a single copula model only. This method is applied at the reference location for an offshore FPV structure to assess the dependence of several hydrodynamic and structural variables, which can be continuous or ordinal.

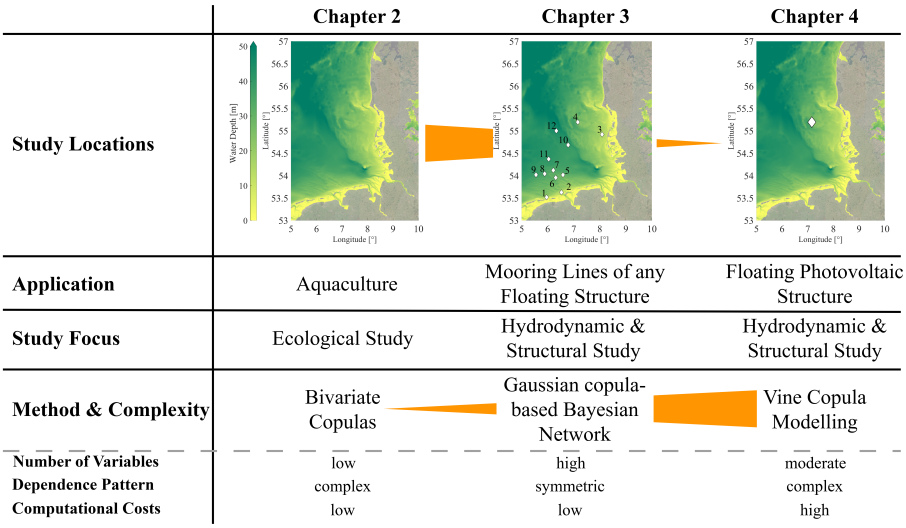


Figure 1.3: Graphical overview of the dissertation chapters, categorising study locations, applications, study focus, methods and their corresponding levels of complexity.

Accordingly, this dissertation progresses from a relatively simple method applied across a broad spatial domain to more complex models focused on a few selected locations, and ultimately on a single site in the North Sea. Figure 1.3 gives initial information on the characteristics and differences of the applied methods. These are discussed further in Chapter 5, both from a conceptual/mathematical perspective and in relation to the applications presented in this dissertation. In addition, Section 5.2.2 provides a roadmap for engineering, helping to select a suitable method for a given application.
This dissertation explores the applicability of probabilistic methods, particularly multivariate copula-based methods, to support the development, spatial planning and structural safety of floating offshore structures, such as offshore aquaculture and floating photovoltaic systems. The overarching research question (RQ) of this dissertation is therefore: *How can multivariate copula-based probabilistic models support spatial planning and structural safety assessment in offshore engineering?* To address this question, the study is structured around three research avenues, answering the questions below, which may differ in scope and specificity from the more detailed questions addressed in Chapter 2 to Chapter 4.

RQ1: How does variable dependence affect the accuracy of aquaculture site suitability maps?

RQ2: How can Gaussian copula-based Bayesian Networks support safety assessment of floating offshore structures by modelling mooring tensions given specific hydrodynamic conditions?

RQ3: How can a vine copula model represent the complex dependence among mixed continuous and ordinal structural and environmental variables of a floating offshore platform?

2

A PROBABILISTIC FRAMEWORK FOR OFFSHORE AQUACULTURE SUITABILITY ASSESSMENT USING BIVARIATE COPULAS

This chapter presents a probabilistic framework for assessing site suitability for offshore aquaculture, focusing on ecological variables relevant to mussel and kelp cultivation. It introduces the use of bivariate copula models to account for dependence among variables, which is often neglected in studies related to marine spatial planning. By contrasting dependent and independent modelling approaches, this chapter highlights how accounting for dependence among variables can significantly influence suitability outcomes by up to 40%. This sets the foundation for the later chapters, which expand the methodology towards more complex multivariate modelling.

2.1. INTRODUCTION

The increasing human world population (Tripathi *et al.*, 2019) leads to an increased demand for sustainable food sources and green energy (Yong *et al.*, 2022). Consequently, the production sector of aquaculture has experienced rapid global growth, continuing to expand despite the pandemic in recent years (FAO, 2022; Michler-Cieluch *et al.*, 2009). Species such as blue mussels and especially sugar kelp have the potential to be used as food, source material for pharmaceutical industry, biomaterial or bioenergy sources such as biogas or bioethanol (Fernand *et al.*, 2017; Kammler *et al.*, 2024; Kerrison *et al.*, 2015; Schiener, 2014), thus contributing to the solution to climate change (Yong *et al.*, 2022). However, available space for aquaculture near-shore is rare. Thus, species cultivation

This chapter has been published in **Aquacultural Engineering** as *A probabilistic framework for offshore aquaculture suitability assessment using bivariate copulas* in 2024, Santjer *et al.* (2024a). Minor editorial changes in this dissertation are implied.

offshore becomes more relevant, even though the available space is limited and the competition at sea among various sectors pose another challenge to the novel sector of offshore aquaculture (Buck and Langan, 2017; Buck *et al.*, 2018). As a result, aquaculture of certain species combined with wind energy production have emerged as a promising multi-use solution to address the above raised challenges as supported by the European Union (European-Commission, 2010; European-Commission, 2012). Moreover, there is evidence to suggest that offshore conditions can lead to increased growth and better product quality of some species (Brenner *et al.*, 2012; Buck, 2002). However, while the offshore wind energy sector is experiencing a rapid development (Michler-Cieluch *et al.*, 2009), offshore aquaculture, as well as the combination of several sectors, still needs to be further investigated and developed. Thus, the presented study provides a framework to assess spatial suitability of offshore aquaculture from an environmental perspective.

Although the importance of co-use and multi-use has been recognised and emphasised by both the research community and policy makers (e.g. Buck *et al.*, 2004; Wever *et al.*, 2015), experience in the field of offshore aquaculture is limited and thus comes with several risks and uncertainties. First steps given by the scientific community point towards the technical and biological suitability of offshore cultivation of certain species (Buck *et al.*, 2008; Maar *et al.*, 2023). Technical feasibility was confirmed by Gagnon and Bergeron (2017), who analysed an exposed submerged mussel long-line structure, 4 km off the Canadian east coast. The measured forces on the structure were significantly smaller than the breaking strength of the lines and no damage such as mussel fall-off was observed. As defined by (Buck *et al.*, 2024) , "offshore" refers to a location with a distance of more than 3 nautical mile from the shore, while "exposed" relates to the physical oceanographic conditions. In an European project (UNITED Project, 2022), a pilot case was conducted to test the cultivation blue mussels and sugar kelp at an exposed offshore site. Regarding the biological feasibility, Brenner *et al.* (2012) and Buck (2002) suggested that offshore conditions can lead to increased growth and better product quality of some species due to the good water quality and oxygen concentrations (Hopkins *et al.*, 1996). Opposite conclusions were reached by Stechele *et al.* (2022), where less productivity was observed offshore than near-shore. However, it should be noted that these authors performed this research for a self-regulating cultivation technique, where wild larvae spat settles and cultivates at the structure. There are other factors in favour of offshore cultivation, such as limited inter-annual variability and lower risk of toxic algae blooms. Other benefits of the open ocean are sufficient water depths and the existence of infrastructure, especially when co-locating aquaculture with wind parks (Buck, 2002; Maar *et al.*, 2023). This allows for shared vessels, operation and maintenance. Thus, offshore aquaculture as single-use, or in the context of multi-use, has a great potential leading to a growing need for tools to support the development of these novel sustainable offshore aquaculture practices (Bergström and Lindegarth, 2016).

In previous studies, different approaches have been used to explore the suitability of offshore aquaculture practices. Benassai *et al.* (2014) focused on the co-location of offshore wind and open-water aquaculture in Danish waters and the southern part of the North Sea. These authors developed a "sustainability" index which provided a first large-scale evaluation of the suitability of aquaculture in offshore wind park sites. This

index was computed through a scoring model which considers physical limitations, such as water depth, wind velocity and water temperature, and biological parameters, such as chlorophyll-a concentrations.
Gimpel *et al.* (2015) developed a marine spatial planning (MSP) tool to evaluate spatial co-location scenarios of offshore aquaculture integrated in wind farms. This tool accounts for environmental, economic, inter-sectoral and social-cultural risks and opportunities and considers 13 species native in the German North Sea, resistant to offshore hydrodynamic conditions and economically interesting for the European market. Di-Tullio *et al.* (2017) also proposed a "sustainability index" to assess the co-location of offshore wind and open-water mussel cultivation using remote-sensing data for physical and biological variables , following the research done in Benassai *et al.* (2014). Geisler *et al.* (2018) conducted a feasibility study on various species at the FINO3 research platform in the German North Sea, located 80 km off the coast. In this feasibility study, a scoring model was applied, considering biological, technical and economical/political aspects for several different species. The aforementioned authors identified the blue mussel *Mytilus edulis* and sugar kelp *Saccharina latissima* as the most suitable species for cultivation at this location in the North Sea. Therefore, to the author's knowledge, methods available in the literature to identify potential locations for offshore aquaculture are deterministic so they do not directly account for the natural variability (uncertainty) of the relevant variables or the probabilistic dependence between them. This is, each variable is included in the analysis through a single statistic, but variables in nature typically present a stochastic behaviour. Moreover, the relationship between variables is not explicitly considered, although knowing information about one variable can provide insights about the distribution of another variable. For example, if the water temperature rises, it is likely that the concentration of dissolved oxygen will decrease but a deterministic estimation of one variable solely based on the other is typically not realistic. This indicates a probabilistic relationship between the two variables. In studies such as Ruiz-Velazco *et al.* (2013) or Santjer *et al.* (2023), explicitly accounting for the relationship between environmental variables was shown to play a relevant role in feasibility assessments. In Ruiz-Velazco *et al.* (2013), a linear relationship between variables was assumed to predict the shrimp production under commercial conditions, highlighting the importance of the interactions between relevant variables. Santjer *et al.* (2023) took the first steps to include the probabilistic dependence between variables and explored the applicability of different bivariate copulas to model the probabilistic dependence between environmental variables to assess the suitability of cultivating blue mussels and sugar kelp in a location in the German North Sea. Best fitting models were identified and a significant influence of the probabilistic dependence was reported. Moreover, bivariate copula models have been widely used in existing literature to model multivariate joint distribution of variables (e.g.: Leontaris *et al.*, 2016; Mares-Nasarre *et al.*, 2024; Ragno *et al.*, 2023), being concluded that considering the dependence between the variables has a severe impact on the results. Thus, there are cases where the probabilistic dependence plays a key role and needs to be accounted for Ju *et al.* (2014). Another recent study made use of copula models, describing the joint dependence between two variables to predict sedimentation processes in aquaculture systems (López-Rebollar *et al.*, 2024). Consequently, the present study aims to provide a probabilistic framework which accounts for the dependence between several ecological

variables using bivariate copulas to assess the suitability of offshore aquaculture. This framework is applied in the south-eastern North Sea (see Figure 2.1) to the cultivation of *Mytilus edulis*, hereon blue mussels, and *Saccharina latissima*, hereon sugar kelp. These two species have been identified as suitable using a deterministic scoring model by Geisler *et al.* (2018). This methodology allows to identify ecologically suitable locations for offshore cultivation and, thus, can potentially contribute to the development of multi-use approaches that could take advantage from offshore environments (Buck, 2002), the design of marine spatial plans and the planning of a more sustainable use of offshore environments.

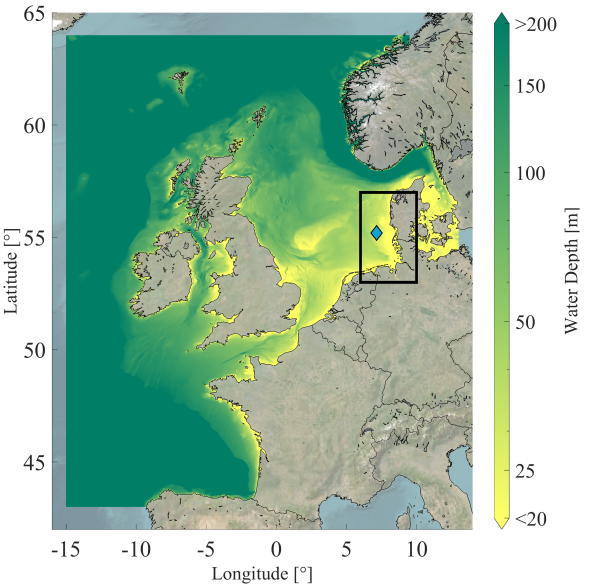


Figure 2.1: Water depths for the European Shelf. The area of interest (AOI) of this study is indicated with a black frame, and the location of the FINO3 research platform is marked by a blue diamond.

The present paper is structured as follows. First, the three-dimensional hydrodynamic and ecological numerical model used to extract the observations of the environmental variables is described in Section 2.2.1. In order to assess the suitability of blue mussels and sugar kelp, three ecological variables per species are selected based on the existing literature in Section 2.2.2. Also, for each of the variables, the optimal growth and critical survival limits are defined. The methodology to build the probabilistic model proposed in this study is described in Section 2.3.1. For each observation point, daily extremes are extracted during the cultivation months of each species. The dependence for each pair of variables is first assessed through correlation analysis in Section 2.3.1 and, later, the best fitting bivariate copula model for each pair of variables is investigated across the AOI in Section 2.3.2). The probabilities of meeting the critical and optimal limits for each species are computed using the developed probabilistic models in Section 2.3.3. These results are

compared to the probability calculations assuming independence between the variables to assess whether the dependence plays a key role in the process. The strengths and limitation of the developed framework, its transferability to other species and areas and the obtained results are discussed in Section 2.4. Finally, conclusions are drawn in Section 2.5.

2.2. MATERIALS AND METHODS

In this section, the steps of the proposed probabilistic framework for performing suitability analysis in aquaculture are described, as well as the case studies used to showcase its application. The methodology is applied to two example species, namely blue mussels and sugar kelp, to assess their spatial suitability for offshore cultivation. A flowchart providing an overview of the method described below is displayed in Figure 2.2.

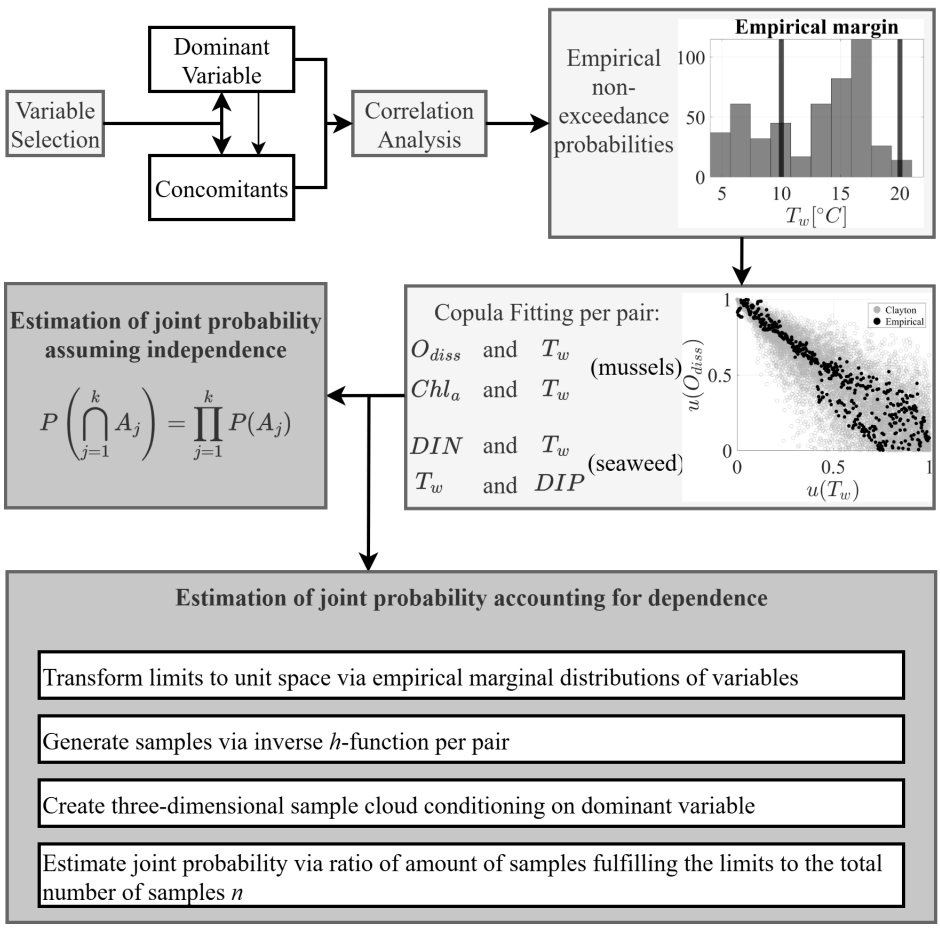


Figure 2.2: Flowchart displaying the main procedures of the described methodology.

2.2.1. 3D HYDRODYNAMIC AND WATER QUALITY MODEL

The environmental conditions in the area of interest, AOI, marked by a black rectangle in Figure 2.1, are extracted from model outputs of the 3D Dutch Continental Shelf Model - Flexible Mesh (3D DCSM-FM) (Zijl *et al.*, 2023). Hydrodynamics and water quality processes are computed using the D-Flow Flexible Mesh (D-Flow FM) component from the Delft 3D Flexible Mesh Suite (Deltares, 2023a). Water quality processes are simulated with the D-Water Quality module (Deltares, 2023b), fully integrated within D-Flow FM.

In this study, the coupled hydrodynamic-water quality 3D DCSM-FM version described in (van Leeuwen *et al.*, 2023) is used. The 3D DCSM-FM model domain covers the entire North-Western European Shelf from 15°W to 13°E and from 43°N to 64°N, including the North Sea (Figure 2.1). The horizontal model grid is coarser near the offshore boundaries and in deep waters (≈ 4x4 nautical miles) and the resolution increases toward the shallower waters and in the Southern North Sea to 0.5x0.5 nautical miles. The water column is represented using a z-sigma layer approach. The top 100 m (or less in shallower areas, such as at the AOI) are divided into a fixed number of 20 uniform layers (sigma-layers). Beneath 100 m-depth the water column is divided into z-layers, at fixed depths over the entire domain, with thicknesses increasing exponentially by a factor 1.19 towards the deeper layers. The hydrodynamic component of the model simulates water levels, currents and tides as well as temperature and salinity. The water quality component of the model simulates the most relevant processes in the cycling of major nutrients (nitrogen, phosphorus and silicate), of organic carbon and dissolved oxygen. Water quality processes parameterisation is based on Blauw *et al.* (2009).

The model was run for 2 years (2014 and 2015), using a 2-year spinup (2012-2013 conditions), allowing to account for some inter-annual variability. These years were selected, as these are the most recent years computed and validated by van Leeuwen *et al.* (2023). Model output is produced at an hourly timestep and a 4 nautical mile spatial resolution for the AOI (from 6°E to 10°E and 53.5°N to 57°N, see Figure 2.1).

2.2.2. VARIABLE SELECTION AND SAMPLING

In this study, the proposed methodology is applied to assess the suitability of offshore cultivation of two species, blue mussel *Mytilus edulis* and sugar kelp *Saccharina latissima*, in the south-eastern North Sea. Here, their growth differences and variables affecting their growth are described for each species. Many factors can impact the growth of these species, such as hydrodynamic variables (waves and currents) and light availability, especially for exposed locations offshore. However, since there is already existing knowledge on the suitability of such offshore aquaculture based on hydrodynamic variables (e.g., Azevedo *et al.*, 2019; Buck and Buchholz, 2005; Geisler *et al.*, 2018), they are disregarded here. Instead, the focus is on ecological variables, as these are as important for assessing the suitability of sites (Schmidt *et al.*, 2018). As salinity is not limiting for the respected species in the North Sea, it is not considered here either.

Three variables per species are considered to assess their cultivation suitability based on the existing literature (references in Table 2.1). The water temperature T_w [$^\circ C$] is selected for both species. In addition, the dissolved oxygen O_{diss} [g/m^3] and chlorophyll-a

concentrations Chl_a $[mg/m^3]$ are considered for the blue mussels. For the sugar kelp, the dissolved inorganic nitrogen DIN $[g/m^3]$ and the dissolved inorganic phosphorus concentration DIP $[g/m^3]$ are selected in addition to T_w. Further information about the variables is given below, while an overview of the selected variables can be found in Table 2.1. This Table also contains information about the optimal and critical limits per variable, which are further described below per species, and the references used to define the variables and their limits. Time series including the limits per variable for both species can be found in Section A.1 for the research platform FINO3 (see location of the blue diamond in Figure 2.1, which is used as a reference location in this study.)

Table 2.1: Overview of the selected three relevant variables and their limits for optimal growth and survival per species, based on literature.

Blue mussels			
Variable	Optimal	Critical	References
T_w $[°C]$	$10-20$	≤ 25	Fly and Hilbish (2013) and Kamermans *et al.* (2022)
O_{diss} $[g/m^3]$	>2	≤ 2	Belivermiş *et al.* (2020) and Tang and Riisgård (2018)
Chl_a $[mg/m^3]$	$3-10$	>0.5	Filgueira *et al.* (2009), Kamermans *et al.* (2022), Pascoe *et al.* (2009), and Riisgård *et al.* (2011)
Sugar kelp			
Variable	Optimal	Critical	Reference
$T_w[°C]$	$5-15$	≤ 20	Kerrison *et al.* (2015)
DIN $[g/m^3]$	>0.35	>0.098	Jevne *et al.* (2020)
DIP $[g/m^3]$	>0.0484	>0.0136	via Redfield Ratio Redfield (1934)

BLUE MUSSEL, *Mytilus edulis*

In this study, the cultivation of the blue mussels *Mytilus edulis* is considered as growing at submerged long-lines at sea. Buck (2007) investigated such cultivation in the German Bight and concluded that these structures for mussel cultivation should be at least 6 to 7 m below mean surface level (MSL) to avoid high impacts of waves. The long-line was installed at a depth of 5 m, while the cultivation range was from 5 to 8 m. Another pilot project in the German North Sea (Strothotte *et al.*, 2021) far offshore cultivated the species at depths of 7 to 11 m below MSL. Therefore, 9 m below MSL is selected in the current study and the selected ecological variables are extracted from the numerical model at that depth. Although the blue mussels are growing throughout the whole year, the main growth season is between March and October. Therefore, data from these months is selected for the analysis.

As already described, the selected variables are water temperature T_w $[°C]$, dissolved oxygen O_{diss} $[g/m^3]$ and chlorophyll-a concentration Chl_a $[mg/m^3]$ based on existing

literature (Bergström and Lindegarth, 2016; Kamermans *et al.*, 2022; Li *et al.*, 2022; Tang and Riisgård, 2018). Blue mussels are relatively resistant to extreme water temperatures (high and low), but optimal growth conditions exist for certain temperature ranges according to Fly and Hilbish (2013). The minimum dissolved oxygen concentration to enable growth of blue mussels should be above 2 g/m^3 (Tang and Riisgård, 2018; Tyler-Walters and Hiscock, 2008). This is similar to the concentrations of Chl_a, which should not fall below 0.5 mg/m^3 to ensure the survivability of the species. Thus, minima are selected to be most critical for both, O_{diss} and Chl_a.

SUGAR KELP, *Saccharina latissima*

The second species considered in the current study is the sugar kelp *Saccharina latissima*. According to Peteiro and Freire (2011), the ropes for the sugar kelp are recommended to be at depths of 2 to 4 m below MSL. In the pilot project in the German North Sea (Strothotte *et al.*, 2021), the sugar kelp is cultivated at depths of 0 to 4 m below MSL. Thus, here 2 m below MSL is selected. As the sugar kelp develops better in colder water temperatures, the main growth season is from September until May. Therefore, data from this period at a depth of 2 m is used in this study.
The growth and survival of the sugar kelp relies heavily on the water temperature T_w, as water temperatures above 20°C are lethal for the selected type of seaweed (Kerrison *et al.*, 2015). The temperature ranges that lead to the highest growth rates are given in Table 2.1. Besides, two additional variables are inevitable for the growth of the sugar kelp according to Buck and Buchholz (2004): dissolved inorganic nitrogen, DIN (composed of both, ammonium, NH_4, and nitrate, NO_3), and dissolved inorganic phosphorus, DIP (phosphate, PO_4), each in [g/m^3]. Since too low concentrations for both, DIN and DIP, are critical for the growth and survival of the species, they are selected in this study.

VARIABLE SAMPLING

In this study, the daily maxima of the water temperatures T_w at the appropriate depth (9 m and 2 m below MSL for blue mussels and sugar kelp, respectively) are selected. The concomitants of 4 hours before and after the maxima of T_w are determined for the other variables and their minimum values are selected.

Figure 2.3 shows an example of the data used for the relevant variables T_w, O_{diss} and Chl_a of the blue mussel case at the FINO3 location. The figure presents scatter plots per variable pair in their respective units, to allow for a comparison of pairwise relationships. The limits from Table 2.1 are represented by dashed lines. On the diagonal of Figure 2.3, the empirical probabilistic distribution functions (pdf's) per variable are shown. Figure A.8 in the Section A.6.1 shows the scatter data of the relevant variables T_w, DIN and DIP for the sugar kelp.

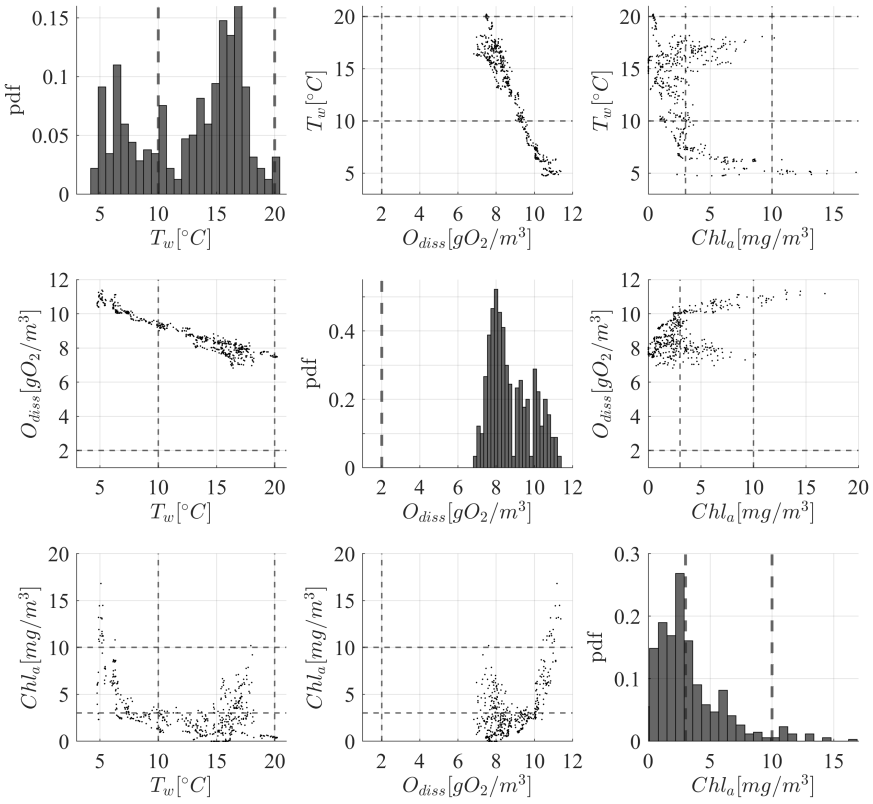


Figure 2.3: Overview of sampled data for blue mussels at the location of the FINO3 research platform together with the optimal limits from Table 2.1, while on the diagonal the empirical probabilistic distribution functions (pdf's) are shown.

2.2.3. SPEARMAN RANK CORRELATION

Spearman's rank correlation coefficient (Spearman, 1987), r, is calculated for each pair of variables as a starting point of the dependence analysis to assess the strength of the dependence between variables and its significance (e.g. Hanea *et al.*, 2006; Mares-Nasarre *et al.*, 2024; Santjer *et al.*, 2023). r computes Pearson's correlation coefficient (Pearson and Galton, 1895) between the ranks of each variable to assess monotonic relationships different to Pearson's which detects linear relationships. Thus, the Spearman's rank correlation coefficient r is used here to assess the degree of monotonic dependence between the random variables and is given by

$$r = \frac{Cov[R(X), R(Y)]}{\sigma_{R(X)}\sigma_{R(Y)}} \tag{2.1}$$

where $Cov[R(X), R(Y)]$ is the covariance of the ranked variates of X and Y, such as T_w or O_{diss}, and $\sigma_{R(X)}$ and $\sigma_{R(Y)}$ are their standard deviations. $r \in [-1, 1]$, where r = 1 and -1 represent the perfect (monotonic) positive and negative correlation, respectively. In order

to determine the significance of the observed correlations, the p-values are calculated. This significance test is done via the student-t distribution. Therefore, first the t-score is calculated (Zar, 1972):

$$t = r\sqrt{\frac{n-2}{1-r^2}} \tag{2.2}$$

where n are the degrees of freedom or sample size. Following, the p-value is defined as the combined area in both tails of the student-t distribution. Therefore, the p-value per tail can be determined via the cumulative distribution function of the student-t distribution evaluated at the t-score from Equation (2.2) with $(n-2)$ degrees of freedom. A given rank correlation is statistically significant if the p-value is below the significant level of $\alpha = 0.05$ and thus, it is unlikely that the observed correlation is due to chance.

The goal of this analysis is to assess the strength of the dependence between the selected variables (see Section 2.2.2) and whether they are significant along the AOI. This can be used as a first assessment of whether incorporating dependence between the variables is significant for suitability analysis.

2.2.4. COPULA MODELING

Copulas are a popular approach to model dependence between random variables (e.g., Genest and Favre, 2007; Joe, 2015; Nelsen, 2006). The concept of copulas is based on Sklar's theorem (Sklar, 1959): any multivariate joint distribution can be described in terms of a set of univariate marginal distributions and a copula that models the dependence between the variables. Thus, copulas are multivariate distribution functions whose one-dimensional margins are uniform on the interval $[0,1]$. For the bivariate case, as used in this study, copulas can be defined as follows:

Let $H(x,y)$ for $(x,y) \in \mathbb{R}^2$ be a joint distribution with univariate marginals $F(x)$ and $G(y)$, then there exists a copula C in the unit square $I^2 = ([0,1] \times [0,1])$:

$$H(x,y) = C(F(x), G(y)) \tag{2.3}$$

Equation (2.3) is satisfied for all $(x,y) \in \mathbb{R}^2$.

EMPIRICAL NON-EXCEEDANCE PROBABILITIES

In order to later calculate probabilities of meeting the optimal and critical limits (see Table 2.1), the univariate uncertainty of each variable needs to be modelled per location. This is done using the univariate empirical distribution functions, as the limits from Table 2.1 are within the range of observations and thus no extrapolation is required. Table 2.2 shows the empirical non-exceedance probabilities corresponding to those limits at the location of the FINO3 research platform. Note that $F_{T_w}(x)$ denotes here the empirical cumulative distribution function of the random variable T_w evaluated at x. Figure 2.4 gives an example of the empirical distribution functions of the three variables T_w, O_{diss} and Chl_a of the blue mussel case at the location of the FINO3 research platform, which are displayed together with the optimal limits indicated by vertical dashed lines.

Also, the empirical copulas are presented. For an example of the case of the sugar kelp, see Figure A.9 in Section A.6.1.

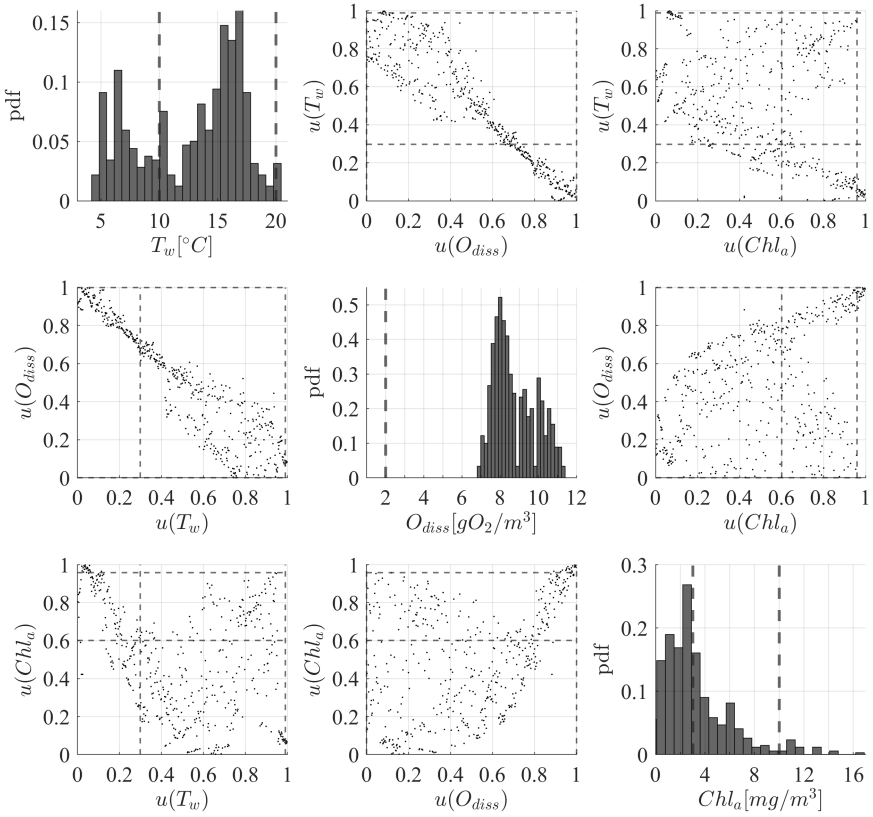


Figure 2.4: Overview of uniform observations of the sampled data for blue mussels at the location of the FINO3 research platform together with the empirical non-exceedance probabilities of the optimal limits (see Table 2.2), while on the diagonal the univariate empirical probabilistic distribution functions are shown together with the optimal limits from Table 2.1.

CHARACTERISTICS OF PARAMETRIC COPULA MODELS APPLIED IN THIS STUDY

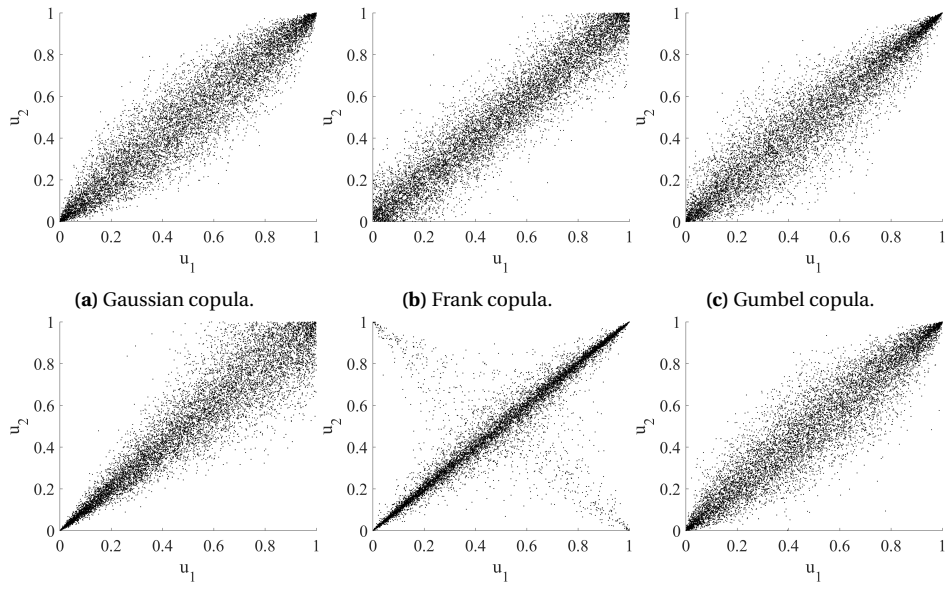
In this study, five commonly used parametric bivariate copula models , with different characteristics are considered to capture the possible asymmetric behaviour that the joint distributions of each pair may exhibit. Copula models can be used to describe these asymmetries, such as tail dependence. Upper tail dependence occurs when the correlation between the high values of the two modelled random variables is greater than that between lower values. Lower tail dependence occurs when the opposite behaviour is observed. For more information see Section A.2.

Table 2.2: Overview of the selected three relevant variables and the univariate probability of meeting the limits (optimal and critical, see Table 2.1), based on their empirical marginal distributions, here for an example location at the FINO3 research platform.

Blue mussels		
Variable	Optimal	Critical
T_w	$F_{T_w}(10) = 0.298$	$F_{T_w}(25) = 1$
	$F_{T_w}(20) = 0.989$	
O_{diss}	$1 - F_{O_{diss}}(2) = 1$	$1 - F_{O_{diss}}(2) = 1$
Chl_a	$F_{Chl_a}(3) = 0.601$	$1 - F_{Chl_a}(0.5) = 0.918$
	$F_{Chl_a}(10) = 0.959$	
Sugar kelp		
Variable	Optimal	Critical
T_w	$F_{T_w}(5) = 0.084$	$F_{T_w}(20) = 1$
	$F_{T_w}(15) = 0.802$	
DIN	$1 - F_{DIN}(0.35) = 0$	$1 - F_{DIN}(0.098) = 0.671$
DIP	$1 - F_{DIP}(0.0484) = 0$	$1 - F_{DIP}(0.0136) = 0.562$

The copula models considered in this study are described below. Figure 2.5 shows samples generated from the five different bivariate copulas considered in this study with the same rank correlation of $r = 0.8$, namely (*a*) Gaussian, (*b*) Frank, (*c*) Gumbel, (*d*) Clayton, and two t-copulas with different degrees of freedom ν in (*e*) and (*f*). Their bivariate cumulative distribution functions (cdf's) are given in Section A.3.

The density of the Gaussian copula is ellipse-shaped and thus symmetric (see Figure 2.5a). The distribution of the t-copula is also elliptical but has a star-like shape and stronger dependence in the tails. Different to the other considered copula models in this study, the bivariate t-copula has two copula parameters. The first parameter is the same as for the Gaussian copula, the correlation coefficient ρ, while the second parameter is the shape parameter ν, describing the degree of freedom. High degrees of freedom make the t-copula converge towards the Gaussian copula (see Figure 2.5f for $\nu = 10$). The density of the Frank copula is also symmetric and has similar characteristics to the Gaussian copula, with the difference that it has greater dependence in both tails than in the centre. The Gaussian, Frank and t-copula are symmetrical and thus are not characterised by any tail dependence (Joe, 2015). Different to that, the Gumbel copula has an upper tail dependence in the top right quadrant, as the correlation is higher in this quadrant of the bivariate distribution, while the Clayton copula has a lower tail dependence in the bottom left quadrant (see Figure 2.5c and Figure 2.5d). Because of the different tail dependencies of these two copula models, rotation might be needed during the copula fitting process. Therefore, the quadrant with the highest correlation of the variable pair is determined and the Gumbel and Clayton copula are rotated clockwise by 90°, 180° or 270°, if necessary, such that their tail dependence is in the quadrant with the strongest correlation.



(a) Gaussian copula. **(b)** Frank copula. **(c)** Gumbel copula.

(d) Clayton copula. **(e)** t-copula with $\nu = 0.3$. **(f)** t-copula with $\nu = 10$.

Figure 2.5: Samples generated for random variables with rank correlation of $r = 0.8$ for different copula families, each for $n = 10,000$ samples.

GOODNESS-OF-FIT MEASURES FOR COPULA SELECTION

Goodness-of-fit (GOF) measures are needed to assess and compare the performance of different parametric copula models. Two GOF measures are applied here. The first one is the Cramér-von Mises (CvM) statistic S_n, as described by Genest *et al.* (2009b). It determines the sum of squared differences between the empirical joint cdf and the joint cdf of the parametric copula model. S_n for a sample length of n is determined as:

$$S_n(\boldsymbol{u}) = n \sum_{|n|} \left\{ C_n(\boldsymbol{u}) - C_{\widehat{\theta}_n}(\boldsymbol{u}) \right\}^2 \quad , \boldsymbol{u} \in [0,1]^2 \tag{2.4}$$

Here, $C_n(\boldsymbol{u}) = \frac{1}{n} \sum_{i=1}^{n} 1\,(U_i \leq \boldsymbol{u})$ is the empirical copula and $C_{\widehat{\theta}_n}(\boldsymbol{u})$ is the parametric copula with estimated parameter $\widehat{\theta}_n$ from the sample. The lower $S_n(\mathbf{u})$, the better the GOF. For the calculation of the CvM statistics, the BANSHEE toolbox version 1.3 was used (see Paprotny *et al.* (2020) and Mendoza-Lugo and Morales-Nápoles (2023) for the MATLAB implementation and Koot *et al.* (2023) for the Python implementation).

The second measure to assess the GOF of the copula models are the semi-correlations, as described by Joe (2015). For this method, the uniform margins are transformed to standard normal $N(0,1)$ margins through the inverse of the standard normal univariate Gaussian distribution. This allows to better evaluate asymmetries such as tail dependence. Then, the correlation coefficients per quadrant are calculated. For positive correlated samples, the upper and lower semi-correlation coefficients are defined as:

$$\widehat{\rho}_N^+ = Cor\left[Z_1, Z_2 | Z_1 > 0, Z_2 > 0\right], \\ \widehat{\rho}_N^- = Cor\left[Z_1, Z_2 | Z_1 < 0, Z_2 < 0\right]. \tag{2.5}$$

The smaller the difference of the semi-correlation coefficients between the empirical and parametric copula, the better the fit. If the GOF results of the two described methods do not align, the results of S_n are followed.

2.2.5. JOINT MULTIVARIATE PROBABILITY CALCULATION

Once the parametric copulas for each pair of variables and locations are selected and fitted, the joint probabilities of meeting optimal growth and critical survival limits are computed for the respective growth period of the species. These probabilities are computed for each location with and without considering dependence between the variables to assess the role of dependence. Assuming independence, the joint probability can be calculated as follows:

$$P\left(\bigcap_{j=1}^{k} A_j\right) = \prod_{j=1}^{k} P(A_j) \tag{2.6}$$

where $P(A_j)$ is the probability of variable A_j meeting the optimal (or critical) condition. These probabilities are determined using the empirical cdf's of the variable data sets (see Section 2.2.4).

When considering dependence between variables, the fitted bivariate copulas are used as in higher dimensions multivariate copula modelling becomes more difficult and loses part of its flexibility. Here, a pair copula construction or a one-tree vine-copula (Joe, 2015) is applied similar to previous studies (e.g. Mares-Nasarre *et al.*, 2024), as shown in Figure 2.6. The one-tree vine-copula consists of three nodes which represent the three relevant variables per species and two edges (C_{ij}), which are the bivariate copulas. Note that the variables X_1 and X_3 are then conditionally independent given X_2.

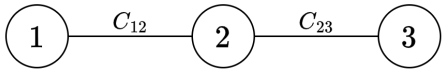


Figure 2.6: One-tree vine with three nodes and two edges.

The model is built using the best fitting bivariate copulas for the pairs T_w and O_{diss}, and T_w and Chl_a for the blue mussels and bivariate copulas for the pairs T_w and DIN, and T_w and DIP for the sugar kelp. This is, the conditioning variable for both species (variable X_2 in Figure 2.6) is set to water temperature T_w, which is the dominant variable. T_w is easy to measure and influences the other selected variables, called conditioned variables.

The joint probability of meeting certain conditions per species and location is calculated here through numerical simulation as follows:

1. The limits defined in Table 2.1 are transformed to unit space via the empirical marginal distributions of the variables (see Table 2.2).

2. For each of the three variables per species, random samples with size $n = 10,000$ in [0,1] are drawn from a uniform distribution, denoted as u_1, u_2 and u_3. u_1 represents random samples of T_w per species, while u_2 represents the conditional probabilities $P(O_{diss}|T_w)$ for blue mussels and $P(DIN|T_w)$ for sugar kelp. Consequently, u_3 represents the conditional probabilities $P(Chl_a|T_w)$ for blue mussels and $P(DIP|T_w)$ for sugar kelp, respectively.

3. A set of n samples per variable pair (u_1 and u_2, and u_1 and u_3, respectively) is generated for each species using Monte-Carlo simulations. This is done via the cdf and the inverse h-function (Aas *et al.*, 2009) for the best fitting copula models. This results in sample clouds per variable pair conditioning on T_w. More information about the inverse h-function can be found in Section A.4.

4. In order to determine the joint probability of these variables satisfying the optimal or critical growth requirements, the samples fulfilling these conditions are counted (see an example in Figure 2.7, where the green samples fulfil the requirements). The ratio of this amount of samples to the total number of samples n provides an estimate of the probability accounting for the dependence between the pairs.

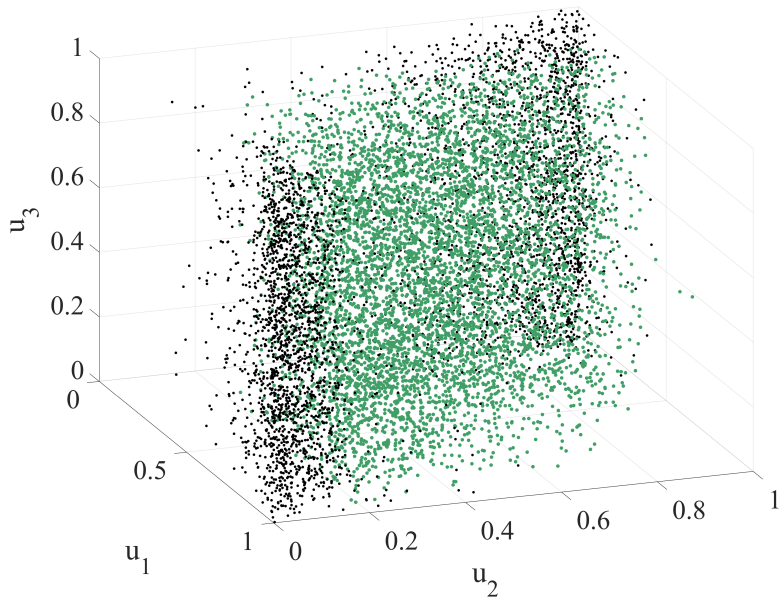


Figure 2.7: Three-dimensional sample cloud for an arbitrary example, with u_1 being the conditioning variable, where the samples in green are meeting the required conditions.

2.3. RESULTS

2.3.1. CORRELATION ANALYSIS OF THE DATA

A correlation analysis is performed via the Spearman rank correlation coefficient as described in Section 2.2.3. The ranges of these rank correlation coefficients for the offshore AOI (see Figure 2.1) are shown in Table 2.3.

Table 2.3: Correlation ranges in the offshore part of the south-eastern North Sea.

Blue mussel	
$T_w - O_{diss}$	-0.95 to -0.9
$T_w - Chl_a$	-0.68 to -0.2
Sugar kelp	
$T_w - DIN$	-0.6 to 0
$T_w - DIP$	-0.7 to 0

As shown in Table 2.3, the correlations are strong for the offshore AOI, mostly linked with p-values below the significance level (here $\alpha = 0.05$) being then significant correlations (see Section 2.2.3). For the variable pairs of the sugar kelp, correlations are occasionally found not to be significant. This indicates that the dependence between the variables might be relevant and, thus, needs to be considered in the suitability analysis. The results in the following sections will mainly focus on the blue mussel due to the stronger dependence between the variables. Partial results for the sugar kelp will be presented; the complete results can be found in the Section A.6.5.

For further results on the spatial variability of the rank correlation coefficients and corresponding p-values, the reader is referred to the Section A.5.1 (Figure A.3 and Figure A.4) for the blue mussels and to the Section A.6.2 (Figure A.10 and Figure A.11) for the sugar kelp.

2.3.2. COPULA SELECTION AND FITTING

As introduced in Section 2.2.4, copulas are used here to model the dependence between the pairs of studied variables. Best copula model is selected by fitting a selection of copula families (Gaussian, Frank, Gumbel, Clayton and t-) and comparing them using two GOF techniques: (1) the CvM statistic, and (2) the semi-correlations, as exposed in Section 2.2.4 [1].

The spatial variability of the best fitting copula model for the variable pairs T_w and O_{diss}, and T_w and Chl_a is displayed in Figure 2.8 for the blue mussel case, where each copula is represented by a different colour. For the sugar kelp, the results for the pairs T_w and DIN, and T_w and DIP are shown in the Section A.6.3 (Figure A.12 (*a*) and Figure A.12 (*b*)).

[1] Because of the extensive volume of results obtained from these two fitting methods, the findings are not presented in this study but can be made available upon request.

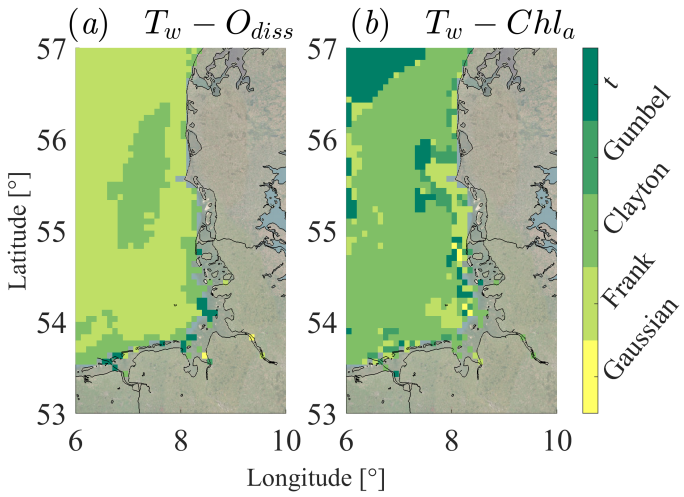


Figure 2.8: Best fitting copulas across the AOI for both variable pairs for the blue mussels: (*a*) T_w and O_{diss} and (*b*) T_w and Chl_a.

In the case of the blue mussels, the Frank copula is the best fit for the variable pair T_w and O_{diss} across most of the investigated area, as displayed in Figure 2.8 (*a*), with a parameter that ranges from -24 to -12 (see Figure 2.9 (*b*)). The Clayton copula is the optimal fit for the central region of the North Sea and its parameter varies between 4 and 11, as shown in Figure 2.9 (*a*).

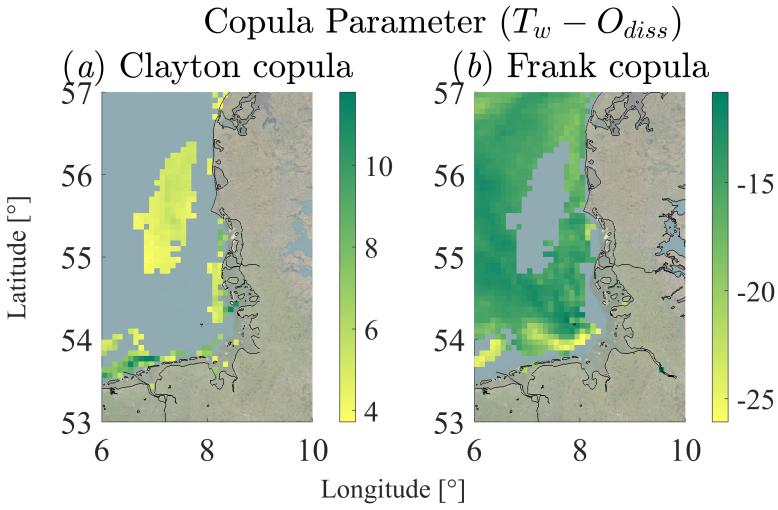


Figure 2.9: The spatial copula parameter for the Clayton copula (*a*) and the Frank copula (*b*) for the variable pair of T_w and O_{diss} for the blue mussel.

