Analysis of Offshore Wind Turbines : Probabilistic assessment of operational design

A thesis submitted to the University of Dublin, Trinity College in candidature for the Degree of

Doctor of Philosophy in Civil, Structural and Environmental Engineering

by

Rui Teixeira

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Under the supervision of Professor Alan O'Connor and Professor Maria Nogal

June 2019

Declaration

I declare that this thesis has not been submitted as an exercise for a degree at this or any other university and it is entirely my work.

I agree to deposit this thesis in the University's open access institutional repository or allow the library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement.

Signature:

Date:

Acknowledgements

First, I would like to express my deepest gratitude to both my supervisors, Prof. Alan O'Connor and Prof. Maria Nogal, for allowing me to embark in this adventure that is a Ph.D work; for their constant guidance, their support and encouragement to overcome all the challenges that arose in past three years. Without them none of this would be possible.

Thank you Alan for you support, for selecting me in first place for such a special position without knowing me, and particularly for opening a new world of opportunities, always encouraging and allowing me to take this to its fullest potential. I believe this once in a lifetime opportunity is really unique, and I am deeply grateful for that. Maria, I do not think there is enough space here to thank you, but I am pretty sure that you know that I am deeply grateful for the numerous hours you have put into this work, for the motivating tone, for being always there, *etcetera etcetera*. Words will never make justice to it. Thank you, both.

I would like to thank also all my friends, from Dublin, Madeira and Porto for being the best in this journey. To all of you, many thanks for all the support during these three years. A special thanks to my colleagues here at Trinity, just for being amazing colleagues. For each one of you, you know. For my colleagues at TRUSS, many thanks, we were the right bunch of people.

Many thanks also to TRUSS team, Prof. Arturo Gonzaléz, Loreto, and Parisa and all the other partners involved. Thanks for bringing this project to the highest standards possible. I extend this gratitude to everyone that supported me in my secondments in London, Aberdeen and Paris, from Lloyd's Register and Phimeca. Thanks James, Antoine and Nandakumar I was lucky to work with you.

Finally, I would like to thank also my mother, my aunt, my cousins and all my family, for their support during all this process. For being far but not far. All I am and potentially achieve started 29 years ago in the small city of Machico, Madeira. My circumstances tell me that I am very lucky to have the family I have. I wish I could have been around much more time than I do. Many thanks for everything.

This Ph.D thesis work has received funding from the European Union Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No. 642453.

Abstract

The demand for renewable energy is unquestionable. Global climate trends have been fomenting the urge for renewable, low carbon-footprint technologies. Wind energy is one of the most prominent alternatives to conventional fossil fuel energy conversion. While the interest on wind energy dates back to the 1980 decade, offshore wind only became highly relevant later in the 2000's, with the realization of the potential within the offshore wind resource. The fact that wind depends on the Sun, makes it a virtually infinite source of energy.

In the case of offshore wind energy, it is remarkable the contribution and role of engineering practices and procedures to enhance the cost-reliability ratio for the harnessing systems. In this sector, research is expected to be the key driver of development up to 2050. While improvement of the techniques applied in the sector is on high demand in order to unlock new breakthroughs that will enable wind energy to become progressively more competitive, the development of new innovative practices is no less important.

In this thesis, design techniques for offshore wind turbines are evaluated and discussed. The work presented is strongly focused on the probabilistic assessment for structural design. A particular focus of the work is also directed to the cost of the design procedures, which recurrently hinder optimization at the design level.

Different contributions are presented in the following document. Extreme waves are evaluated using different probability models, motivated by an identified lack of consistency in the approach to wave extrapolation. Three probability models, Weibull, Exponential and Generalised Pareto, are analysed in a Peak-over-threshold analysis. An innovative choice of threshold methodology is introduced. This procedure relates to the density function shape. A logarithmic transformation to significant wave data is applied jointly with the Generalised Pareto model. Five fit indicator are used to compare the results, showing no evidence to reject any of the models studied.

Stress-cycle fatigue design, and its probabilistic basis is also discussed. Stress-cycle fatigue has been a widely studied problem in the wind engineering sector. Different probabilistic approaches are compared to assess it. Sample size influence on the loading spectra approximation is investigated and a bootstrapping procedure is applied to quantify the loading sample uncertainty.

The large design effort associated with the stress-cycle fatigue assessment motivates then the application of meta-modelling techniques for fatigue design. Gaussian process predictors, also known as Kriging models, are researched as an alternative for efficient stress-cycle fatigue assessments. Due to their inherent probabilistic character, these are also evaluated as potential quantifiers of uncertainty. A new indicator for space reduction considering uncertainty is presented in a global sensitivity analysis. Gaussian process predictors are applied using different methods to define the design of experiments. An innovative search function is introduced, that relates to the problem of stress-cycle fatigue analysis.

The developments presented are supported by an extensive literature review. Application of the design procedures are studied on a 5MW monopile baseline turbine.

Results from the work developed showed that some engineering practices still demand improvements in regard of their probabilistic comprehension. Are examples; the analysis of the probability model support; design approach and its probabilistic relation to the problem; size of representative design samples; or demand for notions of improvement that relate to the physical problem when meta-modelling.

It is of relevance to highlight that developments on uncertainty assessment for the wind energy sector have been steadily occurring through small contributions from multiple authors that spend significant efforts in order to comprehend what are the sources of error for each of the design steps. Aligned with this, the proposed work adds new insights to the fields of offshore wind turbine analysis and uncertainty characterization.

Contents

List of Figures

List of Tables

List of Symbols

α_{PP}	Linear regression slope angle.
$\mathbf{\Lambda}$	Probability model parameters.
$\boldsymbol{\theta}$	Gaussian process predictor hyperparamters.
Δt	Interval of time.
δ_i	*moment-free* measure of statistical divergence .
Δx	Imposed change in the generic variable x.
λ	Number of *exceedances*.
$\mathbb{E}[\,]$	Expected value.
$\mathcal{L}$	Likelihood function.
P[]	Probability.
μ	Statistical mean .
μ_{D_t}	Statistical mean of the short term damage rate .
μ_G	Gaussian process predictor mean.
ϕ_w	Wind direction angle relatively to rotor axis ($^\circ$) .
ψ	Infill criterion that uses W and P_s.
σ	statistical standard deviation.
σ_c^2	Gaussian process predictor constant process variance.
σ_G	Gaussian process predictor standard deviation.
σ_U	Standard deviation of U.
τ	Standard deviation of the Gaussian process predictor DoE noise.
$\boldsymbol{\beta}$	Linear regression coefficients.
$\mathbf{C}$	Covariance matrix.
$\boldsymbol{f_p}$	Linear regression basis functions.
$\mathbf{F}$	Regression matrix.
$\mathbf{R}$	Gaussian process predictor correlation function.
$\mathbf{X}$	Support sample in x.
$\mathbf{Y}$	Support sample image, $g(\mathbf{X})$.
θ_w	Wind direction.
φ	Probability model scale parameter.
ζ	Probability model shape parameter.
d	Dimension of the x space.

D_i Damage indicator .

D_t Fatigue SN damage in time t.

D_{KL} Kullback-Leibler discrimination .

$F(x, \Lambda)$ Theoretical probability distribution function.

$f_Y(x)$ Probability distribution function of Y that depends on x .

F_{ecdf} Empirical probability cumulative distribution.

$f_{Y^*(x)}$ Generic probability distribution function of Y^* that depends on x parameters.

h Transformation function.

$G(x)$ Gaussian process predictor of $g(x)$.

$g(x)$ True function that depends on x.

H_s Significant wave height (m) .

H_s Significant wave height.

I Turbulence Intensity (%) .

k Size of $\mathbf{X}$ variable.

L Load range (kNm).

l Number of divisions of the LHS space.

L_m Mean load range.

L_u Ultimate load strength.

L_{eq} Equivalent load range.

M Number of *exceedances* of load range data.

m SN fatigue exponent.

N_f Number of cycles until failure.

N Sample size.

n Number of cycles.

n_{eq} Number of equivalent cycles

$N_{f_{L_i}}$ Allowed number of cycles at load range L.

n_L Number of cycles at load range L.

n_{surf} Number of sampled surfaces.

p Number of terms of the polynomial regression.

P_s Search penalty spatial function.

r Number of years for return level estimation (years).

r_s Spearman correlation coefficient.

S Stress range (MPa).

S_m Mean stress (MPa).

S_u Ultimate stress strength (MPa).

S_0 Stress range at mean load 0.

T Lifetime design reference value.

t Time.

T_p Peak wave period.

T_r	Return Period.
t_{dc}	Declustering time.
U	Mean wind speed (m/s) .
u	Threshold value.
W_c	Cramér von-Mises GoF.
x	Environmental parameters.
x_{H_s}	Exceedances of H_s data in relation to a threshold value u.
z^*	Critical value for t-student distribution.
$Z_{1-\alpha}$	Standard normal distribution $1-\alpha$ probability associated value.
H_s	Significant wave height.
H_{s_r}	H_s with a return probability associated with r.
2D	Two-dimensional.
3D	Tri-dimensional.
ACF	Auto-correlation function.
AD	Anderson-Darling.
BB	Block bootstrap.
BEM	Blade element momentum.
CDF	Cumulative distribution function.
CFD	Computational fluid dynamics.
CI	Confidence interval.
COV	Coefficient of variation.
DEL	Damage equivalent load range.
DFF	Fatigue design factor.
DLC	Design Load Case.
DOE	Design of Experiments.
DOF	Degree-of-freedom.
DTU	Denmark Technical University.
EI	Expected Improvement.
EU	European Commission.
FEM	Finite-Element-Methods.
GED	Generalised wake model.
GEV	Generalised extreme value distribution.
GoF	Goodness-of-fit.
GP	Generalised Pareto.
GSA	Global Sensitivity Analysis .
IAHR	International Association for Hydro-Environment Engineering and Research
J	Kullback-Leibler divergence.
KL	Kullback-Leibler .
KS	Kolmogorov-Smirnov.

LCOE Levelized Cost of Energy
LHS Latin Hypercube Sampling.
LSE Least squares estimation.
MBB Moving block bootstrap.
MC Monte Carlo sampling.
MLE Maximum likelihood estimation.
MoM Method of moments.
MRL Mean residual life.
MSL Mean sea level
NTM Normal Turbulence Model.
P-GSA Probabilistic Global Sensitivity Analysis .
P[] Probability.
PDF Probability density function.
PWM Probability weighted moments.
RF Rainflow counting.
RMSE Root mean square error.
RNA Rotor-nacelle assembly.
SN Stress-cycle.
TP Transition piece.

Part I

Introduction

Chapter 1

Introduction

Contents

1.1 Motivation

The demand for renewable energy is unquestionable. Global climate trends have been fomenting the urge for renewable, low carbon-footprint technologies. Wind energy is one of the most prominent alternatives to conventional fossil fuel energy conversion. While the interest on wind energy dates back to the 1980 decade, offshore wind only became highly relevant later in the 2000's, with the realization of the potential within the offshore wind resource. The fact that wind depends on the Sun, makes it a virtually infinite source of energy.

Offshore wind energy experienced an exponential growth in installed power since the beginning of the current century. In particular, during the last decade, the increase on development of wind energy has been substantial. While this growing trend is expected to continue, further growth of the sector imposes more demanding, complex and accurate engineering methods. It is of major importance to ensure reliable operation in the sector.

In the case of Offshore Wind Turbines (OWT), the contribution and role of engineering practices and procedures for cost reduction and reliability enhancement is remarkable. In fact, IRENA [2016] highlights that innovation in wind turbine design and operation has been a key driver for competitiveness in the wind energy sector. Furthermore, Iván et al. [2017] identified that research, along with regulatory framework, are expected to be the key drivers for offshore wind up to 2050. In this context, while improvement of the techniques applied in the sector is in high demand in order to unlock new breakthroughs that will enable wind energy to become progressively more competitive, the development of new innovative practices is no less important.

Figure 1.1 shows that, despite the significant developments of the two last decades, the Levelized Cost of Energy[1] (LCOE) for OWT installations is just now achieving the level of economic competitiveness of the most conventional alternatives used to convert energy

[1] A ratio that includes the operational lifetime costs and divides it by the benefit of energy production.

(*e.g.* that use fossil fuels). Wind, nevertheless, is one of the most powerful resources available on earth.

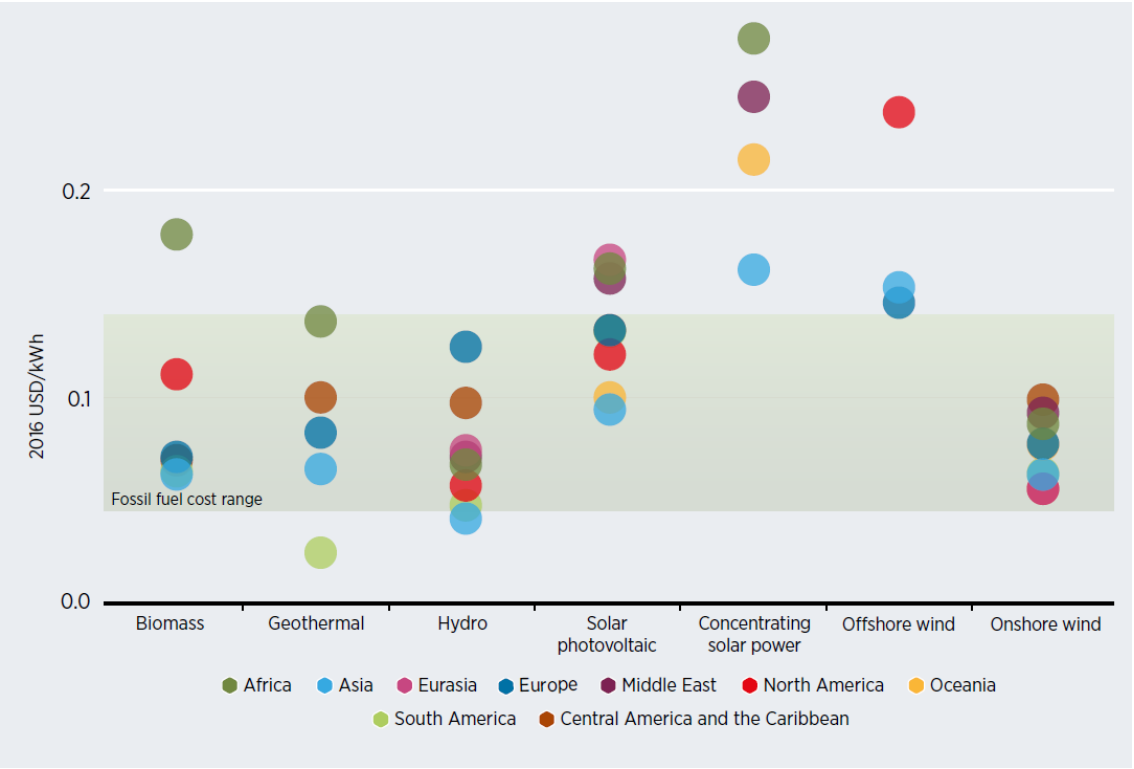


Figure 1.1: Levelized Cost of Energy (LCOE) for different sources of renewable energy divided by region. Source: [IRENA, 2017]

In the last decade, the scale-up of wind turbines has been the major driver of economic competitiveness for OWT. Larger turbines means more installed capacity and access to a larger resource, which contributes to amortise the project development costs faster. In particular, increase on the tower component height has been a major driver of the decrease in the LCOE for wind turbines [IRENA, 2017]. The fact that the technological solution to harness the wind power is well established, enables the sector to improve competitiveness through scale-up. Larger turbines, however, reflect on increasingly challenging engineering designs.

As the sector develops, an area of particular concern is the need to address uncertainty in the analysis and design of OWT. With better uncertainty characterization and understanding, the sector walks towards robust and optimized designs. For the designer, uncertainty characterization unlocks a new dimension of comprehension. Perception of the potential deviations experienced by the design variables, enables a more complete understating of the risk associated with the operation of an OWT.

Enclosing uncertainty in the design is usually enabled by probabilistic analysis, whose application to OWT is not a new practice. A brief analysis of the standards for designing OWTs (IEC 61400 class) or some recommended guidelines, such as DNV guidelines DNV [2014a, 2011], shows that assessing uncertainty is a recurrent process inside the OWT design chain. Although, and as mentioned, an important requirement still prevails in the field to progressively improve the established techniques, and to develop new methodologies that address topics where the subject, uncertainty, is not yet accounted for.

In regard to the OWT structural performance, some components are not allowed to fail. Other components should not, but may fail. This distinguishes two levels of relevance for the system's survivability per component, critical and non-critical for survival. Components that are not allowed to fail enclose most of the turbine cost breakdown. These are structural key elements, of larger dimensions, such as blades or tower. Failure of one of these elements is very likely to result in total loss of the system. Blades and tower alone enclose more than 40% of the costs for a 6 MW turbine [BVG, 2014]. On the other hand, turbines, according to IRENA [2017], comprise 64 to 84% of onshore, and 30 to 50% of offshore, wind energy projects costs. Foundations alone may account for $1/5^{th}$ of the project total cost. Design optimization and improvement for these components requires full comprehension of their response.

The motivation for the present PhD thesis is therefore to discuss some of the design methodologies that intervene in the design of these critical components, with particular focus on probability techniques and uncertainty characterization. The research conducted is applied to the tower component. Nevertheless, the discussion is extended on how the methods presented and discussed may be applied to other components. Uncertainty characterization is many times hindered by the costs needed to perform it. The existing design procedures for OWT are already highly expensive in regard of the amount of effort demanded for their implementation. Uncertainty characterization increases these design requirements exponentially. Uncertainty characterization, which is intrinsically related to probability theory, addresses repeatability of the design procedure up to a number of times were the designer may be confident that he/she comprehends the problem in-hand and the different uncertainties that affect it. This may range from a simple comparison to observed data, to analysis of the results consistency and repeatability.

Through efficient characterization of uncertainty further increase of competitiveness for the OWT sector may be achieved. Not only economic, but also in performance. Uncertainty though, is a broad and wide field of analysis and affects every single design step. Developments on uncertainty assessment have been steadily occurring through small contributions from multiple authors that spend significant efforts in order to comprehend what the sources of error are for each of these steps. Aligned with this, the proposed work adds new insights to the fields of OWT analysis and uncertainty characterization.

1.2 Context and current practice

OWT are designed accordingly to standards such as IEC [2009, 2005], DNV [2014a] and recommended practices such as DNV [2011]. To certificate OWT operation to be robust to a range of potentially threatening conditions, different loading scenarios need to be analysed. These are expected to cover foreseen operational conditions.

Commonly, failure of OWT structural vital components are divided in two main categories: ultimate strength (U) and fatigue (F) failure. Ultimate strength failure relates to a single event that exceeds the structural capability of the OWT to respond. Fatigue failures relate to common and repetitive occurrences that cumulate in time and cause the OWT system to fail.

The challenge with the ultimate strength failure is to foresee when this event, which may load the turbine over its capability to resist, may occur, and what is the expected intensity for it. Statistics provide the tools that helps designer in the definition of the intensity and the so-called return period of events that are beyond current records. This is just one example on how probabilistic analysis may allow the extrapolation of present

knowledge to characterize the "unknown". Probability techniques may enable uncertainty quantification, however, if not applied correctly, they may add further uncertainty to the procedure. Characterization of these low-occurrence high-intensity events is one of the current challenges in the engineering world, which is aggravated by the ongoing trend on the increase of frequency and intensity for these events due to climate change.

For fatigue analysis the problem is no less challenging. Although the characterization is easier due to their frequent character, loading conditions response and their relation to fatigue mechanisms are not understood to their full extent. An example of such is the fact that most of the fatigue calculation employed today use the cumulative damage rule, which has strong empirical basis (response to loading is calibrated through the usage of a stress-cycle curve). More recent theories try to approach fatigue by using multi-scale models, however, applicability of these methods to complex systems is still not close to being state-of-the-art.

When designing OWT, the effort needed to evaluate the fatigue design is a recurrently highlighted limitation. Lack of consistency in the design approach is also commonly identified.

The present PhD thesis debates issues related to both fatigue and ultimate strength failure analysis. In some probability principles, fatigue and ultimate strength analysis touch each other. Not only from the point of view of repeatability and consistency, but also on some of the probabilistic techniques applied.

It is known that the OWT sector demands reliable designs. These are strongly enforced by the need to characterize the system's response, and its uncertainty in the design phase. As a result, research on probability applied to OWT design is expected to contribute also to improvements on the cost required to perform design procedures. With no limitation on time and cost, uncertainty on an OWT design could be reduced to almost zero (although, never zero). However, in a competitive engineering world, this is not the case, and limitations in cost are significant. Therefore, the present PhD work addresses uncertainty with the awareness that efforts allocated to uncertainty characterization should not increase in the design phase, if not strictly necessary. Only within this context, new methodologies may be considered to further develop the sector.

To note that, despite the analysis performed being centred on a monopile foundation, the methods discussed may be applied to any other component. The only requirement for such is that a specified analysis per component is developed. Nonetheless, monopile OWT represented 87% of all OWT installations in 2017 [Iván et al., 2017]. These type of structures have played a major role in the development of the offshore wind sector.

1.3 Objectives

The ultimate goal of this thesis is to contribute to the understanding of uncertainty and statistical characterization within the design procedures for OWT. Uncertainty may have many sources. The designer approach to the design process was highlighted as one of the main sources of uncertainty for some design procedures during the emergence of the wind energy sector in 1990's.

The research conducted aims to highlight some of the shortages identified in the current design practices, but also to propose the application of new emerging scientific knowledge in the design of OWT. This is the case of all the techniques that intend to simplify engineering designs by exploiting the growing trend in computational power and availability. Computers are powerful machines that outpace humans in the capability of doing repeti-

tive tasks.

Three global objectives for the presented PhD work were outlined in a earlier phase of its development in order to tackle some of the gaps in the knowledge identified. They are:

- **Extrapolation of physical quantities.** The need to extrapolate physical quantities to lengths of time much larger than the available data is a probabilistic problem that concerns multiple fields of knowledge. Usually a data-set, if available for design, may cover few years. However, the designer of an engineering project is required to design a structure to survive periods of time much larger than the reference period of time of the data used to undergo design calculations. Situations are rare where long data-sets are available.

  Extrapolation, however, has its own challenges. If applied incorrectly, predictions for the events that are out of the data-set may be strongly conservative or non-conservative. While non-conservative designs are a concern for the designer, conservative designs imply that some extra, non-required, costs are being allocated unnecessarily. Therefore, extrapolation procedures should be accurate and robust. Accuracy and robustness are discussed in the present work, and extrapolation is discussed in both the context of ultimate strength analysis and fatigue failure analysis. In particular, issues concerning some of the existing design procedures and recommendations are highlighted, such as extrapolation of waves and loads or fatigue calculations.

  The goal of such an analysis is to contribute to the comprehension of the extrapolation procedures in the context of engineering design.

- **Uncertainty based on statistical procedures.** The simple fact that a probabilistic procedure is applied does not necessarily contribute to more accurate and robust designs. Design recommendations for OWTs enclose multiple statistical techniques that try to characterize uncertainty, extrapolation is only an example of such techniques. In some design procedures, authors such as Tarp-Johansen [2003] highlight that uncertainty in the design may be quantified by simply asking some experts to design a wind turbine and comparing the result. Heterogeneity in the calculations would be as large as the number of designers. Other authors, such as Sutherland and Butterfield [1993] highlight that results for a probability-based design procedure, in case of fatigue, could vary by multiple orders of magnitude.

  The second goal of the PhD thesis is therefore to discuss important aspects from statistical characterization. Literature review allowed the identification of questionable practices in the current design procedures. Limited interest is given to important probabilistic features, such as; bounded/unbounded character of probability models, identification of what is a representative sample, or truncated vs full data-set fitting goals. This PhD work tackles some of these important probabilistic issues in the context of the OWT design.

- **Effort reduction for design procedures.** Finally, the third goal of the underpinned work is to answer to the demand for procedures that enable designs to be optimized. Or, in other words, for procedures that cut the significant cost associated with some of the present design practices. The cost of some design procedures may hinder design optimization. In such cases, the computational cost required for

optimization is too large for the necessary number of iterations required to be performed. Optimization is then implemented using a series of assumptions that enable it to be practicable.

The current state-of-the-art in computational methods, however, indicates that some advanced modelling techniques can be applied to enable efficient design procedures without the need to compromise accuracy. New implementations on learning techniques and meta-models are at the present making their way to successfully solve complex engineering problems. Such techniques and methods, with particular emphasis on meta-modelling, will be researched as an alternative for the existing design practices.

One final consideration that is of major relevance for new implementations in the OWT sector is to relate all the discussed topics with the current design practices. None of the discussions aforementioned are valid if not contextualized in the light of the current design procedures.

The following section presents the organization of the thesis, structured in order to answer the proposed objectives.

1.4 Organization of the thesis

The following PhD thesis is divided in 7 Chapters, 4 that discuss the State-of-the-art and 3 that present the original contributions from the work developed.

- **Chapter 2: Design of OWT - An introduction**

  The aim of the Chapter 2 is to introduce the topic of OWT design in its broadest meaning. It gives an overview of the OWT design problem. OWT design is a multidisciplinary procedure that relies on know-how from multiple fields of engineering. These different fields are presented in Chapter 2 with the purpose of establishing the framework for the presented work. Moreover, current OWT design practices are significantly backed up by computational codes that allow cost-efficient design procedures. The original contributions are not directly related with the limitations that still hinder the accuracy of the computational modelling results. Nonetheless, the work presented is significantly backed-up by simulation codes. Hence, a detailed overview of computational simulation and the codes used to simulate OWT is given in this chapter.

- **Chapter 3: Characterization of high return period waves**

  Chapter 3 presents the state-of-the-art for high return period waves modelling. Some ambiguity on the probabilistic modelling for high return waves was identified in the literature. This motivated the necessity to extensively discuss the topic of wave extrapolation, and the different techniques used for it. Different works and standards that relate to wave extrapolation are discussed. The distinct models used to approximate extreme wave data are presented, with particular focus on the modelling of *exceedance* data. The considerations required for an accurate implementation of extrapolation methodologies are also discussed. The purpose of this chapter is to present the context that motivated the original contribution from Chapter 6.

- **Chapter 4: Fatigue Design of OWT**

Chapter 4 discusses the fatigue assessment for OWTs. Fatigue design is recurrently highlighted as challenging. Multiple techniques to analyse to fatigue are available, and are accepted in a design context. For many years wind turbine fatigue design has concerned those who implemented the calculation procedures. The main root cause for such concerns was the lack of consistency, when it comes to fatigue design for wind turbines. The most relevant works that addressed wind turbine fatigue design are extensively discussed in this chapter. Then, a comprehensive review of the design methodologies and the different variables involved in the fatigue design of OWT is presented. The sources of uncertainty that affect the fatigue design are discussed in detail in order to enable the reader to grasp the complexity inherent to the fatigue design. Allowing, therefore, him or her to understand further developments of the thesis in the circumstance where fatigue design is discussed.

- **Chapter 5: Meta-modelling of complex engineering problems**

  It was mentioned earlier that the computational time was a limitation that concerned OWT design. Dealing with computational cost is particularly challenging for fatigue design. Meta-modelling is one of the alternatives that showed the best potential to reduce the computational cost of engineering design procedures. It has been applied with success to many engineering fields, including OWT analysis.

  Chapter 5 presents and discusses different alternatives for meta-modelling in the context of OWT engineering analysis. The idea of a meta-model is to replace the computationally expensive codes that account for the most of the effort spent on the design procedure. In the case of fatigue analysis, the meta-model may be applied to take advantage of the fatigue statistical properties, such as convergence to the mean value. With the definition of a meta-model it is possible to interpolate the response of the OWT on its full-field operation without the need to eventually run the expensive code that characterizes its response. This is of major relevance today, since OWT design and optimization strongly rely on computational simulation. This chapter backs-up the implementation of a new approach to fatigue calculations presented in the original contributions.

  The following chapters present the original contribution of the work developed, following the gaps in knowledge identified in the previous chapters.

- **Chapter 6: Extrapolation of extreme waves with the Peak-over-threshold methodology**

  Chapter 6 is the first chapter of the original contributions and discusses the application of the Peak-over-threshold for significant wave extrapolation. The Peak-over-threshold is the recommended technique for extrapolation of significant wave data. It allows the extrapolation of waves even when limited data is available. In response to the identified divergences in regard of the methodological steps to extrapolate significant wave data, an extensive discussion on the extrapolation of this variable is presented. Data from four meteorological buoys installed in Irish waters are used to compute the results on extrapolation. Different fitting techniques are studied in order to compute robust results. An innovative methodology to establish the threshold value that defines the *exceedances* is proposed. It uses the PDF function shape in order to characterize the point where the tail region is expected to start.

  Three preferential probability models are compared for extrapolation. A logarithmic transformation applied to the wave data is also used to compute the results. Robust-

ness is also guaranteed by studying the influence of the sample size and the decoupling technique. Extrapolation results allow the identification of significant issues that relate to the physical problem of waves, such as the bounded and unbounded influence of some of the probability models commonly applied to extrapolate waves.

- **Chapter 7: Comparative fatigue design of 5MW baseline turbine**

  Chapter 7 discusses the current practices to design OWT for fatigue. It compares the different techniques used for this end. Amongst them, fitting a probability model to the data, truncation, and simple scale-up. The ultimate goal is always to approximate the fatigue of a much higher than t time, T using results from the t time. The fatigue of the tower component is studied.

  Evaluation of different SN fatigue calculation techniques considers a comprehensive discussion on its application, and on their purpose (for applying such procedure). The influence of small amplitude loading is researched. Analysis of the fatigue calculations considers 11 operational points on a first iteration. Approximation of the SN fatigue induced on approximately 170 hours of operation using results from a single hour of operation, is evaluated. Further research on different operational points is developed with an additional of 25 points that consider also random turbulence conditions.

  Sample size and the representative behaviour are identified to be of major importance to produce accurate results. These are also studied in detail, identifying that the minimum sample size recommended by IEC does not produce robust results for SN fatigue calculations.

  Finally, uncertainty due to sample size is studied. A bootstrap methodology is proposed to compute confidence intervals for the mean. Bootstrapping approximates well the confidence intervals and is particularly interesting for small number of simulations per load case. Asymptotic results of bootstrapping hold for different sample sizes. However the biggest advantage of such a methodology is the fact that it has no initial assumptions in regard of the loading distribution. Results are presented and discussed in the chapter.

- **Chapter 8: Meta-modelling in the design of OWT for fatigue**

  Chapter 8 explores the advantages of meta-modelling to design OWT to fatigue. The fatigue of the tower component is studied. It was mentioned that meta-modelling approximations showed promising results when surrogating complex problems. As in any meta-modelling problem, space reduction is first studied by analysing different operational points and their SN fatigue damage. Uncertainty is enclosed in the analysis using a probability distribution sensitivity indicator. A transformation of variable of particular interest for uncertainty characterization and reliability applications is implemented. Meta-modelling for fatigue design is more efficient if its implementation focuses on the variables that have relevant influence on the problem. According to Saltelli et al. [2008], in high dimensional spaces, frequently only a limited set of variables have influence on the output of the problem.

  A procedure to calculate fatigue using meta-modelling is then designed. LHS is studied as an alternative to set the design of experiments in order to apply a meta-model as a surrogate of the SN fatigue damage. It is noted that, when applying a meta-model, a "notion" of improvement is expected to improve the surrogate performance. A search function that uses a "notion" of improvement is compared

with the LHS and with the calculation of SN fatigue using binned environmental conditions. Results show that relating the meta-model with the problem of SN fatigue design improves the convergence of the results. A full one-year loading assessment is used to validate the results.

The meta-modelling researched considers a Gaussian process predictor. These models have showed large potential due to their flexibility as an interpolator for complex problems. Moreover, their Gaussian behaviour allowed the definition of a "notion" of improvement when defining the design of experiments to analyse. The Gaussian behaviour also has the potential to quantify uncertainty in the surrogate approximation, minimizing potential errors of the meta-modelling approximation. The approach to quantify uncertainty, and the findings of its implementation, are presented in the same chapter.

- **Chapter 9: Conclusions**

  The final chapter that closes this thesis presents the main conclusion of the work developed. The more promising points of the analysis developed are highlighted, and recommendations for future works presented. The synergies between the proposed methodologies and other areas of OWT analysis are also highlighted in this chapter.

Part II

State of the Art

Chapter 2

Design of offshore wind turbines: An introduction

Contents

2.1 Design of OWT

2.1.1 OWT design overview

Design of OWTs is regulated by different standards and endorsed support documents that specify a set of requirements for the design. The most widely accepted standards are given by the International Electrotechnical Commission (IEC) series 61400, which are complemented by other standards, such as Det Norske Veritas - Germanischer Lloyd (DNV GL) standards and recommendations, or the International Organization for Standardization (ISO) standards.

The IEC [2005] sets the requirements for the design of inland Wind Turbines, while the IEC [2009] sets the design requirements for OWT. There are design overlaps within these standards, it is not uncommon for the IEC [2009] to refer to the IEC [2005]. Furthermore, it is also common for the IEC [2009] to direct some of the more detailed analysis regarding OWT design to other standards such as ISO. Due to the inherent complexity of the OWT design, the terms "appropriate, reliable or adequate" are also frequently used in the design standards. This implies that different techniques may be used during the design phase, and these should have technical requirements verifiable by an independent third party.

Regarding the OWT design in operational conditions two major issues can be highlighted: the loading and the response. In the scope of the loading, the OWT analysis concerns the variables that are going to induce loading on the wind turbine, including their interaction with the system. All the fields of analysis that concern the response of the OWT to the loading variables are in the scope of response.

Figure 2.1 presents a simplified representation of the loading, response, some pivotal engineering concepts, and their interaction. Within the loading scope there are the dif-

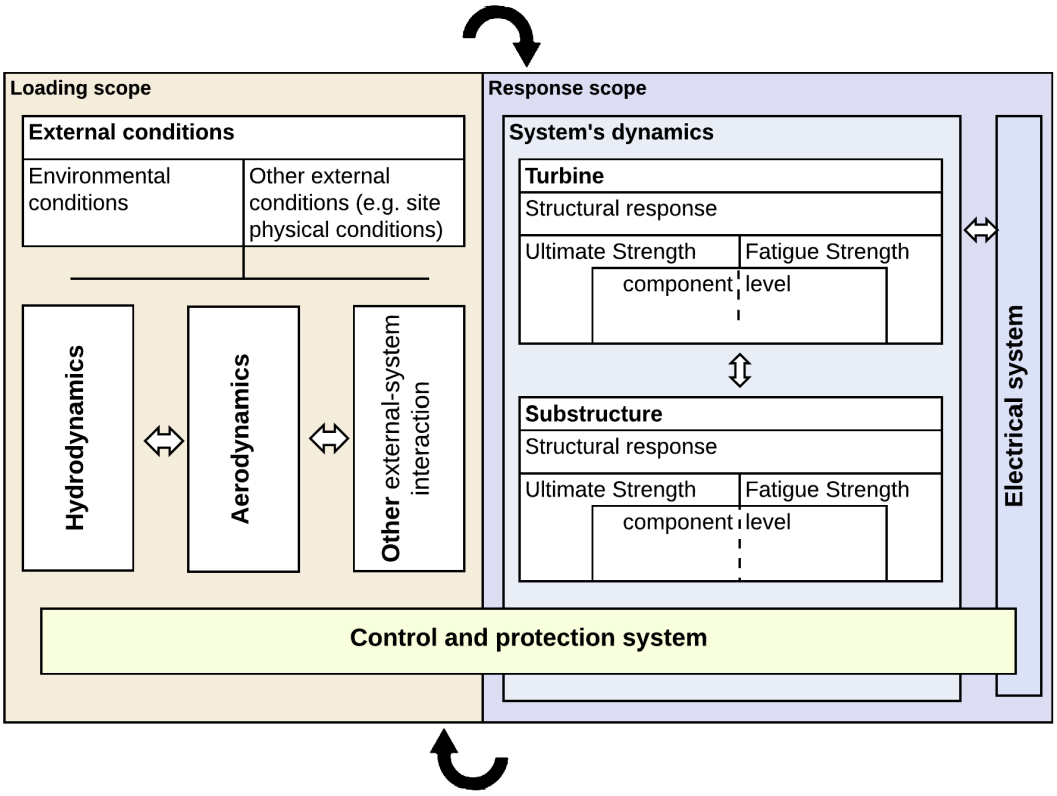


Figure 2.1: Design interaction of the loading and response for OWT operation regimes.

ferent environmental conditions that will interact with the system. These manifest in loading through hydrodynamics, for sea conditions, aerodynamics, for wind conditions or other external environmental or non-environmental conditions that generate other types of physical interactions (e.g. ice loading, vessel loading), and their coupled behaviour.

The response tackles the analysis, as the name indicates, of the system's response, and its dynamics. Here, most of the analysis addresses the response and its reliability. For design purposes structural response is frequently divided in two types of ultimate failure analysis; ultimate strength and fatigue. The "feedback" from the structural response is what then results in an electrical response, e.g. production of energy. The control and protection system overlaps over both, knowledge areas and their interactions. It can be identified on the system's dynamics by mechanisms such as yaw control, or pitch control; or in the electric system through generation control and; also in the environmental variables through protection control strategies such as the shut-down control.

The idea of defining the presented schematic division was to highlight the wide knowledge demanded for the design of OWT. It is no surprise for a design of a wind turbine to comprise extensive teams with a wide set of complementary know-how in different areas of knowledge.

The following chapters focus on some of the areas highlighted in Figure 2.1, which are particularly related, directly or indirectly, to the structural design of the tower and substructure component.

IEC [2005, 2009] recommends for the structural analysis of OWTs to be based on ISO [2015]. It was referred before that many OWT standards build on the ISO standards. DNV [2014a] for OWT, presents an organized structure of assumptions for the design, also

with strong support from the ISO standards. DNV also offers a range of recommended practices such as DNV [2011, 2014b] for the design of offshore structures that are of interest for OWTs.

In order to produce robust designs, the IEC61400 defines a set of Design Load Cases (DLC) that define the operational scenarios for the design procedure. The DLC enclose a set of operational conditions that may occur during the lifetime of the OWT. These are:

- DLC1 - Power production
- DLC2 - Power production plus occurrence of fault
- DLC3 - Start up
- DLC4 - Normal shutdown
- DLC5 - Emergency shutdown
- DLC6 - Parked (standing still or idling)
- DLC7 - Parked and fault conditions

These 7 main DLC classifications comprise a set of 31 subdivision that sum for a total of 31 DLC, and whose specifications can be consulted in more detail in IEC [2009]. The different DLC consider different operational conditions of loading and study the response of the designed turbine in the listed conditions. What is pursued with the design analysis is to infer on the survivability of the OWT to two types of structural failures: fatigue and ultimate strength. And what is tested is the loading that an OWT suffers under different assumptions, such as extreme mean wind speeds and turbulence intensity (I), gusts, misalignments or extreme waves and wave states. Given that these need to be defined, several assumptions were built over the years in order to set up a robust design procedure.

The standards must be seen as a dynamic document that builds up on cumulated knowledge on the field. Many times, new contributions to standard improvement come from research works. A very representative example can be found in Cheng [2002], where the author investigated extreme loads and concluded that the most significant operational loads for a pitch-controlled wind turbine (WT) occurred at mean wind speed (U) values slightly larger than the turbine's rated U. To characterize this problem of extreme loading a methodology for robust design in extreme loading is introduced. This methodology found its way to the IEC standard and is the current methodology to calculate operational extreme loading at the Rotor Nacelle Assembly (RNA), DLC1.1.

Moriarty et al. [2004] study on fatigue design using extrapolation techniques also appears on IEC [2005] - Annex F, where the methodology implemented is recommended for operational fatigue calculations.

The IEC 61400-1 edition 3 was updated since its launch date with different amendments, and the 4^{th} edition was recently published. Relevant new corrections to the standard involve examples such as the introduction of contemporaneous loads, which are the loads that a certain component is subjected at the same time as it is under an extreme loading event in one of its Degrees of Freedom (DOF) [Natarajan et al., 2008]. An updated version of this IEC61400-1 is expected to be published in 2018.

The most updated version of IEC61400-3 dates from 2009, it is its edition 1. Due to its high connection to offshore standards and the aforementioned IEC61400-1, the IEC61400-3 builds on wide knowledge basis. Yet, it is still under active discussion, and new techniques

to improve the design of OWT are demanded and continue to be formulated. Two major topics that relate to the design will be discussed in the present work.

In the design context, it is frequent for OWTs to have an experimental component, however, increasingly more often, computational simulation models are used to design OWTs, and in particular during the early stages of new developments. During the design phase multiple iterations may be demanded to build an OWT. Computational modelling is of interest due to its relative low cost and resource consumption when compared with other types of modelling. It has been a powerful tool to promote cost-efficient wind energy. Nevertheless, experimental validation of the simulation models is still required and mandatory for fully commercial developments. After the Technology Readiness Level 8 [1], development and approval of new wind farms is achieved using certified computational simulation models.

2.1.2 OWT computational simulation

The computational codes that analyse OWTs are commonly referred as aero-hydro-servo-elastic codes. The demand for these to be accurate is to some extend related to the fact that they show the potential to be significant enablers of further developments in the sector. Significant efforts to develop accurate aero-hydro-servo-elastic codes have been developed since the establishment of the wind energy sector, and resulted in the emergence of different codes to supply the necessity for simulation models. Reference computational codes in the field of OWT simulation are the FAST, HAWC2, GH Bladed, ADAMS and the FLEX5.

As the turbine moved from land to the sea, the complexity of the analysis increased. In a similar way than other technologies implemented offshore, different engineering solutions appeared to tackle the same challenge of installing the turbines at the sea.

This motivated the need to perform a comparative analysis of the different existing codes to design OWT in order to introduce consistency in the design aided by computational codes. Despite the existence of different codes, only a limited number of these were fully validated for OWTs. This is related to the relatively recent history of the sector and consequent limited availability of data.

One important benchmark initiative in this context was the promotion of the Offshore Code Comparison (OC3) project by the International Energy Association (IEA) [Passon et al., 2007], and further continuation with the OC4 and OC5 projects [Popko et al., 2012, Robertson et al., 2017].

Even considering that these codes may be relatively accurate in predicting the operational behaviour of OWTs (for example Veldkamp [2006] analysed aerodynamic modelling uncertainties related to fatigue design by comparing experimental and simulation data for the fore-aft moment of the tower using 500 operational 10 minutes simulations, inferring that computational prediction approximated the experimental data after correction due to the meteo-mast positioning in relation to the turbine) active research is being conducted to improve the approximation between experimental and simulated models. While two decades ago the effort was to achieve global predictions of the turbine behaviour, current trends in research involve the analysis of more specific details such as multiple turbine interaction, and detailed wind-structure interactions or wind-wave interactions.

Wang et al. [2016] present an overview of the different challenges faced by the aero-hydro-servo-elastic codes applied to model OWTs when modelling the blade component,

[1] Compreheensive definition can be found in the european association of research and technology organisation TRL reference document.

and justified by the ever increasing size of the wind turbine blades. One of the limitations identified is the limited capability of the blade element momentum (BEM), the most widely used modelling techniques for blades, in providing detailed aerodynamic information (such as characterization of the flow-field around the blade). It is important to highlight that although the blades being a critical component of the turbine, aero-hydro-servo-elastic codes comprise a much higher number of relevant components, such as the tower, the foundation, the generator, or the mooring system in the case of floating OWT. Table 2.1 presents relevant problems that have been the target of active research in the recent years, divided by main areas of analysis, in the simulation of horizontal axis OWTs.

Table 2.1: Recent relevant trends in the analysis of OWT modelling.

Field of research	**Research trends**	**Relevant works**
Wind simulation and Aerodynamics	Resource characterization, offshore wind physical description, wake modelling, eddy and vortex modelling	[Luo et al., 2015, Martínez-Tossas et al., 2015, Gupta, 2016, Kecskemety and McNamara, 2016, Holtslag et al., 2017, Sedaghatizadeh et al., 2018]
Wave theory and hydrodynamics	Non-linear wave modelling, wind-wave interaction	[Karimirad, 2013, Marino et al., 2017]
Geotechnics	Soil-structure interaction	[Carswell et al., 2015]
Structural Modelling	Fluid-structure interaction, FEM coupling, structural optimization, computational efficiency	[Hsu et al., 2014, Fischer et al., 2014, Bazilevs et al., 2015, Nezamolmolki and Shooshtari, 2016, Marino et al., 2017, Murcia et al., 2018]
Control systems	New control techniques for structural load reduction, and power optimization in individual or array analysis	[Fleming et al., 2016, Cortina et al., 2017, Petrovic and L. Bottasso, 2017]
Array systems	Turbine placement, Aerodynamic interaction	[Stevens, 2016, Cortina et al., 2017, Stevens et al., 2018, Zergane et al., 2018]

One of the identified trends is the application of more complex computational fluid dynamics (CFD) and Finite Element Methods (FEM) techniques in order to analyse the turbine dynamics. The application of both is expected to further increase the cost of the already expensive computational analysis of OWTs. Furthermore, in terms of structural analysis the increasing size of OWTs may also introduce important non-linearities in the structural response due to higher deflections, which in consequence, may increase the demand for more complex modelling techniques such as the ones mentioned.

Some works were identified that tackle the need to maintain designing effort practical [Marino et al., 2017, Murcia et al., 2018]. These relate not only to the fact that an increase is foreseen in computational demand, but also due to the current status of computational demand and offer. Despite the significant computational developments experienced in the last decades, significant time is still required to design an OWT using a computational simulation model. The design process is a problem of repetition, where multiple evaluations are demanded at the same operational conditions. Moriarty et al. [2004] performed 4725 computational simulations to achieve an accurate analysis of fatigue on a stall and a pitch controlled WT. Despite the high number of total simulations, only 9 simulations were performed per environmental state defined on a two dimensional grid of U and I.

As the dimension of the environmental grid increases, the number of simulations to be performed increases exponentially.

NREL's 5MW baseline wind turbine has played a major role in the development of new modelling and design techniques. This turbine has been extensively researched in the OWT context using NREL's code FAST and different foundations. A detailed overview of this code is given on Section 2.1.3. Understanding NREL's FAST is of major importance to comprehend the work presented in this thesis.

2.1.3 NREL's FAST aero-hydro-servo-elastic code

The FAST is a aero-hydro-servo-elastic code developed by the National Renewable Energy Laboratory (NREL). It uses multiple coupled modules to analyse the OWT dynamics, Figure 2.2. The software has been certified for wind engineering calculations for over 10 years [GmbH, 2005]. This code, which is open source, has been one of the most important stimulator of further developments in the research of WTs and OWTs. Five main sub-

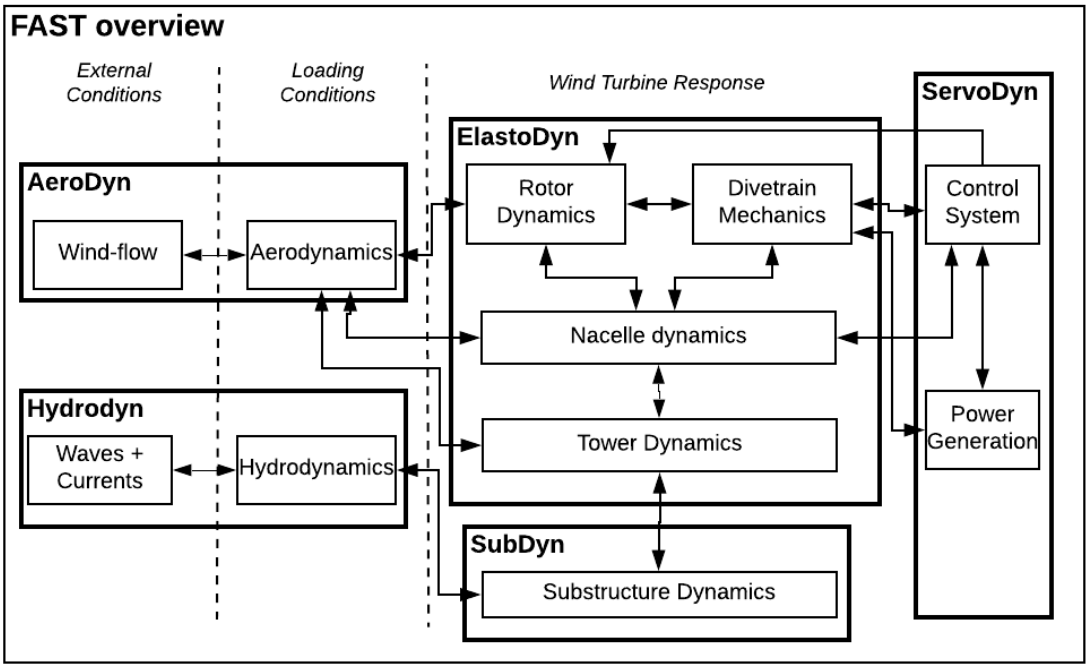


Figure 2.2: NREL's OWT modelling FAST (v8) software overview [Jonkman, 2013].

modules are currently part of FAST: the AeroDyn, the Elastodyn, the ServoDyn, the HydroDyn and the SubDyn. An overview and some important consideration on these are presented.

AeroDyn

AeroDyn is a time-domain aerodynamics code that uses BEM and a generalized wake model (GED) to calculate the loads on the blades and a potential flow solution around a cylinder plus, as well as its influence on the surrounding incoming flow. A non-linear drag term that depends on the square of the velocity is also used to calculate the loads due on the tower component.

Four main sub-models can be highlighted in the code composition: rotor wake/induction, blade airfoil aerodynamics, tower influence on the wind and tower drag.

The calculations in AeroDyn are performed using a local two-dimensional (2D) discretization at the cross sections both in the rotor and tower. These are then integrated along the length.

In the coupling with FAST, AeroDyn will assess the instantaneous structural position, orientations, and velocities in all the components of the turbine in order to iterate the solution to the aerodynamic problem. By joining the information provided by FAST's structural model with the free-stream wind velocities, the aerodynamic non-linear iterative calculations take place and are returned to the coupled FAST code to feed the ElastoDyn module. A detailed overview of the AeroDyn module is given on Jonkman et al. [2016] and Moriarty and Hansen [2005].

ElastoDyn

The Elastodyn module performs the structural analysis. It applies a modal or a multi-body dynamics analysis to calculate the structural dynamics of the coupled turbine system. It considers 24 DOF for three bladed horizontal OWTs in its calculations. Modal calculations are used for the tower and the blades considering the first vibration modes in the fore-aft and side-to side positions for the tower; a single mode for edge displacements of the blade and two modes for flap displacements, see Figure 2.3.

Full specification of the modal shapes and system matrices are demanded as an input for the structural dynamic analysis. The system dynamics is solved using a numerical integration to solve a set of equations of motion. Displacements, velocities, accelerations and reaction loading are output from the code to feed the remaining FAST modules.

The structural assumptions of the ElastoDyn present important limitations, such as the lack of axial or torsional DOFs, or the aforementioned assumption of small deflections. For the current state-of-the-art turbines these assumptions still produce accurate results. However, future trends may demand higher complexity in the structural simulation.

ServoDyn

ServoDyn module addresses the implementation of the control system, and its actuators. Four main independent control methods can be found on wind turbines and can be implemented in the ServoDyn module: pitch control, yaw control, high-speed shaft, blade-tip brake, and torque control.

The pitch control is one of the most common forms of control found in a wind turbine. In the past, turbines were controlled using stall (blade geometry was fixed and induced aerodynamic stall at high wind speeds) control, but, rapidly the pitch control overtook the sector for the multi-megawatt turbines. In pitch control each blade has an actuator that changes its angle of attack depending on the incoming wind and on the rotor angular velocity, controlling as a result the loading of the blades. The yaw control is responsible to position the RNA in relation to the wind speed. The high speed shaft and blade-tip brakes as the name indicates induce a braking torque on the shaft or in relation to the blade tip angle that can be used to simulate these brake systems in the case the simulated turbine has one.

The torque control is a system of major importance for wind turbines. It enables to impose a generator torque that optimizes the power extracted from the wind below rated U, and, with the pitch control, to stabilize it above the rated U. The energy available on

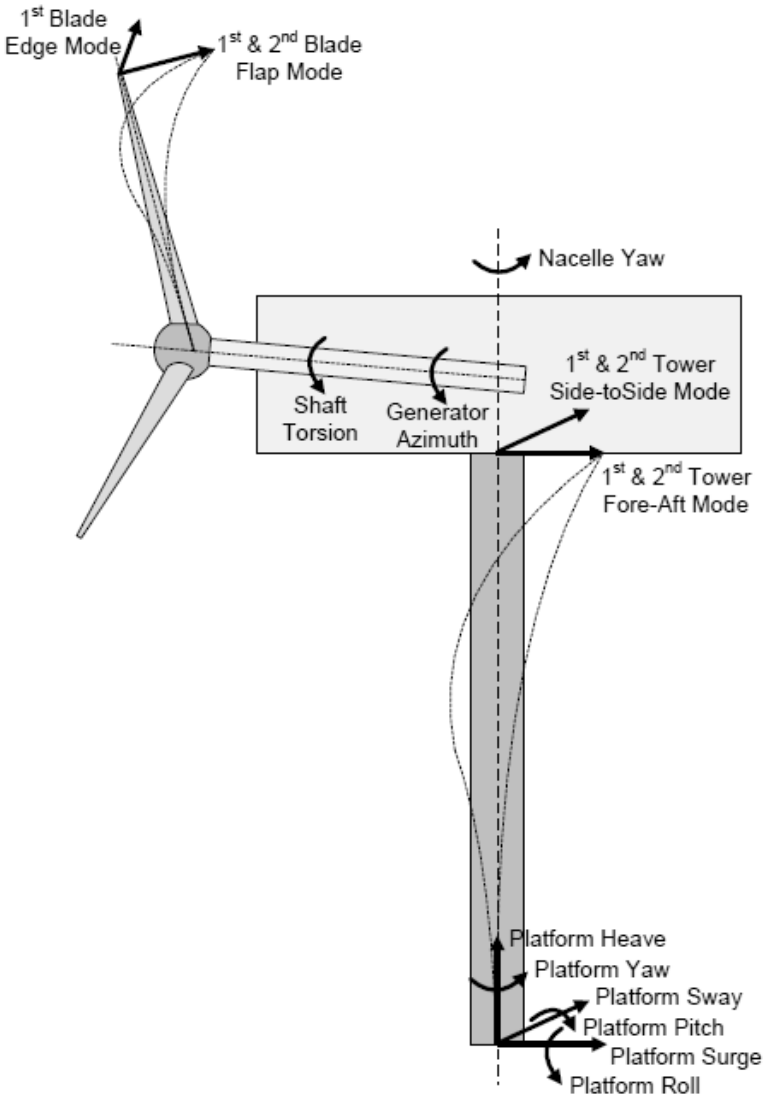


Figure 2.3: Overview of the DOF used to analyse the turbine dynamics in ElastoDyn (adapted from [Jonkman and Buhl Jr, 2005]).

the wind increases depending on U^3, however the turbine system has either structural or aerodynamic (noise due to rotation) physical limitations, which demands for the torque, rotor rotation and power to be constant above a certain wind speed. An example of power production control for an OWT is presented in Figure 2.4. The steady state control curve is complemented by an incremental time domain analysis of the turbine where the dynamic lag of the control response can be identified. In this case an incremental rate of $0.1m/s/s$ and a uniform wind field were used to compute the results.

In Figure 2.4 it can be seen that there are 3 main control regions for control during operation: 1 - start-up, 2 -optimization of power capture, 3 - constant power generation. In 1, there is no power production and the only objective is to accelerate the rotor. In 2 the goal is to maintain a constant tip-speed ratio so that power capture is optimized. In 3 the generator power is held constant by controlling the rotor speed, and the generator speed and torque. This is achieved using an inverse proportional gain control on these last

two quantities. The 1-1/2 control region is a transition control zone where the aim is to define minimum generator speed where energy production starts. The 2-1/2 is a direct gain region that relates to the induction generator that is needed to limit the turbine tip-speed ratio.

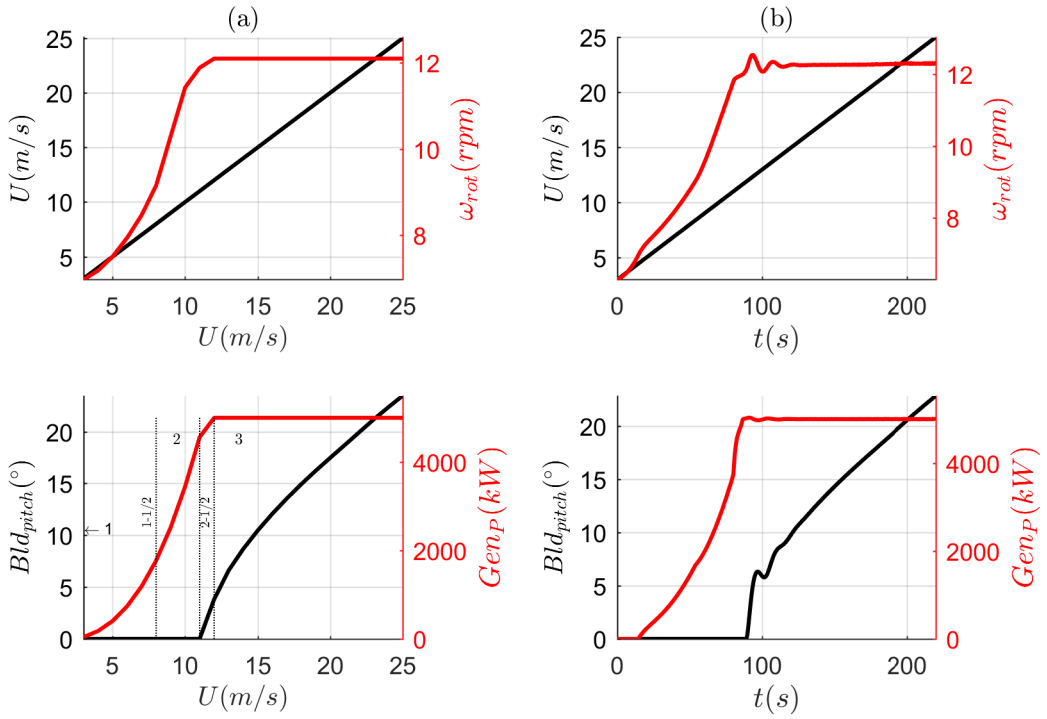


Figure 2.4: Example of OWT control procedure for power production. (a) Steady response control for [Jonkman et al., 2009]. (b) Example of time domain incremental U simulation and control response. ω_{rot} refers to the rotor speed, Bld_{pitch} to the blade pitch angle, and Gen_P to the generated power. Region 1 - start-up; Region 2 - optimization of power capture; Region 3 - constant power generation.

Further detailed information on the control system implemented in ServoDyn is given on Jonkman and Buhl Jr [2005].

HydroDyn

The Hydrodyn [Jonkman et al., 2014] module addresses the calculations of hydrodynamics for OWT analysis. It allows analysis of hydrodynamic bodies considering different approaches. Generation of waves uses 1^{st} (linear Airy) or 1^{st} plus 2^{nd} order theories (Sharma and Dean). It allows for regular and irregular waves.