The main difference in the results for these two areas is the correlation in the upper

left tail quadrant: if the correlation in this quadrant for the pair of T_w and O_{diss} is below -0.9, the Clayton copula provides the best fit. When correlation goes above -0.9, the symmetrical Frank copula becomes the best fit. Note that T_w and O_{diss} are negatively correlated so it is required to rotate the Frank and Clayton copula. As this variable pair is characterised by a strong correlation in the upper right quadrant, the rotation angle for the Clayton copula is consistently 270° for the entire area. The rotation angle of the Clayton copula for the pair of T_w and O_{diss} can be found in Figure A.5 (*c*) in the Section A.5.2. Different to the Clayton copula, the Frank copula is symmetrical and is thus rotated by 90° for negative correlations.

Regarding the pair T_w and Chl_a, the Clayton copula provides the best fit (see Figure 2.8 (*b*)) with the parameter increasing gradually from approximately 0 in the south-east to 1.3 in the north-west of the area, similar to the correlation pattern (see Figure A.3 (*b*) in the Section A.5.1). Again, the variables are negatively correlated, so rotation of the copulas is required. Different to the pair of T_w and O_{diss}, the degree of rotation angle for the Clayton copula is less uniform, although 270° is still the most prominent (see Figure A.5 (*d*) in the Section A.5.2).

The t-copula is the best fitting copula model for the pair T_w and Chl_a in the north-western part, where the correlation is the strongest. The bivariate t-copula presents two parameters: ρ (correlation coefficient) and ν (degrees of freedom); ρ varies between -0.6 and -0.7, while the parameter ν ranges from 2 to 10 (see Figure 2.10 (*b*) and (*c*)). However, note that for a $\nu > 10$, the t-copula approximates the Gaussian copula. Finally, near the coastline, where the correlation is not found significant, the fitting of copula models is difficult and spread between the different considered models.

Regarding the sugar kelp, the fitting of copula models for the variable pair T_w with DIN does not present one dominant copula model across the entire studied area (see Figure A.12 (*a*) in the Section A.6.3) due to the low observed correlations. In regions with a stronger correlation near the coastline and further offshore, the Frank or Gaussian copula models appear to be the most suitable fit. Thus, no significant tail dependence is observed between the variables. The variability of the parameters for these two fitted copulas is shown in Figure A.13 in the Section A.6.4. For a great part of the studied area offshore, the t-copula is predominant. However, the copula parameter ν reaches values up to 10^7, approximating a Gaussian copula.

Different to the previous variable pair, the fitting of copula models is less varied across the area for the variable pair T_w with DIP (as depicted in Figure A.12 (*b*) in the Section A.6.3). In regions with correlations between -0.4 to 0, the t-copula is the best fit with very high ν values (see Figure A.14 in the Section A.6.4). In regions where the correlation is stronger (near the coast and further offshore), the Gumbel and Frank copulas are identified as the best fit. For the Gumbel copula, the parameter varies between 1 and 2.2, while for the Frank copula ranges between -8 and -2. Note that the samples for the Gumbel copula are rotated by 270 $^\circ C$ (see Figure A.15 (*b*) in the Section A.6.4) and for the symmetrical Frank copula by 90 °, as the correlation of this variable pair of T_w and DIP is negative.

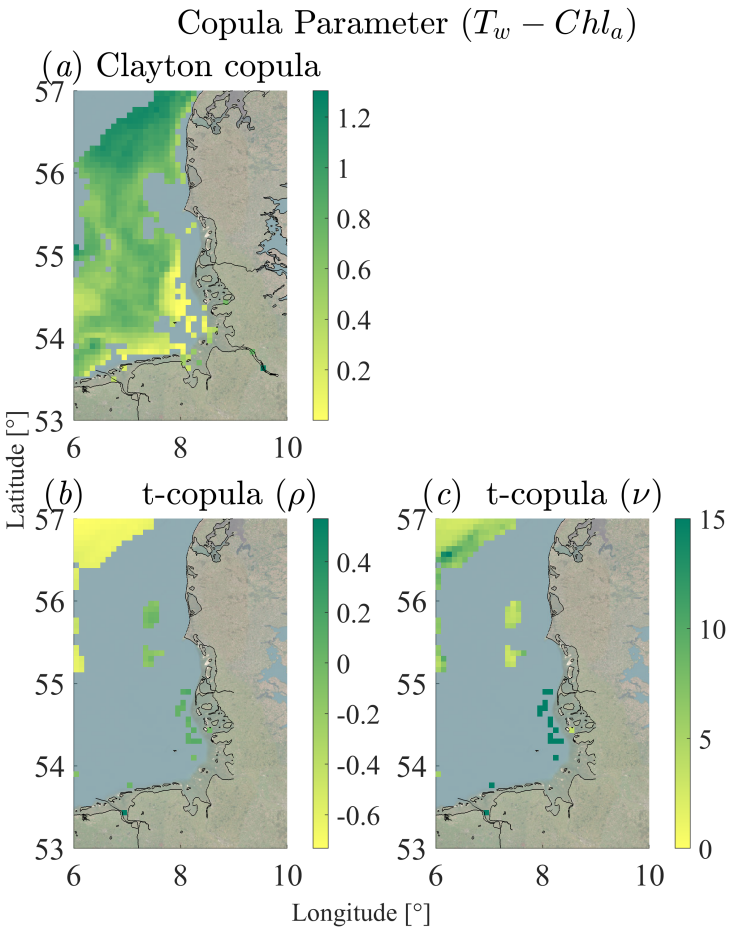


Figure 2.10: The spatial copula parameter for (*a*) the Clayton copula, (*b*) *rho* and (*c*) ν of the t-copula for the variable pair of T_w and Chl_a for the blue mussel.

In summary, the fitting of copula families in the blue mussel case within the offshore AOI is rather uniform. In the nearshore area, however, where the nearshore processes take a relevant role, the correlation between variables is lower leading to less clear patterns. Similarly, for the sugar kelp, where correlations are in general lower, the fitting and selection of the best copula model becomes more challenging.

2.3.3. PROBABILITY CALCULATION AND SUITABILITY MAPS

After selecting the copula models to fit the dependence between the variables, the probabilities of meeting the optimal and critical thresholds for the three variables during the main growth season of each species is calculated as described in Section 2.2.5. These probabilities are also calculated under the independence assumption to investigate whether incorporating the dependence between variables improves the estimations.

First, the probabilities of meeting the critical limits (see Table 2.1) are investigated. For the blue mussels, these probabilities are for most of the area close to 1, regardless of whether dependence is considered or not (see Figure A.7 (*a*) and (*b*) in the Section A.5.3). For the sugar kelp, it can be seen that the probability of meeting the critical limits is high close to the coast but decreases with increasing distance to the shore, as shown in Figure A.17 in the Section A.6.5. In Figure A.17 (*c*), the differences between the probabilities computed accounting for the dependence and disregarding it are displayed, reaching a maximum of about 0.09. Close to the coast, the differences are around 0, while offshore, the differences in probabilities are negative for the area with tail dependence (Gumbel model) between DIP and T_w and positive where symmetrical copula models (i.e., Frank and t-copulas) are the best fit. Note that these differences become more remarkable for the areas with low probabilities of meeting the critical thresholds (northwestern part of the AOI). When computing the difference relative to the actual probability considering dependence, the difference can be up to 40%.

With regard to the optimal growth conditions for the blue mussels, the highest probabilities of meeting them are observed near the coast, exceeding 0.4, as displayed in Figure 2.11 (*a*). As the distance from the shore increases, the probability gradually decreases, reaching approximately 0.2 for most of the offshore area. In an elongated area running from south to north along the German and Danish coast, a similar probability to that near the shore is observed. This pattern is caused by the differences in the univariate probabilities; the probability of meeting the optimal conditions is relatively consistent and high for T_w and O_{diss}, while the probability of Chl_a meeting the optimal conditions is more variable. It is highest close to the coast and within the aforementioned elongated shape, ranging from 0.6 to 0.8. In the rest of the area, the probability is relatively low, ranging from 0.2 to 0.4. This leads to Chl_a being the most limiting factor, as already pointed out by Geisler *et al.* (2018).

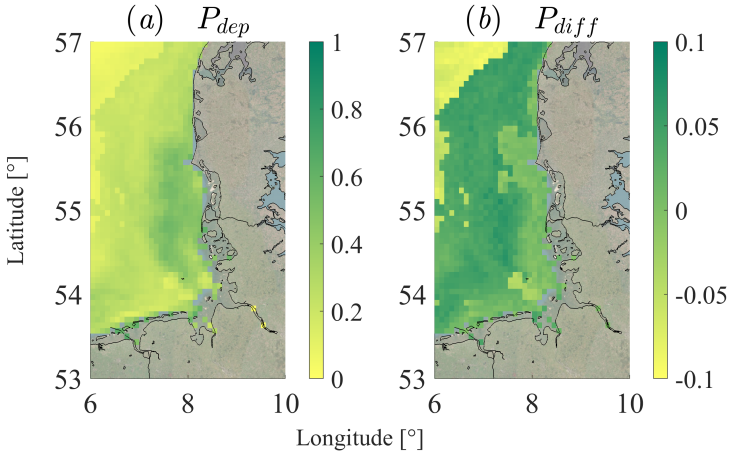


Figure 2.11: Probabilities of meeting the optimal conditions for the suitability of blue mussel *Mytilus edulis*: (*a*) probabilities computed using the dependence models, and (*b*) (absolute) differences between the probabilities computed accounting for the dependence and assuming independence between the variables.

Regarding the difference between considering or disregarding the dependence between the variables, a maximum absolute difference of about 0.1 is found, as shown in see Figure 2.11 (*b*). The darker area in Figure 2.11 (*b*) indicates a positive difference between the dependent and independent approach, thus ignoring the dependence between variables would underestimate the joint probability. In contrast to that, the lighter area northwest indicates a negative difference, such that assuming independence might overestimate the joint probability for these areas.

For the sugar kelp, the probabilities of meeting the optimal growth conditions across the entire area are very low regardless considering or not the dependence between the studied variables (see Figure A.16 (*a*) and (*b*) in the Section A.6.5). While the probability of meeting the optimal conditions for T_w all along the domain and DIN close to the coast is relative high, DIP marginally fails to meet the specified conditions. The probability of DIP meeting the optimal concentrations is close to 0 for the entire area, except in estuaries and rivers, where the probability reaches a maximum of more than 0.6. This could be primarily explained by the nutrient influxes from coastal regions for both, DIN and DIP. The results of the present study align with the findings from previous studies. For instance, Gimpel *et al.* (2015) found that the offshore cultivation of blue mussels is in general feasible for the AOI, while for sugar kelp, regions with higher water temperatures and nutrient rich waters due to river inflows are mostly suitable.

In summary, the critical limits for the blue mussel species are generally met differently to those for the sugar kelp, especially further offshore. For the optimal growth conditions, the overall suitability across the entire area is higher for the blue mussels compared to the sugar kelp, as DIP marginally fails to meet the optimal limits. Besides, the strengths of the correlation between variables has a direct impact on whether the dependence makes a difference in the computed calculations and, thus, needs to be accounted for. If the correlation is not significant, the simplified independent probability calculation may be sufficient. However, accounting for the dependence when asymmetries such as tail dependence are present might be needed when dealing with extremes, namely minimum and maximum observations of the variables. Also, in cases where low probabilities are computed (e.g. the optimal requirements for the blue mussel species in Figure 2.11 or critical conditions for the sugar kelp in Figure A.17 in the Section A.6.5), the differences between accounting for the dependence or disregarding it can be up to values between 30 and 40%. This highlights the importance of considering the dependence between variables in suitability studied.

2.4. DISCUSSION

In this study, a probabilistic framework is proposed to assess the suitability of offshore aquaculture accounting for the dependence between relevant ecological variables. This methodology is applied as a case study to two species, the blue mussel *Mytilus edulis* and the sugar kelp *Saccharina latissima*. The growth offshore in the North Sea is investigated by computing the probability of occurrence of the optimal and critical conditions of these species. Such conditions are parameterised using three variables per species during their growth season: temperature, dissolved oxygen and chlorophyll-a concentrations

for the blue mussels, and temperature, dissolved inorganic nitrogen and phosphorus concentrations for the sugar kelp. Thresholds describing critical conditions for the survival of the species and optimal conditions for their growth are defined based on the existing literature. Empirical distribution functions are applied here to model the univariate uncertainty of each variable (see Section 2.2.4), while bivariate copula models are recommended in Section 2.2.5 to account for the probabilistic dependence between variables. Empirical distribution functions were used here since the optimal and critical limits are generally for most of the variables within the observed ranges and there is no need to extrapolate out of them. However, if those thresholds would be changed towards more extreme probabilities, parametric univariate margins would be needed.

The probabilities to meet the critical survival and optimal growth limits for each species were computed both accounting for the dependence using bivariate copulas and disregarding it. The results show that this may lead to significant differences in the computed probabilities and, thus, in the conclusions of the suitability study. Here, five parametric families of bivariate copulas were considered, namely Gaussian, Frank, Gumbel, Clayton, and t-copula. Such parametric copulas were selected as they are commonly used and they cover different shapes of symmetric dependence, as well as the asymmetry of tail dependence. However, no other asymmetries were considered within the considered models. This is, for a positive dependence, asymmetries between the upper right and lower left quadrant are considered, but not between the upper left and lower right quadrant. Thus, further extensions of this work may focus on analyzing the role of such asymmetries in the computation of the probabilities (Jaeger and Morales-Nápoles, 2017).

Regarding the results, high probabilities of meeting the optimal or critical requirements for the growth of the blue mussels are obtained. On the contrary, the probabilities of meeting optimal conditions for the sugar kelp are low due to the variables marginally failing to meet the specified optimal growth conditions. This is, the computed probabilities are sensitive to the defined thresholds. However, slightly different limits are provided for some of the studied variables depending on the source and there is no general consensus about the relevance of some ecological variables for the species to grow (Bergström and Lindegarth, 2016; Kamermans *et al.*, 2022). Moreover, there is limited literature to define limits for DIN and DIP for sugar kelp. In this study, the conditions for optimal and critical DIN concentrations are taken from Jevne *et al.* (2020), where the growth of small sporophytes under controlled conditions in tanks for 20 days of experimental time was investigated. These experimental results were also used to determine the conditions for DIP via the molar Redfield ratio of 16:1 ($molN/molP$) (Redfield, 1934). Therefore, field observations would be desirable to validate such limits, as well as further research on defining the critical and optimal threshold for growth. However, the goal of the applications here is to showcase the proposed probabilistic methodology and, thus, the computed probabilities can be recalculated in light of new literature.

Possible extensions of this study could include hydrodynamic factors like waves and currents that can be crucial for species (Kerrison *et al.*, 2015) and the safety of the installed structures for their growth. Also, interactions between species, such as integrated multi-

trophic aquaculture (IMTA) or influences of large-scale cultivation systems on e.g. the water quality can be included in extensions of this study. If the incorporation of further variables leads to higher dimensional models, copula-based Bayesian Networks (e.g. Hanea *et al.*, 2006) or Vine Copula models (Joe, 2015) might be applied. Moreover, long-term effects such as the increase in water temperature or potential changes on the nutrients for the species due to climate change can be investigated and, if found significant, modelled using non-stationary margins (Ragno *et al.*, 2019). Finally, future studies could aim to account for the relationship between consecutive extremes as, for example, the sugar kelp *Saccharina latissima* can tolerate short-term exposure to high DIP concentrations without negative effects for a few days (Lubsch and Timmermans, 2019). This is not investigated here, as this study analyses daily extremes independently, rather than in sequence.

The proposed method is highly transferable to other species, such as fish or oysters, to define cultivation suitability through probabilities by accounting for the probabilistic dependence structure among variables. Beyond aquaculture, renaturation or species reintroduction, not just in the marine context, are a possible field of application.

2.5. CONCLUSION

The offshore aquaculture sector is still in early stages. Some methodologies exist for spatial planning, but they do not account for the dependence between the relevant growth variables of species. In this study, a method accounting for this dependence between variables is proposed and its results are compared to the results when disregarding it (independent case). To model the joint distributions of the studied variables, bivariate copulas are applied. Using the developed models, the spatial suitability for the cultivation of two species, specifically blue mussels *Mytilus edulis* and sugar kelp *Saccharina latissima*, is assessed in the south-eastern North Sea. This is done by determining the probabilities of meeting certain conditions for three ecological variables per species. The results show that there is a significant difference in the probabilities when incorporating the dependence compared to assuming probabilistic independence. This is especially remarkable for areas with low probabilities (less suitable areas for cultivation) as not accounting for dependence can underestimate the suitability by up to 40%. However, if the correlation between the variables is low, the independent approach may be sufficient, such as in most of the studied area for the sugar kelp. Besides, the results of this study indicate that both species' cultivation is generally suitable in the defined area with regard to their survival conditions. For the optimal growth conditions, an offshore north-south elongated area along the German and Danish coast seems to be most suitable for blue mussel cultivation next to near-coast areas. Different to that, optimal conditions for the sugar kelp are only met in rivers and their estuaries.

3

GAUSSIAN COPULA-BASED BAYESIAN NETWORKS FOR DYNAMIC LOADS IN MOORING SYSTEMS

Building on the importance of accounting for variable dependence shown in the previous chapter, this chapter extends the modelling approach from bivariate to multivariate dependence. It introduces Gaussian copula-based Bayesian Networks (GCBNs) to model the joint behaviour of up to 13 hydrodynamic and structural variables related to mooring tensions in floating offshore structures. By enabling conditional analysis of mooring tensions under specific hydrodynamic scenarios, the GCBN approach improves understanding critical load conditions and supports more informed probabilistic safety assessments of floating offshore structures.

3.1. INTRODUCTION

The offshore marine environment has enormous economic potential. For example as a source of renewable energy and proteins obtained from aquaculture. Thus, making efficient use of this space is increasingly important. Not only is there an increasing demand for proteins (Yong *et al.*, 2022), but offshore wind energy is also expected to continue to grow in the coming years (Soares-Ramos *et al.*, 2020). This in turn will help replacing energy production sectors based on fossil fuels and reducing carbon emissions, thus potentially contributing to limit global warming. Osman *et al.* (2023) reviewed several aspects of renewable energy under climate change. Prices for energy from offshore wind, biomass production, and other green energy sources are expected to drop drastically in the future while having a comparably minimal impact on the environment. However, the space at sea, especially near-shore, is limited and therefore there are conflicts between sectors and users (e.g., Buck and Langan, 2017; Buck *et*

This chapter has been published in **Applied Ocean Research** as *Gaussian copula-based Bayesian networks for dynamic loads in mooring systems* in 2025, Santjer *et al.* (2025). Minor editorial changes in this dissertation are implied.

al., 2018; Pınarbaşı *et al.*, 2017). These conflicts include spatial, environmental, and societal challenges, between traditional users such as fisheries, tourism, and ports. Unlike nearshore, offshore areas have different stakeholders and a lot of potential for multi-use of several sectors, such as offshore wind, solar and aquaculture (Buck and Langan, 2017; Maar *et al.*, 2023).

However, offshore structures face harsh environmental conditions that risk the safety of the structures. Thus, mooring systems are essential to ensure the stability of various floating structures, including floating wind turbines, floating solar panels, ships, installation vessels, and aquaculture structures. Especially during extreme events, tensions in single mooring lines are critical to structural safety. Even more critical are snap forces, which can cause early failure of mooring lines (Governo *et al.*, 2023). snap forces are sudden increases in tensions, which occur after the mooring line becomes temporarily slack. These dynamic snap tensions can exceed static tensions by a multiple, even up to 6-times higher than the static pretension (Guo *et al.*, 2017; Hsu *et al.*, 2017; Miškov *et al.*, 2023).

Research has been done to understand the behaviour and expected tensions in mooring lines under snap loads, using field measurements, numerical simulations and model experiments. Research on field experiments to investigate tensions in mooring line, especially with regard to snap loads, is however limited. Nguyen *et al.* (2019) measured the tensions in the moorings of an oyster cage along the coast of Maine. The results show that the mooring line tensions correlate with the tide-induced water level variations. Currents have a significant influence at low to mid tidal levels, but not at high tides. Similar results are obtained in Nasyrlayev *et al.* (2023). A Finite Element Model is used to improve the design of anchors and other structural elements, especially for high-energy regions. The results show that the current magnitude and its angle of incidence do have a much higher impact on the tensions than those of the waves. Lumped-mass models are extensively used in studies related to aquaculture systems (Feng *et al.*, 2021; Hermawan and Furukawa, 2020), highlighting the complexity of interaction between the environment and the structures. Another Finite Element Model is used by Desiré *et al.* (2023) to investigate the seabed interaction with mooring lines in complex bathymetries. The model is validated against experimental scale tests for both cases with and without snap loads. Following the analyses of Desiré *et al.* (2023), experimental tests (Somoano *et al.*, 2024) confirm that bathymetry is another factor influencing the mooring line dynamics. While the quasi-static loads are not impacted by the shape of the bathymetry, the dynamic loads, such as snap loads, are significantly influenced by the slope of the bathymetry.

Experimental investigations at a 1:50 scale, corresponding to a full-scale water depth of 28 m, were conducted to investigate the relationship between snap loads and the damping of mooring lines (Hsu *et al.*, 2019). The results show that snap loads occur more frequently in shallow water settings due to shorter mooring lines and reduced damping. Also, a positive correlation was found between the pretension of a mooring line due to the self-weight with the mooring line damping. The intensity and occurrence rate of snap loads were also investigated in Hsu *et al.* (2017) through experiments at a 1:50 scale, simulating a floating offshore wind turbine in conditions for very high return periods. The results show that dynamic tensions caused by snap events are 37-68%

higher than non-snapping tension peaks. In addition, these snap loads follow a different exceedance probability distribution than non-snapping peak tensions. A composite Weibull distribution, which has different parameters for lower and higher tension rates, is found to be a good fit to account for snap loads, while the conventional Weibull distribution tends to underestimate how often these extremes occur (Hsu *et al.*, 2017).

Various probabilistic models are applied to assess the reliability and uncertainty of mooring line tensions under stochastic environmental conditions. Regarding peak tensions in mooring lines, Zhao and Dong (2023) identified the Gumbel distribution as the best model to describe them. The study investigates uncertainty of extreme mooring loads of floating systems using environmental contours, which are based on the joint distribution of variables. Li *et al.* (2016) assesses mooring line fatigue damage using probability density functions (pdf's) for short-term mooring line tension amplitudes, which result from single sea states or short durations, and for long-term amplitudes, which are obtained by integrating the short-term distributions while considering the probability of various sea states. Several different pdf's have been investigated, while the Gamma distribution is a good fit for both short and long-term tension amplitudes, for the latter, the Gamma distribution is extended by a location parameter. Thus, the mooring line tension response is non-Gaussian.

To assess the tensions of mooring lines while considering the relation between variables, Wu *et al.* (2022) employs conditional probabilities of tensions and Monte-Carlo simulations. The results show that, in case of failure, the tension responses in the remaining mooring lines can be described by the Generalised Extreme Value (GEV) distribution. Conversely, Montes-Iturrizaga and Heredia-Zavoni (2016) evaluates the reliability of mooring lines by investigating the impact of the statistical relationship between the significant wave height and peak spectral period. Therefore, copula functions are used to describe their dependence. The results show that accounting for the dependence structure among environmental variables has a significant impact on the reliability of mooring lines. Modelling the multivariate probability distribution via for example multivariate copula models is determined by Montes-Iturrizaga and Heredia-Zavoni (2016) as an important further step to investigate how environmental variables affect the reliability of floating structures.

Considering the multivariate dependence of variables is crucial for improving the reliability and safety of mooring line systems due to complex environmental conditions and the interactions with the structures. To the authors' knowledge, taking the multivariate dependence of variables into account to assess mooring line tensions due to environmental impacts has not been done before. In this study, the multivariate dependence between hydrodynamic variables and tensions in mooring lines is modelled using a Gaussian copula-based Bayesian Network (GCBN). The nodes of the Bayesian Network represent continuous random variables and the arcs between the nodes describe the (conditional) probabilistic dependence using bivariate Gaussian copulas. Due to the application of copula models, the GCBN has no restrictions on marginal distributions. Thus, the GCBN is a fully probabilistic model accounting for uncertainty. The advantages of a GCBN are many-fold. GCBNs are computationally efficient (see e.g. study Barros *et al.* (2024) using a GCBN with 879 variables). Their graphical representation of the complex multivariate dependence of a set of variables allows to visually give information about

(conditional) independence of variables. Additionally, the model allows for incorporation of prior or new knowledge and thus can be updated given new observations. That is, a GCBN can efficiently and accurately enable inference i.e. providing distributions of variables given observational data as input. This is especially of importance for scenario analysis (e.g. conditionalising on certain variables and analysing the remaining conditional distributions), offering a more comprehensive view of the problem, enabling risk assessment and informed decision-making. The main disadvantage, however, is that the joint dependencies between variables are restricted to the Gaussian copula. This entails that GCBNs cannot reproduce certain asymmetries in the joint distribution such as tail dependence for example.

Regarding the tensions in the mooring lines, this study focuses on extremes of tension rates for two reasons: 1) snap loads, characterised by high tension rates, inducing a strong dynamic response of the floater; 2) operational limitation related to the application of the floater. For example, operation of offshore wind turbines require a constrain on the acceleration of the tower to minimise damage to the blades (Dueñas-Osorio and Basu, 2008). This is similar to another example on aquaculture, where extreme accelerations lead to damages such as mussel fall-off (e.g., Gagnon and Bergeron, 2017; Knysh *et al.*, 2020).

This study takes the mooring line dimensions exemplarily from a seaweed aquaculture structure, specifically a net structure. This reference structure is a pilot designed for offshore conditions in the North Sea, 80 km off the coast, for a water depth of 23 m (Strothotte *et al.*, 2021). Note however, that the methodology as well as the results presented in this study can be transferred to any floating structure, which requires a mooring line.

Environmental data used in this study come from a three-dimensional large-scale hydrodynamic model covering the European Shelf, while a Finite Element (FE) model is used to get observations about tensions within the mooring lines. The mooring line of the aquaculture structure is implemented in the model and adjusted for different water depths. Based on these observations, GCBNs are set up for a few different locations in the south-eastern North Sea. The aim of this study is twofold. The multivariate dependence is analysed and compared among different locations with different characteristics. This helps to assess which hydrodynamic variables are the main driver for snap loads under different conditions. Secondly, this study shows how the GCBN can be used to help reliability engineering by assessing extreme tension rates conditioned on certain hydrodynamic conditions. Not only are offshore hydrodynamic impacts highly relevant for such structures, but also ecological variables are crucial, particularly in the case of aquaculture. The ecological suitability for the same structure and area was investigated using bivariate copula models in Santjer *et al.* (2024a).

The present study is structured as follows. First, the hydrodynamic model, including the study locations and selected hydrodynamic variables are described in Section 3.2.1 and Section 3.2.2, while the Finite Element model for mooring lines is described in Section 3.2.3. Section 3.3.3 describes the probabilistic modelling approach, including copula models and copula-based Bayesian networks, in Section 3.3.2 and Section 3.3.3, respectively, as well as the processing of the numerical tension data (Section 3.3.4). The results are described in Section 3.4.5. First, Bayesian Networks are set up and described

in Section 3.4.1 for the hydrodynamic variables, which output serves as input for the FE mooring line model. The post-processing of the tension data is for one location in the North Sea exemplarily presented in Section 3.4.2, while the GCBNs consisting of environmental variables and tension rates are described in Section 3.4.3. Final results analysing extreme tension rates in mooring lines, also under given environmental conditions, are described in Section 3.4.4 and Section 3.4.5, respectively. Discussion of the results can be found in Section 3.5 and conclusion of the presented study are drawn in Section 3.6.

3.2. MODELLING FRAMEWORK AND STUDY SETUP

3.2.1. 3D HYDRODYNAMIC MODEL

Wave and current characteristics at the locations of interest are necessary to determine the tensions in the mooring line of a floating structure for each of these locations. In this study, data is extracted from a large-scale three-dimensional hydrodynamic model called Dutch Continental Shelf Model (DCSM) (Zijl *et al.*, 2023). The DCSM makes use of the D-Flow FM component of the Delft 3D Flexible Mesh Suite (Deltares, 2023a) in which the grid is based on a flexible mesh, as the grid resolution is spatially variable. In deep oceanic waters, the grid is coarse with 4×4 nm, while it increases towards shallow areas and to coast up to 0.5×0.5 nm. Cells that change resolution are connected by triangles. In total, the DCSM consists of 630,000 cells. The domain of the model is between 17°W to 13°E and 43°N to 64°N and thus covers the northwestern European continental Shelf. To account for variations in water depth across the model domain, z-sigma layers are employed for vertical resolution, combining two different types of layers. In the upper 100 m, sigma-layers are utilised, consisting of a fixed number of 20 layers, with varying thickness depending on the water depth. This allows for a high resolution especially in shallow areas. Beneath the sigma-layers, z-layers are employed, which have a fixed depth throughout the domain, while their thickness increases exponentially by a factor of 1.19 towards the deeper layers. Variables describing the current flows are derived from this model with a 10 minutes resolution, while variables describing the waves are derived from the third general spectral wave model SWAN (simulating waves nearshore (Booij *et al.*, 1999)). SWAN is added on top of the D-Flow FM component and solves shallow-water processes. Data from the models are extracted for 5 years (2013 until 2017).

3.2.2. STUDY LOCATIONS AND VARIABLE SELECTION

The current study focusses on the south-eastern part of the North Sea, which is well represented by the model domain, which covers the entire northwestern European shelf. Here, 12 spatially distributed locations l are chosen (see Figure 3.1, in which the reference location (RefLoc 4) is highlighted). Some of these locations are located close by or within offshore wind farms, others are research platforms or measurement buoys. The close distance to offshore wind farms is relevant, especially in the context of marine multi-use or co-location of several sectors (Buck, 2002; Buck and Langan, 2017; Maar *et al.*, 2023). Furthermore, the selected sites span a variety of depths in this area from approximately 13 m to 45 m (see Table 3.1) and experience different environmental conditions due to their varying distances from the shore.

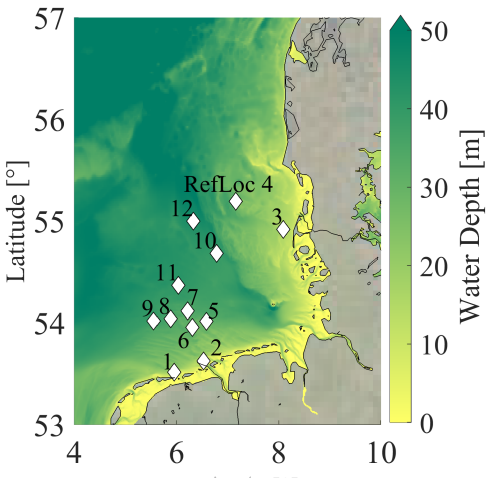


Figure 3.1: Map of the south-eastern North Sea, showing 12 study locations (l), with the reference location RefLoc 4 highlighted.

Table 3.1: Overview of the total water depth (h) in [m] per location l together with depths i of current magnitude [m/s] data throughout the water column, where $d1$ is 1 m above the seabed, while $d5$ is close to the surface.

Location l	Total depth (h)	$d1$	$d2$	$d3$	$d4$	$d5$
1	12.6	1	4	6	9	12
2	14.5	1	4	7	10	14
3	15.0	1	4	7	11	14
4 (RefLoc)	22.8	1	6	11	17	22
5	30.1	1	8	15	22	29
6	30.2	1	8	15	22	29
7	33.6	1	9	17	25	33
8	35.7	1	9	18	26	35
9	38.6	1	10	19	28	38
10	39.2	1	10	20	29	38
11	40.7	1	11	20	30	40
12	44.7	1	12	22	33	44

The main environmental impacts relevant for the estimation of the tension in the mooring lines of a (submerged) floating structure are waves and currents, whose characteristics for selected study locations are provided by the DCSM model (see Section 3.2.1). Specifically, significant wave height, H_S, and peak wave period, T_P, are selected to describe wave characteristics (see Figure B.1 in Section B.1 for an overview of the magnitude of both variables for each location), while for currents, horizontal flows $c_{d,i}$ are considered at different depths i, which are selected according to the water depth per location. The top and bottom 1 m of the water column are excluded from consideration, while the

remaining portion of the water column is equally divided into 5 depths. The water depth per location and depth selected for the extraction of the current velocity magnitude is shown in Table 3.1.

3.2.3. MOORING LINE MODEL

In this study, the mooring line model reported in Agarwal *et al.*, 2025 is employed. This model is specifically developed for flexible cables and incorporates a geometrically non-linear analysis, allowing for large displacements and finite strains. Consequently, it extends beyond the conventional small-strain assumptions. Figure 3.2 presents a schematic of the mooring line, along with the external forces considered in this analysis.

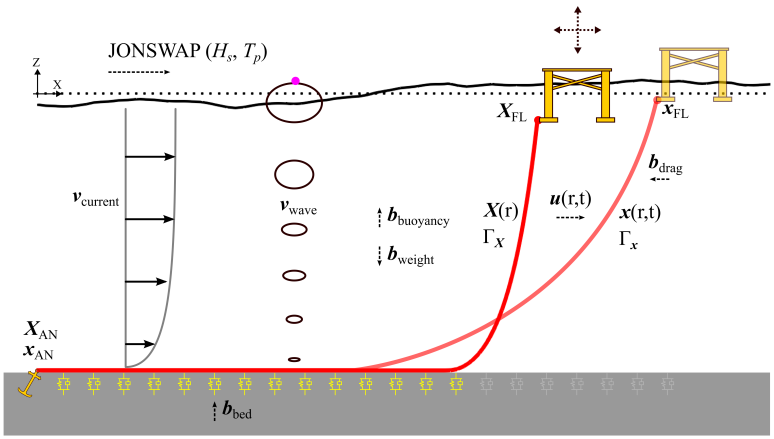


Figure 3.2: Schematic of the mooring line, along with the external forces acting on the line.

The mooring analysis is described in terms of the internal and external forces acting on the line. It is considered to be a curved manifold in 2D XZ space, with $z = 0$ at the mean-sea level. The initial undeformed shape $\Gamma_{\boldsymbol{X}}$ of this manifold is given by a function field $\boldsymbol{X}(r)$. Here r represents a 1D reference domain Ω along the line, and $\boldsymbol{X}(r)$ is its map to the 2D space. For instance, $r = 0$ and $r = L_0$ denote the two end points of the curved manifold $\Gamma_{\boldsymbol{X}}$. The anchor is located at $\boldsymbol{X}(r = 0) = \boldsymbol{X}_{\text{AN}}$, and the fairlead is at $\boldsymbol{X}(r = L_0) = \boldsymbol{X}_{\text{FL}}$. The dynamic deformed shape $\Gamma_{\boldsymbol{x}}$ of the manifold is given by function field $\boldsymbol{x}(r, t)$. The difference between undeformed and the dynamic deformed shape is given by the displacement field $\boldsymbol{x}(r, t) = \boldsymbol{X}(r) + \boldsymbol{u}(r, t)$. The equilibrium of the mooring line is always defined on the dynamically deformed configuration, $\Gamma_{\boldsymbol{x}}$. The equilibrium equation per-unit length is expressed as

$$\rho A \ddot{\boldsymbol{u}} - \text{div}_{\Gamma} \boldsymbol{\sigma}(\boldsymbol{u}) A - \boldsymbol{b}(\boldsymbol{x}) = 0 \quad \forall \boldsymbol{x} \in \Gamma_{\boldsymbol{x}} \tag{3.1}$$

Here A is the cross-sectional area, ρ is the material density and $\rho \ddot{\boldsymbol{u}}$ represents the inertia of the line. The term $\boldsymbol{\sigma}(\boldsymbol{u})$ is the Cauchy stress tensor and $\text{div}_{\Gamma} \boldsymbol{\sigma}(\boldsymbol{u})$ is its *surface divergence*. This term characterises the variation in stress within the line along its curved

manifold, effectively capturing the internal material resistance to deformation. Finally, the term $\boldsymbol{b}(\boldsymbol{x})$ represents the external forces acting on the line.

The Cauchy stress tensor is defined based on the Saint Venant-Kirchhoff solids theory. For cables made of material with Young's modulus E and Poisson's ratio ν, the corresponding Lamé constants are $\lambda = 0$ with $\mu = \frac{E}{2}$. Consequently, the first and second Piola-Kirchhoff stress tensors $\mathbf{K}$, $\mathbf{S}$ on $\Gamma_{\boldsymbol{X}}$ and the Cauchy stress tensor $\boldsymbol{\sigma}$ on $\Gamma_{\boldsymbol{x}}$ are expressed as follows,

$$\boldsymbol{\sigma} = \frac{1}{\Lambda} \cdot \mathbf{F}_\Gamma \cdot \left(2\mu \cdot \mathbf{E}_{\text{tang}}\right) \cdot \mathbf{F}_\Gamma, \tag{3.2}$$

where $\mathbf{F}_\Gamma$ is the *surface deformation gradient*, and $\mathbf{E}_{\text{tang}}$ tangential Cauchy-Green strain tensor. For further details regarding the structural model, see Agarwal *et al.*, 2025.

The external force acting on the line include self weight, buoyancy of the surrounding fluid, drag due to relative motion against the fluid and the effect due to the sea-bed. Since the external forces are a function of the position on the line, the external forces are defined per-unit length. Self weight of the line is given by

$$\boldsymbol{b}_{\text{weight}}(\boldsymbol{x}) = -\rho\, A\, g\, \hat{k}, \tag{3.3}$$

where g is the acceleration due to gravity and $\hat{k}$ is a unit vector along the vertical axis. The force of buoyancy is given by the weight of water displaced by the line,

$$\boldsymbol{b}_{\text{buoyancy}}(\boldsymbol{x}) = \rho_{\text{water}}\, A_{\text{disp}}\, g\, \hat{k}, \tag{3.4}$$

where ρ_{water} is the density of water, A_{disp} is the displaced volume per unit-length of the line.

The effect of the sea-bed is implemented using a damped spring, which is activated only for the portion of the line in contact with the sea-bed, i.e.

$$\boldsymbol{b}_{\text{bed}}(\boldsymbol{x}) = \begin{cases} -k_b\,(z - z_{\text{bed}})\,\hat{k} - c_b \dot{z}\,\hat{k} & \text{if } z \le z_{\text{bed}} \\ 0 & \text{if } z > z_{\text{bed}} \end{cases} \tag{3.5}$$

The drag forcing on the line is due to three components, 1) self drag, 2) action of the current, and 3) wave action. If $\boldsymbol{v}_{\text{wave}}(\boldsymbol{x})$ is the local velocity of the wave and $\boldsymbol{v}_{\text{current}}(\boldsymbol{x})$ the local current velocity, then the drag force on the line is given by

$$\boldsymbol{b}_{\text{drag},n}(\boldsymbol{x}) = \frac{1}{2}\rho_{\text{water}}\, C_{d,n}\, \boldsymbol{v}_{r,n}(\boldsymbol{x})|\boldsymbol{v}_{r,n}(\boldsymbol{x})| \tag{3.6}$$

$$\boldsymbol{b}_{\text{drag},t}(\boldsymbol{x}) = \frac{1}{2}\rho_{\text{water}}\, C_{d,t}\, \boldsymbol{v}_{r,t}(\boldsymbol{x})|\boldsymbol{v}_{r,t}(\boldsymbol{x})| \tag{3.7}$$

$$\text{where } \boldsymbol{v}_r(\boldsymbol{x}) = \boldsymbol{v}_{\text{wave}}(\boldsymbol{x}) + \boldsymbol{v}_{\text{current}}(\boldsymbol{x}) - \dot{\boldsymbol{u}}(\boldsymbol{x}) \tag{3.8}$$

$$\boldsymbol{v}_{r,t}(\boldsymbol{x}) = \left(\boldsymbol{v}_r(\boldsymbol{x}) \cdot \hat{t}(\boldsymbol{x})\right) \hat{t}(\boldsymbol{x}) \tag{3.9}$$

$$\boldsymbol{v}_{r,n}(\boldsymbol{x}) = \boldsymbol{v}_r(\boldsymbol{x}) - \boldsymbol{v}_{r,t}(\boldsymbol{x}) \tag{3.10}$$

In the present analysis, it is not accounted for the influence of added mass and added damping of the mooring line. This is based on the findings in Hall and Goupee, 2015,

which demonstrate the limited impact of added mass on the overall system response.

The definition of the boundary value problem is completed by specifying the boundary conditions at the two ends of the mooring line. A Dirichlet-type boundary condition is applied, with the anchor fixed at $\boldsymbol{u}(r=0,t)=\boldsymbol{u}_{\mathrm{AN}}=[0.0,0.0]^T$, and a prescribed motion for the fairlead $\boldsymbol{u}(r=L_0,t)=\boldsymbol{u}_{\mathrm{FL}}(t)$.

The governing equations are numerically solved using the finite-element method. The definition of surface gradients along curved manifolds is done using the *Tangential Differential Calculus* approach presented in Fries and Schöllhammer, 2020. The formulation of the variational problem is done using the standard Galerkin approach. The reference domain Ω is discretised using linear Lagrangian elements, while the time-stepping is implemented using the implicit Generalised-α method (Chung and Hulbert, 1993). A detailed description of the mooring line model is given in Agarwal *et al.*, 2025. It also presents rigorous investigation of numerical convergence properties, validation of the model for conventional chain mooring lines and a detailed study of non-linear hardening of taut mooring lines.

The properties of the line modelled in this study are listed in Table 3.2. These properties represent an engineer design for a steel-wire type mooring line. However, the approach for the mooring model and the probability analysis can be directly applied to chain, synthetic or composite mooring lines.

The initial shape $\Gamma_{\boldsymbol{X}}$ of the mooring line is defined by solving the catenary equation. The anchor is located at $\boldsymbol{X}_{\mathrm{AN}}=[0,-h]^T$, where h represents the total water depth, as detailed in Table 3.1. The fairlead is positioned at $\boldsymbol{X}_{\mathrm{FL}}$, with its coordinates listed in Table 3.3. For each location, the fairlead is placed 3 m below the mean sea level. The horizontal position of the fairlead is determined to ensure consistent surge restoring stiffness k_{xx} of the mooring line across all locations. Table 3.3 presents the surge restoring stiffness values for each mooring line. To calculate this surge restoring stiffness, we evaluate the horizontal restoring force at the fairlead for two displaced positions: $\boldsymbol{X}_{\mathrm{FL}}+[0.01h,0]^T$, yielding $\boldsymbol{f}^+_{\mathrm{FL}}=[f^+_x,f^+_z]$, and $\boldsymbol{X}_{\mathrm{FL}}-[0.01h,0]^T$, yielding $\boldsymbol{f}^-_{\mathrm{FL}}=[f^-_x,f^-_z]$, where h denotes the water depth. The surge restoring stiffness is then computed as $k_{xx}=\frac{f^+_x-f^-_x}{2\times 0.01h}$.

Table 3.3 also includes the calculated length of the suspended portion of the line, L_{susp}, in the initial configuration. It is important to note that the total initial length of the line, L_0, is selected solely to ensure that the line does not become fully suspended during the tested sea states. Consequently, for locations 10–12, which correspond to greater water depths, the total line length is increased from 75 m to 105 m to prevent the line from becoming fully suspended.

The loads on the line are determined by the environmental conditions. The velocity profiles for the current and wave are the key parameters controlled in this study. The current velocity profile along the depth is specified at five discrete depths listed in Table 3.1, $\boldsymbol{v}_{\mathrm{current}}(z=-di)=c_{di}$, and subsequently interpolated across the entire depth using the piecewise cubic Hermite interpolating polynomial (PCHIP) method. The wave environment is defined using a JONSWAP spectrum with the values of significant wave height (H_s) and peak time-period (T_p). The wave particle velocity $\boldsymbol{v}_{\mathrm{wave}}$ is then calculated by first obtaining a wave time-series via an inverse Fourier transform (IFFT) of the input spectrum, followed by application of the linear wave theory. The direction of propagation for the wave and the current are assumed to be along the mooring line, since this

Table 3.2: Properties of mooring line.

Property		**Value**
Diameter		0.068 m
Cross-sectional area	A	0.003619 m²
Mass per unit length	ρA	28.23 kg m^{-1}
Axial Stiffness	EA	232.7 MN
Unstretched length	L_0	75 m or 105 m
Normal Drag coefficient	$C_{d,n}$	2.6
Axial Drag coefficient	$C_{d,t}$	1.4
Sea-bed stiffness	k_b	3×10^6 Pa m^{-1}
Sea-bed damping	c_b	3×10^5 Pa s m^{-1}

Table 3.3: Position of the anchor and the fairlead for all locations.

Location l	Total depth (h) m	L_0 m	L_{susp} m	$\boldsymbol{X}_{\text{AN}}$ m	$\boldsymbol{X}_{\text{FL}}$ m	k_{xx} N m^{-1}
1	12.6	75	10.72	$[0.0, -12.6]^T$	$[67.7, -3.0]^T$	218
2	14.5	75	12.79	$[0.0, -14.5]^T$	$[66.3, -3.0]^T$	218
3	15.0	75	13.40	$[0.0, -15.0]^T$	$[65.9, -3.0]^T$	218
4 (RefLoc)	22.8	75	22.14	$[0.0, -22.8]^T$	$[60.0, -3.0]^T$	218
5	30.1	75	30.27	$[0.0, -30.1]^T$	$[54.5, -3.0]^T$	218
6	30.2	75	30.36	$[0.0, -30.2]^T$	$[54.4, -3.0]^T$	218
7	33.6	75	34.16	$[0.0, -33.6]^T$	$[51.8, -3.0]^T$	218
8	35.7	75	36.46	$[0.0, -35.7]^T$	$[50.2, -3.0]^T$	218
9	38.6	105	39.78	$[0.0, -38.6]^T$	$[78.0, -3.0]^T$	218
10	39.2	105	40.43	$[0.0, -39.2]^T$	$[77.6, -3.0]^T$	218
11	40.7	105	42.15	$[0.0, -40.7]^T$	$[76.4, -3.0]^T$	218
12	44.7	105	46.54	$[0.0, -44.7]^T$	$[73.4, -3.0]^T$	218

represents the extreme scenario for external loads acting on the line. The bed properties used in this study are also listed in Table 3.2.

The final component of the problem definition is the motion of the fairlead, $\boldsymbol{u}_{\text{FL}}$. The motion of the fairlead is governed by the dynamics of the floating body to which the mooring line is attached. In this analysis, we assume the mooring line is connected to a weightless buoy. Therefore, the characteristics of the floater, the forces acting on it, and, as a result, its dynamic response are not considered in the scope of this analysis. Consequently, the fairlead is assumed to follow the motion of the local wave particle. This assumption is made for two key reasons: (1) it represents the most extreme scenario for the motion of the floater; and (2) it is easily replicable in numerical models or through simple experimental setups, such as those described in Bergdahl *et al.*, 2016, without the typical challenges associated with scaled physical modelling of floaters. Accordingly, the motion of the fairlead is prescribed using an inverse fast Fourier transform (IFFT) and linear wave theory applied to the input JONSWAP spectrum, with the wave particle position calculated around $\boldsymbol{X}_{\text{FL}}$.

We would like to note that the current approach can be readily extended to include floater dynamics. If floater dynamics are incorporated, it would then be necessary to account for the second-order forcing acting on the floater within the analysis. We would like to highlight that the second-order forcing on the mooring line, resulting from wave and current-induced drag, is already included in the present study via Equation (3.6) and Equation (3.7). Incorporating floater properties would make the resulting Bayesian network specific to those floater properties. Therefore, to address a more general case, it would be necessary to consider a range of floater characteristics and include them as inputs to the Bayesian network. While this is beyond the scope of the current manuscript, we intend to explore this direction in future work.

3.3. MODELLING APPROACH

In this section, the methodology is described. An overview of the detailed methodology is presented in Figure 3.3.

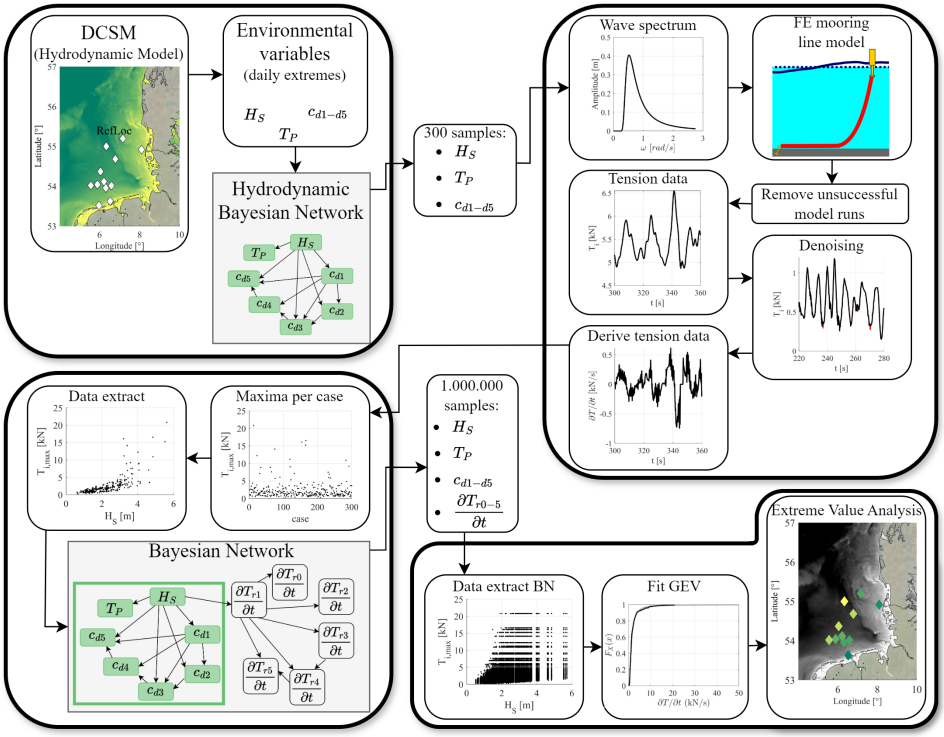


Figure 3.3: Detailed framework of the methodology applied in this study.