On the hydrodynamic interaction, both the large body or slender body hydrodynamics can be assessed. While for large bodies there is a requirement for an auxiliary potential flow code that calculates the hydrodynamic coefficients for the body being analysed, radiation and diffraction effects. For the slender body calculations a strip-theory solution using the Morison's empirical formulation is applied to calculate both, the inertial and the viscous loading components.

SubDyn

The Subdyn [Damiani et al., 2015] is a FAST module developed for the analysis of non-floating foundations of the type, monopile, tripod or jacket. It uses a multi beam element model, combined with Craig-Bampton reduction.

For monopile foundations, or other foundations that include a transition piece (TP), the analysis above this component is not performed on Hydrodyn. In such cases, the dynamics of the substructure is "glued" to the ElastoDyn module at the TP through a rigid connection. It is relevant also to highlight that in this module soil-foundation interaction is not considered, and the OWT substructure is considered to be clamped to the seabed. Further developments in SubDyn are expected to tackle this limitation.

In the light of research, NREL 5 MW baseline turbine analysis using FAST has been extensively studied and constitutes a reference system for innovative studies on WT and OWT.

2.1.4 NREL 5MW baseline turbine on monopile foundation

The NREL 5MW baseline turbine has been seen as a reference turbine for OWT research since its definition in Jonkman et al. [2009]. The initial rationale behind this turbine was to set a reference size turbine to enable OWT studies. DTU would later introduce a 10 MW reference turbine [Bak et al., 2013], and the EU LEANWIND project would promote a 8 MW reference turbine [Desmond et al., 2016]. Nonetheless, regarding availability of information and incidence on the research field, NREL's 5MW OWT has captivated significant pertinence. It has been the reference turbine for offshore code comparison studies OC3, OC4, and more recently OC5 Robertson et al. [2017]. In particular, the application of this turbine on a monopile foundation has pioneered reference studies on OWT analysis, such as [Jonkman and Musial, 2010, Marino et al., 2017, Morató et al., 2017].

The main characteristics of the baseline turbine implemented on a monopile foundation are presented in Table 2.2.

Table 2.2: NREL's monopile OWT main generic characteristics.

Type OWT	Horizontal Axis with 3/63m blades
Rated Power	5MW
Cut-in and cut-out speed	3m/s, 25m/s
Hub height	87.6m above mean sea level (MSL)
Tower base height	10m above MSL
Tower section	Tubular
Foundation Tower interface (TP)	Rigid connection
Diameter at base of the tower	6m
Thickness at base of the tower	0.027m
Control system	generator-torque and a full-span rotor-collective blade-pitch control.
Foundation	Monopile 20m below MSL, clamped connection

The fact that it has been extensively compared within frameworks for the validation of OWT design aided by computational tools is of interest for research validation.

NREL's 5MW baseline turbine will be used to carry on the analysis proposed. Most turbines have a cut-in and cut-out wind speed, which gives the interval for mean wind

speed where the turbine is converting wind energy to power. Control sequences are demanded for start-up and shut-down of the turbine. These are not implemented on the baseline turbine and as such, attention should be given when considering the environmental conditions to analyse the turbine operation to not perform analysis that are physically inconceivable. The controller for the turbine's yaw action is also not implemented. The generator control is defined therefore and valid in between U values of 3 and 25 m/s.

To conclude, it is important to highlight that these preliminarily introduced concepts regarding OWT design procedure and modelling are of significant importance for the understanding of the following chapters and developed analysis.

Chapter 3

Characterization of high return period waves

Contents

3.1 Introduction

In the previous chapter, an overview of the OWT design techniques and some important computational simulation models were given. This mixed analysis was of interest due to the growing influence of the computational techniques on the design of this type of equipments.

It was also seen that OWT analysis and design involve multiple areas of knowledge which are permanently under discussion. The environmental trends indicate that the areas that deal with the definition of environmental conditions are expected to be one of these areas.

When the designer uses the current standards for implementation, (s)he should have some insight on the limitations that are behind the different techniques used. It is not infrequent to find recommendations such as, "the design should be reliable", in the standards, which is connected to this demand for the designer to have a comprehensive understanding of the problem and recognition by the standard that a technique may not be universal.

The DLCs foresee the usage of different environmental conditions, uncommon or normal, in production or faulty operational states. The calculations with these have the ultimate goal of finding the loads experienced by the OWT. While Cheng [2002] identified the importance of the normal operation for the WT loading, Morató et al. [2017] showed

that the DLC1.6a, which considers normal wind conditions jointly with severe sea state conditions, caused the highest loads on the substructure. Severe sea states are an example of an uncommon environmental conditions, which is expected to occur in this case once every 50 years.

The current chapter addresses the environmental conditions within the loading knowledge area, Figure 3.1. In particular, it is of interest to study wave conditions and the definition of high-return period waves which are expected to have a significant role in the design of OWT. In the design phase, the wave state that is expected to occur every 50 years may be required to characterize severe sea state conditions. It is very unlikely for 50 years of recorded or hindcast data to be available. As a result, definition of uncommon sea states is usually performed using statistical extrapolation techniques. To characterize

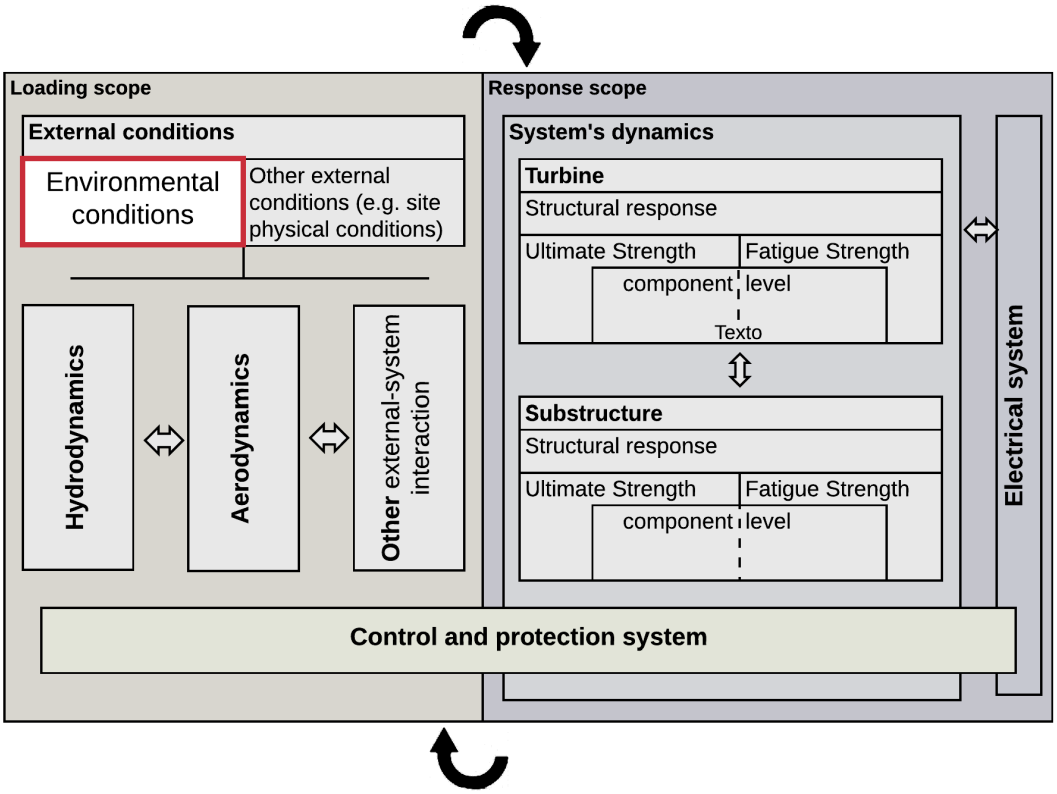


Figure 3.1: Design interaction of the loading and response for OWT operation regimes and positioning of analysis presented.

long-term statistics of sea states three main techniques are identified:

- Initial distribution method.
- Generalised extreme value theory of order statistics.
- *Exceedance* modelling.

The full distribution method is commonly described as the global model, both the extreme theory for order statistics and the *exceedance* modelling are classified as event models. A detailed discussion of the different methodologies used to analyse extreme sea states is presented in the present chapter, also known as high-return period sea states. Section 3.2 introduces the different techniques used to estimate high-return period sea states The

main relevant works and practices on the field are also discussed in this section, motivating the analysis on the Peak-over-threshold method presented in Section 3.3. Different considerations on the implementation of the Peak-over-threshold methodology are presented in between Section 3.3.1 and Section 3.3.6. Finally the main conclusions are drawn in Section 3.4.

3.2 Modelling high return period sea states

The problem of characterizing extreme occurrences is a statistical problem that has been widely discussed due to its significance. Extreme occurrences are located in the tail region of the probability density function (PDF). This tail region is mainly characterized by its limited statistical density that usually translates in limited availability of information. As limited statistical density exists in the tail region, it is understandable that limited data points may be available to characterize it when sampling from a population that follows a known distribution. To characterize extreme occurrences for design, the designer is expected to extrapolate the results from a limited data-set. This means that, when designing for an event that is expected to occur in 50 or 100 years (return period (T_r) of 50 or 100 years), it is very uncommon 50 or 100 years of recorded data to be available. In ideal conditions this should be the case, however, in most of the cases a much smaller time frame of recorded data will be available. Moreover, in some specific fields, such as climate change, even the existence of large data sets does not eliminate the need to work with complex statistical techniques to predict future trends.

For the case of significant wave height (H_s), application of different probability distributions can be found in the literature [Castillo et al., 2014, Nogal et al., 2016]. Due to the mandatory need to discriminate adequately the tail behaviour, it is common to consider subsets of data to characterize the tail. Using the whole data available from oceanographic monitoring records may result in an inaccurate approach to the tail region.

Global model

The global model considers the usage of all the data available to approach the distribution tail region. Considering all the data available increases the probability modelling accuracy, as more points are available to perform a probability characterization. The characterization of the tail region is usually poor within this method. As the relative density of the tail occurrences is smaller when compared with the density of common occurrences, less weight is given to the characterization of the tail region when performing a probability analysis (e.g fitting a probability model to characterize long-term wave statistics). This techniques might not be adequate for seas that are calm most of the time and experience storm events.

The event models address the need for accuracy in the tail region, and are currently the main techniques to characterize high-return period waves.

Order Statistics

Characterizing extreme events in a probabilistic way can be conducted through Extreme Value Theory. According to Extreme Value Theory, the maxima occurrences from a sample of order statistics with size n coming from a population with a given distribution function follow a known distribution. For instance, if a variable is assumed to follow

a Gaussian distribution, then its maxima follows a Gumbel distribution [Castillo et al., 2005]

The procedure for the usage of order statistics to define extreme occurrence is to identify a set of maxima values that can be extracted from a sample. This can be achieved by using techniques such as the Global Maxima, where a maxima is extracted from each available sample (e.g. yearly sample) or, the Block Maxima, where the sample is divided in subsets and then the maxima of each subset is extracted. The Global Maxima and the Block Maxima are similar techniques to extract maxima values from a time series. The Block Maxima can be understood as a modification of the Global Maxima technique in order to assess limitations that may be originated by the availability of data.

The main advantage of using this technique is that, for the limit case of a sample of independent maxima taken from a population, the relations between parent distributions and domains of attractions of extreme value distributions are fully characterized [Castillo et al., 2005]. The Generalised Extreme Value distribution (GEV) characterises the three forms taken by the extreme value distributions. An example of a GEV for a the standard normal distribution is presented in Figure 3.2. It is known that the empirical distribution for the normal maxima approaches asymptotically a Gumbel distribution, therefore the GEV distribution approaches its Gumbel form. This Gumbel distribution represents the probability distribution that the maxima from a given normal standard sample of size n is expected to follow.

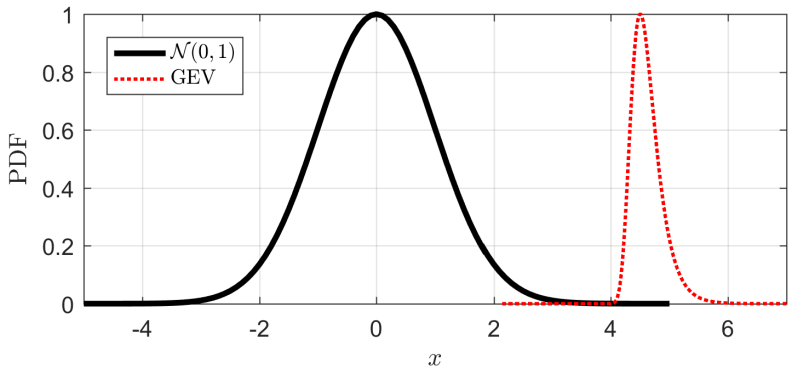


Figure 3.2: Standard normal distribution and associated generalized extreme value distribution (GEV) of maxima (PDF is divided by the maxima of PDF for representation).

However, it is not always of interest to use the limit cases for order statistics. An efficient characterization of the tail region can be obtained by focusing only on the analysis of occurrences that exceed a certain threshold value of u. Similarly to the development of the Block Maxima, this alternative is of relevance to model extreme occurrences with precision as it may allow the usage of more data points when compared with the limit distributions of order statistics.

Exceedance data

When considering significant wave heights (H_s) taking into account large values other than the maximum is important to overcome any limitation that may be originated by the high temporal variability naturally subjacent to the ocean waves. Considering a single or even multiple maxima disregards the fact that during some time intervals (e.g. a year) the maxima can be smaller than other order statistic for alternative time intervals. If

the aim is to analyse extreme wave data it is important then to analyse the *exceedances* over a specified threshold value u.

[Nogal et al., 2016] discusses the probability modelling of different engineering variables, including maxima wave data and *exceedance* wave data, referring that the Generalised Pareto (GP) distribution should be applied to model *exceedances* of significant wave data.

The GP distribution is commonly associated with the modelling of *exceedance* data. This is due to the fact that when u is sufficiently large, the exceedances are expected to follow a GP distribution [Pickands III, 1975].

The most common technique to model extremes considering *exceedances* over a certain threshold u is the Peak-over-threshold (POT). This technique involves truncating independent occurrences that respect the condition $> u$. An alternative less common method is the usage of stastitical distribution to model the waves within a stormy sea. Such an approach may be of interest to estimate individual wave heights.

POT and GEV

The need to characterize extreme waves using field data backs to the 1960s [Draper, 1963], where a "design wave" was characterised using full records of data. Since then, the topic of extreme wave or "design waves" remained under discussion for the decades to come [Nordenstrøm, 1969, Petruaskas et al., 1970, Battjes, 1972, Jahns et al., 1972, Forristall, 1978].

A comprehensive review comparing the different methods to extrapolate wave heights was presented in Muir and El-Shaarawi [1986]. The usage of GEV and storm data to approximate extreme waves is discussed in the paper. The authors compare 6 methodologies to estimate extreme waves concluding that none is indubitably more robust for universal applications. The application of the GP distribution to fit the probability tail of wave data was mentioned as untried and unproven in the context of fitting the wave heights within a storm. Modelling *exceedances* over a certain threshold u is discussed through the application of compound distributions, such as the Poisson-Rayleigh or the Poisson-Gumbell distributions.

At the time, different methodologies could be found to estimate extreme sea states, which motivated Goda [1989] to highlight the need for consistency in the calculation of extreme significant wave heights. The author discusses then extrapolation of wave heights with specific incidence on "design waves", comparing the usage of the POT within different probability models, and stating that the POT application should be prioritized if enough data are available for statistical analysis.

Due to the persistence of a lack of an uniform approach to model significant wave data and its extreme occurrences, in 1994 experts in the statistical treatment of wave data were gathered by the International Association for Hydro-Environment Engineering and Research (IAHR) to discuss and standardize the modelling of extreme H_s. The POT methodology and Weibull distribution were recommended as the reasonable choices for most oceans. Although, it is highlighted in the technical document produced that there is no theoretical argument to indicate a preferred model distribution to fit the observations of maxima. The main findings were published in a recommended practice for extreme wave analysis [Mathiesen et al., 1994].

Currently, several standards and practices are found to guide the design of offshore structures. While, some of them present generic considerations related to the definition of the extreme H_s occurrences, e.g. emphasizing the need of reliable and robust estima-

tions, other codes provide specific recommendations about the techniques to apply when modelling extreme H_s. Det Norske Veritas recommended practice on environmental conditions and environmental loads DNV [2014b] accepts the usage of both, the global and event model. It previously rejected the application of the GP to model exceedances of H_s, recommending instead the usage of the Weibull and exponential distributions. This recommendation, used as reference for offshore standards, such as [DNV, 2014a], was recently amended. Nevertheless, the application of the GP distribution continues to not be equated at the same level of efficiency of other alternatives, and is characterized as of sensitive choice.

Ferreira and Guedes Soares [1998] compared the application of the Exponential and the GP distribution to fit H_s POT data and decided on the application of the former.

In opposition to DNV recommendations, Hawkes et al. [2008] state that the GP distribution appears to be the optimal regional fit for the extreme wave heights in the North Sea. Regional Frequency Analysis is applied in Van Gelder and Mai [2008] to analyse extreme wave heights in Dutch North Sea Coasts in which five different probability distributions were evaluated in a fitting process. Wave data are analysed with POT. Results showed that GP distribution fitted better the data analysed when comparing with the other distributions tested.

In Méndez et al. [2006] the POT is used in combination with the GP distribution to evaluate the long-term trends in the frequency and intensity of severe storm waves. With the same basis, in Cañellas et al. [2007] the GP distribution is applied to probabilistically model extreme wave heights in the Balearic Sea. Teena et al. [2012] compared the POT with the GEV, concluding that the first is recommended in areas where storms are prevalent.

Relevant works that discuss the characterization of the long term H_s statistics are presented in Table 3.1.

Table 3.1: Relevant works that discuss the modelling of extreme significant wave data.

Method	**Distribution**	**Relevant works**
Initial distribution method	lognormal	[Draper, 1963]
Initial distribution method	Weibull	[Nordenstrøm, 1969, Battjes, 1972]
Exceedance data	Gumbel, Weibull	[Petruaskas et al., 1970, Goda, 1989]
Exceedance data	Rayleigh	Jahns et al. [1972]
Exceedance data	Rayleigh, Weibull	[Forristall, 1978]

Continued on next page

Continued from previous page

Order statistics and *exceedance* data	GEV, lognormal, Weibull fit to maxima and Rayleigh, Poisson-Gumbel fit to exceedances	[Muir and El-Shaarawi, 1986]
Initial distribution method, Order statistics, *exceedance* data	Weibull, Gumbel	[Mathiesen et al., 1994]
Exceedance data	GP, Exponential	[Ferreira and Guedes Soares, 1998]
Exceedance data	GP	[Castillo and Hadi, 1997, Méndez et al., 2006, Hawkes et al., 2008, Bommier, 2014]
Exceedance data	Weibull	[Stephens and Gorman, 2006]
Order Statistics and *exceedance* data	GP and GEV	[Cañellas et al., 2007, Teena et al., 2012, Young et al., 2012]
Exceedance data	GEV, GP, lognormal, Pearson type III, Generalised logistic	[Van Gelder and Mai, 2008]
Order Statistics	GEV	[Menéndez et al., 2009]
Exceedance data	Weibull	[Goda et al., 2011]
Initial distribution method, Order statistics, *exceedance* data	Weibull, lognormal, GEV, GP	[Holthuijsen, 2010]

It is interesting to notice that even considering some rejection from reference documents, the GP distribution is still widely applied to characterize extreme waves. An increase in the usage of the POT after Mathiesen et al. [1994] paper can be identified. The POT is currently widely accepted as a methodology to model long-term H_s data, and is the most applied for the effect.

The interest of using *exceedances* is that during the analysis it is possible to focus on the tail region taking the most of the data available. It mitigates the lack of robustness of the limitation of the initial distribution method in the approach of the tail region, and

the limitation of the limit distributions of order statistics disregarding important H_s data to characterize the tail region (e.g. similar extremes in the same interval used to account for the maxima).

Important considerations for the usage of the POT are given in the following section.

3.3 POT probability characterization of H_s

If H_s is a random variable with maxima m_{H_s}, then for all $u < m_{H_s}$, the function

$$F_u(x_{H_s}) = P[(H_s - u) \leq x_{H_s} | H_s > u], \qquad x_{H_s} \geq 0 \tag{3.1}$$

is called the function of *exceedances* of H_s, which can be also defined as

$$F_u(x_{H_s}) = \begin{cases} \frac{F(x_{H_s}+u)-F(u)}{1-F(u)}, & \text{if } x_{H_s} \geq u \\ \\ 0, & \text{if } x_{H_s} < u \end{cases} \tag{3.2}$$

a normalised representation of $F_u(x_{H_s})$. Analysis of x_{H_s} comprises the definition of a subset $H_{s(>u)}$ of H_s that respects the condition $H_s > u$. Pickands III [1975] showed that the limit probability distribution of this subset approaches a GP distribution. Nonetheless, despite the GP mathematical affinity with the POT methodology, POT *exceedances* are also modelled in the literature with alternative probability models, such as Weibull or exponential probability models.

3.3.1 Probability models to characterize exceedance data

The GP distribution models the conditional probability of an observation exceeding a value of x_{H_s}, given the condition that $H_s > u$. It is common for the GP distribution to appear associated with a probability process where the occurrences are assumed to follow a Poisson distribution with a λ expected number of occurrences. This fact will be discussed in the following section and is related to the independence of data.

Its continuous CDF is characterized for having two main forms that vary accordingly to

$$F(x_{H_s}, \varphi, \zeta) = \begin{cases} 1 - \left(1 + \frac{\zeta x_{H_s}}{\varphi}\right)^{-\frac{1}{\zeta}}, & \zeta \neq 0, \\ \\ 1 - \exp\left(-\frac{x_{H_s}}{\varphi}\right), & \zeta = 0, \end{cases} \tag{3.3}$$

ζ and φ being respectively the shape and scale parameters of the function and x_{H_s} the exceedances $H_s - u$ as previously highlighted. If $\zeta = 0$ the GP distribution becomes the exponential distribution, which means that the GP distribution approximates the exponential distribution when $\zeta \to 0$.

ζ also defines the GP distribution support, which is of relevance to model extreme waves or any physically limited phenomena,

$$\text{GP support} \begin{cases} x_{H_s} \geq 0, & \text{if } \zeta \geq 0 \\ \\ 0 \leq x_{H_s} \leq u - \frac{\varphi}{\zeta}, & \text{if } \zeta < 0 \end{cases} \tag{3.4}$$

The GP distribution is stable to truncations from the left [Castillo et al., 2005]. If x_{H_s} follows a GP distribution, then $x_{H_s} - u_i$ with $u_i > u$ is also a GP distribution. u_i are any

larger than u, and smaller than m_{H_s}, truncation values. Other interesting properties of GP distribution are:

- if $\zeta = 1$, the GP is the Uniform distribution $U[0, \varphi]$.
- if $\zeta = -0.5$, $Var(x_{H_s}) = \infty$.
- if $\zeta < 0$, the GP becomes the Pareto distribution.

An alternative identified and currently widely used to model POT *exceedances* is the Weibull distribution. It is also the distribution recommended to be applied in conjunction with the POT by reference documents such as, [Mathiesen et al., 1994] or [DNV, 2014b]. The Weibull distribution, introduced in Weibull [1951], is described as a probability model of wide applicability. Its most general formulations are characterized by two and three parameters. Both consider shape and scale parameters, however, the 3-parameter Weibull has an additional location parameter. In the Extreme Value Theory of order statistics the 3-parameters form is also the specific case of the GEV distribution for minima. For maxima the GEV, has the denominated reverse Weibull form.

Its continuous CDF formula is given by

$$F(H_s - u, \varphi, \zeta) = \begin{cases} 1 - \exp\left[-\left(\frac{H_s - u}{\varphi}\right)^{\zeta}\right], & H_s - u \geq 0 \\ 0; & H_s - u < 0 \end{cases} \tag{3.5}$$

where ζ and φ are the shape and scale parameters. For applications with the POT for H_s data analysis, the idea of *exceedance* should be used and therefore the location parameter will take the value of u (as shown in Equation 3.5). Consequently the Weibull with 3 parameters, reduces to a 2 parameter probability model where the location parameter is assumed to be known *a priori*. $H_s - u$ is replaced as x_{H_s}, i.e. exceedance data obtained from the POT methodology.

The Weibull distribution has support in $[0, +\infty)$.

3.3.2 Estimation of ζ and φ

The application of the probability models described depends on a set of parameters, also known as the probability model DOFs. Different parameter estimating techniques may be applied to define the DOFs of the different probability models. Extensive literature addressing the topic of parameter estimation can be identified. Castillo et al. [2005] present an extensive discussion of the probability distribution fitting methodologies. Castillo and Hadi [1997] also extensively discuss the different methodologies to calculate the DOFs, having a special emphasis on the application of GP distribution.

The estimation techniques used to define the probability models may have a key role in the analysis. In the design process, the sensitivity to these may be addressed. Sánchez-Arcilla et al. [2008] compare the effects on the extrapolation results caused by the application of two different estimation techniques, a maximum likelihood estimation (MLE) and a Bayesian approach. In the example considered, no major divergence was found in the extrapolated levels for both techniques.

Three common techniques of simple implementation are used to estimate the probability model parameters are: the MLE, the method of moments (MoM), or the least squares estimation (LS).

MLE

The MLE technique may be applied to find the DOFs of any of the models described. The MLE estimates the unknown parameters of a given continuous statistical distribution based on its likelihood function, defined for x_{H_s} as in 3.6.

$$\mathcal{L}(\mathbf{\Lambda}|x_{H_s}) = \prod_{i=1}^{n} f(x_{H_s}, \mathbf{\Lambda}), \tag{3.6}$$

$\mathbf{\Lambda}$ is the vector of parameters to estimate, which for the Weibull and the GP is a vector composed of ζ and φ. $f(.)$ is the derivative of $F(.)$, the PDF. The MLE problem is defined as,

$$\max_{\mathbf{\Lambda}\in\mathbb{R}} = \mathcal{L}(\mathbf{\Lambda}|x_{H_s}) \tag{3.7}$$

This method maximizes the likelihood functions $\mathcal{L}(\mathbf{\Lambda}|x_{H_s})$ to find $\mathbf{\Lambda}$. For practical applications it is of interest to consider the log-likelihood functions, Table 3.2. Maximizing the likelihood provides the same result as maximizing its logarithm. Most of the MLE

Table 3.2: MLE estimation functions.

Distribution	**log-likelihood function**
Weibull	$n\ln\zeta - n\zeta\ln\varphi - \sum_{i=1}^{n}\left(\frac{x_{H_s}}{\varphi}\right)^{\zeta} + (n-1)\sum_{i=1}^{n}\ln x_{H_s}, \quad \Lambda = [\varphi, \zeta]$
GP($\zeta \neq 0$)	$-n\ln(\varphi) + \left(\frac{1}{\zeta} - 1\right)\sum_{i=1}^{n}\ln(1 - \zeta x_{H_s}/\varphi), \quad \Lambda = [\varphi, \zeta]$
GP($\zeta = 0$)	$-n\ln(\varphi) - \varphi\sum_{i=1}^{n} x_{H_s}, \quad \Lambda = [\varphi, \zeta]$

estimation problems are only solvable numerically. In other cases a closed form function is enough to estimate $\mathbf{\Lambda}$, e.g exponential distribution. Its applicability may have limited $\mathbf{\Lambda}$ domain, nevertheless the MLE is the most efficient parameter estimation methodology. By analysing the form of the likelihood function, and its dependence on the PDF, it is possible to understand that what the MLE is searching is for the best combination of parameters of a PDF and CDF function that better describe the density of the points in the x_{H_s} sample. The MLE can also be understood as the the calculation of the parameters $\mathbf{\Lambda}$ that approximate the likelihood of getting the observed data through the theoretical distribution used.

MoM

The MoM uses the statistical moments of x_{H_s} and the theoretical distribution in order to find $\mathbf{\Lambda}$. Knowing that the first two moments of x_{H_s} are $\hat{\mu}_{x_{H_s}}$ and $\hat{\sigma}^2_{x_{H_s}}$ respectively, the estimator of MoM are given in Table 3.3.

Table 3.3: MoM estimation functions for the Weibull and GP distributions (Γ is the gamma function).

Distribution	$\hat{\mu}_{x_{H_s}}$	$\hat{\sigma}^2_{x_{H_s}}$
Weibull	$\varphi\Gamma\left(1+\frac{1}{\zeta}\right)$	$\varphi^2\left[\Gamma\left(1+\frac{2}{\zeta}\right)-\left(\Gamma\left(1+\frac{1}{\zeta}\right)\right)^2\right]$
GP($\zeta \neq 0 \ \cap \ \zeta < \frac{1}{2}$)	$\frac{\varphi}{1-\zeta}$	$\frac{\varphi^2}{(1-\zeta)^2(1-2\zeta)}$
GP($\zeta = 0$)	–	$\frac{1}{\varphi}$

Despite its easy implementation the MoM has some important limitations. It is frequently biased, and in some cases, such as the GP distribution, its applicability is limited. A popular alternative variation of the MoM it the Probability Weighted Moments method (PWM) [Greenwood et al., 1979]. This technique uses an alternative definition of the statistical moments in order to estimate $\mathbf{\Lambda}$.

LS

The LS estimation involves solving an optimization problem that minimizes the sum of the square of residuals from a sample x_{H_s} of size n, Equation 3.8.

$$\underset{\Lambda}{\text{Minimize}} \sum_{i=1}^{n} (x_{H_s} - x_{H_{s_{F_p}}})^2 \tag{3.8}$$

The parametric model is compared with the data from x_{H_s} using an assumption of plotting positions to estimate the quantile function that defines $x_{H_{s_{F_p}}}$. Examples of plotting positions are the usage of $F_p = (i)/(n+1)$ or $F_p = (i-0.4)/(n+0.2)$. This methodology is of interest due to its simple application, however it demands the usage of an algorithm to minimize the residuals.

Other methodologies can be identified to estimate $\mathbf{\Lambda}$, such as the Element Percentile Method [Castillo and Hadi, 1997]. Teena et al. [2012] compare different estimation techniques when using the POT and conclude that the estimation techniques have comparable performance. Nonetheless, the implementation of any statistical model to approximate a data-set should comprise an identification of the suitable estimation techniques.

Furthermore, an additional consideration when analysing physical variables, such as H_s, is that for some models, $\mathbf{\Lambda}$ estimations are of major importance to ensure that the probabilistic model is an adequate representation of the variable. This is the case of unbounded distributions in the right tail, which may not be adequate to model naturally bounded events such as H_s. The parameter estimation in the GP distribution defines its support. Previous examples can be found where H_s modelling was constrained to

be upper bounded, highlighting the importance of representing H_s as an upper bounded process [Ortego et al., 2012]. It is straightforward to understand that the waves cannot grow infinitely due to their physical character.

3.3.3 Independence

The POT method implies that over u the number of *exceedances* follows a Poisson process. This implies that in a reference period of time Δt the number of *exceedances* can be characterized by a Poisson distribution with mean λ. Occurrences over a certain u shall follow a purely random process, or in other words, each occurrence should be independent. This consideration is of major importance to predict long-term occurrences with the POT.

If *exceedances* are not independent, the statistical model used to characterize it may underpredict or overpredict long-term occurrences. If two occurrences are not independent between each other, they belong to the same cluster of occurrences (e.g. storm) and, as a result, if both are considered in the sample extracted with the POT the statistical model accuracy will be influenced by the fact that the same cluster of occurrences is taken into account twice. Furthermore, as the model is in the assumption of a Poisson process, the mean of the process will be overestimated.

For some physical variables *exceedances* over a certain threshold u are very likely to occur in clusters. Dependence between occurrences is highly frequent in natural phenomena. In the case of H_s it is understandable that sea states are expected to occur in clusters, if a high H_s is recorded, it is more likely for the next H_s on record to have also large values. When analysing POT data it is then mandatory to guarantee that each sampling point, x_{H_s},(when the threshold u level is surpassed) is independent from the remaining x_{H_s}.

Different methodologies are available to ensure independence within the POT dataset. One of the most frequent techniques used is the consideration of a minimal time lag between two events. The most common techniques consist in defining a temporal parameter t_{dc}, declustering time. This is also the simplest technique to decluster data. For H_s a t_{dc} of two to four days is recommended as sufficient to guarantee independent data in Mathiesen et al. [1994]. When modelling H_s using the POT, Van Gelder and Mai [2008] applied a t_{dc} filter of 48 hours to extract independent wave data. In a review of the framework for dealing with environmental extremes, Bernardara et al. [2014] states that a time lag of 24 hours is sufficient to guarantee independence between storm events for North sea data. It is also referred that a time lag larger than 24 hours may lead to undesirable loss of information which should be avoided when modelling extreme occurrences. Additional examples of application to extreme H_s are presented in the same work. Van Gelder [2000] uses a value of 48 hours to model *exceedances* in several North sea locations.

It is important to highlight that in the case of H_s some authors define the t_{dc} as the minimum time between two clusters of x_{H_s}, while others define it as the minimum time between two POT maxima.

An alternative to establish t_{dc}, is the calculation of the auto-correlation function (ACF) associated with a Δt within the time series data. The auto-correlation function identifies the correlation between two points separated by a time lag characterized by a Δt. The ACF is the ratio of covariance (Cov) of a point x_{H_s} with a point separated by Δt with

the unconditional variance of the process (which is the covariance at $\Delta t = 0$).

$$\begin{aligned} ACF(x_{H_s}; \Delta t, t) &= \frac{Cov(x_{H_s}(t + \Delta t), x_{H_s}(t))}{Cov(x_{H_s}(t), x_{H_s}(t))} \\ &= \frac{\sum_{t=1}^{n-\Delta t}(x_{H_s}(t) - \hat{\mu}_{X_{H_s}})(x_{H_s}(t + \Delta t) - \hat{\mu}_{X_{H_s}})}{\sum_{t=1}^{n}(x_{H_s}(t) - \hat{\mu}_{X_{H_s}})^2} \end{aligned} \tag{3.9}$$

Mathiesen et al. [1994] recommended when using the ACF to decluster H_s data. In such case that t_{dc} should not be smaller than the value of Δt for which the ACF has a values of 0.3-0.5.

3.3.4 Defintion of the threshold

Definition of u is one of the most significant and challenging tasks when modelling long-term statistical variables using POT. Embrechts et al. [1999] describe the process of setting the threshold for a POT analysis as challenging and complex.

Different methodologies and rules of thumb can be identified to define u. Even within a single analysis, multiple approaches can be used to estimate u. Defining u can be constrained to physical considerations, such as the physical variable level that may be considered as a threat, or simply to straight mathematical considerations [Lang et al., 1999].

Scarrott and MacDonald [2012] review the process of choice of threshold, with reference to multiple threshold approaches and automated threshold methodologies. The authors present a comprehensive review of the choice of threshold highlighting the main advantages and drawbacks of each.

Bommier [2014] also presents a comprehensive review of the procedure to define the u to model *exceedances* when using a GP distribution, and with specific application to environmental data. Sánchez-Arcilla et al. [2008] propose a thoughtful selection of the threshold. By using a Bayesian approach, the authors show that uncertainty can be reduced in the process and therefore support to the management of coastal infrastructures enhanced.

Some of the techniques used to define u are discussed in Table 3.4.

Table 3.4: Techniques to estimate u.

Threshold estimation	**Methodology**
Rules of thumb	Rules of thumb are very popular in the estimation of u. They are of simple approach and have high reproducibility. Examples of rule of thumbs is the application of a u value corresponding to order statistics equivalent to high quantiles, such as 90% or 95%. Despite being considered theoretically inadequate [Scarrott and MacDonald, 2012], rules of thumb appear in guidance documents, such as IEC [2005], where the 95% quantile is recommended to set u in the extrapolation of loads. Other examples can be found that use rules of thumb, such as the usage of a square root rule $u = \sqrt{n}$.
Parameter stability plot	The parameter stability plot is a graphical analysis to estimate u. It uses the truncation properties of the GP distribution in order to defined the level u. If H_s is a GP distribution with parameters $\Lambda = [\varphi, \zeta]$ then $H_s - u$, given $H_s > u$ is a GP distribution with parameters $\Lambda = [\varphi - \zeta u, \zeta]$. The idea is then to search in the plots the point u_0 for which the plots given by the two pairs, $$[u, \varphi - \zeta u] \quad [u, \zeta], \ \forall u > u_0 \cap u < m_{H_s}$$ are graphically stable, or in other words, have constant values for progressively increasing values of u.

Continued on next page

Continued from previous page

Mean residual life (MRL) plot	The MRL plot also builds on the idea that if the GP distribution is valid for the *exceedances* over u_0 then it is also valid $\forall u > u_0 \cap u < m_{H_s}$. Therefore, as GP mean for u_0 is given by $\varphi/(1-\zeta)$, then for $\forall u > u_0 \cap u < m_{H_s}$, $$\mathrm{E}[H_s - u | H_s > u] = \frac{\varphi + \zeta u}{1-\zeta}$$ Therefore, if the GP conditions are true, $\forall u > u_0$ the expected value ($\mathbb{E}[\ ]$) is a linear function of u. The value of u_0 defines this linear trend and should be identified in a plot that compares the mean of the *exceedance* with different values of u. A main advantage of this methodology is that the mean of *exceedances* can be calculated directly from x_{H_s}. $$(H_s - u | H_s > u) = \frac{1}{n(u)} \sum_{i=1}^{n(u)} (H_{si} - u), \quad u < m_{H_s}$$ $n(u)$ is the sample size depending on u. The graphical interpretation of the MRL plot can be challenging. The plot also becomes unstable for very high values of u due to the progressively decrease in $n(u)$.
Hill plots	Hill plots are used under the estimation that the *exceedance* data follows a Pareto distribution ($\zeta > 0$). It uses a graphical method to identify the minimum value of u for which the Hill estimator is stable. Despite its different formulation, this estimator has essentially the form of a MRL estimator with consideration of log-transformed data Scarrott and MacDonald [2012].
Dispersion Index plot	Introduced in Ribatet [2006] it also uses a graphical method to estimate u. It is based on the extreme value assumption that data POT data comes from a Poisson process. By using the knowledge that for this distribution $\mu = \sigma^2$, it calculates then the Dispersion index by calculating the ratio σ^2/μ, σ^2 being the intensity of the Poisson process. When this ratio is significantly close to 1, the corresponding level of threshold for which λ was estimated is not rejected and can be applied for the POT.

Continued on next page

Continued from previous page

Mixture models	Most of the techniques discussed focus only on the tail data. The mixture models comprise the techniques of selection of threshold that weight the usage of the tail and "main body" data. These were introduced motivated by the fact that information about a sample is comprised, not only in the tail, but also on the "bulk" of the sample. The most common types of mixture models are the parametric models. These consist in a spliced distribution of two parametric components, one for the "bulk" and one for the tail model. Semi-parametric and non-parametric mixture models can be also found in the literature. An extensive discussion on these is given in Scarrott and MacDonald [2012], Bommier [2014]. One of the main drawback pointed to these models is their limited history of implementation, which contributes to a lack of a deeper understanding of them.

Table 3.4 shows that many u fixation techniques, mainly the most popular, depend on a graphical interpretation. In fact, Scarrott and MacDonald [2012] highlight the strong subjectivity that still persists in the process of estimating u.

It is noted that the discussion presented here is focused on stationary extremes. These are usually the ones used to design engineering systems. Non-stationary extremes may involve calculation on extreme values that are expected to change with time, such as the extremes related to climate change.

3.3.5 Validation of the model

The importance of validating the probability model is many times disregarded when using probability techniques such as the one presented. Frequently, multiple probability models and techniques are used to guarantee robust results.

Nevertheless, estimation of the DOFs of a probability model and their validation should also consider a verification process in order to ensure robustness of the results. An efficient approach to guarantee robustness in the statistical fitting involves splitting the dataset into more than one subset (usually two), and using one for fitting the parameters, and the other for validation of the fitting. A leave-one-out analysis can be also performed in order to infer on the robustness of the models calculated in relation to specific points.

If availability of the data is a limitation probably a more simple cross-validation methodology is of interest, where a small number of points are used to validate the statistical model predictions. An alternative for validation that uses less data is particularly relevant for high values of u. In this case, availability of data may be an issue.

3.3.6 Return Level estimation

It was mentioned previously that the ultimate goal of using the techniques here discussed is to estimate a return level (H_{s_r}) within a specified T_r. T_r is usually described as the

period of time for which a rare *exceedance* event is expected to occur once, usually r years are considered. For OWT, the T_r for H_s is of 50 years [IEC, 2009]. For other applications other T_r can be specified.

To estimate the return level the assumption of independence is of major importance, as it will define the number of occurrences that can be expected in r years.

In the case of H_s what is sought is the level of H_s that is exceeded in every r years. This can be also rewritten as the level of H_s that is exceeded every λ *exceedances*. However, this is not so straightforward to comprehend given that λ may be dependent on the period of time for which it is specified. If λ is defined for 1 year, then $r\lambda$ *exceedances* are expected in r years. A unbiased estimator for λ is given by

$$\hat{\lambda} = \frac{M}{r_{ref}} \tag{3.10}$$

where r_{ref} is the number of times the data set available comprises a measure of reference time that relates to period r (*e.g.* if r is in years then r_{ref} is the number of years of records.). Or, if considering years, the number of years of recorded data. Naturally, the higher this number the more accurate the estimate of $\hat{\lambda}$. M is the number of *exceedances* recorded in r_{rec}.

The probability associated with the T_r, $\mathrm{P}[T_r]$, is then calculated as

$$\mathrm{P}[T_r] = \frac{1}{\hat{\lambda}\, r} \tag{3.11}$$

$\hat{\lambda}$ is the estimated mean of the Poisson process, which, without further knowledge, is taken for extrapolations as the true vale λ. This straightforward definition of probability within events can be used in the statistical distribution models applied in the POT in order to estimate the return level. For H_s, the H_{s_r} that is exceeded one in every r years is obtained solving for the GP distribution,

$$H_{s_r} = \begin{cases} u + \frac{\varphi}{\zeta}\left[(\lambda r)^{\zeta} - 1\right], & \text{if } \zeta \neq 0 \\ \\ u + \varphi \ln(\lambda r), & \text{if } \zeta = 0 \end{cases} \tag{3.12}$$

and for the Weibull distribution of *exceedances*,

$$H_{s_r} = u + \varphi \left[-\ln\left(\frac{1}{\lambda\, r}\right)\right]^{\frac{1}{\zeta}} \tag{3.13}$$

3.4 Conclusions

The need to extrapolate physical quantities is an engineering and statistical problem that has been discussed for decades. Extrapolated physical quantities allow the designer to establish reference statistical values for threatening occurrences in order to ensure reliable and robust engineering designs. The demand for extrapolation has its origin in the frequent limitations regarding availability of data. An engineering system is commonly designed to operate for at least 20 years. It is understandable that it is unlikely for 20 years of data to be available in a particular location so that an extreme occurrence with probability of one in every r years can be characterized with minimum uncertainty, even considering that the trend of monitoring and recording big amounts of data is relatively recent.

For H_s data, there are different methodologies that approach the calculation of extreme wave states. This chapter reviewed these different techniques, with special focus on the Peak-Over-Threshold methodology. Despite the initial lack of interest for the field when extreme waves started to be researched, the peak-over-threshold technique has gained high relevance since the 1980s, and in particular, after IAHR promotion of a working group on the topic of extreme sea states [Mathiesen et al., 1994]. Currently the peak-over-threshold is the preferred technique for the calculation of extreme sea states. However, it is not without flaws.

The current chapter evaluated the progressive steps undertaken when developing a peak-over-threshold analysis for H_s data. Within the discussion presented, the following main conclusions were drawn:

- There is still disagreement in regard of the application of the Generalised Pareto distribution to model *exceedance* H_s data. While it was shown by Pickands III [1975] that the Generalised Pareto distribution has strong affinity with the peak-over-threshold method (it is the limit distribution function for *exceedances* when a high enough threshold is considered), in the case of H_s there is still no consensus in regard of which distribution should be applied to H_s *exceedance* data. While some authors, such as [Hawkes et al., 2008, Van Gelder and Mai, 2008], refer that the Generalised Pareto provides the ideal fit for peak-over-threshold H_s data; other mention alternative statistical models [Ferreira and Guedes Soares, 1998]. In particular, it is interesting to note the position of DNV [2014b], which discourages the usage of the Generalised Pareto distribution to model H_s data. These highly relevant recommendations encourage the usage of the Weibull and exponential distributions. It is of interest to evaluate this recommendation in a comparative study. Furthermore, when the exponential is a particular case of the Generalised Pareto distribution in its more constrained form (only 1 DOF).

- The procedure to select the threshold level is a topic of permanent discussion. According to Embrechts et al. [1999], the selection of the "optimal" threshold value is, in the author words, "difficult (if not impossible)". Most of the techniques identified to estimate the threshold value to use in a peak-over-threshold analysis use graphical methods (MRL or the Parameter Stability plot), or, empirical rules. Rules of thumb and empirical rules are theoretically inappropriate [Scarrott and MacDonald, 2012]. Nevertheless, these appear in major standard documents such as [IEC, 2005]. Graphical methodologies are prone to strong bias from the implementation. Recent developments in the estimation of the threshold value try to mitigate the limitation coming from these techniques, e.g. usage of some mixture models. However, most of these still have limited applicability and track record, and the main techniques to estimate the threshold still rely significantly on graphical procedures or rules of thumb. For this reason, a demand for further research on clear, quantitative and unbiased threshold estimation techniques exists in the field of extrapolating physical quantities with the peak-over-threshold method.

- There is limited discussion on the importance of the probability model support when modelling extreme waves. As physically limited occurrences, the expectation is for the waves to not increase in height infinitely. In particular, when wave theories define the criteria for maximum height of a wave. Statistically, when using probability models to extrapolate H_s this fact should be a concern. The main probability models used with the peak-over-threshold have support in $[0, \infty)$, except for the

Generalised Pareto with $\zeta < 0$. This characteristic is of interest in the design and in particular case of deciding the probability model to apply. A more comprehensive discussion on the effects of the statistical support when modelling extreme H_s is therefore necessary.

- There is not a systematic approach to validate peak-over-threshold implementation. Neither to validate the probability models used. In the industry applications an *ad-hoc* approach is adopted to mitigate this lack of a systematic approach. Usually, multiple models, estimation techniques and threshold are used and the designer decides then an a "reliable" estimate of the $x_{H_{s_r}}$. Definition of a systematic approach into the peak-over-threshold analysis has the potential to mitigate bias in the implementation.

Chapter 4

Fatigue design of OWT

Contents

4.1 Introduction

It was seen that the division of the OWT analysis and design in knowledge areas provide a structured understanding of the problem in-hand enabling an easier comprehension of the dimension and limitations of the problem. The previous chapter discussed the definition of the H_s loading environmental conditions and the definition of its long-term occurrences.

The current chapter introduces the discussion of the problem of fatigue design for OWT at the component level, Figure 4.1.

There has been more interest on the evaluation of high return period loads for wind turbines than on the estimating fatigue loads and life. To a significant extent, both problems are related. Nonetheless, fatigue loading is still a major problem for OWTs which has its own specific character.

Fatigue is an important engineering problem. Gurney [1979] estimated that 90% of failures of engineering components could be attributed to fatigue. In the particular case of wind turbines, Sutherland [2000a] states that wind turbines are fatigue critical systems.

Arabian-Hoseynabadi et al. [2010] evaluated some of the failure modes and root causes for wind turbines, and fracture was identified to be the second most common failure mode following material failure, which is also to some extent related to fatigue.

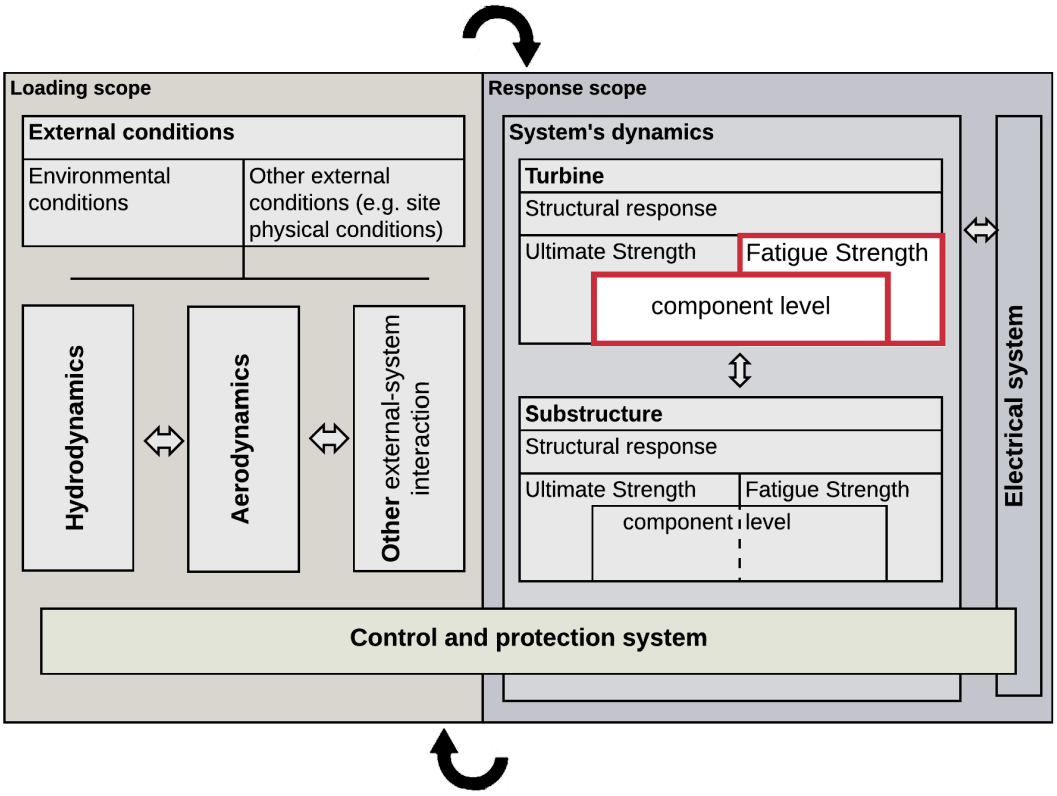


Figure 4.1: Design interaction of the loading and response for OWT operation regimes and positioning of analysis presented in the current Chapter.

Recently, Jiang et al. [2017] reviewed the structural reliability of wind turbines and identified that fatigue was a frequent cause of failure in almost all the components analysed (except for mooring lines and gravity-based foundation).

Moriarty et al. [2004] studied wind turbine loading both for ultimate strength and fatigue and inferred that low SN fatigue exponent materials' life is expected to be dominated by fatigue considerations.

It is known that fatigue is a problem of repetition where the contribution from many small and large operational loads cumulates in time to cause failure of the system. Due to its cumulative behaviour, fatigue characterization is challenging. It depends on multiple loading and material considerations that are demanding to represent accurately. As a result, fatigue design is affected by many uncertainties [Veldkamp, 2006].

The identification of fatigue as a major problem for dynamic systems, such as OWTs, motivates an increasing interest on the improvement of the fatigue design practices. It is of interest to quantify fatigue uncertainty in the design phase in order to produce robust fatigue designs.

The following aspects of fatigue applied to OWT are presented and discussed in the current chapter:

- In Section 4.2 an overview on the fatigue assessment for OWTs and main approaches used is given. In the same section, are presented and discussed the main relevant works on the fatigue assessment of OWT.

- Section 4.3 presents the design procedure for OWT fatigue design, with main em-

phasis on the loading. The loading is a substantial problem of OWT fatigue design. The different approaches found in the literature to define the loading spectra are discussed, and fundamental knowledge on the sector, such as counting algorithms, mean load correction or design equivalent loads, is presented.

- Section 4.4 discusses further sources of uncertainty in the fatigue design of OWTs.
- Section 4.5 discusses the fatigue design from the design-effort perspective. Designing effort is frequently highlighted as a main constrain in the fatigue design for OWT, hindering optimization processes and motivating the application of assumptions to maintain the design procedure practical, such as *e.g.* application of lumped environmental data.
- Finally the main conclusions of the chapter are drawn in Section 4.6

4.2 Structural fatigue analysis of OWT

Fatigue design of OWT towers is focused on the loads generated during the operation of the turbine. It was seen that the current design practices for design of OWT are regulated by standards such as [IEC, 2005, 2009, DNV, 2014a]. The design rules for operational fatigue are no exception.

IEC [2009] states that the design of OWT components shall be determined in conformity to a recognized offshore standard. It demands for it be reliable. Generic applicability methodologies are presented to assess fatigue design along with recommendations to other design standards, such as the analysis of the IEC61400-1 for the extrapolation of fatigue loads.

DIN [2007] recommends spectral methods for linear structures with significant wave loading, but also distribution based methods as an approximation of the long-term stress ranges. Frequency domain analysis is widely applied for offshore structures hydrodynamics. It is advantageous in terms of designing effort. Notwithstanding, an OWT is more accurately described as a system where non-linear loading is important.

DNV [2014a] and DNV [2011] recommend a systematized design procedure for steel structures that uses a statistical distribution, the Weibull distribution, to define the long-term loading spectra. DNV [2011] presents several recommended practices for a comprehensive fatigue design.

Several methodologies can be identified to assess structural fatigue. In fact, one of the constant problems raised in regard of the OWT fatigue assessment is the lack of consistency that exists in the design. Tarp-Johansen [2003] even recommends for the fatigue design of a component to be performed by multiple engineers as a measure to assess its uncertainty.

IEC [2009] distinguishes 7 DLC to address fatigue design, DLC 1.2, 2.4, 3.1, 4.1, 6.4, 7.2 and 8.3. Each of these concerns fatigue generated under different circumstances. Respectively:

- **DLC 1.2** addresses the fatigue that is expected to occur during power production. The fatigue that occurs during power production and that is associated to highly frequent operational conditions, such as common wind and sea states.
- **DLC 2.4** addresses the fatigue caused during power production but when there are fault conditions. This DLC addresses the fatigue generated during scenarios where fault conditions do not cause the immediate turbine shut-down. In such circumstances significant fatigue damage may occur in the turbine.

- **DLC 3.1** addresses the fatigue associated with the start-up process. All relevant fatigue damage resulting from the transient process where the turbine goes from idling or standstill conditions to power production should be considered in this DLC.
- **DLC 4.1** addresses the fatigue originated during the shut-down process. Similarly to DLC 3.1, the DLC 4.1 addresses the damage originated from the transient process where the turbine goes from power production to a standstill or idling position.
- **DLC 6.4** addresses the fatigue caused by common parking condition (e.g. turbine standstill or idling). For this an expected number of hours of non-production needs to be considered and all the critical components in this situation addressed (e.g. blade weight loading, low aerodynamic damping resonance).
- **DLC 7.2** addresses the fatigue caused by the turbine parked with fault conditions. In this DLC significant scenarios of non-operation with fault conditions should be addressed. Fault conditions can occur in the networks or in the turbine itself. This DLC assumes common environmental conditions.
- **DLC 8.3** addresses the fatigue that is associated to the transport, assembly and maintenance operations. Some of the conditions for these may be manufacturer specified, other need to be addressed by the designer. Particular consideration should be given to the assembly phase. In some circumstance the partial completed structure (e.g. withou RNA) may stand for significant periods of time *in-situ*.

In all the cases it is recommended for important environmental states to be considered. For the fatigue DLCs important environmental states are the environmental conditions of common occurrence, normal wind and sea states. Frequent occurrences are the operational conditions that are expected to contribute to decrease most of the system's fatigue life. The analysis of these already comprises a significant amount of designing effort, further limiting any potential analysis of less common occurrences.

In terms of design it is more frequent to find fatigue design recommendations that use SN curves. Fracture mechanics is also commonly considered as an alternative to design or to complement the design to fatigue. The main shortcoming of the fracture mechanics is related to the definition of the initial crack formation and size, which still hinders its application in the design phase [Chryssanthopoulos and Righiniotis, 2006]. Fracture mechanics is frequently seen as a *post*-design engineering technique.

Other techniques can be found for fatigue design. The *strain-based*, or the *energy-based*, methods are examples of alternative approaches. An overview of some of the established techniques to assess structural fatigue is given in Cui [2002]. For OWT the stress-based and the fracture mechanics are seen as the state of the art design methodologies.

SN or stress-based fatigue

The SN fatigue method is the most widely used methodology to assess fatigue in the design phase. It is also the most established. It is related with the usage of an Stress-Cycle curve (SN), which is pre-specified by the material manufacturer. This curve relates the number of cycles (N_f) to failure to each stress range (S). A SN curve reads the number of cycles that a specified material can sustain at a certain S level with a specified level of survivability, e.g. Figure 4.2 for load S of 150 MPa (representative figure). The finite-life

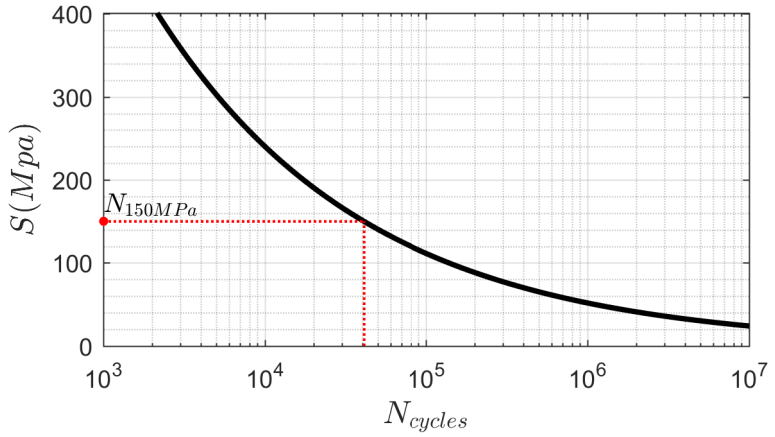


Figure 4.2: Example of SN curve. The limit number of cycles to failure N_f is read on the N_{cycles} axis.

region of a SN curve is usually represented as

$$\log N_f = \log Q - m \log S \tag{4.1}$$

where $-1/m$ is the curve slope on a log-log scale and m is the fatigue exponent. $\log Q$ is the intersection in the $\log N_f$ axis. Where m and $\log Q$ are empirical specifications of the material being considered. These are to be fitted to the experimental testing data. When S has the value of the material's ultimate strength (S_u), Q should have the value of 1 according to Equation 4.1. In many cases the best fit to the SN data does not occur for $Q = 1$, and a multi-segmented curve is applied [Sutherland, 1999]. The result is the definition of SN curves that discriminate low and high cycle regions.

As $S \to 0$ the N_f grows substantially. In practice, a horizontal or vertical line, depending on the representation, appears in this region, and is described as the fatigue endurance limit. It consists on a loading area where the material virtually does not fail to fatigue. Recent findings showed that this fatigue endurance limit in reality does not exist [Sakai, 2009, Stanzl-Tschegg, 2014].