First, data is extracted from the DCSM hydrodynamic model for the months September to April for each location l of interest. This winter period is chosen, as this is the main growth period of the species Saccharina latissima, whose cultivation structure at RefLoc 4

serves as a foundational basis for this study. The daily maxima of the significant wave height, H_S, are selected together with the corresponding peak wave period, T_P. This is because wave heights have a greater influence on the mooring line tensions than current flows v_{di}. The maximum of current flows v_{di}, within a time window of ±4 hours around the daily maximum of the wave height, at the depths describes above are selected as concomitants.

In the next step, Spearman's rank correlations (Spearman, 1987) between the variable pairs per location in the North Sea are analysed. Based on the underlying physics and correlation matrices, structures of individual Bayesian Networks (BNs) are set up, denoted as $BN_{h,l}$. Therefore, two goodness-of-fit measures are applied, the d-calibration score and the Cramér-van-Mises method (CvM). Both, the BNs and goodness-of-fit measures, are further described in Section 3.3.3. From each $BN_{h,l}$, $n = 300$ samples are drawn. Wave spectra per sample set are generated and, along with steady current flow magnitudes, serve as input for the finite-element model for flexible mooring lines.

The flexible mooring line model produces time series data of tensions at the anchor, T_{r0}, fairlead, T_{r5}, and four intermediate rope positions p, T_{r1-4}. In certain scenarios, the numerical model exhibited poor convergence, particularly in cases involving line folding or the presence of compression. Additionally, the tension time-series contains occasional instances of noise introduced by the spring-damper seabed. The handling of these issues is summarised in Section 3.3.4. The tension time-series are then derived and maximum tension rates, $\frac{\partial T_{r,p}}{\partial t}$, at the same rope positions p are selected.

For each location l in the North Sea, extended BNs are established, $BN_{ht,l}$, incorporating hydrodynamic data from $BN_{h,l}$, which served as input for the flexible mooring line model, and maxima of tension rate data at different rope locations, p, $\frac{\partial T_{r,p}}{\partial t}$. These $BN_{ht,l}$ are also assessed via two goodness-of-fit measures and generate a sample set of $n = 1.000.000$ samples per location l. The structures of $BN_{ht,l}$ provide further understanding of the underlying relationships among variables and their spatial differences.

Finally, the tension rate data $\frac{\partial T_{r,p}}{\partial t}$ generated by $BN_{ht,l}$ are fitted to a Generalised Extreme Value (GEV) distribution, to obtain values corresponding to certain probabilities for each study location.

3.3.1. SPEARMAN'S RANK CORRELATION COEFFICIENT

Between each pair of variables, Spearman's rank correlation coefficient, ρ, is calculated (Spearman, 1987). This acts as a foundation for the dependence analysis and the structure of the Bayesian Network for each location. Spearman's rank correlation coefficient ρ is Pearson's correlation coefficient (Pearson and Galton, 1895) computed with the ranks of each variable. It assesses the degree of monotonic dependence between the random variables and is given by

$$\rho = \frac{Cov[R(X), R(Y)]}{\sigma_{R(X)}\sigma_{R(Y)}} \tag{3.11}$$

where the covariance of the ranked variates of X and Y is denoted $Cov[R(X), R(Y)]$, and $\sigma_{R(X)}$ and $\sigma_{R(Y)}$ are their standard deviations. $\rho \in [-1, 1]$, where $\rho = 1$ and -1 denote perfect (monotonic) positive and negative correlation, respectively. The significance of

observed correlations is assessed by computing p-values using the student-t distribution (Zar, 1972):

$$t = \rho \sqrt{\frac{n-2}{1-\rho^2}} \tag{3.12}$$

where n represents the sample size or degrees of freedom. The p-value is defined as the combined area in both tails of the student-t distribution. The p-value per tail is obtained from the cumulative distribution function of the student-t distribution, evaluated at the t-score from Equation (3.12) with $(n-2)$ degrees of freedom. A rank correlation is statistically significant if the p-value falls below the significance level of $\alpha = 0.05$.

3.3.2. BIVARIATE COPULAS

Copula models are a popular approach to model the dependence structure between random variables, independent from their marginal distributions. Any joint multivariate distribution can be described in terms of a set of one-dimensional univariate marginal distributions in [0,1] and a copula that models the dependence between the random variables Sklar (1959). For the bivariate case, let $H(x, y)$ for $(x, y) \in \mathbb{R}^2$ be a joint distribution with univariate marginals $F(x)$ and $G(y)$, then there exists a copula C in the unit square $I^2 = ([0,1] \times [0,1])$:

$$H(x, y) = C(F(x), G(y)) \tag{3.13}$$

Equation (3.13) is satisfied for all $(x, y) \in \mathbb{R}^2$ (Genest and Favre, 2007; Joe, 2015; Nelsen, 2006).

3.3.3. GAUSSIAN COPULA-BASED BAYESIAN NETWORK

A Bayesian Network is a probabilistic graphical model to represent high-dimensional and complex dependence structures (Hanea *et al.*, 2006). Bayesian Networks are directed acyclic graphs (DAG), representing the joint distribution of a set of variables. Each node represents a random variable, while the arcs between the nodes describe (conditional) probabilistic dependence between the variables. The predecessor and successor of a node are called parent and child nodes, respectively. In this study, this dependence is described via one-parameter conditional copulas, which are parameterised by Spearman's rank correlation coefficient ρ (Hanea *et al.*, 2015). Here, the Gaussian copula is selected for its computational efficiency, enabling fast inference when handling large and complex problems. Thus, the dependence structure of the variable set is described by a multivariate normal distribution. Note, however, that for a copula-based Bayesian Network, any other one-parameter copula could be used to describe the joint dependence of variables. Each node represents a random variable described by its empirical distribution function. Absence of arcs between nodes indicates conditional independence between two variables (Hanea *et al.*, 2015). The joint probability density function (pdf) of a Gaussian copula-based Bayesian Network (GCBN) with m variables can be described as:

$$f_{1,\dots,m}(x_1,\dots,x_m) = f_1(x_1)\prod_{v=2}^{m} f_{v|Pa(v)}(x_v|x_{Pa(v)}) \tag{3.14}$$

where the joint density of m variables is denoted as $f_{1,\dots,m}$, the marginal densities are denoted by f_v and the conditional densities are denoted by $f_{v|}$.. Note that each node v represents a variable X_v, while the parent nodes of v form the set $Pa(v)$ (Hanea *et al.*, 2015).

To assess the goodness of fit of the Gaussian copula models per variable pair, the Cramér-von Mises (CvM) statistic S_n is used (Genest *et al.*, 2009b). The sum of squared differences between the empirical joint cumulative distribution function (cdf) and the joint cdf of the parametric copula model is determined. S_n , the CvM test statistic, for a sample length of n is determined as:

$$S_n(\boldsymbol{u}) = n\sum_{|n|}\left\{C_n(\boldsymbol{u}) - C_{\widehat{\theta}_n}(\boldsymbol{u})\right\}^2 \quad , \boldsymbol{u} \in [0,1]^2 \tag{3.15}$$

Here, $C_n(\boldsymbol{u}) = \frac{1}{n}\sum_{i=1}^{n} 1\,(U_i \leq \boldsymbol{u})$ is the empirical copula and $C_{\widehat{\theta}_n}(\boldsymbol{u})$ is the parametric copula with the parameter $\widehat{\theta}_n$ estimated from the sample. The lower $S_n(\mathbf{u})$ for the Gaussian copula compared to other copula models, the better the Gaussian copula-based Bayesian Network (GCBN) captures the joint dependence among variable pairs. Another goodness-of-fit measure, known as d-calibration score (Morales-Nápoles *et al.*, 2014), is applied. The dissimilitude between two correlation matrices is determined via the Hellinger distance, where a score of 0 indicates that the correlation matrices differ element-wise, while a score of 1 signifies identical matrices. The d-calibration score, however, is sensitive to the sample size. For more information, the reader is referred to Morales-Nápoles *et al.* (2014).

In this study, the open-source PYTHON library BANSHEE, version 1.3, is used for implementing GCBNs and for the calculation of the goodness-of-fit measures d-calibration score and CvM statistics (see Koot *et al.* (2023) and Paprotny *et al.* (2020) and for the MATLAB implementation see Mendoza-Lugo and Morales-Nápoles (2023)).

3.3.4. PROCESSING OF NUMERICAL TENSION DATA

The finite-element model for flexible mooring lines, described in Section 3.2.3 provides time series of tension data in the mooring line for different positions p in the rope, while the tension close to the anchor, T_{r0}, fairlead, T_{r5}, and 4 intermediate depths, T_{r1-4}, are chosen for the analysis. These time series have a resolution of 0.1 s and a duration of 23 minutes, of which the first 3 minutes are used as ramp-up time and are disregarded from the analysis. If the finite-element model fails to converge at a given time-step due to unphysical folding or compressive forces, the corresponding data point, along with 1 second before and after, are removed from the analysis in post. If a model run does not converge for more than 5 times, it is disregarded from the analysis. As a result, each location l in the North Sea has between 286 and 300 cases. Location 1, which has the shallowest water depth among all locations, was excluded from the analysis. In this location, the mooring line frequently became slack, primarily because the selected line properties and the catenary configuration were unsuitable for such shallow water

depths. Ideally, a multi-segment mooring system incorporating a buoy would be more appropriate for this location to reduce the required line length. A single-segment catenary system alone would necessitate an impractically long line.

Before transforming the tension data, $T_{r,p}$, to the relevant tension rates, $\frac{\partial T_{r,p}}{\partial t}$, outliers are removed from the time series. These outliers primarily result from the damped-spring implementation of the seabed. The spring and damper coefficients for the seabed, listed in Table 3.2, were selected to function effectively in most cases. However, in certain instances, they may be slightly stiffer or more compliant than necessary, leading to occasional outlier values. An outlier is detected if its value exceeds three-times the moving-median of the tension time series. Spikes are removed only if they occur for a single time step. This ensures that potentially realistic snap loads, which typically persist over multiple time steps, are not erroneously filtered out. To properly analyse snap loads, which are characterised by high tension rates, $\frac{\partial T_{r,p}}{\partial t}$, they are in a next step obtained by applying the second order central difference approximation to the tension data $T_{r,p}$. The processing of the tension rate is performed in post-processing for two reasons: 1) it enables the filtering of outliers and irregularities in the tension time series, providing a more informed calculation of the tension rate, and 2) the second-order accurate central difference approach is more general and does not require exporting data from intermediate steps.

In a next step, extremes of tension rates, $\frac{\partial T_{r,p}}{\partial t}$, are selected for further analysis. Therefore, two approaches are compared: block maxima (BM) and peaks-over-threshold (POT). According to the Fisher-Tippett theorem (Fisher and Tippett, 1928), the probability distribution of a sample of the maxima per block (here, a block is defined as a case of 20 minutes simulation time) converges to generalised extreme value (GEV) distributions. While, according to Pickands' theorem (Pickands, 1975), the probability distribution of a sample of peaks-over-threshold converges to generalised Pareto distributions (GPD). There is equivalence between these distributions, which have the same shape parameter. Thus, the GPD is compared with the GEV for the tension rates at the anchor, $\frac{\partial T_{r0}}{\partial t}$, for the reference location (RefLoc) 4, as this position p in the rope exhibits most extremes. To fit the GPD, a threshold is defined via an analysis of the stability of parameters of the GPD for different tension rate percentiles. The 90th percentile is found to be a proper threshold. Design values for the GEV and GPD are calculated and found to be close to each other. Because of this and because the data used in this study is already available as independent blocks in a sufficiently high amount, the block maxima are chosen to obtain extreme tension rates for further analysis. Additionally, the GEV is also recommended by Xu and Guedes Soares (2021), who analysed different distribution models for extreme mooring tensions. However, mixture distribution models might offer a better fit under certain conditions (see also Xu *et al.* (2022)).

3.4. RESULTS

3.4.1. HYDRODYNAMIC GAUSSIAN COPULA-BASED BAYESIAN NETWORKS

Before generating tension data with the FE mooring line model described in Section 3.2.3, hydrodynamic input data are required for the model. Therefore, the daily extremes from

the DCSM model are used to set up Gaussian copula-based Bayesian Networks (GCBNs), which generate a set of $n = 300$ data. This is done as the selected data from the DCSM could under-represent extremes, as it only spans five years and thus may not capture rare but critical conditions. Thus, the GCBNs generate data whose dependence is described by the Gaussian copula. These data, account for diverse conditions, while accounting for the underlying dependence structure of the hydrodynamic variables.

In this section, the structures of the GCBNs of the hydrodynamic data for each location l are described, denoted as $BN_{h,l}$. Therefore, seven variables are considered, as described above: significant wave height H_S, peak wave period T_P and current magnitudes $c_{d,i}$ at five different depths i throughout the water column (see Table 3.1). Figure B.2 in the Section B.2 presents the hydrodynamic data from the DCSM for RefLoc 4, which serves as input for $BN_{h,4}$. The data has been transformed to uniform [0,1] to facilitate its use in the copula models (see Section 3.3.2). The uniform marginal distributions are shown on the diagonal. The pairwise Spearman rank correlation coefficients, ρ, are displayed on the upper-right part of the Figure B.2, while the scatter plots in the lower-left section of the figure show the relationship between variables. Visually, some asymmetries in the bivariate pairs are observed. In order to further investigate this observation, we use the CvM statistic, which is described in Section 3.3.3. According to CvM, the dependence structure of the variable pair of H_S and T_P is best captured by the symmetric Frank copula, while for other variable pairs, the Gaussian copula is generally the most suitable fit; however, there are instances where the Frank or asymmetric Gumbel copulas provide comparable fits. Given that the results of the CvM statistic predominantly indicate that the symmetric Gaussian or Frank copulas provide the best fit, the GCBN (see Section 3.3.3) is deemed an acceptable method for modelling the dependence among hydrodynamic variables.

Figure 3.4 shows the underlying structure, which is valid for all 12 locations l in the south-eastern North Sea (see Figure 3.1). For each arc, the correlation between the variable pair is $r \geq 0.2$, while the significance level is below the threshold ($\alpha \leq 0.05$). In general, there is a strong correlation between the wave variables H_S and T_P. Regarding the current variables, the current magnitude near the seabed, denoted c_{d1}, has arcs to the current magnitudes throughout the water column, c_{d2-5}. To ensure that the $BN_{h,l}$ adequately captures the correlation matrix among current magnitudes, additional arcs between c_{d3} and c_{d4}, as well as between c_{d4} and c_{d5} are necessary. In addition, the northern locations 3, 4 and 12 (see Figure 3.1) require an arc between c_{d2} and c_{d3}, despite their different water depths. The single $BN_{h,l}$ structures per location l are shown in Figure B.3 in the Section B.3, marked in green.

Further differences in the structures of $BN_{h,l}$ among locations are observed on how the wave variables are connected with the current magnitudes. An arc consistently connects H_S with the current magnitude near the seabed c_{d1} or the current magnitude near the surface c_{d5}, or sometimes both. Specifically, the following patterns are observed. For all locations l in the area of interest, except locations 1 and 3, an arc connects H_S with the current component with the highest correlation and $\alpha \leq 0.05$, which is mostly c_{d5}, as this is the current close to the surface. Different from that, an arc between H_S and c_{d1} is added for all locations except for location 2 and 11 (shown in Figure B.3(b) and (i)). As shown in Figure 3.4, c_{d1} points to the other current variables c_{d2-5}.

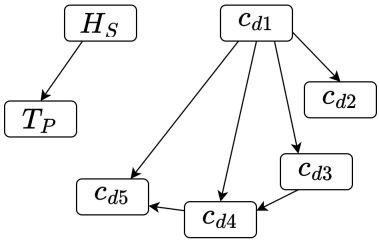


Figure 3.4: Underlying structure of GCBNs applicable to all locations with respect to hydrodynamic variables. For each location l, certain arcs are added to account for location-specific dependence structures.

In general, the following procedure is observed among the offshore locations 4 until 12, while the GCBN structure of location 11, $BN_{h,11}$, requires the least amount of arcs (Figure B.3(i)).

- Adding an arc between H_S and c_{d1} leads to the GCBN structures of location 5, 6 and 7, $BN_{h,5-7}$, (see Figure B.3(e)).
- The structure of $BN_{h,5-7}$ results in the GCBN structure of location 8 and 9, $BN_{h,8-9}$, by adding another arc between H_S and c_{d3} (see Figure B.3(f) and (g)).
- The addition of an arc between H_S and c_{d2} leads to the structure of location 10, $BN_{h,10}$, (see Figure B.3(h)).
- Another additional arc between c_{d2} and c_{d3} represents the structure of the GCBN for location 4, $BN_{h,4}$, and 12, $BN_{h,12}$, (see Figure B.3(d) and (j)).

For the nearshore locations, certain arcs of $BN_{h,4}$ and $BN_{h,12}$ can be removed to get a satisfactory fit of the correlation matrix of the GCBN for these locations.

- Based $BN_{h,4}$ and $BN_{h,12}$, removing arcs between H_S and c_{d3} and H_S and c_{d5} results in $BN_{h,3}$.
- Additionally, $BN_{h,1}$ results from $BN_{h,11}$ by removing an arc between H_S and c_{d5} and adding an arc between H_S and c_{d1}, which is necessary, despite the slightly lower correlation compared to the correlation between H_S and c_{d5}. A connection between these variables degrades the fit, possibly because c_{d1} instead of c_{d5} points to the other current variables.
- $BN_{h,2}$ also builds on the structure of $BN_{h,11}$ by removing an arc between H_S and c_{d1}.

Summarising, it can be observed that the smaller the distance of a location to the shore and the more it is affected by strong (tidal) currents (e.g. between islands or close to an estuary, see locations 1 to 3, 5 to 8 and 11), the fewer arcs are required compared to locations further offshore (e.g. locations 4, 10 and 12).

3.4.2. METHODOLOGICAL APPLICATION ON TENSION DATA EXEMPLIFIED BY THE REFLOC 4

In this section, the described methodology in 3.3.4 regarding the processing of the tension data, T_{r0-5}, of the FE model is applied and presented for the RefLoc 4 as an example. Similar type of analyses may be performed on other locations.

DENOISING OF TENSION DATA

First, the denoising process of the tension data $T_{r,p}$ described in Section 3.3.4 is demonstrated. Although all tension data time-series throughout the mooring line undergo the described denoising process, the anchor rope position is selected here to showcase the process. This is, because for the tension data at the anchor, T_{r0}, unrealistic spikes due to numerical errors are most frequent and intense, compared to other positions p along the rope at different depths. Figure 3.5 *(a)* shows the denoised time-series of T_{r0} for the shallow location 2 under moderate conditions ($H_S = 3.1m$, $T_P = 8.1s$) and *(b)* shows the denoised time-series of T_{r0} under more intense conditions for the RefLoc 4 ($H_S = 3.6m$, $T_P = 11.0s$). Both time-series have a duration of 20 minutes.

Figure 3.5 *(c)* and *(d)* demonstrate the difference of the raw time-series (in red) and the denoised time-series (in black). *(c)* shows a segment of *(a)*, and *(d)* shows a segment of *(b)*, each for 50 s respectively. The denoised time-series are obtained by detecting and removing the outliers, which are no longer than one timestep, i.e. 0.1 s, and interpolating linearly among them. This is necessary since the tension rates, $\frac{\partial T_{r0-5}}{\partial t}$, for later analysis are derived from the time-series of T_{r0-5} via the second order central difference approximation. Thus, peaks in T_{r0-5}, which stem unphysically from the seabed spring-damper system, lead to unrealistically high tension rates.

The obtained tension rates, $\frac{\partial T_{r,p}}{\partial t}$, for a shallow location in Figure 3.5 *(a)* are shown in *(e)*, while the tension rates for the RefLoc 4 from *(b)* are displayed in *(f)*. The maxima of the time-series of $\frac{\partial T_{r0-5}}{\partial t}$ are selected per case for each location l in the North Sea.

ANALYSIS OF VARIABLES IN PHYSICAL UNITS

These extreme tension rates, along with the hydrodynamic data, which was used as input for the FE model, are subsequently utilised to set up extended GCBNs. Here, the data for the RefLoc 4 is presented in Figure 3.6. On the diagonal, the empirical distributions are shown as histograms. On the lower-left part of the figure, the scatter plots per variable pair are presented, which allow for a comparison of pairwise relationships. Among the current magnitudes, c_{di} at different depths i, strong correlations are visible, especially among current magnitudes that are closely aligned depth-wise. A similar, even stronger pattern can be observed for the tension rates, $\frac{\partial T_{ri}}{\partial t}$, at different positions in the mooring line. On the top-right part of Figure 3.6, Spearman's rank correlation coefficients ρ are shown, which confirm these observations. These correlations are further analysed in Section 3.4.2.

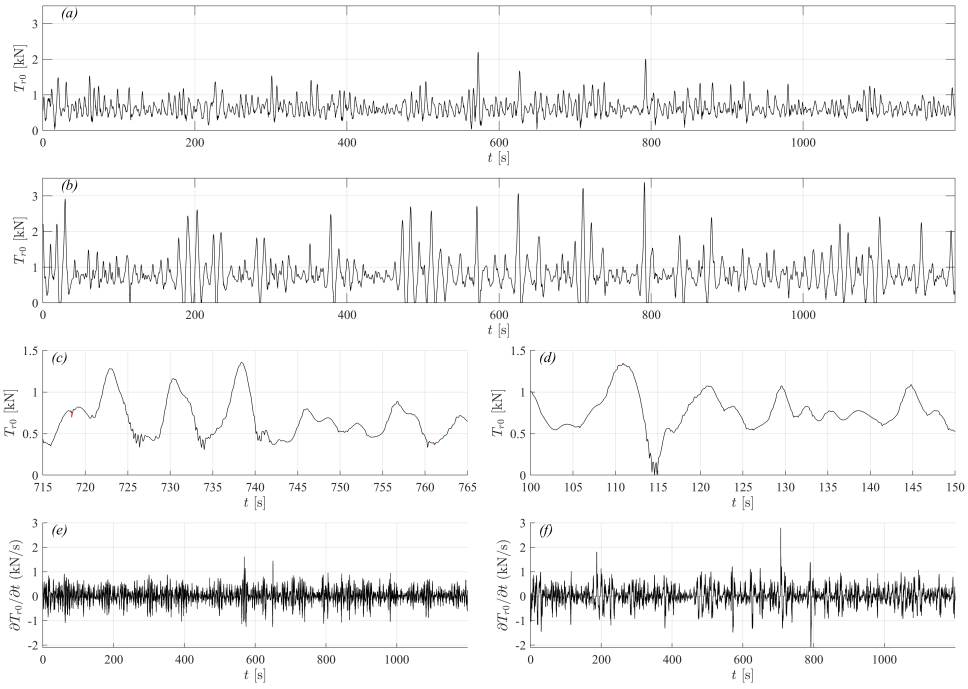


Figure 3.5: Anchor data visualisation: (*a*) Denoised tension data for shallow location 2 under moderate conditions ($H_S = 3.1m, T_P = 8.1s$); (*b*) Denoised tension data for RefLoc 4 under storm conditions ($H_S = 3.6, T_P = 11.0$); (*c*) Close-up view of raw versus denoised tension data for shallow location 2 under the same conditions as in (*a*); (*d*) Close-up view of raw versus denoised tension data for RefLoc 4 under the same conditions as in (*b*); (*e*) Tension rate data for shallow location 2 under the same conditions as in (*a*); (*f*) Tension rate data for the RefLoc 4 under the same conditions as in (*b*).

As described above, the denoising process aims to remove unrealistic spikes in the data. However, occasionally, more extreme values tend to occur at greater depths. Less extreme tension rates in upper rope positions, such as at the fairlead, $\frac{\partial T_{r5}}{\partial t}$, are due to the weight of the mooring line, which might dampen sudden movements and reduces extreme tension rates in these rope positions.

ANALYSIS OF VARIABLES IN $[0,1]^2$

In a next step, the 13 variables are transformed into a uniform distribution on the interval $(0,1)$. This is important as the applied GCBN is based on using Gaussian copulas (see Section 3.3.3), which operate in this interval.

Figure 3.7 shows the ranked variables for the RefLoc 4 (FINO3) as pairwise scatter plots on the lower-left part, together with the one-dimensional uniform margins on the diagonal and their density contour plots on the top-right part of the figure. This allows one to better see the above described correlations between variables independent from their margins, different from Figure 3.6, which shows the variables in their physical units.

Here, a high correlation between the significant wave height H_S and the tension rates, is visible. This is also evident in the corresponding density plot, where the contours have a narrow and symmetric shape. This indicates that H_S is the main driver for the tension rates in this case, as the correlations between T_P or $c_{d,i}$ with tension rates are significantly lower.

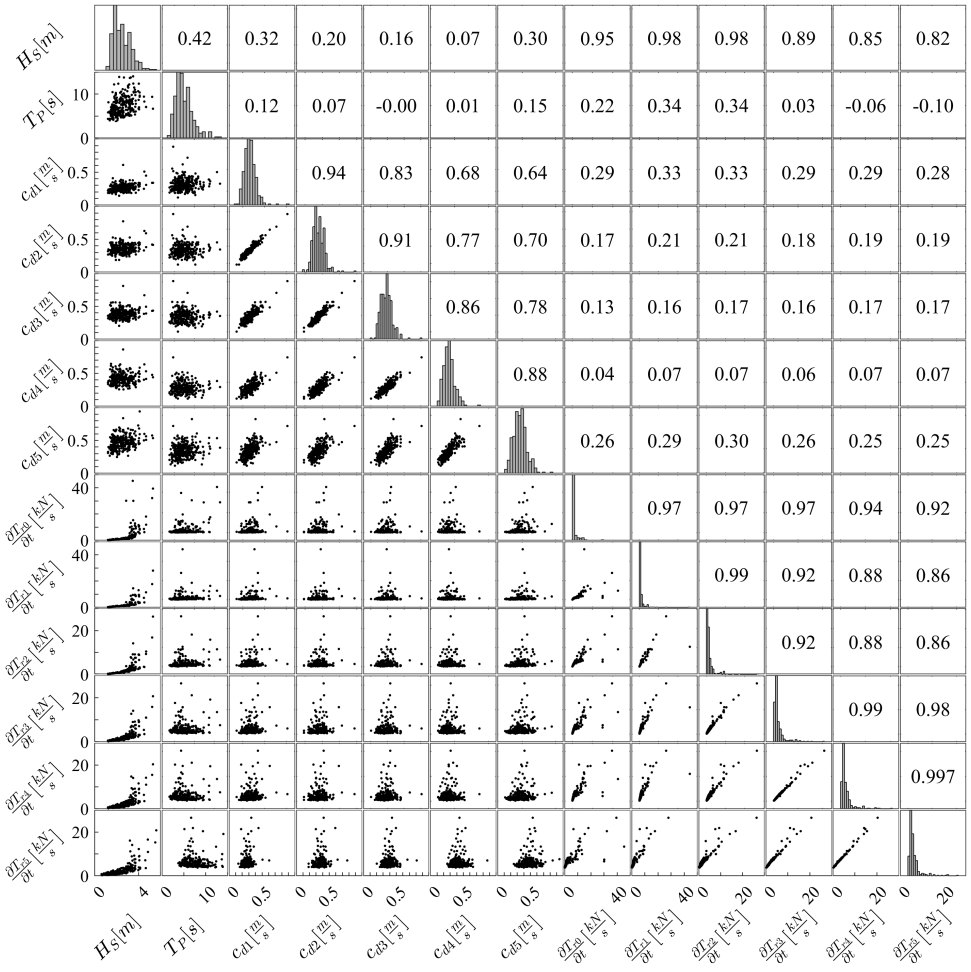


Figure 3.6: Pairwise scatter plots of the studied variables in their physical units for RefLoc 4 are presented in the lower-left part, together with their empirical marginal distributions on the diagonal and rank correlation coefficients above the diagonal.

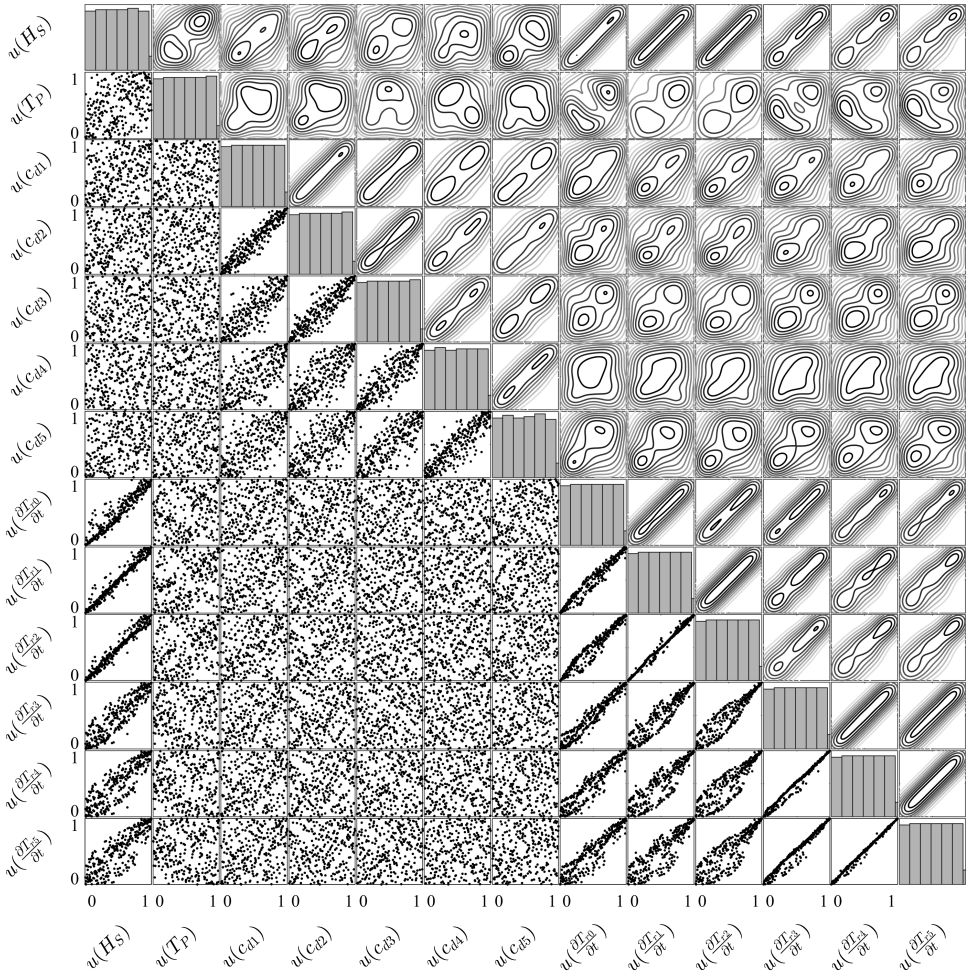


Figure 3.7: Pairwise scatter plots of the studied variables transformed to unit space for RefLoc 4 are presented in the lower-left part, together with their one-dimensional uniform margins on the diagonal and their densities as contours above the diagonal.

3.4.3. BAYESIAN NETWORKS OF HYDRODYNAMIC VARIABLES AND TENSION RATES

The GCBN structures of the hydrodynamic input variables described in Section 3.4.1 are extended by the derivatives of the tensions, $\frac{\partial T_{r,p}}{\partial t}$, at different positions p in the mooring line. The structures of the GCBNs for each location l in the North Sea, denoted $BN_{ht,l}$, can be found in Section B.3.

For each location l, different structures have been compared and assessed, to identify the one that best represents the dependence structure of the empirical data, hydrodynamic data from the DCSM and maxima of tension rates, while aiming to minimise

complexity by keeping the number of arcs necessary low. The structures of $BN_{ht,l}$ are assessed via the d-calibration score and CvM statistic. For sample sizes of 80 to 300, the d-calibration score between the correlation matrix of $BN_{ht,l}$ and the empirical data varies between 0.5 and 0.8. Note however that this method is known to be sensitive to the sample size. The Gaussian copula for each variable pair connected by an arc is additionally verified via the CvM statistics. For the results of CvM among the hydrodynamic variables, see Section 3.4.1. An example of the empirical data for RefLoc 4, which serves as input for the $BN_{ht,4}$, is given in Figure 3.7. For this location, but also other locations with medium distance to shore (3 to 9), it appears that there is some upper tail dependence among the tension rates. According to CvM, however, the joint dependence can be mostly described via the symmetric Gaussian copula, or the Frank copula. For some variable pairs, such as among tension rates, the Gaussian copula fits to a similar degree as the Gumbel copula, or even better. This is the same for the joint dependence structure among currents for location 3.

The correlation matrix of $BN_{ht,l}$ should approximately match the correlation matrix of the empirical data. In addition to the Spearman correlation coefficients ρ, p-values are determined as a measure for statistical significance of ρ. The left side of Figure 3.8 shows the correlation matrix for RefLoc 4, while the right side presents the corresponding p-values. Note that for both, top-right above the diagonal black line corresponds to the empirical data, while parts below the line correspond to $BN_{ht,4}$. [1]

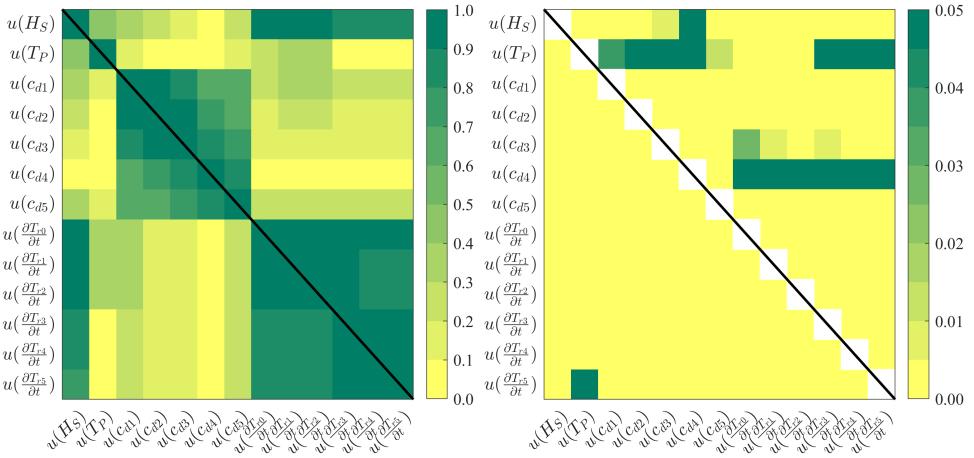


Figure 3.8: Left: Spearman's rank correlation coefficients for the studied variables at RefLoc 4. Right: Corresponding p-values. In both plots, values above the black diagonal correspond to empirical data, while those below correspond to data from the GCBNs.

For the RefLoc 4, all correlation coefficients between the empirical variables for this location are positive, except for the pairs of T_P with the current magnitudes c_{d3} and maxima of tension rates, $\frac{\partial T_{r4-5}}{\partial t}$, both close to the surface. However, these correlations

[1] The rank correlation coefficients and p-values are similar across the selected locations in the North Sea and thus not all are presented in this study, but can be made available upon request.

are close to zero and their p-values are > 0.05, thus the correlations are found not to be significant. Generally it can be seen that the correlation matrix of $BN_{ht,4}$, lower left, does match well the correlation matrix of the empirical data, upper-right. Slight difference in the correlations of > 0.1 are observed for RefLoc 4 for the variables pairs of T_P with $\frac{\partial T_{r,0}}{\partial t}$ and $\frac{\partial T_{r,0}}{\partial t}$ with $\frac{\partial T_{r,3}}{\partial t}$. For both, empirical correlation matrix and the one of $BN_{ht,4}$, there are strong correlation coefficients between H_S and tensions rates, as well as among current magnitudes and among tension rates. The p-values in for the empirical data are above thresholds for the variables pairs of H_S with c_{d4} and T_P with current magnitudes and tension rates close to the surface. However, these variable pairs have very low correlations and are thus not connected by an arc in the $BN_{ht,4}$.

The described characteristics for RefLoc 4 are generally observed across all locations l. However, for locations 2, 3, 5 to 8, and 12, some arcs between variables are needed despite their low correlation linked with a high p-value. Removing these arcs would make the fit worse.

Generally, it is observed that the structure among the tension rates $\frac{\partial T_{r,p}}{\partial t}$ depends on the distance to shore. The structure for the offshore locations 10 to 12 with high water depths is identical. Adding two arcs among the upper tension rates in the rope, the structure becomes identical for the medium-distant locations 4 to 9. Adding two more arcs to this structure among the tension rates T' results in the structure of the near-shore and shallow locations 2 and 3.

3.4.4. EVALUATION OF TENSION RATES AT VARIOUS PERCENTILES

The above described structures for each $BN_{ht,l}$ are used to generate $n = 1.000.000$ samples for seven hydrodynamic variables and five tension rate variables per location l in the south-eastern North Sea. These sample sets are used to assess expected tension rates and to investigate the spatial differences across the south-eastern North Sea.

Figure 3.9 presents the tension rate values for the 90th percentile across the studied locations, with the left side showing values at the anchor ($r0$) and the right side showing values at the fairlead ($r5$). Thus, thresholds are shown, indicating that 90% of the tension rates in the mooring line fall below these values. As previously mentioned, results for location 1 are not reported, because the mooring line simulations have not been successful due to the insufficient water depth. The background in Figure 3.9 indicates the water depths. For detailed information on the water depth in the area, see Figure 3.1, and for the precise water depths per location l, see Table 3.1. The 90th percentile is determined by fitting the generalised extreme value (GEV) distribution to the selected maxima tensions, as described in Section 3.3.4. Figure B.4 in Section B.4 provides an overview of the maxima of the tension rates at the anchor for RefLoc 4. It compares the empirical cumulative distribution function (cdf) with the cdf of the fitted GEV distribution. The inverse cdf of the GEV is used to determine the value corresponding to a probability of 0.9. The tension rates at the anchor, $\frac{\partial T_{r0}}{\partial t}$, in the left map in Figure 3.9 are similar among the locations 9 to 12. These locations have the greatest water depth among those studied, with depths exceeding 38.6 m. For the fairlead (right side of Figure 3.9), location 12 stands out by experiencing significantly higher tension rates compared to the other locations. This could be due to the water depth, as location 12 is the deepest, with a water depth of

about 45 m. Additionally, mooring lines modelled for locations 4 to 8 also exhibit similar tension rates at both the anchor and fairlead, with water depths ranging from 22.8 to 35.7 m. Similarly, the tension rates for the mooring lines at the shallow locations 2 and 3, with depths below 15 m, are also comparable.

Following these observations, the rank correlations between the tension rate values for the 90th percentile and the water depths across the studied locations are calculated. For the anchor, the correlation is 0.83, while for the fairlead, the correlation is found to be 0.91. These high correlations indicate that the water depth is one of the main variables influencing the tension rates in the mooring line.

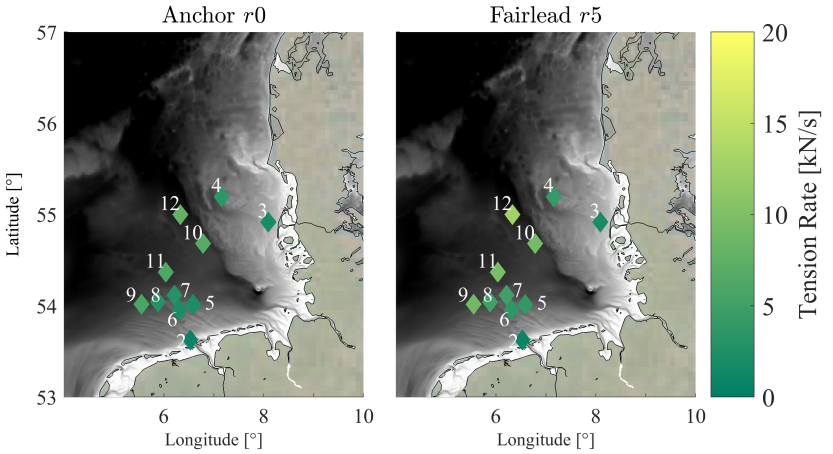


Figure 3.9: Design values of tension rates for $P(X \leq 0.9)$ at the anchor (left) and at fairlead (right).

To better visualise the differences among locations l with increasing depth d, as well as among several positions p in the mooring line, Figure 3.10 presents the 90th percentile of tension rate values for three different conditionalisation scenarios, indicated by different markers. These tension rates are calculated for the anchor ($r0$), four intermediate rope locations, and for the fairlead ($r5$). For comparison, the results for the 95th percentile are shown in Figure B.5 in Section B.5.

The unconditionalised scenario is marked by a triangle ($\triangledown$). The results of this scenario at the anchor (top-left plot of Figure 3.10) correspond to the map on the left, while the results of the fairlead correspond to the map plot on the right in Figure 3.9. Note that at locations 2 and 3, the water depth is low and thus, the vertical resolution is not sufficient to separate the rope into 6 positions. Therefore, intermediate rope position $r4$ overlaps with the fairlead $r5$, so no data is shown for $r4$ for these locations. For the unconditionalised scenario ($\triangledown$), it can be seen that the differences in tension rates across mooring line positions, p, are generally small. Tension rates appear to be higher at the fairlead $r5$, compared to lower-located rope positions. Locations 2 and 3 are an exception to this due to their low water depth. For the shallow locations 1 to 8, the differences in tension rates for the 90th and 95th percentile for the unconditionalised scenario are comparably small (see Figure B.5 in Section B.5). This is different for the deeper locations 9 to 12, where the differences in tension rates for the 90th and 95th percentile for the

unconditionalised scenario marked by ▽ are relatively large, despite using the symmetric Gaussian copula, to generate the sample sets. Therefore, the distribution of tension rates for deep locations appears to have a heavier tail.

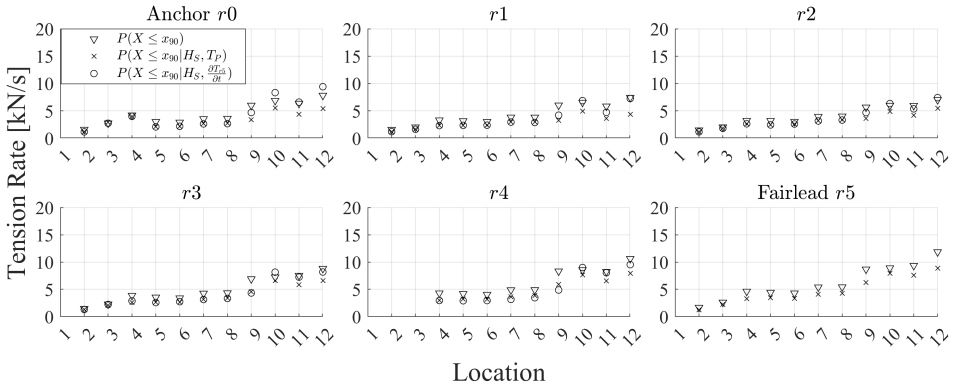


Figure 3.10: The 90th-percentile design values of tension rates are presented for six rope positions (r), including the anchor and fairlead, for three conditionalising cases.

Figure B.6 in Section B.5 presents the 90th percentile results, normalised by the static pretension per location l. Similar to the non-normalised Figure 3.10 and Figure B.5, a general trend between water depth and tension rate is visible as well, except for two locations. Location 3 and 4, which are relatively shallow with a water depths of 15.0 m and 22.8 m respectively, have significantly higher tension rates, particularly at the anchor, compared to other locations with similar depths. This can be attributed to their exposure to the dominant south to south-west storm direction (de Winter *et al.*, 2013; Sušelj *et al.*, 2010) and the influence of a long wind fetch. Same applies to the deeper locations 9 to 12, whereas locations 1, 2 and 6 to 8 are relatively sheltered by the coastline.

3.4.5. EVALUATION OF TENSION RATES FOR DIFFERENT CONDITIONALISATION SCENARIOS

The results above in Section 3.4.4 are here compared to two conditionalising cases. The power of the GCBN to conditionalise on certain variables is done here for several reasons. First, conditionalising on certain variables provides further insights into the characteristics of the conditioned extreme tension rates and thus helps to assess the reliability of a floating structure under certain conditions. Therefore, by holding certain variables constant, uncertainties in the analysis of extreme tension rates can potentially be reduced. Furthermore, if information about specific variables is available, its impact on tension rates in the mooring lines can be assessed. The two conditionalising cases per location l considered here are:

1. Conditionalising on the 80th percentiles of H_S and T_P
2. Conditionalising on the 80th percentiles of H_S and $\frac{\partial T_{r5}}{\partial t}$

The first case is selected as waves are earlier identified as most influential for the tension rates. Additionally, they are comparably easy to measure. The second case is chosen by assuming that information about the tension at the fairlead is available, e.g. due to an installed sensor between the mooring line and the floating structure. This helps to receive information about rope positions p below the fairlead, including the anchor. Regarding the first case, it is important to note that for the RefLoc 4, conditionalising on T_P has no influence on the tension rates, if H_S is known. This is due to conditional independence of tension rates and T_P given H_S (see the structure of $BN_{ht,4}$ in Figure B.3d in Section B.3). BANSHEE (Koot *et al.*, 2023; Paprotny *et al.*, 2020) is used to generate sets of $n = 1.000.000$ samples conditionalised on certain variables and values per location l, which are shown in Table 3.4.

Table 3.4: 80th percentile values for variables used for conditionalisation.

Location l	H_S [m]	T_P [s]	$\frac{\partial T_{r5}}{\partial t}$ [kN/s]
1	-	-	-
2	2.84	8.49	2.88
3	2.63	8.20	1.80
4	2.98	8.73	2.93
5	2.84	8.49	2.88
6	2.86	8.60	2.89
7	3.11	8.89	3.21
8	3.00	8.44	3.35
9	3.16	8.57	5.00
10	3.09	8.74	5.23
11	3.11	8.64	5.77
12	3.34	9.22	7.20

Similar to the results described in Section 3.4.4 for the unconditionalised case, GEV distributions are fitted to the tension rates and the obtained values corresponding to 90th and 95th percentile are analysed. In Figure 3.10 and Figure B.5 in Section B.5, the results for the conditionalising cases 1, indicated by $\times$, and 2, indicated by $\circ$, are presented for the 90th and 95th percentiles, respectively. With regard to the water depths, a similar behaviour for the conditionalised tension rates for the 90th percentile is observed compared to the unconditionalised case. The larger the water depth, the higher the 90th percentile of tension rates. Additionally, for most of the locations l and rope positions p, the 90th percentile of conditionalised tension rates are lower than those of the unconditionalised case (marked as $\triangledown$). Details of the RefLoc 4 are displayed in Table 3.5. The reason for these unexpected observations lies in the heavy-tailed distributions of the tension rates. Therefore, there is a (relatively) high probability of extreme values and thus large variance, as the variance is sensitive to the square of the deviations from the mean per variable. By holding certain variables constant at their 80th percentile, extremes are eliminated and thus, the variance of tension rates is reduced and generally (uninventively) lower compared to the unconditionalised case.

The difference between the two conditionalising cases is most significant for the deep

locations 9 to 12 (see Figure B.5 in Section B.5). This is similar to the above described results, as the tension rate distributions at these locations seem to have a heavier tail.

Finally, comparing the conditionalising cases 1 and 2, differences in tension rates for the lower rope location r_{0-4} are marginal. Thus in this case, where H_S is found to have the most influence on the tension rates, it is sufficient to conditionalise only on H_S in order to analyse extreme tension rates under specific conditions.

Table 3.5: Tension rate values [kN/s] for 90th and 95th percentile for RefLoc 4 for three conditionalisation cases.

	$r0$	$r1$	$r2$	$r3$	$r4$	$r5$
$P(X \leq x_{90})$	4.22	3.31	3.23	3.85	4.30	4.62
$P(X \leq x_{90}\|H_S, T_P)$	3.95	2.33	2.63	2.61	3.02	3.27
$P(X \leq x_{90}\|H_S, T_P, \frac{\partial T_{r5}}{\partial t})$	4.01	2.31	2.64	2.87	2.96	-
$P(X \leq x_{95})$	7.42	5.08	5.00	5.80	6.35	6.67
$P(X \leq x_{95}\|H_S, T_P)$	5.19	2.58	3.00	3.06	3.47	3.71
$P(X \leq x_{95}\|H_S, T_P, \frac{\partial T_{r5}}{\partial t})$	5.19	2.53	2.98	3.05	3.03	-

3.5. DISCUSSION

Gaussian copula-based Bayesian Networks (GCBNs) were used in this study to model the multivariate dependence of hydrodynamic variables, i.e. wave and current variables, and tension rates at different positions in a mooring line. It helped to analyse extreme tension rates as an indicator of snap loads in a mooring line of a floating structure. This was done exemplarily for a seaweed cultivation structure for several different study locations across the south-eastern North Sea, each with different offshore conditions and water depths. The results presented here are, however, transferable to any other floating structure.

Some considerations potentially influencing the results are not addressed in this study. Even though study locations with different water depths are selected, the bathymetry at each location is not considered further. This might be relevant because according to Desiré *et al.* (2023), accounting for the bathymetry has a significant effect on snap loads. Furthermore, the present study assumes that waves and currents are consistently in-line with the structure. Mooring lines move in plane of the incoming wave and current field, see Figure 3.2. However, this way the complexity of the environment is not fully taken into account (Feng *et al.*, 2021; Hermawan and Furukawa, 2020). Additionally, in this study, one single mooring line is considered, while different tensions are expected for different mooring lines of a floating structure. For example, in aquaculture structures, an increase in biomass results in higher tensions in several lines, particularly in the downstream mooring lines, but reduces the maximum tensions in the upstream mooring lines (Feng *et al.*, 2021).

It was found that influences of hydrodynamic variables on the tensions or tension rates appear to be different for certain conditions. While Nguyen *et al.* (2019) observed in field experiments that currents significantly influence mooring line tensions in low to mid-tidal levels, this study determined the significant wave height to be the main driver for tensions. However, the average water depths of the field experiments conducted by

Nguyen *et al.* (2019) was 2.5 m. In the presented study, the depths varied between 12.6 to 44.7 m. The correlation between current magnitudes and tension rates was found to be mainly between -0.19 and 0.35, different from the correlation between significant wave heights and tension rates (0.81 to 0.98). Thus for snap loads, the current magnitudes can be neglected, as they are rather relevant for static pre-tensions than snap loads.

This study calculates mooring line tension rates using GCBNs, which account for the multivariate dependence between hydrodynamic variables and tension rates. These results are then compared to tension rates obtained under two conditionalising cases: one conditioned on the 80th percentile of significant wave height (H_S) and peak wave period (T_P), and the other conditioned on H_S and anchor tension rate. At the RefLoc 4, the results for the first conditionalising case are independent of T_P given H_S. Thus, in this specific case it is sufficient to conditionalise only on H_S. Therefore, computational efforts can be minimised given conditional independent variables for certain GCBN structures. Additionally, the results show that tension rates for the two conditionalising cases are lower compared to the unconditionalised case. This can be attributed to the heavy-tailed distributions of the tension rates, as there is a high probability of extreme values. Thus, conditionalising on certain variables leads to an "elimination" of extremes and therefore, both the 90th and 95th percentiles are lower compared to those of the unconditionalised case. Although this was observed in the presence of heavy-tailed tension rate distributions, the mooring line design should be based on a conservative case, considering most extreme tension rates. Uncertainties are reduced when information about specific hydrodynamic variables is available, and conditionalising on this reduces uncertainties leading to increased safety and reliability of the structures.