The SN curves are usually specified for target levels of reliability. Accordingly to DNV [2014a], the SN curves should be specified for quantiles of survivability of 97.7%. Depending on the manufacturer, the SN curves can be also defined in terms of range or amplitude depending on the provider. Their analysis demands therefore a careful approach so that no significant errors are committed during the design.

When applying the SN method, a counting of the loads is performed and compared with the SN curve. Naturally, it is very unlikely to have a loading scenario where only a single loading amplitude or range occurs. The reality is frequently more chaotic, and the loading is usually irregular (multiple loading amplitudes or ranges). Therefore, the comparison of the loading with the SN curve requires a criteria that defines when failure occurs. The most well-known criteria to assess the SN curve and a system's structural fatigue is the Palmgren-Miner's rule [Palmgren, 1924, Miner et al., 1945]. The Palmgren-Miner rule, which is presented in more detail in Section 4.3, may seem simplistic to solve such a complex problem as structural fatigue. Other alternative indicators emerged to face the Palmgren-Miner rule limitations, such as the sequence of loading. An extensive discussion on alternative fatigue models and indicators is presented in Fatemi and Yang [1998].

Despite the criticism, the SN method and the Palmgren-Miner's rule are the most used techniques to design engineering systems during the design phase. OWTs are no exception.

Fracture mechanics

Fracture mechanics is an alternative formulation that addresses, instead of an expectation of failure based on S, the actual growth of the fatigue crack. To understand structural fatigue it is important to understand the three stages that are usually distinguished in the fatigue mechanism. These are commonly called, stage I - initiation, stage II - propagation and stage III - fracture. Other authors distinguish more detailed classifications in order to promote further discussion on the subject Cui [2002].

Here, the interest is to understand globally what happens within these 3 stages. Initiation involves the beginning of the crack formation. Cracks usually start at the surface and propagate through the material's crystal planes in stage I. The initiation stage involves crack grow at very small rates up to the point where the crack propagation starts to be perpendicular to the applied stress. This is stage II crack propagation. It involves small and long crack growth until the point where the crack becomes unstable. In this moment stage III begins and fracture occurs.

The most popular model to predict the fatigue crack growth was introduced in Paris and Erdogan [1963] and gives the advancement of a fatigue crack da per cycle (dN), as a function of an amplitude stress intensity factor (ΔK). This relationship, so-called Paris law, is given by,

$$\frac{da}{dN} = C(\Delta K)^c \tag{4.2}$$

C and c are the Paris constants. ΔK has a geometrical dependence on the component. To note that, despite being recurrently mentioned on fracture mechanics, this crack propagation model has limited applicability. It is commonly associated with accurate crack growth prediction at stage II, where crack growth has a linear relationship to ΔK. This limited applicability fomented the appearance of other models to cover the limitations of the Paris law, e.g. Forman model [Forman, 1972], and McEvily Groeger model [McEvily and Groeger, 1977].

Even considering the increasing interest on fracture mechanics for fatigue analysis, crack initiation still undermines a more wide use of this approach in the design phase. As mentioned, fracture mechanics appears frequently in a sequential application after the design with Miner's rule in order to ensure robustness during the operational lifetime of the designed component.

Recently, multiple spatial scale models have been tackling the limitations of fracture mechanics. These try to analyse the crack propagation in multiple scale, using a range of micro to macro spatial models.

4.2.1 Research on fatigue assessment of OWT towers

A significant amount of the work in the fatigue assessment of OWT derivatives from other areas of knowledge. A relevant amount of the knowledge in the fatigue assessment of offshore structures was transferred to the OWT sector from the offshore oil and gas sector. Know-how concerning fatigue assessment of steel materials (usually characterized as low SN m materials) and its uncertainty for offshore applications was widely developed for oil and gas applications.

As the turbine system is fundamentally the same as the one used on inland applications, other significant contributions to the fatigue assessment of OWTs came from previous developments on the inland WT sector.

The problem of fatigue assessment of offshore wind turbines is seen from two different perspectives: the perspective of the low m materials, and the perspective of the high m materials. Fatigue in high m materials approaches a problem of extreme loading. The fatigue damage density being the % of the total fatigue damage that is caused by a stress range S, in high m materials most of the fatigue damage density is condensed in the high S region. For lower m materials, the damage density is spread over different S. Figure 4.3 shows the fatigue density of S data for $m = 3$ and $m = 10$. In the presented example,

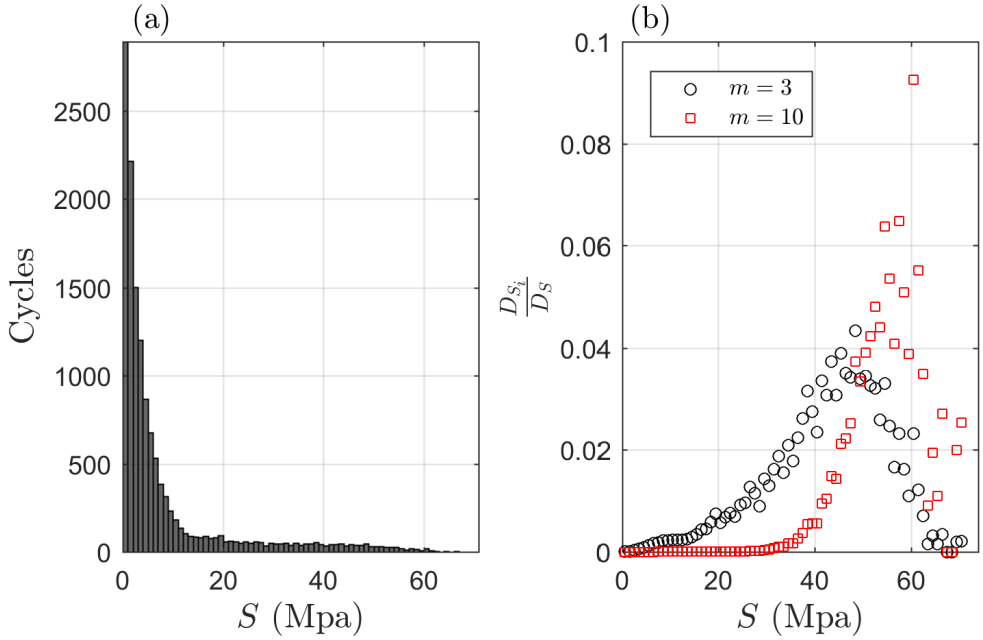


Figure 4.3: Damage density for the same fatigue binned data function of m. (a) loading spectra. (b) damage densities. $\frac{D_{Si}}{D_S}$ represents the fraction of damage D associated to load range S_i.

while for $m = 10$ approximately all of the SN damage density is located at $S > 40MPa$, for $m = 3$, $S < 40MPa$ account for relevant damage density. The problem of fatigue design for both materials is fundamentally the same, however, due this consideration their analysis may diverge in approach.

It was seen that the SN method is the most established approach to assess OWT to fatigue [IEC, 2005, 2009, DIN, 2007, DNV, 2014a]. From the OWT perspective, the SN method design problem is mainly related to the accurate prediction of the lifetime stress/cycles experienced by the OWT. The accurate definition of the SN material properties is also a major topic for fatigue design, however, it is intrinsic to all the engineering problems that demand fatigue assessment.

The complexity of the fatigue design in wind turbines motivated the earlier establishment of a working group to discuss this specific problem [Sutherland and Butterfield, 1993]. The usage of composite materials in the turbine RNA, due to their lower weight, motivated the need to establish efficient techniques to deal with the concentration of the damage density at high quantiles of S. Sutherland [1994] summarizes the main findings,

highlighting that different approaches were commonly used to assess fatigue of WT. In particular three approaches were identified, direct scale-up of cycle counts for the operational design lifetime T, statistical fits of the cycles and stresses, and truncation of data to more accurately characterize the high S region. Despite none of the methodologies being considered non-adequate, a significant concern on quantifying uncertainty was raised. It is highlighted that in one of the case studies presented, the usage of a Rayleigh and a exponential distribution to calculate fatigue (and approaching the tail region) resulted in a difference of the order of 10^3 in the estimated fatigue life for the blade material.

During that time, other working groups were developed in order to promote consistency in the field of WT fatigue assessment Van Grol and Bulder [1993].

These organized efforts represented a milestone on the fatigue assessment of WT and originated a series of works that dealt with fatigue uncertainty characterization.

Lange [1996] researched WT fatigue design and introduced a Generalized Weibull distribution to improve the loading description in the tail region. With 4 parameters, this distribution improves the fit of the tail region. In particular when compared with the most applied distribution to date, the exponential distribution. The author also proposed a method for the reliability design that used parametric uncertainty.

Veers and Winterstein [1998] further discuss uncertainty in wind turbine loading for fatigue analysis by using parametric statistics. The fit of different statistical models is compared using a probability plot. The authors present a simplified reliability methodology that accounts for uncertainty in both, the SN curve and loading spectra.

Manuel et al. [2001] studied parametric modelling of blade loading and its uncertainty using material m values between 3 and 9. The characterization of long-term loading ranges (L) is performed using a truncation of the data at different L values. The author justifies the truncation due to the necessity to eliminate low L cycles, which are highly frequent and account for little damage density, that distort the approximation of the tail region. The authors also highlight the lack of conformity in fatigue calculation for WTs. Winterstein et al. [2001] further extend the work on parametric models using a damage-based model that also uses a Weibull distribution fit. Kelley and Sutherland [1997] complemented the work on parametric modelling by researching on what constitutes a representative sample.

One of the most comprehensive works for the fatigue assessment of WT is presented in Sutherland [1999]. The author discusses the different variables, methodologies, and assumptions frequently undertaken within a fatigue design-basis.

Tarp-Johansen [2003] also extensively discusses uncertainty in the fatigue assessment, and highlights that uncertainties related to the stress characterization have major influence on the fatigue lifetime. Code calibration and optimization for wind turbine blade flapwise fatigue were performed in Ronold et al. [1999], Ronold and Christensen [2001]. The authors quantify uncertainty by comparing empirical data with statistical fits for different operational conditions.

The methodology currently recommended by IEC [2005] to assess the design loads was implemented in Moriarty et al. [2004]. Extrapolation of loads and cycles for the calculation of the design fatigue is performed. Results presented indicate that the design of wind turbine components with low SN fatigue exponent is expected to be dominated by fatigue considerations.

Later, Veldkamp [2006] developed one of the most comprehensive works discussing uncertainty in the fatigue assessment for WTs. Different sources of uncertainty from both, loading and resistance, are thoroughly discussed by the author.

In recent years, research on wind turbine fatigue has addressed detailed specificities. Dalili et al. [2009] discussed the effects of surface phenomena on the wind turbine loading

performance, indicating icing as a relevant threat in terms of fatigue survivability. Dimitrov et al. [2017] studied the effect of the turbulence length scales on the fatigue loading spectra. Murcia et al. [2018] characterized the lifetime fatigue equivalent loads and its uncertainty for the DTU 10MW reference wind turbine, and addressed the usage of surrogate models in order to maintain the uncertainty quantification procedure practicable. Further analysis on OWT loading has addressed other physical phenomena, such as atmospheric stability [Holtslag et al., 2016] or hydrodynamic theories [Marino et al., 2017]. All these works highlight the significant uncertainty that still persists in the fatigue assessment for OWT during the design phase.

Table 4.1 summarizes important works on fatigue assessment that relate to the OWT design and uncertainty assessment. Categorizing the different works is a challenging task. Structural fatigue is transversal to many areas of knowledge. Nevertheless, in Table 4.1, 3 classifications are used: Load characterization, Material and OWT modelling. A significant research effort was identified in the definition of the SN load spectra that specifies the system fatigue life. In particular, as highlighted, efforts were conducted to promote consistency in the field in the 1990s. The material SN resistance is a broad topic. Many material considerations for OWT fatigue are the same as for fatigue of other offshore systems using the same materials. Therefore, it is difficult to narrow the research on material's fatigue properties to only the OWT sector. Even considering this fact, relevant research was conducted explicitly for OWT applications. The design process depends progressively more on accurate modelling. This fact fomented the recent trends in the research of modelling details. Again, it is difficult to separate the modelling from loading and resistance considerations. However, since relevant efforts have been developed to improve the modelling, relevant works on the field are presented.

Table 4.1: Relevant works that discuss the fatigue design of OWT. (components: blade; nacelle (+machinery); tower; substructure)

Research	**Component**	**Reference**
Load characterization		
Load spectra statistical modeling and reliability analysis	blade	[Lange, 1996, Veers and Winterstein, 1998, Winterstein et al., 2001]
Load spectra sample size representation	blade	[Kelley and Sutherland, 1997]
Environmental conditions and loading spectra moments correlation	blade	[Kashef and Winterstein, 1999]
"Best practices" for fatigue analysis components	all	[Sutherland, 1999]
Fatigue reliability using a random factor characterized by load spectra uncertainty at different environmental conditions	blade	[Ronold et al., 1999]
Definition of fatigue equivalent loads for design	blade	[Freebury and Musial, 2000]
Statistical definition of the tail region using truncation	blade	[Manuel et al., 2001, Moriarty et al., 2004]

Continued on next page

Continued from previous page

Fatigue reliability using state of the art uncertainty considerations	nacelle	[Tarp-Johansen, 2003]
Fatigue reliability and optimized code calibration	blade	[Ronold and Christensen, 2001, Márquez-Domínguez and Sørensen, 2012]
Computation of loads in gearbox system	nacelle	[Heege et al., 2007]
Fatigue reliability and optimized code calibration	tower, substructure	[Sørensen et al., 2008, 2011]
Fatigue analysis and reliability of gears	nacelle	[Nejad et al., 2014]
Material		
Fatigue material considerations	all	[Sutherland, 2000b, Sutherland and Mandell, 2005, Marin et al., 2009]
Composite modelling for fatigue failure analysis	blade	[Shokrieh and Rafiee, 2006]
Material surface engineering considerations	all	[Dalili et al., 2009]
Fracture analysis of adhesive joints	blade	[Ji and Han, 2014]
OWT modelling		
Comprehensive review of fatigue design uncertainty	all	[Veldkamp, 2006]
Simulation time convergence	all	[Haid et al., 2013]
Reduction of fatigue design load cases	substructure	[Zwick and Muskulus, 2016]
Atmospheric stability influence on fatigue loads	blade, tower	[Holtslag et al., 2016]
Environmental characterization influence on fatigue life prediction	blade, tower, nacelle	[Toft et al., 2016]
Wave-theory influence on fatigue assessment	substructure	[Marino et al., 2017]
Influence of turbulence spectral parameters on fatigue life.	blade, tower	[Dimitrov et al., 2017]
Surrogate characterization of uncertainty	blade, tower	[Murcia et al., 2018]

Despite all the efforts on the load characterization, a limited amount of works that research on loading for other components than the blades was identified. In particular, the analysis of Figure 4.3 shows that fatigue of high m materials is highly dominated by high values of L/S, but also that for small m materials relatively high load ranges can also have important contributions to the fatigue life of the component, especially, if the representative sample for fatigue assessment is not extensive enough.

Definition of the loading spectra is highly dependent on the loading time series obtained using computational or experimental models. It is therefore also related to the OWT

modeling.

In order to further discuss uncertainty in the fatigue design it is important to understand in detail how fatigue design is performed.

4.3 Procedure to design OWT to fatigue

The current design practices for the tower component involve defining operational loads by running multiple simulations and establishing characteristic stress range distributions. By defining stress range distributions it is possible to assess fatigue of the component by using the Palmgren-Miner damage summation rule presented in Equation (4.3).

$$D_t = \sum_{L_i} = \frac{n_{L_i}}{N_{f_{L_i}}} \tag{4.3}$$

D_t is the damage generated in a specified period of time t, for which n_{L_i} is the recorded number of cycles, or repetitions, of a L_i load/stress and $N_{f_{L_i}}$ the allowed number of cycles at L_i given by a pre-specified SN curve. For evaluation of the design, t should correspond to the design lifetime T. In case of OWT, $T = 20$ years are specified.

The Palmgren-Miner's rule here presented uses a formulation based on L. In the literature it is more common to find it as a function of S. However, for OWT design, curves presented as a function of the loading instead of the stress can be found, in so-called MN curves [Freebury and Musial, 2000].

Moreover, the relation between between L and S for many OWT components is approximately linear (e.g. tower, shafts). Hence, L and S are indiscriminately used in the current document to refer to the fatigue calculations. When this dependence is not linear, or the fatigue curves being used are defined in terms of stress-cycles, the calculations should be performed in the S domain.

Ideally, the designer should replicate the 20 years of operation in order to have an optimum design (bounded in the accuracy to the experimental or computational techniques applied). However, this is usually not feasible and so, extrapolation from a reference time (t) shorter than T is applied to approach OWT fatigue during T. Operation at the different environmental conditions, characterized by x environmental parameters, is expected for OWTs. n_{L_i} has contributions from multiple x states. As a result, it is more convenient to characterize n_{L_i} as a function of x and combine it with the loading spectra conditional on x.

$$D_t(x) = \int_L \frac{f(L|x)n_t(x)}{N_{f_L}} \tag{4.4}$$

$f(L|x)$ is the loading spectra conditional on x, and $n_t(x)$ is the number of cycles experienced in the reference time t at x. The unconditional value of damage D_t can be calculated through integration with respect to x and L,

$$D_t = \int_x \int_L \frac{f(x)f(L|x)n_t(x)}{N_{f_L}}\, dL\, dx \tag{4.5}$$

with $f(x)$, the joint distribution of environmental conditions, which equates the contributions from different environmental conditions. Attention should be given to the definition of $n_t(x)$, in case it comprises half-cycles the presented formulations needs to be applied carefully to not over-estimate high load quantiles where half-cycles commonly appear. To

extrapolate D_t to D_T, t is directly scaled-up to T by scaling the number of times the representative sample is expected to occur within the design time T.

$$D_T = \frac{T}{t} \int_x \int_L \frac{f(x) f(L|x) n_t(x)}{N_{f_L}} \, dL \, dx \tag{4.6}$$

Within the fundamental approach, that applies $f(x)$ and defines $f(L|x)$ and $n(x)$, three main approaches exist to estimate fatigue life, Figure 4.4. The common steps to any of the DLC that involve fatigue calculations are, simulation of different x states, cycle-stress counting, account for local fatigue specificities, *e.g*, mean stress effects, characterization of the loading distribution function of x and sum the fatigue contributions.

While the material characteristics have crucial influence in the fatigue design, an erroneous estimation of the loading and loading cycles may result in conservative or non-conservative designs. It is noted that in the middle step of defining the loading spectra,

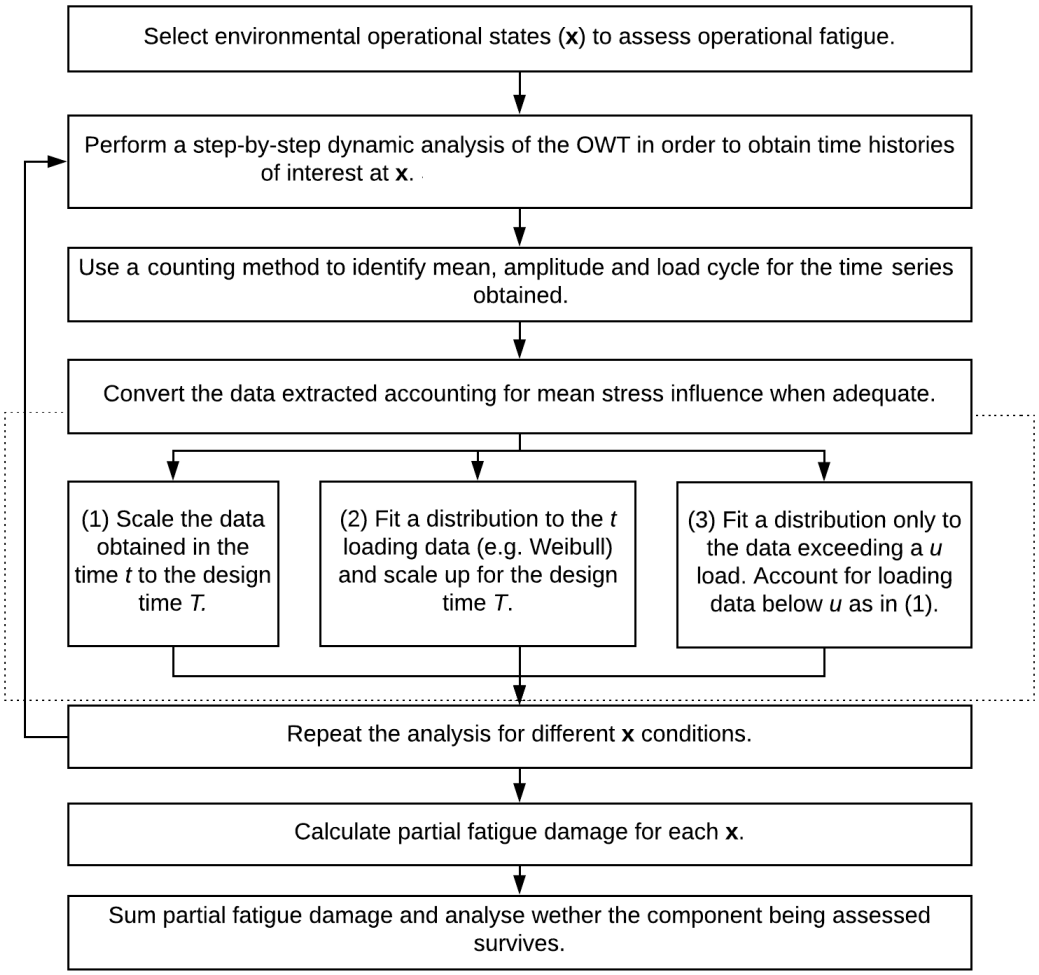


Figure 4.4: Fatigue design methodology to design OWT components to structural fatigue.

different approaches were identified. It is noted that the application of distinct parallel techniques to design OWTs is not uncommon. Teixeira et al. discusses this fact when analysing reliability methods.

The simplest approach, (1), is just the scale-up of the cycles from t to T. In this case D_t is defined as,

$$D_t(x) = \sum_i \frac{n_i}{N_{f_{L_i}}}, \quad i = 1, \ldots, l \tag{4.7}$$

with l being the number of L in the representative sample of t at x. n_i can take the value of 0.5 or 1 in case half-cycles are considered. It is common to use a discrete approach in the calculation of D_T, e.g. [Hayman and Buhl Jr, 2012]. D_t and D_T are given by the sum of $D_t(x)$ at different x.

Method (2) uses a similar approach but using a probability model to approximate $F(L|x)$. Common probability models are the Weibull and generalised Weibull [Lange, 1996].

The final alternative is to use a truncation to approximate the tail region, dividing the problem of calculating D_t in two sub-problems, below u and above u. Structural fatigue linear cumulative damages has two parts, below ($D_{L<L_u}$) and above ($D_{L>L_u}$) u. L_u is the load corresponding to u. Below u the cycles are scaled up accordingly to Equation 4.7. Above u a probability model is fitted and the extrapolation of L takes place. If $L = L_1, \ldots, L_l$ are sorted loads from a time series t, then D_t is calculate using Equations (4.8 - 4.10).

$$D_t = \int_x f(x)(D_{L \leq L_u}(x) + D_{L>L_u}(x))\, dx \tag{4.8}$$

$$D_{L \leq L_u}(x) = \sum_i \frac{n_i(x)}{N_{f_{L_i}}} \quad i = 1, \ldots, l_u \tag{4.9}$$

$$D_{L>L_u}(x) = \int_{L>L_u}^{\infty} \frac{f(L > L_u|x) n_u(x)}{N_{f_L}} d_L \tag{4.10}$$

l_u is the index of L_u in the L sample. When addressing Equation 4.10, extrapolation considerations are demanded for an accurate definition of $f(L > L_u|x)$.

IEC [2005] recommends, if not specified in another way, that at least six 10 minutes simulations (updated in edition 4 to 15) to be taken at each x in order to characterize D_t. This is the case for all the DLCs that involve fatigue calculations.

4.3.1 Counting Loads and Cycles

The SN method requires the assessment of the operational time series of loading, and the extraction of the loads and cycles that occur on it. Different methodologies exist to extract the fatigue loads and cycles from a loading time sequence. The rainflow counting algorithm (RF) introduced in Downing and Socie [1982] is the most widely used counting technique applied to OWT fatigue analysis. The RF methodology is the reference practice for multiple standards, such as [IEC, 2005, DNV, 2014a, CEN, 2005].

The RF involves setting the time series as a sequence of "peaks" and "valleys", *i.e.* maxima and minima, and decomposing the time series by analysing these "peaks" and "valleys" as a sets of "roofs". An example of a RF is presented in Figure 4.5. To apply the RF, the time series is transformed in a series of "peaks" and "valleys". In order to understand how the RF is performed it is necessary to look at Figure 4.5 - (a) and rotate 90° clockwise. If a drop of water falls at Point 1 location, imagining the time series

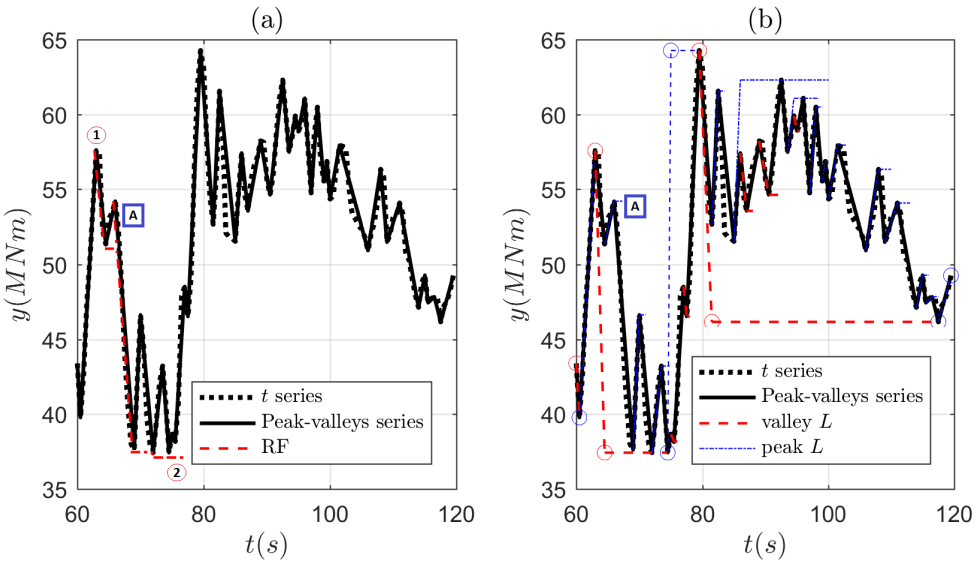


Figure 4.5: Example of the RF algorithm applied to a time series of loading y. (a) Example of counting algorithm. (b) Complete count of the time series. Note: the "peak" and "valley" lines with a circle marker corresponding to half-cycle counts.

as "roof", it will move accordingly with the (red) trimmed line through the time series crossing multiple "roofs". This will account for a L cycle with range equal to the distance between 1 and the maximum position attained by the drop, 2. The areas within the roof not covered in between 1 and 2, e.g. A, are also counted. In case of A, it originates a full cycle with its peak counterpart, accounted in Figure 4.5 - (b).

This process should be repeated until all the "valleys" and "peaks" in the time series are covered. Attention should be given to the fact that when a drop crosses multiple "peaks" or "valleys", it should stop if the "peak" of the following roof is higher than the "peak" of the initial roof, or contrariwise for "valleys". This is of importance to not overestimate the "peak" or "valley" range and counting.

Figure 4.5 - (b) presents the L counting for the presented time series. Half-cycles are identified by the trimmed lines with a circular marker. Half-cycles occur when no counter-part counting exists for a certain L.

The RF procedure is, efficient, and relatively straightforward to understand and apply. However, there are alternative methodologies to count for load and cycles and ranges. Different alternative examples are presented in Lalanne [2010].

4.3.2 Mean load effect

The motivation to implement load cycle counting techniques is related to the irregular character of the loading time series. Figure 4.5 loading sequence is not what the material would expect to experience when being tested for fatigue. Fatigue testing is performed in laboratories and with controlled conditions.

Apart from the random character of the loading within the times series (that results in different L), the loading rarely fluctuates around a zero load. A mean loading component is usually identified in the real loading sequences. The L midpoint in Figure 4.5 is offset in value of approximately 50 units from 0.

Moreover, SN curves found in the literature are frequently characterized for a mean

value of 0. Nonetheless, it is comprehensible that different fatigue damage may be expected for a structure if it is subjected to a L of 10 MNm with a mean value of 0 MNm, rather than if this same structure is subjected to the same S with a mean value of 60MNm. This midpoint offset adds additional structural considerations that should be accounted for in the design phase.

The most common approach to account for mean stress effects in fatigue is the usage of constant life diagrams, Figure 4.6. Constant life diagrams are material dependent representations of fatigue life in different loading scenarios, in particular, load ranges and mean loads. The constant life diagrams are built for different ratios (R) between the minimum and the maximum values of S, and enclose different isolines that characterize the number of cycles for R at specified S and mean stress S_m. They can be understood as a representation of multiple SN curves for different mean loading and amplitude ratio testing conditions.

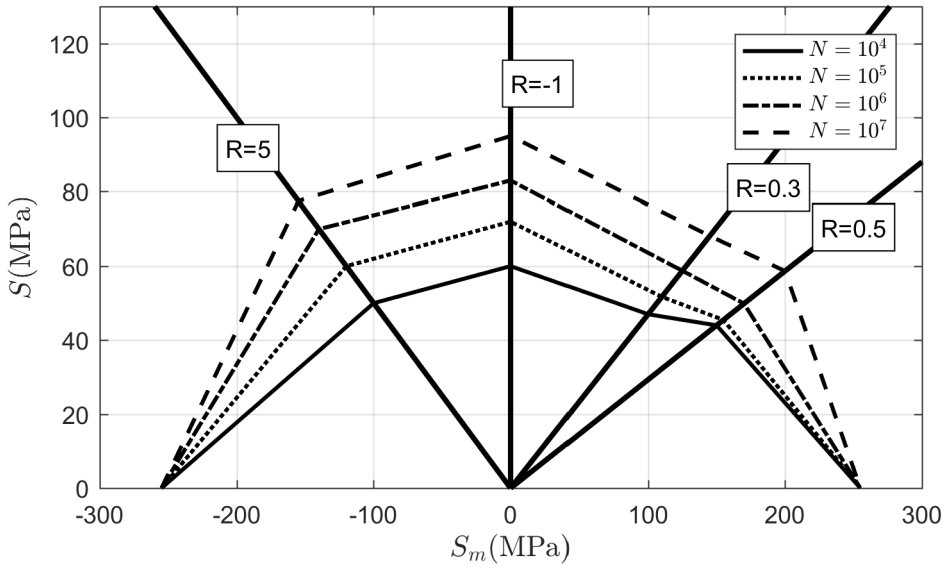


Figure 4.6: Constant life diagram for fatigue calculations at different S_m and R.

The interest of the constant life diagram is that the fatigue number of cycles for a ratio R, a mean load S_m and different S are all available in a single representation.

It is common to collapse this diagram in a more simple relation that characterizes the mean stress effects. A rough estimate of the mean load effect is given by Equation 4.11, proposed by Smith [1942].

$$S = S_0 \left[1 - \frac{S_m}{S_u}\right]^c \tag{4.11}$$

Usually the c exponent has the value of 1, but other values may improve the mean load corrections. This equation is used to estimate the equivalence between a stress S_0 at 0 mean load with a stress S at a mean stress S_m. S_u is the ultimate strength of the material. Other authors propose a modification of this to use the yield strength instead of the S_u, *e.g.*, Soderberg mean load correction.

Moriarty et al. [2004], Hayman and Buhl Jr [2012] use a generalised formula to account for the mean loading. In this case, an equivalent range L_a at an equivalent mean load

L_{m_a} is given by

$$L_a = L\left[\frac{L_u - L_{m_a}}{L_u - L_m}\right]^c \tag{4.12}$$

which simplified for $L_{m_a} = 0$ becomes

$$L_a = L\left[\frac{L_u}{L_u - L_m}\right]^c \tag{4.13}$$

that is equivalent to Equation 4.11, when S and L have linear dependence. L_u is the ultimate strength of the material. L_m and L_{m_a} are respectively the mean load and the equivalent mean load for the transformed data.

It is noted that the mean load correction are not always necessary. Mean load corrections may be disregarded for some fatigue details, such as welds.

Current OWT fatigue design practices also apply Design Equivalent Loads (DEL) to account for the variability of the loading spectra.

4.3.3 Damage Equivalent Loads (DEL)

The idea of damage equivalent loads is also based on the need to deal with different loading conditions. The principle of the DEL is to determine the number of equivalent cycles (n_{eq}) at a certain constant equivalent stress S_{eq} that induces equal SN fatigue damage as the whole load spectrum, or loading of interest.

Damage equivalent loads emerged in the wind engineering sector as a solution for testing of components and also to account for the irregularity of the loads (e.g. correction of L_m). By defining a damage equivalent load that generated the same damage as a load spectra, testing of WT components was simplified [Freebury and Musial, 2000].

Equation 4.1 can be rewritten in a linear representation in a log-log scale and non-dimensional form.

$$N_f = Q\left(\frac{S}{S_u}\right)^{-m} \tag{4.14}$$

The damage caused by l loads is given by Equation 4.7. The damage generated by n_{eq} at S_{eq} is therefore

$$\frac{n_{eq}}{QS_{eq}^{-m}} = \sum_{i=1}^{l} \frac{n_i}{QS_i^{-m}} \tag{4.15}$$

Correction for mean load S_m can be included by merging Equations 4.11 and 4.14 with $c = 1$, in such case correction for mean load is introduced in Equation 4.15.

$$\frac{n_{eq}}{Q}\left(\frac{S_u - S_{m_{eq}}}{S_{eq}}\right)^{-m} = \sum_{i=1}^{l} \frac{n_i}{Q}\left(\frac{S_u - S_{m_i}}{S_i}\right)^{-m} \tag{4.16}$$

Due to the linearity character of the SN curve, DEL appears more commonly represented as

$$S_{eq} = \left(\frac{\sum_i^l S_i{}^m n_i}{n_{eq}}\right)^{\frac{1}{m}} \tag{4.17}$$

This is the DEL calculation for the simplest case where the same S_m, m and S_u are used. Naturally there are two DOF that can be controlled in the formulation presented, n_{eq} and S_{eq}. The same number of cycles $\sum_i^l n_i$ can be used as a reference for n_{eq}.

Alternative definitions may use a measure of frequency to set n_{eq} and then estimate S_{eq}, e.g. [Hayman and Buhl Jr, 2012] where a frequency is defined depending on the length of the load time series.

4.3.4 Frequency Domain Fatigue

The fatigue estimation techniques discussed rely on time-series to define a fatigue loading spectra, a representative sample of L/S. These time series can be obtained from experimental or analytical setups.

OWT fatigue analysis can be also performed in the frequency domain. To perform fatigue analysis in the frequency domain it is necessary to define a frequency spectra that decomposes the time series in multiple frequency components. This spectra should comprise the energy or amplitude of loading for each frequency component.

The frequency spectra multiple frequency components can be merged to create synthetic time-series. This practice is widely used for offshore engineering hydrodynamic analysis. Wave theory shows that irregular waves are well approached by using linear superimposition of regular waves. Therefore, each loading component can be decomposed and related to each wave component, and consequently, frequency component.

The principles for frequency analysis of fatigue in wind turbines are similar. A frequency spectra is defined from an initial time series and is applied to create synthetic loading time series from which load ranges and cycles can be counted to assess fatigue. While for wave climates, multiple spectra exist that describe accurately the significant wave heights (e.g. JONSWAP spectra), for OWT the loading spectra needs to be defined from the loading information about the OWT being analysed.

One of the main methods used to create the spectra and synthetic time series is the application of the Fast Fourier Transform algorithms and their inverse. Sutherland [1996] applied this methodology to calculate a blade loading frequency spectra.

Ragan and Manuel [2007] compared the application of time domain and frequency domain methods to estimate a 1.5MW wind turbine where the frequency spectra is directly used to estimate the PDF of the loads. A good approximation of the tower loading was attained, whereas the same was not verified for the blade loading. There is direct connection between the results attained and the damage densities related to the components (high m are more influenced by low frequency loading). The comparative calculations were performed in DEL. The main advantage of estimating directly the PDF is that there is no need to define the synthetic time series which demand the establishment of a random phase component (the spectra does not include phase for the superimposition of the frequency components). Definition of the spectra from the initial dataset implies that the time component is lost.

Nevertheless, the usage of DEL for a comparative analysis is not a reasonable choice. The error of the DEL approximation results in much higher errors of approximation in terms of Palmgren-Miner's cumulative damage, see Figure 4.7.

Tibaldi et al. [2016] recently evaluated the frequency domain analysis using a linear model of the NREL's 5MW reference turbine. This frequency approach is developed in order to perform an optimization procedure without compromising computational cost. Nevertheless, the authors are the first to highlight the limitations of the frequency domain approach when comparing with the non-linear time domain load calculation.

Usage of the frequency domain techniques to analyse fatigue of OWTs has received limited interest. Frequency domain analysis does not approach well many OWT signals. One of the main limitations of using the frequency domain analysis to assess fatigue is that the initial frequency spectra is bound in accuracy to the time series used to define it. Synthetic times series can give new insight into important loading for L considerations, however, these insight are not necessarily right as they are bound to the initial time sample. The frequency spectra are limited in the amount of frequency components by the

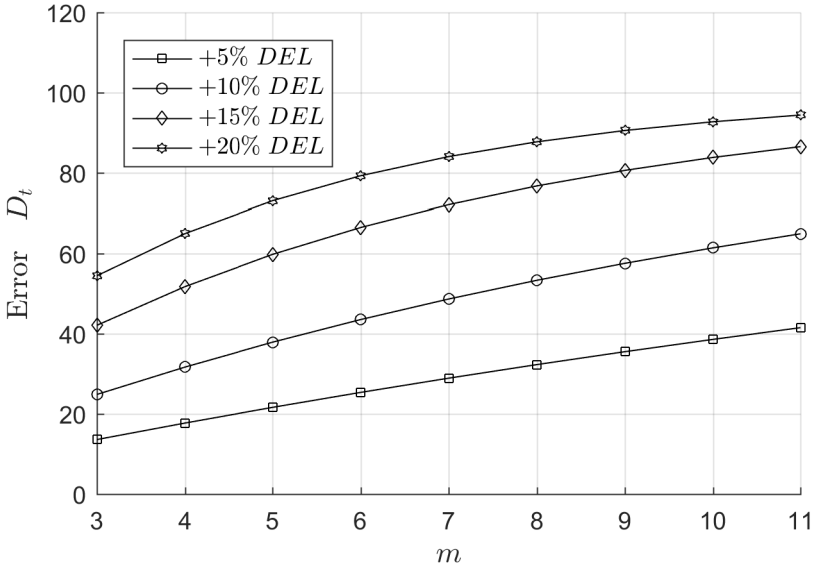


Figure 4.7: Variation in the Miner's cumulated fatigue damage function of different variation in the value of the DEL.

length of the time series. Moreover, some important non-linear effects may be disregarded by the definition of the frequency spectra and the synthetic times series. Some of these are of importance to OWT's fatigue. Frequency domain methods have been recently used to analysis floating OWTs.

4.3.5 Representative Sample

Fatigue is a problem of the mean, where the accuracy of the T fatigue estimation is highly related to the dependence on the mean estimation.

Sutherland [1999] highlighted previously the SN fatigue's dependence on the mean value. The author mentions that, when a loading sample is representative, the accuracy of the fatigue estimation can be considered independent of T.

Therefore, it is not problematic that 20 or more years of estimation are demanded. In ideal conditions these could be fully assessed. However, it is not feasible to estimate the whole lifetime of fatigue loading, it is sufficient to define a representative sample and use it in order to estimate the T fatigue.

Converging the mean of the loading distribution is a challenging task. In particular due to the high variability of the damage that is induced in the reference analysis time t (usually 10 minutes) Sutherland and Veers [1995]. The authors remark, through experimental tests on a 34 meters blade, that fatigue damage calculated in 10 minutes could vary by two orders of magnitude. Despite being applied in a 34 meters blade, which is not representative of the current state-of-the-art OWT dimensions, the large variability identified in this experimental test is still characteristic of today's multi-megawatt turbines, as shown by Tibaldi et al. [2015] when comparing DEL for NREL's 5MW turbine.

Figure 4.8 shows the convergence results for the $t = 10$ minutes cumulated damage on the tower component. The variability of the D_t (with a $t = 10$ minutes) when using an increasing number of seeds in the calculation of the mean value is presented.

A lower number of seeds naturally results in more variability of D_t and its mean. It

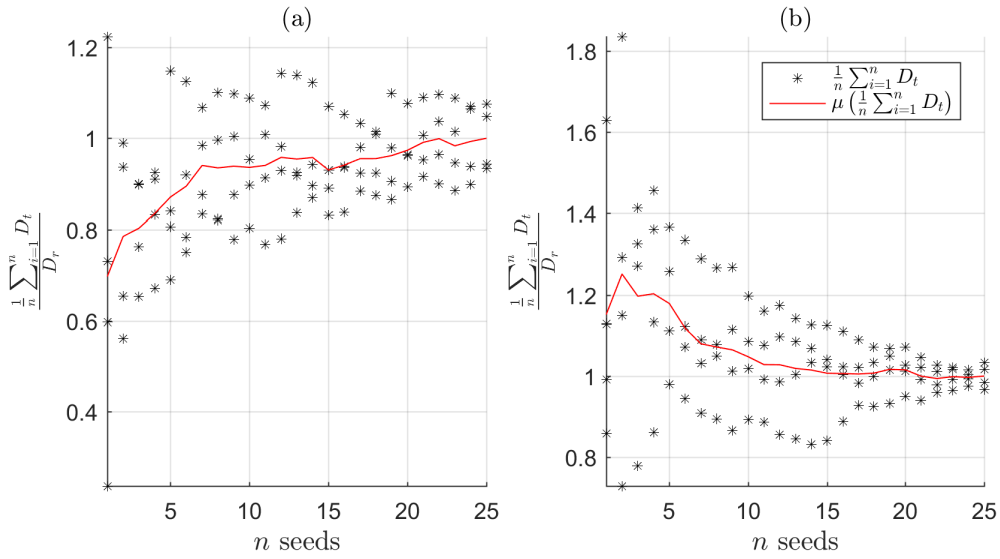


Figure 4.8: Convergence of the sample mean D_t in the tower component (with SN $m = 5$) with the increasing number of simulation seeds. (a) $U = 8m/s$; $H_s = 2.8$; $10.0s$ collinear wind and waves and I modelled NTM. (b) $U = 17m/s$; $H_s = 3.8$; $11.2s$ collinear wind and waves and I modelled NTM. D_r is considered the real D_t, i.e. mean estimation of D_t with high n seeds.

is similar as sampling from a distribution. The bigger the sample the more accurate the statistical representation of the mean. The lower the number of seeds the wider spread of the mean estimation points. Nonetheless, even for a large number of seeds, more than 20, there is still considerable variability in the estimation of D_t. When the number of seeds increases, so too does the computational cost. The T estimation at a certain x is bounded in accuracy to the number of different seeds used to estimate D_t.

Representative samples are not only demanded for extrapolation of high m material, but also to estimate the fatigue of low m components. IEC [2009] sets a lower limit for the representative character of the sample as 6 seeded simulations per x for all the cases that involve fatigue calculations. Sutherland [1999] states that a representative sample may include "several hours of operational simulations using different seeds for high m materials".

4.4 Further considerations on uncertainty in the fatigue design of OWT

An extensive discussion on the methodology to establish the fatigue loading spectra and sources of uncertainty related to it was presented.

Design practices that relate directly to the definition of load-cycles and loading spectra are substantial sources of uncertainty in the design. In particular because many of these still depend on the designer preferences. However, there are multiple additional sources of uncertainty that affect the OWT fatigue design analysis. These were extensively discussed by Veldkamp [2006].

For the OWT designer it is of major importance to comprehend the root cause of all sources of uncertainty that affect the fatigue design. A brief complementary discussion on

other relevant contributors of OWTs fatigue design uncertainty is presented in Table 4.2.

Table 4.2: Relevant sources of uncertainty in the OWT fatigue design

Wind	
Wind environmental parameters	It is common to use generalised statistical fits to approximate wind distributions of mean speed, direction and in some cases also turbulence (e.g. Weibull). Despite being a practical approximation of the environmental statistics, these models are only an approximation of the exact site statistics. Moreover, many distribution models are bounded in accuracy to the length of the initial time series used to define them. When extrapolation measures are demanded to estimate long-term occurrences of any of the parameters related to the wind, the probability models used to approximate these variable have a major role. While data may be available for many sites, it is rare for long-term measurements to be always available. Even if measurement campaigns are developed in the installation site, these are unlikely to last for more than 1 year. Spatial extrapolation is frequently used to tackle this limitation and site long-term wind data is characterized by using established techniques such as the WAsP [Troen and Petersen, 1989].
Wind shear	Collection of offshore environmental wind data is challenging. The costs of assessing environmental wind parameters offshore is significantly higher than inland. Most offshore existing wind records were taken at height of 3 to 5 meters. Recently, LiDAR technology slightly changed the paradigm of wind data collection offshore, however the cost of this laser technology is still substantial. The wind needs to be corrected from the sea level to the OWT hub height. More than one model exists for the effect, being the most applied, the power law and the logarithmic profile. The last relates U at height z_2 and z_1 using a surface roughness parameter z_0, $$\frac{U(z_2)}{U(z_1)} = \frac{\ln(z_2/z_0)}{\ln(z_1/z_0)} \tag{4.18}$$ This representation of the logarithmic profile disregards atmospheric stability that was shown in Holtslag et al. [2016] to be of importance for OWTs. Moreover, estimation of the terrain roughness and the approximation of the wind profile do not always approximate accurately the site conditions.
Turbulence Intensity(I)	The wind speed is not always constant over the recorded time. Thus, a measure of its mean variability is used to characterize the U fluctuation. The I is the complementary of U which characterize the standar deviation of U (σ_U) in the measured time. It is defined as, $$I = \frac{\sqrt{\left[\frac{1}{t}\int_t (U(t)-U)^2 dt\right]}}{U} = \frac{\sigma_U}{U} \tag{4.19}$$ The accurate estimation of I is challenging, both, technically and economically. Nevertheless it has high influence on OWTs fatigue. It is highly site specific and estimation of I usually use approximations, *e.g.*, IEC classes. Türk and Emeis [2010] showed that [IEC, 2009] turbulence classes are conservative for offshore applications.

Continued on next page

Continued from previous page

Wind field description	The definition of the different parameters that characterize the wind is not enough to replicate the wind field. Mainly because the wind field is not well described as a laminar flow, instead it is highly turbulent. In similar to what was discussed for the frequency spectra for waves, an elementary input to characterize the wind field is its power spectral density. It is defined by a spectra formulation that relates the U scales. The Kaimal model turbulence model [Kaimal et al., 1972] is commonly recommended for OWT applications. $$\frac{fs_k(f)}{\sigma^2_{U_k}} = \frac{4fL_k/U_{ref}}{(1+6fL_k/U_{ref})^{5/3})}, \quad with \quad \sigma^2_{U_k} = \int_0^\infty s_k(f)d_f \tag{4.20}$$ where s_k is a spectral component and k is the index of the spatial dimension. f is the frequency and L_k is a length scale. U_{ref} is the mean wind speed at hub-height. Spectral components cannot be sampled without any consideration of correlation between points, otherwise the wind field may be physically impossible even considering that it respected the spectral conditions. To account for the correlation within the space a coherence function complements the spectral model. An exponential correlation is recommended by IEC [2005] to use with the Kaimal spectrum. Other spectral models can be used to characterize the wind field spatial conditions, some alternative examples are presented in Jonkman and Kilcher [2012].
Sea	
Sea environmental parameters	The estimation and characterization of sea conditions, similarly to the wind, depend on *in-situ* recording. When these are not available, spatial extrapolation is used and sea conditions from close site or hindcast data are used. For fatigue considerations of OWT, waves are expected to have major influence in the substructure. Similarly to wind conditions, long-term wave conditions need to be estimated from data, an extensive discussion on the methods for extrapolation of H_s was presented in the previous chapter. H_s is usually taken as the mean of the 1/3 biggest waves in a wave state [Holthuijsen, 2010]. Some uncertainty can be found in the definition of the wave period. IEC [2009] recommends the complementary period to be taken as a range value for normal and severe sea states. Wave characterization usually uses a spectral formulation, a wave spectrum described by different regular waves. A wave spectrum can be defined using wave records or hindcast data. As alternative an existing wave spectrum can be implemented, examples are the Jonswap or Pierson-Moskowitz spectrum. As the dimension of the environmental parameters increases, fatigue analysis can become very time-consuming. A common approach to avoid unmanageable design times is to define lumped load cases [Kühn, 2001], *e.g.*, [Fischer et al., 2010]. In this case, only a few relevant sea state parameters are simulated in pair with the wind conditions without significant loss of accuracy according to the author.

Continued on next page

Continued from previous page

Wave field description	Recurrent descriptions of the physical wave field use the linear wave theory. It considers that regular waves can be superimposed to create a irregular wave field, in which the amplitudes are defined by the wave spectrum components. Wave kinematics are described by orbital motions that decrease in radius exponentially with the depth. Despite its wide application it as some fundamental assumptions that are not verified in practice. The seabed is expected to be horizontal and impermeable. It assumes ideal inviscid fluid and the theory is only valid in deep water where there is no interaction between the wave and the seabed. Moreover, waves are assumed to be two dimensional, which in practice does not occur (waves are usually three-dimensional). Furthermore, important non-linear wave effects may be disregarded by the linear wave theory, such as breaking waves. Several other wave theories can be applied as an alternative, each with its range of applicability in terms of wave steepness and depth wave length ratio. For coastal waves higher order non-linear models are preferred in relation to linear wave theory, *e.g.* Marino et al. [2017] study on NREL's 5MW baseline on monopile foundation.
Hydrodynamic Interaction	There are two types of hydrodynamics parameters, the ones related to the inertial loading and the ones related to with viscous loading. Definition of the inertial parameters depends on the relative size of the characteristic length of the structure in relation to the incident waves. OWTs are usually slender cylinders and their analysis depends on two empirical parameters C_D and C_M, the drag and inertia coefficients respectively, which relate to the loading through the Morison's equation: $$L(z) = \frac{1}{2}\rho_w C_D D(z)\lvert v\rvert v + C_M \dot{v}\frac{\pi}{4}D(z)^2 \tag{4.21}$$ where ρ_w is the water density, $D(z)$ is the cylinder diameter at z height and v and $\dot{v}$ are the velocity and acceleration of the wave particles in the direction perpendicular to the cylinder. Some value of C_D and C_M are chronic, and their estimation for each case is difficult. When the characteristic length of the body is large in comparison with the incident waves other hydrodynamic effects need to be accounted, e.g. diffraction or radiation. In practice these effects, even for relatively small bodies may introduce some inaccuracy in the estimated hydrodynamic forces.
Current and Tide	The sea surface level and the currents appear in a secondary role when concerning OWT fatigue. Very simple assumptions are undertaken for these during the calculations. Current is assumed to be negligible when compared with the wave particle speed. Veldkamp [2006] defines an equivalent mean water level to account for sea surface variations.
Aerodynamics	

Continued on next page

Continued from previous page

Wind-blade interaction	Blade element momentum (BEM) is the most used technique to describe the interaction between the blades and the wind. In the BEM the blade is modelled as discrete elements where the loading in each element depends on the its local aerodynamic components, drag and lift. The local loading components are then integrated over all elements. Tangential and axial flow conditions are accounted using induction factors (axial and angular). Despite its wide usage, the original BEM has several limitation which were overcome through considering empirical corrections. Examples are the corrections for tip-loss which are related to the vortices shedding from the tip of the blade; corrections for turbulent wake-state that are related with the physical instability for large tip speed ratios; or radial flow corrections to account for radial flow in the calculation of aerodynamic loads. The success of the BEM is highly based on its practicability in terms of computational effort, alongside its proven accuracy [Wang et al., 2016]. Recently, interest in the usage of CFD models for wind engineering has increased. CFD uses a system of Navier-Stokes equations solved through the definition of a discrete fluid domain (most commonly solved using finite volumes). The cost and the complexity of OWT analysis with CFD still hinders a wider applications of CFD codes in the wind industry.
Wind-tower interaction	The dynamic interaction of the wind with the tower is not commonly seen as a major source of uncertainty in the modelling. Apart from the tower loading, the tower component also introduces a disturbance on the wind field that interacts with the passing blades. Drag of the tower depends on empirical coefficients based on the classical solution for a potential flow around a cylinder. Some models account for tower loading independently of the inflow disturbance and local aerodynamic interaction. At the tower-wind interaction level, empirical corrections were proposed to account for local uncertainties, many of these specified for inland applications.
Wake interaction	The calculations of loading in OWT frequently consider the individual loading of the wind turbine. In reality, an OWT is expected to be installed in a wind farm. The wake of each turbine is therefore expected to interact aerodynamically with near OWTs. Because the energy available in the wind field right after the turbine is lower, OWT are positioned in order to take into account the farm layout aerodynamic interactions. Nevertheless, due to its high uncertainty, wake interaction is to expected to be under active research in the coming years [Barthelmie et al., 2009].
Turbine System	
Control system	Most of the commercial turbines installed today use at least 3 control systems: yaw, pitch and generator. The existence of the control system increases the complexity of the fatigue analysis. At the same time, control systems are important dampers of it. When designing OWTs, absolute threshold values are used to feed the control systems, e.g. operation U limits, yaw angles, wind speeds. However most of these variables suffer from inherent uncertainty from the measurement which may affect the estimated fatigue. Furthermore, control is highly related to the electronics of the system, and power generation, which also "feedbacks" on the structural response.

Continued on next page

Continued from previous page

Structural response	Many OWT simulation codes consider beam models. These can be applied because OWTs structures have slender shapes. Furthermore, beam models are accurate and computational efficient. Because OWTs blades and towers have low deflections it is frequent to encounter a modal approach to model these components. Definition of the structural behaviour in terms of their 1^{st} two modal components is frequently accepted to produce accurate results. Nevertheless, modal decomposition has inherent uncertainty as only a limited number of modes of vibration is used to describe the structural response. Additionally, with the increasing size of OWTs, structural modelling techniques need to account for potential large deformations. Multi-body-dynamics are implemented as an alternative of modal analysis. For fatigue calculations the increase in computational time from multi-body-dynamics does not justify the gains in accuracy of the system's response. A more comprehensive approach to aero-elastic simulation comprises the consideration of Finite-Element-Methods (FEM) e.g. [Passon et al., 2007]. While there is an increase in one-dimensional FEM models, three-dimensional FEM codes are powerful tools to accounts for design details which are of major relevance for OWT fatigue. FEM models are still relatively expensive for a more wide applicability in the OWT sector. Furthermore, accurate application of FEM involves precise definition of the boundary conditions and mesh elements which may contribute to increase uncertainty in the modelling. Structural modelling uncertainty also accounts for additional assumptions related to the turbine system itself that are challenging to characterize. Are examples, eigenfrequency definition, geometric uncertainty or dynamic coupling.
Foundation	Many OWT certified modelling systems still disregard the foundation dynamics. Aparent-fixity models are the most common type of foundation encountered in OWT modelling, [Passon et al., 2007]. Other alternatives to model foundation systems include discrete-springs, lateral, axial, or rotational. These models are a simplistic approximation to the modelling of foundations and are highly dependent on the values assumed to reproduce the soil-structure behaviour.

Continued on next page

Continued from previous page

Material

Most of the discussion presented concerns the definition of the loading spectra for fatigue calculations. These are intrinsically related to the materials properties through the SN curve. Moreover, looking at the DELs it is possible to infer that these are independent of the geometry but are material dependent. Materials play a significant role in fatigue design.
Some of the concepts related to material properties were discussed when discussing loading, such as DEL and mean loading effects, or SN curves.
Research on fatigue in general has dedicated a considerable amount of time in solving material-related problems. Even if all the variables of loading discussed up to this point were certain for the designer, uncertainty would still occur due to material considerations.
Even within a laboratory controlled context, SN curves present scatter, mainly due material and geometric imperfections. Material uncertainty considerations for an OWT in real operational scenarios add on top of the uncertainty resulting from testing. Though that related to the structural response, stress concentration effects are an example of uncertainty originated on the system description and operation. Assessment of fatigue material uncertainty its a whole field of research on its own.
In case of SN curves, it was seen that these are usually defined for a specified level of survivability. Typical reference values for the slope of the SN curve are (*Sources:* [Mandell and Samborsky, 1997, Freebury and Musial, 2000, CEN, 2005, 2011]):

Material	**SN Slope** (m)
Steel	3-5
Aluminium	3-7
Composite (glass fiber)	6 - 12

The value of m is dependent not only on the material, but also, on the structural detail (*e.g.*, welds) and cycling regime (high or low cycling). OWT, due to their complexity, enclose different materials, selected accordingly to their applicability and cost. Structural critical turbine elements, such as turbine blades are mainly manufactured in composite material. Glass fiber reinforced polyester and epoxy resins are widely common. Carbon fiber may be also applied. Towers and foundations are mostly built on steel. New research trends have been studying the applicability of composite towers [van der Zee et al., 2017]. IRENA [2012] presents a comprehensive overview of the manufacturing materials applied in the wind energy sector. Detailed information on some of the materials applied to blades is presented in Burton et al. [2011]. Garrett [2011] when performing life-cycle assessment showed that steel and iron composed approximately 82% of the manufacturing materials for a V90-2MW wind turbine. Whereas, composite materials composed approximately 8.5% of the total manufacturing materials for the same turbine.

The considerable amount of sources of uncertainty that is involved in OWT fatigue assessment motivates conservationism in the design practices. An example is the specified levels of survivability for the SN curves.

The fatigue design safety factors, or fatigue design factors, found in guiding documents, such as [DNV, 2014a] (minimum of 2, 1 can be assumed but with a regular inspection plan), or research works such as [Sørensen et al., 2011] (mininmum of 2.5) are a strong evidence

of the high uncertainty associated with the fatigue design of OWT. In the latter, when assessing calibration of Design Fatigue Factors (DFF), the authors apply representative values for offshore oil and gas platforms which may be a conservative approach.

One of the origins of the conservative approach to the fatigue assessment is related to the high (computational or experimental) costs involved in the procedure, which inhibit more comprehensive analysis. Another is the significant uncertainty that still exists in the fatigue in general. The case of very high cycle fatigue that only recently started to be faced as a "real" fatigue failure problem is a relevant example of the generic uncertainty that still affects the fatigue assessment.

4.5 Cost of fatigue calculations

One of the most recurrent limitations pointed to the OWT fatigue design procedure is the amount of effort required to perform the analysis. For loading calculations the joint distribution of environmental parameters is divided in bins. If only two parameters are used, let's say U and I, and divided in 10 bins then 100 combinations of U and I are possible. If H_s, and Peak Period (T_p) are added using 10 bins, then 10^4 combinations are possible. If the wind direction (θ_w) is added, 10^5 and further on. This increase in number of combinations when dimensions increase is related to the curse of dimensionality. The difficulty of covering all the possible combinations in between variables in space increases exponentially when new dimensions are added.