The structure chosen for this study is based on a pilot case at RefLoc 4. Consequently, as outlined in Section 3.2.3, the mooring line dimensions for other locations are scaled according to their respective water depths to ensure a consistent surge restoring stiffness across all locations (refer to Table 3.3). Thus, the results will differ in magnitude for different design parameters or even different moored floating structures, such as offshore wind turbines, vessels or floating solar panels. However, the characteristics are expected to be similar and the method can be applied to any moored floating structure. The method presented here is chosen for several reasons. It not only identifies which hydrodynamic variables significantly influence tension rates but also enables spatial analysis to compare variations across locations and estimate expected tension rates. Additionally, GCBNs offer computational efficiency in modelling the multivariate dependencies among variables, which Montes-Iturrizaga and Heredia-Zavoni (2016) identified as a crucial next step in the reliability assessment of mooring lines, as dependencies among variables often affect failure mechanisms. This has been further shown by using the GCBN to conditionalise on certain variables and analyse the resulting tension rates.

For the GCBN, the Gaussian copula is selected to model the multivariate distribution among variables. The selection of the Gaussian copula might not be the best chosen copula model. The results of Montes-Iturrizaga and Heredia-Zavoni (2016) show that selecting the Gaussian copula model can over- or underestimate the failure probability. In general, the results are sensitive to the choice of copula model. However, the results of the Cramér-von Mises statistic, which was applied as one out of two goodness-of-fit measures (see Section 3.4.1 and Section 3.4.3), show that the Gaussian copula is

a reasonable fit to model the dependence among the studied variables. Additionally, the d-calibration score was used to assess whether the GCBN properly represents the multivariate dependence structure of the selected variables, together with a pairwise correlation comparison (see Section 3.4.3). Both conclude that the individual GCBN per studied location are a suitable fit to model the multivariate dependence of studied variables. The presented GCBN enables to determine marginal distributions of tension rates influenced by multiple factors, or given specific conditions (see Section 3.4.5). Thus, GBCN's are employed to quantify the sensitivity of tension rates at different rope positions to certain hydrodynamic factors. Since empirical data is used as input to the GCBNs, the model can be updated given new observations to improve its accuracy. To account for different copula models among variable pairs and also to model tail dependencies such as those observed among tension rates, vine copula models (Joe, 2015; Morales-Nápoles *et al.*, 2023) are a potential next step. However, due to their complexity and fitting all possible structures of vine copula models to the data, they are limited in application to a small number of variables. Considering 13 variables, as done in this study via GCBNs, would make the investigation of all possible vine copulas for the data unfeasible.

3.6. CONCLUSION

Gaussian copula-based Bayesian Networks (GCBNs) have been used in this study to model the multivariate dependence and thus the influence of hydrodynamic variables on the tension rates in a mooring line of a floating structure. This has been done to gain more understanding of critical snap loads, which put the structure at risk.

The presented method allows for inference for unconditional and conditionalising cases. GCBNs rely on a single one-parameter copula, which is the Gaussian copula.

It has been applied at 12 study locations, conditionalising on variables highly correlated with tension rates, which have an effect on the marginal distributions of the tension rates. This approach is particularly of advantage when data for such variables are available, as it reduces uncertainty in predicting tension rates under specific conditions. The results show that conditionalising on harsh hydrodynamic conditions, the snap loads are counter-intuitively lower due to their heavy-tailed marginal distributions. This highlights the need to account for the dependence among hydrodynamic variables and tension rates, to improve structures' safety.

The current study establishes a framework for constructing a Bayesian Network that correlates sea state conditions with mooring line tension. The analysis was conducted under a simplified assumption that the floater moves with the wave particles. This assumption facilitates reproducibility and ensures that the analysis is not constrained by specific floater properties. In future work, we intend to extend the Bayesian Network to incorporate floater properties and their corresponding dynamic responses. These properties will serve as additional inputs to the network, enabling a more comprehensive and realistic representation of the system behaviour.

4

VINE COPULA MODELLING OF MIXED CONTINUOUS AND ORDINAL DATA IN OFFSHORE FLOATING SOLAR PLATFORMS

This chapter advances the multivariate dependence modelling introduced in the previous chapter by adopting vine copula models to analyse the structural response of an offshore floating platform under extreme wave conditions. In contrast to Gaussian copula-based Bayesian Networks, vines offer greater flexibility by allowing any bivariate copula as building blocks, enabling the representation of asymmetric and complex dependence patterns that are typical for interactions between environmental loads and structural responses. Additionally, vine copula models are implemented not just for continuous data, but also for mixed continuous and ordinal data, which frequently arise in engineering applications. Consequently, this chapter aims to answer the third sub-research question on how a vine copula model can represent the complex dependence among mixed continuous and ordinal structural and environmental variables of a floating offshore platform. Beyond characterising the dependence pattern, this chapter investigates the extent to which these patterns can be generalised across different structural configurations, how they vary between them, and what this implies for the transferability of probabilistic models for offshore floating platforms.

4.1. INTRODUCTION

The transition towards carbon-neutral energy remains crucial to mitigate climate change. While offshore wind farms are already an established technology, increasing spatial competition and environmental protection requirements, such as the Biodiversity Beyond National Jurisdiction (BBNJ) Agreement mandating 30% marine protection by 2030 (United Nations, 2023), limit further single-use developments. As a result, combining marine sectors through co-location and multi-use concepts is gaining attention. Among

these, floating photovoltaic (FPV) platforms have emerged as a promising complement to offshore wind, offering synergies such as reduced costs by shared operation and maintenance (Fan *et al.*, 2025; Maar *et al.*, 2023).

Environmental context of offshore floating photovoltaic systems

Floating photovoltaic (FPV) systems have been widely investigated with respect to their environmental implications, including associated benefits and potential drawbacks. While the present chapter focuses on structural response, the main environmental challenges and opportunities related to the installation and operation of (offshore) FPV systems are briefly summarised below to provide contextual background.

On the one hand, several uncertainties and potential negative impacts are associated with offshore FPV systems. The floating structures interact with both wind and waves and may alter local hydrodynamic conditions (Karpouzoglou *et al.*, 2020; Nalmpanti *et al.*, 2025). These changes can negatively affect marine species on a local scale (Benjamins *et al.*, 2024) and potentially influence marine ecosystems more broadly, for example by impacting corals and seagrass (Chayma *et al.*, 2024; Hooper *et al.*, 2021). Additional environmental effects may arise from increased biofouling caused by sea surface coverage (Hooper *et al.*, 2021; Nalmpanti *et al.*, 2025). Not least due to their visual impact on the environment, social acceptability remains a challenge and social impacts need to be investigated (Benjamins *et al.*, 2024; Hooper *et al.*, 2021).

On the other hand, offshore FPV also offer several potential benefits for both, environment and society. While sea surface coverage can promote biofouling, the associated reduction in light availability leads to limited primary production and algae growth, but it may also help preventing harmful algae blooms (Haas *et al.*, 2020; Karpouzoglou *et al.*, 2020; Sahu *et al.*, 2016), ultimately improving local water quality (Nalmpanti *et al.*, 2025; Sahu *et al.*, 2016). Additionally, evaporation losses are reduced, leading to water conservation (C.J. *et al.*, 2024b; Djalab *et al.*, 2024; Sahu *et al.*, 2016; Temiz and Javani, 2020). Moreover, such floating structures create artificial reefs and restore habitats (Hooper *et al.*, 2021; Nalmpanti *et al.*, 2025). Another benefit is the cooling effect of the water, which enhances energy efficiency (Chayma *et al.*, 2024; Djalab *et al.*, 2024; Dörenkämper *et al.*, 2021). These benefits are most pronounced for inland water bodies but are also relevant in marine environments. Overall, offshore FPV can contribute to achieving global carbon neutrality (Benjamins *et al.*, 2024; Fan *et al.*, 2025; Nalmpanti *et al.*, 2025). While additionally, different to land-based photovoltaic systems, FPV installations do not require land, a particular valuable advantage in densely populated coastal regions (Chayma *et al.*, 2024; Djalab *et al.*, 2024; Fan *et al.*, 2025; Temiz and Javani, 2020).

State of the art in offshore FPV structural analysis

Different from other floating offshore structures, such as wind turbines, the structural behaviour of FPV systems remain comparatively unexplored. Offshore FPV installations are typically deployed as floating platforms and have, to date, not progressed far beyond pilot stage. This is largely due to harsh environmental conditions at sea, which pose challenges to structural integrity and operational safety (e.g. Chayma *et al.*, 2024; C.J.

et al., 2024b; Djalab *et al.*, 2024; Fan *et al.*, 2025; Hooper *et al.*, 2021; Sahu *et al.*, 2016).

Recent literature has increasingly focused on evaluating the performance and environmental implications of FPV structures through field experiments, laboratory testing, and simulations. To date, field research on FPV has primarily focused on freshwater applications, particularly in lakes across Europe and Asia (Dörenkämper *et al.*, 2021; Karatas and Yilmaz, 2021; Kjeldstad *et al.*, 2022; Manoj Kumar *et al.*, 2022). However, according to (Hooper *et al.*, 2021), results from studies on FPV in freshwater cannot be directly applied to marine environments due to different conditions, while research in marine settings remains limited to pilot-scale projects, such as those funded by the European Union. Small scale experiments in a wave basin or flume were done by (Ji *et al.*, 2025) and (Sree *et al.*, 2022), respectively. In contrast to the limited number of field research or experimental studies, numerous simulations of FPV have been conducted. (Ikhennicheu *et al.*, 2021) chose an analytical approach for mooring design, because according to the authors, numerical modelling is rarely done in practice, despite them being more accurate. Other studies use commercial software (C.J. *et al.*, 2024a; Dörenkämper *et al.*, 2021; Ji *et al.*, 2025; Jifaturrohman *et al.*, 2024). However, according to (Manoj Kumar *et al.*, 2022), many of these tools are not correct as they do not account for e.g. fluid-structure interaction.

Due to the fluid-structure interaction, the behaviour of FPV structures is highly complex, particularly under harsh and uncertain offshore environmental conditions. A probabilistic approach can be beneficial in understanding the dependence between variables and understanding the influence of environmental variables on the structure. This helps assessing risks and improving the structures' reliability. Thus far, limited research has addressed FPV within a probabilistic framework. (Bru *et al.*, 2025) did a life cycle cost analysis of FPV via Monte Carlo simulations. Different to that, (Zhou *et al.*, 2022) predicted probabilistically the photovoltaic power based on a neural network and copula models, but indicate that a more complex method, such as vine copula models, might be able to better capture the dependence pattern of relevant variables.

Scope and objective

Given high complex interaction between uncertain offshore environmental conditions and structural response of a flexible offshore FPV structure, a probabilistic method based on copula theory has been chosen in this study to investigate the dependence between hydrodynamic and structural response variables.

Different probabilistic methods based on copula theory have previously been employed in engineering studies to assess loads or structural responses. For example, Torres-Alves *et al.* (2022) used a copula-based approach to assess expected structural responses and stresses of a submerged floating tunnel. Santjer *et al.* (2025) applied copula-based Bayesian Networks to assess tension rates in mooring lines of floating structures in the North Sea. A different study (Li and Zhang, 2020) made use of a vine copula model with 6 environmental variables to predict long-term design loads. In addition to that, Djalab *et al.* (2024) concluded in a review study on FPV that further research is needed on the risk assessment of marine FPV, while according to (Sahu *et al.*, 2016) further research is needed to understand the behaviour of the structure to improve the design of its anchoring system.

In this chapter, structural response due to offshore wave impact is investigated via variables such as longitudinal stress and shear forces for different structural configurations of an offshore FPV structure. The data are derived from a numerical FPV model that accounts for fluid-structure interaction (Alcañiz *et al.*, 2025), representing a potential offshore installation located in the German North Sea. A wide range of extreme sea states and structural characteristics are combined to capture the diversity of realistic offshore operating conditions. To analyse these interactions, a copula-based approach is chosen for multivariate dependence modelling among selected environmental and structural variables. Copula models allow to describe the dependence pattern independently from the marginal distributions of the individual variables. They have been successfully applied in a broad range of disciplines, including finance (e.g. Cherubini *et al.*, 2004; Genest *et al.*, 2009a; Kole *et al.*, 2007), healthcare (e.g. Cameron *et al.*, 2004; Candio *et al.*, 2021), and environmental and engineering sciences (e.g. Genest and Favre, 2007; Phillips *et al.*, 2022; Santjer *et al.*, 2024a; Santjer *et al.*, 2025), as well as in oceanographic contexts such as sea state modelling (e.g. Fazeres-Ferradosa *et al.*, 2019; Fazeres-Ferradosa *et al.*, 2018; Jäger *et al.*, 2019; Zhang *et al.*, 2015). However, in higher dimensions, copula models become less flexible, increasingly complex and less suitable for capturing asymmetries such as tail dependence. To overcome these limitations, the present study employs vine copula models, which construct higher-dimensional dependencies through a sequence of bivariate copulas as building blocks. In the FPV application considered here, some structural variables, such as the number of floaters, are ordinal variables, thus, vine copula models are employed for mixed continuous-ordinal data. Literature shows that several studies exist applying vines with ordinal data only, for instance in social sciences (e.g. Loaiza-Maya and Smith, 2019; Panagiotelis *et al.*, 2017) or healthcare (e.g. Panagiotelis *et al.*, 2012). Whereas mixed continuous and ordinal data in vine copula applications have been done in the fields of infrastructure and technology (e.g. Şahin and Joe, 2024; Zilko and Kurowicka, 2016) and biology and healthcare (e.g. Chang and Joe, 2019; Pan and Joe, 2022; Stöber *et al.*, 2015).

Building on these foundations, the present study applies vine copula modelling to describe and analyse the dependence patterns among key hydrodynamic and structural variables of offshore FPV systems, which, to the authors' knowledge, has not previously been addressed. The approach aims to identify whether generalised dependence pattern can be established across different structural configurations using a copula-based method, and to evaluate their predictive capability in reproducing and forecasting maximum longitudinal stress and shear forces, an analysis that goes beyond existing deterministic or configuration-specific studies.

The chapter is structured as follows. Section 4.2.3 outlines the probabilistic background, with a focus on copula modelling and the application of vine copulas to both continuous and mixed continuous–ordinal data. The study location and input data are introduced in Section 4.3.1, and the applied hydroelastic fluid–structure interaction model is described in Section 4.3.2. The results are presented in Section 4.4.5. First, hydrodynamic input data for the FPV model are prepared, and model outputs are analysed to provide an initial understanding of the structural response (Section 4.4.2). Subsequently, vine copula models are fitted and evaluated for various cases, as detailed in Section 4.4.3 to Section 4.4.5. The findings are discussed in Section 4.5, and conclusions are summarised in Section 4.6.

4.2. PROBABILISTIC MODELLING APPROACH

4.2.1. (BIVARIATE) COPULA MODELLING

Understanding the dependence pattern between random variables is a fundamental task in multivariate statistical analysis. Copula models are a popular and powerful tool for this purpose, as they describe the dependence between variables separated from the marginal distributions.

The concept of copula models relies on Sklar's theorem (Sklar, 1959) which states that for a d-variate joint distribution F_X, a set of univariate marginal distribution functions F_{X_i}, $i = 1,\dots,d$, in $[0,1]$ and a d-dimensional random vector $X = (X_1,\dots,X_d)$, the joint distribution function can be expressed as (Czado, 2019; Joe, 2015):

$$F_{\mathbf{X}}(x_1,\dots,x_d) = C(F_{X_1}(x_1),\dots,F_{X_d}(x_d)) \,, \quad x_i \in \mathbb{R} \tag{4.1}$$

Where $C : [0,1]^d \to [0,1]$ is a d-dimensional copula, which is unique for absolutely continuous distributions. Due to copulas' complexity in higher dimension, they are predominantly applied for the bivariate case.

While the theoretical definition of copulas and many of their applications traditionally assume continuous random variables, copula models can also be extended to accommodate ordinal variables, which are frequently encountered in practical settings. The main disadvantage, however, is the lack of uniqueness. In that case, the copula is unique only on the set (Genest and Nešlehová, 2007; Joe, 2015)

$$\text{Range}\,(F_{X_1}) \times \ldots \times \text{Range}\,(F_{X_d}) \tag{4.2}$$

Where the marginal distributions of ordinal (discrete) variables F_{X_i} have discontinuities, their inverse have plateaus. This also means that the Fréchet-Hoeffding upper and lower bounds are not necessarily perfect monotone functional dependent. In addition, while the correlation measures of Kendall's tau (τ) and Spearman's rho (ρ) for the copula C are identical with τ and ρ of the random variables X_1 and X_2, this is different in the ordinal case, as τ and ρ do not account for ties, i.e. pairs of observations with identical values. This results in a less intuitive interpretation and thus, estimation of copula parameter. Because these two correlation measures are margin-free measures of dependence, they can be used with ordinal variables to estimate the dependence pattern (i.e. copula parameter), particularly when the copula belongs to a specific parametric family (Genest and Nešlehová, 2007).

Kendall's τ_b, an adjusted version of Kendall's τ, is less affected by ties than Spearman's ρ. While τ_b evaluates pairwise concordance and discordance, Spearman's ρ is more sensitive to rank distances. τ_b is given by

$$\tau_b = \frac{P - Q}{\sqrt{(P + Q + T_X)(P + Q + T_Y)}} \tag{4.3}$$

Here, P is the number of concordant pairs and Q is the number of discordant pairs. Suppose there are two observations (X_1, Y_1) and (X_2, Y_2), then they are concordant if $X_1 < X_2$ and $Y_1 < Y_2$, or if $X_1 > X_2$ and $Y_1 > Y_2$, and they are discordant if $X_1 < X_2$ and

$Y_1 > Y_2$, or if $X_1 > X_2$ or $Y_1 < Y_2$. T_X and T_Y are defined as the numbers of ties only in variable X and Y, respectively (Kendall, 1945).

Despite the above mentioned challenges, the use of copula models in the context of discrete distributions is a valid approach and has been applied in numerous studies, such as healthcare (e.g. Klein *et al.*, 2019; Nikoloulopoulos and Karlis, 2008; Varin and Czado, 2009), or socio-economic studies (e.g. Hoff, 2007; Yang *et al.*, 2020).

4.2.2. VINE COPULA MODELLING FOR CONTINUOUS VARIABLES

As copula models are getting very complex in more than two dimensions, different methods can be applied to model multivariate dependence of variables. This study makes use of vine copula models, which employ bivariate copula models as building blocks. The advantage of this method is that the bivariate copulas can be selected independently and thus capture tail asymmetries and represent the multivariate dependence by appropriate conditioning (Aas *et al.*, 2009). In other words, vine copula models decompose complex multivariate dependencies into a structured sequence of simpler relationships, beginning with dependence between variable pairs and progressively incorporating conditional dependencies through hierarchical levels. This is achieved by expressing the joint density as a product of conditional densities using bivariate copulas (Czado, 2019). Relying on Sklar's Theorem (Equation (4.1)), the joint density of a d-dimensional distribution can be factorised into the product of its univariate marginal densities and a sequence of bivariate (conditional) copula densities, which serve as the building blocks of the vine copula model:

$$f_{1,\dots,d}(x_1,\dots,x_d) = f_{1|2,\dots,d}(x_1|x_2,\dots,x_d) \cdot f_{2|3,\dots,d}(x_2|x_3,\dots,x_d) \cdot \dots \cdot f_d(x_d) \tag{4.4}$$

Vine copulas are based on a graphical representation known as a vine tree hierarchy (Czado and Nagler, 2022a), which consists of a sequence of trees where each tree contains nodes $N = 1,2,\dots,d$ and edges E. The edges are a subset of node pairs, which form an acyclic graph. A vine is called regular vine, or R-vine, if two edges, which share a common node, become a node in the following tree (Bedford and Cooke, 2001; Bedford and Cooke, 2002). This is called the proximity condition. A R-vine can have different types of hierarchies, two of them are canonical vines (C-vines) and drawable vines (D-vine). In total, a d-dimensional R-vine can have $\frac{d!}{2} \times 2^{\binom{d-2}{2}}$ different configurations (Morales-Nápoles, 2010a; Morales-Nápoles, 2010b).

Definition 1. *(Regular (R-) vine tree sequence) The set of trees $\mathcal{V} = (T_1,\dots,T_{d-1})$ is a regular vine tree sequence on d elements if:*

1. *Each tree $T_j = (N_j, E_j)$ is connected, i.e. for all nodes $a, b \in T_j$, $j = 1,\dots,d-1$, there exists a path $n_1,\dots,n_k \subset N_j$ with $a = n_1$, $b = n_k$.*

2. *T_1 is a tree with node set $N_1 = \{1,\dots,d\}$ and edge set E_1.*

3. *For $j \geq 2$, T_j is a tree with node set $N_j = E_{j-1}$ and edge set E_j.*

4. *(Proximity condition) For $j = 2, \dots, d-1$ and $\{a, b\} \in E_j$ it must hold that $|a \cap b| = 1$, i.e., for nodes connected by an edge in T_j, the corresponding edges in T_{j-1} share a common node.*

Applying recursion on Equation (4.4), the joint density for a R-vine in the continuous case can be written as (Bedford and Cooke, 2001; Bedford and Cooke, 2002):

$$f(x) = \prod_{i=1}^{d-1} \prod_{e \in E_i} c_{e_a, e_b; D_e}(F_{e_a|D_e}(x_{e_a}|x_{D_e}), F_{e_b|D_e}(x_{e_b}|x_{D_e})|x_{D_e}) \prod_{i=1}^{d} f_i(x_i) \tag{4.5}$$

where e_a and e_b are two conditioned variables of the edge $e = \{e_a, e_b, D_e\}$ and D_e is a conditioning set of variables. Note that in practise, a simplifying assumption is made such that the pair copulas do not directly depend on the values of variables in the conditioning set (i.e. they do not depend on x_{D_e}).

The conditional distribution functions can be evaluated according to (Joe, 1996) as:

$$F_{X|Y_j, \mathbf{Y}_{-j}}(x_j|y_j, \mathbf{y}_{-j}) = \frac{\partial C_{X, Y_j|\mathbf{Y}_{-j}}(F_{X|\mathbf{Y}_{-j}}(x|\mathbf{y}_{-j}), F_{Y_j|\mathbf{Y}_{-j}}(y_j|\mathbf{y}_{-j}))}{\partial F_{Y_j|\mathbf{Y}_{-j}}(y_j|\mathbf{y}_{-j})} \tag{4.6}$$

where X is a random variable, $\mathbf{Y}$ is a random vector and Y_j is any scalar element of $\mathbf{Y}$. $\mathbf{Y}_{-j}$ is its complement with Y_j not an element of $\mathbf{Y}$.

To apply vine copula models in practice, an appropriate configuration needs to be determined first. As mentioned above, the number of possible configurations of a R-vine depends on d. As the number of variables d increases, the number of possible vine configurations grows rapidly, making exhaustive comparison infeasible already for about 7–8 variables. To allow for a feasible configuration selection, the search algorithm of the maximum spanning tree (MST), also called sequential method, is commonly applied (Dißmann *et al.*, 2013). This algorithm employs the MST for each tree, i.e. it selects a tree configuration which maximises the sum of pairwise dependence (absolute Kendall's τ). This is to ensure that the strongest dependencies among variables are captured. Once a potential configuration is determined, copula families can be selected for each tree based on the Akaike Information Criterion (AIC) (Akaike, 1974) and their parameters can be estimated. The AIC value is commonly used to assess the fit of bivariate copula models and vine copula models, where k denotes the number of estimated parameters and $\hat{L}$ corresponds to the maximised value of the model's likelihood function:

$$AIC = 2k - 2\, ln(\hat{L})\,. \tag{4.7}$$

For further information on the concept of vine copula models, the reader is referred to (Czado, 2019; Czado and Nagler, 2022b; Joe, 2015). Note that in the remainder of this study, R-vines are referred to as vines.

4.2.3. VINE COPULA MODELLING FOR MIXED CONTINUOUS AND ORDINAL VARIABLES

As described in Section 4.1, random variables are not always continuous, but can be ordinal. In that case, where vine copula models are generally also applicable, the probability mass function (pmf) for the ordinal case can be decomposed in a similar way as the pdf for the continuous case in Equation (4.4):

$$\begin{aligned} Pr(Y_1 = y_1, \dots, Y_d = y_d) \;=\; & Pr(Y_1 = y_1 | Y_2 = y_2, \dots, Y_d = y_d) \\ & \times Pr(Y_2 = y_2 | Y_3 = y_3, \dots, Y_d = y_d) \times \dots \\ & \times Pr(Y_d = y_d) \end{aligned} \tag{4.8}$$

Equation (4.8) allows to express the entire multivariate pmf in terms of bivariate copula functions and marginal distributions. Further, the conditional distribution functions in the ordinal case satisfy the following recursion (analogue to Equation (4.6) in the continuous case):

$$\begin{aligned} F_{X|Y_j,\mathbf{Y}_{-j}}(x_j|y_j,\mathbf{y}_{-j}) = [C_{X,Y_j|\mathbf{Y}_{-j}}(F_{X|\mathbf{Y}_{-j}}(x|\mathbf{y}_{-j}), F_{Y_j|\mathbf{Y}_{-j}}(y_j|\mathbf{y}_{-j})) \\ - C_{X,Y_j|\mathbf{Y}_{-j}}(F_{X|\mathbf{Y}_{-j}}(x|\mathbf{y}_{-j}), F_{Y_j|\mathbf{Y}_{-j}}(y_j - 1|\mathbf{y}_{-j}))] \;/\; Pr(Y_j = y_j|\mathbf{Y}_{-j} = \mathbf{y}_{-j}) \end{aligned} \tag{4.9}$$

For simplicity and without loss of generality, we assume $\mathbf{Y} \in \mathbb{N}^d$ here. $y_j - 1$ indicates the left limit of the cumulative distribution function (CDF) of the ordinal variable Y_j.

For further details on vine modelling in the case of ordinal data, the reader is referred to (Panagiotelis *et al.*, 2012).

4.3. STUDY SETUP

Following the description of the probabilistic foundations, this section presents the overall study setup. The corresponding workflow is illustrated in Figure 4.1.

This study focuses on a floating photovoltaic (FPV) system located in the south-eastern North Sea. For this purpose, the research platform FINO3 has been selected, due to its exposure to offshore conditions with a water depth of approximately 23 m. The site is surrounded by offshore wind farms, making it particularly relevant for potential multi-use applications.

Data on significant wave height (H_S) and peak period (T_P) for the years 1990–2020 are obtained from a large-scale three-dimensional hydrodynamic model. Extreme sea states are subsequently selected from this dataset. Using a bivariate copula approach, 300 samples are generated to serve as input for the FPV model, as further described in Section 4.3.1.

The hydroelastic model to determine the response of the FPV platform is described in Section 4.3.2. Particularly, maximum longitudinal stress and shear forces are analysed probabilistically, using vine copula models as described in Section 4.3.3.

4.3.1. LOCATION SELECTION AND INPUT DATA

For the analyses of the FPV study, the FINO3 research platform is chosen (see top-left of Figure 4.1). It is located 80 km off the German coast in a water depth of around 23 m, experiencing harsh environmental conditions with a significant wave height of a 50-year return period of 9.2 m. This location has been chosen because it is located close to offshore wind parks and has been chosen as a reference location in Santjer *et al.* (2024a) and Santjer *et al.* (2025).

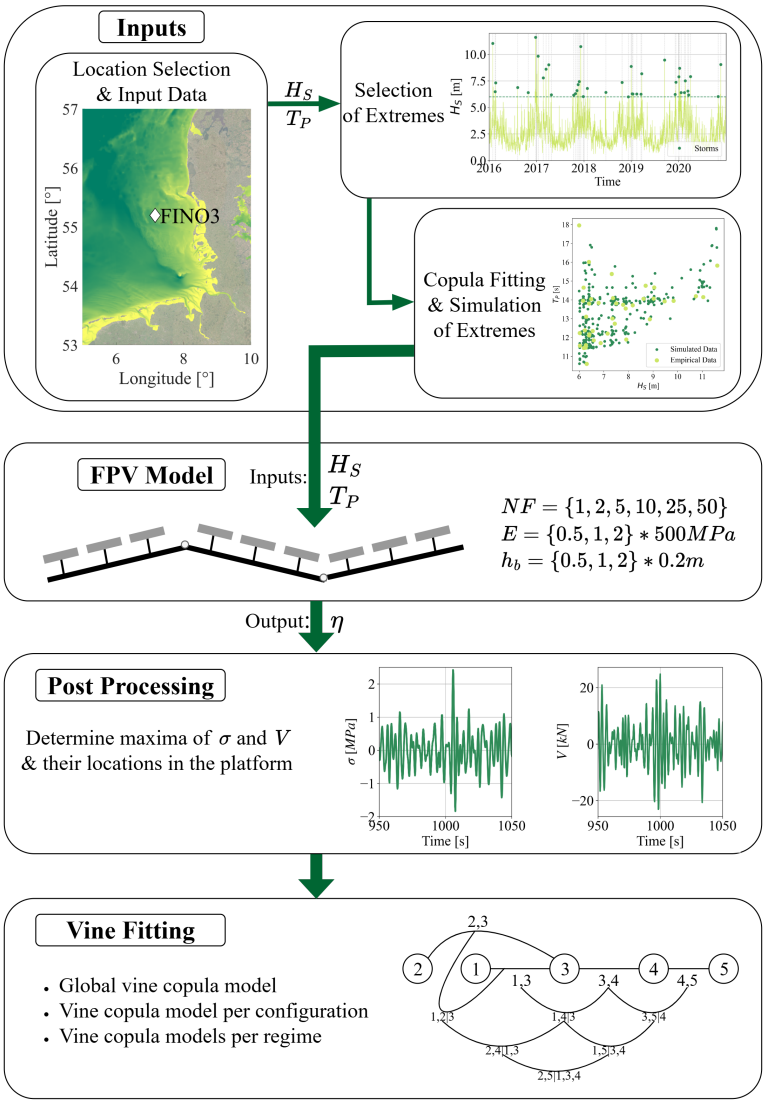


Figure 4.1: Detailed framework of the methodology applied in this study.

Data at this location for significant wave height H_S and wave period T_P used for the present study comes from the Dutch Continental Shelf Model (DCSM) (Zijl *et al.*, 2023) for the years 1990 to 2020 in an hourly resolution. The DCSM uses the D-Flow FM component of the Delft 3D Flexible Mesh Suite (Deltares, 2023a) and covers the northwester European continental shelf. Since the FPV will be analysed for extreme cases, these need to be identified in the available data. Therefore, thresholds and declustering time need to be selected, to identify extremes of H_S and corresponding T_P. Declustering times of 12, 24, 48, and 72 hours were selected for analysis, as storm durations in comparable North Sea conditions, such as those reported for the southwestern Dutch coast by Caspers *et al.* (2025), typically range between 37 and 78 hours. At first, an appropriate threshold h needs to be selected for H_S. For each considered declustering time, a Generalised Pareto distribution (GPD) is fitted to peaks above different threshold and a parameter stability analysis is conducted (Coles, 2001; Davison and Smith, 1990). This means that a threshold h needs to be determined up to which the GPD is valid, i.e. its shape parameter ξ and modified scale parameter σ^* remain constant. Commonly, the modified scale parameter σ^* is used, because it helps normalise the scale parameter σ relative to the threshold h, making it easier to detect true stability (Coles, 2001). It is defined as

$$\sigma^* = \sigma - \xi\, h \tag{4.10}$$

Next, an appropriate declustering time needs to be selected. Therefore, a Poisson distribution is fitted to the number of selected extremes per year, to assess whether the identified extremes are i.i.d., i.e. whether they follow a single Poisson distribution with a constant parameter Λ. This is defined as our null hypothesis H_0 for a likelihood ratio test (LRT), to assess which declustering time is the best choice (Davison and Smith, 1990; Neyman and Pearson, 1928). If each year, however, has its own rate parameter $\lambda_{p,i}$, thus the number of selected extremes do not follow a single distribution, this constitutes the alternative hypothesis H_1. For both, the null hypothesis H_0 and the alternative hypothesis H_1, the corresponding log-likehoods, L_0 and L_1, are computed. The statistic of the LRT is then defined as

$$\Lambda = -2\, log\, (\frac{L_0}{L_1}) \tag{4.11}$$

The smaller Λ is, the better does H_0 fit the number of selected extremes per year, than H_1. Once an appropriate threshold and declustering time are determined, extreme H_S and their corresponding T_P are selected to continue with the fitting of a copula model using PYVINECOPULIB package in PYTHON (Nagler and Vatter, 2025). All parametric copula models implemented in the package are considered and the model yielding the lowest AIC (see Equation (4.7)) is selected as the best fit. Based on the selected copula, 300 samples are simulated, a sample size shown to be sufficient by Li and Zhang (2020), applied in Santjer *et al.* (2025), and still computationally feasible for subsequent simulations of the FPV model in the present study.

4.3.2. HYDROELASTIC MODEL FOR FPV

The case study is done for a hypothetical design of a floating photovoltaic (FPV) platform. The total length of the platform is $L_p = 100\,\text{m}$. The total platform is constructed using NF number of floaters, each of length $L_f = L_p/NF$, and connected using free hinges. Each floater is assumed to have a constant thickness h_b, and is constructed using high-density polyethylene (HDPE), with density $\rho_{HDPE} = 952\,\text{kg m}^{-3}$. The study simulates various scenarios for the FPV design, by varying the number of floating $NF \in \{1,2,5,10,25,50\}$, thickness of floater $h_b \in \{0.5,1,2\} * 0.2\,\text{m}$ and the Young's modulus $E \in \{0.5,1,2\} * 500\,\text{MPa}$. Figure 4.1 shows a schematic of the FPV platform and its properties.

The FPV platform is floating in a fluid domain with still-water depth $d = 22.8\,\text{m}$. The platform is exposed to a sea-state defined by a JONSWAP spectrum with significant wave-height H_S and peak time-period T_P, based on the extreme value analysis described in Section 4.3.1. For each state, the response of the FPV platform is computed for a duration of 1200 s at a time-resolution of $\Delta t = 0.25\text{s}$. A detailed description of the procedure for computing the response of the FPV platform for a input sea-state is given in Alcañiz *et al.*, 2025, however, a brief description is given in this section.

The structural analysis of the FPV platform is performed using the hydroelastic model described in Alcañiz *et al.*, 2025. This model is a monolithic finite element (FE) formulation for fluid–structure interaction, where the structure is represented by an Euler–Bernoulli beam and the fluid is modelled using potential flow theory with a linearised free-surface boundary condition. The framework is capable of handling both two- and three-dimensional problems and can be applied in either the time domain or the frequency domain. The present study, however, is restricted to a two-dimensional frequency-domain analysis, in which the FPV platform is represented by a one-dimensional beam equation and the fluid is governed by a two-dimensional potential-flow formulation with a linearised free-surface boundary condition.

A configuration of the FPV platform is defined based on the choice of $\{NF, h_b, E, d\}$. For each configuration, the FE model results in a complex-valued response-amplitude operator $\text{RAO}(\mathbf{x},\omega)$ for each point $\mathbf{x}$ on the FPV platform and input ocean wave frequency ω. Here, the RAO is computed for a wide range of ocean-wave frequencies $\omega \in [0.2, 6]\,\text{rad s}^{-1}$.

Using the RAO for a given FPV platform configuration, the transverse deflection of the platform under an incoming sea state characterised by H_S, T_P is obtained using the following formulation for the wave elevation:

$$\eta(\mathbf{x},t) = \sum_{j=1}^{j=n} \text{RAO}(\mathbf{x},\omega_j)\sqrt{2S(\omega_j)\Delta\omega}\exp\left(i\left(k_j\mathbf{x} - \omega_j t + \alpha_j\right)\right). \tag{4.12}$$

Here $S(\omega)\,[\text{m}^2\,\text{s}]$ is the power spectral density function for the JONSWAP spectrum, which is a measure of the intensity of each wave frequency in the sea state. Additionally, k_j is the wave-number for frequency ω_j, calculated using the dispersion relationship, while α_j is the uniformly randomised initial phase.

Once the transverse deflection is obtained, the bending moment BM, longitudinal stress σ and shear force V at each point in the platform is obtained using the following equations,

$$BM = EI\,\eta_{xx}, \tag{4.13}$$

$$\sigma = BM\frac{y_c}{I}, \tag{4.14}$$

$$V = EI\,\eta_{xxx}. \tag{4.15}$$

where $I = \frac{1}{12}h_b^3$ is the cross-section Moment of Inertia per unit width and $y_c = h_b/2$ is the normal distance between the rotation axis for bending and the location on the cross-section.

Therefore, using this procedure, we compute the transverse deflection and subsequently the stresses in the structure for each FPV configuration exposed to each extreme sea-state. The maxima of σ and V, together with their location in the platform, are determined to build the dataset for the probabilistic analysis. Finally, for each FPV configuration, we define a characteristic hydroelastic wavelength λ_c as

$$\lambda_c = 2\pi\left(\frac{EI}{\rho g}\right)^{1/4}. \tag{4.16}$$

The ratio λ_c/λ_j, where λ_j is the wavelength of the incoming ocean wave with frequency ω_j, serves as a useful indicator for categorising the elastic behaviour of the structure.

4.3.3. VINE FITTING APPROACH

In this study, vine copula modelling is applied to analyse how hydrodynamic and structural variables of offshore FPV systems depend on each other. The aim is to identify potential generalised dependence structures across different configurations and to assess their predictive skill in capturing and forecasting maximum longitudinal stress and shear forces. The vine fitting is carried out using parametric copulas as implemented in the PYTHON package PYVINECOPULIB (Nagler and Vatter, 2025). Parametric copulas are chosen because they are computationally efficient, capable of capturing complex dependence features such as tail dependence, and relatively robust to small sample sizes. They also tend to be easier to interpret and less prone to overfitting compared to fully non-parametric approaches.

Most variables considered in this study are continuous, yet some are treated as ordinal due to their discretised sampling. While E and h_b are physically continuous quantities, in the dataset they take only a limited set of values (three values each) while NF, by contrast, is discrete and takes six categories here. This is the reason to treat them as ordinal variables in the vine fitting process, resulting in a mixed continuous-ordinal data setup. The PYVINECOPULIB package supports such mixed data, but requires specification of both the left and right limits of the empirical CDF for ordinal variables to properly account for discontinuities in their distributions (see Section 4.2.3).

Vine copula models are fitted according to the sequential method (described in Section 4.2.2) for different sets of structural configurations to identify potential generalised dependence structures and to assess whether certain configurations can be grouped based on similar dependence behaviour. These generalised vines are then compared with

vines fitted individually to datasets for each number of floaters (NF). Model performance is evaluated by comparing the AIC values (see Section 4.2.2) of the generalised vines to those obtained for vines per NF case. This comparison is motivated by the assumption that NF acts as the dominant structural variable influencing the overall system response. The scatter plots shown in Section 4.4.2 (Figure 4.3) and Figure C.3 (after transformation to the uniform scale) reveal distinct clusters, which are likely associated with the different NF configurations.

4.4. RESULTS

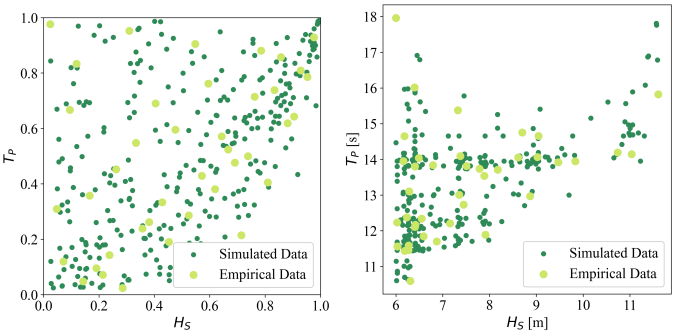
4.4.1. SELECTION OF EXTREMES AND PREPARATION OF INPUT DATA

First, input data for the FPV model need to be prepared by identifying extremes of H_S in the dataset for the FINO3 location. Thresholds are determined via parameter stability analysis of Generalised Pareto distribution (GPD) and an appropriate declustering time is determined via the Poisson distribution (see Section 4.3.1). Suitable thresholds are determined and compared for different declustering times of 12, 24, 48 and 72 hours. The results are exemplarily shown for a declustering time of 72 hours in Figure C.1a in C.1. It can be seen that the two parameters, the shape parameter ξ and modified scale parameter σ^*, have a plateau in H_S between around 2 to 6 m. This is similar for all investigated declustering times, therefore, the threshold is set to 6 m, as we are interested in the extremes.

A suitable declustering time is chosen via the fitting of a Poisson distribution, to assess whether identified extremes are i.i.d.. Again, this is done for all 4 selected declustering times with a fixed threshold of 6 m. The estimated parameter of the Poisson distribution, $\hat{\lambda_p}$ is set to the mean of the number of identified extremes per year. The cumulative distribution function (CDF) of the Poisson distribution is compared with the empirical CDF and shown exemplarily for a declustering time of 72 hours in Figure C.1b. Additionally, a p-p-plot, where values of the CDF are plotted against the values of the empirical CDF, is shown in Figure C.1c. The Poisson distribution seems to be a good fit. To assess, which of the considered declustering times is the best, a LRT is conducted (see Section 4.3.1). The results are presented in Table C.1. A declustering time of 72 hours appears to have the smallest Λ and the highest p-value. Thus, there is no strong evidence that λ_p varies over time and therefore, H_0 cannot be rejected. With a determined threshold of 6 m and a declustering time of 72 hours, this results in 44 identified extremes in the dataset spanning 30 years (see Figure C.1).

In a next step, a bivariate copula model is fitted to this data of extreme H_S and corresponding T_P. Therefore, the samples first need to be transformed to uniform space, see Figure 4.2a. All parametric copula models implemented in the PYTHON package PYVINECOPULIB (Nagler and Vatter, 2025) are fitted and the model with the lowest AIC value (see Equation (4.7)) is selected as the best fit. This results in the Tawn copula with the asymmetry parameter of 1, the scaling or boundary parameter of 0.41 and the parameter for dependence strength of 3.5. The Tawn copula is based on an asymmetric transformation of the Gumbel copula (Joe, 2015; Tawn, 1988). It should be noted, however, that according to Jaeger and Morales-Nápoles (2017), the Tawn copula does not impose limitations on thresholds relevant to wave breaking. Nonetheless, alternative

copula models such as the gamma 1-factor model and the skew-t copula, which may offer improved flexibility in this context, are currently not available in the PYVINECOPULIB implementation. Despite that, the Tawn copula is found to be a suitable copula model to capture the dependence pattern of extreme H_S and corresponding T_P, therefore, 300 samples are drawn from the fitted Tawn copula and transformed to physical space using the empirical CDF of the selected data of H_S and T_P (see Figure 4.2b). Finally, this simulated data represent extreme conditions of H_S and T_P at the location study and serves as input data for the FPV model.



(a) Uniform space. **(b)** Transformed to physical units.

Figure 4.2: Empirical and simulated data of H_S and T_P for a threshold of 6 m and a declustering time of 12 hours.

4.4.2. HYDRODYNAMIC AND STRUCTURAL DATA OVERVIEW

As described in Section 4.3.2, the main output of the FPV model is the second spatial derivative of the deflection, η_{xx}, which represents the curvature of the structure. By differentiating this quantity once more, the third spatial derivative, η_{xxx}, is obtained, corresponding to the rate of change of curvature along the platform (i.e., the structural slope gradient). From η_{xx}, the bending moment (BM) is calculated (Equation (4.13)), which is then used to determine the longitudinal stress σ (Equation (4.14)), while the shear force V is calculated from η_{xxx} (Equation (4.15)). Both σ and V are the main variables of interest describing the structural response of the floating platform, and therefore used for further analyses. Their respective time series are illustrated in Figure C.2 in C.3. Note that, for the subsequent analysis, the maxima of σ and V are extracted from the structural response to a sea state characterised by H_S and T_P and are therefore not temporally coincident with H_S and T_P.

An overview of all simulation results is shown in Figure 4.3. The lower-left section of the figure presents scatter plots for all considered cases (i.e., for varying NF, E, and h_b) and all 300 combinations of H_S and T_P, while the upper-right section shows the corresponding Kendall's τ values of the variables in physical units. The diagonal contains the empirical marginal distributions of each variable. Due to the discrete sampling of E, h_b (and thus the Moment of Inertia I) and NF (with 3, 3, and 6 values, respectively), these variables exhibit a discrete-like appearance in both the scatter and density plots.

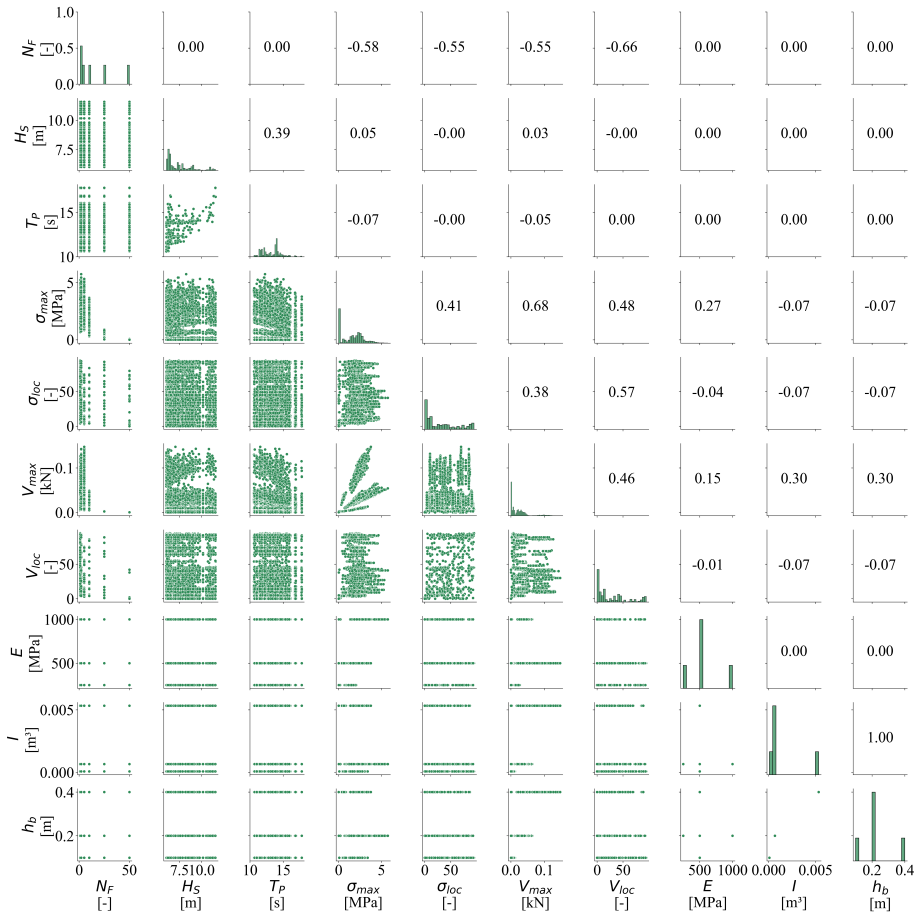


Figure 4.3: Pairwise scatter plots of the studied variables in their physical units are presented in the lower-left part, together with their empirical marginal distributions on the diagonal and Kendall's τ correlations, computed from the raw data in physical space, above the diagonal.

The strongest correlations are observed between NF and the maxima and locations of both stress and shear (σ_{max} and V_{max}). Furthermore, the latter four variables also exhibit strong correlations among themselves. Since I is directly derived from h_b, their correlation equals 1. Thus, I can be excluded from further analysis, as h_b already captures its influence in the FPV model settings, also, because h_b is one of the configuration variables changing for each run of the FPV model. Among the structural design variables, h_b (and therefore I) shows the strongest correlations with V_{max}, while E primarily correlates with σ_{max}, as expected from how V and σ are calculated (see Equation (4.15) and Equation (4.14)). Additionally, distinct modes are visible in the scatter plots, particularly for the variable pair of V_{max} and σ_{max}, which can be attributed to the different combinations of E and h_b in the FPV model configurations. The empirical marginal distributions on the

diagonal show that both σ_{max} and V_{max} contain many small values, which likely occur for larger NF, due to more rigid floaters. The location of maximum stress and shear (σ_{loc} and V_{loc}, respectively) predominantly occur in the frontal region of the platform, where waves impact the structure first.

Since copulas operate in the unit hypercube $[0,1]^2$, each variable is transformed to uniform. This allows to better see the dependence pattern among variables independent of their empirical margins. The resulting data are shown in Figure C.3 in C.3. On the diagonal, the uniform marginals are displayed, while the lower-left part shows scatter plots and the upper-right part the corresponding densities. Here, the discrete-like appearance of some variables is also visible, particularly emphasised in the density plots. Analysing the dependence pattern, it can be seen that a lower-tail dependence appears between the maxima of shear V_{max} and stress σ_{max}, suggesting that low values tend to co-occur, even though multiple modes are obvious again. Additionally, some tail dependence is observed between NF and σ_{max} and V_{max}. The asymmetric dependence pattern between H_S and T_P, already discussed in Section 4.4.1, also becomes evident here. Overall, H_S and T_P show comparatively weak correlations with the maxima and locations of σ and V, as reflected by the low Kendall's τ values in Figure 4.3.

4.4.3. GLOBAL VINE MODEL

As a next step, a single global vine was fitted with the objective of assessing whether the dependence pattern remains the same across FPV configurations with different numbers of floaters NF. The dataset for fitting a global vine comprised the following variables: H_S, T_P, σ_{max}, σ_{loc}, V_{max} and V_{loc}, NF, E and h_b. The latter three variables are considered ordinal in the vine fitting process due to their possible values (3 categories for E and h_b, 6 categories for NF), although E and h_b are physically continuous. This results vine fitting to mixed continuous and ordinal data (see Section 4.3.3).

The following preprocessing was applied: for each fixed NF case, the variables were transformed to uniform margins and then compiled to form a single dataset for the global vine fit. This allows to investigate whether there is a common dependence pattern for each NF case. Therefore, the performance of the global vine was compared to separate vine models fitted per NF case. The models are compared using the AIC (see Equation (4.7)), where lower AIC indicates a better fit. In the graphical comparison in Figure 4.4, smaller difference between the AIC values of the global vine (•) to those of the case-specific AICs (×) indicate a better approximation of the global model.

The comparison shows that the global vine provides a better fit to the vines for lower NF values, whereas the differences increase for large number of floaters (e.g., $NF = 25$ or 50). Thus, while a single global vine might be able to reasonably approximate the dependence pattern for some configurations, it does not achieve full suitability across all NF cases. This suggests that vine fitting for individual cases is necessary, as a global vine does not provide a sufficiently good fit.