This motivated the development of many strategies to maintain the fatigue designing effort for wind turbine bearable, such as [IEC, 2009] recommendation for the sea states to be conditional on the wind speed when calculating fatigue loads. The idea of limiting the sea states conditional on the wind parameter for fatigue calculations has its origin in Kühn [2001] which introduced the use of lumped scatter diagrams. Despite reducing computational time, inaccuracies are expected from not considering $f(x)$ to its full extent. Veldkamp [2006] remarks that this approach is uncontroversial. However, it seems a simplistic approach for design when hydrodynamics (due to direct interaction) may be of high importance.

Nevertheless, it is not uncommon to find in the industry disassociation between the foundation design and the turbine loading calculations, performed by different contractor, who then exchange information at the foundation-turbine interface level.

For certification purposes more than 12000 time domain simulations were identified in the industry to certify a OWT to fatigue design, the most of all the DLCs.

To conduct a very accurate analysis of fatigue and extreme loads, Moriarty et al. [2004] performed more 4725 thousand simulations to characterize a two-dimensional U and I grid. Even considering the high number of simulations the authors only consider 9 seeds per U and I bin. They then compare the results obtained with a much smaller sample of 189 simulations. The subset samples overestimate extreme loads and under-predict fatigue life for both, stall and pitch regulated turbines.

An alternative identified to reduce the computational effort is the usage of the frequency domain methods. Despite approaching the design of some components relatively well, it was seen that frequency domain fatigue analysis has very limited interest for other. Cosack [2010] remarks that, and even considering the already poor acceptance of frequency domain fatigue for WT systems, it is unlikely for fatigue frequency analysis to be further developed in the future as complexity of the system is expected to increase.

The need to decrease the designing effort on OWTs has been an active topic of research.

Meta-modelling has shown the highest potential to reduce the simulation costs of high-fidelity computational or experimental analysis. Meta-models for OWTs can be applied as surrogates of OWT expensive evaluation models, such as the aero-hydro-servo-elastic codes. Application of meta-models as surrogate models for OWTs has increased recently [Maki et al., 2012, Yang et al., 2015, Morató et al., 2016, Teixeira et al., 2017b, Murcia et al., 2018].

After their definition, these models are costless to evaluate and produce accurate predictions. Moreover, the demand for solutions that enable the reduction of the design efforts for OWT gains significant relevance due to the identified trends in the increase of computational complexity for design purposes, such as incremental application of CFD and FEM models.

4.6 Conclusion

Fatigue design still constitutes one of the challenges in the design of OWT components. The root causes for the challenging OWT fatigue design is the high number of uncertainties that are involved in the design process. Given that fatigue is a cumulative event, in the case of OWT with the contribution from millions of load cycles, its characterization is difficult. The uncertainty existing in the fatigue assessment can be identified in its higher safety factors when comparing with ultimate loading assessment. Cosack [2010] highlights that wind turbines in terms of fatigue assessment are conservatively designed by default.

Moriarty et al. [2004] showed that low m materials' life in WT is expected to be dominated by fatigue considerations. Uncertainty in the fatigue assessment of these materials is therefore a topic of high relevance.

Fatigue design analysis faces two major challenges:

(a) Fundamental definition of the fatigue assessment procedure. Different approaches may be applied to assess the structural fatigue life in the design phase and during the OWT life. Despite the strong criticism, SN fatigue assessment is still the most implemented practice for fatigue design.

(b) Characterization of load ranges, cycles, material properties and all variables that concern their definition. Different sources of uncertainty (*e.g.,* environmental conditions; material properties; modeling approximations; random representation of operation; among others) may contribute to a large overall uncertainty in the fatigue design and non-accurate risk of failure predictions. The multi-disciplinary character of the OWT design, and the large number of variables involved in the design procedure, poses a challenge to accurate probabilistic life predictions, and to the creation of techniques that fully enclose extensive uncertainty analysis in the design life.

    The recommended procedure to define the OWT fatigue loading is to perform time domain simulations and apply a counting technique in order to characterize the occurring load ranges, means and cycles. Alternatively formulation on the frequency domain can be applied, with considerable loss of accuracy.

The current chapter discussed the fatigue analysis of OWTs with particular focus on OWT loading definition. Resistance properties are an extensive area of engineering research that is common to multiple sectors of research, with some specificities to OWTs. Nevertheless, it is difficult to completely disassociate the problem of loading definition from material

considerations, e.g. structural response may be affected by material physical considerations. In relation to OWT fatigue loading, independent research lines can be identified that tackle particular OWT fatigue problems. Relevant future areas of research on OWT fatigue assessment and fatigue uncertainty reduction involve:

- Despite the increasing interest on fracture mechanics, SN method is still the "standard" approach to assess OWT fatigue. Limitations of fracture mechanics in the first stage of fatigue, initiation, still hinder its wider application. Fracture mechanics is frequently faced as a sequential alternative to SN fatigue assessment [Chryssanthopoulos and Righiniotis, 2006]. Multi-scale problems that have gained notoriety in the aerospace industry and may be a long-term solution for this problem.

- In the SN method, lack of consistency was identified in the approach to define the long-term load ranges and cycles. It was seen that different techniques were applied to define long-term structural fatigue. Nowadays, statistical extrapolation is the preferred technique. Definition of the long-term structural fatigue requires scaling-up from a time t much shorter than T. Sutherland [1999] highlights the importance of having a representative sample to perform this approximation, highlighting that a representative sample for fatigue calculations may comprise extremely consuming computations.

  In direct relation to the fatigue assessment the characterization of the loading spectra has been the main area of research for WT and OWT. With the amount of x cases to be analysed during the design phase having a sample of several hours for each x may be impossible. Given that fatigue is a problem of mean, having a representative sample may be more relevant than having a unique uncontroversial approach. The estimation of the uncertainty related to a loading spectra during the design phase is of relevance for the fatigue assessment and its optimization.

- Physical modelling still encloses significant uncertainty. In particular, computational models have been relevant contributors to improve the established practices in the sector. Yet, they are affected by multiples sources of uncertainty. An improvement of the current models, decreasing uncertainty in the design phase, may be a driver of further economical competitiveness in the sector. Fortunately, many of OWT modelling challenges are shared by extreme response and fatigue problems and therefore can be developed *in tandem*. While some problems are directly applied to fatigue calculations, such as [Marino et al., 2017], knowledge from other OWT research fields can be also applied to fatigue calculations.

- Definition of the material properties, usually given to OWT designer in form of SN, MN curves or other type of diagrams, are one of most important areas of structural fatigue uncertainty. Even with the best possible definition of all the variables that concern the OWT loading, significant uncertainty in the fatigue design would still arise from the material's fatigue properties.

- Fatigue is an effort demanding assessment. Many works were identified that research on reducing the amount of designing effort spent to assess fatigue in the design phase using simplification of the current approach or new modelling techniques e.g. [Maki et al., 2012, Zwick and Muskulus, 2016]. Future trends in the computational simulation of the OWT, increasing usage of CFD and FEM, may further increase the demand for techniques that enable for a practical assessment of fatigue in the

design phase. Frequency domain methods were developed to tackle the limitations imposed by the design effort, however with limited success.

Chapter 5

Meta-modelling of complex engineering problems

Contents

5.1 Introduction

Computational analysis is widely used in applied mechanics. While models and computers steadily grow in complexity the balance is not always favourable to the reduction of computational effort. Engineering problems can become unreasonably effort-consuming during the design phase. Recurrently, important engineering tasks, such as optimization, are hindered by the cost of the design process. In particular, system evaluation tools are increasingly expensive to process as the complexity of the engineering system increases, being an inhibitor of more comprehensive designs processes.

In the framework of making complex problems manageable, meta-models have gained significant importance. Metal-models, or surrogate models, are a common engineering approach that defines inexpensive "black-box" models to evaluate effort-consuming processes. Their definition involves finding the best emulator of the procedure being analysed. This is achieved by finding the best set of x input and y output points that approximates the global definition of the problem through a meta-model. Their implementation is highly dependent on the field of application.

In engineering activities meta-models have been widely applied to work as a surrogate of high accuracy computationally expensive models, such as the CFD or FEM models. Nevertheless, their application is not restricted to these. Meta-models, can be also applied in other contexts, such as experimental analysis or data-base reduction. The motivation

to use these as surrogates of high fidelity codes and results is related to the fact that after their definition, meta-models offer virtually no-cost to evaluate.

Even if no specific model relates x and y, meta-modelling can still be applied to build empirical surrogates that replicate the behaviour between the variables of interest. Dubourg [2011] distinguishes two sub-disciplines of meta-modelling, regression and classification. The distinction is related with the variable y. Regression approximates a continuous y to its respective x within the x continuous space. Classification also surrogates the x space, but attributing discrete labels to y.

In the context of the work proposed, the discussion presented will be focused on meta-modelling for regression problems, and with particular focus on OWT meta-modelling. Popular meta-models for regression are polynomial surfaces, support vector machines, or Gaussian process predictors. The last present interesting potential for meta-modelling when uncertainty is of relevance.

With the particular aim of discussing meta-modelling applications to surrogate complex models:

- Section 5.2 discusses the main types of meta-modelling techniques. The focus of the analysis is directed to Gaussian process predictors, also known as Kriging models, a type of meta-model with relevant interest for structural engineering problems.
- Section 5.2.1 presents and discusses the main works on OWT analysis that apply Gaussian process predictors.
- Section 5.3 introduces the mathematical background of the Gaussian process predictors; Section 5.3.1 discusses the model definition and its parameters; Section 5.3.2 discusses the case of noisy x; finally, Section 5.3.3 examines the importance of having a measure of improvement when defining the meta-model.
- Finally, the main conclusions of the review performed are presented in Section 5.4.

5.2 Meta-modelling

Meta-modelling refers to the mathematical techniques that emulate the relationship between a set of input parameters $x \in \mathbb{R}^{\mathrm{d}}$ and some output $y \in \mathbb{R}^{\mathrm{w}}$, with d and w being the respective dimensions, using a subset $X \subset x$ and the respective solutions $Y \subset y$, and taking into account some pre-established assumptions. Within the discussion of the previous chapters, a meta-model of an OWT would relate x, defined previously as environmental variables, to an output variable, such as L. However, in the framework of OWT analysis, a meta-model of an OWT could comprise more than the environmental variables. For sake of simplicity, x will refer in the current chapter to generic input variables, whereas, for the remaining of the present document, x refers to the OWT environmental loading variables.

Important regression meta-modelling for engineering applications comprise, general linear regression, polynomial chaos expansion, support vector machines or Gaussian process predictor.

General linear regression

General linear regressions are the most common type of meta-models, are also many times described as polynomial approximations. General linear regression are approximated by

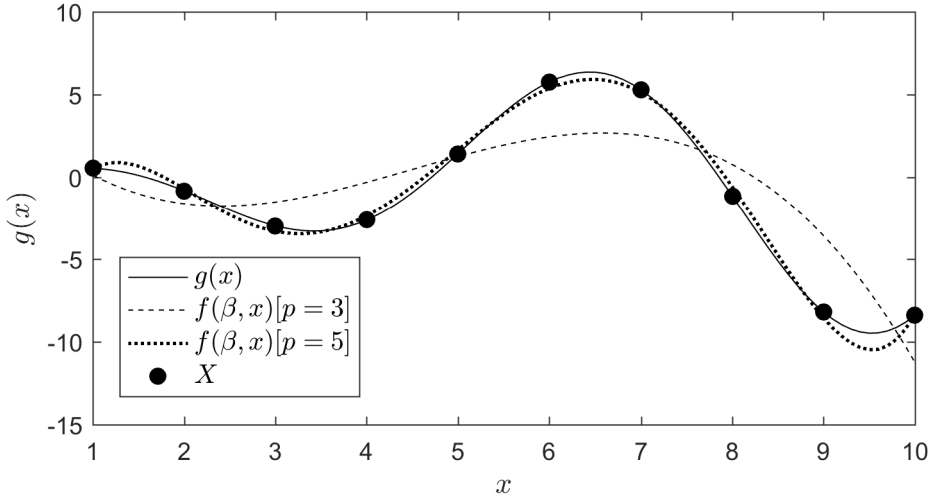


Figure 5.1: Representation of two polynomial approximations to function $g(x)$ using $p = 3$ and $p = 5$ basis functions and respective X sample used to define β.

a linear combination of p basis functions $\mathbf{f} = f_1, \ldots, f_p$.

$$f(\boldsymbol{\beta}; x) = \sum_{i=1}^{p} \beta_i f_i(x) \tag{5.1}$$

and $\boldsymbol{\beta} = [\beta_1, \ldots, \beta_p]^T$, a set of weight functions dependent on the X dataset used to define $f(\boldsymbol{\beta}; x)$. There are different techniques to find $\boldsymbol{\beta}$, the most widely known technique for the effect being the least-squares regression discussed in Section 3.3.2. Polynomial regressions will appear more commonly with a constant term β_0, which can be easily incorporated by considering an additional term in the basis function that is equal to one. A common variation that has close relation to the linear regression approximation is the response surface methods, which are surrogate approximations that consider a 1^{st} or 2^{nd} order polynomial approximations using coded variables and an error term.

Polynomial Chaos Expansion

A particular case of polynomial approximation that has been widely used as surrogates is the Polynomial Chaos Expansion. This polynomial approximation resembles the formulation presented for the generic linear functions, but instead it defines the approximation of $g(x)$ with dependency on stochastic inputs. In some applications the input variables of $g(x)$ follows a distribution function given by f_x, their PDF. If $g(x)$ is a complex finite variance model, then it may not be sufficient to introduce the mean of f_x in the computations to calculate the mean of $g(x)$.

The polynomial chaos method consists then in finding a polynomial family that is orthogonal with respect to the input distribution of x, f_x, and fitting a polynomial response surface taking into account the statistical behaviour of the input variables.

Considering that x is described by a joint probability density function f_x. The polynomial chaos expansion of $g(x)$ is defined as:

$$g(x) = \sum_{i=1}^{P} a_i \varrho_i \tag{5.2}$$

where a_i are a series of deterministic coefficients and ϱ_i are a set of multivariate orthogonal polynomials. The coefficient P represents the truncation order of the polynomial approximation. Different methods exist to fit $g(x)$, being the most popular the projection estimation, which uses quadrature methods.

Recently, Murcia et al. [2018] applied polynomial chaos expansion to quantify uncertainty using it as surrogate of DTU's 10 MW OWT.

Some of the main families of orthogonal polynomials that relate to the most important distributions are well defined, see Table 5.1.

Table 5.1: Examples of orthogonal polynomial families.

PDF(x)	**Orthogonal polynomials**
Uniform	Legendre
Gaussian	Hermite
Gamma	Laguerre
Beta	Jacobi

Support Vector Machines

Support Vector (SV) machines is popular learning technique initially created for classification problems, and later extended to regression problems.

Its definition resembles a problem of classification where points are classified in a $c \pm 1$ category in order to find a linear boundary between the two classes given by the hyperplane $SVM(x)$ that maximizes a distance $2/||\boldsymbol{w}||^2$ between the points that define the support vectors, see Figure 5.2. Where $\boldsymbol{w}$ is a vector of weights.

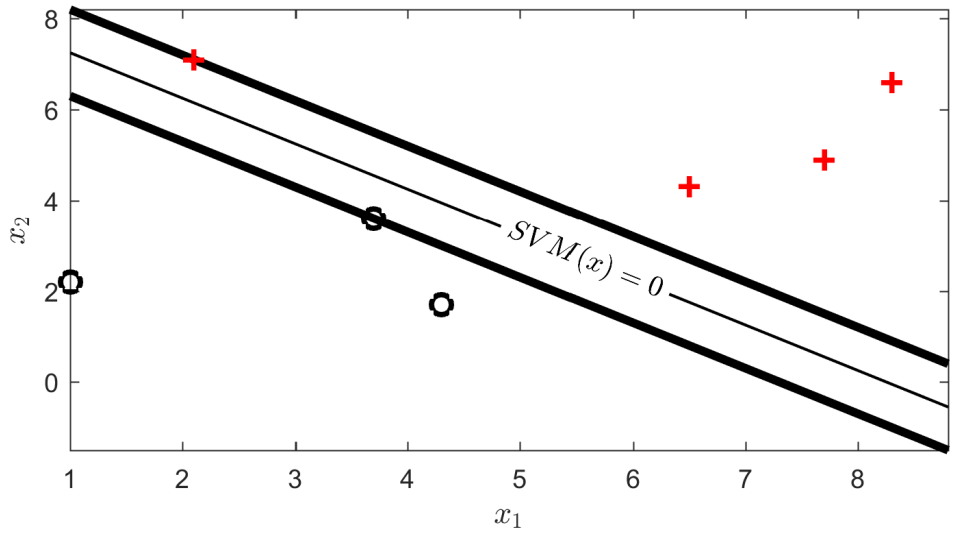


Figure 5.2: Example of support vector machines problem. Plus marker - $c = +1$ and round marker - $c = -1$.

The support vectors are the points identified with SV. These points define the margin, which is the distance between the two classes (grey area), and also the distance given by

$\boldsymbol{w}$. The best estimate for this hyperplane $SVM(x)$ is attained when the width of the margin is maximized.

The $\boldsymbol{w}$ can be calculated with information about α_s Lagrange multipliers and s SV points, making irrelevant the information about the other non-SV points [Smola and Schölkopf, 2004]. $SVM(x)$ is then defined as,

$$SVM(x) = \sum_{i=1}^{s} \alpha_i c_i x_i \tag{5.3}$$

This is a very simple example of a linearly separable support vector machine application. In this case it was straightforward to infer on the hyperplane that separates the classified points. However, the separation is not always so simplistic (e.g. commonly there are mixed classification points above and below $SVM(x)$) and a change in the variables space is demanded. For such scenarios a non-linear mapping is recommended.

This is achieved using a *kernel trick* [Aizerman, 1964], which projects the support data in a higher dimensional space. In this case, when dealing with a sample of variables x and a classification $c \pm 1$, the support vector machines considers that there is a space $\mathbb{H}(K)$, called the feature space, where c is linearly separable. The extension to the feature space is obtained applying a Ψ transformation such that the inner product ($\langle \cdot, \cdot \rangle$) in this space can be computed using a kernel function K in the former space:

$$\langle \Psi(x_i), \Psi(x) \rangle = K(x_i, x); \tag{5.4}$$

The main advantage of this consideration is that there is no need to explicitly know Ψ. Instead $SVM(x)$ can be calculated using K.

$$SVM(x) = \sum_{i=1}^{s} \alpha_i c_i \left(K(x_i, x) - b \right) \tag{5.5}$$

An example of a structural application of this technique is the definition of limit state functions with a separation between failure and non-failure through $c \pm 1$. In this case the margins can be used to further improve the $SVM(x)$ accuracy. b is a bias term. The main disadvantage of support vector machines in this form is that to define $SVM(x)$ the nature of the problem should admit classification.

To perform support vector machine regression the same theoretical background is used, but instead of using the maximum margin condition, the requirement is to find the flattest function which performs a regression with a pre-specified deviation ϵ. More details on the usage of support vector machines for regression and its derivation is presented in Smola and Schölkopf [2004].

Kriging or Gaussian process predictor models

Kriging models, also known as Gaussian process predictor, can be seen as exact interpolators based on the idealization of the approximated model response as the realization of a Gaussian stochastic process [Echard et al., 2011]. They perform as some of the described regression models, but consider Gaussian uncertainty in between the points of the support sample X, Figure 5.3.

The Gaussian process predictor, or Kriging model, $G(x)$ has μ and σ components that approximate uncertainty in the regression of $g(x)$ as a Gaussian process. This elegant regression technique has its origin in geostatsitics applications and was introduced in 1951 by Krige in his pioneering work on mining Krige [1951].

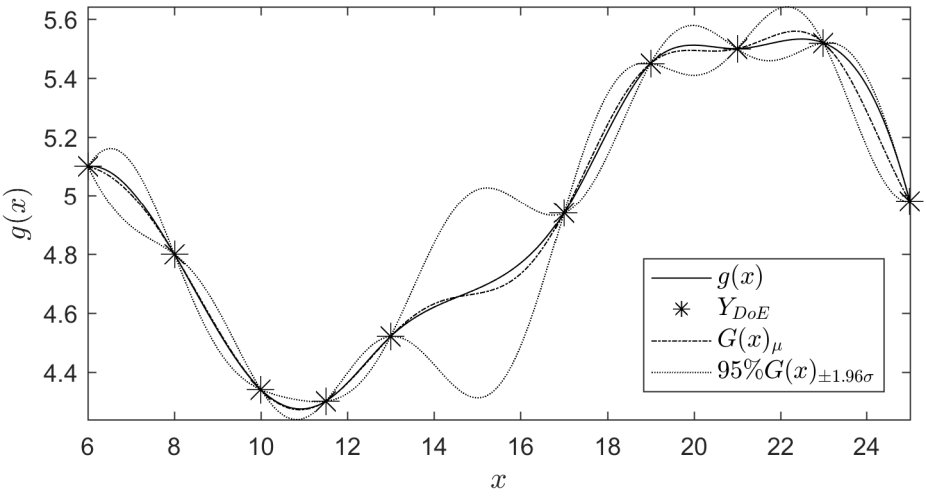


Figure 5.3: Example of a Kriging regression ($G(x)$) approximation to function $g(x)$.

Apart from the meta-modelling techniques discussed, other powerful techniques can be used to surrogate complex models. Some approaches do not even relate directly to regression problems, such as artificial neural networks. This technique has received significant interest for multiple applications, including engineering applications (see Section 5.2.1). However, artificial neural networks commonly demand high training data-sets to perform accurate predictions. Their application is limited when the goal is to create surrogate models that mitigate the effort of evaluating the originally complex model.

Within the work proposed, application of the Gaussian process predictor models as surrogates of the OWT is researched. In the author's view these meta-models have the potential to be applied as surrogates of the OWT complex model. They can also be applied as surrogates of the expensive limit state function evaluation. Their intrinsic statistical character is of interest for some practical applications, such as quantifying uncertainty. Any of the other models could be used as an alternative for OWT surrogate, but the Gaussian uncertain character of these regression models will be seen to be major interest for OWT fatigue calculations; and eventually to characterize its uncertainty too.

The application of Gaussian process predictor models for OWT applications is not new. Its relevance for the field can be identified in the growing number of research works that try to exploit the potential of these regression models to surrogate WT or OWTs.

5.2.1 OWT structural analysis using Gaussian process predictor

Examples of the application of the Gaussian process predictor models for structural engineering problems can be found in [Bichon et al., 2008, Echard et al., 2011, 2013, Gaspar et al., 2014, Yang et al., 2015, Zhang et al., 2015, Tian et al., 2018]. In the particular case of OWT analysis, the application of Gaussian process predictors as surrogate models is relatively new. Only a limited number of works that merge wind engineering and Gaussian process predictors can be identified.

The application of Gaussian process predictor for OWT structural analysis was first highlighted in Yang et al. [2015], where a reliability-based probabilistic optimization of a Tripod foundation is developed. The variable of optimization considered was the weight of the structure which was subjected to specific constraints, with reliability being checked

through Monte-Carlo simulations. The computational advantage of using a surrogate model is highlighted, and a more comprehensive design is obtained while maintaining reduced computational effort. Latin Hypercube Sampling (LHS) was used to set the model due to its efficiency in covering x.

Following, in Morató et al. [2016] the response of an OWT is modeled using a Gaussian process predictor. The authors approximate the response of an OWT model with a regression surface also motivated by the reduction of the computational time. Two limit states functions are considered in the analysis. In both works, Latin Hypercube sampling is used to set the design of experiments (DoE). The DoE is the set of evaluation the support the definition of the meta-model.

While these models steadily gained interest for analysing OWTs, their usage in the wind energy industry dates from before. In Maki et al. [2012] the WT response is approximated with a Gaussian regression model. The Gaussian process predictor predictions/model replace(s) the computational expensive coupled dynamic code, enabling a multi-level optimization approach 1MW wind turbine blades.

In fatigue design, the usage of Gaussian regression is not common. However, the potential for its application is high. Echard et al. [2013] reduced the computational time of a fatigue reliability analysis from more than 100 days to approximately 9 hours by using a surrogate model and adaptive sampling methodologies, showing therefore the huge benefits in reducing computational cost resulting from the application of a Gaussian regression model, and without compromising accuracy in the results.

In Yang and Wang [2012] a structural fatigue optimization of a bending stiffener is presented. Three metamodels are compared, among which a Gaussian regression model, to reduce computational time.

Teixeira et al. [2017b] applies a Gaussian regression to assess fatigue by creating fatigue damage surfaces. In this work a simple one-dimensional implementation was developed without assessment of the DoE infill criteria or model variance. The surrogate was self corrected for estimation of the lifetime fatigue damage.

5.3 Gaussian process predictors

5.3.1 Modelling basis

The Gaussian process predictors approximate a true state function $g(x)$ that depends on an input space $x \in \mathbb{R}^{\mathrm{d}}$, in a d dimensional space, with an approximate mathematical model $G(x)$ that considers uncertainty in the approximation.

Assuming that the true function $g(x)$ can be defined $\forall x$, the process of defining $G(x)$ demands a sample of k support points or observations that are usually designated as Design of Experiments (DoE); with this $DoE = [\boldsymbol{X}, \boldsymbol{Y} == g(X)]$, with $\boldsymbol{X} = [x_1, x_2, \dots x_k] \subset x$ representing a vector of realisations of x and $\boldsymbol{Y}$ being the respective true evaluations of $g(x)$ at $\boldsymbol{X}$.

Using a Gaussian process predictor as a surrogate model assumes that the true response function $g(x)$ can be approximated as

$$G(x) = f(\boldsymbol{\beta}; x) + Z(x), \quad \forall x \in \mathbb{R}^{\mathrm{d}} \tag{5.6}$$

$$f(\boldsymbol{\beta}; x) = \beta_1 f_1(x) + ... + \beta_p f_p(x), \quad \forall x \in \mathbb{R}^{\mathrm{d}} \tag{5.7}$$

with $f(\boldsymbol{\beta}; x)$ as a deterministic function characterized by a regression model with p ($p \in \mathbb{N}^+$) basis trend functions ($f_p(x)$), and p regression coefficients ($\boldsymbol{\beta}$) to be estimated using the known sample $\boldsymbol{X}$. The p number of coefficients is function of the approximation intended in the regression.

$Z(x)$ is a Gaussian stochastic process with zero mean:

$$\mathbb{E}[Z(x)] = 0, \quad \forall x \in \mathbb{R}^{\mathrm{d}} \tag{5.8}$$

which relates to a stationary covariance matrix:

$$\mathbf{C}(x_i, x_j) = \sigma_c^2 R(x_i, x_j; \boldsymbol{\theta}) \qquad with\ i,\ j = 1, 2, 3, ..., k \tag{5.9}$$

the covariance matrix $\mathbf{C}$ relates generic $\boldsymbol{X}$ points using; σ_c^2 which is a scale parameter called constant process variance and a correlation function $R(x_i, x_j; \boldsymbol{\theta})$.

For the structural analysis, $\mathbf{C}$ frequently takes the so-called, and widely applied in computer experiments [Roustant et al., 2012], *separable* form in Equation (5.10). Other types of correlation can be applied, such as *ellipsoidal* correlation functions [Rasmussen, 2004, Marelli and Sudret, 2014].

$$R(x_i, x_j; \boldsymbol{\theta}) = \prod_{i=1}^{d} R(h_i; \theta_i),\ \boldsymbol{\theta} \in \mathbb{R}^{\mathrm{d}} \tag{5.10}$$

The correlation function depends then on $h = [h_1, \ldots,\ h_d]$, a set of incremental values of type $x - x_i$ type and $\boldsymbol{\theta}$ hyperparameters. For a given sample of support points the problem of prediction can be solved using a generalised least squares formulation, where the estimators for $\boldsymbol{\beta}$ and σ_c^2 depend on $\boldsymbol{\theta}$ and are given by:

$$\boldsymbol{\beta} = \hat{\boldsymbol{\beta}}(\boldsymbol{\theta}) = (\mathbf{F}^T \mathbf{C}^{-1} \mathbf{F})^{-1} \mathbf{F}^T \mathbf{C}^{-1} \boldsymbol{Y} \tag{5.11}$$

$$\sigma_c^2 = \hat{\sigma}_c^2(\boldsymbol{\theta}) = \frac{1}{k} \left(\boldsymbol{Y} - \mathbf{F}\hat{\boldsymbol{\beta}} \right)^T \mathbf{C}^{-1} \left(\boldsymbol{Y} - \mathbf{F}\hat{\boldsymbol{\beta}} \right) \tag{5.12}$$

$\mathbf{F}$ is the k x p regression matrix in which the rows are trend functions $f_p(x)$ evaluated at the k support points. A Maximum Likelihood formulation can be used to optimize the likelihood of the observations $\boldsymbol{Y}$. The likelihood function based on the assumption that $\boldsymbol{Y}$ points follow a Gaussian distribution is given by,

$$\mathcal{L}(\boldsymbol{Y}|\beta, \sigma_c^2, \boldsymbol{\theta}) = \frac{1}{(2\pi\sigma_c^2)^{k/2} [\det \mathbf{C}]^{1/2}} \exp\left[-\frac{1}{2\sigma_c^2} (\mathbf{Y} - \mathbf{F}\boldsymbol{\beta})^T \boldsymbol{C}^{-1} (\mathbf{Y} - \mathbf{F}\boldsymbol{\beta}) \right] \tag{5.13}$$

Merging Equations (5.11) and (5.12) with the likelihood function allows the calculation of the $\boldsymbol{\theta}$ parameters that maximize the likelihood. This is achieved through the minimization of the negative natural logarithm of $\mathcal{L}(\boldsymbol{Y}|\beta, \sigma_c^2, \boldsymbol{\theta})$,

$$\boldsymbol{\theta} = \arg\min(-\log \mathcal{L}(\boldsymbol{Y}|\beta, \sigma_c^2, \boldsymbol{\theta})) \tag{5.14}$$

This optimization can be also conveniently represented using a reduced likelihood function [Dubourg, 2011],

$$\Omega(\boldsymbol{\theta}) = \sigma_c^2 [\det \mathbf{C}]^{1/k} \tag{5.15}$$

where Ω is called the reduced likelihood function.

The solution to the presented optimization problem cannot be solved analytically and usually requires the application of global optimization techniques such as the ones applied in Marelli and Sudret [2014] and Roustant et al. [2012].

Any prediction for the true realisation $g(u)$ in a point $u \in x$ in the space is then given based on the Kriging expected value μ_G and variance σ_G^2:

$$\mu_G(u) = f(u)^T \, \boldsymbol{\beta} + \mathbf{c}(u)^T \mathbf{C}^{-1}(\boldsymbol{Y} - \mathbf{F}\boldsymbol{\beta}) \tag{5.16}$$

$$\sigma_G^2(u) = \sigma_G^2 \left(1 + D(u)^T (\mathbf{F}^T \mathbf{C}^{-1} \mathbf{F})^{-1} D(u) - \mathbf{c}(u)^T \mathbf{C}^{-1} \mathbf{c}(u)\right) \tag{5.17}$$

$$D(u) \equiv \mathbf{F}^T \mathbf{C}^{-1} \mathbf{c}(u) - f(u); \tag{5.18}$$

where $\mathbf{c}(u) = c(u, x_i), \; i = 1, 2, \ldots, k$ is the correlation vector that relates the realisation to be evaluated with the known points, and $f(u)$ is the vector of trend functions evaluated at u. $D(u)$ is introduced for the sake of brevity.

Equations (5.16) and (5.17) offer the best unbiased predictor of u. Proof and more details are presented in Dubourg [2011].

Table 5.2 presents the main correlation families used in the definition of the Gaussian process predictors. In the framework of structural analysis the exponential Gaussian correlation function has been widely implemented.

Table 5.2: Examples of correlation families used with the Gaussian process predictors. $|h|$ refers to the absolute value of the difference h.

Gaussian	$R(h, \boldsymbol{\theta}) = \exp\left(-\left(\frac{h}{2\theta}\right)^2\right)$
Matern $v = 5/2$	$R(h, \boldsymbol{\theta}) = \left(1 + \frac{\sqrt{(5)}\lvert h\rvert}{\theta} + \frac{5h^2}{3\theta^2}\right) \exp\left(-\frac{\sqrt{(5)}\lvert h\rvert}{\theta}\right)$
Matern $v = 3/2$	$R(h, \boldsymbol{\theta}) = \left(1 + \frac{\sqrt{(3)}\lvert h\rvert}{\theta}\right) \exp\left(-\frac{\sqrt{(3)}\lvert h\rvert}{\theta}\right)$
Exponential	$\exp\left(-\frac{\lvert h\rvert}{\theta}\right)$
Power-Exponential	$\exp\left(-\left(\frac{\lvert h\rvert}{\theta}\right)^p\right)$

These correlation families have a decreasing form. The value of the correlation decreases with the increase of h, see Figure 5.4. Other correlation families can be found in Rasmussen [2004], Marelli and Sudret [2014].

The correlation family may have important influence on the $G(x)$ predictions, see Figure 5.4. Nevertheless, a more significant influence on the $G(x)$ predictions can be attributed to $\boldsymbol{\theta}$. The influence of the other model parameters is also related to $\boldsymbol{\theta}$. Two one-dimensional prediction models are considered for two values of $\boldsymbol{\theta}$, and two correlation families. The most common approach is to use a search algorithm for the selection of the most suitable $\boldsymbol{\theta}$.

$\boldsymbol{\theta}$ is therefore the most important parameter in the definition of $G(x)$ because it directly defines the remaining model variables. If the goal is to approach a stochastic field instead

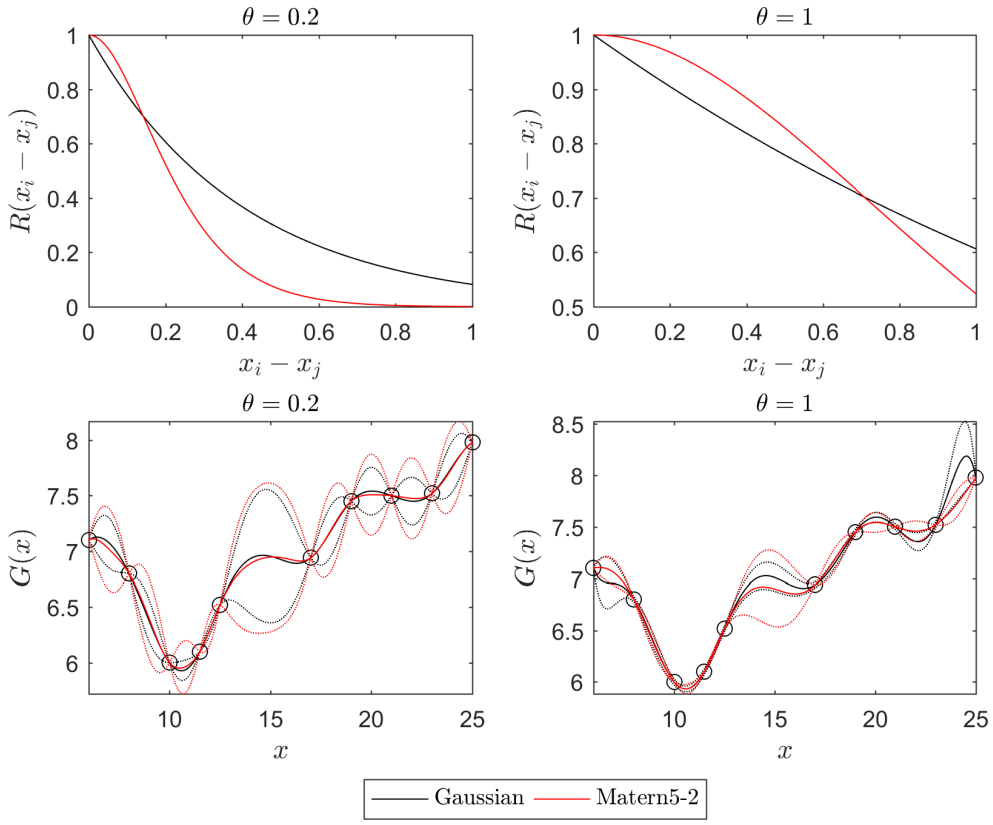


Figure 5.4: Gaussian process predictor function of the correlation function for $\theta = 0.2$ and $\theta = 1$. On the top: correlation family as a function of h. On the bottom: Predictor function.

of a $g(x)$ function it may be of interest to perform the $\boldsymbol{\theta}$ selection manually. However, such a procedure is not common in $G(x)$ applications.

For polynomial approximations, many applications use a zero order polynomial function. In this case Equation (5.7) reduces to $f(\beta, x) = \beta_1$ and the Kriging is called *ordinary Kriging*.

Gaspar et al. [2014] showed that higher order polynomials may not contribute to improve accuracy of the estimations in the particular case of structural reliability engineering problems. Teixeira [2018] also analysed the influence of model parameters on predictions, with similar findings in relation to the polynomial approximation.

One important consideration regarding the $G(x)$ prediction is that, in a Gaussian prediction model a function is approximated in a way that it is exactly predicted at a point u if this point belongs to $\boldsymbol{X}$. This means that $\mu_G(u)$ gives a deterministic prediction of $G(u)$ for which $\sigma^2_{G(u)}$ is null. This assumption in the Gaussian process model is strong and it does not depend on the input parameters. It is always verified if the convergence of the optimization problem is successful.

5.3.2 Noisy or stochastic DoE

A local variance of the DoE can be introduced in the $G(x)$ formulation in order to account for uncertainty in $\boldsymbol{Y}$. In such a case each component of $\boldsymbol{Y}$ will no longer be a deterministic

true realisation of the $g(x)$ function, but instead a stochastic variable Y_i defined as,

$$Y_i = Y_e + \epsilon_i \tag{5.19}$$

with Y_e being the expected value of a true realisation Y_i and ϵ_i an additional variable which represents a Gaussian process with mean 0 and variance equal to $\sigma^2_{Y_i}$. This variance component is evaluated at $\mathbf{C}$.

$$\mathbf{C}(x_i, x_j) = \mathbf{C}(x_i, x_j) + \boldsymbol{\delta}\boldsymbol{\tau}^{\mathbf{2}}. \tag{5.20}$$

where $\boldsymbol{\tau}^{\mathbf{2}}$ is the vector of variance $\sigma^2_{Y_i}$ of the realisations of $\boldsymbol{Y} \in g(x)$ used to define the surrogate model. $\boldsymbol{\delta}$ is the identity matrix of size k. The Gaussian process predictor model

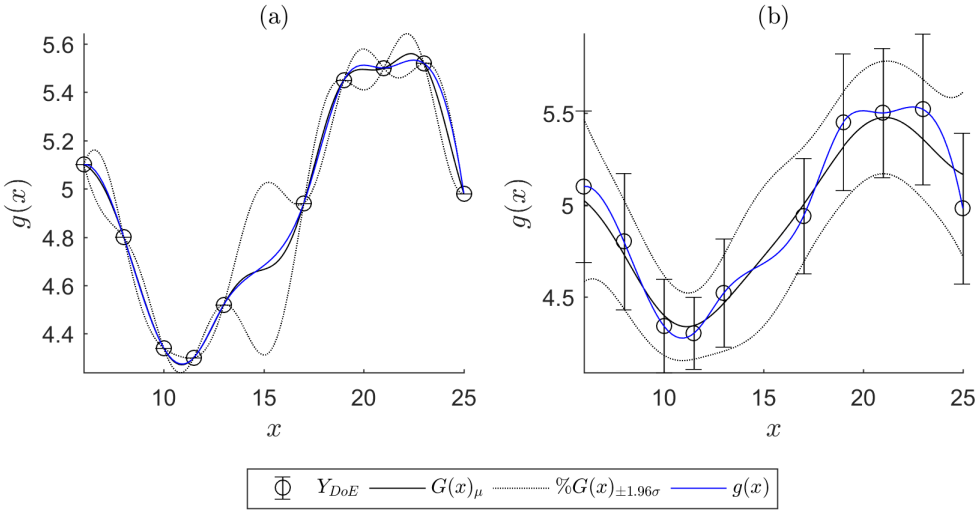


Figure 5.5: Example of non-noisy and noisy Gaussian process predictor model considering a non-stochastic and stochastic $\boldsymbol{Y}$ respectively.

considering $\boldsymbol{\tau}^{\mathbf{2}}$ does not have a deterministic prediction in the $\boldsymbol{X}$ points used to define it. At the same time, the local variance along with the constant process variance approach better the local predictions of variance.

5.3.3 Selection of the DoE

The DoE is of significant importance to generate an accurate Gaussian process predictor. It defines the points in x that will be used to build the model.

When defining a meta-model to replace a high fidelity code it is generally not efficient to simply pick random points from x. Meta-models applied efficiently to engineering analysis consider some *learning* criteria that is responsible to select new points in the DoE.

It is possible to use simple techniques to define the DoE, such as factorial designs, Monte Carlo sampling (MC) or Latin Hypercube Sampling (LHS).

Factorial design sampling, as in Saltelli et al. [2008], are an example of deterministic DOE definitions. If a full factorial design is applied, it involves running all possible combinations over a grid of j divisions of the x space. This method is severely affected in applicability by the dimension of the x space. With $j = 10$ and $d = 4$, 10^4 evaluations

of $g(x)$ are demanded to defined the full factorial sample. Variations of the full factorial include the fractional factorial, however their interest is limited when the analysis cost-efficiency and random character are important.

In such scenarios random sampling is more interesting. The most popular random sampling techniques are the MC and LHS. MC consists in sampling a space x using a random generator that follows the probability distribution of x. One natural limitation of the MC is related to the difficulty of efficiently covering areas of x with low probability of occurrence. This method is often seen as of expensive use when compared with other random sampling alternatives. LHS also considers the probability distributions of x, but guarantees that at least one point is taken at each equivalent probability division of the x space. A comparative example of both is presented in Figure 5.6.

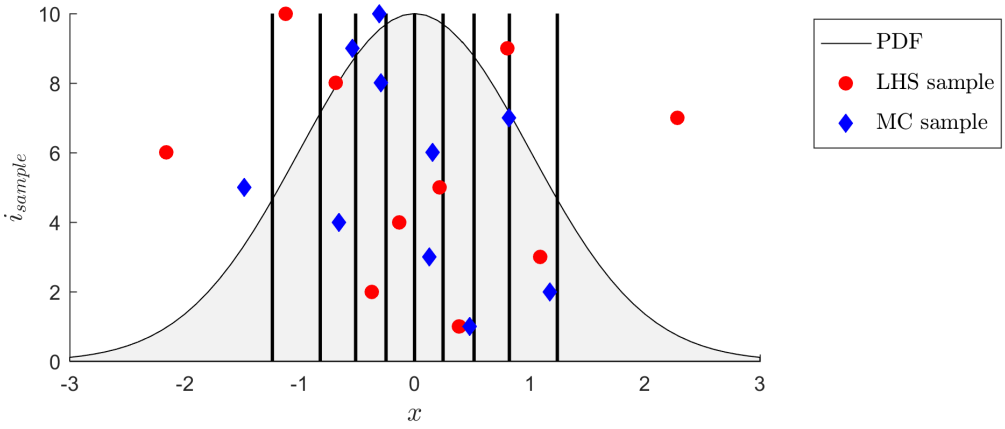


Figure 5.6: Representative example of a 10 points ($i = 1, \ldots, 10$) LHS and MC sample from a $\mathcal{N}(0, 1)$.

The LHS guarantees a better coverage of the x space, while the MC samples naturally concentrate most of its sample in the middle of the x distribution. Efficiency with the MC is commonly only achieved with very high sample sizes, and its interest is limited when computational efficiency is of relevance. Other simple alternatives for sampling are presented in Saltelli et al. [2008].

When sampling the DoE space in pursuit of a surrogate that minimizes the effort in approximating $g(x)$, the methods described are rarely of interest. They can be used as a starting point to apply more efficient DoE characterization techniques, such as the adaptive DoE search. One of the particularities of the Gaussian process predictor is the possibility of using it (as it is quantified in uncertainty by its $Z(x)$) as a measure of further improvement of the model.

Picheny et al. [2013] discussion on the infill criteria for applications on the noisy case highlights the importance of having a notion of improvement for the implementation of efficient Gaussian process predictors. This rationale refers to the need of considering the problem as a whole when selecting the criteria to infill the DoE.

In order to have efficient implementations it is common to start with a subset of k points and incrementally infill the DoE with new points using some form of search algorithm. This guarantees that the implementation effort is directed to the required points to build an accurate approximation of $g(x)$.

The Kriging prediction variance increases as the prediction region moves away from known DoE points. These are regions where less is known about the true function $g(x)$ and hence the built model is more likely to be inaccurate.

The idea of using the predictor to model new iteration in the DoE was highly fomented with the appearance of the *efficient global optimization* (EGO) concept. The most relevant work in this context was the introduction of the Expected Improvement (EI) criteria in Jones et al. [1998].

Expected Improvement

The EI algorithm [Jones et al., 1998] uses a global search function denoted by:

$$EI(x) = \mathbb{E}\left[(min(\boldsymbol{Y}) - G(x))^{+} | G(X) = \boldsymbol{Y_n}\right] \tag{5.21}$$

which can be written as,

$$EI(x) = (min(\boldsymbol{Y}) - \mu_G(x))\Phi\left(\frac{min(\boldsymbol{Y}) - \mu_G(x)}{\sigma_G(x)}\right) + \sigma_G(x)\phi\left(\frac{min(\boldsymbol{Y}) - \mu_G(x)}{\sigma_G(x)}\right) \tag{5.22}$$

The EI is the expected difference between the minimum currently best-known prediction $\boldsymbol{Y}$ at iteration i and the conditional $G(x)$ built on the assumption that $G(X) = \boldsymbol{Y_i}$, with i denoting the iteration index and X the DoE points used to build $G(x)$. Φ and ϕ are respectively Gaussian cumulative and probability density functions.

The $k+1$ index point in the DoE that should be used to update the Kriging model should be selected using a search in,

$$x_{k+1} = \arg\max\left[EI(x)\right] \tag{5.23}$$

and the new DOE iteration becomes,

$$\boldsymbol{X} = [x_1, \ldots, x_k, x_k + 1] \tag{5.24}$$

By construction, the EI criteria is always non-negative and its magnitude increases with σ.

From its publication date, many modifications of the EI emerged, some are discussed in Table 5.3.

True $G(x)$ assumption

A natural DoE infill criteria to use is to search for the quantities of interest by assuming that $G(x)$ is the true realisation of $g(x)$. Such an example can be identified in optimization problems. In search of a function minimum, the minimum of $G(x)$ can be taken by the expectation that it will close to the minimum of $g(x)$ and hence $x_k + 1$ is defined as,

$$x_{k+1} = \min G(x), \quad \forall x \in \mathbb{R}^{\mathrm{d}} \tag{5.25}$$

This method is less efficient when compared with th EI [Jones, 2001]. Nevertheless, its implementation is simple and hence its application may be of interest in some scenarios.

Several other criteria can be identified as infill sampling, Table 5.3.

Table 5.3: Infill criteria.

Infill criteria	Application	Reference
non-Noisy DoE		
σ_G^2 Error	One simple alternative approach to the true $G(x)$ assumption is to use the point x_{k+1} point were σ_G^2 is maximum.	[Liu et al., 2012]
Maximum Probability of Improvement	Uses $G(x)$ uncertainty to define the probability that a certain point will be a better estimate on a search for minimum of $g(x)$. The function to maximize I as $$I(x) = \Phi\left(\frac{min(\boldsymbol{Y}) - \mu_G(x)}{\sigma_G(x)}\right) \quad (5.26)$$ It is a method applied to the problem of maximum/minimum search.	[Kushner, 1964]
Margin indicators (U function)	The idea of margin indicator was implemented for reliability applications with $G(x)$. The U function is an example of a criteria that uses these for reliability, and is defined as $$U(x) = \frac{\lvert\mu_G(x)\rvert}{\sigma_G^2(x)}$$ It predicts the distance to the real limit state function $g(x)$. A set of initial DoE is evaluated and fitted, the next x_{k+1} points is selected based on the minimization of U. The minimum is found in two circumstances, or $G(x)$ is close to 0, which means that it is close to the limit state function, or $\sigma_{G(x)}$ is large.	[Echard et al., 2011]

Continued on next page

Continued from previous page

Surprise minimization	The surprise minimization criteria aims to minimize the probability that a true value deviates significantly from its predicted value. It is written as $$Wb = \min_{x_{k+1}} \max_{v} [\sigma_G(x|\boldsymbol{X}, x_{k+1})]$$ The x should be selected at the minimum x that maximizes the variance at a unobserved location v. This criteria was introduced initially to select the x following candidate points assuming that x_{k+1} was known.	[Watson and Barnes, 1995]
Noisy DoE		
Minimum quantile improvement	Similar to the assumption that a quantity of interest is close to the quantity of interest in $G(x)$, the minimum quantile improvement uses this assumption to compute x_{k+1}. If a minimum is sought the search function is: $$x_{k+1} = \arg\min_{x \in \mathbb{R}^{\mathrm{d}}} \mu_G(x) + \phi(\beta)^{-1} \sigma_G(x)$$ where $\Phi(\beta)^{-1}$ increases with the number of iterations.	[Cox and John, 1997]
Augmented EI	Augmented EI is a variation of the classical EI that uses a penalty function. The augmented EI is calculated through, $$AEI(x) = EI(x) \left(1 - \frac{\tau}{\sqrt{\sigma_G^2(x) + \tau^2}} \right)$$ the application of the penalty function is justified due to its capability of penalising DoEs with small $\sigma^2(x)$, and promoting exploration.	[Huang et al., 2006]

Continued on next page

Continued from previous page

Reinterpolation	The reinterpolation consists in using an auxiliary model to select the new Gaussian process DoE iteration in order to use the EI with noisy observations. This auxiliary model is therefore a non-noisy model used to predict a new point using the EI criteria. In the computation of noisy DoEs, the reinterpolation guarantees that the new prediction is not a previously predicted point.	[Forrester et al., 2006]
Expected quantile improvement	Assumes that for a noisy model a $G(x)$ prediction may be closer to the true value than the original data. Therefore, assuming a quantile $$I_{EQI}(x) = \left(\min_{1<i<k} \left(Q_k(x_i) - Q_{k+1}(x) \right) \right)^+$$ where the Q are pre-defined quantiles for the k known noisy data and $k+1$ update at $G(x)$ considering $x = x_{k+1}$.	[Picheny et al., 2010]
Approximate knowledge gradient	The knowledge gradient function aims at defining the global effect of a new measurement on $G(x)$ by measuring a mean function M. $$I_{AKG}(x) = \min[M_k(x_{\boldsymbol{X}_{k+1}})] - \min[M_{k+1}(x_{\boldsymbol{X}_{k+1}})]$$ where M_k and M_{k+1} represent the mean of $G(x)$ at $G(\boldsymbol{X}_{k+1}\|\boldsymbol{X} == \boldsymbol{X})$ and at $G(\boldsymbol{X}_{k+1}\|\boldsymbol{X} == \boldsymbol{X}_{k+1})$ assuming that X_{k+1} is updated at $x_{k+1} = x$.	[Scott et al., 2011]

There are other criteria that can be identified to infill the DoE space, e.g. Dubourg [2011]. No DoE search criteria was identified to address the application of the Gaussian process predictor as a surrogate of a complex model. Most of the work developed on infill criteria is directed to the application of the Gaussian process predictors for optimization problems. The examples found in the literature for OWT analysis, applied LHS to define the DoE.

5.4 Conclusion

An overview of meta-modelling with Gaussian process predictors to approximate complex simulation models was given in the current chapter. A brief discussion of different meta-modelling techniques was presented. Within the diverse existing meta-models, Gaussian process predictors have been target of active research for structural engineering problems.

One of the main advantages identified in the usage of Gaussian process predictors is their capability to interpolate not only the response of the model for which they are

defined as surrogates, but also its uncertainty. The Gaussian process predictors can then be described as an infnite number of potential curves within a single model.

This capability of comprising prediction and uncertainty is of most interest for engineering applications. Nevertheless, some limitations can be identified when applying Gaussian model predictors for engineering applications:

- Limited interest was identified in the literature, in the context of engineering applications, in the identification of the sensitivity of the model to the definition of the correlation function or hyperparemeters convergence algorithm and interval. These are recurrently neglected in respect of their influence in the $G(x)$ approximation.

- Gaussian process predictors are widely used for surrogate problems, but limited research was performed on the infill criteria for surrogate complex models. Commonly, a LHS sample is applied and assumed to suffice an accurate approximation of x, consequently $g(x)$. As Picheny et al. [2013] highlights, the application of these models requires a measure of improvement in order to achieve efficient implementations.

The last bullet point is of major relevance if efficient surrogate algorithms are sought. In the particular case of fatigue analysis, and when the Gaussian predictor is used as an approximation to model fatigue, important fatigue contributions may be neglected using a Gaussian process predictor generated with a random sampling and validated with cross-validation.

Part III

Original contributions

Chapter 6

Extrapolation of extreme waves with the Peak-Over-Threshold method

Contents

The analysis presented in the following chapter was published in Teixeira et al. [2018a]

6.1 Motivation and original contribution of this chapter

The following chapter debates the extrapolation of wave data using the POT methodology. The goal of the chapter is to discuss the distinct phases of the extrapolation process using

the POT. The motivation for such analysis emerges from the low agreement identified in the literature regarding the probability modelling of *exceedance* H_s data.

To analyse *exceedance* data, it is very common to find the POT methodology applied in combination with a Generalised Pareto - Poisson probability model. However, when it comes to the modelling of H_s data, literature that advises against the application of the Generalised Pareto in detriment of the two parameter Weibull and the Exponential distributions can be found. Since the Generalised Pareto is the limit distribution for *exceedance* data, it is of interest to research on the applicability of the Generalised Pareto as an alternative to fit H_s *exceedance* data.

Real H_s data from four oceanographic buoys installed in Irish waters is considered to research on the extrapolation using *exceedance* H_s data. The challenging procedure of choice of threshold, motivated the introduction of an innovative methodology to characterize the threshold level u when analysing H_s data. Five evaluators of the probabilistic approximation are applied in the analysis. A sensitivity analysis is conducted to ensure robust results. Finally, the return periods are estimated for the probability models compared, and the comparative results are extensively discussed.

The fundamentals of the POT here discussed can be extended to the extrapolation of other physical quantities.

6.2 Introduction

The analysis of *exceedance* data is a very efficient alternative to deal with extreme values. Extreme values, as mentioned, are of difficult characterization due to their scarcity. If availability of data is a limitation, the difficulty of characterizing extreme values is accentuated. The POT is an efficient alternative to model extreme occurrences due to its capability of enclosing more data in the analysis than similar alternatives that extrapolate physical quantities.

The main challenge of the extrapolation is the need to define occurrences based on what is currently known from the dataset. If 50 years of data are available, defining the 50 years maxima is straightforward.

However, having such a long dataset is rarely the case. Monitoring and recording data is a relatively recent trend, and as a result, defining the 50 years maxima can be a challenging task. Additionally, even if records for a similar or close quantity are available (e.g. H_s data from a near site can be available), it is not necessarily true that this dataset will be a good approximation of the quantity of interest.

In the case of the POT a truncation technique is used to increase the relative statistical weight of the tail region, increasing the likelihood of having an accurate statistical description of the tail region. A recommended POT procedure is presented in Figure 6.2.

There are different stages in the analysis of data with the POT. A threshold level u needs to be defined in first place. This u should be carefully selected depending on the data set. The data is truncated according to u, and independence needs to be ensured in between the truncated points. A probability model should be used to assess occurrences in the tail region. Depending on the analysis, different probability models may be used. Commonly, for practical applications, the probability models are recommended *a priori* depending on the physical measures being analysed. These recommendations emerge from previous studies regarding the physical quantity being discussed. If a more accurate analysis is needed, different strategies can be used, such as, analysing the fitting. Ideally, if enough data are available, for validation, the extrapolation can be compared with a

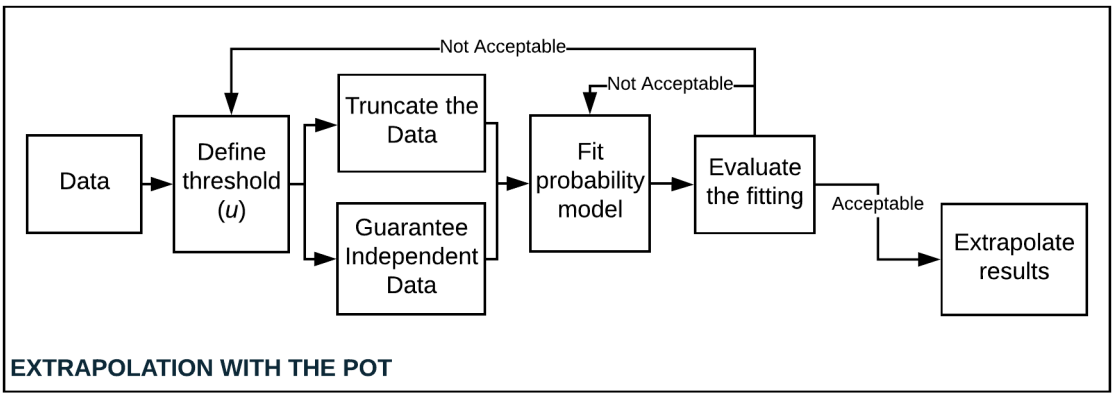


Figure 6.1: POT methodology implementation procedure.

measure of real occurrence.

Several challenges emerge when applying the POT, such as, how to guarantee independence of the data, what probability models to use, and how to compare them.

It was seen that for H_s extrapolation several standard and practices could be found. Some present specific recommendations about the type of techniques to use when modelling extreme H_s. One case of particular interest is the DNV recommended practice on environmental conditions and environmental loads [DNV, 2014b] that used to reject the application of the GP and POT to model *exceedances* of H_s. This recommendation, used as reference for offshore standards [DNV, 2014a], was recently amended, and the application of the GP distribution is now limited to a "careful approach". On the other hand, it was also identified that many experts recommend the usage of the GP distribution to model *exceedance* of H_s data. Moreover, the GP was introduced due to its character of being the limit distribution for *exceedances*.

In detriment of the GP distribution, the 2 parameter Weibull and the Exponential distributions are referred as the reference distributions for the problem of fitting the H_s POT data. While the Weibull is a Generalised Extreme Value distribution in one of its forms for minima, the exponential distribution is one of the forms taken by the GP distribution, namely when ζ is 0.

Since the Exponential distribution is a form taken by the GP distribution under specific conditions, it is very relevant to analyse the fitting of GP for data over a certain u. The GP distribution should present a solution for fitting, at least in the limit case when it tends to the Exponential distribution. Compared with the Exponential, the GP distribution has the additional advantage of using 2 unknown parameters to fit the data. Using two parameters is an indicator that the GP is likely to produce a better fit to the real distribution, than the Exponential, with just one parameter..

The present chapter settles the discussion related to the validity of applying the GP to model *exceedances* of H_s by comparing its application with other widely accepted models such as the 2-parameter Weibull or the Exponential distributions. As highlighted previously, records from four oceanographic buoys located in Irish waters are applied in the assessment.

In order to develop a comprehensive analysis and produce robust comparative results, Section 6.2 presents the data-set used for the comparative analysis. Then, building on

the knowledge of extrapolation presented in Chapter 3, Sections 6.3.1 to 6.3.4 discuss the fundamental concepts of the extrapolation procedure needed to analyse the data presented. Assumptions for independence, selection of threshold and evaluation of the theoretical model approximation are extensively discussed in this section. Section 6.4.3 introduces an innovative approach to calculate u. Section 6.4 presents the comparative results of the application of the 3 statistical models. Robustness of these results is researched in this section and return levels of H_s are presented and compared. Finally, the main conclusion of the work underpinned are compiled in Section 6.5.

6.3 In-field measured wave data

In-field data were collected in four buoys located around the Irish coast by Met Éireann, presented in Figure 6.2. The period of collection comprised an interval of time of 14 years. The exact[1] location of the buoys M1, M4, M5 and M6 are presented in Table 6.1.

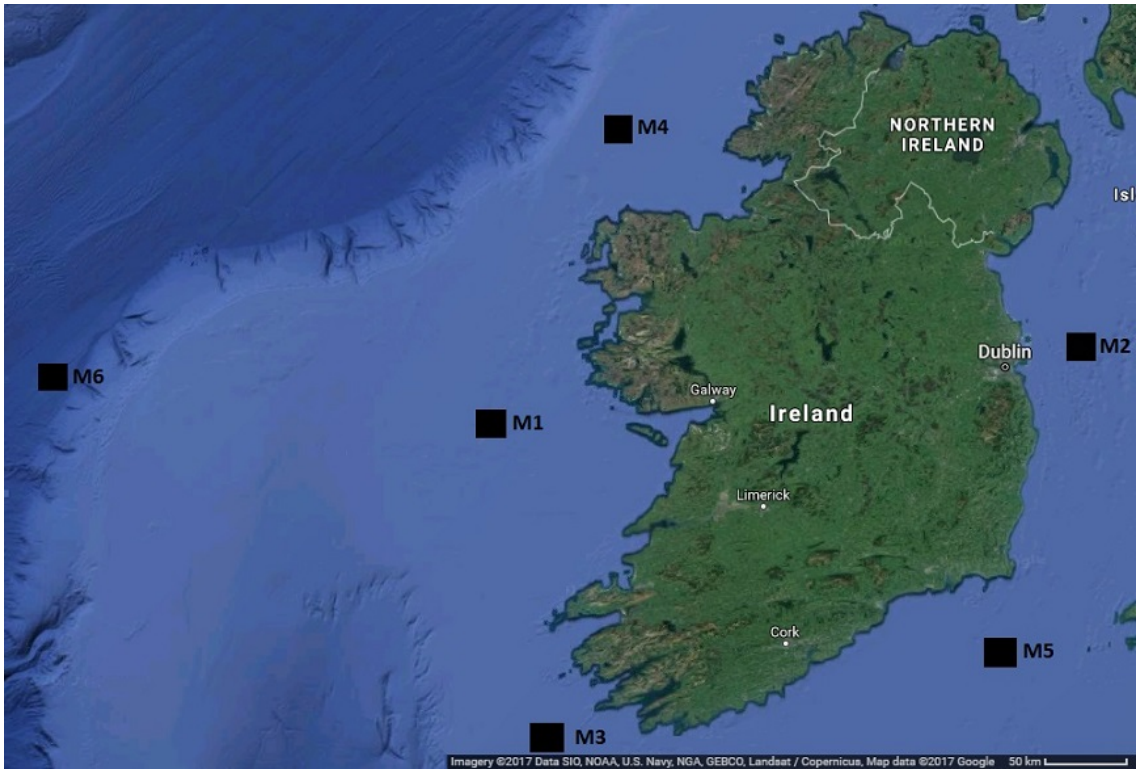


Figure 6.2: Approximate locations of the Met-Éireann oceanographic buoys. Note: Research involved only the analysis of the M1, M4, M5 and M6 buoys. (*Map adapted from Google Maps.*)

M6 is representative of Atlantic deep waters. M1 and M4 is likely to represent a similar wave climate but in shallower waters. M4 is located north of M1 and M6, where the wave climate is expected to be more energetic (relatively higher waves and periods). M5 is representative of the wave climate in the Irish Sea. Due to its oceanographic conditions, the Irish Sea is characterized for having less energetic sea states when compared with the

[1] in Degrees, Minutes and Seconds

remaining locations analysed. The less energetic sea state of M5 can be identified by the smaller H_s values that occur for this buoy, Figure 6.3.

For all of the buoys the most energetic sea state events occurred mainly in two main periods of the year, January-April and October-December.

Table 6.1: Met Éireann buoys location, periods of operation and availability.

Buoy	M1	M4	M5	M6
Degree Minutes	53° 07' 36" N	55° 00' 00" N	51° 41' 24" N	52° 59' 09" N
Seconds(DMS)	11° 12' 00" W	10° 00' 00" W	06° 42' 16" W	15° 52' 00" W
General Location	Off the Galway coast	Off the Donegal coast	Off the Wexford coast	Deep Atlantic coast
Period of Activity	2000-2007	2007-2015	2004-2015	2006-2015
Availability	89%	69%	76%	78%
Sample Size (no. of points)	47065	51343	81667	58187

The operational time and the availability instrumentation equipment are also presented in the Table 6.1. The data records start in the year of 2000 and end in the year of 2015. Availability of data is not 100% due to some periods of interruption in operation for which the buoys could not be fixed, or due to logistic operational problems (e.g. buoys M1 would replace M6 at its location in 2007). Data was recorded in hourly averages and registered at HH[2].

The histograms of occurrences that characterize the H_s data are presented in Figure 6.3.

The statistical analysis of wave data records demands a careful approach due to its physical character and nature.

Wave data is frequently heteroscedastic. To remove heteroscedasticity [Sánchez-Arcilla et al., 2008] and [Vanem et al., 2012] model H_s using a logarithmic transformation. The use of a logarithmic transformation is expected to improve the trend of the analysis for the extreme values. Therefore, it is of interest, when comparing different statistical models to model H_s *exceedance* data, to consider a logarithmic transformation of H_s in the analysis. In this context, the logarithmic transformation is applied as presented in [Vanem et al., 2012], where it is combined with the GP distribution.

Different u values are considered in the analysis. As u increases low H_s occurrences, that may contribute to overestimate the tail region, are eliminated. In its methodology the POT already accounts for the annual description of H_s, which is characterized by its probability of exceeding a return level in a full year.

Figure 6.2 presents independence of data as one of two relevant initial steps when developing a POT analysis. This is due to the fact that H_s *exceedance* data occurs in clusters that are related to storms events. Therefore, if observations are not guaranteed to be independent, long-term occurrences are very likely to be over or underestimated.

Nevertheless, when dealing with *exceedances* in the POT, the first step is to analyse the value of the threshold u that should be used to truncate the data.

[2]Averaged records from 15 minutes before the hour; e.g. 01:00 PM record will correspond to the average of the previous 15 minutes.

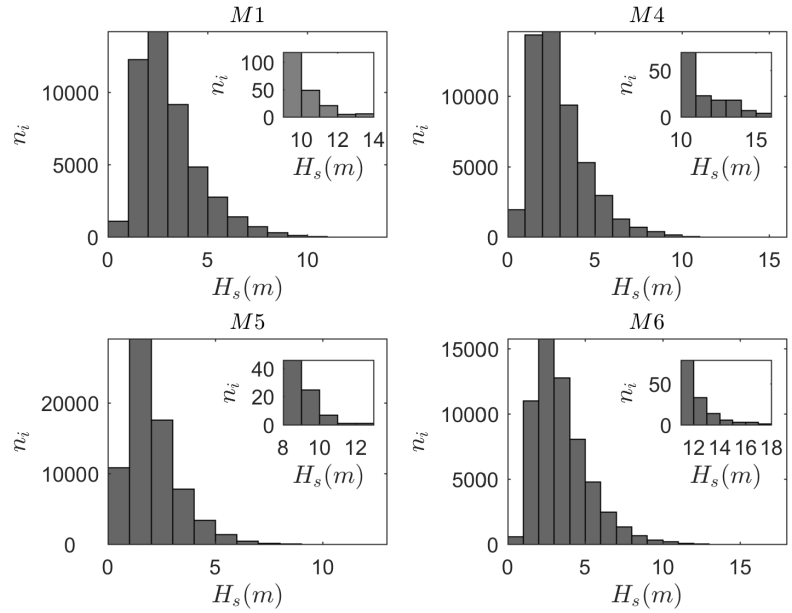


Figure 6.3: Histograms of occurrences of H_s.

6.3.1 H_s independence

In the analysis a time period of 48 hours between two storm events is used to decluster POT data. For the proposed goal of fitting a probability model to POT data, the value of 48 hours ensures that a sufficient margin over the 24 hours is used so that the number of observations above the threshold are independent and the erroneous loss of data is not significant.

To confirm on the adequacy of the 48 hours to guarantee independence of data, the average values of the threshold were compared with the minimum average of 12 hours in between storms, Figure 6.4. When the decrease in average H_s for the period of 12 hours was relevant (approximately 60% of the storm average value during 12 hours), the storm was considered finished.

In-between storms average values of H_s decrease when using 96 hours of declustering time, however, even for 48 hours, a significant gap between the storm and the in-between storms average values can be identified. Nevertheless, usage of 96 hours of declustering results in approximately 30% less storms when comparing with the 48 hours of decluster time. It is common in periods of multiple days in the Northwest Atlantic for the average value of H_s to not decrease below 3 meters even if storm events do not occur.

As the threshold was increased for further studies, independence was ensured even if two peaks occurred separated by more than 48 hours within the same storm using *a priori* knowledge about the storm duration. If less than 48 hours exist between two events, then they are considered the same storm. A sensitivity analysis of the effect of the decluster time is presented later to complement the results on the different probability models.

6.3.2 Choice of threshold u

Chapter 3.3.4 presented different methodologies to calculate the value of u. Common methodologies to define u are the parameter stability plots and MRL plots.

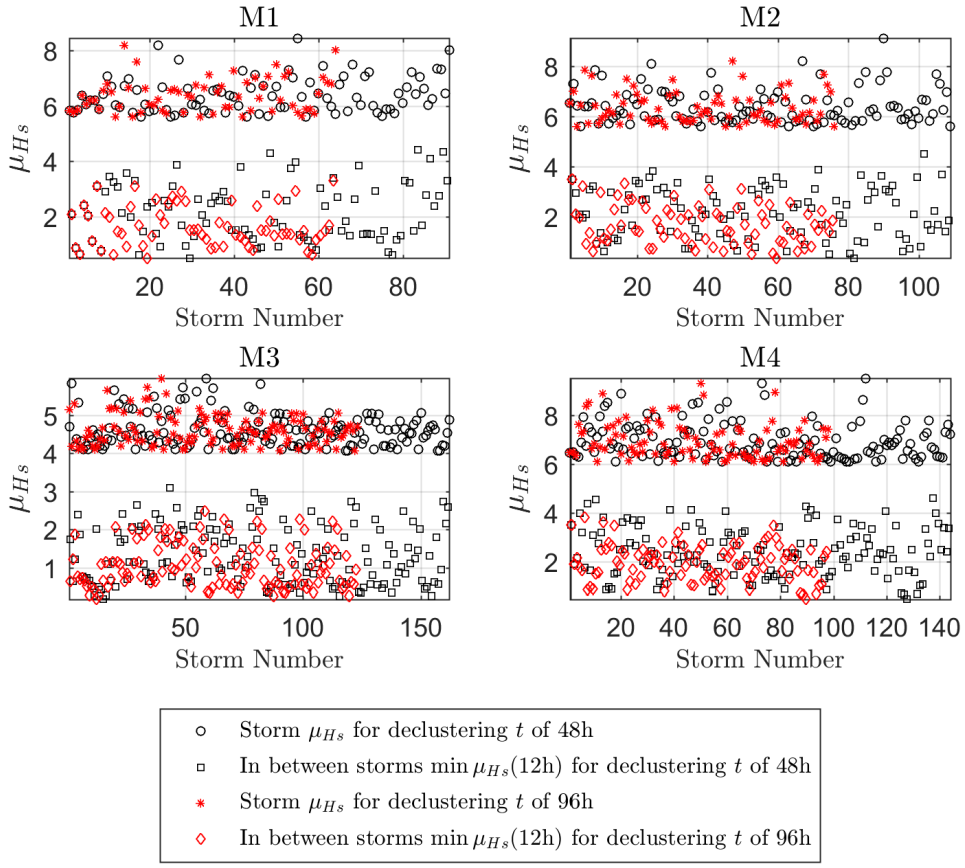


Figure 6.4: Comparison of the declustering time used to guarantee independence between storms. μ_{H_s} refers to the value of H_s in and in-between storms.

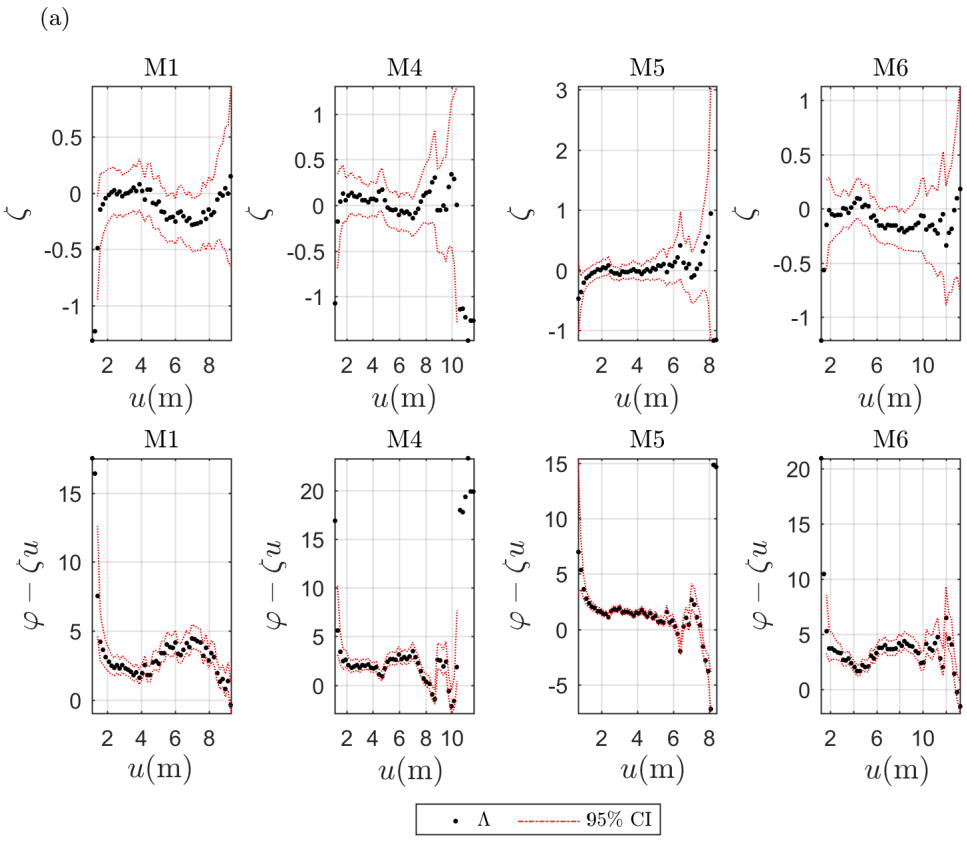
The first is applied to the GP distribution and is used to choose the threshold in combination with a set of fitted parameters that characterizes this distribution model. In this case, the threshold is defined as a function of the probability distribution parameters Λ. The value of u is defined as the threshold for which the Λ values stabilize.

A graphical analysis of the stability of Λ is shown in Figure 6.5(a) with the respective 95% confidence intervals for the estimation. As the value of u increases, the number of data points extracted decreases, the stability plots become then unstable for high values of u. Uncertainty in the estimation of u increases significantly as higher values are evaluated.

A graphical alternative to the stability plots widely used to define u is the MRL plot [Davison and Smith, 1990], Figure 6.5(b). The MRL plot consists in selecting the lowest level of u for which all higher threshold *exceedance* samples mean is consistent with a straight line in a plot [Scarrott and MacDonald, 2012].

Not ignoring the fact that graphical methodologies have a wide application, these can be challenging to interpret. Analysing the graphical representations of Figure 6.5 distinct u levels may be obtained. If for the MRL plot this choice falls into a multiple threshold category with minimum value; for the stability plots u is almost impossible to set without uncertainty relative to the implementation. In both cases the level u, as a graphical analysis, interpretations are likely to be influenced by analytical bias.

In order to avoid any type of bias, a quantitative measure can be used to define u.



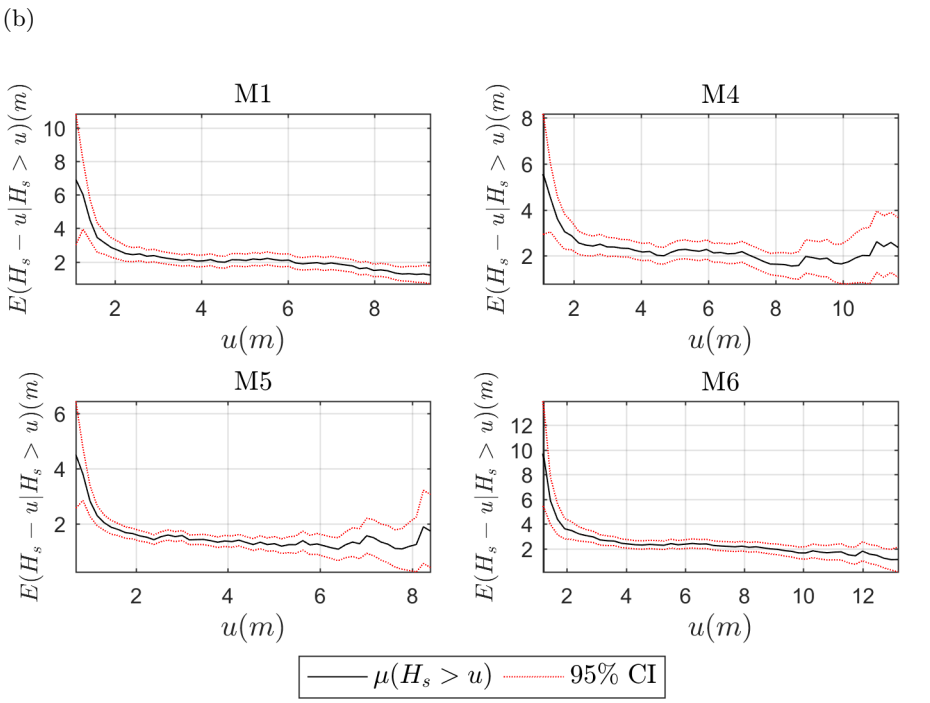


Figure 6.5: M1, M4, M5 and M6 choice of threshold results for the parameter stability of the GP distribution (a) and the MRL plot (b). In the notation Λ in (a) refers to the fitted probability model parameters. $\mu(H_s > u)$ refers to the mean of the *exceedances*. CI is the confidence interval.

A frequent and widely applied rule of thumb to define u is the consideration of occurrences over a certain order statistic, e.g., a 90% quantile of data. It is not uncommon to find applications where a certain quantile is proposed in a POT analysis, such as the recommendation of the 95% quantile by IEC [2005]. Implementing a quantile of 90% in the present example is expected to increase the u relatively to the threshold that may be obtained with the previous methodologies presented. M1, M4, M6 in this case would a have a value of u of approximately 5 and M5 of 3.5 meters. Although of straightforward implementation, using fixed quantile was shown to be theoretically non-adequate by [Scarrott and MacDonald, 2012].

6.3.3 Selection of threshold using the derivatives of the empirical PDF

A new methodology to calculate the value of u is built on the lack of uniformity identified in the current methods to characterize the threshold value. Its benefits are introduced and discussed in this section.

It is assumed that a tail region of a PDF starts in the point where its curvature is maximum positive, conditional on the requirement for the maxima to occur for a value of x higher than the value of x corresponding to the peak of the PDF. As a result, at a certain value of x an inflection in the curvature of the PDF will occur motivated by the approach of the tail region. In addition, as x increases it becomes more distant from the "bulk" of the data. To calculate u, the maximum value of the curvature, after the inflection point, is selected.

The idea that supports such an approach is not only the need to assure a negative slope which is required for the tail, but also to ensure that the tail of the PDF follows the same curvature trend as the theoretical model used to fit the *exceedances*. The calculation of u is performed using the combination of the empirical PDF and a finite differences scheme to calculate higher order derivatives. Results of the implementation for M1, M4, M5, M6 are shown in Figure 6.6. The calculated values of u based on the implementation described are presented in the figure. It can be seen that these values are pointed by the maxima of the second order derivative, that also identifies the beginning of the tail region. It is noted that the u levels calculated show good agreement with the other methodologies. In addition, by directly using the empirical PDF it is ensured an explicit analysis of what is called the tail region. This has the advantage of ensuring that every analysis is case specific. Moreover, a better comprehension and reduction of the risk of erroneously selecting a u level that is inappropriate for the shape of the empirical PDF tail is guaranteed in the process. By comparing with the methods presented in the previous chapter, it is possible to perceive that a significant advantage is also the possibility of obtaining a quantified u value with minimum bias. Implementation of the methodology is simple and uses a search procedure that limits the influence of the designer and his(her) subjective choice.

In order to ensure robustness of the results to be developed, different levels of u are considered in the following analysis. Table 6.2 presents the number of *exceedance* data depending on the level of u considered in the following analysis of Goodness-of-fit (GoF).

Truncation should consider data independence, and therefore should run *in tandem* with it. If independence is not guaranteed, the POT analysis may be compromised in efficiency.

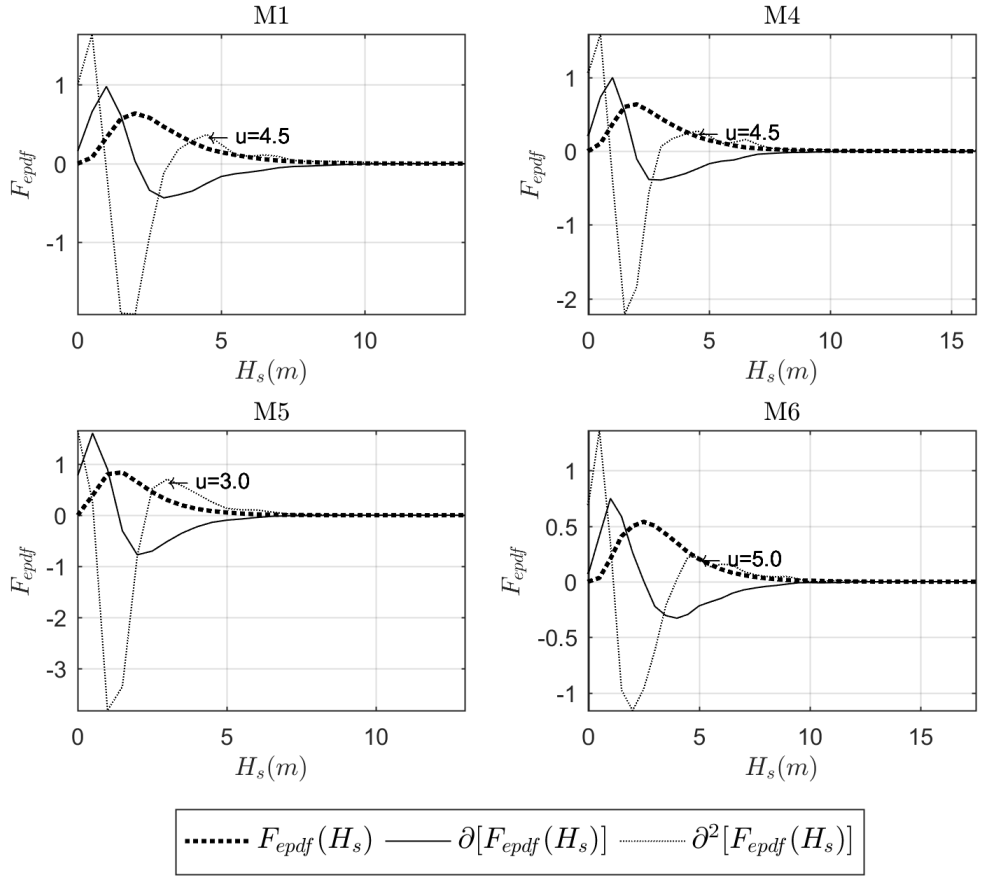


Figure 6.6: Calculation of u for the M1, M4, M5 and M6 buoys using the second derivative of the empirical PDF ($\delta^2[\mathrm{eF_{PDF}(H_s)}]$).

6.3.4 Evaluation of the fitting

The more adequate process to evaluate the fitting of a probability model to a set of data is still a topic under discussion. There is not a clearly identified "best method" to evaluate the goodness of a fitting. Evaluating the adequacy of fitting a probability model usually compares the real distribution against a fitted theoretical distribution. The tests that evaluate the confidence with which a statistical model approximates a set of data are denominated as Goodness-of-fit (GoF) tests. Sheskin [2003] presents a comprehensive discussion on GoF tests to evaluate the adequacy of applying different probability models to characterize a dataset.

A GoF test involves evaluating two hypothesis, H_0 and H_1, that a certain sample of data may come or not from a parent known distribution. Under this assumption the test may be either accepted or rejected with a specified level of confidence p the hypothesis that is being tested. The definition of H_0 and H_1 depends on the test being implemented.

As an alternative other efficient indicators that do not use this hypothesis approach may be used to compare the quality of the fit of a certain probability model to a set of data.

Five methodologies to validate the adequacy of the different statistical models to rep-

Table 6.2: Number of *exceedance* data depending on the level of u considered for truncation.

$u(m)$	3	4.5	5	6	7	8	9
M1	-	124	93	51	35	25	-
M4	-	148	107	62	37	27	-
M5	235	125	67	32	32	-	-
M6	-	-	180	111	71	49	36

resent a set of H_s data are implemented, three GoF tests and two comparative fitting indicators. Since the main objective of the proposed analysis is to infer on the validity of using different probability models to characterize *exceedance* data, this number of comparative measures guarantees robustness of the comparative results.

The most common GoF tests are the Chi-Squared (χ^2) GoF test and the Kolmogorov-Smirnov (KS) GoF test. When dealing with scarce data, such as data in the tail region, the application of the first is limited. In the tail regions of the probability distributions occurrences are scarce and the applicability conditions of the test cannot be entirely fulfilled (minimum 5 occurrences per cell accordingly to the test specifications [Sheskin, 2003], other authors may specify different minimum requirements).

On the other hand, the KS GoF test, is mainly used to compare continuous models (such as the probability models studied). This GoF is based on the comparison between the i and $i-1$ order statistics of the empirical CDF with the respective i statistic given by the estimated theoretical model. For a double-sided test the KS follows (6.1).

$$\begin{gathered} KS_{stat} = \max(D^+, D^-), \\ D^+ = |F_{ecdf}(x_i) - F(x_i, \Lambda)| \quad \text{and} \quad D^- = |F_{ecdf}(x_{i-1}) - F(x_i, \Lambda)|, \end{gathered} \tag{6.1}$$

where $F_{ecdf}(x_i)$ is the empirical cumulative distribution and $F(x_i, \Lambda)$ the statistical cumulative distribution used to describe the empirical data. D^+ and D^- are the test statistic and are calculated from comparison of the empirical and theoretical models in the i and $i-1$ order statistics.

It is known that the KS statistic overestimates the quality of the fitting when the theoretical model parameters are estimated from the set of data that is used to evaluate the fitting. In order to tackle this limitation Stephens [1970] and Lilliefors [1969] present two distinct methodologies. In the case of Lilliefors, the critical values are more likely to reject the probability model fit.

The discrepancy between the theoretical and empirical distributions can be also characterized by the application of quadratic statistics. The Cramér von-Mises (W_c) GoF test is representative of an evaluator of fitting that uses quadratic statistics to compare the empirical and the theoretical distributions.

$$W_c = \int_{-\infty}^{\infty} (F_{ecdf}(x) - F(x, \Lambda))^2 \psi(x) dx, \tag{6.2}$$

$\psi(x)$ is a weight function that has the value of 1 for the case of the the Cramér von-Mises GoF. The weight function is of major importance in the definition of other GoF tests, such

as the Anderson-Darling (AD). The AD is a variant of the Cramér von-Mises GoF test. In the definition of the AD GoF the $\psi(x)$ function is set to be more sensible to variations in the tail of the distribution [Stephens, 1979], taking the following form,

$$\psi(x) = [F(x,\Lambda)(1 - F(x,\Lambda))]^{-1} \tag{6.3}$$

which counteracts the fact that the difference between $F_{ecdf}(x)$ and $F(x,\Lambda)$ is necessarily becoming smaller in the tails as both tails approach 1 at the extremes values.

An extensive discussion of the different these GoF tests, their computational forms and asymptotic values for distinct sample sizes was established in different works. Stephens [1976] discusses the asymptotic results for the exponential distribution with unknown parameters; Richard et al. [1985] evaluate the Weibull distribution with three parameters and present the asymptotic values of its respective test statistic; and Choulakian and Stephens [2001] study the limit test statistic results for the GP distribution with unknown parameters.

If the process of defining GoF tests, with further calculation of their test statistics, are not the most straightforward choice for specific analysis; alternative fitting indicators can be used.

Probability-Probability (P-P) and Quantile-Quantile (Q-Q) plots are frequent evaluators of the fitting between theoretical and empirical statistical results. Castillo and Hadi [1997] use P-P and Q-Q plots to compare fitting results in different data sets, including wave data.

One of the main difficulties of evaluating the fit with Q-Q or P-P plots is that the these are visual methods, the plots are compared based on a graphical analysis. Therefore, in order to ensure a comparative indicator between results that is not subjective as the case of visual evaluation, a quantitative measure may be extracted from the P-P or Q-Q plots. A slope indicator of fit is proposed as a quantitative measure of fit in the P-P and Q-Q plots.

Ideally the slope in a Q-Q or P-P plot should be of 45°. By evaluating the slope, considering that the P-P or Q-Q curve is of type $y = x$, it is possible to quantify how good the fit is resulting from a P-P or Q-Q plot analysis.

$$\alpha_{PP} = \arctan(s) \tag{6.4}$$

with

$$s = \left(\frac{\sum_i F_{ecdf}(x_i) F(x_i,\Lambda)}{\sum_i \left(F_{ecdf}(x_i)\right)^2} \right) \tag{6.5}$$

s is the slope obtained using a minimum least squares formulation assuming that the curve to be fitted needs to be of the type $y = sx$. This slope can then be quantified in an angle using the arctan() function.

Forcing the curve to go through the origin and calculating its deviation when compared to the ideal slope is applied in the current work as a measure of the fit. With the $F(x,\Lambda)$ in the x axis; if the deviation is positive, the fitted model is underestimating $F_{ecdf}(x)$; if the deviation is negative, the fitted model is overestimating $F_{ecdf}(x)$.

As the aim in the current analysis is to produce comparative results α_{PP} will not be further investigated. It is of interest to infer on the sensitivity of this measure as a tool to accept or reject statistical theoretical models.

As an alternative to these plots and to the W_c GoF test in using quadratic differences, the fit measurement is also possible using the the root mean square error (RMSE). The

RMSE is a measure of dispersion of the results that is widely used to measure data fitting. In a comparative scenario it is of interest due to its simple implementation, Equation (6.6). Since a comparative analysis uses relative results, RMSEs are easier to analyse. If the analysis with RMSE is not relative, the measure is not directly transformed in a confidence level.

$$RMSE = \frac{1}{N}\sqrt{\sum_{i}\left(F_{ecdf}(x_i) - F(x_i,\Lambda)\right)^2} \tag{6.6}$$

It is important to highlight that the last two indicators of the quality of the fit discussed are not under a hypothesis testing.

6.4 Long-term wave heights

Results for the comparative evaluation of the long-term H_s using different probability models are now discussed.

6.4.1 Goodness-of-fit results

The minimum value of u for analysis of the statistical fitting was set to be variable depending on the analysis of Section 6.3.3. The influence of the selection of different thresholds is also discussed in terms of the robustness of the fitting. Table 6.3 and Table 6.4 present the results for the statistical fitting parameters for each case and the results for the three GoF and the two fitting indicators considered.

Table 6.3: Model parameters obtained from the fitting process for the different distributions and values of u.

$u(m)$	**Weibull**		**Exponential**	**GP**		$\log(x_{H_s})$ **GP**	
	φ (scale)	ζ(shape)	φ	ζ(shape)	φ(scale)	ζ(shape)	φ(scale)
			Buoy M1				
4.5	2.009	0.987	2.021	-0.009	2.039	-0.322	0.442
5	2.149	1.04	2.115	-0.096	2.32	-0.381	0.446
5.5	2.273	1.077	2.209	-0.192	2.646	-0.453	0.455
6	2.599	1.277	2.411	-0.353	3.28	-0.597	0.508
6.5	2.369	1.179	2.247	-0.336	3.024	-0.566	0.442
7	2.498	1.437	2.261	-0.423	3.242	-0.651	0.445
7.5	2.003	1.191	1.883	-0.256	2.385	-0.463	0.315
8	2.021	1.245	1.878	-0.352	2.585	-0.507	0.326
			Buoy M4				
4.5	1.963	0.93	2.033	0.13	1.774	-0.19	0.39
5	2.193	0.975	2.218	0.043	2.122	-0.257	0.412
5.5	2.202	0.972	2.23	0.05	2.12	-0.245	0.38
6	2.519	1.072	2.449	-0.072	2.627	-0.349	0.422
6.5	2.398	0.988	2.41	-0.063	2.563	-0.336	0.387

Continued on next page

Continued from previous page

7	2.918	1.242	2.717	-0.256	3.426	-0.527	0.472
7.5	2.379	1.016	2.363	-0.106	2.619	-0.361	0.348
8	2.583	1.109	2.492	-0.236	3.103	-0.503	0.39
			Buoy M5				
3	1.683	1.188	1.581	-0.082	1.709	-0.314	0.496
3.5	1.61	1.161	1.523	-0.065	1.621	-0.28	0.417
4	1.547	1.1	1.492	-0.054	1.571	-0.255	0.362
4.5	1.465	1.129	1.397	-0.019	1.423	-0.206	0.299
5	1.419	1.117	1.356	0.014	1.336	-0.166	0.257
5.5	1.437	1.051	1.407	0.027	1.369	-0.154	0.242
6	1.323	1.018	1.311	0.115	1.161	-0.059	0.192
6.5	1.456	0.985	1.466	0.123	1.288	-0.07	0.199
7	1.856	1.107	1.782	-0.069	1.907	-0.314	0.28
7.5	1.375	0.908	1.45	0.331	1.011	0.138	0.14
			Buoy M6				
5	2.297	0.976	2.322	0.019	2.277	-0.295	0.442
5.5	2.504	1.038	2.466	-0.072	2.645	-0.35	0.453
6	2.457	0.984	2.474	-0.094	2.71	-0.356	0.427
6.5	2.784	1.138	2.66	-0.211	3.229	-0.428	0.45
7	2.777	1.178	2.624	-0.23	3.232	-0.432	0.421
7.5	2.7	1.181	2.549	-0.236	3.155	-0.426	0.387
8	2.692	1.189	2.542	-0.268	3.223	-0.445	0.372
9	2.433	1.185	2.297	-0.254	2.885	-0.412	0.303

Table 6.4: GoF results for the three test statistic K-S, W_c and A-D and the two GoF indicators studied for the Weibull, Exponential and GP Distribution. The p-values in % are given for the test statistics performed. Underlined values correspond to the appropriate level u as specified in Section 6.3.3. Results for the log transformation are shown for the GP distribution between brackets.

$u(m)$	3	3.5	4	4.5	5	5.5	6	6.5	7	7.5	8	9
					Weibull Distribution							
					Buoy M1							
KS(p%)	-	-	-	67.8	58.0	55.5	72.8	57.8	85.3	98.4	90.7	-
W_c(p%)	-	-	-	58.2	46.3	46.2	65.7	62.4	57.3	65.2	49.1	-
AD(p%)	-	-	-	11.1	13.9	14.7	44.0	21.0	38.9	47.6	40.1	-
α_{PP}(°)	-	-	-	44.99	45.01	45.13	45.06	45.2	44.64	44.93	44.91	-
RMSE	-	-	-	0.027	0.034	0.038	0.031	0.038	0.037	0.032	0.047	-
					Buoy M4							
KS(p%)	-	-	-	79.8	94.8	90.2	75.0	87.2	96.5	91.8	87.3	-
W_c(p%)	-	-	-	56.7	62.1	54.1	49.7	54.9	59.7	62.0	55.9	-
A-D(p%)	-	-	-	13.0	35.7	28.5	28.7	32.1	46.5	39.7	33.4	-
α_{PP} (°)	-	-	-	44.83	44.86	44.91	44.88	45.2	44.64	45.08	44.82	-
RMSE	-	-	-	0.022	0.022	0.027	0.032	0.033	0.030	0.035	0.039	-
					Buoy M5							
KS(p%)	17.3	13.7	56.4	46.3	27.2	75.4	77.7	72.4	61.5	66.4	-	-

Continued on next page

Continued from previous page

$W_c(p\%)$	62.5	60.7	66.1	58.7	34.8	37.3	22.0	35.7	14.2	11.4	-	-
AD($p\%$)	< 1.0	1.6	4.7	12.5	< 1.0	9.2	8.1	18.4	10.6	9.3	-	-
$\alpha_{PP}(^\circ)$	44.69	44.66	44.69	44.38	44.11	44.15	43.68	43.82	43.77	43.72	-	-
RMSE	0.024	0.028	0.028	0.029	0.046	0.045	0.057	0.064	0.010	0.104	-	-
					Buoy M6							
KS($p\%$)	-	-	-	-	56.9	46.1	70.9	96.4	92.3	87.9	85.0	77.4
$W_c(p\%)$	-	-	-	-	52.2	44.8	55.0	76.7	68.3	59.5	71.3	61.2
AD($p\%$)	-	-	-	-	11.6	10.5	10.7	61.6	51.9	37.2	50.5	45.2
α_{PP} ($^\circ$)	-	-	-	-	45.02	45.03	45.33	45.15	45.1	45.12	45.26	45.26
RMSE	-	-	-	-	0.023	0.027	0.025	0.017	0.023	0.029	0.026	0.035
					Exponential Distribution							
					Buoy M1							
KS($p\%$)	-	-	-	52.3	48.3	29.6	7.1	7.4	2.9	86.2	70.1	-
$W_c(p\%)$	-	-	-	65.6	70.9	70.8	47.9	53.6	35.0	82	68.9	-
AD($p\%$)	-	-	-	26.1	40.3	39.5	15.2	22.3	3.3	48.9	42.3	-
α_{PP} ($^\circ$)	-	-	-	44.97	45.07	45.25	45.41	45.43	45.29	45.34	45.49	-
RMSE	-	-	-	0.028	0.029	0.034	0.056	0.058	0.081	0.035	0.046	-
					Buoy M4							
KS($p\%$)	-	-	-	21.7	81.4	71.2	86.9	81.4	32.6	80.8	51.6	-
$W_c(p\%)$	-	-	-	34.3	65.3	56.8	70.9	66.2	40.7	72.2	57.3	-
AD($p\%$)	-	-	-	6.4	45.3	37.3	50.4	47.1	26.1	56.1	39.5	-
α_{PP} ($^\circ$)	-	-	-	44.61	44.79	44.82	45.08	45.17	45.32	45.14	45.17	-
RMSE	-	-	-	0.036	0.025	0.032	0.027	0.034	0.053	0.035	0.050	-
					Buoy M5							
KS($p\%$)	< 1.0	6.5	6.2	28.6	12.2	50.2	63.1	66.5	33.1	40.1	-	-
$W_c(p\%)$	73.2	72.5	64.3	70.1	58.6	53.0	43.6	49.7	40.7	18.9	-	-
AD($p\%$)	< 1.0	1.4	7.9	10	7.1	25.4	26.9	35.9	30.3	18.0	-	-
α_{PP} ($^\circ$)	44.82	44.81	44.8	44.6	44.37	44.31	43.76	43.74	44.29	43.1	-	-
RMSE	0.037	0.036	0.040	0.039	0.049	0.049	0.055	0.065	0.096	0.121	-	-
					Buoy M6							
KS($p\%$)	-	-	-	-	29.5	56.7	53.7	79.7	76.4	69.2	55.6	55.7
$W_c(p\%)$	-	-	-	-	54.9	71.5	64.8	67.1	61.9	62.1	53.2	69.0
AD($p\%$)	-	-	-	-	20.4	37.7	25.5	44.3	36.9	37.7	39.0	47.1
α_{PP} ($^\circ$)	-	-	-	-	44.98	45.12	45.29	45.47	45.51	45.53	45.64	45.53
RMSE	-	-	-	-	0.027	0.022	0.027	0.027	0.031	0.032	0.041	0.040
					Generalised Pareto Distribution							
					Buoy M1							
KS($p\%$)	-	-	-	50.6 (53.9)	36.1 (40.3)	53.0 (60.5)	68.2 (62.7)	60.9 (56.6)	28.8 (25.8)	88.3 (81.4)	70.3 (62.0)	-
$W_c(p\%)$	-	-	-	54.5 (51.3)	43.8 (45.2)	45.0 (49.0)	67.8 (60.3)	61.9 (55.5)	37.1 (25.7)	61.9 (55.0)	41.6 (30.0)	-
AD($p\%$)	-	-	-	13.7 (12.1)	18.3 (20.1)	22.5 (26.1)	51.2 (39.6)	38.7 (32.4)	9.5 (2.8)	46.6 (39.9)	34.0 (23.9)	-
α_{PP} ($^\circ$)	-	-	-	44.95 (44.83)	44.86 (44.75)	44.82 (44.7)	44.61 (44.46)	44.63 (44.45)	44.21 (43.95)	44.62 (44.42)	44.42 (44.09)	-
RMSE	-	-	-	0.028 (0.029)	0.035 (0.033)	0.037 (0.033)	0.028 (0.028)	0.034 (0.035)	0.053 (0.051)	0.035 (0.039)	0.051 (0.057)	-
					Buoy M4							
KS($p\%$)	-	-	-	67.0 (67.8)	93.6 (90.3)	84.3 (82.7)	77 (81.0)	71.7 (73.2)	63.0 (60.1)	88.5 (88.4)	81.3 (83.6)	-
$W_c(p\%)$	-	-	-	64.0 (60.5)	66.9 (65.0)	58.2 (55.4)	58.1 (55.8)	54.3 (51.8)	42.7 (29.8)	60.2 (53.2)	39.5 (24.8)	-
A-D($p\%$)	-	-	-	28.4 (23.7)	48.2 (45.2)	40.5 (37.0)	40.9 (36.8)	33.3 (30.6)	34.2 (22.9)	44.1 (37.7)	31.7 (21.0)	-
α_{PP} ($^\circ$)	-	-	-	45.01 (44.91)	44.92 (44.83)	44.98 (44.86)	44.85 (44.73)	44.96 (44.81)	44.43 (44.19)	44.76 (44.55)	44.3 (43.96)	-
RMSE	-	-	-	0.021 (0.021)	0.021 (0.021)	0.027 (0.027)	0.029 (0.028)	0.035 (0.034)	0.038 (0.044)	0.035 (0.037)	0.047 (0.054)	-
					Buoy M5							
KS($p\%$)	1.7 (< 1.0)	12.6 (4.2)	12.5 (5.0)	31.7 (22.8)	11 (8.8)	43.9 (40.2)	45.6 (44.7)	50.8 (51.8)	40.3 (44.7)	76.2 (75.2)	-	-
$W_c(p\%)$	74.0 (52.8)	70.2 (56.7)	61.1 (47.5)	62.7 (55.1)	49.0 (44.4)	42.2 (37.2)	43.8 (41.0)	56.3 (52.7)	23.4 (15.7)	41.5 (35.7)	-	-

Continued on next page

Continued from previous page

AD(p%)	27.8 (< 1.0)	3.6 (58.3)	3.8 (1.7)	2.2 (< 1.0)	< 1.0 (2.3)	14.4 (9.7)	20.9 (17.4)	35.7 (32.2)	19.1 (14.0)	35.5 (30.6)	-	-
α_{PP} (°)	44.70 (44.69)	44.69 (44.70)	44.69 (44.69)	44.55 (44.55)	44.42 (44.42)	44.41 (44.42)	44.23 (44.22)	44.29 (44.24)	44.01 (43.76)	44.47 (44.39)	-	-
RMSE	0.029 (0.027)	0.030 (0.028)	0.034 (0.033)	0.037 (0.033)	0.049 (0.042)	0.051 (0.052)	0.055 (0.055)	0.055 (0.056)	0.098 (0.097)	0.086 (0.090)	-	-
					Buoy M6							
KS(p%)	-	-	-	-	38.5 (38.5)	33.0 (47.1)	39.8 (47.4)	98.3 (95.6)	97.9 (98.3)	96.2 (97.5)	95.9 (96.3)	86.8 (85.8)
W_c(p%)	-	-	-	-	50.1 (49.7)	47.4 (56.4)	43.8 (51.5)	78.5 (81.3)	75.0 (76.9)	70.6 (73.6)	77.9 (75.4)	70.1 (71.9)
AD(p%)	-	-	-	-	15.2 (15.3)	16.8 (22.9)	3.9 (8.7)	73.0 (70.1)	64.2 (59.7)	55.3 (54.3)	71.2 (67.1)	62.2 (61.5)
α_{PP} (°)	-	-	-	-	45.03 (44.91)	44.92 (44.87)	45.02 (44.98)	44.88 (44.96)	44.88 (44.95)	44.91 (44.96)	44.98 (45.02)	44.96 (44.94)
RMSE	-	-	-	-	0.025 (0.024)	0.028 (0.023)	0.031 (0.026)	0.016 (0.014)	0.019 (0.017)	0.022 (0.020)	0.020 (0.020)	0.028 (0.026)

The KS GoF gives a measure of the biggest deviation in the fit, being very useful to detect outliers. Considering that it aims to identify the biggest difference between the real and the estimated curves, in the present case its application is well complemented by the tests that work with mean values and weighted averages of the error, which is the case of the remaining tests applied. An increased value of the test statistic is representative of a smaller $p-value$, and consequently a worse fit to the data.

The KS GoF for a single sample was initially developed for the case of a completely specified Λ. The need for a correction of the KS statistic is justified by the fact that since the Λ vector was obtained from the sample it is very likely for the estimated model to approximate better the real distribution when compared with the case where Λ is not estimated from the data. The Lilliefors test $p-values$ are then expected to decrease for the same asymptotic values of the test statistic. For the current implementation it is expected that the results of the KS will hold relevant comparative results of GoF. Although it is important to understand that for practical applications of not rejecting any of the statistical models as valid for representation of the empirical distribution, the GoF results need to consider the estimation of Λ.

The $p-value$ was also used as a reference to compare the GoF results for the W_c and AD statistics. Extensive tables that evaluate the W_c and the AD GoF are available in the literature.

The calculation of the $p-value$ for each test statistics involves interpolating and extrapolating techniques. A 3^{rd} degree polynomial function was fitted to the asymptotic points to calculate the $p-values$, as their variation with the test statistic is clearly non-linear. The variation of the confidence of the fit with the test statistic for the considered distributions can be found in Stephens [1976, 1977, 1979] and Spineili and Stephens [1987].

The 2-parameter Weibull distribution, when the location parameter is known, needs to be tested against a type I GEV distribution [Lockhart and Stephens, 1993]. A transformation of variable needs then to be applied to the *exceedance* values to test against a type I GEV.

It is interesting to notice that for the buoy M4 in the reference threshold the GP distribution produces better fitting results in four of the five indicators used for comparison. Nevertheless, the GP distribution is unbounded for this case as $\zeta > 0$. For M5, the same happens in two of the five indicators. In the case of buoys M1 and M6, the two-parameter Weibull and the Exponential distributions achieved, respectively, better fitting results in three and two of the five indicators used for comparison respectively.

One of the concerns highlighted when using the GP to model physically limited events, is to ensure that the distribution is bounded. As mentioned, this happens when the shape (ζ) parameter is smaller than 0. In the present case, no significant improvement of the fitting indicators occurs when a logarithmic transformation is applied (case $\log(x_{H_s})$ with H_s). Still, one of the interesting points of using a logarithmic transformation is that it ensures the ζ parameter to be negative, thus bounded for the Generalised Pareto. As example, GP fits for non transformed data with unbounded domain occur for values of u close to the reference value in the case of the buoys M4 and M6. This should be avoided on a design basis approach.

It is important to notice that the Weibull and Exponential distributions have support in the interval $[0,+\infty)$. Therefore, due to the physical character of the waves, even considering the following comparison of GoF the GP should be applied primarily when its right tail is bounded.

In regard of the other threshold values considered, cases where the GP distribution present better indicators than the other two statistical distributions can be found, e.g. all the cases above the threshold of 6.5m meters in the buoy M6, where the GP distribution is bounded.

Nonetheless, relevant levels of u where the GP has no better value of fitting in the five indicators used can be identified; instances are M1 with $u = 4.5$m or M6 $u = 5$m, among other examples.

Under the discussed fluctuation in the results, it is of additional interest to understand how the results vary with different decoupling times in the application of the POT. In that way, potential limitations of using the GP distribution as a probability model for the *exceedances* of H_s can be identified.

6.4.2 Influence of the decoupling time

The analysis of the decoupling process that guarantees independent data is of major importance to ensure the validity of the study and the robustness of the results obtained. It was seen that the methodology to ensure independence involved the application of a minimum period of time between storm events.

One GoF test and a fitting indicator were considered here for studying the influence of the decoupling time in the GoF results, i.e., Kolmogorov-Smirnov and the RMSE. Figures 6.7 and 6.8 show, respectively, the results for the variation of the GoF with the variation of the decoupling time.

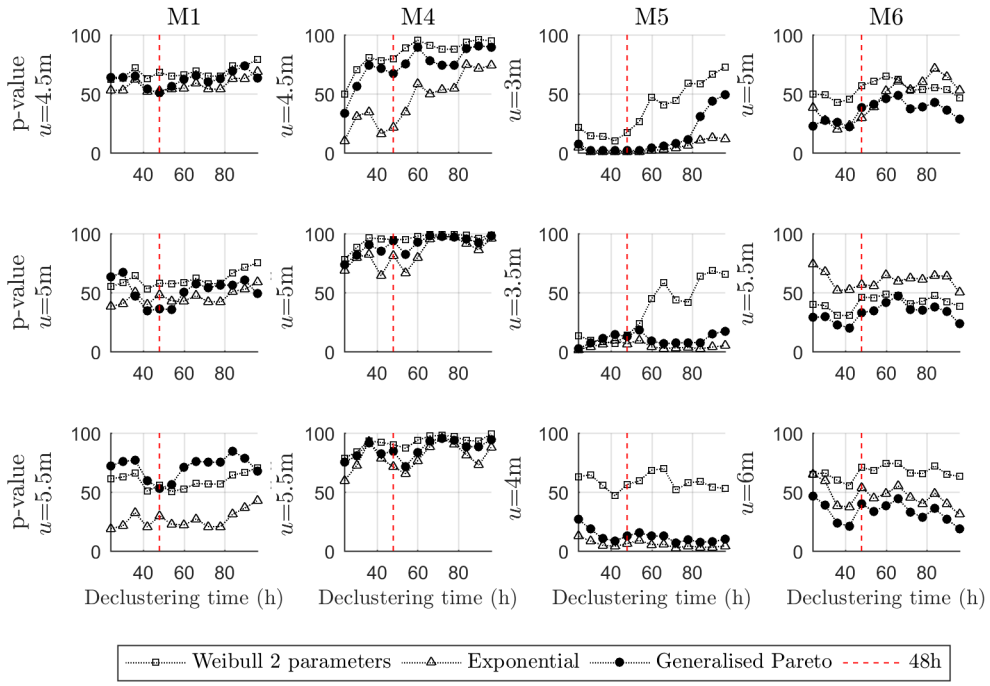


Figure 6.7: Influence of the decoupling time in the KS results. The trimmed vertical line represents the reference value of 48 hours.

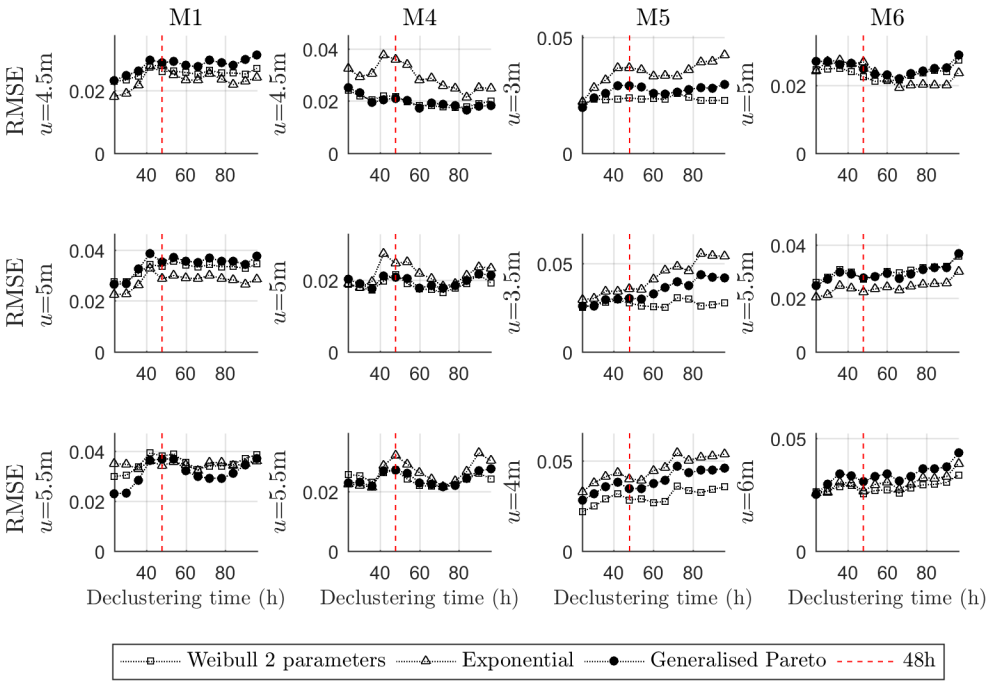


Figure 6.8: Influence of the decoupling time in the $RMSE$ results. The trimmed vertical line represents the reference value of 48 hours.

The choice of these two test statistics is justified by the two quantities evaluated in the tests, while the KS test evaluates the maximum deviation from the real function and calculates the confidence of the fit based on that value, the RMSE considers an averaged value of deviation from the real distribution, taking into account all the points of the CDF.

The results show similar trends for all the probability distributions. In M1 the $p-values$ are stable to the point that the acceptance of the fitting is never doubted (never inferior to 25% of p value). In M4, the increase of the decoupling time increases the quality of the fit in the KS test. The GP has lower averaged deviation from the empirical distribution. In M5, the results indicate that probably a more detailed analysis of choice of the u may be needed as the quality of the fit in average seems to improve substantially with the decoupling time. This buoy is particularly distinct from the other buoys due to its location. It has a low number of points with H_s bigger than 8 meters and high number of points with H_s between 3 and 5 meters. For M5, as the decoupling time increases the GP averaged deviation decreases. For the decoupling time used, the Weibull distribution presents better average deviation results. Finally, in M6 the quality of the fit is more variable with the decoupling time but not enough to question the acceptance of any of the distributions studied. The RMSE results for M6 do not change significantly for different decoupling times.

The quality of the fit fluctuates and all of the three distributions compared may produce a better fit than the others under specified circumstances.

The analysis of the decoupling time is a methodology that to some extent trains the statistical model. Even though, on a recurrent design basis the training methodology, as

presented in Section 3.3.5, may already complement the analysis of fitting robustness. This can contribute to more robust results and confidence in the fitting procedure underpinned.

6.4.3 Limited number of data points

As the dataset was extracted using the POT methodology, the same relatively high number of truncated data points was used to fit the three statistical distributions analysed. It is of interest, to draw robust conclusions, to also analyse the trends in the results when limited data points are used.

A progressive cyclic methodology consisting in randomly removing one point from the initially truncated data-set (N_0) with consequent re-fitting was applied.

A set of 10 data points was considered as the minimum sample size for fitting. The results are represented as a function of the relative number of points from the initial data truncated by the POT methodology, Figures 6.9 and 6.10.

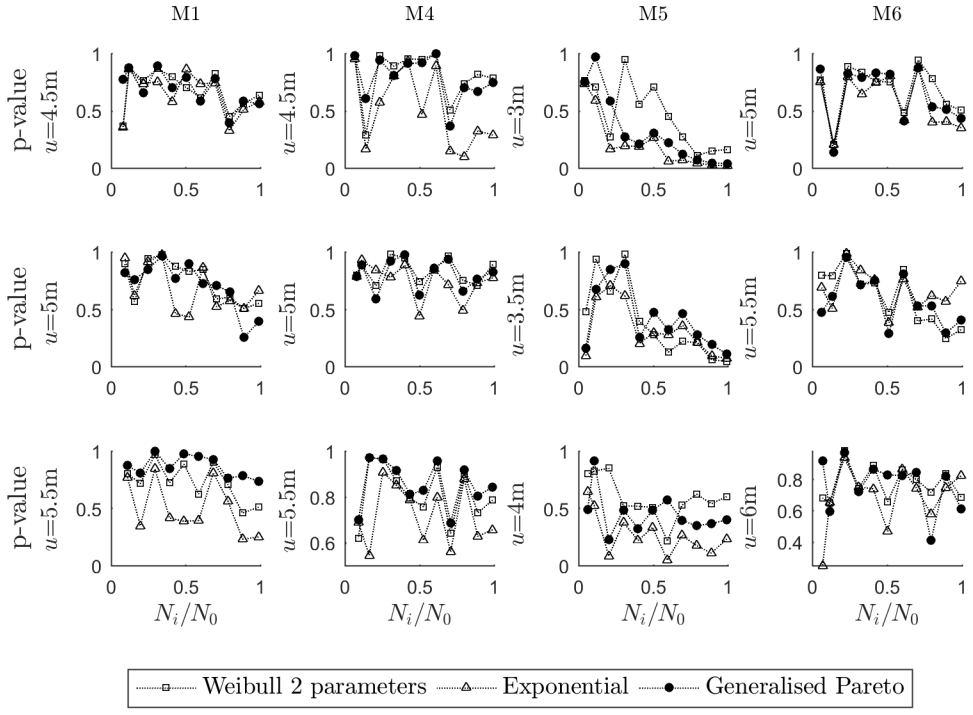


Figure 6.9: Influence of decreasing the sample in the KS results. N_i/N_0 is the ratio of remaining points in comparison to the initial truncated sample.