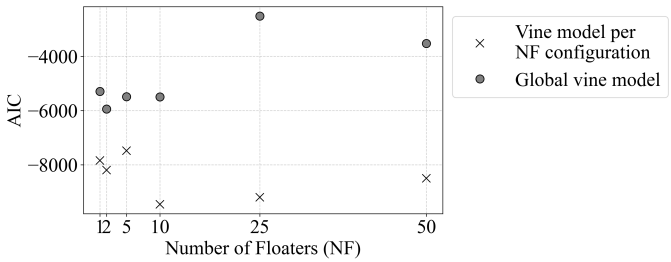


Figure 4.4: Comparison of AIC values for the global vine model (•) and the case-specific vines (×) across different NF configurations.

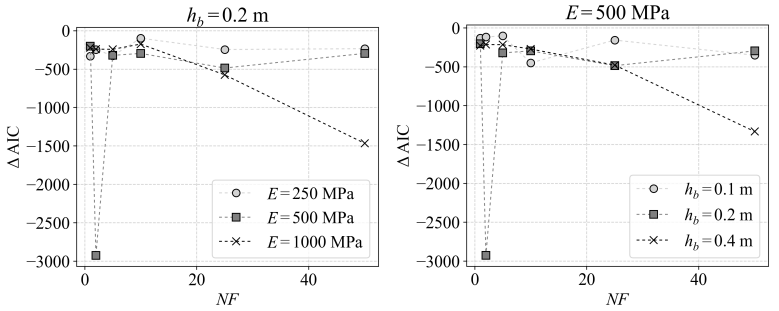
4.4.4. VINE MODELS PER CONFIGURATION

As shown previously, a single global vine copula model does not adequately represent the dependence pattern across all configurations. Therefore, in this section, a separate vine is fitted for each configuration, defined by the combination of Young's modulus E and floater thickness h_b. Similar to the approach presented in Section 4.4.3, one vine is first fitted across all NF cases and then compared with vines fitted individually per NF. The objective is to evaluate whether, for a given configuration of E and h_b, a single vine can sufficiently describe the dependence pattern of the floating system, independent of NF. Consequently, only the continuous variables H_S, T_P, σ_{max}, σ_{loc}, V_{max} and V_{loc} are included in the fitting. Following the same procedure as for the global vine, the data of these variables are transformed to the uniform per NF and subsequently combined to fit one vine per E and h_b configuration. This approach makes it possible to assess whether the dependence pattern, independent of the marginal distributions, can be generalised for a given configuration.

Figure 4.5 shows the results for constant floater thickness $h_b = 0.2$ m (*a*) and constant Young's modulus $E = 500$ MPa (*b*). Light grey circles (∘) indicate the more elastic structures, while crosses (×) represent the more rigid ones. Unlike in previous section, where absolute AIC values were shown, here the differences in AIC (ΔAIC) are presented to allow for a clearer comparison between the configuration-specific and the NF-specific vines. The closer ΔAIC is to 0, the better the configuration-wise vine represents the corresponding NF case. Although using ΔAIC rather than absolute AIC values can slightly change the interpretation, it remains a widely accepted, non-relativised measure of model fit and allows consistent comparison across cases. It is trivial that all ΔAIC values are below zero, indicating that the NF-specific vines fit their data better than the more generalised configuration vines.

For small NF (i.e. long floaters), ΔAIC values are relatively close to zero across all E and h_b configurations. This suggests that the underlying dependence pattern are similar across a configuration case and may be generalisable. An exception is observed for the reference case ($E = 500$ MPa, $h_b = 0.2$ m) for $NF = 2$, where the floater length is about 50 m and the dominant wavelength roughly 200 m. This configuration is neither particularly rigid nor particularly elastic, and its dependence pattern differs notably from the other analysed cases. Moreover, the rigid configurations (marked by ×) show an increasingly distinct dependence pattern as the floaters become shorter (i.e. as NF

increases), whereas for the elastic and reference cases, ΔAIC does not vary significantly with NF. This indicates that more rigid structures reflect the waves (particularly for high wave frequencies), leading to a different hydroelastic response.



(a) Constant floater thickness h_b = 0.2 m. **(b)** Constant Young's modulus E = 500 MPa.

Figure 4.5: Comparison of ΔAIC values between configuration-specific vines and NF-specific vines.

The maximum spanning tree (MST) used in the sequential fitting algorithm is based on the absolute values of Kendall's τ computed from the uniform-transformed (i.e. '*pseudo*') data. From this point onwards, these will be referred to as correlations. The resulting first trees are shown in Figure 4.6 together with the corresponding Kendall's τ values and the selected copula families (including potential rotation) identified via lowest AIC, using the parametric copulas implemented in the PYVINECOPULIB package. The colours highlight variable pairs that consistently appear in the first tree across the different E and h_b configurations.

It is important to note that Kendall's τ values obtained from the fitted vines differ from the empirical τ values derived directly from the data. The former reflect the model-implied dependencies of the fitted pair-copulas, while the latter are purely data-driven. Comparing these two provides an indication of the model's fidelity, i.e. how well the fitted vine reproduces the empirical dependence pattern.

Across all configurations, the fitted vines take the form of either C-vines or D-vines, determined primarily by the MST algorithm from the sequential method (see Section 4.2.2). For the C-vine cases, the root node is consistently V_{max}, as this variable exhibits the strongest overall dependence with the others. The analysis below focuses on the first tree level, as it captures the dominant dependencies.

Several consistent patterns emerge in the first trees across the configurations:

- V_{max} and σ_{max}: This pair shows the strongest dependence (τ = 0.65 to 0.73) across all configurations and therefore forms the first edge according to the MST algorithm. This is because of their calculation (see Equation (4.15) and Equation (4.14)). The relationship is modelled with the BB1 copula for the elastic case of E = 250 MPa and h_b = 0.2 m. For all the other configurations, it is modelled by the BB7 copula,

rotated by 180°, except for the thick case with h_b = 0.4 m. Both copulas exhibit strong lower tail dependence and moderate upper tail dependence.

- H_S and T_P: The second-strongest correlation (τ = 0.39) appears consistently across all configurations, modelled by a Tawn copula (see Section 4.4.1). This is expected since H_S and T_P are identical input variables across cases.

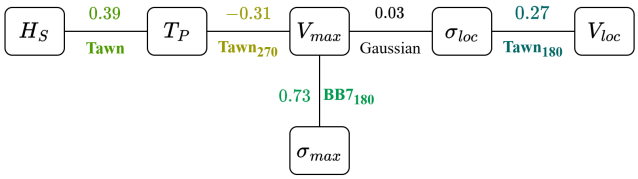
- T_P and V_{max}: A moderately negative dependence (τ = -0.33 to -0.31) is captured by rotated Tawn copulas (90° or 270°). T_P is a key driver for both V_{max} and σ_{max}. However, since its correlation with V_{max} is slightly stronger than with σ_{max}, this variable pair forms a consistent edge in the first tree.

- V_{loc} and σ_{loc}: This variable pair, with a correlation of τ = 0.11 to 0.27, forms an edge in the first trees of all configurations (modelled by a Tawn copula, partly rotated by 180°), except for one rigid configuration (E = 500 MPa, h_b = 0.4 m), as for this configuration, the correlation between V_{loc} and σ_{loc} is very weak (τ =0.02). Note that these correlations are obtained from the uniformed variables, where the influence of NF has been removed. In contrast, the correlations calculated in the physical space, which include NF, show that σ_{loc} and V_{loc} are more strongly correlated for the rigid configurations. This behaviour can be explained by the fact that, in rigid systems, both maxima are more likely to occur at the hinges, whereas in more flexible cases, σ_{loc} tends to occur further within a floater rather than at the hinge. This indicates that NF has a significant effect on the locations of both maxima, an influence that is intentionally removed in the uniformed analysis. The purpose here is to investigate whether there exists a consistent dependence pattern for each configuration defined by E and h_b, independent of NF.

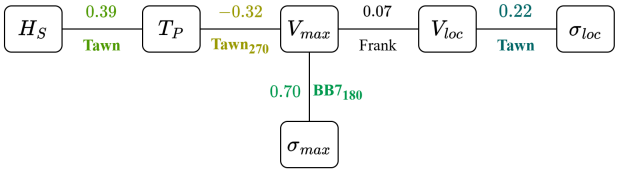
The remaining edges in the first tree follow directly from the MST algorithm but exhibit relatively weak correlations and are therefore not further discussed here. Notably, however, Figure 4.6 shows that most first-tree pair-copula families are asymmetric, particularly for variable pairs exhibiting meaningful dependence.

As mentioned above, the vines determined via the sequential method are either C-vines, or D-vines. It appears that for the configurations where E = 500 MPa, C-vines are determined via the MST, where V_{max} serves as the core node. Different to that, for configurations with E = 250 MPa or 1000 MPa, the MST approach results in D-vines, where V_{max} is no longer directly connected to V_{loc} or σ_{loc}. This can be explained by the MST algorithm: in most configurations, V_{loc} and σ_{loc} form a strong pair. For E = 500 MPa with h_b = 0.1 or 0.2 m, the highest correlation to connect this pair with the remaining variables of the tree is through V_{max}. For E =250 MPa or 1000 MPa, however, the strongest connection occurs via either H_S or σ_{max}. In the thick, rigid case (h_b = 0.4 m), V_{loc} and σ_{loc} do not form an edge due to much weaker dependence as in the other configuration cases.

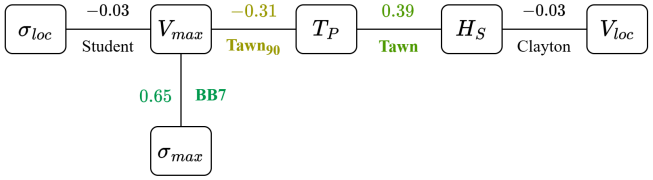
(a) First vine tree for $E = 250$ MPa and $h_b = 0.2$ m.



(b) First vine tree for $E = 500$ MPa and $h_b = 0.1$ m.



(c) First vine tree for $E = 500$ MPa and $h_b = 0.2$ m.



(d) First vine tree for $E = 500$ MPa and $h_b = 0.4$ m.

(e) First vine tree for $E = 1000$ MPa and $h_b = 0.2$ m.

Figure 4.6: First vine trees for each (E, h_b) configuration obtained using the sequential fitting method. Pair-copula families (including rotations, where applicable) and corresponding Kendall's τ values (computed from the *pseudo* data) are indicated. Colours highlight variable pairs that consistently appear in the first tree across different E and h_b configurations.

Overall, while some consistent dependence patterns can be observed across these five configuration cases analysed here, the results from Section 4.4.3 show clearly that a single global vine fails to capture those as well. Nevertheless, these findings indicate that certain generalised vine configurations might exist, describing the dependence pattern for groups of configurations, which would allow for more efficient application of vine copula models for FPV structures in practice.

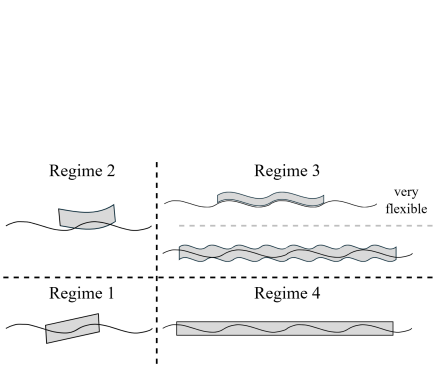
4.4.5. VINE MODELS PER REGIME

In the previous section, a separate vine was fitted for each configuration of the Young's modulus E and the floater thickness h_b. The results indicated that rigid configurations with shorter floaters (i.e. higher NF values) exhibit a noticeably different dependence pattern compared to the elastic and reference cases. This observation is not unexpected, as the structural flexibility strongly influences the relation between hydrodynamic impacts and structural response.

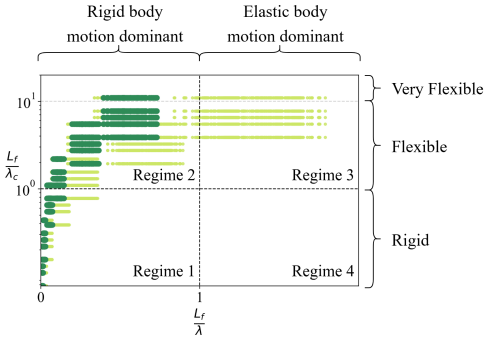
To further explore this behaviour, the vine fitting is now performed based on a classification of the global response of floating structures into four regimes (Alcañiz *et al.*, 2025; Zhang and Schreier, 2022). The regimes are defined using two non-dimensional ratios: the ratio of the floater length L_f to the hydroelastic length λ_c (see Equation (4.16)), and the ratio of L_f to the dominant ocean wave length λ. The resulting four regimes can be described as follows:

- Regime 1: $L_f/\lambda_c < 1$ and $L_f/\lambda < 1$
- Regime 2: $L_f/\lambda_c > 1$ and $L_f/\lambda < 1$
- Regime 3: $L_f/\lambda_c > 1$ and $L_f/\lambda > 1$
- Regime 4: $L_f/\lambda_c < 1$ and $L_f/\lambda > 1$

Regimes 1 and 2 are typically dominated by rigid-body motion, whereas regimes 3 and 4 are characterised by elastic-body motion (see Figure 4.7a). Figure 4.7b illustrates the distribution of these regimes as determined from the FPV model. The dark-green markers indicate regimes based on the peak period T_P, while the light-green markers correspond to classifications based on the 10th and 90th percentiles of the wave period (T_{10} and T_{90}) of the JONSWAP spectrum. The latter are included as the floating structure is excited by a range of wave periods across the full spectrum rather than by a single monochromatic wave at T_P.



(a) Global response of floating structures per regime.



(b) Dark green indicates regime boundaries based on T_P, while light green corresponds to classifications derived from T_{10} and T_{90}.

Figure 4.7: Classification of the global response of floating structures into four regimes (Alcañiz *et al.*, 2025; Zhang and Schreier, 2022).

For the subsequent analyses, the regime classification is based on T_P, while T_{10} and T_{90} are included as continuous explanatory variables. Accordingly, one vine is fitted per regime. However, since the classification is based on T_P, only regimes 1 and 2 occur in the present dataset.

The variables considered for these regime-based vines are as follows:

- Continuous: H_S, T_P, σ_{max}, σ_{loc}, V_{max}, V_{loc}, T_{10} and T_{90}
- Ordinal: NF and h_b (treated as ordinal due to their limited number of possible values)

Thus, similar as for the global vine, this results in a dataset of mixed continuous and ordinal variables for the vine fitting process. Also, it should be noted that some variables such as H_S, T_P and E are used to define the regimes, but they are nonetheless retained as variables in the subsequent vine fitting and dependence analyses.

VINE FITTING PER REGIME

The goal of this section is to examine whether a single vine copula model per response regime, as defined above, can adequately capture the dependence pattern of the floating structures across different configurations. In other words, the question is whether the statistical dependence among considered variables can be generalised within each regime, rather than requiring a separate vine for each configuration of E, h_b, or NF. The procedure follows the same principle as in the previous sections. For each regime, the data of the considered variables are first transformed to the unit scale. A vine copula model is then fitted to these uniformed data using the sequential selection method. Additionally, for each value of NF, considered variables are uniformed and a separate vine is fitted to each. The comparison between the regime-level vine and the vines fitted per NF allows assessment of how well a generalised regime vine, independent of the marginal distributions, represents the dependence pattern of floating structures with different number of floaters NF.

The results are presented in Figure 4.8. Similar to before, the figure shows the differences in AIC values (ΔAIC) between the vines fitted per regime and the specific vines fitted per NF. ΔAIC values close to zero indicate that the regime-based vine reproduces the dependence pattern of that particular NF case well. In the figure, squares (□) represent regime 1 (where $L_f/\lambda_c < 1$ and $L_f/\lambda < 1$), and circles (∘) represent regime 2 (where $L_f/\lambda_c > 1$ and $L_f/\lambda < 1$), see also Figure 4.7b. A good agreement is found for regime 2, i.e. ΔAIC are relatively close to 0, especially for cases with a low number of floaters ($NF < 5$), corresponding to longer individual floaters. For regime 1, the fit is reasonable for configurations with $NF \geq 25$, although the agreement is generally weaker compared to regime 2 for the lower NF. A likely explanation is that with increasing NF, the structure consists of more, but shorter floaters, resulting in a larger number of hinges and consequently higher overall flexibility, while each individual floater behaves more rigidly. Consequently, when the floaters are longer (i.e. fewer hinges), the structure behaves more as a single rigid body, making the dependence patterns among the studied variables more consistent and easier to generalise. The case with $NF = 10$ does not clearly belong to either regime. This

is consistent with the regime classification, as some of the corresponding cases fall into regime 1 and others into regime 2.

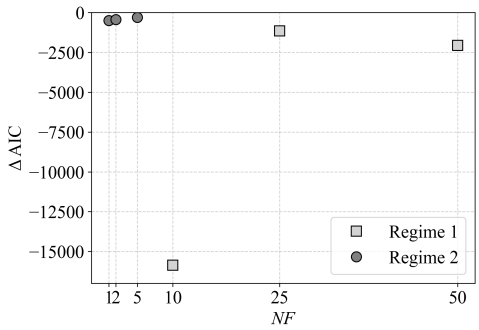
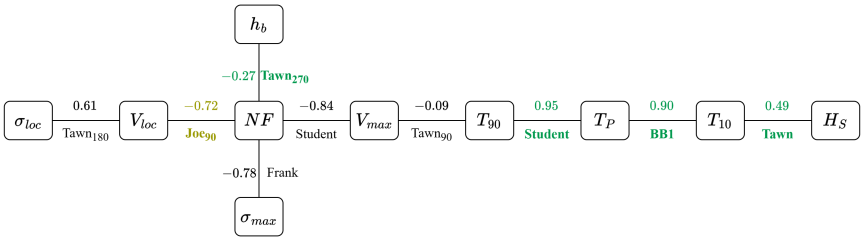


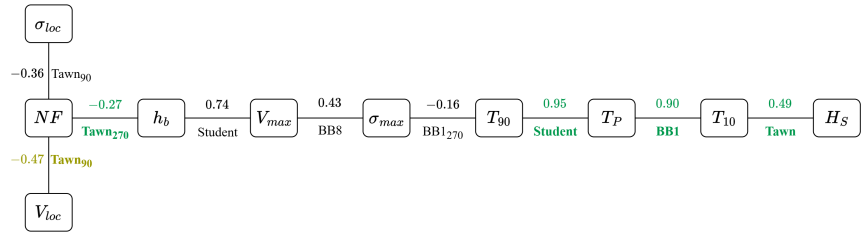
Figure 4.8: Comparison of ΔAIC values between regime-specific vines and NF-specific vines.

To gain further insight into the dependence pattern, the first trees of the fitted vine copula models for each regime, representing the strongest pairwise dependencies identified by the sequential fitting algorithm, are presented in Figure 4.9. Each edge in the first tree is annotated with the empirical Kendall's τ and the selected parametric copula family (including potential rotations) determined via the lowest AIC. Common variable pairs appearing in the first trees of the vine copula models of both regimes are marked in colour. Both regime-based vines are found to be C-vines, with NF acting as the central node. This means that NF serves as the conditioning variable for all pair-copulas in the first tree, highlighting its central role in governing the structural response.

For detailed analyses of the first trees, the scatter plots of the data (in physical units) per regime are presented in Figure C.4b and Figure C.4a in the C.3. The strongest correlations are observed between T_P and T_{90}, as well as between T_P and T_{10}, which is expected. According to the MST algorithm, H_S and T_{10} form an edge, as this variable pair exhibits the overall highest correlation. These relationships are identical for both regimes, as these input variables are shared. The second strongest absolute correlations, however, differ between the regimes. For regime 1, NF and V_{max} show a strong negative correlation ($\tau = -0.84$), described by a symmetric Student-t copula (see Figure 4.9a), while for regime 2, this correlation is weak ($\tau = -0.14$, see Figure 4.9b). Here, however, the variable pair of V_{max} and h_b exhibit a strong positive correlation ($\tau = 0.74$) and is also described by a student-t copula, whereas this correlation is much weaker in regime 1 ($\tau = 0.22$). These patterns suggest that in regime 1 (shorter floaters), the length of the floaters (i.e. NF) has a significant influence on the maximum shear force V_{max}, while in regime 2 (longer floaters, $NF \leq 5$), the floater length has a much smaller effect. Instead, the thickness h_b becomes the dominant factor influencing the shear force response. This is consistent with the physical interpretation that shorter floaters behave more rigidly, so changes in thickness have less impact, whereas for longer and more flexible floaters, the structural thickness has a stronger influence on the shear forces.



(a) Regime 1.



(b) Regime 2.

Figure 4.9: First vine trees for each regime obtained using the sequential fitting method. Pair-copula families (including rotations, where applicable) and corresponding Kendall's τ values are indicated. Colours highlight variable pairs that consistently appear in the first tree across different regimes.

In general, the correlations between NF and the other considered variables are stronger for regime 1 than for regime 2. This indicates that the more rigid the overall floater behaviour, the more sensitive the structural response is to the floater length. Conversely, in regime 2, where the behaviour is more flexible, variations in NF have a less pronounced influence on the response variables. This also explains why both regime-based C-vines differ in their layout, even though NF remains the central node in both. It confirms that NF continues to be the key configuration variable that drives the overall structural response, particularly longitudinal stresses and shear forces. Consequently, comparing the generalised regime-based vines with the individual NF vines proves to be a meaningful and appropriate approach to evaluate the generalisability of dependence structures across different floater configurations.

VINE PERFORMANCE AND PREDICTIVE CAPABILITY

Having established the dependence patterns and structural differences between the vines fitted per regime, the next step is to assess how well these regime-based vines can actually perform, i.e. whether they are able to reproduce or forecast the main response quantities of interest. In particular, the focus here is on the maximum shear force $V_{\max}$ and the maximum longitudinal stress σ_{max}, as these represent key indicators of the structural response. The central question is whether a single vine per regime can adequately capture the behaviour of floating structures across different configurations and conditions.

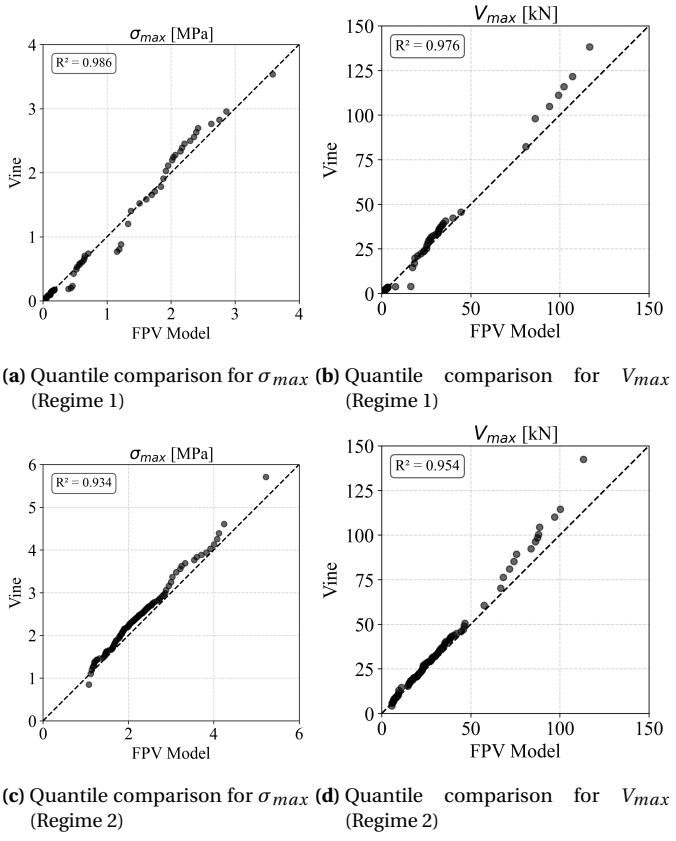
To test this, the distribution of σ_{max} obtained from the FPV model is compared with

that simulated from the corresponding vine model. Both are conditioned on the median of the peak wave period (namely T_P = 13.59 s), which is relatively high, as the input data for the FPV model represent extreme conditions. Note that the correlation of T_P with σ_{max} is generally very low (-0.05 and -0.14 for regime 1 and 2, respectively), meaning the distribution of σ_{max} does not change significantly through this conditionalisation process. For this comparison, 100,000 samples are simulated from each fitted vine, conditioned on the 50th percentile in unit space. These simulated samples are then transformed into physical space using the empirical cumulative distribution function. For the empirical FPV model results, an equivalent conditioning procedure is applied directly in physical units based on percentiles.

In both cases, an "error window" around the median of ±2 % is defined. This results in sample sizes of roughly 4 % of the total data, corresponding to around 4,000 simulated samples from the vine and approximately 180 empirical samples from the FPV model. Due to this difference in sample size, a quantile-quantile (QQ) plot is used to compare the two distributions. The QQ-plot compares the quantiles of σ_{max} and V_{max}, respectively, simulated from the vine with those obtained from the FPV model, allowing for a visual assessment of how well the vine reproduces the observed distribution. It should be noted that no bootstrapping was applied in this analysis. Consequently, the simulated results vary slightly with each run, since each simulation from the vine involves random sampling. More specifically, this process involves generating a matrix of independent uniform random variables, which are then transformed through the inverse Rosenblatt transform. This step converts independent uniform variables into dependent ones that follow the multivariate dependence structure specified by the fitted vine. The results are shown in Figure 4.10 for Regime 1 and Regime 2, both for σ_{max} (Figure 4.10a and 4.10c) and $V_{\max}$ (Figure 4.10b and 4.10d). In addition to this visual comparison, a quantitative measure of goodness of fit is required to assess how well the vine-based simulations reproduce the FPV model results. While classical tests such as Kolmogorov–Smirnov (KS) or Cramér–von Mises (CvM) could be applied, comparing either the maximum or overall distance between cumulative distributions, here the coefficient of determination R^2 is chosen. This is because, as shown earlier in Figure 4.8, cases such as NF = 10 do not fit well into either regime. As a result, the KS and CvM tests would not yield meaningful values, while R^2 provides a more direct measure of how well the quantiles align between both distributions.

Overall, the results indicate a good agreement between the vines and the FPV model. For Regime 1, R^2 exceeds 0.976, and for Regime 2, it remains above 0.934. One might expect Regime 2 to perform slightly better, based on the ΔAIC values discussed earlier. However, those values only indicate how well a vine per regime fits each specific NF case. Regime 2 applies primarily to cases with $NF \leq 5$, where ΔAIC remains finite, whereas Regime 1 covers all NF but performs poorly for $NF < 10$, where ΔAIC becomes very large to even infinite. Examining the correlations between NF, σ_{max}, and $V_{\max}$ helps explain these differences. In Regime 1 (the more rigid regime), correlations between NF and both σ_{max} and $V_{\max}$ are strong and negative (-0.78 and -0.84, respectively). In contrast, in Regime 2 (the more elastic regime), these correlations are very weak (0.05 and -0.14, respectively). The more flexible the floaters (i.e. the longer the floaters), the less sensitive does its response become to the exact length of the floaters and thus variations in NF no

longer significantly affect σ_{max} or $V_{\max}$.



(a) Quantile comparison for σ_{max} (Regime 1)

(b) Quantile comparison for V_{max} (Regime 1)

(c) Quantile comparison for σ_{max} (Regime 2)

(d) Quantile comparison for V_{max} (Regime 2)

Figure 4.10: Quantile–quantile (QQ) plots per regime comparing σ_{max} and V_{max} simulated from the vine model with those obtained from the FPV model. The coefficient of determination (R^2) values are provided to quantify the goodness of fit.

The correlation between σ_{max} and $V_{\max}$ itself is also notably different between the two regimes, being 0.77 for Regime 1 and 0.43 for Regime 2. This suggests that in the more elastic regime, the structural response is generally more complex and therefore more difficult to model or predict accurately. This may explain why Regime 1 slightly outperforms Regime 2 in reproducing σ_{max} and $V_{\max}$. The response of σ_{max} and $V_{\max}$ also depends on the ratio of the wave length (or equivalently, the wave period) to the length of the individual floaters, as this ratio governs the structural excitation and, in most unfavourable cases, approaches the system's natural frequency. Despite these nuances, the overall predictive performance of the vines per regime can be considered good. While the extremes, particularly for $V_{\max}$, tend to be somewhat overestimated, the vines reproduce the lower and more moderate ranges of the maximum longitudinal stress and shear forces remarkably well.

CONDITIONAL ANALYSIS OF STRUCTURAL CONFIGURATIONS

It was demonstrated above that the vine model allows conditionalising on a specific variable value, such as the median wave period T_P = 13.59 s, to obtain the corresponding distributions of the other variables. This ability to condition on arbitrary variable values while preserving the complex dependence structure among all variables is one of the major advantages of the vine copula models. Building on this, the following section serves as a demonstration case to illustrate how the vine model can be applied to identify the configurations per regime, that perform best in keeping the maximum longitudinal stress σ_{max} low. Since the floating structures considered here are complex systems operating in a highly dynamic environment, it is not straightforward to determine which combination of variables provides the most favourable structural response. The performance depends on several interacting factors, such as the relation between the floater length and the dominant wave length, or the floater thickness h_b. For instance, doubling h_b not only doubles the structural mass but also the surface area exposed to wave loading. In general, more elastic structures tend to experience higher stresses compared to more rigid ones.

To assess the most suitable configuration per regime, the analysis is conditionalised on a selected threshold of the maximum stress. Specifically, the 66th percentile of σ_{max} is chosen as a reference value, denoted as $\sigma_{max,S}$. For each configuration, the relative frequency (occurrence [%]) of cases where the simulated σ_{max} remains below this limit is determined. In Figure 4.11, results are indicated by squares (□) for Regime 1 ($\sigma_{max,S}$ = 0.6 MPa) and circles (•) for Regime 2 ($\sigma_{max,S}$ = 2.6 MPa). The plot is divided by two vertical dashed lines, corresponding to the different floater thicknesses: h_b = 0.2 m (left), h_b = 0.1 m (middle), and h_b = 0.4 m (right).

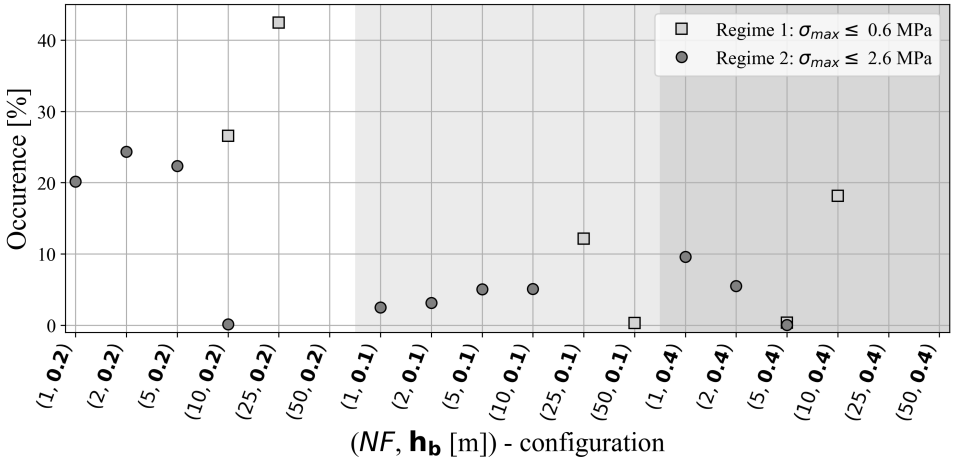


Figure 4.11: Occurrence of σ_{max} below the 66th percentile ($\sigma_{max,S}$) for each configuration, where squares indicate Regime 1 ($\sigma_{max,S}$ = 0.6 MPa) and circles indicate Regime 2 ($\sigma_{max,S}$ = 2.6 MPa). The background shades indicate configurations with equal floater thicknesses h_b (marked in bold).

Most configuration cases belong clearly to one of the two regimes, but there are excep-

tions. For instance, the case with $NF = 5$ and $h_b = 0.4$ m is represented by both a square and a circle, as it can fall into either regime. In contrast, for some configurations neither a square nor a circle is visible. Either because they do not belong to Regime 1 or 2, or because σ_{max} is always above $\sigma_{max,S}$. Note also that while the Young's modulus E is not displayed directly in the figure, it is implicitly included in the regime classification (see beginning of Section 4.4.5).

For the chosen limits, the most favourable configurations are identified as follows: in Regime 1 (rigid), the structure with 25 floaters and a thickness of 0.2 m achieves the highest frequency of σ_{max} remaining below the set threshold; in Regime 2 (elastic), this applies to the configuration with 2 floaters and a thickness of 0.2 m. Interestingly, in both regimes the best-performing configurations are neither the most nor the least extreme in terms of number of floaters or thickness. This again highlights the complexity of the system behaviour and the interplay between structural flexibility and hydrodynamic forcing.

4.5. DISCUSSION

The main objectives of this study were to describe and analyse the dependence patterns among hydrodynamic and structural response variables of an offshore FPV system. A further objective was to examine whether generalised dependence structures can be established across different structural configurations. Finally, the predictive capability of suitable vine copula models was tested to assess whether the fitted vines can replicate the observed system behaviour.

Overall, the results show that vine copula modelling provides a statistically powerful framework for capturing complex multivariate dependencies in offshore floating structures by capturing nonlinear and asymmetric relationships through different copula families. Within the boundaries of the explored problem, it is shown that vine copulas result in an improved understanding of the interaction between environmental and structural variables. This allows for translating these statistical relationship into directly actionable engineering guidance.

The applied fitting approach follows the sequential method of Dißmann *et al.* (2013), which constructs vines through a maximum spanning tree (MST) based on maximising the sum of pairwise absolute correlations. This ensures computational efficiency but risks failing to determine a sufficiently accurate vine configuration. Essentially, analysing the tree configuration of the vine determined via the sequential method yields little additional information compared to a correlation analysis. However, the added value comes by selecting pair-copulas, which are able to capture non-linear and asymmetric dependencies. Still, it is important to note, however, that the sequential method does not explore all possible vine configurations. It determines only one configuration via the MST. An exhaustive search across all possible vine structures could, in principle, yield a better vine configuration and thus a more accurate description of the underlying multivariate dependence pattern, but this becomes computationally infeasible for more than about 7 to 8 variables (Morales-Nápoles, 2010b). Even for moderate dimensions (4 to 6 variables), the benefit of such a search may not justify the effort, as improvements in fit might be marginal relative to the computational cost and depend on the application

(see Section 5.1.6 for a comparison of different vine copula models with three variables related to an application of this dissertation). Therefore, the commonly applied sequential method represents a practical and balanced compromise between interpretability and computational feasibility.

A key aspect of this study was the comparison of generalised vines, fitted across multiple configurations, with vines fitted specifically for each number of floaters (NF). Using NF as the primary configuration variable proved to be a meaningful choice, as it was shown to be the dominant variable influencing both shear force and longitudinal stress (see Section 4.4.5). The comparison through AIC demonstrated that a single, global vine model is insufficient to describe the full range of structural behaviours. As shown in Figure 4.4, the global vine fitted better to configurations with smaller NF (more rigid floaters) than to larger NF (more elastic floaters). This reflects fundamental differences in the underlying physical behaviour: while rigid configurations are governed primarily by hinge effects and global motions, elastic configurations are dominated by distributed bending and local deformations. Consequently, both the marginal distributions and the dependence patterns among variables vary substantially between configurations, making a single global model physically and statistically unsuitable.

The analysis presented in Section 4.4.4, where vines were fitted for each configuration of (E, h_b), served as an intermediate diagnostic step. Earlier results had shown that NF is the dominant variable governing the structural response and that a single global vine across all NF values cannot adequately represent the system behaviour (Section 4.4.3). The configuration-based analysis therefore aimed to examine whether variations in E and h_b introduce additional systematic differences in the dependence pattern that might justify a more detailed modelling approach. While the results did not indicate an influence of E and h_b comparable to that of NF, they did reveal a clear separation between elastic and rigid structural responses. This distinction subsequently motivated the regime-based modelling strategy in Section 4.4.5. The findings thus suggest that generalised vine models are feasible, but only when configurations are grouped into physically meaningful regimes, such as the rigid and elastic behavioural domains.

The modelling framework used in this study involved mixed continuous–ordinal data, as several structural design variables are considered ordinal, particularly NF, E and h_b (see Section 4.3.3). Their ordinal treatment in the vine fitting process ensures a mathematically consistent representation of discontinuous distributions, as supported by PYVINECOPULIB. In the present analysis, this aspect primarily served as a methodological requirement: although treating these variables as ordinal within the vine fitting process ensures a consistent handling of their limited categorical resolution, this aspect did not influence the substantive findings, which were dominated by the physical differences between elastic and rigid regimes rather than by the discretisation of design variables. Thus, future analyses could further explore the influence of these ordinal variables on the fitted vines with a different dataset.

The conditionalisation analysis illustrates the potential of vine copula models for engineering applications. In this study, for each structure configuration, stress thresholds

$\sigma_{max,S}$ were set to the 66th percentile of σ_{max}. This threshold can also be set to other critical stress values given e.g. the breaking strength of the overall structure. Conditioning σ_{max} on $\sigma_{max,S}$, configurations could be identified which most frequently remain below this$\sigma_{max,S}$. The presented approach effectively demonstrates the flexibility of vine copula models in exploring conditional dependencies and response sensitivities across complex structural regimes.

While such an evaluation could in principle also be conducted using physical data, the limited sample size would restrict its reliability. In contrast, vine copula models overcome this limitation by enabling the generation of arbitrarily large sample sets that preserve the underlying dependence pattern. This enables robust conditional analyses and probabilistic assessments even with relatively sparse original data. In the present work, conditional distributions were obtained through random sampling from the fitted vine, followed by empirical conditioning and transformation to physical space. As this is done via empirical CDFs, which requires interpolation in the sparsely populated tail regions, one needs to be aware that this increases the uncertainty in the tails. However, since a vine model provides a full decomposition of the joint probability density function (Aas *et al.*, 2009), theoretically an even more rigorous approach would be to compute conditional distributions directly from this density. Given that both the vine configuration and the parametric pair-copula families are known, such distributions could be derived numerically without relying on sample generation, thereby reducing uncertainty and improving computational efficiency. However, yet to date, there is no global algorithm for such an analytic approach.

The analysis was intentionally limited to extreme environmental conditions in significant wave height and peak period. It is, however, possible that certain non-extreme combinations of these variables could lead to even higher structural responses, particularly if they coincide with resonance effects related to the system's natural frequencies. Such cases were not intentionally captured here and represent a potential area for future work. Moreover, at the highest ranges of H_S and T_P, the empirical data from the DCSM model showed a discrete appearance, likely due to limited availability of extreme events even across 30 years of simulation data. As a result, the fitted Tawn copula also produced relatively few extreme values, giving the impression of discretisation in those ranges.

The transferability of the vine copula models developed here remains an open question. The analysis was performed for one specific site in the German North Sea. It would be valuable to test, whether the identified dependence pattern remain stable across different offshore locations or environmental regimes. If so, the vines defined here could be used to predict structural responses, such as shear force and longitudinal stress, based on site-specific wave characteristics, offering a potentially powerful and generalisable tool for FPV design.

Finally, it should be emphasised that the present work serves primarily as a methodological demonstration rather than a detailed physical interpretation of each individual configuration case. A full discussion of the underlying physical mechanisms would require a deterministic, physics-based modelling approach, which is beyond the scope

and intention of this study. Instead, the focus here lies on demonstrating the capability of vine copula models to represent and explore the probabilistic dependencies within complex offshore floating systems.

4.6. CONCLUSION

This study investigated the dependence structure between hydrodynamic variables and structural response measures (maximum longitudinal stress and shear force) for various configurations of offshore floating photovoltaic (FPV) structures. Vine copula models were employed to capture these relationships, including cases involving mixed continuous-ordinal data arising from discretised structural variables.

A central aim was to identify generalisable dependence patterns across different structural configurations. Therefore, global, configuration-specific, and vine models per regime (i.e. elastic or rigid structure behaviour) were compared. The results showed that a single global vine is not suitable, as it cannot represent the full range of structural behaviours. However, when configurations were grouped into physically meaningful regimes, consistent and generalisable dependence patterns could be identified. The few remaining deviations likely stem from eigenfrequency effects or from configurations that do not clearly align with a single regime.

Overall, the findings demonstrate that vine copula models provide a powerful framework for probabilistic analysis of offshore floating structures. Their ability to account for complex multivariate dependencies, combined with the option to generate large synthetic datasets, enables detailed conditional assessments, such as identifying configurations likely to remain below prescribed stress thresholds.

Finally, the transferability of the developed models presents a promising direction for future research. First, it needs to be further investigated whether the identified dependence patterns persist across different sites. If this holds, the vine copula models could be applied to new locations without refitting the dependence structure, requiring only an update of the marginal distributions to local environmental conditions. This would enable the prediction of structural responses at new sites and further strengthen the role of vine copula modelling in FPV design and assessment.

5

DISCUSSION & CONCLUSION

The content of this chapter is divided into a discussion section and a final conclusion. The discussion is divided into two sections, each addressing a distinct aspect of the overarching dissertation. Section 5.1.7 forms the core of this chapter, providing a detailed reflection on the applied methodologies, with particular emphasis on the advantages and limitations of Gaussian Copula-based Bayesian Networks and vine copula models, including their performance in capturing extreme values. Section 5.2.2 addresses the practical application of these methods and provides a guide for engineers in selecting an appropriate method. The second part, Section 5.3, presents the conclusions of the dissertation and addresses the research questions outlined in Section 1.6.

5.1. METHODOLOGICAL REFLECTIONS

5.1.1. VARIABLE SELECTION AND SAMPLING

A fundamental challenge in statistical environmental modelling and throughout this dissertation is the appropriate selection of variables, which strongly depends on the focus of investigations, but also on the availability of data. For example, for holistic structural analyses (see Chapter 3 and 4), the wind impact might be relevant. However, in this dissertation, the focus lies exclusively on the hydrodynamic impact, still, including other relevant variables might improve the accuracy of the results. Additionally, the sampling of selected variables has an influence on the obtained results. Here, the emphasis is on extreme values, for example extreme ecological conditions or storm events, as these usually pose the highest risks. To account for temporal dependence among variables to a certain degree, concomitant sampling was applied, i.e. sampling variables within a ± 4 h window around the samples of the dominant variable. Note that the choice of window size is crucial and might have an impact on the results. For instance, one variable may have an influence on another with a temporal lag, or multiple variables may simultaneously exhibit extreme values, indicating potential interdependencies or joint extreme behaviour. If the window for selecting concomitants is too narrow, relevant time-shifted extremes may be missed. If it is too wide, extremes that are not physically related may be falsely captured. This issue reflects the broader risk of incorrectly assessing joint extremes without careful consideration or evaluation of the most appropriate sampling

strategy. Note that there is a difference regarding sampling across this dissertation. While Chapter 3 relies on daily maxima over a relatively short four-year dataset, Chapter 4 instead selects extremes from a long-term dataset of 30 years and thus also captures extremes with lower return periods. This approach on a long-term dataset would have been more suitable in Chapter 3 as well, however, data for fours years was available and assumed to be sufficient.

The sampling strategy applied to variables can significantly influence the resulting empirical distributions. These empirical margins are sensitive to the sample size and thus there is the risk that long-term characteristics such as climate change effects might not be captured properly. This is particularly relevant for methods such as Gaussian copula Bayesian Networks (GCBN), as applied in Chapter 3. The GCBN relies on empirical margins of the variables for both, set-up of the structure and characterisation of the nodes. While such non-parametric approaches allow for flexible updating when new data become available, they remain highly sensitive to the sample size.

In this dissertation, empirical margins are also relevant in Chapter 2, where they are used to transform thresholds in physical units to unit space, to allow for the calculation of probabilities of meeting these thresholds. In Chapter 2, only two years of data have been used which do not capture long-term characteristics. In such cases, parametric margins may be preferable for threshold definition in unit space, as they can incorporate long-term characteristics more robustly. However, in Chapter 2, suitability maps were design at the current status, which did not aim to account for long-term changes.

5.1.2. SIGNIFICANCE TESTING

Throughout this dissertation, significance testing such as the use of p-values is employed, however, it has been subject to considerable criticism in the research community. The main disadvantage of the p-value is that it depends on the sample size. When applying the p-value in this dissertation to assess determined correlations, the sample size was relatively high, between 286 and 915. A common critique of significance testing is that it primarily assesses whether the sample size is sufficient to detect statistical relationships, rather than confirming substantive insights. Therefore it is essential to emphasise that statistical significance is not equivalent to scientific significance. Usually, the threshold α is set to 0.05. If the p-value falls below this threshold, the determined rank correlation is statistically significant, which means that the likelihood of this correlation occurring by chance is very low, or in other words, it is unlikely that the null hypothesis (i.e. there is no correlation) is true. If the p-value is above the threshold, then the observed correlation for the dataset could indeed have occurred by chance.

Often however, this threshold, usually chosen to be 0.05, is overemphasised. If the p-value is just above the threshold, this can lead to the dismissal of potentially relevant findings. Conversely, practices such as p-hacking (i.e. manipulating data or repeating analyses until a significant result emerges) undermine scientific integrity and are particularly common in pseudoscientific contexts. In this dissertation, thresholds were applied conservatively to assess statistical correlations, but with the awareness that such results should not be confused with their scientific or practical importance.

5.1.3. COPULA MODELS

Copula models were applied throughout this dissertation to model dependence patterns between environmental and structural variables. For Chapter 2 and Chapter 3, the analyses were restricted to a set of the five most commonly used bivariate copula families, together with their rotations. These included the Clayton, Gumbel, Frank, and Gaussian copulas, with the addition of the Student-t copula in Chapter 2. The selected families and their rotations were chosen because they represent a broad spectrum of dependence structures and thereby capture the key characteristics of the environmental and structural variables investigated in Chapter 2 and 3. Thus, they were considered sufficient for the analyses presented. A noteworthy finding that emerged later in Chapter 4 is that the Tawn copula provided an improved description of dependence between extreme values of the peak period (T_p) and significant wave height (H_s). Nevertheless, in Chapter 2 and 3, the focus was on one-parameter copulas for computational efficiency, together with the Student-t copula in Chapter 2. Both chapters involved analyses of large spatial domains or multiple sites, and extending the study to higher-parameter copula families would have been computationally challenging. While other one-parameter copulas, such as the Joe copula, exist, they were not included. In practice, the Joe copula exhibits similar behaviour to the rotated Clayton copula, which was already considered, and was therefore excluded from the analysis.

In Chapter 3, an additional restriction arose from the application of Gaussian Copula-based Bayesian Networks (GCBN). These models assume that the dependence structure between the connected variables is Gaussian or at least symmetric. In Chapter 3, the GCBN was applied twice: first, to provide environmental input data for the finite element model of the mooring line, and second, to analyse the model output and the interdependence with hydrodynamic variables. If the underlying dependencies were strongly asymmetric, the use of GCBN could distort the dependence structure, thereby propagating errors into later steps of the analysis. However, the goodness-of-fit (GoF) tests based on the Cramér-von Mises (CvM) statistic indicated that symmetric copulas, such as the Frank or Gaussian copulas, provided an adequate fit to the data, justifying the use of GCBN in this context.

This approach differs from Chapter 4, where the analysis focused on a single location in the North Sea. This enabled the use of the full range of parametric copula families available in the PYVINECOPULIB library in PYTHON (Nagler and Vatter, 2025). Nevertheless, across all chapters of this dissertation, the primary focus has been on extreme values of environmental and structural variables. In such cases, asymmetries in the dependence structure, particularly in the form of tail dependence, are common. This is illustrated again in Chapter 3: although the CvM test indicated that the Gaussian copula was a sufficiently good fit to the data, closer inspection of Figure 3.7 revealed the presence of upper tail dependence among tension rates, that the Gaussian copula is unable to capture. The practical implication of this is that predictions based on Gaussian copulas may underestimate dependence in the extremes. For example, if tension rates are known at one rope position of a mooring system and used to infer tension rates at another, relying on the Gaussian copula would neglect the tail dependence and thus lead to systematic underestimation of extreme tension rates. In Chapter 2 and 3, the GoF of copula models was assessed using non-parametric tests such as the CvM and the semi

correlation (SC) test. These tests evaluate how well a parametric copula reproduces the empirical dependence structure. The CvM statistic is based on the sum of squared differences between the empirical joint distribution function and that of the parametric copula. For the SC test, data is first transformed to standard normal margins to allow for a better evaluation of asymmetries, before correlation coefficients across quadrants are compared. The smaller the difference between the parametric and the empirical copula, the better the fit. While both tests assess the fit of a parametric copula to the empirical copula, they do not account for model complexity. In the context of Chapter 2 and 3, this limitation was of minor importance, since only one-parameter families (plus the student-t copula in Chapter 2) were considered. In contrast, when a broader set of families is explored, such as in Chapter 4, complexity becomes relevant. In that chapter, model selection was therefore performed using information criteria such as the Akaike Information Criterion (AIC), which explicitly balance model fit against parsimony.

5.1.4. GAUSSIAN COPULA-BASED BAYESIAN NETWORKS

In Chapter 3, Gaussian Copula-based Bayesian Networks (GCBNs) were employed to represent multivariate dependencies. The network structure was determined iteratively, beginning with arcs between highly correlated variable pairs and extending step by step until the implied correlation matrix provided a sufficiently good match to the empirical one and goodness-of-fit (GoF) tests were satisfied. This pragmatic approach aims at achieving a "good enough" structure without the need to explore all possible alternatives. In principle, one might compare all feasible network structures, as is increasingly common in the context of vine copulas, but this becomes computationally challenging in higher dimensions and was therefore not pursued here.

Among the applied GoF measures, the d-calibration score was included as a means to assess how well the correlation structure of the fitted GCBN aligns with that of the empirical data. While this metric can offer a rough indication of suitability, its practical usefulness is limited. It is highly sensitive to sample size, and, in most cases of moderate dimension, it provides little additional insight. d-calibration may become increasingly relevant in contexts involving high-dimensional data or when multiple GCBNs need to be constructed and analysed, such as across numerous locations in the North Sea. In such cases, d-calibration can serve as a practical and computationally efficient tool to provide a preliminary assessment of whether a given dependence structure is likely to be an appropriate or well-fitting model.

5.1.5. VINE COPULA MODELS

In contrast to the Gaussian Copula-based Bayesian Networks (GCBN), regular vine copula models offer substantially greater flexibility, as any bivariate copula family can be used as a building block, rather than being restricted to e.g. the Gaussian copula only. This flexibility allows vine copulas to capture asymmetric dependence structures and therefore makes them particularly suitable for the analysis of extremes. This, however, also means that vine copula models are sensitive to the specific choice of copula families.