Using a limited number of data points allows us to investigate if any of the distributions could perform inconsistently when few measurements are available to perform the fitting. The KS and the RMSE were again used to analyse the variation of the GoF when a limited number of data points are available. The selection of the GoF tests followed the criteria mentioned for the selection of the decoupling time. In all the cases the MLE was used to estimate the parameters of the statistical models.

As expected, the results show some variation of the quality of the fit caused by the extraction of points. In particular, the exponential distribution is more sensitive to the variation of the number of points in both of the indicators used for comparison. Examples

of a more erratic behaviour by the exponential distribution can be found in the cases of M4 with $u = 4.5$m or M5 with $u = 4$m where some singular peaks can be identified in the results.

The slight upward trend identified in M5 is again very likely to result from the previously mentioned characteristics of its data. The KS GoF test is highly sensitive to areas of the fit which have very low density of points. As the number of points is randomly decreased, large H_s may be removed, resulting in an increases of the KS $p - value$.

As in the previous section, the results do not show any recurrent trend that supports the hypothesis that the application of the GP should be to some extent rejected to model extreme wave heights.

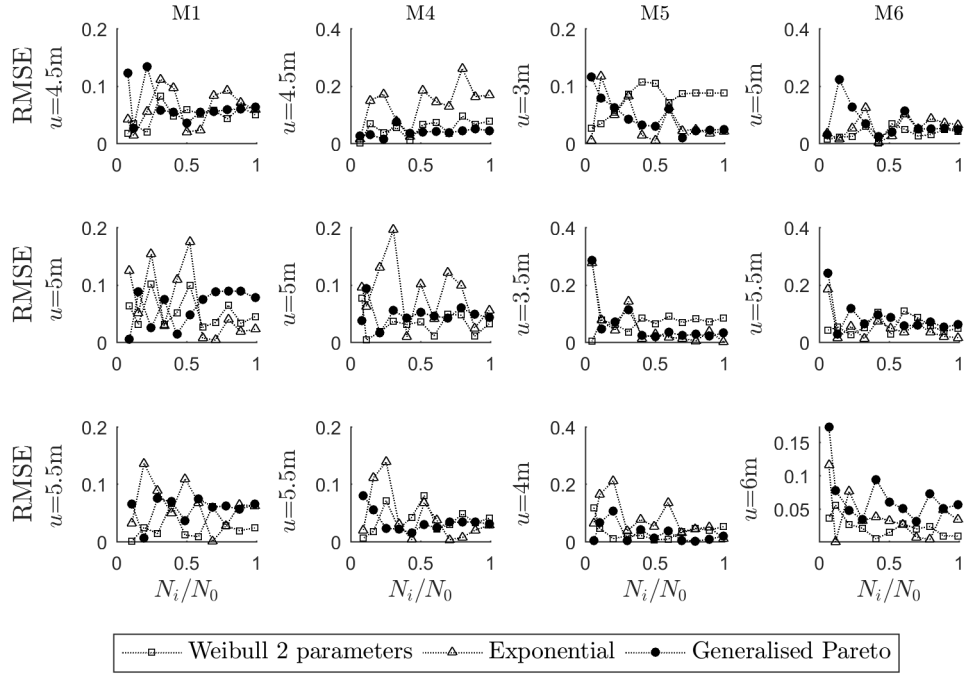


Figure 6.10: Influence of decreasing the sample size in the RMSE results. N_i/N_0 is the ratio of remaining points.

6.4.4 Return periods of H_s

Frequently the necessity to fit a probability model to a known set of data is created by the need to extrapolate the occurrence of a certain event to a pre-specified return period T_r. This is the case for offshore wind turbines, where some of the standards that regulate the design of these equipments e.g. [IEC, 2009], demand the extrapolation of the environmental conditions to return periods of between 20 and 50 years.

Extrapolating long-term occurrences is achieved by knowing the yearly average number of events λ. With the knowledge of λ it is possible to estimate the probability of exceeding in average a certain H_s level in 20 or 50 years of occurrences. No higher return levels were considered due to the amount of data available in the records. As a rule of thumb in the industry applications, it is commonly considered that the minimum amount of the data to extrapolate a certain T_r corresponds to a period of time of at least its value divided by four. The 50 years period, which is out of this limit, is considered due to its importance

in most of the offshore sectors.

Different values of H_s with return level T_r depending on u are shown in Table 6.5. The reference u level for each buoy is highlighted in the table.

Table 6.5: $20yr$ and $50yr$ return levels H_{s_r} of H_s as a function of u and the statistical model applied. (Shadowed values correspond to the appropriate level u as specified in Section 6.3.3. Changes (in %) for the Weibull and Exponential are computed in relation to the non-logarithmic GP results. Logarithmic GP results are complementary to the comparative analysis and also uses a GP model.)

	Weibull		**Exponential**		**GenPareto**			
					H_s		$\log(x_{H_s})$	
u(m)	$T_{r=20}$	$T_{r=50}$	$T_{r=20}$	$T_{r=50}$	$T_{r=20}$	$T_{r=50}$	$T_{r=20}$	$T_{r=50}$
Buoy M1								
4.5	17.12 (+2.7%)	19.03 (+3.3%)	16.90 (+1.4%)	18.75 (+1.7%)	16.67	18.43	14.68	15.42
5	17.03 (+9.3%)	18.79 (+12.3%)	17.68 (+13.5%)	19.62 (+17.27%)	15.58	16.73	14.29	14.8
5.5	17.17 (+16.2%)	18.87 (+21.7%)	18.37 (+24.4%)	20.39 (+31.5%)	14.77	15.50	13.99	14.32
6	16.18 (+21.2%)	17.44 (+15.2%)	19.79 (+40.8%)	22.0 (+52.9%)	14.05	14.39	13.65	13.85
Buoy M4								
4.5	18.51 (-14.0%)	20.74 (-18.3%)	17.16 (-20.3%)	19.02 (-25.1%)	21.52	25.40	18.72	20.7
5	18.87 (-4.0%)	21.03 (-5.4%)	18.39 (-6.4%)	20.43 (-8.1%)	19.65	22.22	17.71	19.02
5.5	19.21 (-4.0%)	21.40 (-6.0%)	18.70 (-7.0%)	20.74 (-8.9%)	20.10	22.77	18.06	19.44
6	18.96 (+2.9%)	20.86 (+4.5%)	20.18 (+9.5%)	22.43 (+12.4%)	18.43	19.96	17.11	17.88
Buoy M5								
3	10.84 (-4.3%)	11.80 (-3.5%)	12.84 (+13.3%)	14.29 (+16.8%)	11.33	12.23	11.63	12.30
3.5	11.13 (-4.5%)	12.11 (-4.0%)	12.77 (+13.3%)	14.17 (+16.8%)	11.65	12.62	11.82	12.57
4	11.71 (-1.5%)	12.8 (-0.9%)	12.73 (+7.1%)	14.10 (+9.1%)	11.89	12.92	12.01	12.83
4.5	11.32 (-7.0%)	12.29 (-7.8%)	12.44 (+2.2%)	13.72 (+2.9%)	12.44	13.33	12.21	13.21
5	11.36 (-8.5%)	12.33 (-10.3%)	12.25 (-1.4%)	13.49 (-1.9%)	12.42	13.75	12.45	13.62
Buoy M6								
5	20.13 (-0.6%)	22.39 (-1.1%)	19.62 (-3.1%)	21.74 (+3.9%)	20.25	22.63	17.68	18.69
5.5	19.93 (+6.8%)	21.99 (+9.0%)	20.69 (+10.9%)	22.94 (+13.7%)	18.66	20.17	17.25	17.98
6	21.37 (+15.2%)	23.73 (+19.3%)	21.02 (+13.4%)	23.29 (+17.1%)	18.54	19.89	17.25	17.98
6.5	19.74 (+15.6%)	21.54 (+18.7%)	22.19 (+27.7%)	24.63 (+35.7%)	17.38	18.15	17.34	18.02

The relative difference to the GP extrapolated non-transformed data value is shown between brackets for the respective 20 and 50 year extrapolations. As expected, in the majority of the cases the return periods do not vary substantially.

Results show that the logarithmic transformation tends to stabilize the extrapolation with the GP (less variability in H_{s_r} and occurrence of large extrapolated values). This may be connected to the shape parameter that, for the logarithmic transformation, Table 6.3, sets the GP to be bounded.

In buoy M1, despite the fact that the GP distribution has no better fitting results for the reference level of u, the values extrapolated are similar for the non-logarithmic results. However, when the value of the threshold increases, the two-parameter Weibull distribution and the Generalised Pareto distribution show a trend to decrease the values extrapolated for the return levels of H_s. This decrease was discussed in Gōda [2010] as one of the points of criticism of the POT. As the threshold increases, the standard

variation of the sample decreases. The fact that these two distributions have one more DOF than the Exponential distribution might be connected to this decreasing trend. This implies an additional capacity to adapt to the dataset to be fitted and may justify further research on the homogeneity of data when comparing with the logarithm transformed results. Considering that the maximum value of H_s in the data set is 13.5m it is very likely that the logarithmic data set is extrapolating the results more accurately in the whole range of threshold while the Exponential distribution is overestimating the return level.

Nonetheless, the reference values are in accordance with the ones given by HSE [2002] indicating that the extrapolation for the reference level may hold valid results for all the cases.

One case of particular concern occurs in the reference u level of M4. Here, the values of H_s extrapolated are substantially higher with the GP statistical model. Accordingly to Table 6.4 and the discussion developed previously, this is one of the cases where the GP distribution has better fitting indicators in four of the five used. Although the GP is, as the other two distributions, unbounded for this case and despite the better fitting indicators, an unrealistic extrapolation level may have been attained. Again, using a logarithmic transformation avoids this unrealistic estimate while maintaining better fitting indicators for all the measures used for comparison.

It is important to notice that, considering that the logarithmic transformed GP produces more realistic results, the Exponential extrapolation results are non-conservative for the reference level and thus, of high risk on a design basis approach.

In M5, the results of both GP cases and the two-parameter Weibull do not deviate much from each other. The Exponential presents higher estimates of the return level. M5 is the other case where the GP distribution demonstrates better fitting indicators while maintaining its bounded character, see Tables 6.3 and 6.4.

In the case of M6, the same decreasing trend was found in the value extrapolated for H_s as the u increases. In this case, the fitness results indicate the application of the two-parameter Weibull or the Exponential distributions. As the maximum occurrence registered is 17.2m, the usage of the logarithmic transformation may result in underestimation of the real return level, whereas, the usage of the Weibull and the Exponential distribution in the overestimation of it. This again may be related to the support of these probability models.

The analysis performed focuses substantially on the descriptive comparison of the different statistical models. Nevertheless, the ultimate goal of the procedure implemented is for it to be applied in a design basis approach. Vrijling and Van Gelder [1999] present a rationale on how to approach the optimal design basis when considering the need to extrapolate values. It is relevant to emphasize that analysing the practical risks of the extrapolation is mandatory for a design approach.

Generalization of the results may be obtained using the methodology proposed in Van Gelder [2000], where a sampling process is implemented to study the selection of a probability distribution to model probability *exceedance* comparing five 3-parameter distributions using samples from a GP distribution.

Considering the results of the return periods when different conditions are met (e.g. increase in u), it is unreasonable to discard the GP as a potential solution to fit *exceedance* data. In fact, the GP distribution, considering both, transformation and no transformation, should be always considered to use with the POT methodology. More critically even when physically limited data are being modelled.

6.5 Conclusions

The main purpose of the current chapter was to investigate the basis for the rejection of the Generalised Pareto distribution over the preferential application of the two-parameter Weibull and the Exponential distributions for the extrapolation of significant wave heights. This was of particular relevance given that the Generalised Pareto distribution is the limit distribution for *exceedances* over a specified threshold. Moreover, the literature review indicated that, while some authors remarked on the ideal fit offered by the Generalised Pareto distribution for significant wave data *exceedances*, other rejected or discouraged its application when paired with the Peak-over-threshold methodology.

Results of the Goodness-of-fit analysis and extrapolation showed that there is no evidence to reject the Generalised Pareto distribution over the two-parameter Weibull or the Exponential distributions. The support of the GP distribution and its bounded character is expected to contribute to a more realistic extrapolation of the return level in the physical context of ocean waves. Nevertheless, attention should be given to the selection and fitting process when applying Peak-over-threshold to model significant wave *exceedances*. Results showed that when using a Peak-over-threshold approach the user should be careful during the selection of the threshold level that will be used to truncate significant wave data.

As main conclusions of the work underpinned it is important to highlight the following:

- The threshold level is the variable of major influence in the outcome results of the POT application. The method used to ensure independence, and the number of the field points, do not show the same decisive influence in the results as the threshold. A contribution to the choice of threshold was added through the introduction of a new methodology to define the threshold level. It consists in settling the threshold value by analysing the three derivatives of the cumulative density function. Its main advantage is that the calculation of the threshold is based on a mathematical procedure that outputs a single solution that is purely defined on the shape of the probability function. The estimations of u obtained were in conformity with the expected values when comparing with other methods, however, with minimum bias (from the user).

- For the studied data set, it is impossible to generalize which distribution, that is, Weibull, Exponential or Generalised Pareto, will produce a better fit. The process should be evaluated in conjunction with a sensitivity analysis of the parameters of the study; threshold level, decoupling time, and fitting methodology; so that a definitive conclusion can be drawn. With the variation of these parameters, no indicators to reject the Generalised Pareto were found. On the contrary, strong evidence was found that backs its application. In particular when analysing the statistical support of the models compared.

- Using a logarithmic transformation of the data when applying the Generalised Pareto distribution to model significant wave data was adequate and produced promising results for the return level. This transformation must be considered for future studies of wave data. In particular it would be interesting to see how the logarithmic transformation approaches the return level when a very long dataset is available. A subset of the long data set could be used to extrapolate the maximum response in r years. In regard of the extrapolated significant wave results, the Generalised Pareto with logarithmic transformation presented return levels that, when compared with

the dataset maxima values, are expected to be the most indicative of the real return levels for the considered sites. In one of the four data-sets studied both the Weibull and the exponential seemed to greatly overestimate the return levels when compared with the recorded maxima. This happened for the buoy M1. In the remaining the return levels extrapolated were comparably similar.

- The analysis of significant wave data associated to specific the return periods showed that special attention should be given when fitting the tail with a multi-parameter distribution as the return levels might be underestimated or overestimated. The current work is representative of a data set of the Irish coast, which in turn is representative of the north Atlantic and the Irish Sea. It is of interest to extended the presented analysis to other areas.

To conclude, the interest of the presented probabilistic-based methodology to extrapolate the values of significant wave data and its relevance for highlighting the role of the different parameters in the extrapolation process is emphasized. The methodology implemented can be used to produce robust results on a design basis approach . Its systematic character, discussing different models and parameters, highlights the relevance of comprehensive analyses in design implementations, that are particularly relevant when the usage of the same modelling techniques with selection of different parameters may lead to significant differences in the estimated results.

Chapter 7

Comparative analysis of OWT fatigue design

Contents

The analysis presented in the following chapter was published in Teixeira et al. [2019b].

7.1 Motivation and original contribution of this chapter

The present chapter discusses the application of fatigue time-domain SN methodologies to design Offshore Wind Turbines. The IEC Design Load Case 1.2 is selected for a comparative analysis of SN fatigue calculations on the tower component.

It was seen that different methodologies could be applied to design OWT to fatigue. Three main methodologies were identified, direct scale up of loads and cycles to the design lifetime, statistical fit of the loading distribution, and truncation of the tail region. These are compared in the following analysis. The procedure to extrapolate the lifetime fatigue is expected to have a significant influence on the fatigue design results. If a sufficiently representative sample is available for the extrapolation methodology, then fatigue calculations are expected to be accurate. Current practices recommend a minimum of six seeded simulations as a representative sample size at each environmental condition.

For materials with large m value, extrapolation of loads and cycles is recommended as most of the fatigue damage density is located in the tail of the loading spectra. For low m materials, extrapolated is not seen as mandatory.

It is of interest to research on the different statistical assumptions for fatigue loading calculations when analysing fatigue of the tower component. In the case of OWT design, it is frequent for the design analysis to be divided in substructure and turbine, with independent calculation procedures. Both analyses are coupled at the interconnection point between the tower and substructure, at the interface piece. The purpose of this comparative analysis is therefore to contribute to the understanding of the influence of the loading spectra characterization on the design of OWT to structural fatigue.

7.2 Introduction

Chapter 4 discussed fatigue assessment for OWT. It was seen that significant efforts were made in the 1990's to promote consistency on fatigue calculations for wind turbines. The issue of fatigue calculations was raised due to the identified heterogeneity in the procedure to calculate fatigue in the design phase.

All the loading assessment fatigue design procedures have common steps, with the main divergence between them being related to the approach to the loading data, where special concern may be given or not to the analysis of the loading distribution tail. While the material characteristics have crucial influence on the fatigue design, an erroneous estimation of the loading and loading cycles may result in an incorrect estimation of the T fatigue.

Limited interest was identified in dealing with OWT structural fatigue for low m material components, such as the tower component. Most of the knowledge applied to assess structural fatigue for this type of components was transferred from previous "know-how" on offshore structures. Nevertheless, the loading of OWT has its own specificities. Moriarty et al. [2004] showed that for a WT, the low m components analysed were fatigue critical components. It is then of interest to evaluate the effect of considering distinct approaches in OWT towers design.

Definition of the loading spectra with direct scale up of loads and cycle, statistical fit to the whole distribution and extrapolation of the tail are compared in the current chapter. Regardless of the approach used to define the loading spectra, different loading conditions need to be considered in the design assessment. Fatigue design involves the characterization of the structural fatigue contributions from multiple environmental conditions (x).

When defining the loading spectra one main question needs to be answered to promote accurate fatigue estimation.

- What is a representative sample for a multi-megawatt OWT in order to accurately define the loading spectra?

Naturally, as fatigue involves extrapolating the t loading to T, some uncertainty is expected due to extrapolation. However, it is known that uncertainty can be mitigated and that above a certain sample size limited gain in accuracy is achieved. The problem of long-term fatigue assessment is a problem of mean value. As long as the mean contribution of the loading at each $D_t(x)$ (short-term SN damage rate at operational state x)to the fatigue life is well defined, so it is D_T.

Within the context described, the present work evaluates different methodologies used for the calculation of the fatigue design of low m materials using computational simulation, and compares the design procedures applied to the IEC DLC1.2 that evaluates operational fatigue.

The techniques are applied to assess fatigue of the OWT tower component. Furthermore, uncertainty in the design procedure is compared by analysing how the results from the reference design methodologies approach larger operational times. It is shown that uncertainty in the SN fatigue design due to sample size can be characterized using a bootstrapping procedure. Its interest resides on its simple application within the current SN fatigue design practices.

To achieve the proposed goal of studying SN fatigue loading:

- Section 7.3 discusses the current SN fatigue design methodologies, using the recommended practice for OWT. Sections 7.3.1 to 7.3.4 introduce the current procedures and some considerations necessary for SN fatigue calculations using different statistical techniques, such as extrapolation. Section 7.3.5 researches on the load range damage density and on how it varies with the loading environmental conditions. Section 7.3.6 compares different techniques to estimate fatigue life, introducing a discussion on Section 7.3.7 about SN fatigue loading sample size.

- Section 7.4 demonstrates how uncertainty can be quantified using non-parametric bootstrapping techniques of simple application. These can be applied to measure uncertainty in the design without being intrusive to the current design procedure. Section 7.4.2 discusses the techniques applied to set confidence intervals in the true value of the SN fatigue. Section 7.4.2 presents the main results for the estimation of uncertainty in the design due to the loading sample.

- Finally, the main conclusions of the work are presented in Section 7.5.

7.3 Comparative fatigue design of 5MW baseline turbine

7.3.1 Tower Component

The baseline wind turbine introduced in Jonkman et al. [2009], mounted on a monopile foundation is used to compute the results presented. The main characteristics of this OWT are presented in Chapter 2. Fatigue is assessed at base of the tower component, on the transition piece (TP) between the tower and the substructure. This is assumed to be the critical loading point for the tower's fatigue response, as it is the location where the bending moment is maximum. Nevertheless, the methods and conclusions presented are valid for any other locations in the tower component.

The tower component is a tapered beam with maximum diameter of 6m (base) and minimum diameter of 3.87m (top). The radius and thickness profiles are linearly tapered from base to the top. Thickness on the component varies from 27mm at the base to 19mm at the top. The Young modulus of the tower is of 210MPa, and the shear modulus

is approximately 81 MPa. The density of the tower component is corrected to account for additional mass introduced by structural components, such as bolted connections, and other additional mass, such as painting.

The tower component as described rests on a monopile with a constant diameter of 6m and a constant thickness of 60mm.

An overview of NREL's 5MW baseline turbine tower component properties is given in Table 7.1.

Table 7.1: NREL's 5 MW installed on a monopile foundation turbine tower component properties.

Tower top height above the MSL (m)	87.6
Tower base height above the MSL (m)	10
Diameter/Thickness at top height(m)	3.87/0.019
Diameter/Thickness at base height(m)	6.0/0.027
Total mass of tower (kg)	237.107
Total mass of tower + monopile (kg)	522.617
Center of Gravity vertical location (m)	37.172
Structural damping ration (all modes)	1%

Distributed tower properties are presented in Jonkman and Musial [2010].

Certified software [Hayman and Buhl Jr, 2012] was applied to validate the D_t calculations. The time series L is decomposed using a rainflow-counting (RC) algorithm.

Correction for mean load estimates are obtained using the Goodman correction. Operational simulations of 600 seconds, or 10 minutes, were considered to compute the results. Sutherland [1999], Veldkamp [2006] indicated that little gain was achieved by increasing the simulation time t to more than 600 second. The same finding occurred in Haid et al. [2013], when analysing a floating concept.

Calculations are based on the IEC DLC1.2. Environmental conditions are studied in between the cut-in and cut-out wind speeds. Nonetheless, Veldkamp [2006] reports that wind speeds from 10 to 25 m/s may account for 98% of the tower's fatigue damage in the fore-aft bending moment (moment originated by the thrust in the rotor) for a specified turbine.

Attention should be given to the fact that every simulation should be characterized by its own seed, different to other simulation performed at the same x. This is of relevance to replicate the random character of the OWT structural response.

7.3.2 Long-term fatigue damage density

Standards ruling the design of OWTs equipments foresee the possibility of using different approaches to design structural fatigue with the SN method. Selecting the adequate design procedure is considered to be highly dependent on the material being analysed. It is common to consider that, for large m materials, most of the fatigue damage density occurs in statistical tail region of the loading distribution which comprises high load ranges, *e.g.* composites. Whereas, for low m materials, the fatigue damage density is spread over the whole loading spectrum, *e.g.* steel.

However, even for low m, e.g., $m = 3$ or $m = 5$, the highest quantiles of loading (Q_{L_i}) may still contribute to significant fatigue damage. Figure 7.1 shows the damage density due to loading during a 1 year of operation of the turbine analysed. For low m contribution

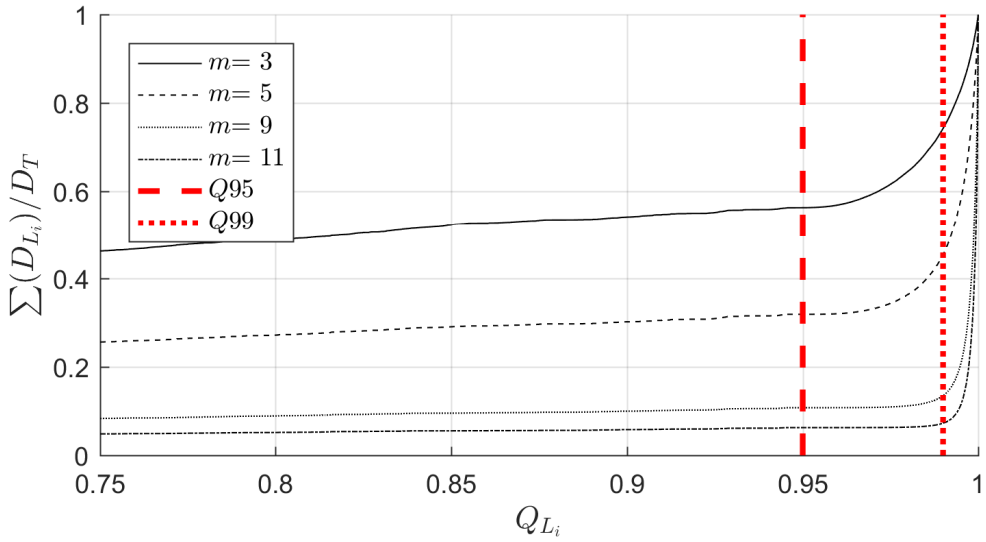


Figure 7.1: Cumulated damage contribution of amplitude L_i in the long-term damage generated in the OWT tower function of m. Results were computed using a full one year OWT operation composed of 51240 operational points.

fatigue damage density divides itself more homogeneously within different load ranges. For $m = 3$ approximately 55% of structural fatigue is induced by the 95% smallest loads. For $m = 5$ this value decreases to approximately 30%. For offshore applications it is common to find double SN slope materials with slopes of $m = 3$ and $m = 5$ [DNV, 2014a].

When designing to fatigue, it is of interest to identify how much effort should be allocate to the analysis of the body of the loading distribution. It is also of interest to research if a detailed analysis of the tail region is required in order to achieve more accurate results.

Fatigue design of the tower is compared using the following methodologies; a 2-parameter Weibull distribution accounting for all loading ranges ($W_{t(ML)}$) and fitted using the Maximum likelihood estimator (MLE); a 2-parameter Weibull distribution accounting for all loading ranges ($W_{t(MoM)}$) and fitted using the method of moments (MOM); truncations at 95% and 99% quantiles (Q) with posterior fit of a Generalised Pareto ($GP_t(> Q95)$ and $GP_t(> Q99)$ respectively) and a 3-parameter Weibull distribution accounting for the tail of the loading distribution ($W_t(> Q95)$ and $W_t(> Q99)$ respectively). In addition to the probabilistic approach, the direct scaling of loads and cycles from t to the T is considered. In all the cases a scaling constant $\frac{T}{t}$ is used.

7.3.3 Statistical distribution of t time fatigue

Lange [1996] describes fatigue as a problem of mean, which is an important consideration to analyse the loading spectra. As the loading repetitions increase the sample size, the sample mean tends to the mean of the short-term damage distribution scaled by a factor $\frac{T}{t}$.

It is known that the loading spectra is a probability approximation of the loads and cycles in a pre-specified t. If D_t is the SN cumulative damage for this load spectra, then D_t also follows a probability distribution that characterizes the statistical distribution of the $D_t|t$. Therefore, as D_T represents the sum of D_t, the cumulative value tends to the

mean of the D_t distribution multiplied by $\frac{T}{t}$, or in other words, extrapolated for T. As points to the right and to the left of μ_{D_t} have equal probability of occurring, multiple repetitions will converge to $\frac{T}{t}\mu_{D_t}$.

For the case of fatigue it is commonly assumed that under constant or random loading a normal or log-normal distribution will describe accurately the variability of the fatigue damage [Wu et al., 1997, Wirsching and Chen, 1988].

To infer on the variability of $D_t(x)$ the statistical output was studied by using probability papers to compare different probability distribution models. The results showed that the lognormal distribution appears to approximate well the statistical behaviour of $D_t(x)$. Lognormal probability papers for four random simulations of x are presented in Figure 7.2. The fact that the damage $D_t(x)$ is always positive makes the lognormal distribution

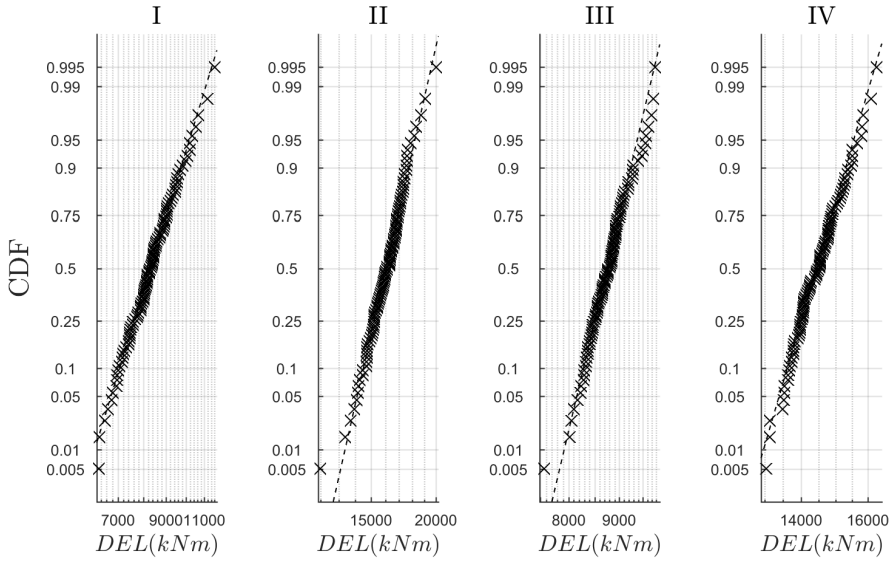


Figure 7.2: Log-normal probability papers for DEL of tower fore-aft moment for four random combinations $x = [U, H_s, T_p, I]$. I - $[8.5, 1.9, 6, 8.9]$. II - $[11.4, 2.2, 8.2, 7.9]$ III - $[18.9, 3.9, 12, 5.8]$ IV - $[24.8, 4.2, 10, 4.0]$. A fixed value of $n = 200$ (equivalent SN damage in 200 cycles) was considered to calculate the DEL in all the IV cases.

especially suitable.

It is known that the logarithm of a lognormal distributed variable follows a normal distribution. It may be more practical for the variable of analysis not to be the SN damage rate in t but instead its logarithm. In such case a transformation of variable can be applied to the lognormal distribution,

$$D_t(x) \sim \log \mathcal{N}(\mu, \sigma^2) \tag{7.1}$$

meaning that,

$$\ln(D_t(x)) \sim \mathcal{N}(\mu,\, \sigma^2) \tag{7.2}$$

the convergence behaviour of D_t, is of crucial importance to understand the problem of OWT fatigue and the importance of having a representative sample of loading.

When using the tail region to perform fatigue design calculations, some considerations on extrapolation are demanded.

7.3.4 Considerations for loading extrapolation

Extrapolation of loads for OWT has been extensively studied in the past [Cheng, 2002, Moriarty et al., 2004, Fogle et al., 2008, Moriarty, 2008, Agarwal, 2008, Teixeira et al., 2019a]. Most of the research works on the characterization of the tail region were directed towards the extreme loads for U design. For fatigue, extrapolation of loads may be of interest to account for the long-term highest load ranges, L, which are expected to account for important structural fatigue damage contributions.

The POT methodology is applied to characterize the statistical tail region. Using the POT, structural fatigue assessment can be divided in two independent problems, the above-u problem and the below-u problem. This division is recommended by IEC [2005].

Some considerations are needed to implement the POT. In particular, independence was seen to be of importance to achieve accurate results. Limited interest was given to independence when extrapolating for fatigue calculations. Fogle et al. [2008] showed that for the analysed turbine, $30s$ intervals between fore-aft moment maxima occurrences guaranteed independence with 99% confidence.

Choice of the threshold (u) is another particularly challenging problem that demands detailed considerations. A comprehensive discussion on the choice of threshold was presented earlier in Chapter 2. In the following comparative study, quantiles of 95% and 99% are considered. The usage of direct quantiles allow simple repeatability of the fatigue analysis method. Embrechts et al. [1999] indicate that these may theoretically inadequate. Nonetheless, so far no rejection of these was found in application to OWTs. In fact, the $Q95$ is recommended by IEC [2005], and the $Q99$ is applied by Moriarty et al. [2004] (which also recommended by IEC as a reference case for extrapolation of loading).

7.3.5 Load range influence

While for low SN slope materials a substantial decrement of the fatigue life occurs due to relatively small loading amplitudes, which are expected to repeat more often, it can be seen in Figure 7.3 that, in the case of the tower, less common L located in the tail region still have an important contribution to decrease the fatigue life.

In order to better understand the results presented in the previous section, SN fatigue variability for a number of seeded (with different seed) simulations (s_0) of 6, or $s_0 = 6$, is discussed. More precisely, ten independent repetitions of $s_0 = 6$ simulations for 11 operational states (see Table 7.2) are presented in Figure 7.3, with a total 60 simulations per simulated case. $s_0 = 6$ is the minimum number of simulations per environmental condition recommended by the IEC standard when assessing SN fatigue loading calculations. Different simulation scenarios, function of 2m/s U bins, are analysed. Values of significant wave height (H_s), peak period (T_p) and turbulence intensity (I) considered for each bin are presented in Table 7.2.

For almost all the simulated cases, the D_t from all the L is similar to the variability induced by the lower $Q99$ L. With exception of Case 3, where the variability of the lower 99% loading contribution is similar to the 100% loading contributions. Furthermore, the smaller the contribution from low L to D_t, the larger the variability of D_t, see case 8 where there is a decrease of contribution of $D_{t:<Q99}$ to D_t.

In Case 3 the tail L dominates D_t. For such scenarios almost all the SN fatigue damage density is located in the tail region, where there is less percentage of loading data, therefore, inducing larger variability in D_t.

Table 7.2: Simulation variables for different operational conditions. Calculations of Table 7.3 are based on these environmental states. The value representative for each U is taken as the midpoint of the bin.

	Simulation Case										
x	1	2	3	4	5	6	7	8	9	10	11
$U(m/s)$	$[3,5]$	$[5,7]$	$[7,9]$	$[9,11]$	$[11,13]$	$[13,15]$	$[15,17]$	$[17,19]$	$[19,21]$	$[21,23]$	$[23,25]$
$H_s(m)$	1.23	1.75	2.86	2.66	3.67	3.84	4.12	3.79	4.34	3.50	5.46
$T_p(s)$	5.80	6.99	11.00	7.28	13.52	11.44	11.84	9.40	16.26	12.7	17.81
$I(\%)$	IEC NTM - class B										

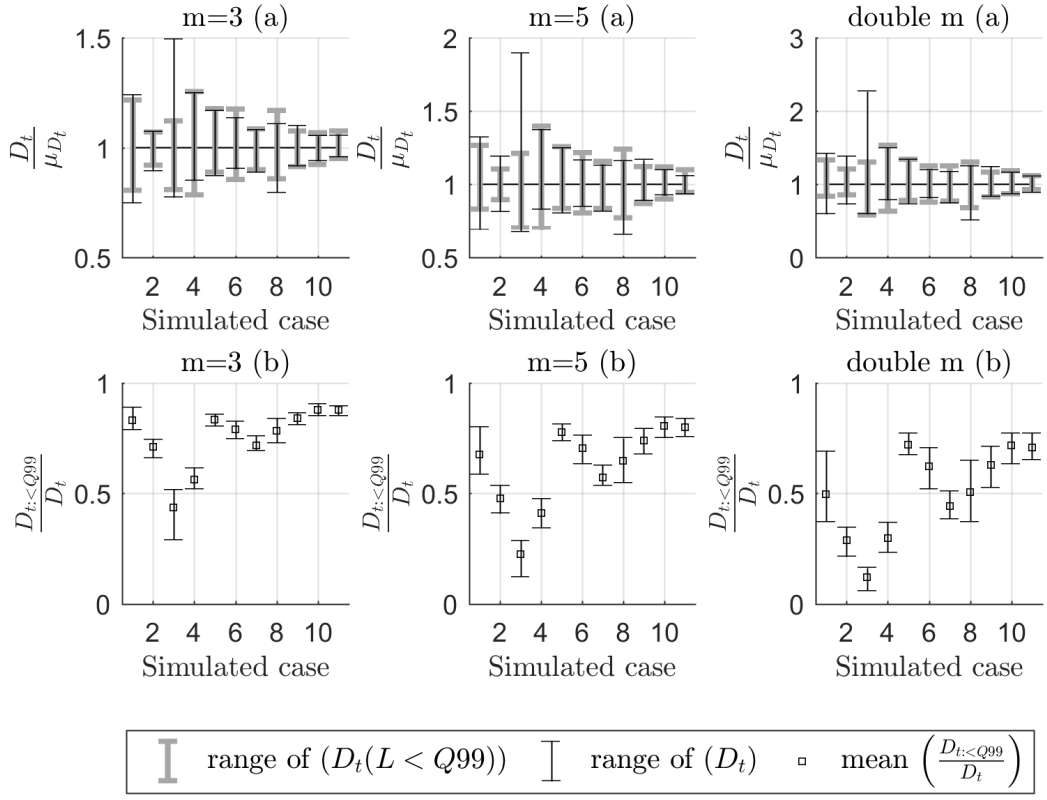


Figure 7.3: Variation of the contribution of t fatigue considering 10 sets of $s_0 = 6$ and a truncation at the 99% L quantile ($Q99$). (a) presents the short term uncertainty for truncated and non-truncated data and (b) the relative percentage of the lower than $Q99$ loading to D_t for $m = 3$, 5 and double m respectively. Each simulated case is a operational state with different environmental conditions.

The loading of the turbine changes considerably from case 4 to 5. In case 5 the turbine is already operating at its rated power. In case 3 and 4 the turbine operates at optimization of the power capture where the tip-speed ratio [1] is maintained constant. For

[1] Ratio between rotor and wind speed.

the operational U value of cases 3 and 4 there is a significant change in the rotor speed. Therefore, the loading regime for these cases is considerably different from the remaining operational U values.

Results show that variability in the cumulated damage contributions from the smaller 99% L can have relative variations to the mean value as large as approximately 25% for $m = 3$. This value increases with the increase of the slope m. The results in subplots (b) show that short-term contribution is not constant with variations in the environmental parameters. Additionally, short-term contributions decrease with the increase of m.

For the points analysed the contribution of the low range loading to D_t was relatively stable above the turbine's rated U (11.4 m/s) with a small increase relatively near the cut-out U. These x states are expected to have the largest contribution to D_T, Figure 7.4.

7.3.6 Application of distinct procedures to assess the fatigue design

Results for the approximation of the structural fatigue occurring on a T time of 100 simulations using a design sample size (s_0) of 6 simulations are presented in Table 7.3 and 7.4. The respective operational states considered for calculations were previously presented in Table 7.2. Table 7.4 presents the parameters of the SN fatigue calculations.

The interest of such analysis is to infer on how different approaches for the analysis of $f(L|x)$ calculated with the recommended minimum of 6 seeded simulations, $s_0 = 6$ approximates the SN fatigue induced by a relatively larger number of seeded simulations.

Reference results for the fatigue damage equivalent loading are presented in Figure 7.4.

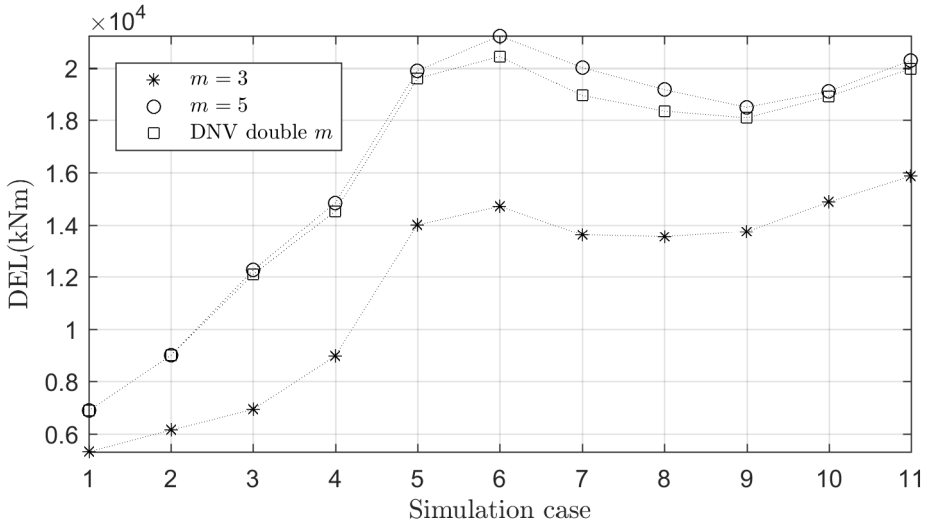


Figure 7.4: Results of fatigue damage equivalent loads (DEL) for the TP moment at the different operational states presented in Table 7.2. 100 simulations were used in the calculations.

The fatigue damage increases with the increase of U up to a U value slightly above the rated speed of the turbine. The DEL then decreases and increases again near the cut-out U.

For $m = 3$ the maximum equivalent load is achieved near the cut-out U, while for $m = 5$ it occurs close to the rated U. The fact that the maxima of loading is expected

to occur near the rated U seems to contribute to the difference between the maxima for $m = 3$ and $m = 5$.

The double SN curve has similar DELs to $m = 5$. However, increasing differences between both DELs can be identified after the rated U, which might be also related to the change in slope for high load ranges.

Table 7.3: Extrapolated values of damage for the different simulation cases presented. The results are presented for the direct scaling of loads, and six probability distributions to account for the total or just the tail region of the loading distribution as introduced in Section 7.3.2. Extrapolation is compared using the SN fatigue from $s_0 = 6$ to predict the SN fatigue in $s_0 = 100$. Double vertical line separates the simulated cases from below to above rated power.

		Simulation Case										
	Methodology	1	2	3	4	5	6	7	8	9	10	11
$m=3$	$\frac{T}{t}D_t$	1.154	0.935	0.852	1.026	0.979	0.96	0.88	1.132	1.073	1.056	1.042
	$\frac{T}{t}D_{W_t}$ MLE	1.425	1.149	0.67	0.682	1.727	1.825	1.204	2.096	1.84	2.518	2.139
	$\frac{T}{t}D_{W_t}$ MOM	1.394	1.119	1.08	1.375	2.912	2.769	1.378	1.833	1.514	1.694	1.397
	$\frac{T}{t}D_{GP_t(>Q95)}$	1.698	1.382	1.431	1.373	1.047	1.053	1.067	1.358	1.734	1.131	1.148
	$\frac{T}{t}D_{W_t(>Q95)}$	1.542	1.296	2.593	1.758	1.054	1.094	1.132	1.446	1.294	1.154	1.16
	$\frac{T}{t}D_{GP_t(>Q99)}$	1.173	0.981	0.891	1.059	1.01	0.993	0.906	1.176	1.124	1.087	1.061
	$\frac{T}{t}D_{W_t(>Q99)}$	1.189	0.988	0.888	1.069	1.008	0.996	0.914	1.177	1.137	1.079	1.067
$m=5$	$\frac{T}{t}D_t$	1.287	0.619	0.62	1.024	0.951	0.912	0.831	1.299	1.527	1.14	1.151
	$\frac{T}{t}D_{W_t}$ MLE	1.412	0.669	0.151	0.303	>5	>5	2.16	>5	>5	>5	>5
	$\frac{T}{t}D_{W_t}$ MOM	1.314	0.617	0.512	1.809	>5	>5	2.996	4.864	3.1	>5	2.844
	$\frac{T}{t}D_{GP_t(>Q95)}$	>5	2.742	1.755	1.685	1.08	1.073	1.327	1.996	>5	1.322	1.513
	$\frac{T}{t}D_{W_t(>Q95)}$	>5	3.016	>5	>5	1.16	1.381	2.29	3.455	3.37	1.619	1.876
	$\frac{T}{t}D_{GP_t(>Q99)}$	1.325	0.699	0.654	1.05	0.998	0.961	0.872	1.378	1.981	1.22	1.196
	$\frac{T}{t}D_{W_t(>Q99)}$	1.603	0.838	0.692	1.125	0.995	0.981	0.928	1.535	2.089	1.2	1.263
double m	$\frac{T}{t}D_t$	1.287	0.619	0.666	0.998	0.977	0.936	0.847	1.222	1.151	1.125	1.095
	$\frac{T}{t}D_{W_t(mle)}$	1.412	0.669	0.162	0.327	2.068	2.093	1.31	2.937	2.726	4.224	3.253
	$\frac{T}{t}D_{W_t(mom)}$	1.314	0.617	0.54	1.285	3.757	3.363	1.577	2.458	2.054	2.506	1.804
	$\frac{T}{t}D_{GP_t(>Q95)}$	>5	2.304	1.422	1.518	1.07	1.054	1.059	1.571	2.064	1.27	1.256
	$\frac{T}{t}D_{W_t(>Q95)}$	>5	2.513	4.08	2.291	1.087	1.117	1.237	1.804	1.608	1.332	1.318
	$\frac{T}{t}D_{GP_t(>Q99)}$	1.325	0.699	0.702	1.047	1.02	0.981	0.886	1.302	1.252	1.187	1.13
	$\frac{T}{t}D_{W_t(>Q99)}$	1.603	0.838	0.718	1.059	1.017	0.984	0.899	1.305	1.275	1.171	1.141

Table 7.3 shows that limited improvement is achieved by fitting probability distributions to the data (methodologies 2 to 7). Above the rated speed, in Cases 8-11, fitting a statistical distribution to the body of the data frequently generates conservative SN

Table 7.4: Parameters of the comparative analysis results for Table 7.3. σ and φ refer respectively to location and shape parameters.

		Simulation Case										
Method	Parameter	1	2	3	4	5	6	7	8	9	10	11
D_t	$\frac{T}{t}$	$\frac{100}{6}$										
$D_{W_t(mle)}$	(ζ)	1.43	1.29	1.06	0.89	0.71	0.71	0.79	0.82	0.86	0.91	0.94
	(φ)	5.91	6.04	4.67	4.79	6.04	6.65	7.21	9.15	9.89	11.65	13.18
$D_{W_t(mom)}$	(ζ)	1.47	1.32	0.87	0.72	0.67	0.68	0.77	0.85	0.92	0.99	1.07
	(φ)	5.96	6.1	4.26	4.15	5.86	6.53	7.11	9.33	10.23	12.12	13.84
$D_{GP_t(>Q95)}$	(ζ)	0.12	0.09	-0.13	-0.36	-0.67	-0.63	-0.08	-0.14	0.21	-0.28	-0.09
	(φ)	2848	3485	9515	16479	11972	14737	9181	10853	5782	7934	7395
	$u(kNm)$	8555	9762	8151	11632	28770	29649	24086	25816	25456	28267	29206
	M	49	54	60	62	54	48	55	50.5	51.5	50	52
	$Max(L)$	43007	46619	54160	54021	46414	52727	76557	75323	127375	52777	69627
$D_{W_t(>Q95)}$	(ζ)	0.83	0.85	0.85	1.15	1.55	1.28	0.94	0.98	0.85	1.06	0.98
	(φ)	2955	3525	7781	12632	7942	9620	8310	9400	6656	6242	6735
	$u(kNm)$	8555	9762	8151	11632	28770	29649	24086	25816	25456	28267	29206
	M	49	54	60	62	54	48	55	50.5	51.5	50	52
	$Max(L)$	41370	47244	92463	85307	57974	75678	94934	97887	96048	69589	81580
$D_{GP_t(>Q99)}$	(ζ)	-0.69	-1.17	-0.78	-0.11	-0.59	-0.43	0.22	-1.19	0.41	-1.09	-0.15
	(φ)	5417	8634	9170	6585	4119	4437	4474	20295	5493	11530	6609
	$u(kNm)$	14484	16573	24767	31281	39853	44156	40182	41241	37079	37870	40293
	M	7	8.5	8	9.5	8.5	7.5	9	7	9	8.5	8.5
	$Max(L)$	22051	23899	36250	55195	46396	53017	76268	58209	111109	48358	62347
$D_{W_t(>Q99)}$	(ζ)	0.81	0.8	1.01	0.9	0.93	0.68	0.71	0.82	0.65	1.08	0.78
	(φ)	2883	3253	4894	5701	2350	2479	4712	7625	6626	4889	5177
	$u(kNm)$	14484	16573	24767	31281	39853	44156	40182	41241	37079	37870	40293
	M	7	8.5	8	9.5	8.5	7.5	9	7	9	8.5	8.5
	$Max(L)$	33304	38421	47001	62416	51967	67639	81030	89891	107626	58082	76659

fatigue design values (over-estimate the value of D_T, ratio is larger than 1 in the table). Below the rated speed, as example Cases 1-4, fitting the whole data-set generates non-conservative results. It is noted that, in general terms, the estimation is progressively more conservative for higher-m, which indicates that this deviation is likely to be connected to the probability tail region.

Even considering that the probability tail is highly influential for high-m materials, it was seen that the tail region has important influence in low-m materials. If the whole distribution is fitted (methodologies 2 and 3), the accuracy of the predictions in the tail region is expected to decrease. The MOM ,when comparing to the MLE, presents similar accuracy for $m = 3$, but is more conservative for $m = 5$. However for the particular simulated cases 3 and 4, it produces a significantly more accurate approximation for the T fatigue damage.

Sutherland and Veers [1995] remarked that in the case of wind turbines, a fitting technique was not a substitute of an adequate loading database.

In the case of the tower, fitting the whole distribution with a Weibull model is not a robust procedure. It is known that fatigue calculations are highly sensitive to the

statistical model when the whole distribution is fitted [Lange, 1996]. It is noted that the quadratic Weibull model, particularly developed to be a flexible probability model for the tail region, was not considered. Nonetheless, considering probability theory, if the interest is to characterize the tail region, the analysis should be focused on it. Even considering the limitation identified when fitting the whole distribution, having all the data compiled in a single model may be of interest for some calculations, such as reliability analysis. In such cases, uncertainty can be circumscribed to the parametric variables.

Truncating the loading data-set using a threshold, and consequent fit of a probability model is motivated by this limitation of dealing with the tail region. This methodology is seen as recommended for high m materials. Results, however, show that application of truncation methodologies does not always result in an improvement of the long-term fatigue estimation. Usage of a $Q95$ produces very conservative results. For the analysed turbine component this u value is not appropriate to analyse the tail region. Confirming the magnitude of the maxima on the reference time can be used to double-check the loading extrapolation, Table 7.4. In particular, the W_t and the GP_t distributions with positive shape parameter $\zeta > 0$ have unbounded character, which may result in non-realistic maxima loading. Clear examples of this can be identified when comparing the extrapolated maximum L with the remaining maxima (e.g. $DGPt(> Q95)$ for case 9).

The application of a $Q99$ slightly improves the estimation of D_T in the case where deterministically scaling from D_t to D_T under-predicts the value of SN damage in T (e.g. Cases 5 and 6, and for the reference T size).

As indicated in Table 7.2, NTM has been used to obtain the turbulence intensity of the wind. The approach proposed by NTM has been criticized before as a very simplistic approach to the turbulence description [Sutherland, 1999]. To corroborate the estimations on T, 25 LHS points extracted using the joint distribution of U, I, H_s, T_p were used to calculate extended comparative results. These are presented in Appendix, Tables A.1 and A.2, and corroborate the results presented in Table 7.3. The GP, the Weibull and the Exponential distributions are applied to extrapolate the tail region. $Q95$ and $Q99$ are similarly applied as the value of u.

This further analysis enables the detection of an important limitation when extrapolating using a u value related to L for SN fatigue calculations. Further discussion on the results is presented in the Appendix A.

Figures 7.5 and 7.6 present the results for the loading spectrum, including tail fit, when approximating a T resulting from 1000 simulations using two initial samples sizes of $s_0 = 6$ and $s_0 = 25$, respectively.

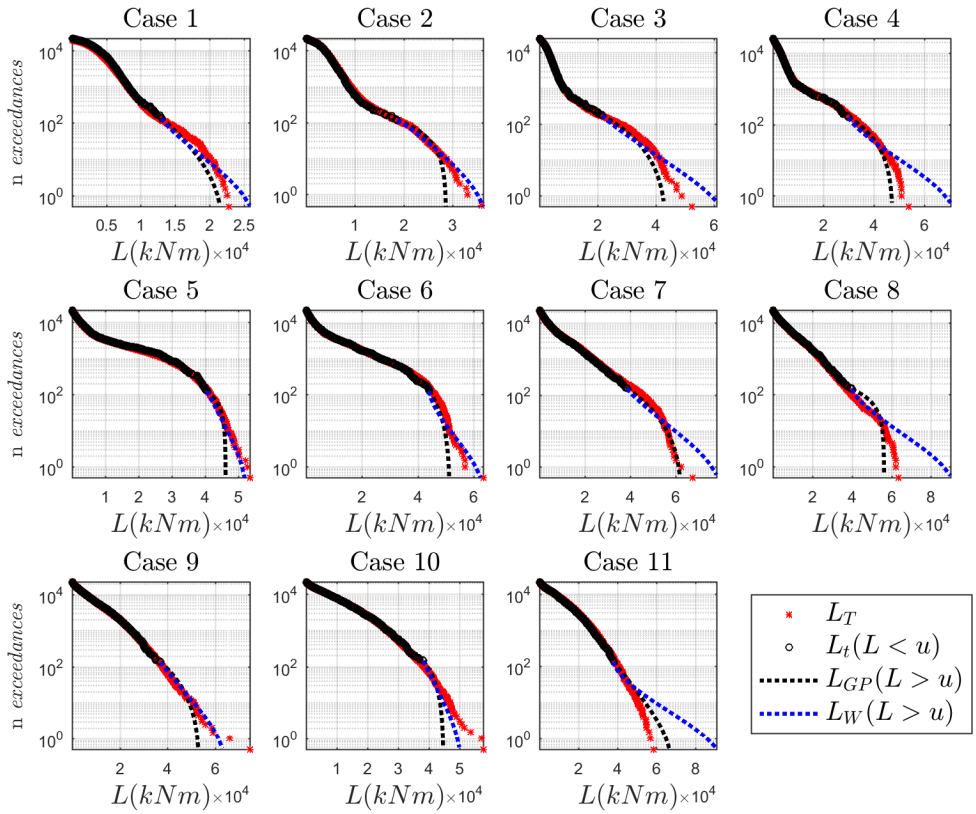


Figure 7.5: Load spectrum approximation considering a $s_0 = 6$ and comparing with a $f(L|x)$ estimated using 100 simulations.

When comparing with the loading spectrum extrapolated from an $s_0 = 6$, the loading characterization using a s_0 of 25 results in a better approximation of the load spectrum characterized 1000 simulations. This approximation is particularly more accurate in the tail region. The Weibull and the GP distributions, with unbounded support, may result in unrealistic loading estimates, which then may result in highly conservative D_T estimations when the tail region is of major importance.

In-line with [Sutherland, 1999], results showed that fitting techniques are no substitute for a reliable sample in order to estimate OWT fatigue in the tower component. A representative sample is necessary to have accurate estimations of fatigue life. The cost of producing a representative sample may be relatively high. However, once it is set, it is expected to approximate well D_T regardless of the T. The following section studies the sample convergence for different material slopes.

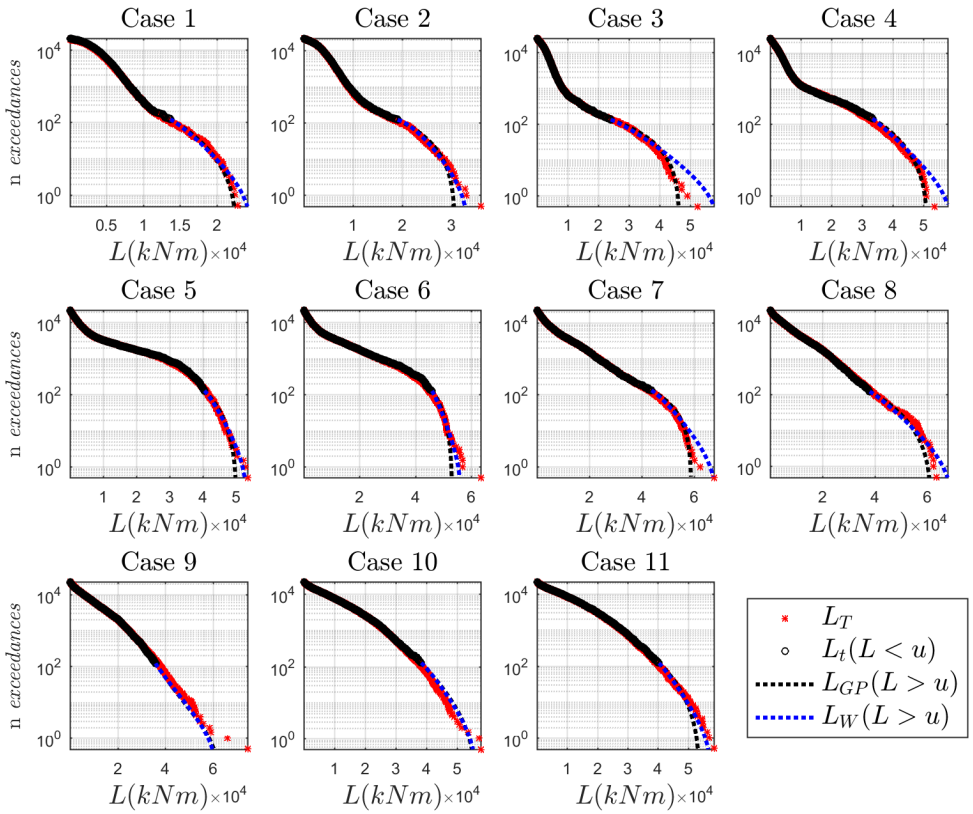


Figure 7.6: Tail approximation considering a $s_0 = 25$ and comparing with a $f(L|x)$ estimated using 100 simulations.

7.3.7 Representative sample

The size of a representative sample is a fundamental question in the fatigue design loading calculations. Previous sections remarked the fact that an initial s_0 of six simulations may not be representative enough to produce accurate predictions of the OWT SN fatigue.

SN fatigue estimations are accurate when a representative sample is defined [Sutherland, 1999]. This is true regardless of the T time, whether it is 5 years or 50 years. Convergence of the sample size is investigated. Figure 7.7 presents results for convergence with an incremental s_0 varying between 1 and 100 simulations and compared with the DEL results for a T of 1000 simulations. The process is repeated 10 times, originating the convergence profiles shown. DEL calculations were performed at $x = [U = 11.5m/s, H_s = 2.2m, T_p = 10.2s]$ and I given by the NTM class B.