A key challenge in applying vine copula models is the selection burden, as one needs to determine both the vine structure and the copula families associated with each edge

(see Section 5.1.3). The number of possible vine structures increases factorially with the number of variables (Morales-Nápoles, 2010b) and therefore vines can become computationally very expensive. Thus, an exhaustive search becomes infeasible beyond relatively low dimensions (around seven to eight variables), while in low-dimensional cases, all possible structures can be compared (Morales-Nápoles *et al.*, 2023).

Several strategies regarding structure and copula selection exist to lower this complexity, however, often at the cost of accuracy. Vine copulas are almost always implemented under the so-called simplifying assumption, which assumes that the functional form and parameters of the pair copulas remain constant and do not vary with the specific values of the conditioning variables. Regarding structure selection, sequential algorithms such as the method proposed by Dißmann *et al.* (2013) have been widely used. This greedy search algorithm constructs each tree by selecting the maximum spanning tree, i.e. the structure that maximises the sum of pairwise dependence. Despite these advances, it is important to note that no heuristic algorithm exists yet to select the best universal vine structure. Another strategy is truncation, where copulas in higher-order trees (i.e. conditional pair copulas deeper in the vine) are set to independent copulas, assuming that their contribution to the overall dependence is less important than that of the first tree(s). Brechmann and Joe (2015) proposed an approach to select the adequate number of levels, ensuring that the truncated vine remains sufficiently flexible to provide a good fit to the data. Similarly, when selecting copula families, practical considerations such as computational resources and dimensionality often require limitations. For instance, analyses may be limited to one-parameter copula families, which reduces computational costs while still capturing a range of dependence characteristics.

The question of which vine model provides the "best" representation of the data can be addressed through several criteria. The log-likelihood provides a direct measure of model fit, but complexity must be taken into account. Therefore, information criteria such as the Akaike Information Criterion (AIC) are commonly applied, as they balance model fit and complexity. The Bayesian Information Criterion (BIC) is also used in some applications, with a stronger focus on model complexity compared to AIC. Finally, it should be emphasised that, while vines offer a powerful tool for representing complex multivariate dependence, their interpretation in applied contexts remains challenging. Although the dependence structure can be visualised through trees, these structures are not necessarily based on physical relationships, which complicates their use in interdisciplinary applications such as environmental or engineering studies.

5.1.6. SENSITIVITY OF VINE STRUCTURES IN LOW DIMENSIONS

As described above, for a small number of variables, vine copula models remain relatively tractable, and it is computationally feasible to enumerate and compare all possible vine structures. This provides a useful setting to investigate the sensitivity of results to structure choice.

In Chapter 2, a one-tree vine (i.e. a vine truncated at level 1, where higher-order conditional dependencies are assumed absent and thus set to independence) was employed to model the dependence among chlorophyll-a concentration (Chl_a), dissolved oxygen concentration (O_{diss}), and water temperature (T_w) (in the mussel case), and compared

against an independent modelling approach. The inclusion of dependence via the one-tree vine resulted in differences of up to 40% in the estimated probabilities of meeting ecological thresholds. While such differences would likely be smaller in higher dimensions, since the joint probability of meeting multiple thresholds generally decreases as more variables are included, this finding highlights the impact of explicitly accounting for dependence even in low dimensions.

Retrospectively, however, the vine copula model in Chapter 2 should have been constructed as a complete vine, i.e. consisting of two trees in the three-dimensional case. For three variables, this means modelling not only the direct pairwise dependencies (first tree), but also the conditional dependence between two variables given the third (e.g. $C_{Chl_a,O_{diss}|T_w}$). This approach was not further explored because probabilistic methods are still unfamiliar within the ecological field and a simpler starting point was needed to ensure the work remained understandable. While the structure in Chapter 2 was chosen based on physics, assuming that both Chl_a and O_{diss} depend on the dominant variable T_w, the optimal vine structure is not necessarily related to or explainable by underlying physics. Since only three variables are involved in this case, all possible vine structures (three in total) can be systematically compared, and the best one can be selected based on statistical criteria such as the Akaike Information Criterion (AIC).

To assess the influence of vine structure with only these three variables (Chl_a, O_{diss} and T_w), three different models are considered: (i) the original physics-based one-tree vine, as applied in Chapter 2, (ii) the same physics-based structure extended to two trees (i.e. a complete vine), and (iii) the best-fitting complete vine structure in terms of lowest AIC. For the latter two, once the vine was defined, samples were generated (n=10.000, as in Chapter 2), transformed back into physical space via empirical cumulative distribution functions (ecdf's), and the resulting probabilities of mussels meeting critical or optimal thresholds were calculated at the FINO3 reference location (by determining the percentage of samples fulfilling the thresholds). Note that a sufficient sample size is required to transform the samples to physical units via the ecdf, lowering uncertainties (see Section 5.1.1).

The results show that the best vine structure in terms of AIC was one in which Chl_a and T_w were conditioned on O_{diss}, which differs from the physics-based structure in Chapter 2 where O_{diss} and Chl_a were conditioned on T_w. For the critical thresholds, the probabilities were nearly identical across all models (92.23% for the one-tree vine (i), 92.03% for the physics-based complete vine (ii), and 91.91% for the 'best' vine (iii)). This lack of sensitivity is consistent with the fact that critical thresholds were set relatively broadly, making them easier to meet regardless of dependence structure.

For optimal thresholds, however, higher difference emerged. While the probability of meeting optimal conditions was 29.87% according to the one-tree vine used in Chapter 2 (i), this value dropped to 18.94% for the complete physics-based vine (ii) and to 20.14% for the 'best' vine (iii). This indicates that the one-tree vine substantially overestimated the probability of optimal growth conditions, and that including the second tree (i.e. the conditional copula) has a significant influence. Interestingly, the difference between the two complete vines, one physics-based (ii) and one statistically optimal (iii), was relatively minor, concluding that in this case the conditional copula in the second tree

has a significant impact and should be considered. It should be noted that for the two complete vine approaches ((ii) and (iii)), the same copula families as in Chapter 2 are considered. However, in this case, the selection of copula families for each variable pair was based on AIC, rather than on CvM statistic and the SC test applied in Chapter 2.

In this analysis, probabilities of meeting certain thresholds are calculated by counting the samples that fulfil the requirements. This approach however depends on the sample size generated from the vine. A more robust approach would be to derive probabilities directly from the joint density of the vine in decomposed form (Aas *et al.*, 2009), since the structure and parametric pair copulas are known. Via numerical integration over this region for each variable, the probability of meeting certain thresholds can be determined. This would yield results independent of the number of generated samples. However, this approach was not applied for the same reason outlined above, namely to ensure that the probabilistic methodology remains accessible for the field of marine ecology.

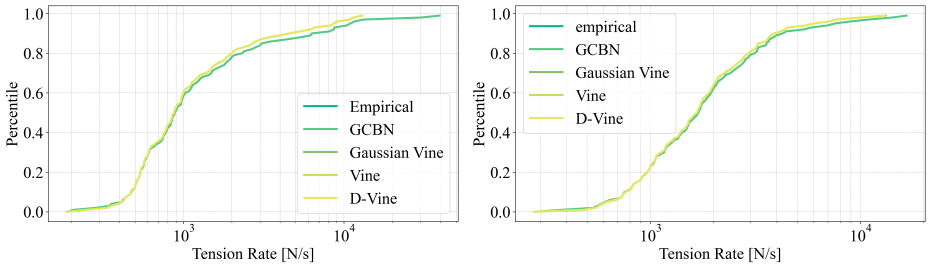
5.1.7. PERFORMANCE OF GCBN AND VINE COPULA MODELS IN CAPTURING EXTREMES

In order to illustrate the implications of different dependence modelling strategies discussed in this dissertation, a comparative case study is conducted. The aim is to highlight potential differences between the empirical approach, the Gaussian Copula-based Bayesian Network (GCBN), and various vine copula models. For this purpose, data from the FINO3 reference location (Chapter 3) are used. This dataset consist of 13 variables, including maximum tension rates at different rope positions (see Figure 3.6), which also served as input for the GCBN analysis in Chapter 3. Prior to modelling, all variables were transformed to uniform by applying the empirical probability integral transform. Three vine-based models were then constructed and compared with the empirical dataset and the GCBN:

- a Gaussian Vine, where the structure was selected using the sequential method of Dißmann *et al.* (2013) and all pair copulas are restricted to the Gaussian family;
- an arbitrary D-Vine, constructed by ordering the variables from 1 to 13 (see Figure 3.6 for the variables and their order)
- a general Vine, based on the sequential method, but allowing for a broader set of copula families.

Note that for the latter two vines, one-parameter copula families are considered (i.e. Gaussian, Clayton, Gumbel, Frank, Joe), which are selected based on the lowest AIC value. Due to the dimensionality (13 variables), it is computationally infeasible to determine the globally best vine structure via exhaustive search, as this would require evaluating approximately 1.12×10^{26} possible configurations. For each of the three vine models, $n = 1.000.000$ samples were generated (same number as for the GCBN in Chapter 2), transformed back into physical units via the empirical cumulative distribution function (ecdf), and subsequently analysed with respect to tension rates at the anchor and fairlead.

The results are illustrated in Figure 5.1a and Figure 5.1b, which compare the cumulative distributions of tension rates across methods. For lower percentiles, differences between the empirical data, the GCBN, and the vine copula models are minimal. From around the 60th percentile onwards, the difference becomes more apparent. Therefore, if the analysis is restricted to moderate values, the computationally more efficient GCBN may be sufficient, as it reproduces empirical percentiles with only minor deviations.



(a) Tension rates at the anchor.

(b) Tension rates at the fairlead.

Figure 5.1: Comparison of the CDFs of tension rates of empirical data, a GCBN and three different vine copula models.

A more detailed comparison of higher percentiles is provided in Table 5.1, which presents relative differences (in %) between the empirical distribution and each modelling approach at the 50th, 90th, and 99th percentiles. For the 50th percentile, differences between the approaches compared to the empirical one are negligible (within around 1%). This however is different for higher percentiles. For example, at the 99th percentile of tail-characterised tension rates at the anchor (see Figure 3.7), the three vine models differ by less than 0.15% from the empirical reference, whereas the GCBN deviates by 203.8%. This indicates that any of the presented vine models is considerably better suited to capture extreme behaviour than the GCBN, especially when asymmetric dependence patterns are present and the focus is on extreme values.

Table 5.1: Relative differences (in %) for selected tension rate percentiles comparing the GCBN and three vine copula models with the empirical distribution.

		50th percentile	90th percentile	99th percentile
GCBN	Anchor	+ 0.82	+ 38.34	+ 203.84
	Fairlead	+ 1.17	+ 10.01	+ 26.48
Gaussian Vine	Anchor	+ 0.01	- 0.24	+ 0.15
	Fairlead	- 0.02	+ 0.11	+ 0.59
D-Vine	Anchor	- 0.16	+ 0.02	+ 0.04
	Fairlead	- 0.50	+ 0.10	- 0.28
Vine	Anchor	- 0.06	- 0.55	- 0.06
	Fairlead	- 0.18	- 0.18	- 0.70

Interestingly, even the Gaussian Vine outperforms the GCBN, despite both being re-

stricted to Gaussian copulas. A potential explanation lies in the modelling semantics: GCBNs are designed around the concept of conditional independence, which implies that two variables are conditionally independent given a third variable. Vine copulas on the other side do not rely on such conditional independence assumptions, but instead model dependence flexibly through each layer of conditional copulas (i.e. across each tree of the vine). This difference in conceptual foundations may explain why vines, even if they only incorporate Gaussian copulas, more accurately capture the dependence structure observed in the data.

Nevertheless, the question remains how much model accuracy is actually required for engineering applications. In particular, it is unclear whether a structurally suboptimal vine (e.g. in terms of AIC) is still "good enough" in practice, or whether more exhaustive structure searches would be needed. Given the factorial growth of possible vine structures with dimensionality, exhaustive search becomes infeasible from around 7–8 variables onwards. Thus, further research is needed to assess whether heuristic algorithms such as the sequential method of Dißmann *et al.* (2013) provide sufficient accuracy for practical use, or whether additional refinements are necessary for high-dimensional engineering problems.

5.2. ENGINEERING IMPLICATIONS

Across the three studies presented in this dissertation, multivariate dependence modelling was applied to different offshore engineering applications. The motivation for moving towards multivariate (probabilistic) approaches comes from the limitations of conventional deterministic or univariate approaches, which treat variables independently. Multivariate approaches allow the joint behaviour of variables to be captured explicitly. In particular, copula-based multivariate approaches allow to investigate the dependence structure among variables independent of their marginal distributions.

The employed methods increased in complexity over the course of the dissertation: starting with a one-tree vine copula for ecological suitability mapping, progressing to Gaussian Copula-based Bayesian Networks (GCBNs) for mooring line dynamics, and finally implementing 'full' vine copula models for floating photovoltaic structures (see Figure 1.3). At the same time, the spatial resolution of the analyses was narrowed from a large-scale spatial application, to 12 selected study locations, and finally to one location in the south-eastern North Sea, i.e. the reference location throughout this dissertation. Given the differences in both methodological approach and application context, a detailed comparison of the chapters is challenging.

Still, the three applications share common features. All address offshore systems subject to harsh metocean conditions, where balancing safety, costs, and sustainability is critical. All point towards the potential of multi-use concepts, where ecological, hydrodynamic and structural variables interact. Combining these variables in future work could provide holistic assessments of offshore systems, though at increased complexity. Dependence-aware methods enable such analyses, allowing asymmetries and extremes to be accounted for, and supporting scenario-based "what-if" investigations through conditionalisation.

Despite the advantages of multivariate modelling, for instance through copula-based approaches, it is not consistently applied in engineering practice, where it would be beneficial. This raises the question: why is it so difficult to implement such methods in real-world applications? This might be explained through the challenges and pitfalls in applying these methods. Modelling the multivariate dependence of variables with the methods presented in this dissertation can become complex and impractical, and therefore can make interpretation difficult. This depends not only on the chosen method itself and knowledge and experience of applying it, but also on the number of variables and their dependence patterns. All methods considered in this dissertation rely on the concept of bivariate copulas, which requires careful selection of families and structures. These choices are rarely straightforward and often demand domain knowledge or extensive model selection procedures. Furthermore, sufficient sample sizes are essential to avoid overfitting, although what constitutes "sufficient" is application-dependent.

Software and computational aspects

Software packages, for example BANSHEE (Koot *et al.*, 2023; Paprotny *et al.*, 2020) for GCBN or PYVINECOPULIB (Nagler and Vatter, 2025) for vine copula modelling, can facilitate the application of these methods, but their effective use still requires substantial methodological understanding. Despite available software and toolboxes, the methods application can become computationally intensive and therefore can limit their application in real-time or larger-scale applications.

Nowadays, increased computational capacity and new approaches such as machine learning is available. Despite it being a powerful tool for prediction and surrogate modelling, they are less suited for applications where the representation and interpretation of multivariate dependence among variables are required. In offshore spatial planning and structural safety assessment, decisions rely not only on expected responses but on joint probabilities, conditional extremes, and joint dependence patterns such as tail behaviour. Copula-based models are therefore better aligned with these objectives.

Generally, at a minimum, practitioners should always assess whether key variables exhibit strong dependence (i.e. high correlations) and whether this dependence appears symmetric or asymmetric. For example, wind speed and wave height are typically strongly correlated and should not be treated independently, whereas dependencies between precipitation and current velocity may be weaker and thus of lower priority. Neglecting important dependencies can result in severe over- or underestimation of risks, which, in the context of offshore engineering, may result in significant safety and cost implications.

5.2.1. A ROADMAP FOR ENGINEERS IN PRACTICE

To support engineers in selecting suitable approaches, this dissertation proposes a decision roadmap (see Figure 5.2). This flowchart is designed as a practical guideline, aligning the choice of method with the data characteristics, the presence or absence of asymmetries, and the complexity that is feasible in a given application. It balances the advantages and disadvantages of different methods, helping to select a suitable method for a given application. The aim is not to select the universal "best" method, but to enable engineers to make informed choices based on data properties and application needs. Below, the reader is guided through Figure 5.2 in the context of the research of this dissertation.

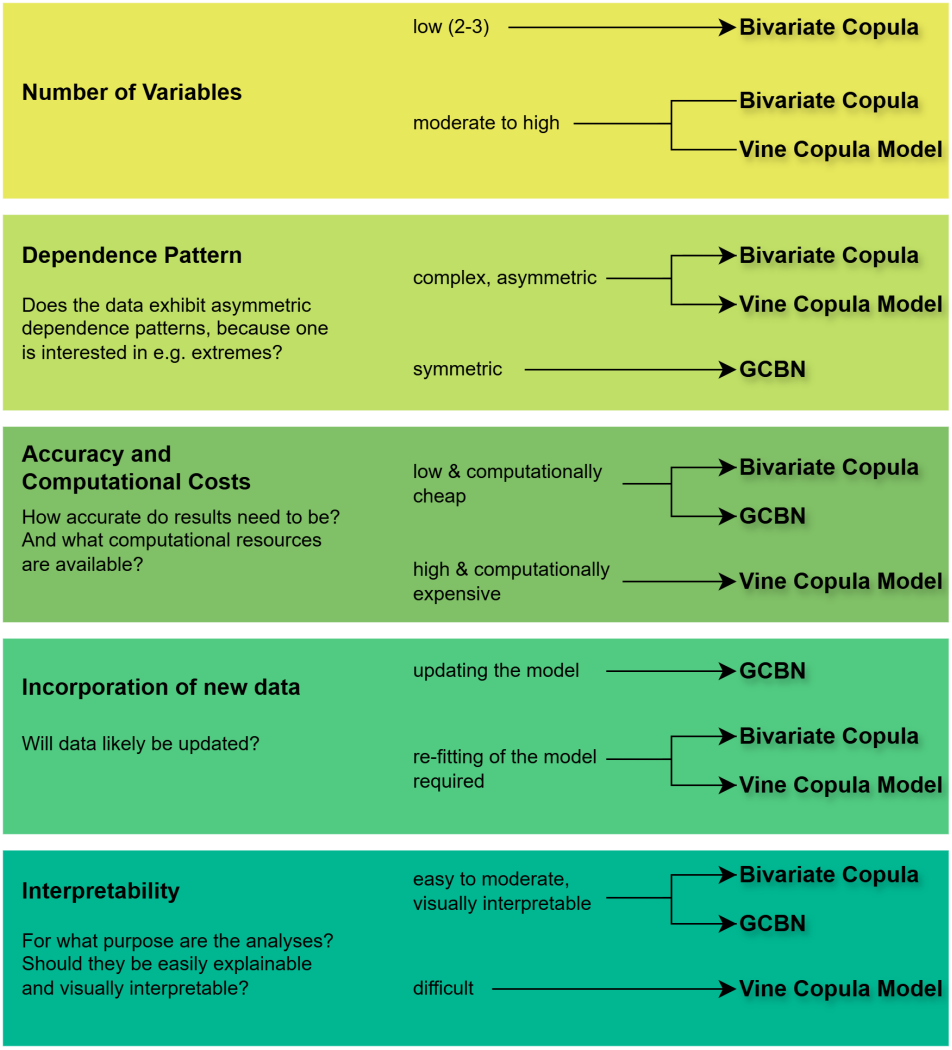


Figure 5.2: Conceptual guideline for selecting modelling methods based on application-specific requirements and limitations.

The **number and type of variables** included depend on the research question and focus of the application, directly influencing which method is most appropriate. For example, in the aquaculture suitability in Chapter 2, only ecological variables were considered, while engineering feasibility aspects (depth, anchoring, maintenance) were excluded. Expanding the analysis could enable investigating large-scale aquaculture impacts on the environment (e.g. on nutrient levels, dissolved oxygen concentration and benthic conditions), biodiversity interactions (e.g. habitat creation or modification), or synergies

in multi-use with e.g. offshore wind farms.

Different to that, Chapter 3 could e.g. be extended by including additional hydrodynamic variables. The analysis is simplified by only considering wave and current impacts in-line with the mooring line. Including the wave and current directions as additional variables, however, could reveal different snap loads due to 3D line dynamics. Further extensions could involve modelling multi-line mooring systems to assess load redistribution under failure scenarios, or moving beyond extreme events to consider long-term fatigue damage of anchors and mooring lines. These examples illustrate that variable choice not only defines the scope of insight but also determines the level of model complexity that is feasible.

The **dependence pattern** among variables is crucial in determining which method to model the dependence among variables is most suitable for a given application. At a minimum, dependence among variables, particularly for pairs with high correlations, should be accounted for. Even a simplified representation of dependence can substantially improve accuracy compared to the assumption of independence.

For the application cases in this dissertation, extreme values are selected as they are most critical for structure safety or survival of species. Therefore, accurate modelling of symmetric, but particularly asymmetric dependencies, such as tail dependence, is essential. The choice of copula family directly influences how these dependence patterns are represented.

Gaussian Copula-based Bayesian Networks (GCBNs) are employed in this dissertation in Chapter 3, however, they rely exclusively on the symmetric Gaussian copula. This makes GCBNs computationally efficient but limits their ability to represent asymmetric dependence structures. In contrast, vine copula models (applied in Chapter 4) offer greater flexibility by allowing any bivariate copula family to be used. Thus, asymmetric dependence patterns can be captured, however, at the cost of increased model complexity. For a detailed comparison of GCBN and vine copula models applied to extreme-value data, see Section 5.1.7. The results demonstrate that, in this case, vine copula models should be preferred, even if the vine structure or copula family selection is not optimally specified.

The first study in this dissertation in Chapter 2 provides further insight into how the strength and pattern of dependence affect model outcomes. While the main focus was to provide suitability maps for present conditions and to assess how a probabilistic approach differs from an independent one, there are several extensions that could further enhance the analysis. The sensitivity of the suitability maps to copula choice was not explicitly addressed, although asymmetric copulas and their rotations were considered. as well as the sensitivity of shifts in thresholds. In addition, potential non-stationarity in dependence, e.g. due to climate change or shifts in nutrient regimes, are not taken into account, i.e. the dependence structure and thus selected copula families could change over time. In that study, the observed correlation strength influenced the selected copula family and thus the resulting probability of meeting defined suitability thresholds. When correlations were strong, the dependence pattern became more pronounced and therefore, the fitted copulas capture it more accurately. Consequently, the difference between dependent and independent modelling reached up to 40 % in threshold exceedance

probabilities offshore.

The spatial patterns of correlation also provided physical insights. For example, correlations between water temperature, T_w, and chlorophyll-a, Chl_a, during the warmer months were generally around -0.6 offshore, where environmental conditions are more stable. This is different to nearshore in the south-western part of the North Sea, where correlations are close to 0, likely due to river inflows and human impacts. The opposite correlation pattern is found for T_w and dissolved inorganic nutrients (DIN, DIP) during winter, with weaker correlations in the central study area compared to nearshore and far offshore regions. This pattern may be explained by the influence of large river outflows, such as those from the Rhine and Meuse. In areas of weaker correlations (between -0.4 and 0), the symmetric t-copula provided the best fit according to Cramèr-van Mises statistic, as tail dependence was not strongly pronounced. In contrast, where correlations were stronger, the asymmetric Gumbel copula performed better offshore, while the symmetric Frank copula was more appropriate nearshore. These findings highlight that dependence structures vary spatially and physically, and that appropriate copula selection, particularly regarding asymmetries, affects the probabilistic estimates, both in the present for ecological application and in other potential use cases.

Accuracy and computational costs are closely linked. GCBNs are easy to apply and computationally efficient but limited to symmetric dependence. Vines, while more complex and computationally intensive, provide robustness in cases where asymmetries and extremes matter. For example, the truncated one-tree vine used in Chapter 2 was relatively simple and efficient but ignored higher-order dependencies. This truncated vine was compared with a 'full' vine in Section 5.1.6. Here, the comparison was done for a 3-dimensional dataset only, however, more variables included results in higher computational and interpretive costs. The choice of method thus depends on the trade-off between the required accuracy per application and available resources.

Another practical consideration is how models handle potential **new input data**, once available. In this case, GCBNs are of advantage, as their structures can often be based on physics and updated with new data without major modifications. This is different to vines, whose structure does not necessarily be based on physics, and thus might require re-specification of the input data changes. However, based on the results of Section 5.1.7, the difference when retaining the same vine structure might be minor, though this may vary depending on the application. Across the applications in this dissertation, environmental data came from large-scale numerical models (particularly the Dutch continental shelf model (DCSM) via D-FLOW FM), but the methods could also be applied to observational or remote sensing data, making them adaptable to any input data.

Interpretability is another crucial factor. GCBNs have an advantage here: their structures often correspond to physical processes and are characterised by the semantics, i.e. it is possible to read from the structure if e.g. two variables are conditionally independent given a third variable. Vines lack such semantics, and their structures are less intuitive, which can hinder communication with stakeholders and decision-makers. This limits their practical usability despite their superior flexibility.

5.2.2. METHODOLOGICAL CONSIDERATIONS FOR ENGINEERING APPLICATIONS

Following the discussion above, the question arises whether multivariate dependence modelling needs to be as complex as the copula-based approaches explored in this dissertation, or whether simpler approximations might often be sufficient in practice. When choosing between specific methods, the trade-offs become evident. GCBNs provide a computationally efficient and relatively simple option, but they rely on Gaussian copulas, which makes them unsuitable for data with asymmetries, as they risk underestimating joint extremes. Different to that, vine copula models can flexibly capture complex and asymmetric dependencies, however, they can become very complex themselves and defining a suitable dependence structure including bivariate copula families is a challenge. Despite several algorithms existing to define a structure for vine copula modelling, the question remains on how good are they and what is sufficient in practice (see Section 5.1.5). Therefore, results from Section 5.1.7 demonstrated that even simple or arbitrarily structured vine models outperform GCBNs in capturing extremes. This suggests that, particularly in the context of extreme values with asymmetric dependence patterns, even "imperfect" vine models may be preferable to GCBNs. Nevertheless, further research is needed to quantify the potential impact of suboptimal choices in vine structure and copula family selection, and what is considered as "good enough" in practical applications. Overall, any method accounting for dependence among variables is preferable to assuming independence, as shown in Chapter 2, where neglecting dependence among only three variables led to deviations of up to 40% in threshold exceedance probabilities.

This discussion underscores the need to move beyond deterministic and purely univariate approaches in offshore engineering design and highlights the importance of integrating such methods into offshore design codes and certification standards (DNV, ISO, etc.). Copula-based methods offer significant improvements in terms of reliability and risk reduction. The decision roadmap described above in Figure 5.2 provides practical guidance for selecting appropriate methods across different applications. Even relatively simple dependence-aware models can meaningfully enhance robustness of offshore engineering analyses and reduce the likelihood of costly failures.

5.3. CONCLUSION

This dissertation demonstrates the application of probabilistic multivariate models based on copulas to a range of emerging offshore technologies, with a particular focus on aquaculture and floating photovoltaic (FPV) systems. The overall objective was to investigate how multivariate copula-based probabilistic models can support spatial planning and structural safety assessment in offshore engineering. This is relevant as sectors such as offshore aquaculture of FPV remain in early stages of development, largely due to the harsh, uncertain, and inherently complex environmental conditions, characteristic for offshore regions. Advanced dependence-aware probabilistic methods, as applied in this dissertation, are able to quantify such uncertainties and to model the joint dependence among environmental and structural variables, some methods even in the case of present asymmetries in joint dependencies. By explicitly representing multivariate uncertainty,

the methods improve tangible decision-making in offshore engineering. Subsequently, the research questions (RQ) formulated in Section 1.6 are addressed.

RQ1: *How does variable dependence affect the accuracy of aquaculture site suitability maps?*

Chapter 2 demonstrates that accounting for dependence can substantially change suitability outcomes. Although it is expected that ecological variables are dependent, the magnitude of the effect in this case is noteworthy: comparing an independence assumption with a one-tree vine copula model using only three variables resulted in differences of up to 40% in the probability of meeting survival or optimal growth thresholds. Such discrepancies mean that areas classified as suitable under an independent model may become unsuitable once dependence is included. This example shows that neglecting dependence can severely under- or overestimate suitability and thus associated risks. Given the complex and often only partially understood behaviour of environmental processes, particularly in offshore settings, assuming independence is therefore an overly simplifying and potentially misleading approach.

RQ2: *How can Gaussian copula-based Bayesian Networks support safety assessment of floating offshore structures by modelling mooring tensions given specific hydrodynamic conditions?*

In Chapter 3, Gaussian copula-based Bayesian Networks (GCBNs) are applied to represent the multivariate dependence among the hydrodynamic variables and tensions in mooring lines. A key advantage of this method is that it imposes no practical limitation on the number of variables, avoiding the need for strict preselection and allowing all (potentially) relevant variables to be incorporated into the analysis. The method of the GCBNs enables conditional predictions, meaning that distributions of tension rates in mooring lines can be obtained for specific hydrodynamic scenarios, for example, conditioning on dominant wave heights or peak periods. This substantially reduces uncertainty in the safety assessment by quantifying how tensions change under given environmental conditions. As a result, GCBNs support e.g. the estimation of exceedance probabilities for critical tension thresholds, or allow for scenario-based risk evaluation, making them a valuable tool for probabilistic safety assessments of floating offshore structures.

RQ3: *How can a vine copula model represent the complex dependence among mixed continuous and ordinal structural and environmental variables of a floating offshore platform?*

In Chapter 4, vine copula models are used to capture the complex multivariate dependence among structural and environmental variables, including both continuous and ordinal variables, such as the discrete number of floaters or structural parameters available only in a few distinct configurations. Vine copulas achieve this by decomposing the joint distribution into a sequence of bivariate copulas, allowing each pair of variables to be modelled with a copula family suited to its

specific dependence characteristics, including asymmetric patterns that cannot be captured by Gaussian-based approaches. Higher-order dependencies are represented through conditional copulas, enabling the model to preserve both pairwise and conditional structures. For ordinal variables, the model accounts for discontinuities in their empirical distributions by incorporating appropriate left and right limits of the cumulative distribution function. Applied to the floating platform, this method allowed the structural-environmental dependence to be represented in detail, even when variables differed in type or exhibited asymmetric interactions. Overall, this method provides a flexible and robust way to model the combined dependence of mixed hydrodynamic and structural variables relevant to analysing the structural response of floating offshore platforms.

ACKNOWLEDGEMENTS

The first years of my PhD trajectory were particularly difficult, starting with the COVID pandemic. Due to the isolation, there were no opportunities to see friends or family abroad. Also, just arriving in a foreign country, getting to know colleagues and building a network is important, but this was not possible either.

At the beginning of my PhD, my ambition and limited experience led me to take on substantial project-related duties alongside my doctoral work. It was difficult to shift the focus back to the PhD *on my own*. However, with the support of a new supervisor I managed to focus on the PhD project again. During the final phase of my PhD, I managed to overcome my last challenge, which lead to another switch of the supervisor(s) once again. Eventually, I finalised my PhD, and for that I want to thank a couple of people who significantly supported me in this, deciding against the formulation of anti-acknowledgements.

I will start with the person who shaped my PhD research the most, my interim supervisor Oswaldo. I am grateful that you became an actual supervisor for me and believed in me when my confidence in myself was nearly gone. I want to thank you for the very good relationship we had. You taught me a lot and managed to get me excited for probability theory and particularly copula models. Despite serious disagreements towards the end, I am very grateful you took care of me and supported me until that moment.

I could not have finalised this PhD journey without the support and dedicated efforts of Andrei. When I came closest to giving up on the PhD near its final stage, you were the one who looked after me. You were always realistic and optimistic with your advices, which was exactly what I needed. Thank you for that.

Next, I would like to express my genuine thanks to Tiziana. After your arrival as a new head of the Department of Hydraulic Engineering, you were consistently committed to finding solutions and ways forward together with Andrei, helping me bring the final part of this thesis to completion.

Special thanks go of course also to Patricia. You helped me significantly in catching up on probability theory and writing my first paper, even though I was not officially in your department. Only then I realised for the first time that I might be able to pursue a PhD after a couple of tough years full of self-doubts. Over the years you became a good friend and I am extremely grateful for that.

For the last two chapters of my thesis, words cannot express my gratitude to Shagun. We had a very close collaboration in the past years, which I enjoyed a lot. You helped realign my research with my background in offshore engineering and extensively helped me to finalise this thesis.

Paul, as my department head for most of my PhD period at Deltares, was one of the people whose support was essential to the success of this work. You supported me in finding an effective balance between my PhD work and my responsibilities at Deltares, including helping to arrange the time needed to complete this thesis. Besides, you acted like a great coach throughout the years, providing advice and guidance for personal and professional development.

Next, I want to mention Anna and Elias, two close PhD fellows at Deltares. I am grateful to both of you for going through this together, sharing experiences, and supporting each other along the way. I extend my gratitude to my colleagues at Deltares: Luis, Sonja, Sonia, Nathalie, David and many more from the DSW and RAP departments. You made the office days so much more enjoyable and joint lunches or walks across the campus were always welcome breaks. Beyond that, I would also like to thank colleagues who helped me content-wise, particularly Julien and Lauriane, e.g. by providing training for Delft3D modelling, provided data or fruitful discussions.

Of course I also would like to mention my PhD fellows at TU Delft: Dorette, Gina, Vidya, Paulina, Marcel and Miguel. You all welcomed me warmly at the other side of the never-finished tram line. Thank you for the catch-ups we had, whether it was about copulas, concerts or knitting. You joined me through ups and downs. And thank you, Paulina, for reading and editing my *samenvatting*.

Through one-on-one sessions, I learned a lot about the mathematics of copulas and particularly vine copula models from Özge, which helped finalising the thesis. Thank you for that.

Several times through my PhD trajectory, I approached Ada as a confidential advisor. Talking with you always provided a safe space. Thank you for always being helpful in finding solutions for the situations I found myself in and your support beyond that.

I stopped counting how often I heard 'work-life-balance' throughout these years. Despite the 'life' came much shorter for me in the past years, a couple of people were crucial and supported me from that side finalising the PhD. First, I would like thank my friends Birte, Tomke, Florian, Wafa and Kathrin, who always needed to listen to my lamentations and positive and negative stories from the PhD.

Next I am grateful to my brother. Aike, thank you for always caring about my well-being, even when it meant being honest with me in ways that were sometimes hard to hear but ultimately exactly what I needed.

I would like to extend my deepest gratitude to my parents, who made this education even possible by financing my studies and believing in me when I didn't. You continued to support me during my PhD in your own unique ways, always providing a safe and comforting place at home. From Mama, I learned how to knit, while from Papa, already as a child, I gained an appreciation for high-quality speakers and the joy of listening to music. Both have been sources of energy and calm, without them I could not have made it through the past years. And of course: thanks, Mama, for the painting of my thesis cover!

Finally, my deepest appreciation to my husband, Tim; words cannot fully express my gratitude. You always supported me throughout these years, despite being (rightly) more than annoyed by my lamentations. You even held yourself back with your enthusiasm in your PhD towards me, to make sure I do not feel worse than I already did. You have always been the best (unofficial) supervisor for me, not only providing guidance and *stimulating* discussions, where possible, but also much beyond that: taking care of myself more than I was able to, making sure I take the best decisions for myself and my well-being. With you by my side, I never felt like a *lone cowboy*.

A

APPENDIX

A.1. TIMESERIES OF THE STUDIED VARIABLES

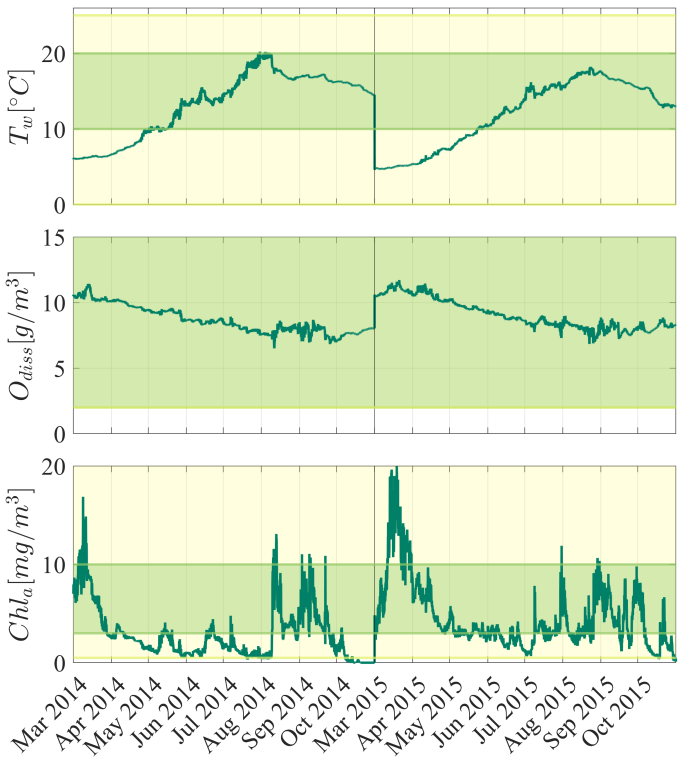


Figure A.1: Time series of the relevant variables for the blue mussels, where the green background marks the optimal growth conditions and the yellow parts mark the critical survival conditions.

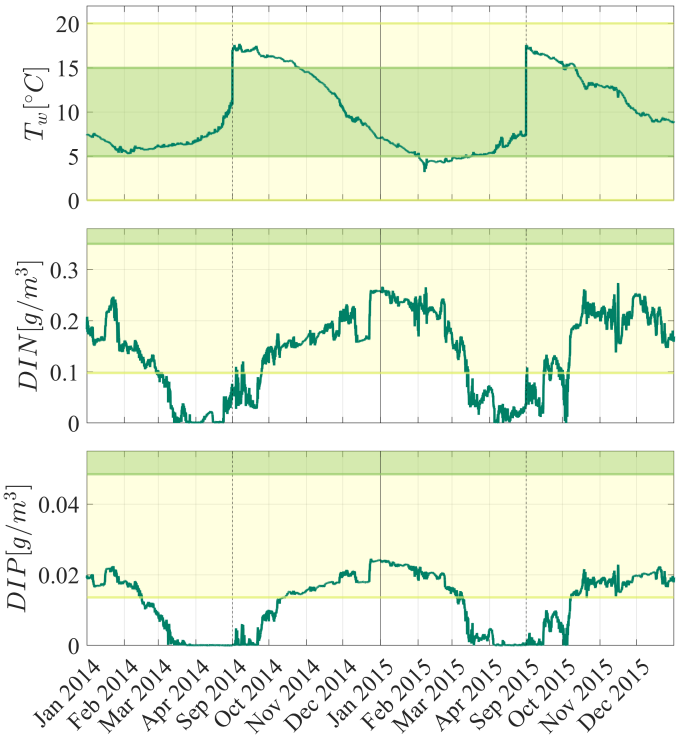


Figure A.2: Time series of the relevant variables for the sugar kelp, where the green background marks the optimal growth conditions and the yellow parts mark the critical survival conditions.

A.2. TAIL DEPENDENCE

Here, the tail dependencies are described.

The upper tail dependence parameter λ_U can be described as follows:

$$\lambda_U = \lim_{u \to 1} P\left[X > F_X^{-1}(u) \mid Y > G_Y^{-1}(u)\right], \quad \lambda_U \in [0,1] \tag{A.1}$$

While the lower tail dependence parameter λ_L is determined by:

$$\lambda_L = \lim_{u \to 0} P\left[X \leq F_X^{-1}(u) \mid Y \leq G_Y^{-1}(u)\right], \quad \lambda_L \in [0,1] \tag{A.2}$$

Where X and Y are continuous random variables with distribution functions $|F|$ and $|G|$. If $\lambda_U > 0$ then X and Y are upper tail dependent while if $\lambda_L > 0$, the random variables are lower tail dependent (Nelsen, 2006).

A.3. CUMULATIVE DISTRIBUTION FUNCTIONS OF BIVARIATE PARAMETRIC COPULAS

Below, the bivariate cumulative distribution function (cdf) of the considered copula families are given (Czado, 2019; Joe, 2015; Nelsen, 2006).

The cdf for the bivariate Gaussian copula is given by

$$C(u_1, u_2; \rho_{12}) = \Phi_2\left(\Phi^{-1}(u_1), \Phi^{-1}(u_2);\ \rho_{12}\right) \tag{A.3}$$

where Φ_2 is the joint cdf of a bivariate normal distribution, Φ^{-1} is the inverse cdf of a univariate standard normal distribution and ρ_{12} being the correlation coefficient.

The cdf for the bivariate t-copula is given by

$$C(u_1, u_2, \nu_{12}, \rho_{12}) = \int_0^{u_1}\int_0^{u_2} \frac{t\left(T_{\nu_{12}}^{-1}(v_1), T_{\nu_{12}}^{-1}(v_2); \nu_{12}, \rho_{12}\right)}{t_{\nu_{12}}\left(T_{\nu_{12}}^{-1}(v_1)\right) t_{\nu_{12}}\left(T_{\nu_{12}}^{-1}(v_2)\right)} dv_1\, dv_2 \tag{A.4}$$

where $T_{\nu_{12}}^{-1}(\cdot)$ is the inverse of the cdf of a univariate t-distribution with ν_{12} degrees of freedom, $t_{\nu_{12}}(\cdot)$ is the pdf of a univariate t-distribution with ν_{12} degrees of freedom and $t(\cdot,\cdot,\nu_{12},\rho_{12})$ is the bivariate pdf of the t-distribution.

The cdf for the bivariate Frank copula is given by

$$C(u_1, u_2, \delta_{12}) = -\frac{1}{\delta_{12}}\ln\left(\frac{1 - e^{-\delta_{12}} - (1 - e^{-\delta_{12}u_1})(1 - e^{-\delta_{12}u_2})}{1 - e^{-\delta_{12}}}\right) \tag{A.5}$$

where $-\infty < \delta_{12} < \infty$ is the copula parameter, describing the dependence, while $\delta_{12} \to 0^+$ means independence.

The cdf for the bivariate Clayton copula is given by

$$C(u_1, u_2, \delta) = \left(u_1^{-\delta} + u_2^{-\delta} - 1\right)^{\frac{1}{\delta}} \tag{A.6}$$

where $0 < \delta < \infty$ is the copula parameter, describing the dependence, while independence occurs with $\delta \to 0$.

The cdf for the bivariate Gumbel copula is given by

$$C(u_1, u_2, \delta) = exp\left[-\left\{(-\ln u_1)^{\delta} + (-\ln u_2)^{\delta}\right\}^{\frac{1}{\delta}}\right] \tag{A.7}$$

where $\delta >= 1$ is parameter of dependence, while independence is described with $\delta = 1$.

A.4. (INVERSE) CONDITIONAL DISTRIBUTION FUNCTIONS

Conditional distribution functions associated with a bivariate copula are used in this study. These are usually referred to as *h*-functions by (Aas *et al.*, 2009). For a bivariate copula C_{XY}, the *h*-function is defined as:

$$h(x, y, \theta_{xy}) = F(x|y) = \frac{\partial C_{XY}(x, y, \theta_{xy})}{\partial y} \tag{A.8}$$

with X being the conditioned variable, Y being the conditioning variable, and θ_{xy} being the set of parameters for the copula of the joint distribution function of x and y. It represents the cdf when X and Y are uniform.
Its inverse with respect to the first variable X, $h^{-1}(X, Y, \theta_{xy})$, is equivalent to the inverse of the conditional distribution function. For the copulas used here (Gaussian, Gumbel, Clayton, Frank, and t-copula), the *h*-function is given by an explicit analytical expression that can be analytically inverted for all pair-copulas except for Gumbel, which requires numerical inversion (Aas *et al.*, 2009; Czado, 2019). Below, the inverse conditional distribution functions (inverse *h*-functions) of the respected copula families are given, as well as the *h*-function for the Gumbel copula (Aas *et al.*, 2009; Joe, 2015).

The inverse conditional distribution function (hinv) for the Gaussian copula is given by

$$\mathrm{h}_{12}^{-1}(u_1, u_2, \rho_{12}) = \Phi\left\{\Phi^{-1}(u_1)\sqrt{1-\rho_{12}^2} + \rho_{12}\Phi^{-1}(u_2))\right\} \tag{A.9}$$

where ρ_{12} is the parameter of the copula, Φ is the cdf of the standard normal distribution and Φ^{-1} its inverse.

The inverse conditional distribution function (hinv) for the t-copula is given by

$$\begin{aligned} &\mathrm{h}_{12}^{-1}(u_1, u_2, \rho_{12}, \nu_{12}) = \\ &t_{\nu_{12}} \times \left\{\mathrm{t}_{\nu_{12}+1}^{-1}(u_1)\sqrt{\frac{\left(\nu_{12} + (\mathrm{t}_{\nu_{12}}^{-1}(u_2))^2\right)(1-\rho_{12}^2)}{\nu_{12}+1}} + \rho_{12}\mathrm{t}_{\nu_{12}}^{-1}(u_2)\right\} \end{aligned} \tag{A.10}$$

where ρ_{12} is the parameter of the copula, ν the degree of freedom, and $t_{\nu_{12}}$ the standard cdf of the t-distribution and $t_{\nu_{12}}^{-1}$ its inverse function.

The inverse conditional distribution function (hinv) for the Frank copula is given by

$$\mathrm{h}_{12}^{-1}(u_1, u_2, \delta_{12}) = -\log\left\{1 - \frac{1-e^{-\delta_{12}}}{(\mathrm{u}_1^{-1}-1)e^{-\delta_{12}u_2}+1}\right\}/\delta_{12} \tag{A.11}$$

where δ_{12} is the parameter of the copula.

The inverse conditional distribution function (hinv) for the Clayton copula is given by

$$h_{12}^{-1}(u_1, u_2, \delta_{12}) = \left\{ \left(u_1 u_2^{\delta_{12}+1} \right)^{-\frac{\delta_{12}}{\delta_{12}+1}} + 1 - u_2^{-\delta_{12}} \right\}^{-1/\delta_{12}} \tag{A.12}$$

where $0 < \delta_{12} < \infty$ is the parameter of the copula, controlling the dependence.

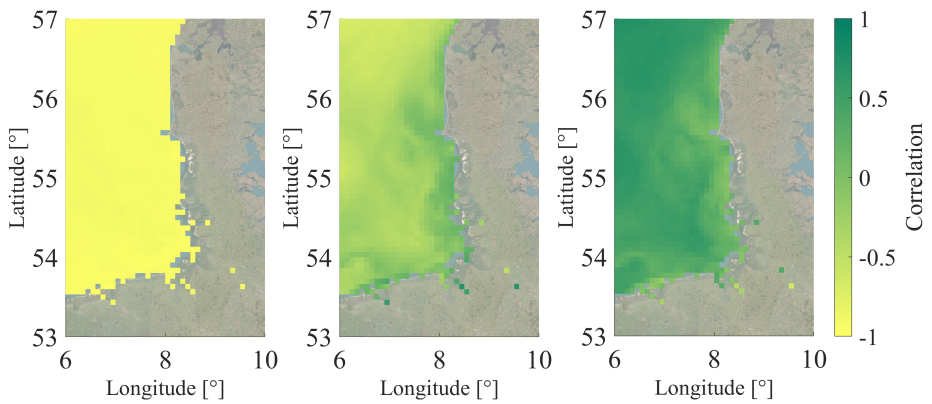
The conditional distribution function for the Gumbel copula is given by

$$h(u_1, u_2, \delta_{12}) = C_{12}(u_1, u_2) \frac{1}{u_2} (-\log u_2)^{\delta_{12}-1} \times \left\{ (-\log u_1)^{\delta_{12}} + (-\log u_2)^{\delta_{12}} \right\}^{1/\delta_{12}-1} \tag{A.13}$$

where $C_{12}(u_1, u_2)$ is the copula, as defined in Equation (A.7) and $\delta_{12} \geq 1$ is the parameter of the copula, controlling the dependence. Note that it is necessary to obtain the inverse of the h-function numerically.

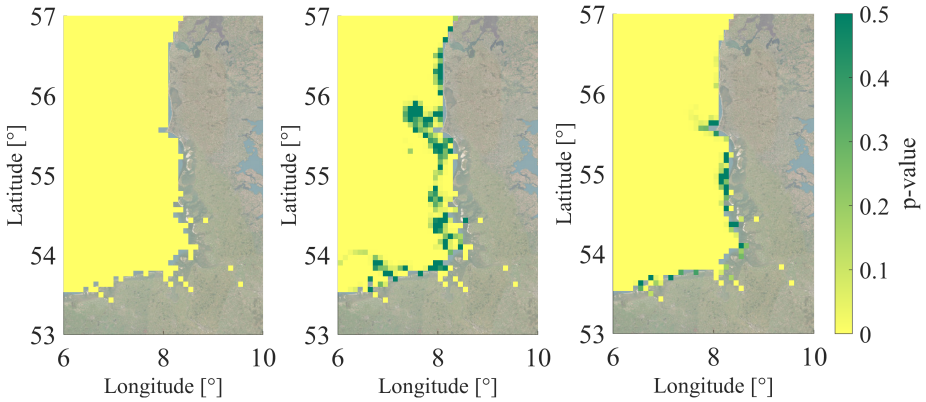
A.5. ADDITIONAL RESULTS FOR THE BLUE MUSSEL CASE

A.5.1. CORRELATION ANALYSIS



(a) T_w and O_{diss}. **(b)** T_w and Chl_a. **(c)** O_{diss} and Chl_a.

Figure A.3: Correlations between the selected three variable pairs for the blue mussel case.



(a) T_w and O_{diss}. **(b)** T_w and Chl_a. **(c)** O_{diss} and Chl_a.

Figure A.4: *p*-values for the correlations between the selected three variable pairs for the blue mussel case.

A.5.2. ANGLE OF ROTATION FOR GUMBEL AND CLAYTON COPULA

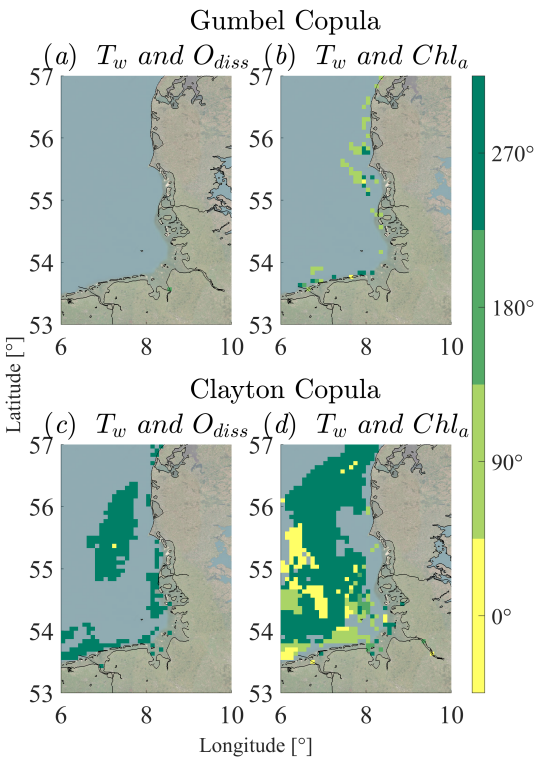


Figure A.5: Rotation of the Gumbel Copula in (*a*) and (*b*) and rotation of the Clayton copula in (*c*) and (*d*) for the variable pairs of T_w and O_{diss}, and T_w and Chl_a, respectively

A.5.3. PROBABILITY CALCULATION

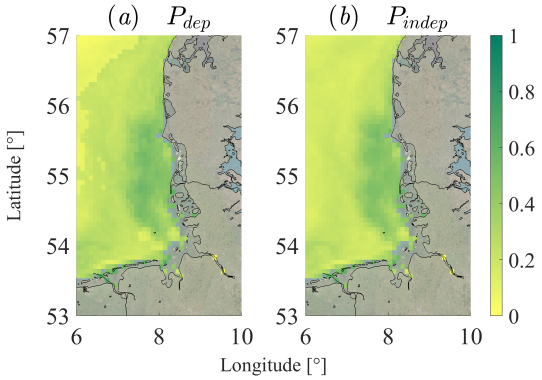


Figure A.6: The spatial probabilities with (*a*) accounting for dependence (P_{dep}) and (*b*) without (P_{indep}), describing the suitability of blue mussel *Mytilus edulis* for the optimal growth conditions

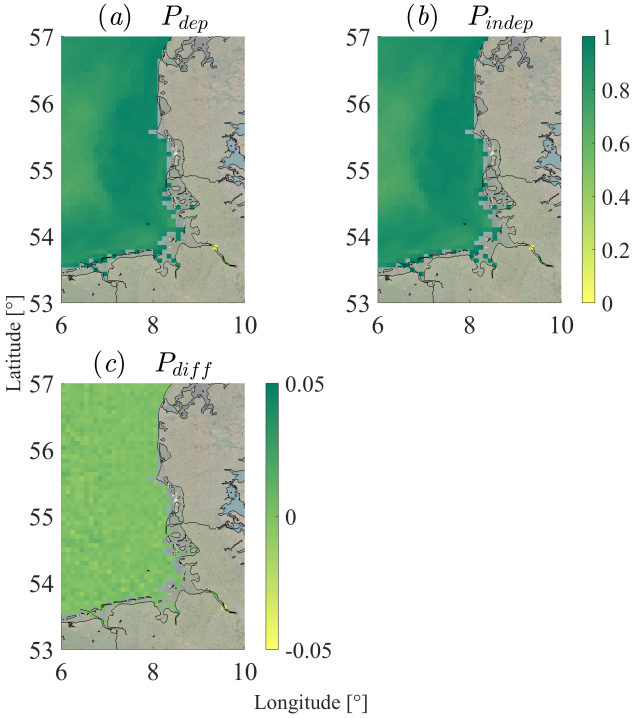


Figure A.7: The spatial probabilities with (*a*) considering (P_{dep}), (*b*) without considering dependence (P_{indep}), describing the suitability of blue mussel *Mytilus edulis* for the critical growth conditions; while in (*c*), the differences between the two approaches is displayed

A.6. ADDITIONAL RESULTS FOR THE SUGAR KELP

A.6.1. EXAMPLE OF THE SAMPLED DATA

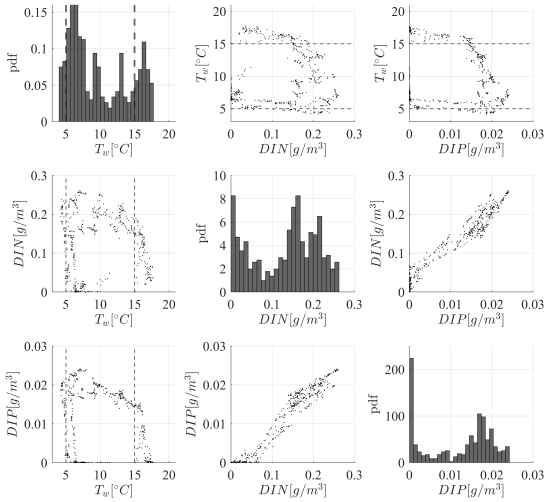


Figure A.8: Overview of sampled data for sugar kelp at the location of the FINO3 research platform together with the optimal limits from Table 2.1, while on the diagonal the empirical probabilistic distribution functions (pdf's) are shown.

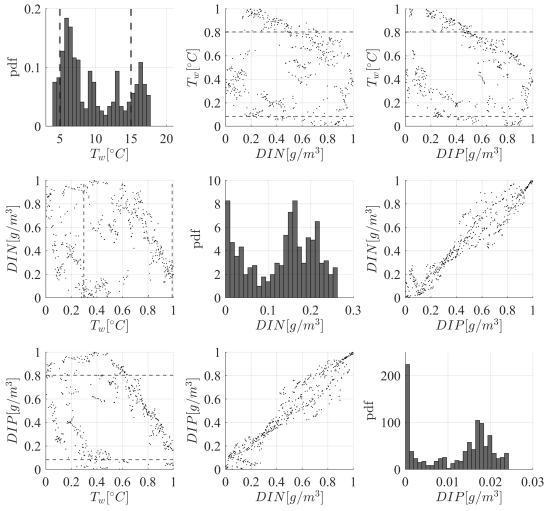
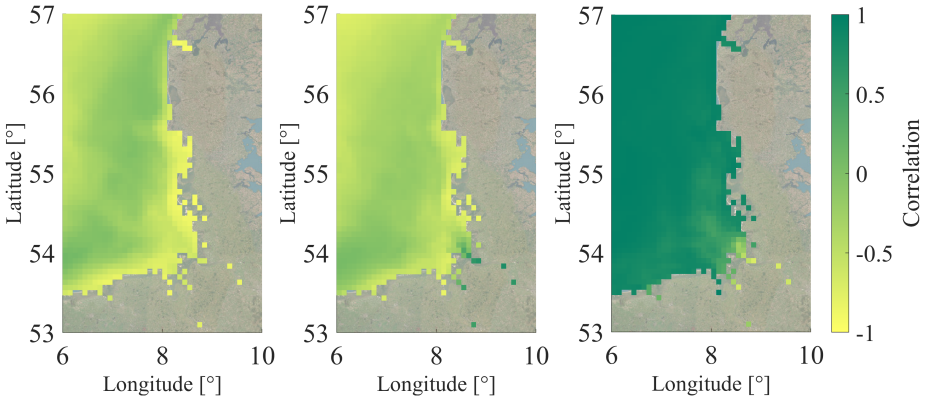


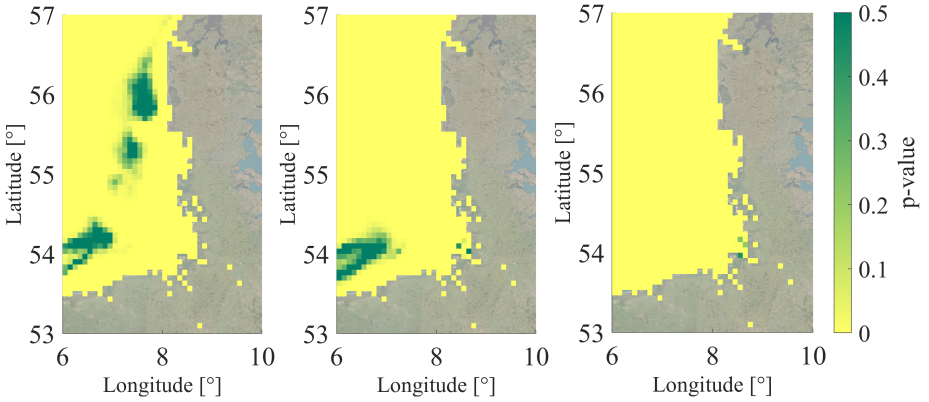
Figure A.9: Overview of uniform observations of the sampled data for sugar kelp at the location of the FINO3 research platform together with the empirical non-exceedance probabilities of the optimal limits (see Table 2.2), while on the diagonal the univariate empirical probabilistic distribution functions are shown together with the optimal limits from Table 2.1.

A.6.2. CORRELATION ANALYSIS



(a) T_w and DIN. **(b)** T_w and DIP. **(c)** DIN and DIP.

Figure A.10: Correlations between the selected three variable pairs for the sugar kelp.



(a) T_w and DIN. **(b)** T_w and DIP. **(c)** DIN and DIP.

Figure A.11: p-values for the correlations between the selected three variable pairs for the sugar kelp.

A.6.3. COPULA FITTING RESULTS

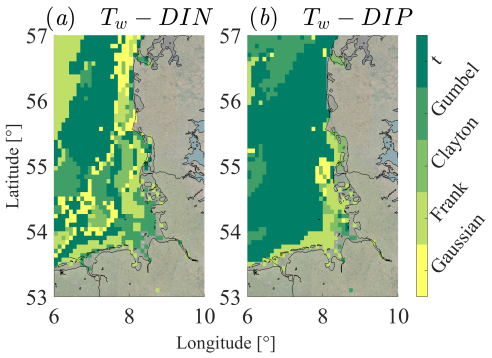
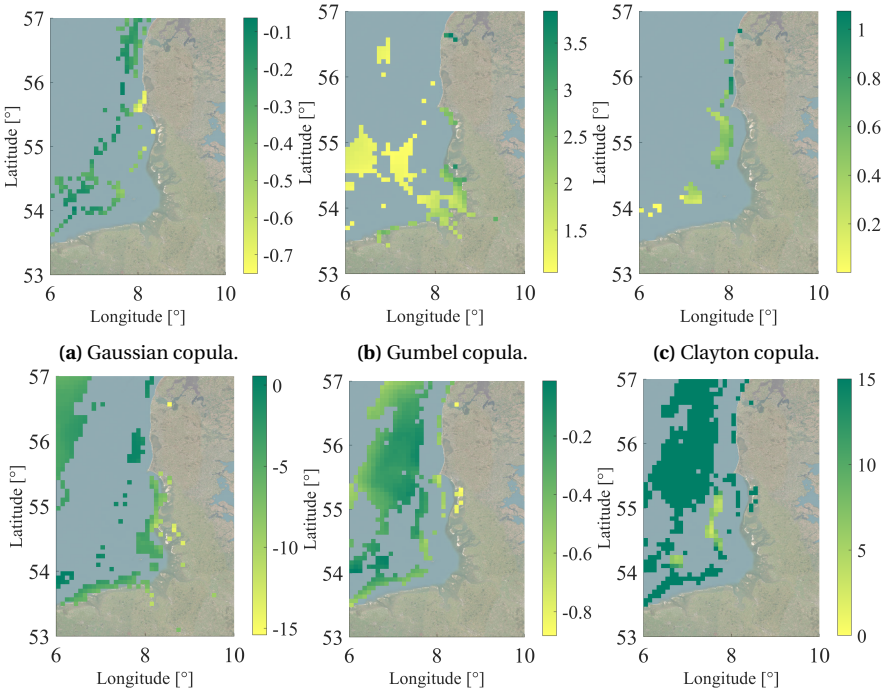


Figure A.12: Best fitting copulas across the AOI for both variable pairs for the sugar kelp: in (*a*) T_w and DIN and in (*b*) T_w and DIP

A.6.4. COPULA PARAMETER AND ANGLE OF ROTATION FOR GUMBEL AND CLAYTON COPULA



(a) Gaussian copula. **(b)** Gumbel copula. **(c)** Clayton copula.

(d) Frank copula. **(e)** t-copula (ρ). **(f)** t-copula (ν).

Figure A.13: The spatial copula parameter for the respected copula families for the variable pair of T_w and DIN for the sugar kelp.

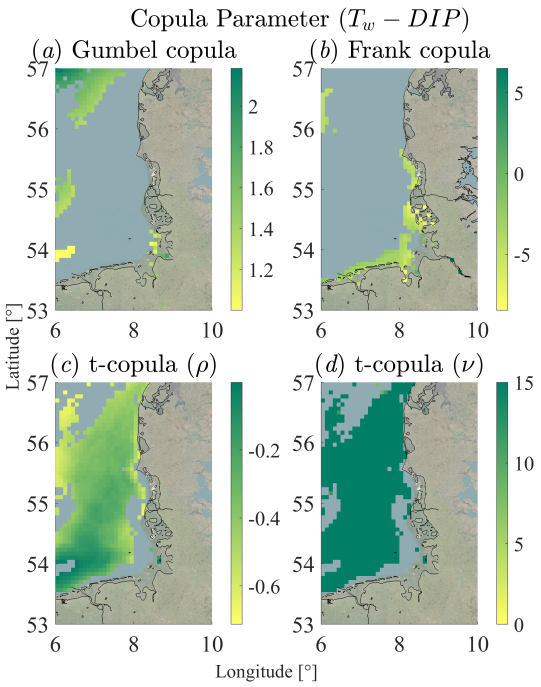
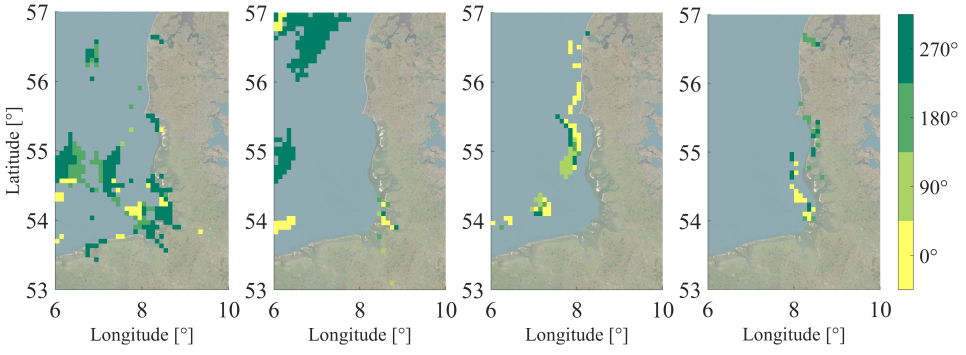


Figure A.14: The spatial copula parameter for the Gumbel, Frank and t-copula for the variable pair of T_w and DIP for the sugar kelp



(a) Gumbel copula for T_w and DIN. **(b)** Gumbel copula for T_w and DIP. **(c)** Clayton copula for T_w and DIN. **(d)** Clayton copula for T_w and DIP.

Figure A.15: Rotation of the Gumbel and Clayton copula for the sugar kelp.

A.6.5. PROBABILITY CALCULATION

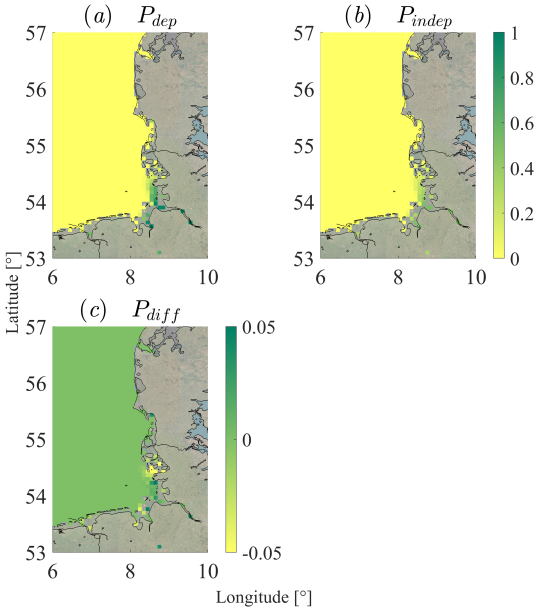
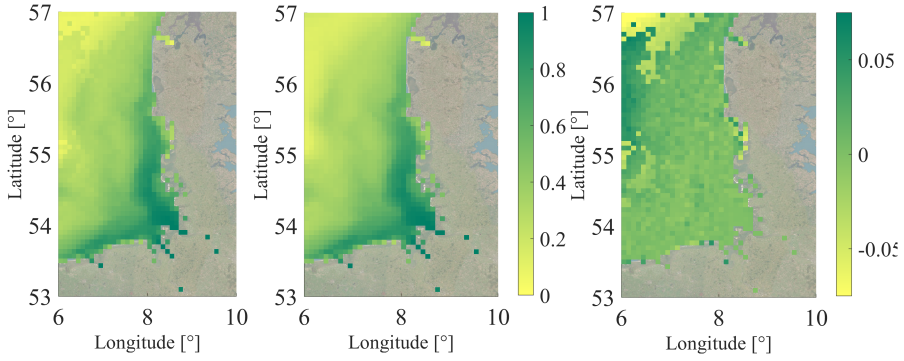


Figure A.16: The spatial probabilities with (*a*) considering dependence (P_{dep}), (*b*) without considering dependence (P_{indep}), describing the suitability of the sugar kelp *Saccharina latissima* for the optimal growth conditions and (*c*) the difference between the two approaches



(a) Considering dependence (P_{dep}).

(b) Without considering dependence (P_{indep}).

(c) Differences between the two approaches.

Figure A.17: Spatial probabilities describing the suitability of blue mussel *Saccharina latissima* for the critical growth conditions, together with the differences between the two approaches.

B

APPENDIX

B.1. OVERVIEW OF WAVE CHARACTERISTICS

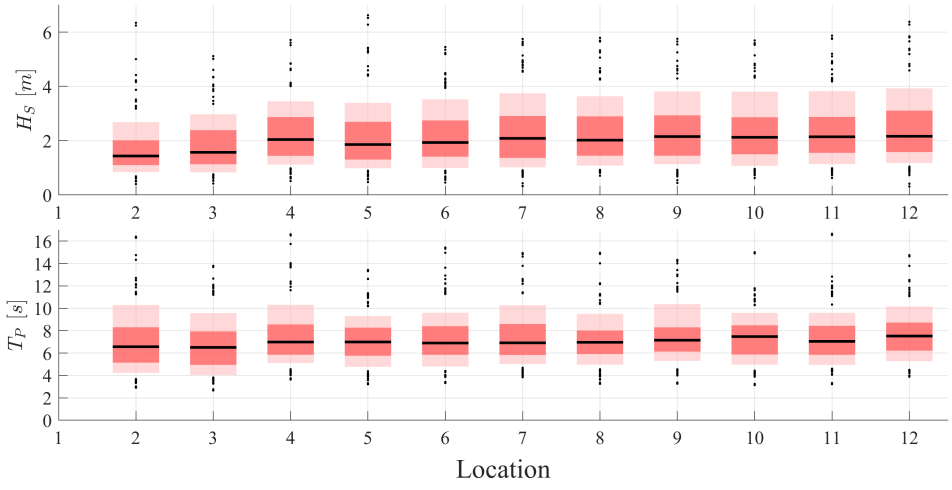


Figure B.1: Overview of significant wave height (H_S) and peak wave period (T_P) distributions for each location l. Light red shading represents the 10th–90th percentile range, dark red shading represents the 25th–75th percentile range, and the median is indicated by a black horizontal line.

B.2. EXAMPLE HYDRODYNAMIC MODEL DATA

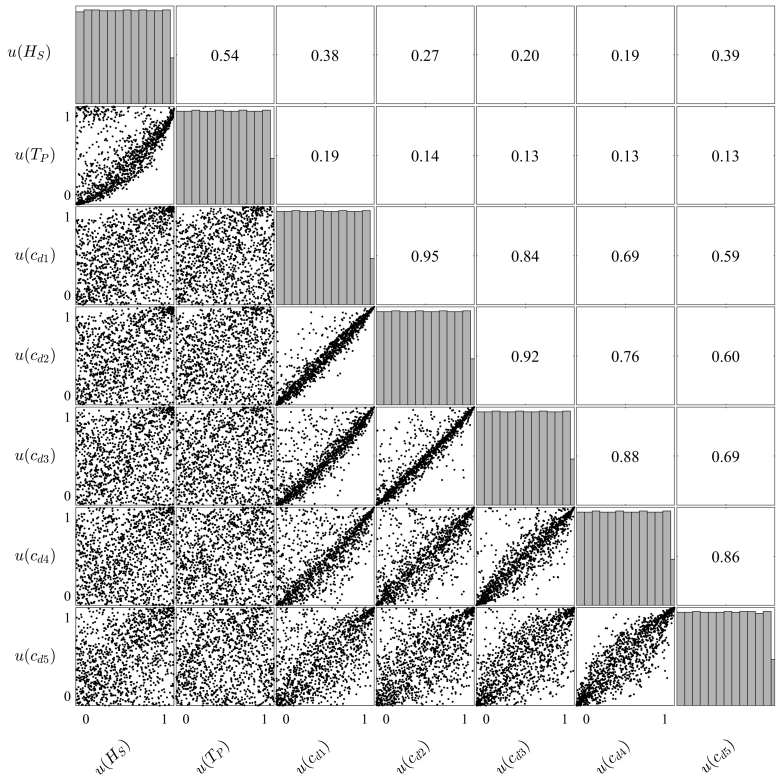
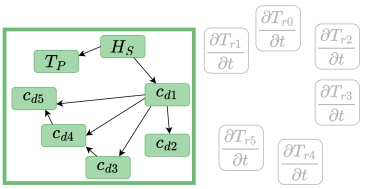
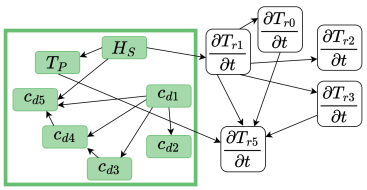


Figure B.2: Pairwise scatter plots of the data from the hydrodynamic model (DCSM) transformed to unit space for RefLoc 4 are presented in the lower-left part, together with their one-dimensional uniform margins on the diagonal and rank correlation coefficients above the diagonal. This hydrodynamic data serves as input for $BN_{h,l}$.

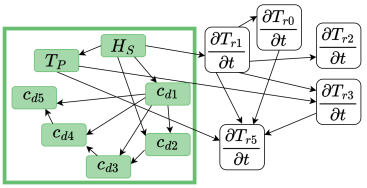
B.3. GCBN STRUCTURES FOR EACH STUDY LOCATION



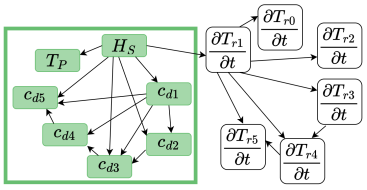
(a) GCBN structure for location 1



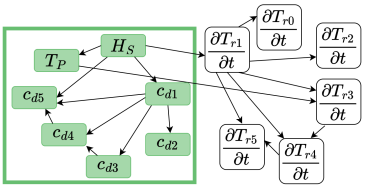
(b) GCBN structure for location 2



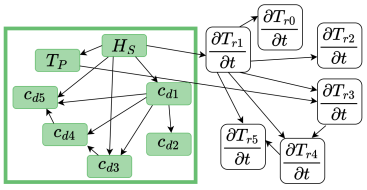
(c) GCBN structure for location 3



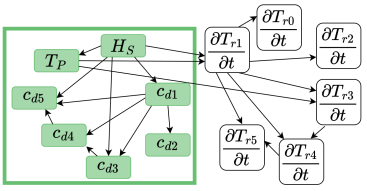
(d) GCBN structure for location 4



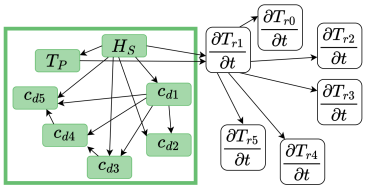
(e) GCBN structure for location 5 to 7



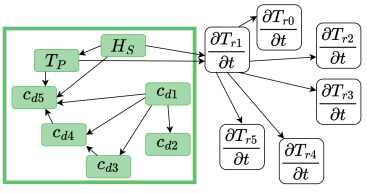
(f) GCBN structure for location 8



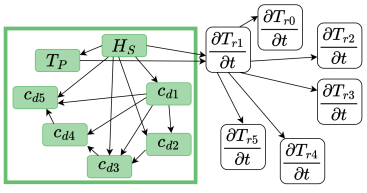
(g) GCBN structure for location 9



(h) GCBN structure for location 10



(i) GCBN structure for location 11



(j) GCBN structure for location 12

Figure B.3: Structure of the GCBNs for each studied location l. The hydrodynamic $BN_{h,l}$, marked in green, are extended by tension rates to receive $BN_{ht,l}$, presented here.

B.4. COMPARING CUMULATIVE DISTRIBUTION FUNCTIONS

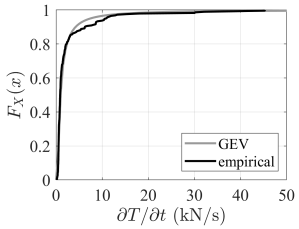


Figure B.4: Comparison of the empirical cumulative distribution function (cdf) for the tension rate at the anchor, derived from $BN_{ht,l}$ with $n = 1{,}000{,}000$ samples, against the cdf of a fitted Generalised Extreme Value (GEV) distribution for RefLoc 4.

B.5. ADDITIONAL RESULTS ON TENSION RATE PERCENTILES

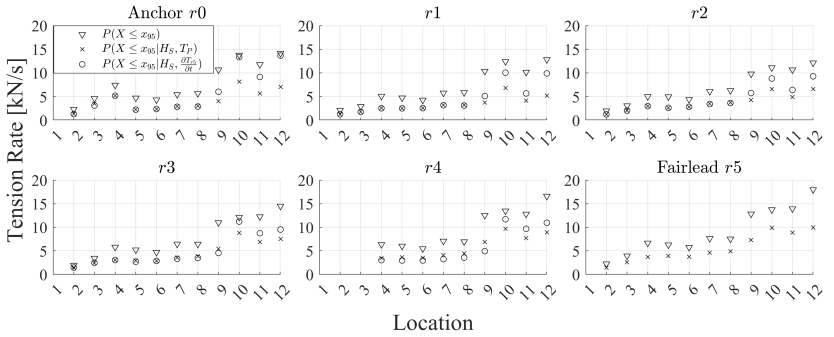


Figure B.5: The 95th-percentile design values of tension rates are presented for six rope positions (r), including the anchor and fairlead, for three conditionalising cases.

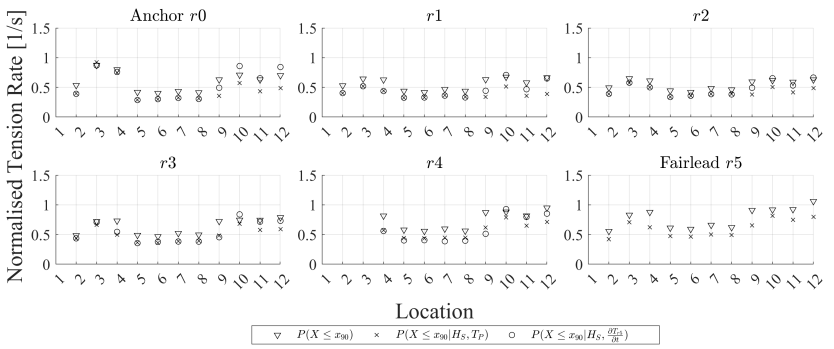
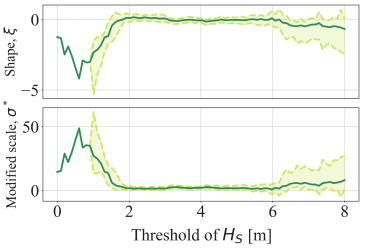


Figure B.6: The 90th-percentile design values of tension rates, normalised by their static pretension, are presented for six rope positions (r), including the anchor and fairlead, for three conditionalising cases.

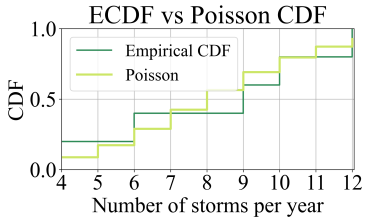
C

APPENDIX

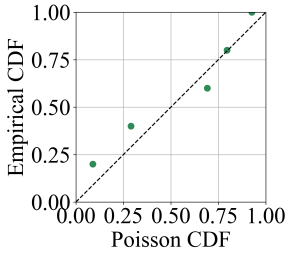
C.1. SELECTION OF EXTREMES



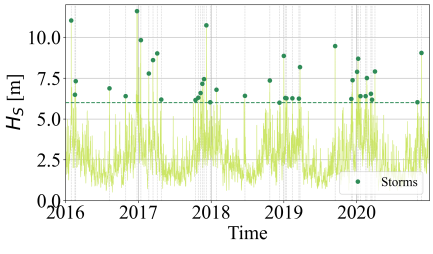
(a) Parameter stability plot for a declustering time of 72h for H_S.



(b) Comparison of empirical CDF and Poisson CDF on the number of storms per year.



(c) Comparison of empirical and estimated non-exceedance probability.



(d) Selected extremes of H_S.

Figure C.1: Results for the selection of extremes of H_S for a declustering time of 72 h.

Table C.1: Number of selected extremes of H_S and corresponding T_P, together with the parameter Λ of the Poisson distribution and their p-values for different declustering times.

Declustering Time [h]	Number of selected extremes	Λ	p-value
12	57	11.014	0.026
24	55	10.917	0.028
48	47	6.348	0.169
72	41	5.292	0.2586

C.2. EXAMPLE OF FPV MODEL OUTPUT

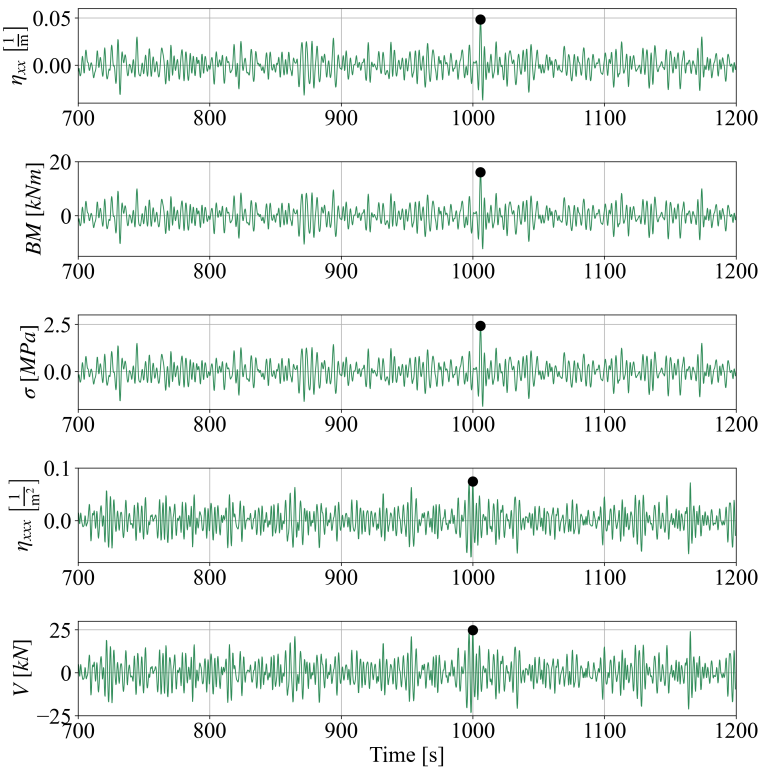


Figure C.2: Time series of FPV model output η_{xx} and derived quantities. η_{xx} is used to calculate bending moment (BM) and longitudinal stress (σ), while η_{xxx} is used to compute shear force (V). The total time series length is 20 min, and circles (•) indicate the selected maxima for each series.

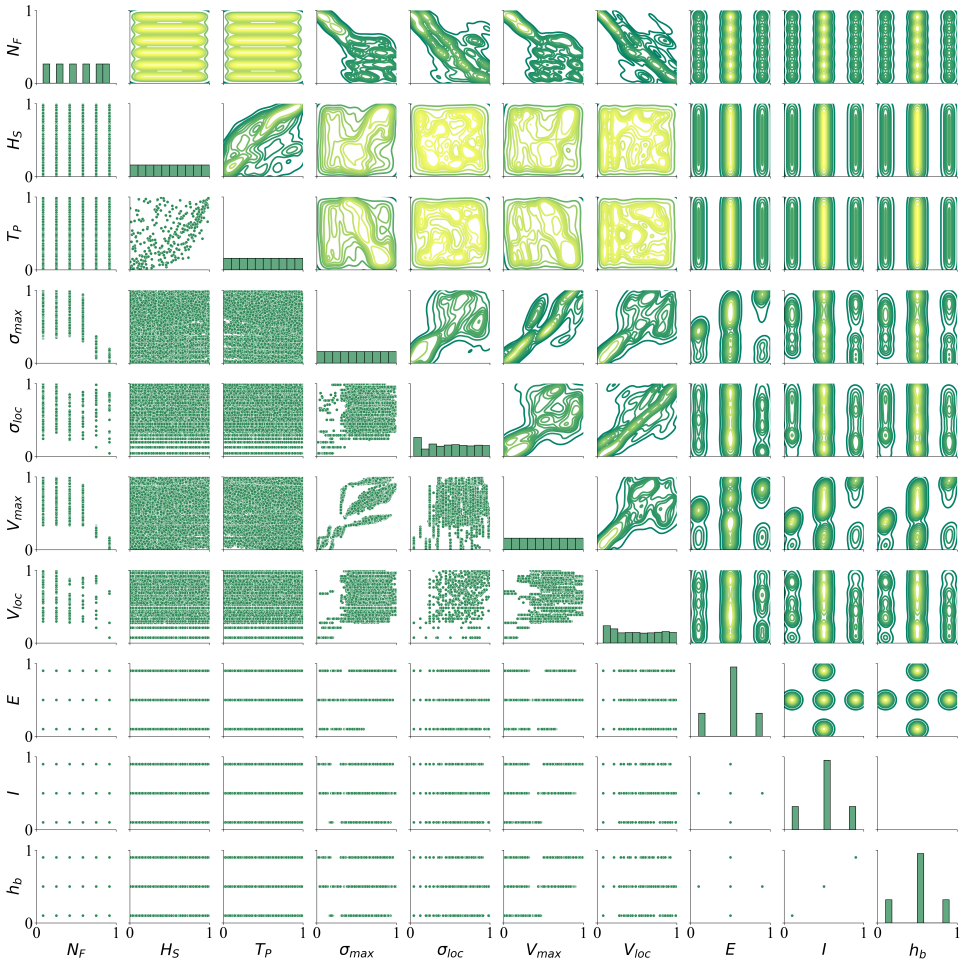
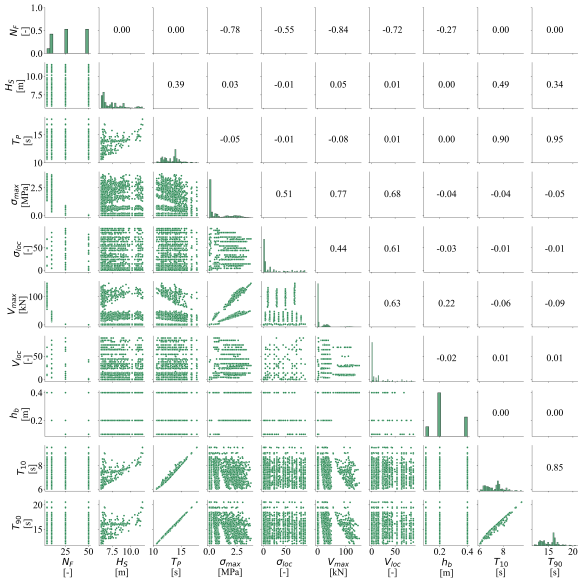
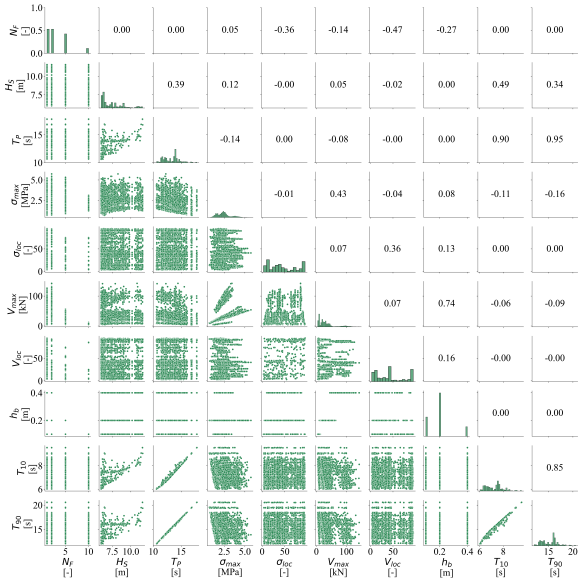


Figure C.3: Pairwise scatter plots of the studied variables transformed to unit space are presented in the lower-left part, together with their one-dimensional uniform margins on the diagonal and their densities as contours above the diagonal.

C.3. DATA PER REGIME



(a) Regime 1.



(b) Regime 2.

Figure C.4: Pairwise scatter plots of the studied variables in their physical units are presented for Regime 1 and 2 in the lower-left part, together with their empirical marginal distributions on the diagonal and Kendall's τ correlations, computed from the raw data in physical space, above the diagonal.

BIBLIOGRAPHY

Aas, K., Czado, C., Frigessi, A., & Bakken, H. (2009). Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics*, *44*, 182–198. https://doi.org/10.1016/j.insmatheco.2007.02.001

Agarwal, S., Gomez, S., & Colomes Gene, O. (2025). *Dynamic finite-strain analysis of mooring lines based on tangential differential calculus* [preprint].

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723. https://doi.org/10.1109/TAC.1974.1100705

Alcañiz, A., Agarwal, S., Tiwald, P., Isabella, O., Ziar, H., & Colomés, O. (2025). Assessment of power losses and structural response of offshore floating solar platform. *Energy*, *335*, 137656. https://doi.org/https://doi.org/10.1016/j.energy.2025.137656

Azevedo, I. C., Duarte, P. M., Marinho, G. S., Neumann, F., & Sousa-Pinto, I. (2019). Growth of saccharina latissima (laminariales, phaeophyceae) cultivated offshore under exposed conditions. *Phycologia*, *58*, 504–515. https://doi.org/10.1080/00318884.2019.1625610

Barros, B., Conde, B., Riveiro, B., & Morales-Nápoles, O. (2024). Gaussian copula-based bayesian network approach for characterizing spatial variability in aging steel bridges. *Structural Safety*, *106*, 102403. https://doi.org/https://doi.org/10.1016/j.strusafe.2023.102403

Bedford, T., & Cooke, R. (2001). Probability density decomposition for conditionally dependent random variables modeled by vines. *Annals of Mathematics and Artificial Intelligence*, *32*(1-4), 245–268. https://doi.org/10.1023/A:1016725902970

Bedford, T., & Cooke, R. (2002). Vines - a new graphical model for dependent random variables. *Annals of Statistics*, *30*(4), 1031–1068. https://doi.org/10.1214/aos/1031689016

Belivermiş, M., Swarzenski, P. W., Oberhänsli, F., Melvin, S. D., & Metian, M. (2020). Effects of variable deoxygenation on trace element bioaccumulation and resulting metabolome profiles in the blue mussel (mytilus edulis). *Chemosphere*, *250*. https://doi.org/10.1016/j.chemosphere.2020.126314

Benassai, G., Mariani, P., Stenberg, C., & Christoffersen, M. (2014). A sustainability index of potential co-location of offshore wind farms and open water aquaculture. *Ocean and Coastal Management*, *95*, 213–218. https://doi.org/10.1016/j.ocecoaman.2014.04.007

Benjamins, S., Williamson, B., Billing, S.-L., Yuan, Z., Collu, M., Fox, C., Hobbs, L., Masden, E. A., Cottier-Cook, E. J., & Wilson, B. (2024). Potential environmental impacts of floating solar photovoltaic systems. *Renewable and Sustainable Energy Reviews*, *199*, 114463. https://doi.org/https://doi.org/10.1016/j.rser.2024.114463

Bergdahl, L., Palm, J., Eskilsson, C., & Lindahl, J. (2016). Dynamically Scaled Model Experiment of a Mooring Cable. *Journal of Marine Science and Engineering*, *4*(1), 5. https://doi.org/10.3390/jmse4010005

Bergström, P., & Lindegarth, M. (2016). Environmental influence on mussel (mytilus edulis) growth - a quantile regression approach. *Estuarine, Coastal and Shelf Science*, *171*, 123–132. https://doi.org/10.1016/j.ecss.2016.01.040

Blauw, A. N., Los, H. F. J., Bokhorst, M., & Erftemeijer, P. L. A. (2009). Gem: A generic ecological model for estuaries and coastal waters. *Hydrobiologia*, *618*, 175–198.

Booij, N., Ris, R., & Holthuijsen, L. (1999). A third-generation wave model for coastal regions, part i, model description and validation. *J. Geophys. Res.*, *104*, 7649–7656.

Brechmann, E. C., & Joe, H. (2015). Truncation of vine copulas using fit indices [High-Dimensional Dependence and Copulas]. *Journal of Multivariate Analysis*, *138*, 19–33. https://doi.org/https://doi.org/10.1016/j.jmva.2015.02.012

Brenner, M., Buchholz, C., Heemken, O., Buck, B. H., & Koehler, A. (2012). Health and growth performance of the blue mussel (mytilus edulis l.) from two hanging cultivation sites in the german bight: A nearshore-offshore comparison. *Aquaculture International*, *20*, 751–778. https://doi.org/10.1007/s10499-012-9501-0

Bru, J., Seland, T. S., Dai, J., & Jiang, Z. (2025). Life cycle cost analysis of an offshore floating photovoltaic concept in the north sea. *Renewable Energy*, *249*, 122981. https://doi.org/https://doi.org/10.1016/j.renene.2025.122981

Buck, B. H. (2002). Open ocean aquaculture und offshore windparks: Eine machbarkeitsstudie über die multifunktionale nutzung von offshore-windparks und offshore-marikultur im raum nordsee. *Reports on Polar and Marine Research*, 254.

Buck, B. H., Krause, G., Michler-Cieluch, T., Brenner, M., Buchholz, C. M., Busch, J. A., Fisch, R., Geisen, M., & Zielinski, O. (2008). Meeting the quest for spatial efficiency: Progress and prospects of extensive aquaculture within offshore wind farms. *Helgoland Marine Research*, *62*, 269–281. https://doi.org/10.1007/s10152-008-0115-x

Buck, B. H., Krause, G., & Rosenthal, H. (2004). Extensive open ocean aquaculture development within wind farms in germany: The prospect of offshore co-management and legal constraints. *Ocean and Coastal Management*, *47*, 95–122. https://doi.org/10.1016/j.ocecoaman.2004.04.002

Buck, B. H., Bjelland, H. V., Bockus, A., Chambers, M., Costa-Pierce, B. A., Dewhurst, T., Ferreira, J. G., Føre, H. M., Fredriksson, D. W., Goseberg, N., Holmyard, J., Isbert, W., Krause, G., Markus, T., Papandroulakis, N., Sclodnick, T., Silkes, B., Strand, Å., Troell, M., … Heasman, K. G. (2024). Resolving the term "offshore aquaculture" by decoupling "exposed" and "distance from the coast". *Frontiers in Aquaculture*, *3*, 1428056. https://doi.org/10.3389/faquc.2024.1428056

Buck, B. H., & Langan, R. (2017). *Aquaculture perspective of multi-use sites in the open ocean*. Springer. https://doi.org/10.1007/978-3-319-51159-7

Buck, B. H., Troell, M. F., Krause, G., Angel, D. L., Grote, B., & Chopin, T. (2018). State of the art and challenges for offshore integrated multi-trophic aquaculture (imta). *Frontiers in Marine Science*, *5*. https://doi.org/10.3389/fmars.2018.00165

Buck, B. H. (2007). Experimental trials on the feasibility of offshore seed production of the mussel mytilus edulis in the german bight: Installation, technical requirements and environmental conditions. *Helgoland Marine Research, 61*, 87–101. https://doi.org/10.1007/s10152-006-0056-1

Buck, B. H., & Buchholz, C. M. (2004). The offshore-ring: A new system design for the open ocean aquaculture of macroalgae. *Journal of Applied Phycology, 16*, 355–368. https://doi.org/10.1023/B:JAPH.0000047947.96231.ea

Buck, B. H., & Buchholz, C. M. (2005). Response of offshore cultivated laminaria saccharina to hydrodynamic forcing in the north sea. *Aquaculture, 250*, 674–691. https://doi.org/10.1016/j.aquaculture.2005.04.062

Cameron, A. C., Li, T., Trivedi, P. K., & Zimmer, D. M. (2004). Modelling the differences in counted outcomes using bivariate copula models with application to mismeasured counts*. *The Econometrics Journal, 7*(2), 566–584. https://doi.org/10.1111/j.1368-423X.2004.00144.x

Candio, P., Hill, A. J., Poupakis, S., Pulkki-Brännström, A.-M., Bojke, C., & Gomes, M. (2021). Copula models for addressing sample selection in the evaluation of public health programmes: An application to the leeds let's get active study. *Applied Health Economics and Health Policy, 19*(3), 305–312. https://doi.org/10.1007/s40258-020-00629-x

Caspers, J. J., Kindermann, P. E., Rongen, G. W., & Geerse, C. P. (2025). Storm surge hydrographs and characteristics along the dutch coast. an analysis from simulation data. *Coastal Engineering, 201*, 104776. https://doi.org/https://doi.org/10.1016/j.coastaleng.2025.104776

Chang, B., & Joe, H. (2019). Prediction based on conditional distributions of vine copulas. *Computational Statistics & Data Analysis, 139*, 45–63. https://doi.org/https://doi.org/10.1016/j.csda.2019.04.015

Chayma, S., Aymen, F., Alkuhayli, A., Ullah, R., & El-Bayeh, C. Z. (2024). A comparison between the ocean and offshore photovoltaic production system into microgrids: Benefits and limits. *Frontiers in Energy Research, Volume 12 - 2024*. https://doi.org/10.3389/fenrg.2024.1466133

Cherubini, U., Luciano, E., & Vecchiato, W. (2004). *Copula methods in finance*. John Wiley & Sons.

Chung, J., & Hulbert, G. M. (1993). A time integration algorithm for structural dynamics with improved numerical dissipation: The generalized-alpha method. *Journal of Applied Mechanics, 60*(2), 371–375. https://doi.org/10.1115/1.2900803

C.J., R., Lim, K. H., Kurnia, J. C., Roy, S., Bora, B. J., & Medhi, B. J. (2024a). Design study on the parameters influencing the performance of floating solar pv. *Renewable Energy, 223*, 120064. https://doi.org/https://doi.org/10.1016/j.renene.2024.120064

C.J., R., Lim, K. H., Kurnia, J. C., Roy, S., Bora, B. J., & Medhi, B. J. (2024b). Towards sustainable power generation: Recent advancements in floating photovoltaic technologies. *Renewable and Sustainable Energy Reviews, 194*, 114322. https://doi.org/https://doi.org/10.1016/j.rser.2024.114322

Coccoli, C., Galparsoro, I., Murillas, A., Pınarbaşı, K., & Fernandes, J. A. (2018). Conflict analysis and reallocation opportunities in the framework of marine spatial plan-

ning: A novel, spatially explicit bayesian belief network approach for artisanal fishing and aquaculture. *Marine Policy*, *94*, 119–131. https://doi.org/https://doi.org/10.1016/j.marpol.2018.04.015

Coles, S. (2001). *An introduction to statistical modeling of extreme values*. Springer-Verlag.

Czado, C. (2019). *Analyzing dependent data with vine copulas - a practical guide with r* (P. Diggle, U. Gather, & S. Zeger, Eds.; vol. 222). Springer. https://doi.org//10.1007/978-3-030-13785-4

Czado, C., & Nagler, T. (2022a). Vine copula based modeling. *Annual Review of Statistics and Its Application*, *9*(1), 453–477. https://doi.org/10.1146/annurev-statistics-040220-101153

Czado, C., & Nagler, T. (2022b). Vine copula based modeling. *Annual Review of Statistics and Its Application*, *9*(Volume 9, 2022), 453–477. https://doi.org/https://doi.org/10.1146/annurev-statistics-040220-101153

Davison, A. C., & Smith, R. L. (1990). Models for exceedances over high thresholds. *Journal of the Royal Statistical Society. Series B (Methodological)*, *52*(3), 393–442.

de Winter, R. C., Sterl, A., & Ruessink, B. G. (2013). Wind extremes in the north sea basin under climate change: An ensemble study of 12 cmip5 gcms. *Journal of Geophysical Research: Atmospheres*, *118*(4). https://doi.org/https://doi.org/10.1002/jgrd.50147

Deltares. (2023a). *D-flow flexible mesh. user manual, deltares*.

Deltares. (2023b). *D-water quality processes library description. technical reference manual, deltares*.

Desiré, P., Rodríguez-Luis, Á., & Guanche, R. (2023). Simulation of mooring lines in complex bathymetries using a finite element method. *Ocean Engineering*, *272*, 113827. https://doi.org/https://doi.org/10.1016/j.oceaneng.2023.113827

Dißmann, J., Brechmann, E., Czado, C., & Kurowicka, D. (2013). Selecting and estimating regular vine copulae and application to financial returns. *Computational Statistics and Data Analysis*, *59*(1), 52–69. https://doi.org/10.1016/j.csda.2012.08.010

Di-Tullio, G. R., Mariani, P., Benassai, G., Luccio, D. D., & Grieco, L. (2017). Sustainable use of marine resources through offshore wind and mussel farm co-location. *Ecological Modelling*, *367*, 34–41. https://doi.org/10.1016/j.ecolmodel.2017.10.012

Djalab, A., Djalab, Z., El Hammoumi, A., Marco Tina, G., Motahhir, S., & Laouid, A. A. (2024). A comprehensive review of floating photovoltaic systems: Tech advances, marine environmental influences on offshore pv systems, and economic feasibility analysis. *Solar Energy*, *277*, 112711. https://doi.org/https://doi.org/10.1016/j.solener.2024.112711

Dörenkämper, M., Wahed, A., Kumar, A., de Jong, M., Kroon, J., & Reindl, T. (2021). The cooling effect of floating pv in two different climate zones: A comparison of field test data from the netherlands and singapore [Special Issue on Floating Solar: beyond the state of the art technology]. *Solar Energy*, *219*, 15–23. https://doi.org/https://doi.org/10.1016/j.solener.2021.03.051

Dueñas-Osorio, L., & Basu, B. (2008). Unavailability of wind turbines due to wind-induced accelerations. *Engineering Structures*, *30*(4), 885–893. https://doi.org/https://doi.org/10.1016/j.engstruct.2007.05.015

European-Commission. (2010). A strategy for smart, sustainable and inclusive growth.

European-Commission. (2012). Blue growth opportunities for marine and maritime sustainable growth.