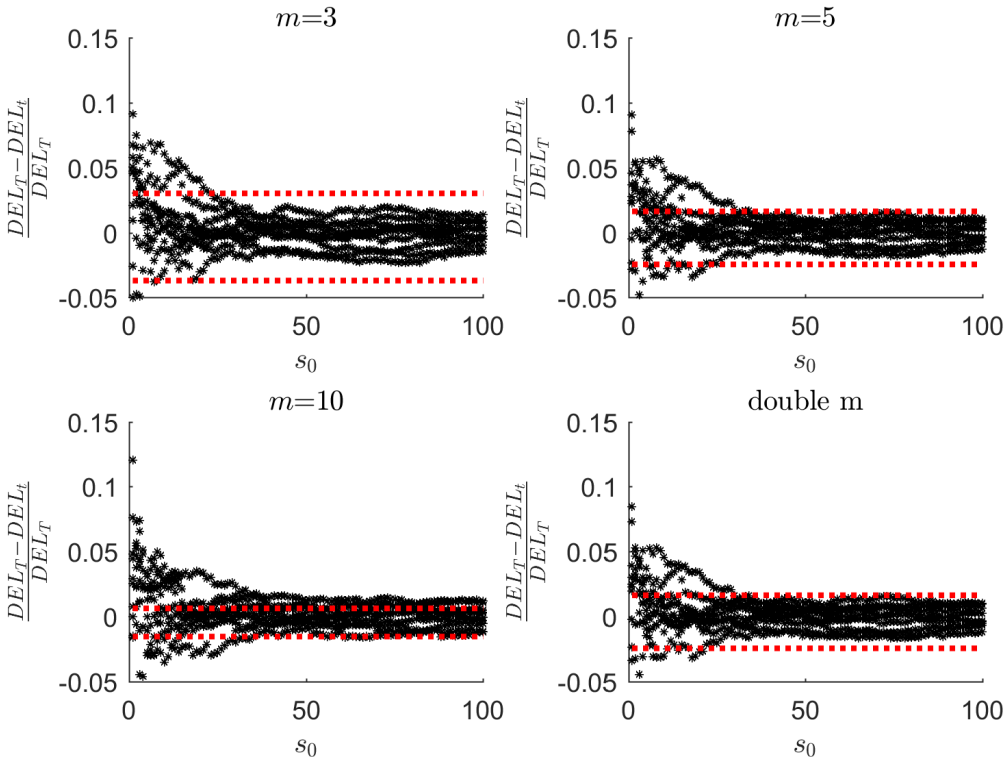


Figure 7.7: Convergence of the DEL with s_0. 10 incremental s_0 simulations are used and D_T is computed using 1000 simulations at $x = [U = 11.5m/s, H_s = 2.2m, T_p = 10.2s]$ and I given by the NTM class B. The red trimmed lines correspond to the variation of DEL that origins a 10% error in the SN fatigue estimation.

Understandably the higher the value of m the more difficult is to converge the t DEL to T DEL. $m = 10$ was added due to its interest for comparative analysis. Currently, the towers are mainly produced in low m, but with the increase of tower height, high m materials may penetrate the tower construction market. Some activities are already being developed that preview this trend, see C-Tower project [van der Zee et al., 2017].

For low m materials, $m = 3$ and $m = 5$, approximately 20 and 30 simulations respectively, seem to suffice to have less than a 10% error in the estimation of D_T for the x simulated. For both cases, the error in the estimation of DEL rarely originated an error of SN fatigue damage exceeding 20% when comparing to the "real" value (assuming that a D_T given by 1000 simulations was representative enough for the error to be approximately 0). The double slope case presents similar results. For $m = 10$, a small s_0 resulted in high uncertainty in the DEL estimation. Even considering an $s_0 = 100$ the error value for the estimated D_T may still comprise a maximum variation of approximately $\pm 10\%$ when comparing with its "true" value.

Nevertheless, for the cases simulated ($x = [U = 11.5m/s, H_s = 2.2m, T_p = 10.2s]$ and I given by the NTM class B) the problem of the T time DEL approximation, even for a high m material, is still a problem highly related to the s_0 accuracy.

Results in Table 7.5 show that extrapolation of t load ranges and cycles to T contributes to an increase of the estimated D_T. Extrapolation adds contribution from larger load ranges. However, the accuracy of the approximation is highly related to s_0, see first and fourth repetition for each approach when $s_0 = 6$.

Table 7.5: Five repetitions of the D_T estimation using a $s_0 = 6$ and a $s_0 = 10$ and extrapolating above $Q99$. Each row denotes a calculation process and represents a repetition within the method used. GP, W and Exp are respectively Generalised Pareto, Weibull and exponential fits to the tail region.

	$s_0 = 6$					$s_0 = 10$				
m	3	5	double m	10	14	3	5	double m	10	14
$T/tD_t[GP(L>u)]$	0.95	0.92	0.94	0.76	0.57	0.95	0.94	0.93	0.99	1.19
	0.97	0.97	0.95	1.29	2.53	1.04	1.04	1.05	1.03	0.91
	0.97	1.06	0.97	>5	>5	1.02	1.02	1.03	0.96	0.82
	0.84	0.78	0.81	0.59	0.4	1.01	1.04	1.03	1.08	1.03
	1.14	1.23	1.22	1.4	1.5	1.11	1.12	1.12	1.05	0.95
$T/tD_t[W(L>u)]$	0.95	0.92	0.94	0.83	0.76	0.95	0.95	0.94	1.2	2.23
	0.98	1.02	0.96	2.54	>5	1.04	1.05	1.05	1.11	1.18
	0.96	0.96	0.96	1.14	2.28	1.02	1.02	1.03	0.97	0.85
	0.84	0.79	0.8	0.63	0.5	1.01	1.04	1.03	1.15	1.26
	1.14	1.26	1.22	1.78	3.48	1.11	1.12	1.12	1.07	1.01
$T/tD_t[Exp(L>u)]$	0.95	0.92	0.94	0.79	0.65	0.95	0.94	0.93	1.06	1.5
	0.97	0.97	0.95	1.27	2.41	1.04	1.05	1.05	1.07	1.05
	0.96	0.94	0.96	0.85	0.75	1.02	1.02	1.03	1.01	0.97
	0.84	0.78	0.8	0.62	0.49	1.01	1.05	1.04	1.24	1.59
	1.14	1.24	1.22	1.52	2	1.11	1.13	1.12	1.21	1.49
T/tD_t	0.93	0.89	0.93	0.73	0.55	0.93	0.92	0.92	0.94	1.05
	0.95	0.93	0.93	1.05	1.41	1.03	1.03	1.04	1	0.89
	0.94	0.92	0.94	0.87	0.84	1.01	1	1.02	0.92	0.78
	0.82	0.76	0.79	0.56	0.38	0.98	1	1.01	1.02	0.96
	1.12	1.2	1.2	1.36	1.46	1.08	1.08	1.1	0.99	0.87

The maximum L only increases slightly with the increase of the T time. Load calculations in Moriarty et al. [2004], for a pitch regulated turbine, show that an increase of 5% occurs in L from a 10^{-4} to a 10^{-5} *exceedance* probability.

The additional SN fatigue damage associated with very large L, of low probability of occurrence, increases therefore slightly with T. This increase is progressively more relevant as the importance of the tail loading increases. Directly scaling up from t to T does not account for the complete tail statistical behaviour. Still, as the M cycles are extrapolated to T, some of the high L are counted T/t times (were in fact they are less likely to occur than T/t times), which compensates to some extent the effect of disregarding the very large load ranges only given by the tail extrapolation.

Furthermore, when using a small s_0, extrapolation may result in significant over or under-estimation of D_T (examples of repetitions 1 and 2 for $s_0 = 10$ and $s_0 = 6$ respectively, or repetition 4 for $s_0 = 6$). The support of the tail extrapolation (maximum range when support is $[0, +\infty)$) should be investigated meticulously when applying extrapolation of loads and cycles.

If an adequate s_0 is considered, extrapolation is expected to be particularly important for materials with large m, such as $m = 10$ and $m = 14$. For these materials, if s_0 does not already comprise statistical information to characterize the tail efficiently, extrapolation will not suffice the problem of characterizing large L and cycles.

The purpose of the table presented was to show indicative results, for more conclusions on the effects of extrapolation demand more repetitions of the extrapolated D_T using more repetitions is required, instead of only 5.

Results for the DEL convergence of Table 7.2 are presented in Table 7.6.

Table 7.6: DEL convergence for Cases 1 to 11 of Table 7.2. The statistical variation of the DEL was obtained using bootstrapping from 100 simulations.

s_0	$m = 3$		$m = 5$		$m = 10$		double m	
	μ	COV	μ	COV	μ	COV	μ	COV
Case simulation 1								
1	5233	0.058	6710	0.06	10166	0.102	11863	0.026
6	5328	0.028	6950	0.04	10797	0.053	12768	0.011
10	5325	0.02	6913	0.027	10706	0.037	12698	0.008
20	5363	0.019	6985	0.019	10911	0.021	12986	0.008
100	5315	0	6980	0	10980	0	13094	0
Case simulation 2								
1	6091	0.085	8928	0.109	14506	0.127	17034	0.03
6	6137	0.028	8917	0.049	14596	0.069	17319	0.01
10	6126	0.021	8894	0.032	14614	0.038	17400	0.007
20	6170	0.013	8984	0.024	14996	0.041	18053	0.004
100	6185	0	9082	0	15693	0	18850	0
Case simulation 3								
1	6695	0.078	11494	0.134	19587	0.163	23125	0.022
6	6921	0.028	12197	0.038	21470	0.043	25819	0.007
10	6810	0.021	11985	0.033	21542	0.05	26221	0.005
20	6971	0.018	12329	0.02	21802	0.024	26438	0.005
100	7018	0	12527	0	22030	0	26515	0
Case simulation 4								
1	9263	0.102	15024	0.102	23772	0.105	27637	0.034
6	8921	0.045	14724	0.055	24022	0.073	28355	0.014
10	8850	0.032	14674	0.033	24154	0.041	28749	0.01
20	8904	0.018	14810	0.018	24532	0.021	29254	0.005

Continued on next page

Continued from previous page

100	8966	0	14927	0	24746	0	29518	0
Case simulation 5								
1	14272	0.041	20007	0.045	27479	0.055	30659	0.019
6	13827	0.033	19598	0.028	27343	0.028	30828	0.015
10	14070	0.019	19987	0.014	27976	0.012	31563	0.008
20	14038	0.013	19964	0.011	27918	0.012	31512	0.006
100	14015	0	19993	0	28134	0	31892	0
Case simulation 6								
1	14861	0.061	21329	0.062	30224	0.058	33995	0.027
6	14547	0.037	21010	0.034	30188	0.034	34280	0.016
10	14751	0.024	21217	0.021	30457	0.026	34728	0.01
20	14628	0.023	21095	0.021	30259	0.018	34389	0.01
100	14673	0	21163	0	30297	0	34379	0
Case simulation 7								
1	13922	0.079	20479	0.119	30773	0.145	35340	0.031
6	13540	0.019	19898	0.019	30631	0.023	35699	0.007
10	13749	0.02	20229	0.024	31078	0.027	36259	0.008
20	13631	0.019	19972	0.022	30652	0.029	35829	0.007
100	13525	0	19865	0	30756	0	36160	0
Case simulation 8								
1	13533	0.065	18862	0.1	28000	0.146	32242	0.027
6	13414	0.028	19036	0.049	29129	0.084	34122	0.011
10	13502	0.024	19125	0.026	29758	0.034	35353	0.009
20	13599	0.017	19285	0.025	30127	0.033	35823	0.006
100	13619	0	19365	0	30450	0	36335	0
Case simulation 9								
1	13862	0.053	18238	0.064	25114	0.09	28330	0.026
6	13686	0.011	18378	0.018	26679	0.04	31050	0.005
10	13746	0.016	18497	0.03	27322	0.092	32074	0.007
20	13718	0.007	18355	0.012	26683	0.029	31313	0.003
100	13840	0	18659	0	28702	0	35750	0
Case simulation 10								
1	14645	0.036	18589	0.044	24908	0.066	28025	0.019
6	14845	0.019	19014	0.021	25954	0.041	29588	0.009
10	14896	0.015	19138	0.017	26462	0.037	30471	0.007
20	14916	0.011	19119	0.011	26360	0.024	30421	0.005
100	14990	0	19315	0	26912	0	31292	0
Case simulation 11								
1	15917	0.033	20268	0.032	27030	0.037	30300	0.018
6	16040	0.013	20513	0.015	28225	0.029	32332	0.006
10	15866	0.014	20258	0.017	27836	0.029	32043	0.007
20	15871	0.011	20308	0.012	28180	0.026	32577	0.005
100	15899	0	20294	0	28017	0	32361	0

Results of convergence improve with the increase of s_0. For $m = 3$ in all the cases a $s_0 = 20$ contributed to improve the results significantly, lowering the Coefficient of Variation (COV). As the sample size increases in average the COV decreases. Although some exceptions can be found, this was the common trend.

As the m increases, a clear trend in regard to the mean DEL value, such as the one found for $m = 3$ cannot be identified. For $m = 10$ the difference between a s_0 of 6 and 10 was not evident. A s_0 of 10 is expected to produce improved results, yet, when considering convergence for large m slopes the gain from computing is not evident. A

$s_0 = 20$ is consistently more accurate than s_0 of 1, 6 and 10. For 8 of the 11 simulated cases, the estimation of DEL for $s_0 = 20$ was more accurate than the remaining s_0. For the cases where the DEL was not more accurate, the $s_0 = 10$ produced the best results.

In the case of $m = 5$, it was already seen (when comparing to a T of 1000 simulations) that $s_0 = 20$ is expected to produce consistent accuracy when comparing with $s_0 = 6$ or $s_0 = 10$. Still, in the calculations presented this is not so evident. Limitations due to sample size may be hindering more conclusive results. Ideally, asymptotic DEL were required for comparison. It is important to remark that asymptotic calculations (such as approximation of 1000 points using a similar large dataset of s_0 calculations) are challenging to establish due to computational limitations.

It was expected for a $s_0 = 20$ to produce more accurate results. However, due to the high number of loading cases that are expected to induce fatigue in an OWT, performing more than 20 simulations per x may not be feasible. The computational effort needed to assess these 20 simulations grows exponentially with the dimension of x.

In order to maintain the designing procedure practicable, and in addition to the complex statistical analysis, it may be relevant to allocate designing time to characterize how much of the loading distribution is not known in the design phase due to s_0 size. By focusing on the uncertainty associated with s_0, it may be possible to quantify the information that it is not known about the loading spectrum, establishing confidence intervals (CI) for the variation of D_T.

In the fatigue code calibrations identified, statistical uncertainty related to s_0 in the definition of $f(L|x)$ and consequent approximation of D_T or DEL for T was taken into account in a limited way. Ronold et al. [1999] characterize load uncertainty by the capability of the quadratic Weibull distribution to fit real loading data at 39 wind climates. Sample size is not specified, and as a result it is difficult to infer if the quantified uncertainty is originated by the expected fit deviations, or also encloses the effect of s_0 at multiple x. Ronold and Christensen [2001] use the same coefficient to account for uncertainty in the loading spectra.

Tarp-Johansen [2003] specifies uncertainty in the loading with a variable χ_{stress} that is specified by the author as the model-only uncertainty coefficient. In the author's words, this coefficient accounts for "bringing reality into a manageable mechanical model, bringing the mechanical model into a computational model, and the engineer's interpretation of results". Sørensen et al. [2008, 2011] use a standard deviation to describe the stress spectrum uncertainty. The stress ranges are assumed to be described by a Weibull model conditional on the variance component that is assumed to be perfectly known in the calculations (defined in terms of expected value).

Márquez-Domínguez and Sørensen [2012] do not consider uncertainty in the loading spectrum. Uncertainty related to loading uses two coefficients that relate to model-only uncertainty and application of the RC as described in Tarp-Johansen [2003]. Finally, Toft et al. [2016] also considers model uncertainty in the loading calculations using the procedure described by Tarp-Johansen [2003]. To mitigate the influence of the sample size in the calculations the authors perform use a large number of seeded simulations.

The following section discusses characterization of uncertainty using the initial set of simulations used for fatigue design. Implementation focuses on DLC1.2, but the same methodology can be implemented to other DLCs that concern fatigue design. The goal is then to produce an approach which is of simple implementation, which considers the current design practices; and that accurately deals with the uncertainty in the fatigue assessment during the design phase.

7.4 Uncertainty characterization in fatigue design loading.

Previous sections showed that the long-term design SN fatigue, applied to the tower component, may vary depending on the implementation procedure. Uncertainty in the case of fatigue is based on how much is not known about the mean of D_T when $T \to \infty$.

The designer approach to the fatigue calculations may be an important source of uncertainty in the fatigue design. In particular, when defining the loading spectra, the number of s_0 performed for loading characterization is a relevant cause for deviations in the estimation of D_T.

Comparison of the different techniques presented in the previous section showed that little improvement from scaling-up the simulation results is achieved by using statistical modelling techniques when s_0 is not adequate. Moreover, if not applied carefully, the statistical modelling may introduce further uncertainty in the design calculations (e.g. error in the fit). One of the main limitations found in the fatigue design procedure is that, to have convergence of the loading spectra, multiple simulations need to be performed in order to have accurate loading characterization.

The present section proposes a new, simple methodology, to calculate uncertainty on fatigue calculations, and which is applicable within the current design practices. It enables to quantify uncertainty due to s_0 at the design phase. The ultimate goal is to achieve robust extrapolation from t to T.

Calculation of confidence intervals (CI) of D_T using bootstrapping techniques is proposed in order to quantify uncertainty in the assessment of the real D_T. By using such approach it is possible to preclude eventual underestimations of the fatigue design damage when using a small s_0. Furthermore, for small m values, directly scaling up the SN damage obtained on a reference t results in relatively small errors in the approximation of D_T. It is highlighted that these are not always on the safe side.

Bootstrap has the advantage of being a non-parametric procedure to estimate CI. It is known that D_t in the present case follows a lognormal distribution. However, it is not always the case that an assumption about the SN fatigue loading distribution exists. For the lognormal distribution, techniques exist that allow to estimate the CI of the mean [Olsson, 2005].

Under the assumption that fatigue does not always follow a lognormal distribution, that D_t may not be the measure of interest to study, and that a limited number of simulations will be available to define s_0, it is possible to, instead of estimating D_T uncertainty through its lognormal characterization, estimate D_T uncertainty using bootstrapping of the loading spectrum.

In order to characterize uncertainty in D_t and consequently D_T, intervals of confidence for the t damage are computed using empirical bootstrapping [Efron et al., 1979], Block Bootstrapping (BB) [Carlstein, 1986] and Moving Block Bootstrapping (MBB) [Kunsch, 1989].

The last two variations of bootstrapping are widely applied in econometrics to compute uncertainty associated with values characterized by autocorrelated time-series data. BB and MBB allow estimation of variance within stationary time series. As the loading sequence is expected to be correlated in time, BB and MBB allow to enclose this correlation in the uncertainty calculations.

7.4.1 Bootstrapping techniques

The bootstrap is a primitive method to estimate uncertainty in a specified statistic when limited data is available. The most widely applied re-sampling procedure is the empirical bootstrap Efron et al. [1979]. The empirical bootstrap is remarkable for its efficiency and simplicity of application.

The empirical bootstrap procedure is defined as follows:

1 $L_1, L_2, \ldots, L_n$ is a sample from a population that follows a certain distribution F.

2 with a statistic computed from the sample, $\hat{q}$, and $\hat{F}$ the empirical distribution of data.

3 A re-sample (sample with replacement) $L_1^*, L_2^*, \ldots, L_n^*$ of the same size as the initial sample can be defined from $\hat{F}$.

4 Similarly, a new q^* statistic can be computed from $L_1^*, L_2^*, \ldots, L_n^*$.

5 Repetition of the previous steps allows to define a sample of the variability of q^*.

6 The bootstrap principle says that the variation of q^* approximates well the variation of q. Hence, the uncertainty of q can be approximated by the computed q^*.

Interest on bootstrapping for fatigue calculations has been previously identified [Bigerelle et al., 2006, da Costa and Sagrilo, 2018]. In the authors view, the attention that these techniques captivated for fatigue calculations has been limited when considering the potential they comprise. In particular for OWT calculations, the cost of the fatigue assessment has been recurrently highlighted as a limitation of the design calculations. Fatigue uncertainty assessment is hindered by this fact.

The concept of BB involves using a time window or block to divide the time series, Figure 7.8, and then to create new random time series based on the previously defined blocks and their bootstrap. Commonly, the time series is divided in blocks of length b

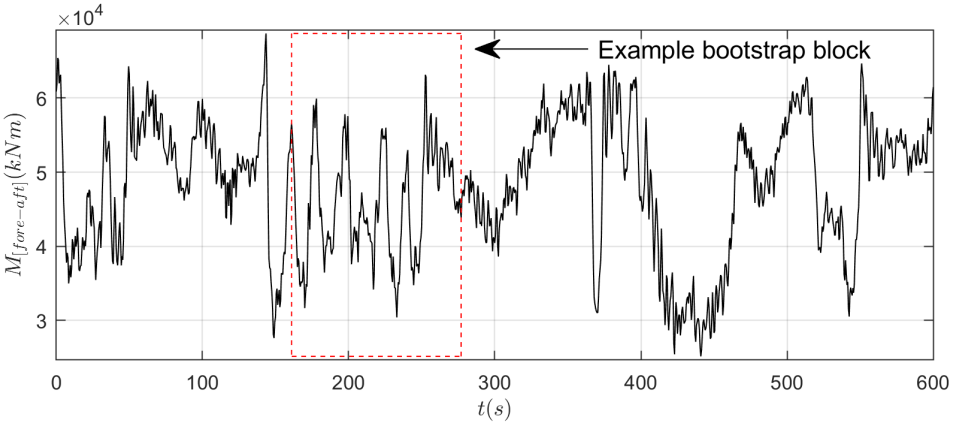


Figure 7.8: Example of block for bootstrapping

which can have a constant length or not, or can superimpose the neighbour block information or not. In the current case, BB is characterized by the definition of constant b and no-superimposition of blocks, while the MBB also considers a constant b, but with superimposition of blocks. Therefore, if $b = 30s$ in a 600s time-series; BB generates 20

blocks and MBB 571 blocks. The number of block in the MBB is $t + 1 - 30$ with t as the length of the time-series.

The basic requirement for using BB techniques is stationarity. A time series defined from time 1 to tf,

$$B = [b_1, b_2, \ldots, b_{tf}] \tag{7.3}$$

is described as stationary if it has the same statistical distribution as

$$B_h = [b_{1+h}, b_{2+h}, \ldots, b_{tf+h}] \tag{7.4}$$

for every integer h. Stationarity is guaranteed by three conditions:

$$I : E[b_t] = \mu \quad \forall t \tag{7.5}$$

$$II : Var[b_t] = \sigma^2 \quad \forall t \tag{7.6}$$

$$III : Cov[b_t, b_{t+h}] = f(h) \quad \forall t, h \tag{7.7}$$

In the case of OWT fatigue, the first two conditions are automatically derived from the fact that L is represented for different environmental conditions by a loading distribution. Moreover, the autocorrelation of L, was already shown to decay with the increase in h [Fogle et al., 2008]. As a result the L time series is expected to be stationary, with the implication that the correlation between x_t and x_{t+h} decays to 0 as points move away from each other in the time series, Equation 7.8.

$$Corr(b_t, b_{t+h}) = 0, for\ h \to \infty \tag{7.8}$$

The fact that t-series of L are expected to be stationary is corroborated by Figure 7.9 autocorrelation function which decays with the number of lags (indicator of a stationary time series). Figure 7.9 shows that loading time series stationary condition cannot be rejected accordingly to Kwiatkowski-Phillips-Schmidt-Shin(KPSS) test. Results are shown for simulation 5 (see Table 7.2), although similar values are exhibited by the other simulated cases. In some cases, the KPSS test may reject the null hypothesis for the t series. However, this is uncommon, and jointly with the reasons highlighted before, stationarity (weak-stationarity) is assumed to hold for all the cases.

When dividing the time series for BB, local L and L that occur within large variations of t, may be lost due to the procedure of the RC (rainflow counting) technique, which uses the so-called "roofs" to count cycles and ranges. da Costa and Sagrilo [2018] identified this limitation when BB for fatigue times series, and corrected it by extending the blocks in the t-domain to the next t where there is a up or down mean-crossing point. The RC is then applied to each independent block, allowing to characterize uncertainty while maintaining information about the correlation of L.

However, in the author understanding, extending to the next mean-crossing point may not be sufficient to avoid the deletion of relevant L. A single RC L can sometimes cross multiple "roofs" and, as a result, also multiple mean-crossing points. This occurs frequently, in particular for L that have important contribution to D_T. These large L are important for fatigue calculations.

In this thesis the RC is counted before the bootstrapping procedure and indexed to the respective blocks. Individually, the RC L may look random, however its sequence of L is only possible due to the related time series autocorrelation. The idea of using the BB and the MBB is then to enclose the temporal dependence of the loading data in the bootstrapping process and infer if there is any influence of it on the results obtained. If a

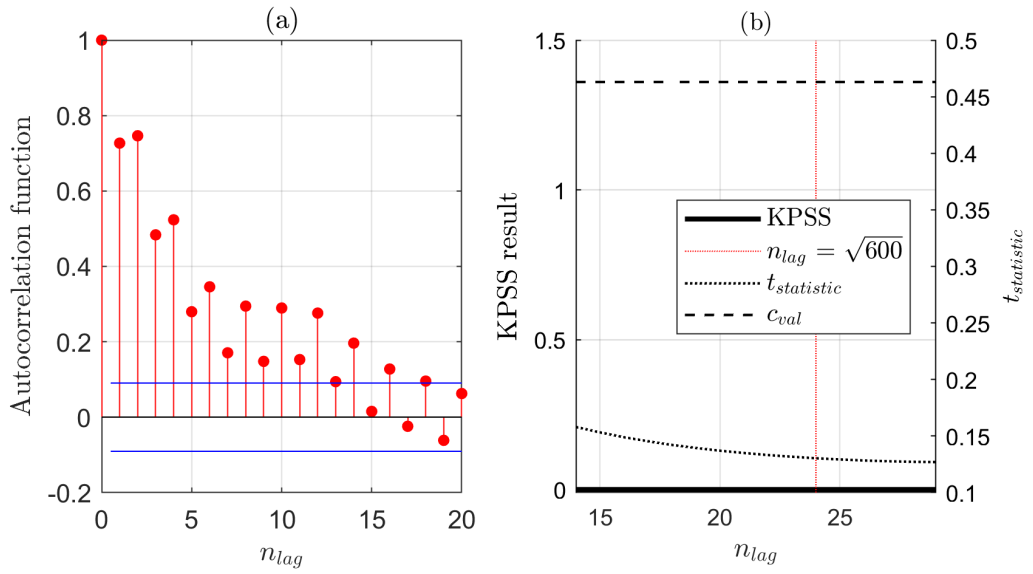


Figure 7.9: Autocorrelation function of the times series (mean trend removed) and results of stationarity hypothesis test with no trend using the Kwiatkowski-Phillips-Schmidt-Shin(KPSS) test. KPSS result is either 0 or 1 (0 means no rejection of null hypothesis that series is stationary); $t_{statistic}$ refers to the test statistic and c_{val} to its critical value; and n_{lag} refers to the number of lags used in the autoregressive models tested with KPSS.

large L crosses multiple blocks, then it is assumed to be part of its starting block in the time series.

The usage of these two time series bootstrapping techniques intends to address the fact the empirical bootstrapping disregards correlation between loads.

7.4.2 Results for confidence intervals in D_T

The concept of confidence interval (CI) of D_T means that; if in 90 out of 100 times an interval for a certain sample statistic includes its "true" value, then the estimated CI corresponds to a 90% confidence in the value of that statistic. It should not be confused with the variability of the studied statistic.

Calculation of CI using bootstrapping is achieved by repeating the bootstrapping procedure multiple times to guarantee convergence of q^*. In the present case the percentile method is applied to calculate CI. It consists in trimming the ordered bootstrap sample at α and $1-\alpha$ in order to set the $1-\alpha$ confidence intervals.

90% confidence intervals are applied for $m = 3$, $m = 5$, $m = 10$ and double $m = 3$ and $m = 5$ curves. The asymptotic convergence to the 90% CI was not well held in the case $m = 10$.

The following sections presents the results for the mentioned bootstrapping techniques, extrapolating these for comparison from a s_0 much smaller than the 1000 simulations used as a reference to compute the "true" statistic being studied. A reference operational state of high importance for the OWT was selected to compute the high fidelity D_T estimation. It considered the NTM, an $U = 11.5$, $H_s = 2.2$, and a $T_p = 8.2$.

Empirical bootstraping

Results for the empirical bootstrapping of the loading distribution are presented in Table 7.7 for different values of s_0. The number of times that in average 100 repetitions of the CI interval calculation did not enclose the true value of D_T were computed using 1000 repetitions of the CI calculation. These are presented in the table for the different s_0. Distinct convergence results for the bootstrap were attained for different SN curves.

For a $m = 5$ the empirical bootstrapping approximated accurately the 90% CI. For the double m curve, similar results were attained. This was expected due to its relation to m=5 curve, however the convergence was less remarkable due to the $m = 3$ slope. For a $m = 3$ the asymptotic results for the bootstrapping were not as accurate as for the m=5, but the results were conservative. Finally, for a $m = 10$ the asymptotic values for 90% CI were underestimated. For all the cases the intervals for a 90% CI resulted in an asymptotic value of 80% capability to enclose the true value of D_T.

Table 7.7: Number of times in average that the empirical bootstrap CI did not enclose the "true" mean in 100 repetitions of fatigue CI estimation. Values were computed using average from 10 repetitions of the 100 CI calculations, or 1000 repetitions. Empirical intervals are obtained after bootstrapping 600 samples.

s_0	$CI_{90}(m = 3)$	$CI_{90}(m = 5)$	$CI_{90}(m = 10)$	CI_{90} (*double m*)
6	5.5	10.6	24.9	9.5
7	4.6	10.1	23.3	8.9
8	4.7	9.6	22.0	8.5
9	3.6	9.5	23.0	9.2
10	2.3	9.6	18.5	8.8
15	1.6	12.0	23.0	7.6
20	2.0	10.5	20.5	10.9
25	0.9	7.3	16	4.2
30	0.5	6.8	15.7	3.8

One interesting feature of the CI results is their relative stability for different s_0 sample sizes. This is to some extent unexpected. As the size of the sample increases the bootstrap results were expected to improve. If the sample is too small, there may not be enough data to represent the CI for D_T. This is even more significant when the sample is not an accurate approximation of the population. In the case of high m, one of the possibilities is that the sample sizes studied are not large enough for bootstrap to be applied. The asymptotic value of the CI confidence slightly decreases with the increase of s_0. Also, as the sample size increases the CI interval decreases in range, as it will be shown in Section 7.4.2.

Regarding $m = 10$, convergence values for the CI_{90} resulted in an approximation of 0.8, or 20 failures in every 100 repetitions. This limit value slightly decreases with the increase of s_0. Sample size influence may be of importance in these results. It is known that bootstrapping is not a substitute of an adequate sample size. However, when the sample size is adequate the bootstrapped CI are a widely accepted procedure to compute uncertainty.

An alternative to improve the capability to comprise the true mean value of D_T in the CI for the case of m =10 is to take the full interval sampled from the bootstrap and

calculate the CI associated with the complete sample. The fact that, regardless of s_0, the CI converged to similar values of confidence indicates that if the whole bootstrapped sample is used to compute CI, it may be expected to converge to a constant CI confidence.

Figure 7.10 presents 100 computations of the CI for the considered m slopes. CI values were computed with a performance approximately similar to the one presented in Table 7.7. In the case of $m = 10$, the full bootstrap sample was considered for reference in the estimations of the CI . Using the whole sampled range, the CI was able to enclose 91 times the "true" value of D_T. However, depending on s_0, the range of the CI may be too large for an efficient application in the context of engineering design. It is noted that a s_0 =6 was used, which might be small for high m materials. The CI range is larger for large m values. Large intervals may not be adequate for efficient design implementation. In order to converge these interval values, the sample size can be increased. Interval reduction with the increase of the sample size may take advantage of the CI relatively stable convergence with s_0.

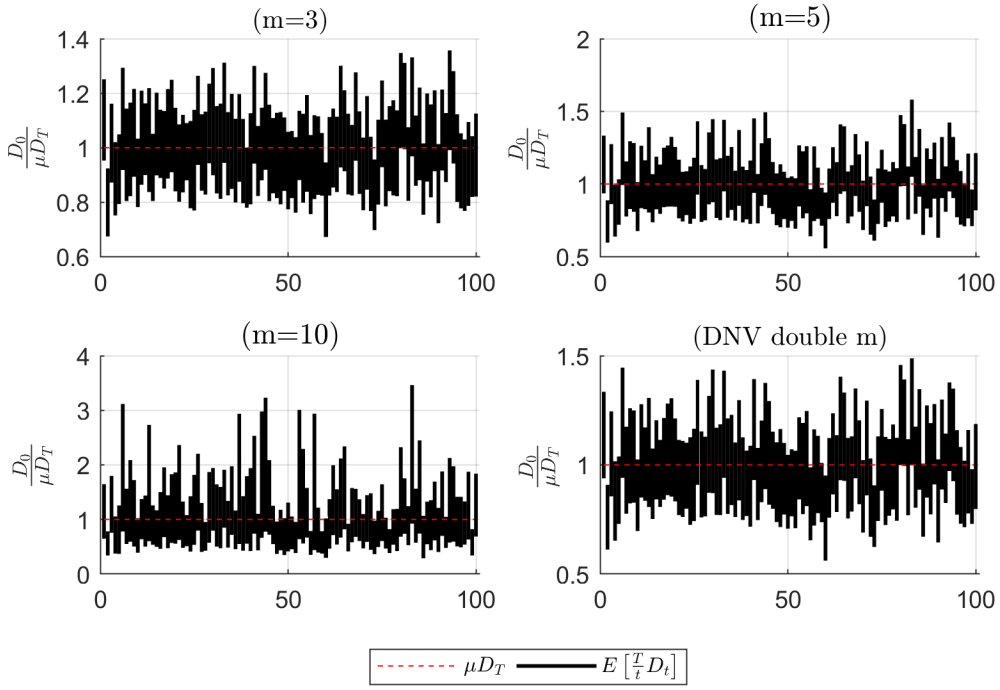


Figure 7.10: Bootstrapping approximation to mean value using different m slope values.

Block bootstraping

It is of interest to study the behaviour of the block bootstrap in comparison with the empirical bootstrap. The fact that whole blocks may be extracted and resampled using a variable b may contribute to better approximate asymptotic convergence of the results from the empirical bootstrapping.

Table 7.8 presents the results of confidence estimation in D_T for different values of b, using $s_0 = 6$, to calculate the confidence in the estimation of fatigue for 1000 simulations.

When comparing to a s_0 of 6, 1000 simulations is equivalent to extrapolating the damage that occurred in 1 hour to approximately 167 hours.

Table 7.8: Results for the number of times the CI does not enclose the "true" mean in 100 simulations using BB, and considering different values of b. Values were computed using average from 10 repetitions of 100 CI calculations and a $s_0 = 6$.

	BB				*MBB*			
$b(s)$	CI_{90} $(m = 3)$	CI_{90} $m = 5$	CI_{90} $m = 10$	CI_{90} *double* m	CI_{90} $m = 3$	CI_{90} $m = 5$	CI_{90} $m = 10$	CI_{90} *double* m
30	5.6	8.8	22.6	6.2	4.9	8.1	20.5	6.6
60	5.3	7.6	20.0	7.1	4.9	7.4	20.4	5.6
75	4.6	8.3	19.6	6.3	4.2	7.7	17.6	7.1
100	6	10.1	18.4	8.5	5.6	10.8	20.8	9.3
120	8.6	12.9	22.1	11.6	10.3	14.5	23.0	12.0
200	2.3	3.7	18	3.5	7.5	12.0	25.5	10.0
300	5.5	6.8	18.9	5.9	8.9	10.1	13.5	9.5

Both the usage of BB and MBB produce accurate results when estimating D_T confidence for $m = 3$, $m = 5$, and double m. When comparing with the empirical bootstrapping, the BB and MBB converged into slightly more accurate CI intervals for $m = 3$. Moreover, the capability of using an extra degree of freedom given by the parameter b, may help to calibrate the results to the application in-hands. Applying large block lengths for $m = 3$ may be an alternative to achieve better convergence of the CI_{90}.

Similarly to the empirical bootstrapping, asymptotic results for $m = 10$ underestimate the level of confidence in the CI. The performance of the MBB when using a b equal to 300s is remarkable. In this case the MBB was able to give accurate predictions of the interval for all the m slopes.

When the b value increases, the CI reference values changes more significantly for MBB than for the BB. The capability to calibrate the CI for the application is particularly notorious in the MBB case. Its overlapping character allows the usage of more flexible values of b. Large b values combined with the MBB resulted in the best CI estimations for $m = 3$ and $m = 10$.

When comparing with the empirical bootstrapping, there is little gain in introducing complexity in the bootstrapping analysis. This may be originated by the fact that, despite being related to the time-series, the RC series has a purely random character. This is evident for small values of b. The autocorrelation of the RC loading series decreases to 0 after the first or second lag value.

Nonetheless, it was shown that accurate confidence intervals for D_T can be estimated using bootstrapping techniques. Both, the independent and the dependent bootstrapping techniques achieve accurate predictions of CI for low m materials. The main limitation of the bootstrapping techniques researched is their computational cost, to converge the CI more than 20 hours are required using a 20 cores Intel Xeon computer with 64GB of RAM.

One alternative to mitigate the effort to compute CI for all x, which can be significant due to the number of possible environmental states, is to calculate CIs for the x that are expected to comprise the largest contributions to decrement the system's fatigue life.

Increase of s_0

If s_0 is increased, despite maintaining the asymptotic value for the confidence of the estimation of the CI, the range of values of the CI becomes narrower. Figure 7.11 presents the convergence of damage confidence with the increase of s_0.

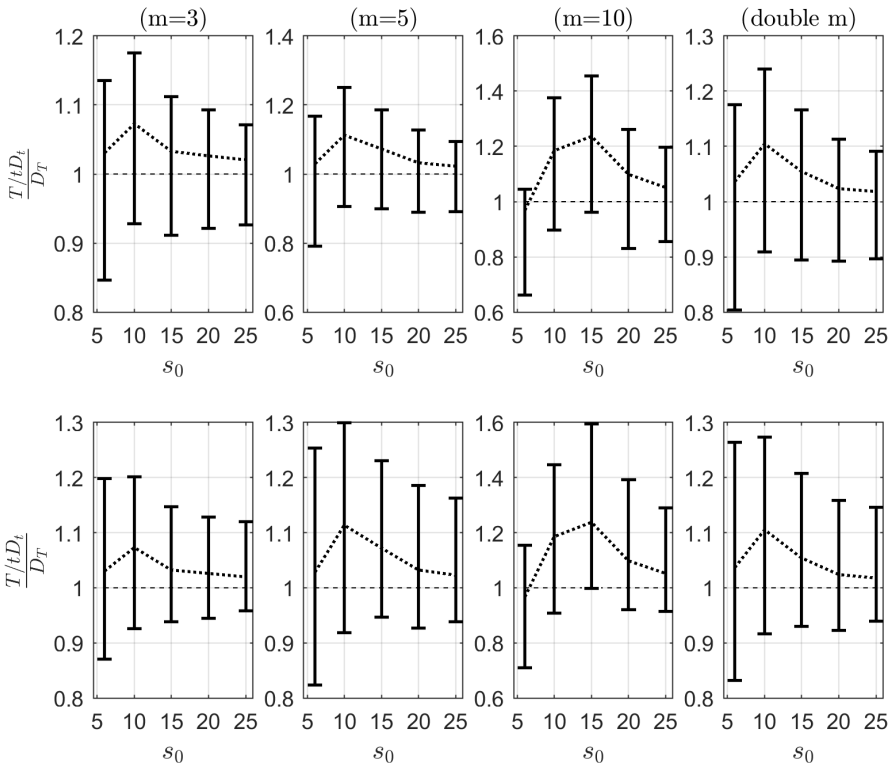


Figure 7.11: Bootstrap confidence intervals considering a $b = 60s$ and different s_0 and comparison in approaching a T equivalent to 1000 simulations of $t = 10min$.

The decrease in the range of the interval may be one of the explanations for the stability of the CI with the increase of s_0. In fact, stability with s_0 is of interest. This allows the increment of s_0 and the reduction of the range of the interval of confidence in D_T, while maintaining a constant level of confidence in the results. For small s_0 the CI may be significantly wide, and therefore, unacceptable in a designing context. Increasing s_0 may mitigate this limitation and allow a narrower range of potential values for D_T, and hence, narrower uncertainty within the same confidence level.

The decision to increase the precision in the estimation of uncertainty for D_T should be weighted along with the increase in the computational effort required.

Increase of s_0 may result in a reduction of the CI interval by more than 50%. This reduction can be identified in Figure 7.11 for cases m=3, m=5 and double m. For the case of m=10, this reduction is not so evident. To note that for these materials, the sample convergence increases slower with the size of s_0. Moreover, for such materials it may be important to have extrapolation techniques. It is of interest to research on the results in the case where the bootstrapping is applied jointly with extrapolation of loads and cycles. In such a scenario, CI may to converge asymptotically to a higher CI confidence in the case of large m.

Block Bootstrap different operational points

BB and MBB results are computed for the simulated cases of Table 7.3 shown in the Appendix Table A.3. The results for each BB and MBB are computed using 3 computations of the confidence interval of D_T, and applying a random s_0 of six points. In this case only 100 simulations were used to compute D_T. For relatively small m, 100 simulation at each x are expected to hold accurate estimation of mean value.

For almost all the cases, the BB techniques predict well the range of values that are expected to include D_T, creating accurate confidence bounds on the uncertainty resulting from the SN fatigue estimations.

The definition of b may be more efficient as a function of the operational environmental conditions. This possibility may be researched in further works. It was seen that different operational regimes may have different loading characteristics (*e.g.*, simulations 2-4). In some circumstances the SN damage density is confined to the tail region. This is a situation that resembles the case of large m, and as such the bootstrapping approximation may generate inaccurate results.

Further developments of the approach presented may be also achieved by increasing the complexity of the bootstrapping procedures. Accounting for bias, comparing with alternative computations of the CI through bootstrapping, where *e.g.* t-bootstrap or BC_a bootstrap (see [DiCiccio and Efron, 1996]) are just examples of different analysis that can be performed to set CI of D_T. The goal of the current calculation was to show how bootstrapping can be applied as an efficient estimator of uncertainty for the OWT tower component SN fatigue design.Such an approach may allow the improvement of the design of the tower components. Nominatively, by reducing the conservationism of the design while maintain a comprehensive uncertainty characterization on the calculations. Therefore, allowing the reduction of the designing costs while producing robust designs.

7.5 Conclusion

A comparative study on fatigue design of OWT with the SN method, with particular application to the tower component, was presented. The interest was to study the fatigue calculations for different SN slope materials. The SN method, despite its limitations, is still the most widely used fatigue design method applied in the design phase. IEC design Load Case 1.2 was considered in the calculations.

Five fatigue calculation methods were compared; directly scaling up the obtained cycles, approximation of loading spectra with a Weibull distribution using two fitting techniques, and truncation at 95% and 99% quantiles (recommended by IEC61400) using the Weibull and the Generalised Pareto distributions to predict large load range occurrences. Four SN material slopes were used to compute results.

From the methodologies tested, it is highlighted:

- Limited interest was found in fitting the complete data set with a theoretical probability model. Low range load cycles occur more often than large load ranges, distorting the fitting of the tail. This distortion can be mitigated with different techniques, such as weighted regression, that are sometimes applied when the "body" of the distribution is of interest to describe the tail trend. This procedure may be of interest if a parametric model is required to address uncertainty, such as in reliability calculations. Nonetheless, the procedure of fitting is an important source of uncertainty, in particular, if considerable SN damage density exists in the probability tail region.

When an adequate threshold is applied, truncation and fitting of the tail produces more accurate approximations.

- Truncation at 95% was identified to frequently overestimate the expected fatigue damage. Truncation at 99% showed better concordance in approaching the real value of fatigue, however the problem of extrapolation of load and cycles is strongly related to the accuracy of defining, not only the loads, but also, the expected cycles in the tail region. All estimates were identified to be bounded in accuracy to the number of initial simulations performed to analyse the fatigue design.

- For low m materials, limited gain in accuracy was achieved by introducing complexity in the analysis procedure (*e.g.* statistical techniques such as fitting models or extrapolation) when comparing with the direct scaling up of results from the initial set of simulations. Nevertheless, directly scaling up the results is a deterministic approach and, as such, has inherent uncertainty. This fact motivated the development of a long-term SN fatigue design uncertainty characterization technique.

A new approach to quantify uncertainty in the long-term fatigue estimations was presented and discussed. It consists of using bootstrapping techniques to establish confidence intervals of the long-term SN fatigue damage mean value. Empirical bootstrap was compared with block bootstrap. The last intends to consider the autocorrelation in the loading times series. Due to the complexity of the rainflow-counting algorithm, it is recommended for the block bootstrapping to be indexed to the time-domain loading. Block bootstrapping was implemented before for fatigue calculations by dividing the time series, [da Costa and Sagrilo, 2018]. However, this procedure may eliminate important loading cycles that occur within multiple blocks.

Time series block-bootstrapping techniques did not contribute for a significant improvement in the accuracy of the results when compared with the less complex empirical bootstrapping. Despite the strong autocorrelation in the loading, RC extracted load ranges can be seen as purely random. Nonetheless, asymptotic convergence for the 90% CI was improved by the block bootstrapping. The empirical bootstrap is an adequate procedure to study uncertainty in D_T estimates for low m materials. However, the asymptotic convergence is not as good as in the case of block bootstrapping, but only slightly.

The bootstrap does not suffice an efficient substitute of an adequate sample to estimate D_T. Nonetheless, it is a simple non-parametric tool to estimate uncertainty which holds efficient prediction of the long-term SN fatigue damage confidence in the design phase for the SN materials with slopes of $m = 3$, $m = 5$ and double m. In both cases, robust CI were attained. For m =10 the convergence results for the CI indicated that the 90% CI were underestimated. More failure in covering the true value of D_T than expected for the CI_{90} were encountered. A conservative alternative to tackle this limitation is to take the full bootstrap sample as the CI and estimate the convergence value for it. For an m =10, the full bootstrapped sample converges to a CI of 95% for the tower component of the considered turbine.

To conclude, although uncertainty in D_T estimations is larger for high m, the bootstrapping technique used has more robust applicability for lower m. For the last, CI were very accurate and relatively narrow. In the case of large m values, and low loading samples, the CI intervals may be too large for consideration in the design phase. This can be solved by using larger sample sizes. Future works should consider research on the performance of the bootstrapping when combined with the extrapolation of loads and cycles

for high m materials. Alternatively, other bootstrapping techniques may be researched in order to improve the identified limitations.

Chapter 8

Meta-modelling in the design of OWT to fatigue

Contents

The analysis presented in the following chapter was published in Teixeira et al. [2019e,c].

8.1 Motivation and original contributions of this chapter

The procedures proposed in the following chapter aim to contribute to the optimization of the SN fatigue design for OWT components. Chapters 4, 5 and 7 recurrently highlighted the high cost associated with operational fatigue design calculations. The need to define accurately the loading spectra that characterize fatigue damage was identified as a time consuming constraint for the design process. Fatigue design procedures demand

the assessment of multiple load case scenarios that increase exponentially with the number of variables considered in the problem. It was seen that the fact that OWT fatigue calculations involve a large set of environmental parameters motivated the development of different techniques to reduce the computational burden of the fatigue assessment.

As the dimension of the environmental variables space increase, the number of load cases for fatigue assessment increases exponentially. This fact limits the application of optimization techniques that consider the full assessment of the SN structural fatigue. Considering wave states as conditional of the wind parameters is an example of assumptions undertaken to make the fatigue analysis problem manageable.

Meta-modelling is implemented in the current chapter as an alternative for efficient fatigue assessment. Meta-modelling has the potential to replace the computationally expensive OWT model and enable a fast and accurate fatigue assessment. In particular, the application of Gaussian process predictors is studied to surrogate fatigue damage. These models have proven before to be efficient surrogates of complex engineering problems.

The motivation of the research presented is to develop an innovative and efficient fatigue design approach, using a surrogate surface of simple application, that uses the same procedure recommended by IEC61400-3 for the fatigue design calculations. DLC1.2 calculations using meta-modelling techniques are considered in the research proposed. It was seen that the cost to design OWT to fatigue limits the amount of effort that can be spent to improve it in the design phase. The ultimate goal of the research presented is, therefore, to produce an innovative robust methodology that allows to decrease the effort required to design OWT towers to fatigue, investigating the potential of meta-modelling techniques within the current design practices.

8.2 Introduction

Meta-modelling was identified as a potential alternative to mitigate the computational burden of the SN fatigue design procedure. In particular, the Gaussian process predictors, or Kriging models, have presented before the capability to efficiently interpolate the response of complex systems. These models merge the classical regression properties with a local measure of uncertainty, which makes them highly flexible in terms of interpolation. The Gaussian term of $G(x)$ guarantees that a prediction is exactly $g(X)$ at X, even with low order models. In the unknown x space the interpolation has Gaussian uncertainty.

In order to have efficient meta-modelling, it is of importance to understand the problem in-hand. If computational burden is of relevance, the analysis should focus on the most important variables for the assessment. It is common for a limited number of input variables to comprise most of the influence on the output of a certain problem [Saltelli et al., 2008]. If efficient implementations are pursued, it is of major importance to focus and reduce the dimension of the input space in order to focus on the variables for which the output has larger sensitivity.

In order to discuss the potential application of $G(x)$ as a surrogate of D_t, Section 8.3 researches the dimensionality of the problem of SN fatigue design for the tower component. Section 8.4 presents the results for the implementation of the SN damage surfaces using $G(x)$. Due to its significant relevance, selection of the DoE is researched. It is common practice to use a LHS scheme to define $\boldsymbol{X}$. However, as the DoE is defined in a single step with the LHS, there is limited notion of improvement in the surrogate approximation. A new infill criterion is therefore proposed in order to have a more comprehensive and robust surrogate approximation. This criterion discussed in Section 8.5.2. Comparative

results for its implementation are given in Section 8.5.3. The results obtained using the LHS, the proposed criterion and the widely applied binning of x are compared with a full year loading assessment for validation. Sections 8.5.4 and 8.5.5 discuss respectively the convergence of the mean and how intervals of uncertainty for D_T can be defined taking advantage of the interpolating capability of the predictor. Finally, the main conclusions of the chapter are presented in Section 8.6.

Before further developments, it is then of interest to study the dimensional variables of SN fatigue assessment for the tower component. In order to have optimum meta-modelling implementations, it is of interest to reduce the design of experiments to the variables of major influence in the problem studied.

8.3 Problem space reduction

As mentioned, in order to have efficient meta-modelling it is of interest to have a global comprehension of the problem in-hand. Such understanding may allow to the efficient allocation of resources during the analysis in the factors of major importance. In the case of complex problems, frequently, if the analysis is not reduced to the variables of major importance, the procedure may become unnecessarily excessive in terms of complexity. If the goal is to improve computational time, then it is mandatory to efficiently cover the DoE.

The problem of dimensional reduction can be achieved performing a sensitivity analysis. One of the most adequate ways to characterize a system's operational behaviour is by conducting a sensitivity analysis, which can be attained by applying efficient Global Sensitivity Analysis (GSA) techniques.

A GSA enables the designer to increase his or her awareness about the operation of a specified system. In the case of OWT, as complex systems, a GSA allows the coverage of its full-field operation. If the goal is to comprehend the system's operational response, a measure of uncertainty needs to be enclosed in the sensitivity analysis. In such a scenario, a probabilistic GSA (P-GSA) may be considered. Saltelli et al. [2008] states that uncertainty quantification and sensitivity analysis, in its more broad understanding of identifying the most relevant variables in the DoE, run *in tandem*.

An efficient strategy to solve the problem of GSA for complex systems is to build importance measures that effectively quantify sensitivity of the output variables in regard of changes in the input field [Caniou, 2012]. For SN fatigue of OWT towers, the analysis demands the creation of an importance measure that reduces the structural fatigue response to shorter time-scales for which the analysis can be underpinned on a feasible basis.

8.3.1 Global sensitivity analysis

Two main initial challenges are encountered when performing a GSA analysis for OWT. The first is related to the difficulty of identifying the range in the space of variables where the sensitivity must be assessed. When analysing a high dimensional space, the design space can become unreasonably high and seriously compromise the assessment. In the case of OWT systems the computational time to characterize the system's fatigue is a limitation. When dealing with highly non-linear problems, whose evaluation commonly requires high computational time, it is not feasible to cover the design space and set the design of experiments (DoE) without an established criteria. The importance of a

balanced approach is discussed in Saltelli et al. [2008], where full insight on how to select the DoE for a GSA is presented.

Sampling techniques described in Chapter 5 are commonly applied to cover the DoE in the context of a GSA. It was seen that LHS is efficient for resource consuming models as it allows covering efficiently the DoE with a relatively small sample size. LHS divides the DoE in a grid with $l > 2$ levels and samples the DoE with the guarantee that a minimum number of samples are taken at each equal probability density division for each variable. A review of the main advantages of the LHS for analysis of multi-scale systems with discussion of its applicability for sensitivity analysis and usefulness to map the DoE can be found in Helton and Davis [2003]. In the specific case of reliability problems, Olsson et al. [2003] discusses its interest as a sampling method. Even considering that LHS introduces a more balanced approach to map the DoE, for complex problems the computational effort can remain quite demanding. Additionally, it is not the perfect approach when focusing on less common occurrences, given the low probabilities cumulated under the tails. Improvements of the LHS use specific criteria to balance the way the points are picked in the probability distributions [Viana et al., 2010].

The second challenge is related to the inherent stochasticity of some of the problems analysed. The OWT tower component's fatigue response is characterized by its loading spectra. Its the case of a model that has inherent random behaviour and where the same set of x generates a non-deterministic output response, and consequently SN fatigue. The focus of the sensitivity analysis should consider the complete statistical behaviour of the output.

A degree of quantification of uncertainty can be achieved with variance-based methods like the ones introduced in Efron and Stein [1981] or Sobol [1993], but as their name states, these are based on the analysis of variance. In the present case the major interest is to infer on how the SN damage varies within the x variables, with the mean having crucial importance. This implies the need to develop a GSA supported by distribution based methods.

The variance assessment may then complement the GSA analysis uncertainty in the measure of interest (mean value). The previous chapter showed that convergence of the mean may demand a very high s_0. Therefore, usage of a distribution method indicator to perform the sensitivity analysis allows us to check the variation of the mean value within x considering a relatively practicable s_0 for each x point. In such a case, the uncertainty in the μ estimate is given by σ.

The use of distribution-based methods to analyse the sensitivity is not a common practice. Nevertheless, different authors identified this necessity and proposed measures to quantify the disparity between two distributions [Kullback and Leibler, 1951, Borgonovo, 2007].

Assuming two Probability Density Functions (PDFs) of an output variable $Y(x)$, $f_Y(x)$ and $f_{Y^*}(x)$, for both, $f_Y(x)$ and $f_{Y^*}(x)$, the Kullback-Leibler (KL) measure of discrimination assesses the relative entropy between the two PDFs and is defined as

$$D_{KL}(f_Y(x)||f_{Y^*}(x)) = \int_x f_Y(x) \ln \frac{f_Y(x)}{f_{Y^*}(x)} dx, \tag{8.1}$$

taking the following form if $f_Y(x)$ and $f_{Y^*}(x)$ are univariate normal distributions with

mean μ_f and μ_{f^*}, and standard deviation σ_f and σ_{f^*} respectively;

$$D_{KL}(f_Y(x)||f_{Y^*}(x)) = \ln\left(\frac{\sigma_f^*}{\sigma_f}\right) + \frac{\sigma_f^2 + (\mu_f - \mu_{f^*})^2}{2\sigma_{f^*}^2} - \frac{1}{2}$$

$$with,\ \mu_f, \mu_f^* \in \mathbb{R};\quad \sigma_f^2, \sigma_f^{*2} \in \mathbb{R}^+ \quad (8.2)$$

The discrimination is a non-symmetric measure of distance between two distributions. The symmetric variant of the discrimination is called divergence J and was defined in Kullback [1997]. In the case of J the discrimination measure is analysed both from $f_Y(x)$ to $f_{Y^*}(x)$, and from $f_{Y^*}(x)$ to $f_Y(x)$.

The idea of using these divergence and discrimination indicators to analyse the sensitivity emerged first in Park and Ahn [1994], where the KL discriminator is used to assess and compare probabilistic sensitivity results. Liu et al. [2006] apply the same measure justifying its application by its capability of enclosing more uncertainty information. Greegar and Manohar [2016] use the KL in a comparison of sensitivity indices. Balesdent et al. [2013] use the divergence as a cross-entropy sensitivity search indicator within a DoE. Hoseyni et al. [2015] apply the same measure as an indicator of sensitivity in a finite mixture models analysis.

D_{KL} is non-monotonic and always positive. It is also symmetric in relation to changes in μ. As it quantifies the discrepancy between probability distribution, in regard of variations in σ it is non-symmetric and has a vertical asymptotic behaviour when $\sigma_{f^*} \to 0$. The latter can impose problems of stability in the calculations.

The sensitivity analysis for an OWT component to SN structural fatigue in the framework described requires the definition of a design "damage indicator" (Di). The definition of a representative Di can be challenging. Similarly to how different related approaches can be found to assess fatigue, stress-based, strain-based and fracture mechanics, the same occurs in the definition of Dis. Mesmacque et al. [2005] highlight the difficulty of defining a representative Di for fatigue calculations. The GSA presented uses D_t as Di for comparison of the results. D_t is a straightforward approximation to quantify the SN long term behaviour using shorter time scales. Moreover, it is aligned with the work presented in the previous chapters. DEL could be applied as alternative.

It was seen that D_t for the tower component depends on a set of parameters $x = [x_1, x_2, \ldots, x_{\check{k}}]$ that influence the OWT loading. Here, x is expressed as a row vector of length $\check{k}$ containing x_i environmental variables that can take multiple values. The role of a sensitivity analysis is then to identify which of these $\check{k}$ parameters are the most relevant for the OWT tower long-term fatigue, as well as to characterize how these affect the uncertainty of the measure of interest. In the case of $D_t(x)$ every singular combination of $x_{\check{k}}$ parameters will originate a conditional response.

GSA Design of Experiments

Five environmental random variables were assumed to be the main contributors for OWT tower loading, $e = 5$, wind speed (U), significant wave height (H_s), wave peak period (T_p), Turbulence intensity (I) and wind direction (ϕ_w) or misalignment with rotor axis.

To cover the DoE, recorded data from the M6 buoy presented in Chapter 6 was used to ensure a realistic approximation to $f(x)$. Sampling resolution at U values above rated power was increased to guarantee more sampling points above U=11.4m/s. The distributions for each input variable are shown in Table 8.1 and are used to model the different input variables. The power law was applied to correct U for the reference height of the

turbine, 87.6 meters. A LHS was implemented to select the $\boldsymbol{X}$ points where to assess

Table 8.1: Random variables considered in the LHS and respective distributions

Random Variable	Statistical Distribution	Relevant Spearman correlation (r_s)
$U(m/s)$	Weibull Distribution	$r_s(U, H_s) = 0.66;\ r_s(U, T_p) = 0.27$
$H_s(m)$	Weibull Distribution	$r_s(U, H_s) = 0.66;\ r_s(H_s, T_p) = 0.79$
$T_p(s)$	Rayleigh Distribution	$r_s(U, T_p) = 0.27;\ r_s(H_s, T_p) = 0.79$
$I(\%)$	Normal Distribution	$r_s(U, I) = -0.30$
$\phi_w(\circ)$	Uniform Distribution	-

fatigue. Ideally a full factorial analysis [Saltelli et al., 2008] should be carried to cover all the DoE, however, the computational time to perform it when the output is not deterministic is unbearable. The LHS is of interest due to its capability of covering the DoE (considering $f(x)$ probabilities) with a balanced amount of computational time.