Fan, S., Ma, Z., Liu, T., Zheng, C., & Wang, H. (2025). Innovations and development trends in offshore floating photovoltaic systems: A comprehensive review. *Energy Reports*, *13*, 1950–1958. https://doi.org/https://doi.org/10.1016/j.egyr.2025.01.053

FAO. (2022). *The state of world fisheries and aquaculture 2022. towards blue transformation*. FAO. https://doi.org/10.4060/cc0461en

Fazeres-Ferradosa, T., Taveira-Pinto, F., Romão, X., Reis, M., & das Neves, L. (2019). Reliability assessment of offshore dynamic scour protections using copulas. *Wind Engineering*, *43*(5), 506–538. https://doi.org/10.1177/0309524X18807033

Fazeres-Ferradosa, T., Taveira-Pinto, F., Vanem, E., Reis, M. T., & das Neves, L. (2018). Asymmetric copula–based distribution models for met-ocean data in offshore wind engineering applications. *Wind Engineering*, *42*(4), 304–334. https://doi.org/10.1177/0309524X18777323

Feng, D., Meng, A., Wang, P., Yao, Y., & Gui, F. (2021). Effect of design configuration on structural response of longline aquaculture in waves. *Applied Ocean Research*, *107*, 102489. https://doi.org/https://doi.org/10.1016/j.apor.2020.102489

Fernand, F., Israel, A., Skjermo, J., Wichard, T., Timmermans, K. R., & Golberg, A. (2017). Offshore macroalgae biomass for bioenergy production: Environmental aspects, technological achievements and challenges. *Renewable and Sustainable Energy Reviews*, *75*, 35–45. https://doi.org/https://doi.org/10.1016/j.rser.2016.10.046

Filgueira, R., Fernández-Reiriz, M., & Labarta, U. (2009). Clearance rate of the mussel mytilus galloprovincialis. i. response to extreme chlorophyll ranges. *Ciencias Marinas*, *35*, 405–417.

Fisher, R., & Tippett, L. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proceedings of the Cambridge Philosophical Society*, *24*, 180–190. https://doi.org/10.1017/S0305004100015681

Fly, E. K., & Hilbish, T. J. (2013). Physiological energetics and biogeographic range limits of three congeneric mussel species. *Oecologia*, *172*, 35–46. https://doi.org/10.1007/s00442-012-2486-6

Fries, T., & Schöllhammer, D. (2020). A unified finite strain theory for membranes and ropes. *Computer Methods in Applied Mechanics and Engineering*, *365*, 113031. https://doi.org/10.1016/j.cma.2020.113031

Gagnon, M., & Bergeron, P. (2017). Observations of the loading and motion of a submerged mussel longline at an open ocean site. *Aquacultural Engineering*, *78*, 114–129. https://doi.org/10.1016/j.aquaeng.2017.05.004

Geisler, R., Schulz, C., Michl, S., & Strothotte, E. (2018). *Feasbility study: Offshore-aquakultur am standort der forschungsplattform fino3*.

Genest, C., & Favre, A.-C. (2007). Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of Hydrologic Engineering*, 347–368. https://doi.org/10.1061/ASCE1084-0699200712:4347

Genest, C., Gendron, M., & and, M. B.-B. (2009a). The advent of copulas in finance. *The European Journal of Finance*, *15*(7-8), 609–618. https://doi.org/10.1080/13518470802604457

Genest, C., & Nešlehová, J. (2007). A primer on copulas for count data. *ASTIN Bulletin*, *37*, 475–515. https://doi.org/10.2143/ast.37.2.2024077

Genest, C., Rémillard, B., & Beaudoin, D. (2009b). Goodness-of-fit tests for copulas: A review and a power study. *Insurance: Mathematics and Economics*, *44*, 199–213. https://doi.org/10.1016/j.insmatheco.2007.10.005

Gimpel, A., Stelzenmüller, V., Grote, B., Buck, B. H., Floeter, J., Núñez-Riboni, I., Pogoda, B., & Temming, A. (2015). A gis modelling framework to evaluate marine spatial planning scenarios: Co-location of offshore wind farms and aquaculture in the german eez. *Marine Policy*, *55*, 102–115. https://doi.org/10.1016/j.marpol.2015.01.012

Governo, A., Henriques, J., & Gato, L. (2023). Modelling mooring line snap loads using a high-order finite-volume approach. *Ocean Engineering*, *275*, 113803. https://doi.org/https://doi.org/10.1016/j.oceaneng.2023.113803

Guo, S., Li, Y., Li, M., Chen, W., & Fu, Y. (2017). *Dynamic response of floating wind turbine under consideration of dynamic behavior of catenary mooring-lines* (Vol. 10). https://doi.org/10.1115/OMAE2017-61689

Haas, J., Khalighi, J., de la Fuente, A., Gerbersdorf, S., Nowak, W., & Chen, P.-J. (2020). Floating photovoltaic plants: Ecological impacts versus hydropower operation flexibility. *Energy Conversion and Management*, *206*, 112414. https://doi.org/https://doi.org/10.1016/j.enconman.2019.112414

Hall, M., & Goupee, A. (2015). Validation of a lumped-mass mooring line model with DeepCwind semisubmersible model test data. *Ocean Engineering*, *104*, 590–603. https://doi.org/10.1016/j.oceaneng.2015.05.035

Hanea, A., Morales Napoles, O., & Ababei, D. (2015). Non-parametric bayesian networks: Improving theory and reviewing applications. *Reliability Engineering & System Safety*, *144*, 265–284. https://doi.org/https://doi.org/10.1016/j.ress.2015.07.027

Hanea, A., Kurowicka, D., & Cooke, R. (2006). Hybrid method for quantifying and analysing bayesian belief nets. *Quality and Reliability Engineering International*, *22*, 709–729. https://doi.org/10.1002/qre.808

Hermawan, Y. A., & Furukawa, Y. (2020). Coupled three-dimensional dynamics model of multi-component mooring line for motion analysis of floating offshore structure. *Ocean Engineering*, *200*, 106928. https://doi.org/https://doi.org/10.1016/j.oceaneng.2020.106928

Hoff, P. D. (2007). Extending the rank likelihood for semiparametric copula estimation. *The Annals of Applied Statistics*, *1*(1), 265–283. https://doi.org/10.1214/07-AOAS107

Hooper, T., Armstrong, A., & Vlaswinkel, B. (2021). Environmental impacts and benefits of marine floating solar [Special Issue on Floating Solar: beyond the state of the art technology]. *Solar Energy*, *219*, 11–14. https://doi.org/https://doi.org/10.1016/j.solener.2020.10.010

Hopkins, D. D., Goldburg, R. J., & Martson, A. (1996). An environmental critique of government regulations and policies for open ocean aquaculture. *Ocean and Coastal Law Journal*, *2*.

Hsu, W., Thiagarajan, K. P., & Manuel, L. (2017). Extreme mooring tensions due to snap loads on a floating offshore wind turbine system. *Marine Structures*, *55*, 182–199. https://doi.org/https://doi.org/10.1016/j.marstruc.2017.05.005

Hsu, W.-Y., Chuang, T.-C., Yang, R.-Y., Hsu, W.-T., & Thiagarajan, K. P. (2019). An experimental study of mooring line damping and snap load in shallow water. *Journal of Offshore Mechanics and Arctic Engineering*, *141*(5), 051603. https://doi.org/10.1115/1.4042535

Ikhennicheu, M., Danglade, B., Pascal, R., Arramounet, V., Trébaol, Q., & Gorintin, F. (2021). Analytical method for loads determination on floating solar farms in three typical environments [Special Issue on Floating Solar: beyond the state of the art technology]. *Solar Energy*, *219*, 34–41. https://doi.org/https://doi.org/10.1016/j.solener.2020.11.078

Jaeger, W. S., & Morales-Nápoles, O. (2017). A vine-copula model for time series of significant wave heights and mean zero-crossing periods in the north sea. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, *3*, 04017014.

Jäger, W., Nagler, T., Czado, C., & McCall, R. (2019). A statistical simulation method for joint time series of non-stationary hourly wave parameters. *Coastal Engineering*, *146*, 14–31. https://doi.org/https://doi.org/10.1016/j.coastaleng.2018.11.003

Jevne, L. S., Forbord, S., & Olsen, Y. (2020). The effect of nutrient availability and light conditions on the growth and intracellular nitrogen components of land-based cultivated saccharina latissima (phaeophyta). *Frontiers in Marine Science*, *7*. https://doi.org/10.3389/fmars.2020.557460

Ji, C., Hao, Y., & Xu, S. (2025). Experimental and numerical study on the hydrodynamic responses of a novel offshore floating photovoltaic system. *Ocean Engineering*, *315*, 119797. https://doi.org/https://doi.org/10.1016/j.oceaneng.2024.119797

Jifaturrohman, M. I., Utama, I. K. A. P., Putranto, T., Setyawan, D., & Huang, L. (2024). A study into the correlation between single array-hull configurations and wave spectrum for floating solar photovoltaic systems. *Ocean Engineering*, *312*, 119312. https://doi.org/https://doi.org/10.1016/j.oceaneng.2024.119312

Joe, H. (1996). Families of m-variate distributions with given margins and m(m-1)/2 bivariate dependence parameters. In L. Rüschendorf, B. Schweizer, & M. Taylor (Eds.), *Distributions with fixed marginals and related topics* (pp. 20–141). Institute of Mathematical Statistics. https://doi.org/10.1214/lnms/1215452614

Joe, H. (2015). *Dependence modeling with copulas*. Monographs on Statistics; Applied Probability, 134. Boca Raton: CRC Press. https://doi.org/10.1201/b17116

Jossart, J., Theuerkauf, S. J., Wickliffe, L. C., & Morris Jr., J. A. (2020). Applications of spatial autocorrelation analyses for marine aquaculture siting. *Frontiers in Marine Science*, *6*. https://doi.org/10.3389/fmars.2019.00806

Ju, Y. R., Chen, W. Y., & Liao, C. M. (2014). Model-based risk assessment for milkfish and tilapia exposed to arsenic in a traditional polyculture system with seasonal variations. *Aquacultural Engineering*, *62*, 1–8. https://doi.org/10.1016/j.aquaeng.2014.07.001

Kamermans, P., Saurel, C., Boudry, P., & Kamermans, P. (2022). Interacting climate change effects on mussels (mytilus edulis and m. galloprovincialis) and oysters (cras-

sostrea gigas and ostrea edulis): Experiments for bivalve individual growth models. *Aquatic Living Resources*, *35*. https://doi.org/10.1051/alr/2022001

Kammler, S., Romero, A. M., Burkhardt, C., Baruth, L., Antranikian, G., Liese, A., & Kaltschmitt, M. (2024). Macroalgae valorization for the production of polymers, chemicals, and energy. *Biomass and Bioenergy*, *183*, 107105. https://doi.org/https://doi.org/10.1016/j.biombioe.2024.107105

Karatas, Y., & Yilmaz, D. (2021). Experimental investigation of the microclimate effects on floating solar power plant energy efficiency. *Clean Technologies and Environmental Policy*, *23*, 2157–2170. https://doi.org/10.1007/s10098-021-02122-y

Karpouzoglou, T., Vlaswinkel, B., & van der Molen, J. (2020). Effects of large-scale floating (solar photovoltaic) platforms on hydrodynamics and primary production in a coastal sea from a water column model. *Ocean Science*, *16*(1), 195–208. https://doi.org/10.5194/os-16-195-2020

Kendall, M. G. (1945). The treatment of ties in ranking problems. *Biometrika*, *33*(3), 239–251. https://doi.org/https://doi.org/10.2307/2332303

Kerrison, P. D., Stanley, M. S., Edwards, M. D., Black, K. D., & Hughes, A. D. (2015). The cultivation of european kelp for bioenergy: Site and species selection. *Biomass and Bioenergy*, *80*, 229–242. https://doi.org/10.1016/j.biombioe.2015.04.035

Kjeldstad, T., Nysted, V. S., Kumar, M., Oliveira-Pinto, S., Otnes, G., Lindholm, D., & Selj, J. (2022). The performance and amphibious operation potential of a new floating photovoltaic technology. *Solar Energy*, *239*, 242–251. https://doi.org/https://doi.org/10.1016/j.solener.2022.04.065

Klein, N., Kneib, T., Marra, G., Radice, R., Rokicki, S., & McGovern, M. E. (2019). Mixed binary-continuous copula regression models with application to adverse birth outcomes. *Statistics in Medicine*, *38*(3), 413–436. https://doi.org/https://doi.org/10.1002/sim.7985

Knysh, A., Tsukrov, I., Chambers, M., Swift, M. R., Sullivan, C., & Drach, A. (2020). Numerical modeling of submerged mussel longlines with protective sleeves. *Aquacultural Engineering*, *88*, 102027. https://doi.org/https://doi.org/10.1016/j.aquaeng.2019.102027

Kole, E., Koedijk, K., & Verbeek, M. (2007). Selecting copulas for risk management. *Journal of Banking & Finance*, *31*(8), 2405–2423. https://doi.org/https://doi.org/10.1016/j.jbankfin.2006.09.010

Koot, P., Mendoza-Lugo, M. A., Paprotny, D., Morales-Nápoles, O., Ragno, E., & Worm, D. T. (2023). Pybanshee version (1.0): A python implementation of the matlab toolbox banshee for non-parametric bayesian networks with updated features. *SoftwareX*, *21*. https://doi.org/https://doi.org/10.1016/j.softx.2022.101279

Landmann, J., Santjer, R., Gieschen, R., Goseberg, N., & Hildebrandt, A. (2019). Laboratory tests on wave forces and accelerations of a three-bay long-line aquaculture system in offshore conditions. *Proceedings of the 29th International Ocean and Polar Engineering Conference, ISOPE*, 2726–2732.

Leontaris, G., Morales-Nápoles, O., & Wolfert, A. R. (2016). Probabilistic scheduling of offshore operations using copula based environmental time series – an application for cable installation management for offshore wind farms. *Ocean Engineering*, *125*, 328–341. https://doi.org/10.1016/j.oceaneng.2016.08.029

Li, H., Du, J., Wang, S., Sun, M., & Chang, A. (2016). Investigation on the probabilistic distribution of mooring line tension for fatigue damage assessment. *Ocean Engineering*, *124*, 204–214. https://doi.org/https://doi.org/10.1016/j.oceaneng.2016.07.024

Li, Q., Zhang, F., & Sun, S. (2022). The survival and responses of blue mussel mytilus edulis to 16-day sustained hypoxia stress. *Marine Environmental Research*, *176*. https://doi.org/10.1016/j.marenvres.2022.105601

Li, X., & Zhang, W. (2020). Long-term assessment of a floating offshore wind turbine under environmental conditions with multivariate dependence structures. *Renewable Energy*, *147*, 764–775. https://doi.org/https://doi.org/10.1016/j.renene.2019.09.076

Loaiza-Maya, R., & Smith, M. S. (2019). Variational bayes estimation of discrete-margined copula models with application to time series. *Journal of Computational and Graphical Statistics*, *28*(3), 523–539. https://doi.org/10.1080/10618600.2018.1562936

López-Rebollar, B. M., Arévalo-Mejía, R., Díaz-Delgado, C., Latif, S., & Ouarda, T. B. (2024). Settling velocity and effective density analysis for aquaculture floc particles: An approach through bivariate parametric copula. *Aquacultural Engineering*, *107*, 102459. https://doi.org/https://doi.org/10.1016/j.aquaeng.2024.102459

Lubsch, A., & Timmermans, K. R. (2019). Uptake kinetics and storage capacity of dissolved inorganic phosphorus and corresponding dissolved inorganic nitrate uptake in saccharina latissima and laminaria digitata (phaeophyceae). *Journal of Phycology*, *55*, 637–650. https://doi.org/10.1111/jpy.12844

Maar, M., Holbach, A., Boderskov, T., Thomsen, M., Buck, B. H., Kotta, J., & Bruhn, A. (2023). Multi-use of offshore wind farms with low-trophic aquaculture can help achieve global sustainability goals. *Communications Earth & Environment*, *4*, 447. https://doi.org/10.1038/s43247-023-01116-6

Manoj Kumar, N., Chakraborty, S., Kumar Yadav, S., Singh, J., & Chopra, S. S. (2022). Advancing simulation tools specific to floating solar photovoltaic systems – comparative analysis of field-measured and simulated energy performance. *Sustainable Energy Technologies and Assessments*, *52*, 102168. https://doi.org/https://doi.org/10.1016/j.seta.2022.102168

Mares-Nasarre, P., van Gent, M. R., & Morales-Nápoles, O. (2024). A copula-based model to describe the uncertainty of overtopping variables on mound breakwaters. *Coastal Engineering*, *189*, 104483. https://doi.org/https://doi.org/10.1016/j.coastaleng.2024.104483

Medina Madariaga, G. P., Falcou-Prefol, M., Oiry, S., Zoffoli, M. L., Fleurence, J., Gernez, P., & Barillé, L. (2025). Site selection for off-shore macroalgae aquaculture on the french atlantic coast. *Aquacultural Engineering*, *111*, 102581. https://doi.org/https://doi.org/10.1016/j.aquaeng.2025.102581

Mendoza-Lugo, M. A., & Morales-Nápoles, O. (2023). Version 1.3-banshee—a matlab toolbox for non-parametric bayesian networks. *SoftwareX*, *23*. https://doi.org/https://doi.org/10.1016/j.softx.2023.101479

Michler-Cieluch, T., Krause, G., & Buck, B. H. (2009). Reflections on integrating operation and maintenance activities of offshore wind farms and mariculture. *Ocean and*

Coastal Management, *52*, 57–68. https://doi.org/10.1016/j.ocecoaman.2008.09.008

Miškov, V., Dragić, M., Tomin, V., & Hofman, M. (2023). Sea trials of sigma wave energy converter – slacking and snapping of tlp mooring lines. *Marine Structures*, *89*, 103378. https://doi.org/https://doi.org/10.1016/j.marstruc.2023.103378

Montes-Iturrizaga, R., & Heredia-Zavoni, E. (2016). Reliability analysis of mooring lines using copulas to model statistical dependence of environmental variables. *Applied Ocean Research*, *59*, 564–576. https://doi.org/https://doi.org/10.1016/j.apor.2016.07.008

Morales-Nápoles, O., Rajabi-Bahaabadi, M., Torres-Alves, G. A., & Hart, C. M. P. '. (2023). Chimera: An atlas of regular vines on up to 8 nodes. *Scientific Data*, *10*, 337. https://doi.org/10.1038/s41597-023-02252-6

Morales-Nápoles, O. (2010a). Bayesian belief nets and vines in aviation safety and other applications.

Morales-Nápoles, O. (2010b). Counting vines. *Dependence Modeling: Vine Copula Handbook*, 189–218. https://doi.org/10.1142/9789814299886_0009

Morales-Nápoles, O., Delgado-Hernández, D. J., De-León-Escobedo, D., & Arteaga-Arcos, J. C. (2014). A continuous bayesian network for earth dams' risk assessment: Methodology and quantification. *Structure and Infrastructure Engineering*, *10*(5), 589–603. https://doi.org/10.1080/15732479.2012.757789

Nagler, T., & Vatter, T. (2025). *Pyvinecopulib* (Version v0.7.1). https://doi.org/10.5281/zenodo.14841456

Nalmpanti, M., Dinandra, T., Maar, M., Larsen, J., & Vlaswinkel, B. (2025). Offshore solar farms as habitats for mytilus edulis: A preliminary modelling study on mussel growth, distribution, chlorophyll-a uptake and bio-deposition in the north sea. *Marine Environmental Research*, *211*, 107372. https://doi.org/https://doi.org/10.1016/j.marenvres.2025.107372

Nasyrlayev, N., Dyson, A. P., Kefayati, G., & Tolooiyan, A. (2023). Modelling the response of offshore aquaculture fish pens to environmental loads in high-energy regions. *Applied Ocean Research*, *135*, 103541. https://doi.org/https://doi.org/10.1016/j.apor.2023.103541

Nelsen, R. (2006). *An introduction to copulas* (Second Edition). Springer Series in Statistics.

Neyman, J., & Pearson, E. S. (1928). On the use and interpretation of certain test criteria for purposes of statistical inference: Part i. *Biometrika*, *20A*(1/2), 175–240.

Nguyen, N. Q., Thiagarajan, K., & Auger, J. (2019). Integrity assessment of an oyster farm mooring system through in-situ measurements and extreme environment modeling. *Ocean Engineering, 172*, 641–659. https://doi.org/https://doi.org/10.1016/j.oceaneng.2018.11.023

Nikoloulopoulos, A. K., & Karlis, D. (2008). Multivariate logit copula model with an application to dental data. *Statistics in Medicine*, *27*(30), 6393–6406. https://doi.org/https://doi.org/10.1002/sim.3449

Osman, A. I., Chen, L., Yang, M., Msigwa, G., Farghali, M., Fawzy, S., Rooney, D. W., & Yap, P.-S. (2023). Cost, environmental impact, and resilience of renewable energy under a changing climate: A review. *Environmental Chemistry Letters*, *21*(2), 741–764. https://doi.org/10.1007/s10311-022-01532-8

Pan, S., & Joe, H. (2022). Predicting times to event based on vine copula models. *Computational Statistics & Data Analysis*, *175*, 107546. https://doi.org/https://doi.org/10.1016/j.csda.2022.107546

Panagiotelis, A., Czado, C., & and, H. J. (2012). Pair copula constructions for multivariate discrete data. *Journal of the American Statistical Association*, *107*(499), 1063–1072. https://doi.org/10.1080/01621459.2012.682850

Panagiotelis, A., Czado, C., Joe, H., & Stöber, J. (2017). Model selection for discrete regular vine copulas. *Computational Statistics & Data Analysis*, *106*, 138–152. https://doi.org/https://doi.org/10.1016/j.csda.2016.09.007

Paprotny, D., Morales-Nápoles, O., Worm, D., & Ragno, E. (2020). Banshee–a matlab toolbox for non-parametric bayesian networks. *SoftwareX*, *12*. https://doi.org/10.1016/j.softx.2020.100588

Pascoe, P. L., Parry, H. E., & Hawkins, A. J. (2009). Observations on the measurement and interpretation of clearance rate variations in suspension-feeding bivalve shellfish. *Aquatic Biology*, *6*, 181–190. https://doi.org/10.3354/ab00123

Pearson, K., & Galton, F. (1895). Vii. note on regression and inheritance in the case of two parents. *Proceedings of the Royal Society of London*, *58*(347-352), 240–242. https://doi.org/https://doi.org/10.1098/rspl.1895.0041

Peteiro, C., & Freire, O. (2011). Offshore cultivation methods affects blade features of the edible seaweed saccharina latissima in a bay of galicia, northwest spain. *Russian Journal of Marine Biology*, *37*, 319–323. https://doi.org/10.1134/S1063074011040110

Phillips, R. C., Samadi, S., Hitchcock, D. B., Meadows, M. E., & Wilson, C. A. M. E. (2022). The devil is in the tail dependence: An assessment of multivariate copula-based frameworks and dependence concepts for coastal compound flood dynamics. *Earth's Future*, *10*(10). https://doi.org/10.1029/2022EF002705

Pickands, J. (1975). Statistical inference using extreme order statistics. *The Annals of Statistics*, *3*, 119–131. https://doi.org/10.1214/aos/1176343003

Pınarbaşı, K., Galparsoro, I., Borja, Á., Stelzenmüller, V., Ehler, C. N., & Gimpel, A. (2017). Decision support tools in marine spatial planning: Present applications, gaps and future perspectives. *Marine Policy*, *83*, 83–91. https://doi.org/https://doi.org/10.1016/j.marpol.2017.05.031

Ragno, E., AghaKouchak, A., Cheng, L., & Sadegh, M. (2019). A generalized framework for process-informed nonstationary extreme value analysis. *Advances in Water Resources*, *130*, 270–282. https://doi.org/https://doi.org/10.1016/j.advwatres.2019.06.007

Ragno, E., Antonini, A., & Pasquali, D. (2023). Investigating extreme sea level components and their interactions in the adriatic and tyrrhenian seas. *Weather and Climate Extremes*, *41*, 100590. https://doi.org/https://doi.org/10.1016/j.wace.2023.100590

Redfield, A. (1934). On the proportions of organic derivatives in sea water and their relation to the composition of plankton. *James Johnstone Memorial*, *176*, 177–192.

Riisgård, H. U., Egede, P. P., & Saavedra, I. B. (2011). Feeding behaviour of the mussel, mytilus edulis : New observations, with a minireview of current knowledge. *Journal of Marine Biology*, *2011*, 1–13. https://doi.org/10.1155/2011/312459

Ruiz-Velazco, J. M., Estrada-Pérez, M., Hernández-Llamas, A., Nieto-Navarro, J. T., & Peña-Messina, E. (2013). Stock model and multivariate analysis for prediction of semi-intensive production of shrimp litopenaeus vannamei as a function of water quality and management variables: A stochastic approach. *Aquacultural Engineering*, *56*, 34–41. https://doi.org/10.1016/j.aquaeng.2013.04.003

Şahin, Ö., & Joe, H. (2024). Vine copula-based classifiers with applications. *Journal of Classification*. https://doi.org/10.1007/s00357-024-09494-y

Sahu, A., Yadav, N., & Sudhakar, K. (2016). Floating photovoltaic power plant: A review. *Renewable and Sustainable Energy Reviews*, *66*, 815–824. https://doi.org/https://doi.org/10.1016/j.rser.2016.08.051

Santjer, R., Agarwal, S., Colomés, O., & Morales-Nápoles, O. (2025). Gaussian copula-based bayesian networks for dynamic loads in mooring systems. *Applied Ocean Research*, *165*, 104809. https://doi.org/https://doi.org/10.1016/j.apor.2025.104809

Santjer, R., Mares-Nasarre, P., Vilmin, L., El Serafy, G., & Morales-Nápoles, O. (2024a). A probabilistic framework for offshore aquaculture suitability assessment using bivariate copulas. *Aquacultural Engineering*, *107*, 102479. https://doi.org/https://doi.org/10.1016/j.aquaeng.2024.102479

Santjer, R., Mares-Nasarre, P., El-Serafy, G., & Morales-Nápoles, O. (2023). A case study of ecological suitability of mussel and seaweed cultivation using bivariate copula functions. *Proceeding of the 33rd European Safety and Reliability Conference*, 1877–1884. https://doi.org/10.3850/978-981-18-8071-1_P172-cd

Santjer, R., Mares-Nasarre, P., El-Serafy, G., & Morales-Nápoles, O. (2024b). Copula based bayesian networks for hydrodynamic load estimation on offshore aquaculture systems. *Proceeding of the 34rd European Safety and Reliability Conference, Advances in Reliability, Safety and Security, Part 3*, 163–172.

Schiener, P. (2014). *Bioethanol production from macroalgae* [Doctoral Thesis]. University of Edinburgh.

Schmidt, W., Raymond, D., Parish, D., Ashton, I. G., Miller, P. I., Campos, C. J., & Shutler, J. D. (2018). Design and operation of a low-cost and compact autonomous buoy system for use in coastal aquaculture and water quality monitoring. *Aquacultural Engineering*, *80*, 28–36. https://doi.org/https://doi.org/10.1016/j.aquaeng.2017.12.002

Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de l'Université de Paris*, *8*.

Soares-Ramos, E. P., de Oliveira-Assis, L., Sarrias-Mena, R., & Fernández-Ramírez, L. M. (2020). Current status and future trends of offshore wind power in europe. *Energy*, *202*. https://doi.org/10.1016/j.energy.2020.117787

Somoano, M., Rodríguez-Luis, A., Blanco, D., & Guanche, R. (2024). Effect of bathymetry irregularities on the energy dissipation of a mooring line. *Ocean Engineering*, *296*, 116932. https://doi.org/https://doi.org/10.1016/j.oceaneng.2024.116932

Spearman, C. (1987). The proof and measurement of association between two things. *The American Journal of Psychology*, *100*, 441–71. https://doi.org/10.2307/1422689

Sree, D. K., Law, A. W.-K., Pang, D. S. C., Tan, S. T., Wang, C. L., Kew, J. H., Seow, W. K., & Lim, V. H. (2022). Fluid-structural analysis of modular floating solar farms under wave motion. *Solar Energy*, *233*, 161–181. https://doi.org/https://doi.org/10.1016/j.solener.2022.01.017

Stechele, B., van der Zande, D., Alvera-Azcárate, A., Delbare, D., Lacroix, G., & Nevejan, N. (2022). Biological site suitability for exposed self-regulating cultivation of blue mussel (mytilus edulis): A belgian case study. *Aquacultural Engineering*, *98*. https://doi.org/10.1016/j.aquaeng.2022.102264

Stöber, J., Hong, H. G., Czado, C., & Ghosh, P. (2015). Comorbidity of chronic diseases in the elderly: Patterns identified by a copula design for mixed responses. *Computational Statistics & Data Analysis*, *88*, 28–39. https://doi.org/https://doi.org/10.1016/j.csda.2015.02.001

Stockbridge, J., Brown, C. J., & Kuempel, C. D. (2025). Mapping the global co-location potential of offshore wind energy and aquaculture production. *Ocean & Coastal Management*, *263*, 107605. https://doi.org/https://doi.org/10.1016/j.ocecoaman.2025.107605

Strothotte, A., Jaeger, M., Pforth, J., Declercq, A., Stechele, B., Nevejan, N., Knoop, J., Kerkhove, T., Petit, S., Pribadi, A., Fernandez, G. V., Lataire, E., Groenendaal, B., Drigkopoulou, I., Brouwers, E., Sørensen, H., & Triest, J. (2021). *Blueprint for the offshore site operation* (Deliverable 7.2). UNITED project: multi-Use platforms, co-locatioN Pilots boostIng cost-effecTive, Eco-friendly, and sustainable proDuction in marine environments.

Sušelj, K., Sood, A., & Heinemann, D. (2010). North sea near-surface wind climate and its relation to the large-scale circulation patterns. *Theoretical and Applied Climatology*, *99*. https://doi.org/10.1007/s00704-009-0149-2

Tang, B., & Riisgård, H. U. (2018). Relationship between oxygen concentration, respiration and filtration rate in blue mussel mytilus edulis. *Journal of Oceanology and Limnology*, *36*, 395–404. https://doi.org/10.1007/s00343-018-6244-4

Tawn, J. A. (1988). Bivariate extreme value theory: Models and estimation. *Biometrika*, *75*(3), 397–415. https://doi.org/10.1093/biomet/75.3.397

Temiz, M., & Javani, N. (2020). Design and analysis of a combined floating photovoltaic system for electricity and hydrogen production [Hydrogen Energy Technologies for Mitigating Global Warming]. *International Journal of Hydrogen Energy*, *45*(5), 3457–3469. https://doi.org/https://doi.org/10.1016/j.ijhydene.2018.12.226

Torres-Alves, G., 't Hart, C., Morales-Nápoles, O., & Jonkman, S. (2022). Structural reliability analysis of a submerged floating tunnel under copula-based traffic load simulations. *Engineering Structures*, *269*, 114752. https://doi.org/https://doi.org/10.1016/j.engstruct.2022.114752

Tripathi, A. D., Mishra, R., Maurya, K. K., Singh, R. B., & Wilson, D. W. (2019). Estimates for world population and global food availability for global health (R. B. Singh, R. Watson, & T. Takahashi, Eds.). *The Role of Functional Food Security in Global Health*, 3–24. https://doi.org/10.1016/B978-0-12-813148-0.00001-3

Tyler-Walters, H., & Hiscock, K. (2008). *Mytilus edulis - common mussel*. Marine Biological Association of the UNITED Kingdom. https://doi.org/10.17031/marlinsp.1421.1

United Nations. (2023). Bbnj agreement: Agreement on marine biological diversity of areas beyond national jurisdiction [Accessed: 2025-08-13].

UNITED Project. (2022). Multi-use offshore platforms demonstrators for boosting cost-effective and eco-friendly production in sustainable marine activities [Accessed: 2024-09-23].

van Leeuwen, S. M., Lenhart, H.-J., Prins, T. C., Blauw, A., Desmit, X., Fernand, L., Friedland, R., Kerimoglu, O., Lacroix, G., van der Linden, A., Lefebvre, A., van der Molen, J., Plus, M., Baroni, I. R., Silva, T., Stegert, C., Troost, T. A., & Vilmin, L. M. (2023). Deriving pre-eutrophic conditions from an ensemble model approach for the north-west european seas. *Frontiers in Marine Science*, *10*.

Varin, C., & Czado, C. (2009). A mixed autoregressive probit model for ordinal longitudinal data. *Biostatistics*, *11*(1), 127–138. https://doi.org/10.1093/biostatistics/kxp042

Weiss, C. V., Ondiviela, B., Guinda, X., del Jesus, F., González, J., Guanche, R., & Juanes, J. A. (2018). Co-location opportunities for renewable energies and aquaculture facilities in the canary archipelago [Maritime Spatial Planning, Ecosystem Approach and Supporting Information Systems (MapSIS 2017)]. *Ocean & Coastal Management*, *166*, 62–71. https://doi.org/https://doi.org/10.1016/j.ocecoaman.2018.05.006

Wever, L., Krause, G., & Buck, B. H. (2015). Lessons from stakeholder dialogues on marine aquaculture in offshore wind farms: Perceived potentials, constraints and research gaps. *Marine Policy*, *51*, 251–259. https://doi.org/10.1016/j.marpol.2014.08.015

Wu, J., Yu, Y., Cheng, S., Li, Z., & Yu, J. (2022). Probabilistic multilevel robustness assessment framework for a tlp under mooring failure considering uncertainties. *Reliability Engineering & System Safety*, *223*, 108458. https://doi.org/https://doi.org/10.1016/j.ress.2022.108458

Xu, S., & Guedes Soares, C. (2021). Mixture distribution model for extreme mooring tension and mooring fatigue analysis due to snap loads. *Ocean Engineering*, *234*, 109245. https://doi.org/https://doi.org/10.1016/j.oceaneng.2021.109245

Xu, S., Ji, C.-y., & Guedes Soares, C. (2022). A semiparametric bayesian method with birth-death markov chain monte carlo algorithm for extreme mooring tension analysis. *Ocean Engineering*, *260*, 111765. https://doi.org/https://doi.org/10.1016/j.oceaneng.2022.111765

Yang, L., Frees, E. W., & and, Z. Z. (2020). Nonparametric estimation of copula regression models with discrete outcomes. *Journal of the American Statistical Association*, *115*(530), 707–720. https://doi.org/10.1080/01621459.2018.1546586

Yong, W. T. L., Thien, V. Y., Rupert, R., & Rodrigues, K. F. (2022). Seaweed: A potential climate change solution. *Renewable and Sustainable Energy Reviews*, *159*. https://doi.org/10.1016/j.rser.2022.112222

Zar, J. H. (1972). Significance testing of the spearman rank correlation coefficient. *Journal of the American Statistical Association*, *67*(339), 578–580. https://doi.org/10.1080/01621459.1972.10481251

Zhang, M., & Schreier, S. (2022). Review of wave interaction with continuous flexible floating structures. *Ocean Engineering*, *264*, 112404. https://doi.org/https://doi.org/10.1016/j.oceaneng.2022.112404

Zhang, Y., Beer, M., & Quek, S. T. (2015). Long-term performance assessment and design of offshore structures. *Computers & Structures*, *154*, 101–115. https://doi.org/https://doi.org/10.1016/j.compstruc.2015.02.029

Zhao, Y., & Dong, S. (2023). Uncertainty analysis of extreme mooring loads associated with environmental contours and peak tension distributions. *Marine Structures*, *89*, 103369. https://doi.org/https://doi.org/10.1016/j.marstruc.2023.103369

Zhou, N., Xu, X., Ya, Z., & Shahidehpour, M. (2022). Spatio-temporal probabilistic forecasting of photovoltaic power based on monotone broad learning system and copula theory. *IEEE Transactions on Sustainable Energy*, *13*(4), 1874–1885. https://doi.org/10.1109/TSTE.2022.3174012

Zijl, F., Zijlker, T., Laan, S., & Groenenboom, J. (2023). *3d dcsm fm: A sixth-generation model for the nw european shelf. technical report, deltares*.

Zilko, A. A., & Kurowicka, D. (2016). Copula in a multivariate mixed discrete–continuous model. *Computational Statistics & Data Analysis*, *103*, 28–55. https://doi.org/https://doi.org/10.1016/j.csda.2016.02.017

SAMENVATTING

Wereldwijd neemt de bevolking voortdurend toe, evenals de vraag naar voedsel en energie. Deze groeiende behoefte en de beperkte mogelijkheden om de grondstoffen op het land verder te benutten, hebben geleid tot een toenemende belangstelling voor de exploitatie van het mariene milieu om het tekort aan hulpbronnen op te vangen. Met name offshoregebieden worden steeds vaker verkend, ondanks uitdagingen zoals barre omgevingsomstandigheden die worden gekenmerkt door sterke variabiliteit en complexe interacties tussen fysische processen.

Terwijl offshore-sectoren zoals windenergie al commercieel zijn ingeburgerd, bevinden andere sectoren, zoals offshore drijvende fotovoltaïsche (FPV) systemen en aquacultuur, zich nog in een vroeg stadium van ontwikkeling. Voor beide sectoren is onderzoek gedaan door middel van kleinschalige veldobservaties, laboratoriumexperimenten of numerieke modellering. Deze benaderingen vereenvoudigen of negeren echter vaak de statistische afhankelijkheid tussen de milieu- en structurele variabelen, waardoor een optimaal ontwerp van offshoreconstructies wordt belemmerd.

Dit proefschrift onderzoekt de toepasbaarheid van probabilistische modellen om dergelijke afhankelijkheden expliciet vast te leggen en laat het zien hoe dergelijke modellen de besluitvorming met betrekking tot verschillende offshore-technologieën kunnen verbeteren, met bijzondere aandacht voor aquacultuur en drijvende fotovoltaïsche systemen. Er worden twee op copula's gebaseerde multivariate modelleringsbenaderingen met verschillende mate van complexiteit toegepast. Enerzijds biedt het op Gaussische copula's gebaseerde Bayesiaanse netwerk (GCBN) een relatief toegankelijk en interpreteerbaar kader, aangezien de grafische structuur ervan kan worden afgeleid uit de onderliggende fysische processen. Anderzijds worden vine-copulamodellen gebruikt om complexe afhankelijkheids patronen flexibeler weer te geven. In tegenstelling tot GCBN's zijn ze niet beperkt tot één copula-familie, maar kan elk paar variabelen worden gemodelleerd met behulp van de copula die hun afhankelijkheid het beste weergeeft. Deze flexibiliteit gaat echter ten koste van een grotere complexiteit van het model en grotere rekenkundige en praktische uitdagingen.

Vervolgens is dit proefschrift gedreven door de algemene onderzoeksvraag hoe multivariate copula-gebaseerde probabilistische modellen ruimtelijke planning en structurele veiligheidsbeoordelingen in de offshore-techniek kunnen ondersteunen. Om deze vraag te beantwoorden, zijn in dit werk drie onderzoeksrichtingen verkend. De eerste vraag onderzoekt de invloed van variabele afhankelijkheid op de nauwkeurigheid van kaarten die de geschiktheid van locaties voor aquacultuur weergeven. Daarom zijn de kansen op het behalen van optimale en kritieke overlevingsdrempels geschat, zowel met als zonder rekening te houden met de afhankelijkheid tussen variabelen. Er zijn verschillen tot

40% waargenomen tussen de twee benaderingen, wat het belang van het meenemen van afhankelijkheid in de modellering van betrouwbare maritieme ruimtelijke ordening onderstreept.

De tweede richting betrof de manier waarop GCBN's de veiligheidsbeoordeling van drijvende offshoreconstructies kunnen ondersteunen door het modelleren van afmeerspanningen onder specifieke hydrodynamische scenario's. Door spanningsreacties afhankelijk te maken van bepaalde hydrodynamische omstandigheden worden onzekerheden binnen de veiligheidsbeoordeling verminderd. De resultaten tonen aan dat onder zware omstandigheden de snapbelastingen contra-intuïtief lager zijn als gevolg van marginale verdelingen met een zware staart.

De derde onderzoeksrichting onderzocht hoe vine copula-modellen complexe afhankelijkheden tussen gemengde continue en ordinale variabelen kunnen weergeven die relevant zijn voor de structurele respons van drijvende offshoreplatforms. De resultaten tonen aan dat vine copula-modellen in staat zijn om het complexe afhankelijkheidspatroon per regime vast te leggen, d.w.z. per reeks parameters, waarbinnen de onderzochte structuur zich gedraagt als een stijf of flexibel lichaam. De vine copula-modellen bieden gedetailleerde inzichten in hoe structurele configuraties het gezamenlijke responsgedrag beïnvloeden en maken conditionele analyses mogelijk die helpen bij het identificeren van configuraties die door golven veroorzaakte spanningen in de structuur minimaliseren.

In alle hoofdstukken tonen de resultaten aan dat het expliciet in meenemen van multivariate afhankelijkheid leidt tot betrouwbaardere beoordelingen op het gebied van offshore aquacultuur, afmeerveiligheid en structurele responsmodellering. Meer in het algemeen benadrukken de bevindingen dat afhankelijkheidsmodellering niet louter een statistische verfijning is, maar een noodzakelijk onderdeel voor het begrijpen en beheren van offshoresystemen die worden gekenmerkt door op elkaar inwerkende milieu- en structurele processen. Door copulagebaseerde methoden in de huidige praktijk te integreren, draagt dit werk bij aan de ontwikkeling van robuustere, transparantere en fysiek onderbouwde kaders voor besluitvorming in opkomende offshore-technologieën.

STATEMENT OF NOVELTY

This dissertation presents a systematic application of probabilistic, dependence-aware modelling approaches based on copula theory to offshore and marine engineering and ecological problems that are traditionally addressed using deterministic methods or simplified dependence assumptions. Rather than proposing new probabilistic methods, the novelty of this work lies in the application, integration, and contextual adaptation of existing probabilistic frameworks to three distinct but related offshore problems. Across all chapters, the dissertation consistently produces decision-relevant probabilistic outputs, rather than deterministic indices or single design values. These outputs include (i) probabilistic ecological suitability maps for offshore aquaculture of mussels and kelp, (ii) a Gaussian copula-based Bayesian Network (GCBN) to describe the multivariate dependence between waves, currents, and mooring line tension rates relevant for snap loads, and (iii) vine copula models to represent complex dependence patterns between environmental forcing and structural response of flexible offshore floating photovoltaic (FPV) structures.

The novelty of the dissertation therefore lies primarily in the application and integration of probabilistic dependence modelling across different offshore domains, rather than in the development of new statistical theory. All methods and applications presented depend on data availability and quality, as well as on the selection of relevant variables. Transferability to other regions, species or structures requires appropriate re-fitting and validation using local data.

Chapter 2 aims to create suitability maps for offshore aquaculture. Existing offshore aquaculture suitability assessments are predominantly based on deterministic scoring systems or sustainability indices. Such approaches do not explicitly account for the stochastic nature of environmental variables or for the probabilistic dependence between them. This dissertation introduces a probabilistic suitability mapping framework for offshore aquaculture that explicitly accounts for dependence between environmental variables using bivariate copula models. The approach enables uncertainty-aware spatial assessments and is applied to the cultivation of Mytilus edulis and Saccharina latissima in the North Sea. To the author's knowledge, this constitutes a novel application within the offshore aquaculture and ecological suitability assessment community. Developed within the EU project UNITED, the methodology is stakeholder-driven and aligned with practical planning needs. The resulting probabilistic suitability maps can be used in practice to further identify suitable areas for mussel or kelp growth, e.g. in research and feasibility studies, before it later finds relevance for commercial-scale deployments. The framework complements, but does not replace, detailed ecological or site-specific assessments.

Chapter 3 focusses on mooring line tension rates under wave and current impacts. Previous studies have assessed mooring line tensions using conditional probabilities and

Monte Carlo simulations, or have modelled bivariate dependence between environmental variables to evaluate reliability. While the importance of dependence modelling has been recognised, the multivariate dependence between hydrodynamic forcing and mooring line tensions has not been explicitly addressed in a unified probabilistic framework. This dissertation introduces a GCBN to jointly model waves, currents and tension rates in mooring lines, allowing conditionalisation and therefore uncertainty reduction for given environmental conditions. To the authors' knowledge, explicitly accounting for multivariate dependence between environmental variables and mooring line tensions in such an approach has not been done before. The presented GCBNs are site- and system-specific. The structures of the GCBNs may be transferable, however, application to other locations requires re-fitting and validation using local hydrodynamic data.

Finally, Chapter 4 addresses the structural response of offshore floating photovoltaic (FPV) systems, which remains comparatively unexplored, particularly under marine conditions. Existing FPV research is largely focused on freshwater applications or deterministic numerical simulations, while only limited probabilistic studies exist and these typically address applications rather than structural response. As a result, probabilistic and dependence-aware frameworks explicitly linking environmental forcing to structural response of offshore FPV structures are lacking. This chapter introduces the application of vine copula models to describe the multivariate dependence between hydrodynamic variables and structural response variables, such as longitudinal stress and shear forces, of an offshore FPV system. The framework enables the derivation of probability distributions of stresses and shear forces, supporting risk-informed assessment and comparison of FPV configurations under uncertainty. Once further validated, the approach can be transferred to other offshore locations by adjusting marginal distributions while maintaining the dependence structure where appropriate. The analysis presented here is limited to wave-induced loading at a single offshore location and supports comparative risk assessment rather than detailed engineering design.

LIST OF PUBLICATIONS

JOURNAL PUBLICATIONS

1. **Santjer**, **R.**, Agarwal, S., Colomés, O., & Morales-Nápoles, O. (2025). Gaussian copula-based bayesian networks for dynamic loads in mooring systems. *Applied Ocean Research*, *165*, 104809. https://doi.org/https://doi.org/10.1016/j.apor.2025.104809
2. **Santjer**, **R.**, Mares-Nasarre, P., Vilmin, L., El Serafy, G., & Morales-Nápoles, O. (2024a). A probabilistic framework for offshore aquaculture suitability assessment using bivariate copulas. *Aquacultural Engineering*, *107*, 102479. https://doi.org/https://doi.org/10.1016/j.aquaeng.2024.102479

CONFERENCE PUBLICATIONS

1. **Santjer**, **R.**, Mares-Nasarre, P., El-Serafy, G., & Morales-Nápoles, O. (2024b). Copula based bayesian networks for hydrodynamic load estimation on offshore aquaculture systems. *Proceeding of the 34rd European Safety and Reliability Conference, Advances in Reliability, Safety and Security, Part 3*, 163–172
2. **Santjer**, **R.**, Mares-Nasarre, P., El-Serafy, G., & Morales-Nápoles, O. (2023). A case study of ecological suitability of mussel and seaweed cultivation using bivariate copula functions. *Proceeding of the 33rd European Safety and Reliability Conference*, 1877–1884. https://doi.org/10.3850/978-981-18-8071-1_P172-cd
3. Landmann, J., **Santjer**, **R.**, Gieschen, R., Goseberg, N., & Hildebrandt, A. (2019). Laboratory tests on wave forces and accelerations of a three-bay long-line aquaculture system in offshore conditions. *Proceedings of the 29th International Ocean and Polar Engineering Conference, ISOPE*, 2726–2732

CURRICULUM VITAE

Rieke Santjer

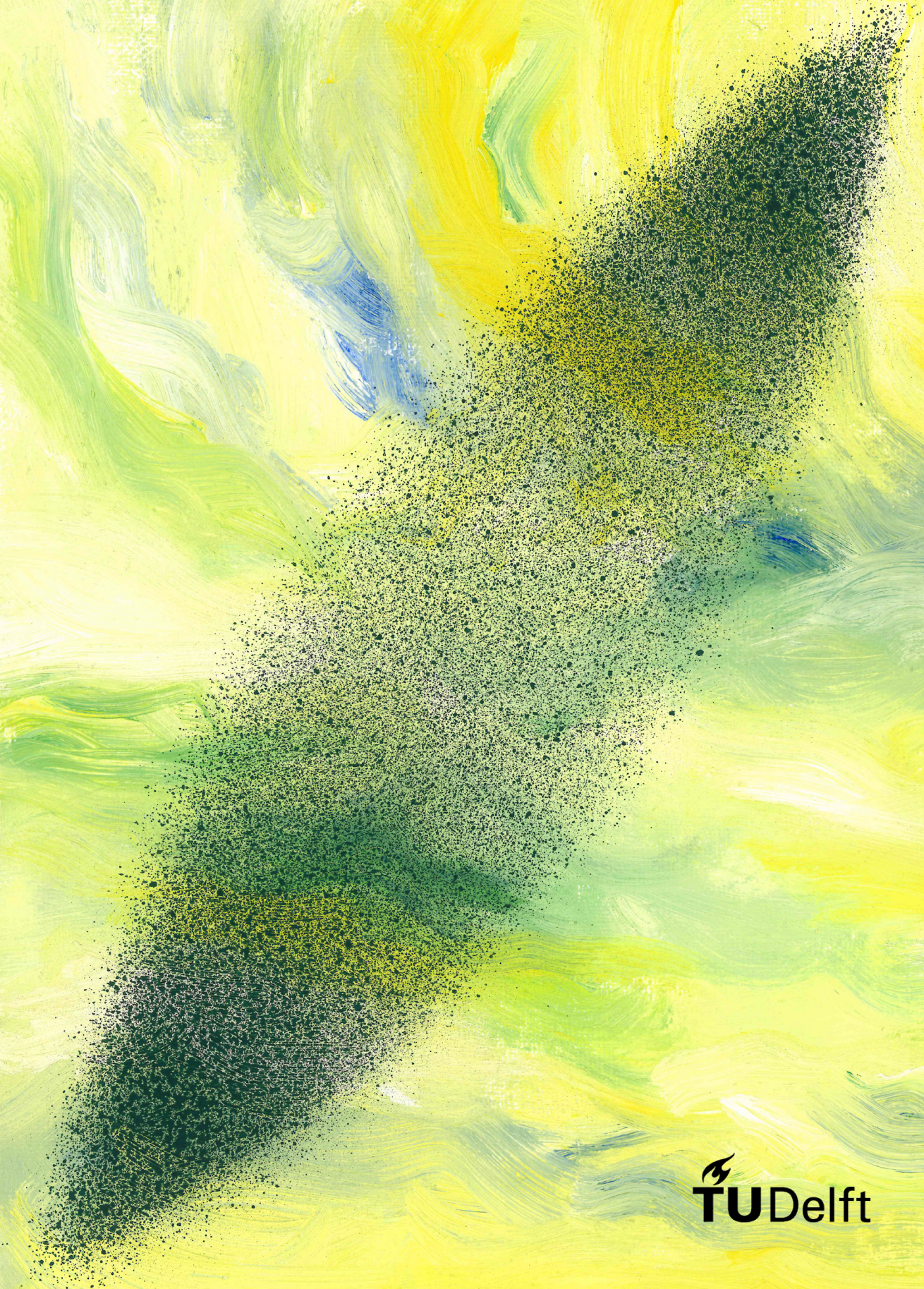
24-06-1995 Born in Aschaffenburg, Germany

EDUCATION

2013–2017 Bachelor of Science
Civil- and Environmental Engineering
Leibniz University Hanover, Germany
Faculty of Civil Engineering and Geodetic Science
Ludwig Franzius Institute of Hydraulic, Estuarine and Coastal Engineering

2016–2019 Master of Science
Water Management, Environmental and Coastal engineering
Leibniz University Hanover, Germany
Faculty of Civil Engineering and Geodetic Science
Ludwig Franzius Institute of Hydraulic, Estuarine and Coastal Engineering

2020–2026 Doctor of Philosophy Research
Probabilistic Multivariate Modelling with Copulas
Delft University of Technology, Netherlands
Faculty of Electrical Engineering, Mathematics and Computer Science
Delft Institute of Applied Mathematics


TUDelft