It is important to understand that some of the variables are correlated and this influences the implementation of the LHS. As the interest is the analysis of SN fatigue, it is relevant to consider correlation within $f(x)$. It is known that there will be a limited amount of $f(x)$ that will contribute to all the SN fatigue life decrement.

Spearman correlation coefficients between the different variables are presented in Table 8.1. The coefficients not presented in the table had a weak correlation and for such case the variables are assumed to be independent. In the present case no data was available for I, therefore the variable I was assumed to be correlated to U and Gaussian distributed with direct correlation to the wind speed accordingly to the data presented in Türk and Emeis [2010], which is representative of an offshore scenario. Monte Carlo simulations were used to define the conditional distribution of I and its correlation to U. This was achieved by generating samples of I conditional on the distribution of U.

Applying the results from the FINO1 [1] platform may introduce some degree of error as they may not be representative of the Irish Atlantic Sea, nevertheless the results are considered to suffice the research undertaken in the present work. Assuming these values instead of the turbulence given by the Normal Turbulence Model(NTM) [IEC, 2005] is expected to generate more accurate results as the NTM tends to overestimate predictions of I for offshore conditions.

The correlation that exists between the x variables increases the complexity of the LHS implementation. The methodology introduced in Iman and Conover [1982] to sample correlated variables is applied to account for the correlation within x. In this methodology, two lower triangular matrices P and Q are obtained using the Cholesky factorization of the Spearman correlation matrix with the target correlations $C = PP'$ and the correlation matrix of the randomly assessed LHS, $T = QQ'$. The transformation is then applied with

$$H^* = H\ S^t \tag{8.3}$$

$$S = PQ^{-1} \tag{8.4}$$

The transformed H^* of the initial H matrix (initial sample in the present case) has a correlation matrix equal to the imposed C. The final sample preserves the univariate distributions of the original LHS. t in this case stands for the transpose.

[1] FINO1 research platform, located in the North Sea to the north of Borkum.

The $\phi_p(D)$ distance criteria introduced in Morris and Mitchell [1995] was used to validate the correlated and non-correlated samples and optimize different DoEs. A non-relevant variation, less than 10% of the original LHS value, in this indicator resulted from creating the correlated LHS.

DoE sample size

Choosing the ideal sample size to perform a GSA is challenging. A minimum sample size of $n+1$ is usually assumed for LHS, but few recommendations are found on the adequate sample size. Iman and Helton [1985] have defined a sample size of $4/3n$ as satisfactory to cover the DoE. Manache and Melching [2008] refers that this measure was compared with $3n$ to test its robustness and positive results of convergence were achieved. Nevertheless many authors do not strictly follow the rule presented in Iman and Helton [1985] and select the sample size which they believe that will produce efficient results. Stein [1987] uses a sample size $l = 100$ for $n = 6$. Helton et al. [2005] uses the same sample size for $n = 31$.

A LHS of size $l = 25$ was selected to perform the GSA. The criteria to select l was to achieve adequate resolution in the variable of the space that is expected to cover a wider physical variation in absolute values, that is, U. Considering a larger value of l may help mitigate some of the limitations of the LHS. The probabilistic space of the variables was divided in $l = 25$ equivalent probabilities. Only values that occur during the operation of the OWT, considering the behaviour of its control system, were considered (e.g. between the cut-in and cut-out wind speeds).

$D_t(x)$ is assumed to follow a log-normal distribution $\forall x$. This consideration is important as $D_{KL}(x)$ can be represented in a closed mathematical form for the case where the statistical moments are defined.

8.3.2 GSA results

It was mentioned that the interest of mapping globally the DoE of an OWT in regard of the D_t was; to identify regimes of operation that contribute for the long-term degradation of the wind turbine; to characterize uncertainty within different areas of operation and; to infer on the most critical variables of the DoE. All the three factors converge in the increasing requirement for accurate reliable designs.

Having the perception of major areas where there is operational risk is not only relevant to grant the designer a better understanding during the design phase of how operation fatigue will cumulate with time, but it can be also relevant on a steady health monitoring basis to update risk profiles of operation. It is complementarity of the GSA procedure.

Transformed sensitivity indicator

A sensitivity analysis to characterize an engineering design variable should produce perceptive results. This means that the results need to be intuitive for the designer and enable her or him to easily draw conclusions.

For some physical quantities the interest is not only to characterize how one distribution behaves in terms of statistical sensitivity, but also to characterize the physical process in regard of its practical effect. For some physical variables, an increase in the standard deviation may result in a decrease of the statistical distance to a reference distribution. However, in terms of practical effect (*e.g.* reliability problem), this increase may represent

a more severe operational condition and therefore should be characterized in terms of sensitivity as that.

D_{KL} characterizes the entropy between two distributions. When using it to characterize the distance between two distributions, it is difficult to infer on the relative location of the sensitivity point on the probabilistic map without further operations.

Consider $\mathbb{R}_{\mu^+}$, $\mathbb{R}^+_{\sigma^+}$ and $\mathbb{R}_{\mu^-}$, $\mathbb{R}^+_{\sigma^-}$ as separate domains of the KL function where the change in mean and standard deviation are positive and negative respectively.

While the KL function is symmetric between both domains, μ^+ and μ^-, the same is not verified in the domains σ^+ and σ^-. Similarly, the KL function is monotonic increasing in both domains of $\mathbb{R}_{\mu}$ and $\mathbb{R}^+_{\sigma^-}$, but non-monotonic in $\mathbb{R}^+_{\sigma^+}$.

For engineering application purposes it may be of interest to have KL as a symmetric and monotonic function in all the division domains introduced in order to have a more perceptive description of the global sensitivity analysis. Despite the KL function being a feasible solution to analyse GSA that was applied in previous works, its original function shape is not very interesting for mapping GSA results in practical examples that need to consider uncertainty.

A transformation to the KL discrimination function is proposed in this thesis in order to enable it to be an efficient indicator of GSA. This transformed function adds complementary information on sensitivity that is of interest for practical applications. To ensure a monotonic and symmetric behaviour in all the domains identified, the KL function should be modified in $\mathbb{R}^+_{\sigma^+}$ using the following transformations h_I and h_{II},

$$h_I : D_{KL} \rightarrow D^I_{KL}, \quad \forall \mu_f \neq \mu_{f^*} \ \cap \ \sigma_{f^*} \in \mathbb{R}^+_{\sigma^+} \tag{8.5}$$

$$h_I = \begin{cases} 2\frac{D_{KL}(\sigma_{f^*}=\sigma_f)}{D_{KL}(\sigma_{f^*})} - 1 & \sigma_f < \sigma_{f^*} \leq \sigma_{f^*_{min}} \\ 2\frac{|D_{KL}(\sigma_{f^*}=\sigma_f) - D_{KL}(\sigma_{f^*}=\sigma_{f^*_{min}})|}{D_{KL}(\sigma_{f^*})} + 1 & \sigma_{f^*} > \sigma_{f^*_{min}} \end{cases} \tag{8.6}$$

with $\sigma_{f^*_{min}}$ being the value of σ^*_f for the respective μ_f and μ^*_f for which the D_{KL} is minimum. Then h_{II} is applied accordingly to,

$$h_{II} : D^I_{KL} \rightarrow D^{II}_{KL}, \quad \forall \mu_f, \mu_{f^*} \ \cap \ \sigma_{f^*} \in \mathbb{R}^+_{\sigma^-} \tag{8.7}$$

$$h_{II} = \frac{D^I_{KL}(\sigma_f + |\sigma_f - \sigma_{f^*}|)}{D_{KL}(\sigma_{f^*})} \qquad 0 < \sigma_{f^*} < \sigma_f \tag{8.8}$$

along the remaining domains, the partial transformations h_I and h_{II} will have the value of one unit as no transformation is expected,

$$h_I = 1, \quad \forall \mu_f, \mu_{f^*} \ \cap \ \left(\sigma_{f^*} \in \mathbb{R}^+_{\sigma^-} \ \cup \ \sigma_{f^*} = \sigma_f\right), \tag{8.9}$$

$$h_{II} = 1 \quad \forall \mu_f, \mu_{f^*} \ \cap \ \left(\sigma_{f^*} \in \mathbb{R}^+_{\sigma^+} \ \cup \ \sigma_{f^*} = \sigma_f\right), \tag{8.10}$$

This transformation could be implemented through a single transformation using

$$[h(I - II) = h_I h_{II}] : D_{KL} \rightarrow D^{II}_{KL} \tag{8.11}$$

respecting the domains of application presented and with respective modification of the domains of the transformation in order to guarantee symmetry and monotonicity in all the domains. $h(I - II)$ is a global transformation function while h_I and h_{II} are local

transformations functions as presented in Equations (8.5 - 8.8). The division in a two step function intends to facilitate the understanding of the proposed implementation.

Figure 8.1 presents the original KL discrimination function, transformation, and the transformed functions for both the cases where $\mu_f = \mu_{f^*}$ and $\mu_f \neq \mu_{f^*}$. The second is representative of both $\mathbb{R}_{\mu^-}$ and $\mathbb{R}_{\mu^+}$ due to the symmetry of the KL function regarding differences in μ. The shape of the D_{KL} function can be identified in Figure 8.1. As stated,

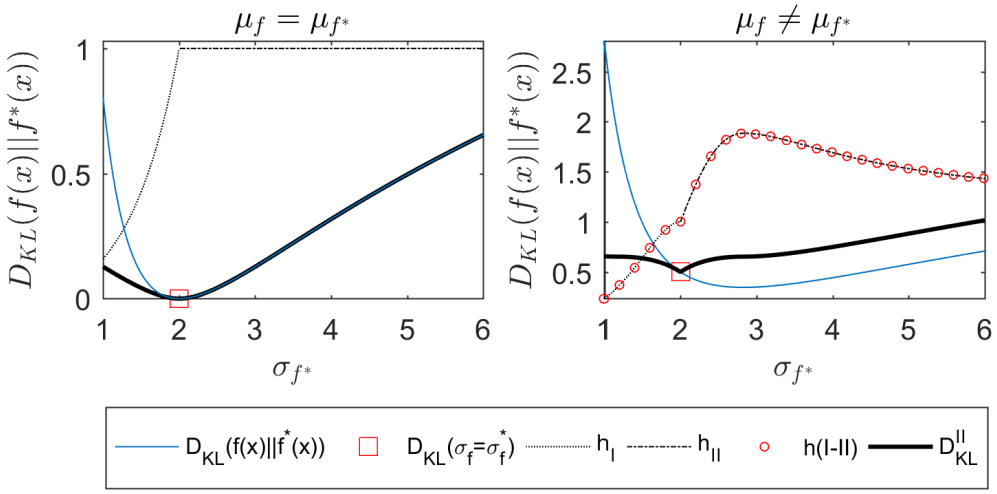


Figure 8.1: Example of the D_{KL} function and transformed function for the cases when $\mu_f = \mu_{f^*}$ and $\mu_f \neq \mu_{f^*}$ for a distribution with $\mu_f = 10$ and $\sigma_f = 2$.

the KL function is not symmetric in regard of changes in σ. When $\mu_f = \mu_f^*$ the minimum of D_{KL} occurs when $\sigma_f = \sigma_{f^*}$, therefore in this case $h(I - II) = h_{II}$. When $\mu_f \neq \mu_f^*$, due to the shift caused by the difference in the means, a similar minimum for $\sigma_f = \sigma_{f^*}$ does not occur. The main purpose of h_I is then to guarantee that the local minimum for $\forall \mu_f \neq \mu_{f^*}$ occurs at $\sigma_f = \sigma_{f^*}$

Two examples of application of D_{KL}^{II} are presented in Figure 8.2. In these, changes in the statistical distribution are studied by evaluating its first two statistical moments. The first (a) presents a univariate data set where the comparative distribution f^* is progressively modified in relation to a reference distribution f.

For the sake of illustration, seven sets with different μ and σ values are shown. Per each set, two σ values have been considered, i.e., the reference one and other selected randomly. Even considering that the f^* is not progressively more distant from f as given by D_{KL}, its statistical moments are. The results for the discrimination are presented to the right. With the single results for D_{KL} it is difficult in terms of sensitivity to identify changes in μ_{f^*} and in σ_{f^*}. This effect is very significant as the ratio between μ_f and μ_{f^*} increases. When this ratio is 1.5, it can be seen that an increase in the σ_{f^*} results in a D_{KL} smaller than when this same ratio is approximately 1.4 and the value of σ_{f^*} closer to σ_{f^*}. On the other hand, D_{KL}^{II} is strictly increasing as the difference between f^* and f statistical moments increases. When conducting a sensitivity analysis this is of interest in order to produce perceptive results and should complement the analysis of D_{KL}. When the changes in the statistical moments are negative, which translates in a negative change in the statistical distribution, the D_{KL} indicator increases rapidly while D_{KL}^{II} increases at the same rate as for symmetrical positive changes. The second example (b) presents

a bi-variate set of data that characterizes a phenomena where the changing behaviour of μ_{f^*} is mainly induced by the modification of the x variable, and the σ_{f^*} is mainly induced by the modification of the y variable accordingly to the figure presented on the left. x and y are independent and can take any value in the field (x, y), μ_{f^*} and σ_{f^*} are constrained to the values taken by x and y. D_{KL} maxima occurs for the points where the difference in μ is maxima but there is no change in σ. D_{KL}^{II} maxima occurs for the points where the change is maxima for both statistical moments. This example shows that, when using D_{KL} as an indicator, if no further information is known about the statistical distributions it is difficult to identify accurately the probabilistic sensitivity results, while D_{KL}^{II} is able to complement this analysis accurately. When the difference in the μ and σ are maxima, D_{KL} takes the same value than when the μ_{f^*}/μ_f is 1.5 with no change in σ.

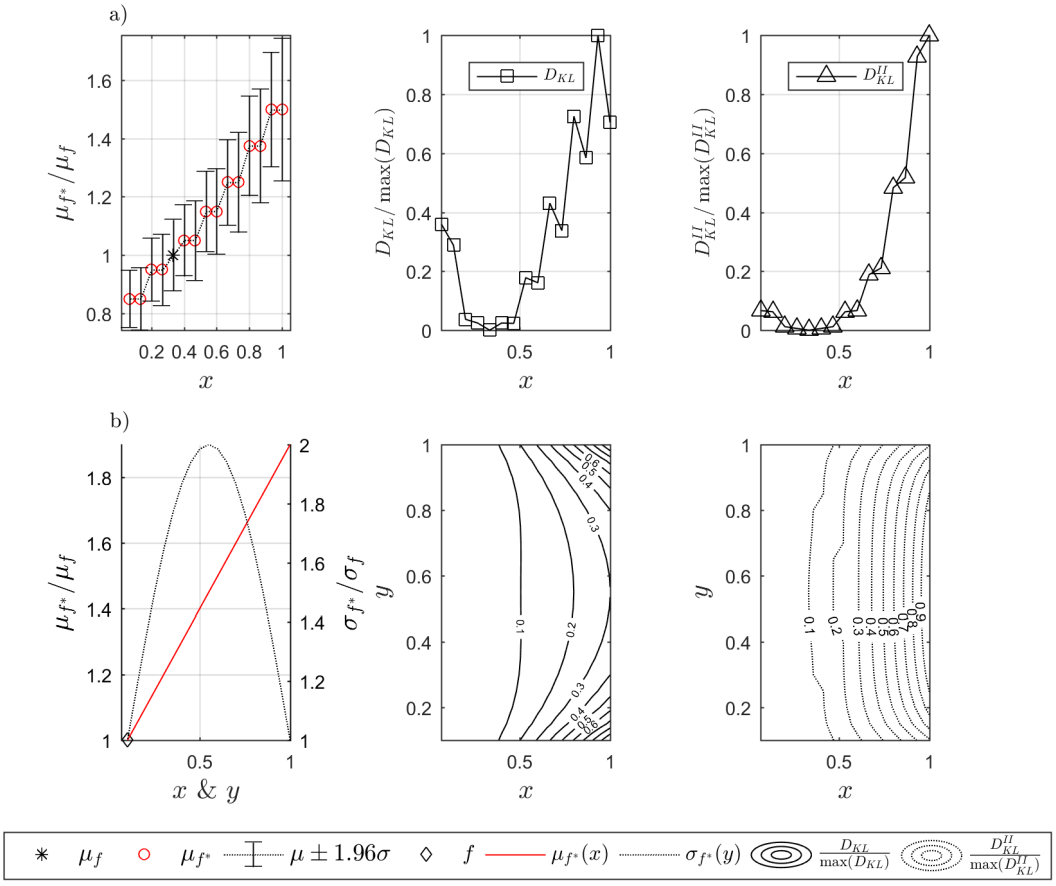


Figure 8.2: Example of implementation of KL discrimination in the context of the P-GSA. a) represents a univariate set of data where both statistical first moments are progressively increased and both D_{KL} and D_{KL}^{II} are calculated to infer on the G-PSA. b) represents a set of data where the sensitivities of μ and σ are mainly influenced by the parameters x and y, respectively.

For some physical processes, such as structural response, the combination of two maxima is expected to be more critical as the μ variable is expected to contribute the most

for the fatigue cumulative effect. In particular, for reliability calculations, the point where both maxima occur is the most dangerous operational point because, with the same μ, more extreme values may be experienced.

While the interest of D_{KL}^{II} might not be so evident in some fields of knowledge such as the ones where the only interest is, for example, to produce statistical distributions that are far from each other; in the case of some physical processes its interest is unquestionable. It is possible to infer that, GSA in a uncertainty quantification (such as mean uncertainty) or reliability context, D_{KL}^{II} enables a more intuitive analysis of the results and complements the analysis with D_{KL}.

In order to have a sensitivity analysis, a reference point is demanded. Given that the sensitivity results have a relative character, absolute probabilistic sensitivity indicators are not of interest. Instead, what is of interest is the representation of sensitivity as a relative measure of the changes in the DoE in a GSA map with its relative physical and statistical meanings. One of the main advantages of using the transformation proposed is that only the reference point is required to implement it. This allows the usage of the transformed KL discrimination function in its discrete form without the need to calculate the statistical moments and assuming the error in the estimation of the discrete distribution functions.

It is noted that the main interest of applying the presented transformation is related to practical applications. The transformation encloses all the probabilistic information in a single variable that is 0 when the PDFs are equal and that increases monotonically with symmetry when one distribution diverges from the other.

If further information is needed about the indicator, the local derivatives can be used to estimate the position of the point in the D_{KL} field. The reference point can also be used to map the D_{KL} and search for an equal discrimination value.

This indicator will be used in the current application to research on the μ_{D_t} sensitivity for the variables considered. σ_{D_t} is considered as an indicator of uncertainty in the μ_{D_t}. It is known that the certainty of the μ_{D_t} is a function of σ_{D_t}. As example , the "naive" calculation of a lognormal variable mean certainty would be given by,

$$CI = \mu_{D_t} \pm z^* \frac{\sigma_{D_t}}{\sqrt{N}} \tag{8.12}$$

with N as the sample size used to converge the statistical moments of the lognormal distribution. To note that calculations use the lognormal's associated normal mean and standard deviation. z^* is the critical value for the CI given given by the standard normal distribution or t-student distribution (function of N). σ_{D_t} relates therefore directly to the uncertainty in the mean for a constant N. The higher the σ the higher the uncertainty in μD_t.

GSA of D_t

The LHS generated for the set of parameters $\boldsymbol{x}$ is presented in Table 8.2. U was constrained between 4 and 25 m/s in order to replicate and efficiently cover power production conditions. The first two statistical moments and two results for D_{KL} and D_{KL}^{II} are also presented in the last four columns of the table.

Given that the sensitivity analysis is based on the relative change of output parameters, the calculation of the statistical discrimination in the context of a GSA for the SN fatigue of the OWT tower demands a reference state of operation. Hence, the definition of the whole GSA map will depend on the reference state used.

In the current example two reference operation points were used to compute the GSA field: (a) the hypothetical more favourable state in terms of fatigue damage possible with

Table 8.2: LHS DoE used in the G-PSA. 10 simulations were performed to characterize each x. The minimum values of μ and σ and Case 1 highlighted in bold were used respectively as reference to calculate the values of D_{KL} and D_{KL}^{II} for the last two columns. D_{KL} and D_{KL}^{II} are represented as relative values in regard of their maxima.

	x					lognormal D_t		$(\mu_{min}, \sigma_{min})$		$Case1$	
$\boldsymbol{x}$	U	H_s	T_p	I	ϕ_w	μ	σ	D_{KL}	D_{KL}^{II}	D_{KL}	D_{KL}^{II}
Case 1	4.41	1.83	6.16	6.11	-0.30	**-22.398**	0.593	< 0.01	< 0.01	0	0
Case 2	8.16	2.86	11.00	6.77	-3.96	-20.83	0.624	< 0.01	0.041	< 0.01	0.027
Case 3	8.52	1.89	5.73	8.87	0.82	-18.49	0.667	0.018	0.255	0.017	0.183
Case 4	9.85	3.29	9.48	9.43	-3.05	-16.84	0.423	0.087	0.509	0.086	0.430
Case 5	11.42	2.19	8.15	7.95	-1.35	-15.27	0.595	0.073	0.85	0.071	0.509
Case 6	12.05	3.67	13.52	8.47	-6.23	-14.43	0.269	0.437	1	0.434	1
Case 7	13.13	2.53	7.16	7.04	9.33	-15.11	0.433	0.143	0878	0.140	0.731
Case 8	13.63	4.05	12.41	7.44	6.01	-15.73	0.43	0.121	0.733	0.119	0.613
Case 9	14.29	3.05	10.36	6.44	8.12	-17.088	0.614	0.039	0.472	0.037	0.300
Case 10	14.86	4.37	17.26	6.91	2.43	-17.18	0.462	0.065	0.451	0.063	0.361
Case 11	15.99	3.24	8.52	9.10	9.14	-16.17	0.411	0.115	0.639	0.114	0.547
Case 12	16.51	5.05	15.32	10.56	4.54	-15.52	0.603	0.066	0.792	0.064	0.487
Case 13	17.18	3.43	14.03	8.65	5.39	-16.48	0.259	0.261	0.548	0.259	0.554
Case 14	17.92	3.79	7.40	5.95	-7.18	-18.01	0.272	0.131	0.304	0.130	0.302
Case 15	18.20	2.41	6.81	5.16	-4.43	-18.80	0.130	0.381	0.164	0.385	0.215
Case 16	18.99	3.98	11.91	5.76	-8.66	-18.35	0.239	0.143	0.252	0.143	0.261
Case 17	19.88	2.77	6.58	7.68	-2.77	-16.67	0.301	0.181	0.524	0.180	0.506
Case 18	20.45	4.19	10.05	8.11	-7.68	-16.32	0.144	0.890	0.469	0.888	0.616
Case 19	20.71	3.10	8.99	7.19	-5.87	-16.89	0.123	1	0.337	1	0.509
Case 20	21.32	4.59	14.58	7.55	-9.87	-16.62	0.302	0.183	0.534	0.182	0.514
Case 21	21.91	3.51	7.71	6.29	7.53	-17.32	0.223	0.258	0.391	0.257	0.415
Case 22	22.47	4.77	12.60	6.60	1.46	-17.15	0.206	0.324	0.409	0.323	0.447
Case 23	23.35	3.84	11.44	5.31	3.11	-17.97	**0.102**	0.935	0.168	0.940	0.331
Case 24	24.13	5.46	17.82	5.66	-0.60	-17.57	0.200	0.290	0.344	0.290	0.379
Case 25	24.82	4.48	9.85	7.78	3.92	-15.82	0.158	0.864	0.58	0.861	0.717

the data generated, and (b) the real less damaging state that effectively occurred (Case 1 highlighted in bold in Table 8.2).

The hypothetical less damaging state, or less damaging uncertain state, is characterized by the combination of the minimum μ and σ (also highlighted in bold in Table 8.2) in the data set considered. This hypothetical state is defined with the intent of having a progressively increasing measure of sensitivity that is associated with the expected risk for the different points of operation considered.

For both D_{KL} results, the change in the sensitivity indicator is minimal. The lower weight given to changes in the standard deviation can be identified as the reason for the similarity between both D_{KL} values. Cases where this indicator does not produce perceptive results in a probabilistic context can be found in multiple examples, such as Case 18 or 19 when comparing with the cases 5, 6 or 8.

When considering D_{KL}^{II} for example, cases 5 and 7 are expected to be very similar in damaging the OWT when considering D_t as an indicator. This is identified when considering the minimum values of μ and σ as a reference point. Nonetheless, the same does not happen when considering the case 1 as a reference point, where case 7 seems more demanding for the OWT tower. This occurs due to the proximity of the mean and standard deviation between Case 1 and 5 when comparing with case 7. As the D_{KL}^{II} measure is more sensitive to changes in the mean, the difference in the resulting indicator between both points of operation increases while they are very likely to be similarly demanding regarding the structural fatigue of the tower component. In addition to using a perceptive measure of GSA, it is then of major importance when building D_{KL}^{II} measures for practical problems to also define a reference point that enables perceptive results. In a design basis, this would help the designer to achieve more comprehensive results.

It is important to highlight that the definition of the reference point is expected to be highly dependent on the problem being analysed. If the goal is to analyse a physical quantity such as the one described, the results should be computed as in the presented case. However, if the problem has, for example, control purposes, it will be of interest to set as the reference a point that is not statistically distant from the intended operational point.

In the current example the first two statistical moments where calculated with relatively low sample sizes due to the fact that computational effort was of importance for the analysis. Furthermore, the interest of using the KL measure was to quantify changes in the mean and its uncertainty (statistical distance increases rapidly with the mean and relatively slowly with the standard deviation). It is noted that depending on the type of analysis implemented and accuracy intended, accurate probability distributions and statistical moments may be needed to characterize the reference point of the D_{KL}^{II} surface, and the sensitivity points. [Greenwood and Sandomire, 1950] presents research results for the convergence of the mean and standard deviation with the sample size. These can be used as a reference in such a scenario.

Figure 8.3 shows the variation of D_{KL}^{II} with the univariate change of each x variable for the case where a hypothetical minimum state of operation is used as reference. The four vertical lines represent, in the order shown in the legend, the rated operation point (case 5), the point with maximum indicator of discrimination (case 6) and the two points with minimum indicator of discrimination at rated operation (cases 15 & 16). These were computed for reference in order to support the comparison of results. The graphical analysis of the results shows two principal trends in the univariate data: U and I are the variables in the DoE that present a stronger indication of being the main contributors to

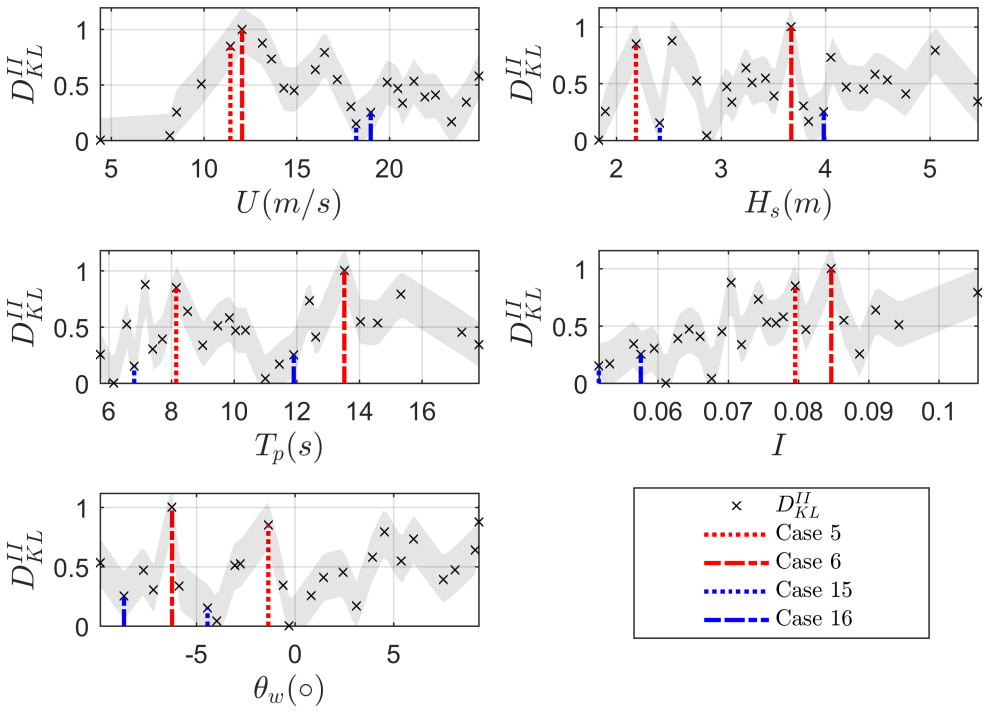


Figure 8.3: D_{KL}^{II} results for each individual θ_i variable. The only purpose of the gray shadowed area is to facilitate the analysis of the results.

changes in D_t. Regarding changes of U, D_{KL}^{II} shows a peak close to the rated operation ($11.4m/s$). Regarding changes of I, there is asymmetry in D_{KL}^{II}. The grey shadow is used to facilitate the identification of these trends when comparing with the remaining variables.

The wave variables do not present any trend indicating significant influence in the D_t. For both H_s and T_p a case where the variable has high and low values with high and low D_{KL}^{II} can be identified, see blue and red trimmed lines of Case 5, 6, 15 and 16. For these, the sensitive behaviour of the D_t is likely to be controlled by U and I, which have more clear asymmetric trends in the results of D_{KL}^{II} also highlighted by the blue and red trimmed lines.

For U it is possible to infer that there is a peak in the sensitivity measure around the rated speed ($11.4m/s$) that occurs independently of the values that any of the remaining variables take. After the turbine is operating above the rated U, D_{KL} tends to decrease. An exception occurs for Case 12, although this is likely to be related to the turbulence. It was mentioned before that, Cheng et al. [2003] showed that the highest loads in a pitch controlled turbine occur slightly above the rated power, and these loads are expected to be related to the increase in the fatigue damage when the OWT enters the rated power region of operation. Such a trend was corroborated by the DEL computed in the previous chapter. Two main regions of operation can be defined due to the large sensitivity to U, above and below the rated power.

Regarding the conditional behaviour of D_t with changes in I, there is a trend for D_{KL}^{II} to increase when I increases. Nevertheless, the peaks of D_{KL}^{II}, and consequently of contribution to the structural fatigue, occur for values of I around 8% and 9%. I and U have a negative correlation and therefore high U values are not expected simultaneously with large I values. This is one of the reasons behind the position of the most sensitive points in the DoE, near the rated operation wind speed. Nevertheless, if large I occurs in combination with large U, as in Case 12, the sensitivity indicator increases significantly.

The only points of operation where a higher turbulence produces a less significant change in the expected D_t occurs for Case 4, and may be connected to the fact that these operational points are below the rated wind speed. By studying the physical description of the system it is possible to understand that, while U is the main variable in the calculation of the average thrust moment in the OWT, I influences how the thrust fluctuates with time. Large I is more likely to occur with low U, where the thrust and consequently moments are lower. Despite the increase in the Gaussian variability of U caused by the higher I, the loading cycles are expected to be less demanding. A similar behaviour may be behind the decrease in D_{KL}^{II} at very high wind speeds where the thrust may be high but the variability and therefore the loading cycles are less demanding.

In the present case it is known beforehand that ϕ_w presents physical symmetry with respect to the OWT tower. Accordingly to the results presented ϕ_w does not seem to have major influence in the output distribution. Cases 7 and 20, are opposite symmetric where 7 and 5 represent a change in D_{KL}^{II} 100% bigger than 20 and 17. Even considering that ϕ_w is expected to have minimal influence in the D_{KL}^{II}, some increase of D_{KL}^{II} is noticed in Figure 8.3 with the increase in ϕ_w yet, no clear trend can be identified. The IEC standard foresees sudden ϕ_w change as one of the design load cases for which a OWT needs to be certified, which is the case where ϕ_w is expected to be have more influence.

The univariate analysis of data, as presented, can be interesting to identify some of the sensitivity patterns. In the case of complex systems, it can be of interest to have a representation that considers the coupling between variables.

One of the main purposes of using a single indicator, such as D_{KL}^{II}, and a minimum

damaging point for reference, is to map the probabilistic field using the GSA results in fatigue operation maps that encloses uncertainty in the variable of interest (in the present case mean), see Figure 8.4. This representation of the results is interesting for dynamic systems as it allows an analysis of the DoE field using a sensitivity and at the same time a design basis perspective. Its representation can replace Figure 8.3, and therefore compile more information about operation in a single figure.

Strong asymmetric or heterogeneous behaviour in the D_{KL}^{II} measure indicates that the output is sensitive in regard of changes in a particular variable. The interaction between variables can be identified when mapping the probabilistic sensitivity as described.

One example is the slight trend of D_{KL}^{II} to increase for positive values of ϕ_w that could be identified in the univariate representation. If with the previous figure it was hard to be certain of the influence of ϕ_w, in the map representation ϕ_w seems to have limited effect in the analysis of the OWT tower structural fatigue. In g) there is a strong asymmetric behaviour that indicates that sensitivity may exist, but this trend is not maintained in d) or i). The asymmetry of ϕ_w is due to its coupling with U in map d).

In the present case only five variables were considered in the DoE and even analysing sensitivity with the usage of tables is not difficult, however this mapping characteristic is of particular interest when a higher number of variables are involved and the analysis is even more complex.

It was mentioned that the wave variables show limited influence in the definition of D_t. Results corroborate the results from [Teixeira et al., 2017a], which performed a preliminary assessment on the sensitivity of D_t for the tower component and concluded that U and I have the largest influence. The other variables modelled were H_s, T_p and wind-rotor misalignment. Despite the high amount of energy that waves carry, common operational waves do not directly interact and load the tower component. The low influence of the wave variables is better identified in the operational map e). The areas of operation where D_{KL}^{II} is maximum seem to be imposed by the values of I and U instead of by the wave parameters. Furthermore, the point of more severe environmental wave conditions, Case 24, results in approximately the same change in D_{KL}^{II} as one of the less energetic sea states, Case 3 (see Table 8.2).

It is possible with GSA mapping representation to infer on how a variable in the DoE is dominant comparing with another. In the current case, as U and I show important influences in the output D_{KL}^{II}, these two variables dominate the location of the most sensitive, thus damaging, points in all the maps. For U, around the rated power ($11.4m/s$), and for I, strong asymmetric behaviour where an large I implies large D_{KL}^{II}.

Variables I and U have a combined effect in the sensitivity indicator. This can be identified in the map c), where for values of U close to the rated U, smaller I implies a decrease in D_{KL}^{II}. At the same time, far from the rated power region, turbulence importantly influences D_{KL}^{II} and consequently, the amount of damage that could be expected in operation for that specific region.

In a case of highly correlated variables, such as U, H_s and T_p, the analysis of sensitivity needs to be taken meticulously or erroneous conclusions may arise. Due to the effect of the correlation some of the grey regions occupy large areas in the plots, e.g. plots a) and e). In the context of fatigue, as a long-term structural event, most of the grey areas are very unlikely to be of interest and therefore they should only be analysed in a design basis if relevant information about structural fatigue may be held in these.

The mapping can be one of the tools used to decide, on a multi-dimensional analysis, what additional points need to be addressed in order to have probabilistic and robust results to sustain the design decisions. It is difficult by analysing Figure 8.3 or Table 8.2

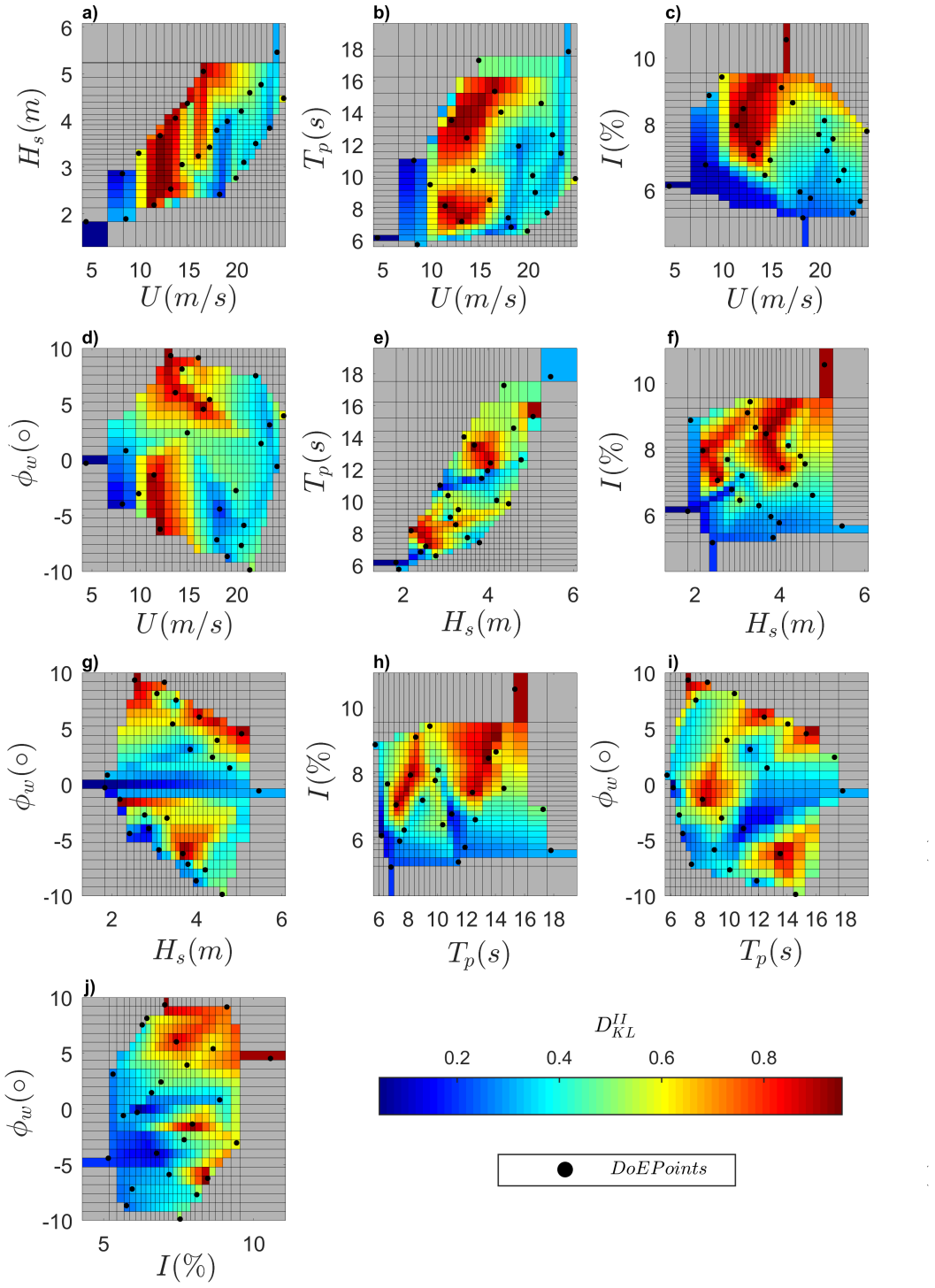


Figure 8.4: G-PSA mapping of fatigue damage operation for all the space of variables simulated (grey areas represent areas not covered by the LHS DoE).

to define new points for analysis without the perception of the DoE space. In the current case the motivation to use this indicator was simpler. The only interest was to reduce the dimensional space for meta-modelling. The KL indicator sufficed this requirement while characterizing variations in the μ, and enclosing σ as a quantifier of uncertainty.

Sensitivity was studied using the first two statistical moments, and achieved robust results that avoid the case where one x variable may have a second order influence, such as individual effect on σ, or combined influence. U and I will be the DoE variables used to build the surrogate model to analyse OWT tower long-term fatigue using meta-models.

8.4 SN Damage surface

A meta-model of D_t is proposed to analyse T fatigue. This meta-model, SN damage surface, works as an interpolator of the SN damage in the x domain. It enables prediction of D_t $\forall x$ using a limited number of evaluation at x. Such an approach is feasible because the SN fatigue damage surface is not expected to have high local non-liner behaviour. Two points that are relatively close in x are expected to have similar mean contribution to the T SN fatigue. SN fatigue was identified to be a problem of mean. A Gaussian process predictor is then used to approximate the mean value and its uncertainty. As D_t was seen to follow a lognormal distribution (see Section 7.3.3) the Gaussian process predictor approach is expected to be adequate. Nevertheless, this is expected not to be a compulsory requirement to use these models as surrogates of D_t.

One of the reasons to apply directly the SN fatigue cumulated damage instead of the DEL to define the damage surface is due to the significant increase in the error of the approximation given by the errors in the DEL approximation. A relatively small error of interpolation of the DEL is magnified to a very high error in D_T.

A diagram of the methodology proposed to analyse the SN fatigue of the OWT tower using a Gaussian process surrogate model is presented in Figure 8.5.

Three main steps can be identified in the procedure presented: evaluation of the simulation model, definition of the surrogate model and SN fatigue assessment. The simulation model is jointly analysed with the surrogate model in order to set a surface of SN fatigue design indicators corresponding to the D_t calculated for X_i. The overall goal is to replicate the currently recommended design procedure by defining a surface that enables evaluation in x without the requirement to run the OWT simulation model. These indicators, the short-term fatigue damage rate D_t, are the SN fatigue that is expected to occur within 600s (in the current case) of operation. They can be defined for other methods or lengths of time. Loads and cycles are counted with a RC algorithm. The built meta-model is then used to predict the long-term fatigue design by defining D_t $\forall x$, or $D_t(x)$ $\forall x$.

Validation of D_t rates calculated was obtained using the certified software [Hayman and Buhl Jr, 2012]. Fatigue loads are calculated at the interface section between the tower and the foundation. This section is assumed to be the critical point concerning the tower component SN fatigue. Nevertheless, the same methodology could be applied to any other section of the tower, or any other component.

If applied to a high m component, particular attention should be given to the characterization of D_t. For the analysis of the tower component, low m materials are considered. The main characteristics for the tower component of the considered turbine were given in Chapter 7. Materials with SN slopes of 3, 5, and double m will be tested in the current implementation.

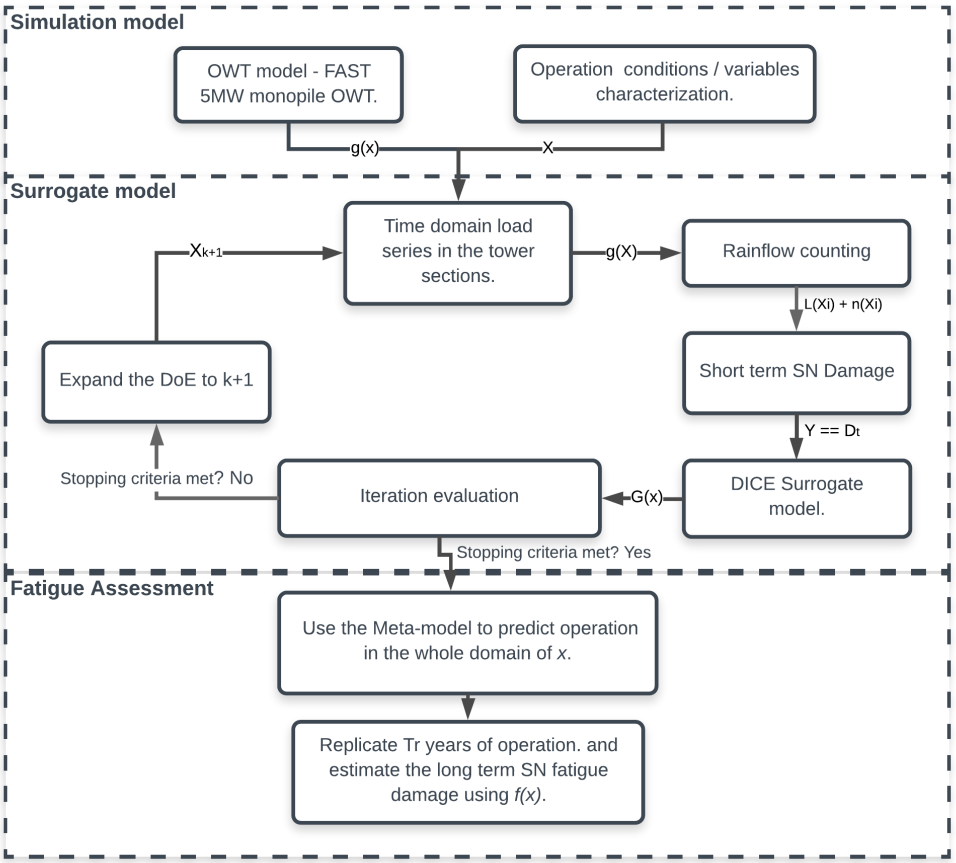


Figure 8.5: Methodology used to calculate the structural fatigue of the 5MW monopile OWT tower.

According to the GSA developed, two relevant environmental variables, U and I, wind mean speed and turbulence intensity respectively, are applied as the main variables of the DoE used to define the operational points and the respective D_t. It is noted that a higher number of variables may be considered when applying the same methodology for other component. Gaussian process models were reported to hold accurate prediction when the number of dimensions in the DoE increases [Gaspar et al., 2014].

8.5 Analysis of OWT towers fatigue operational loads

The algorithm presented in Roustant et al. [2012] is used to build the Gaussian process surrogate, after the DoE and the respective D_t are assessed. Definition of each D_t with respect to DoE is achieved performing multiple time-domain simulations. For reference, Figure 8.6 presents three t-domain simulations and consequent D_t assessment using a double SN slope curve from [DNV, 2011]. The statistical moments of D_t are assessed in each x in order to feed the surrogate model. No fatigue details were considered. Nonetheless, these can be applied, both in the surrogate surface or *a posteriori* to the prediction. The main results for the implementation described are presented in the following sections.

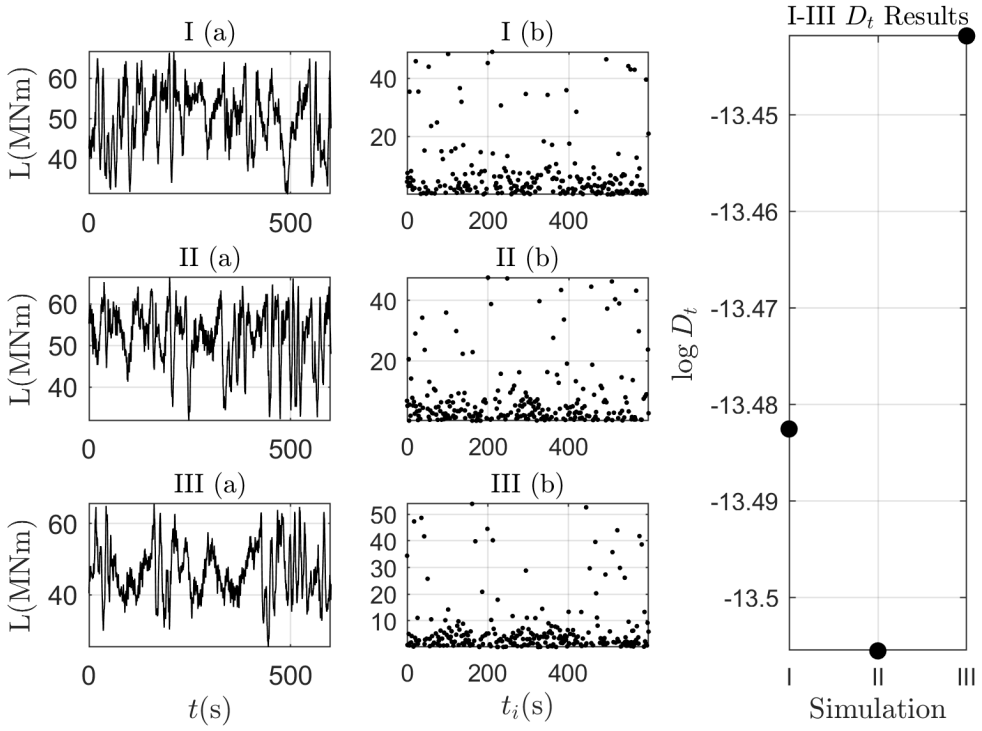


Figure 8.6: Reference example of time domain (named I-III) simulations used to assess fatigue at $x = [U = 11.4m/s, H_s = 2.2m, T_p = 8.2s]$ and NTM class B. Load is assessed at the interface between the tower and foundation (+10m MSL). A time domain simulation is performed (a), rainflow is applied to assess load ranges (b) and D_t calculated.

8.5.1 Yearly simulation data for validation

A one-year operation simulation of fatigue was completed in order to validate the results presented. The full 1-year simulation comprised 51140 simulations of the described OWT model. Every evaluation approximately demands 25 to 30 minutes in average. The computations were performed in a double 10 cores Xeon CPU running 24 simulators in parallel. Even though, the full-assessment still demanded approximately 1100 hours of computational time to complete the evaluation. Dealing with the cost to perform a similar assessment in a standard desktop computer is expected to be challenging.

Histograms of the occurrences for the environmental variables are presented in Figure 8.7.

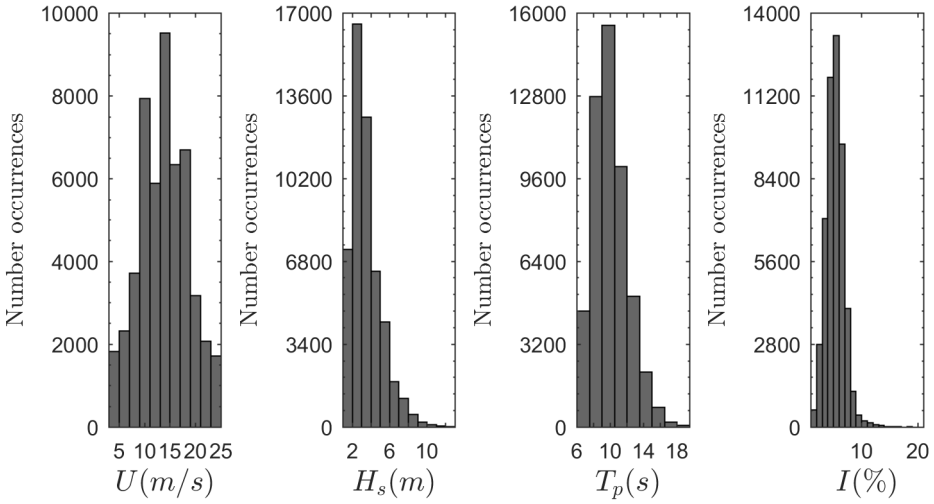


Figure 8.7: Histogram of occurrences used to simulate a full one-year operational loading of the OWT turbine. Wind and waves were assumed collinear in all the cases.

As mentioned, SN fatigue originated in all the environmental states given form these histograms are used to validate the meta-modelling approach proposed.

8.5.2 Selection of the DoE

Chapter 5 discussed the importance of the DoE in the definition of a Gaussian process predictor as a surrogate model. The DoE has critical relevance for an accurate definition of $G(x)$. It was seen that one of the particularities of the Gaussian process predictor model is that uncertainty is high in the points of the space where little is known about the DoE. Furthermore, uncertainty increases with the Euclidean distance to the existing points in the DoE.

The physical description of the problem in-hand may also add specificities to the procedure.

The LHS methodology is characterized by its efficient implementation and description of random fields. Previous applications of Gaussian process predictors to model OWTs applied LHS to define the DoE [Yang et al., 2015, Morató et al., 2016]. In the LHS the initial DoE is defined in the beginning of the procedure and used to define the $G(x)$ function that will be used for further predictions. Despite being efficient and of simple application, the fact that the LHS is defined in one step means that there is no measure of improvement in the approximation of $g(x)$. It is shown in Section 8.5.3 that the application of the

LHS on its own is not sufficient to produce a robust prediction of $g(x)$ using the Gaussian process predictor. Particularly when points with large contribution to cumulative fatigue D_T are missed by the LHS, prediction of D_T using $G(x)$ may comprise large errors.

When applying Gaussian process predictors, a notion of improvement of $G(x)$ should be considered in order to achieve accurate results. An alternative criterion to define $\boldsymbol{X}$ is proposed. It uses a search function that tackles three factors of importance for an efficient and accurate implementation of Gaussian process predictors for SN fatigue calculations:

- Global search of x.
- Relation to the physical problem of SN fatigue analysis.
- Repetition of x points in the DoE.

The proposed infill criterion combines the maxima search function of the EI, to guarantee that no important SN damaging operational points are left out in the analysis (it is expected that some x points will comprise significant contributions to D_T); with an euclidean penalty function, that penalises points that are "too close" in x and prioritizes selection of unexplored areas.

The fact that the EI criterion converges quickly to the value of 0 motivates its implementation. It allows to balance the amount of points that are taken at the high D_t regions, before global exploration. Moreover, if new high D_t regions are encountered during the search, the EI is expected to search these new regions and eventually select new samples that will refine the surface in those regions.

The relation of the search problem and the SN fatigue physical problem is given by $f(x)$. $f(x)$ guarantees that the search follows the joint distribution function and that D_t relates directly to D_T.

The first and the third factors may be covered by the LHS, however little relation to the problem in-hand exists in a LHS procedure.

Spatial maxima search for selection of X_{n+1}

When applying many of the infill criterion discussed in Chapter 5, the algorithms do not always prioritise the characterization of areas of x where little information is available. Under certain circumstances, the $n+1$ point may have little contribution to improve the precision of the approximation, Figure 8.8. This is due to the relation of the currently existing infill criterion and the problem of optimization.

It was seen that few problems consider improvement in regard to the application of $G(x)$ as a surrogate of a complex computational model. Nonetheless, as it is relevant not to miss areas of $G(x)$ that have significant D_t contribution to D_T, some of the search functions may be adequate to calculate fatigue. This is the case of the EI, and the motivation behind its choice to search a SN fatigue space. An EI search was implemented to compute the results from Figure 8.8. It can be seen that the EI predictions occur in a region of the DoE that is expected to have limited contribution to improve the accuracy in approaching the D_t field. This happens because there is already a close DoE point in the same region as X_{n+1}. D_t is not expected to have strong local non-linearities (very close points in the DoE are expected to have similar value of D_t). Therefore, the trade-off in computational effort of defining this point, in the context of $G(x)$ as a surrogate of SN fatigue, is not positive. Computational efforts were allocated to define a point that has little contribution for improving the accuracy of the $G(x)$ surface.

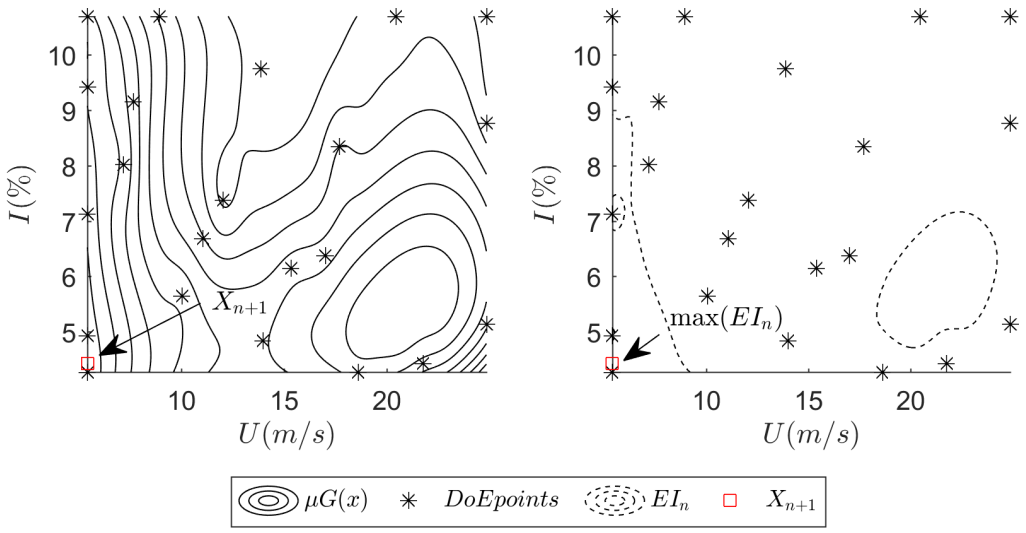


Figure 8.8: Example of reinterpolation of X_{n+1}. The EI is augmented for illustrative purposes (areas are representative of the next potential points).

The analysis of the EI formulation in Chapter 5 shows that the EI criteria is highly related to the known minima and to the local statistical characteristics. This behaviour can be of interest for SN fatigue, as important damaging D_t states x should not be missed when setting $G(x)$. With the aim of accurately assessing structural fatigue it is of interest to change the problem of EI infilling in order to increase the likelihood of search in regions of x where D_t is large. Such a search behaviour is achieved by inverting the damage surface. Therefore, instead of having a search of minima, a search of maxima is promoted.

In a surrogate problem such as the one described, the EI search is of interest to find local maxima points but does not suffice efficient DoE search. In order to optimize the infill process a spatial penalty component is proposed to complement the search with the EI. Equation (8.13) is the search function transformed so that a spatial criterion is introduced.

$$\psi(x) = [EI(x) + P_s(x)]\,, \qquad \forall x\colon \quad 0 < W(x) \leq 1, \quad P_s(x) \in \mathbb{R}^+ \tag{8.13}$$

$P_s(x)$ is the minimum squared Euclidean normalized distance in the space x and is defined as,

$$P_s(x) = \min \mathbf{d}(x|x = x^c)\colon \ \mathbf{d}(x|x = x^c) = \sum_{i=1}^{d} (\boldsymbol{X}_{ji} - x_i^c)^2, \quad j \in [1, \ldots, k]\ ;\ \forall (x^c) \in x \tag{8.14}$$

with the x^c being the potential candidates for the $n+1$ DoE. The maximum values of x in the d dimensions are used to normalize the Euclidean distance.

This spatial condition is added to the EI creating a new search function ψ to be maximized. The argument of the maximum is now selected based on the new search function,

$$X_{n+1} = \arg\max\left[\psi(x)\right] \tag{8.15}$$

The P_s spatial criteria has always a value inferior to the unit. In fact, it will frequently assume a value very close to 0 due to the fact that P_s sums squared distances between 0 and 1. The interest when approaching a probabilistic field, such as the one being studied, is connected to the need to cover most of the x space with the minimum number of iterations.

P_s penalises the computation of points in the DoE that are close to points already in $\boldsymbol{X}$, prioritizing new areas of x where uncertainty is expected to also exist. Enclosing maxima regions of SN fatigue damage is of relevance, but having points well spread in x is no less important. Uncertainty in the $G(x)$ surrogate is commonly higher far from the DoE points. As $P_x(x)$ is close to 0, the individual properties of the EI are dominant in the search, prioritizing search of local maxima. Furthermore, this property takes advantage of the fact that the EI is expected to decrease to a value very close to 0 in few iterations. Therefore, global search with $P_s(x)$ is promoted when it is assumed that the maxima D_t regions are covered. If a new region that has large contribution to D_t is encountered, the EI promotes search in the referred region. It is noted that the EI also has a σ component that promotes global search.

An additional consideration when using a mathematical model as a surrogate model of a complex system's SN fatigue is the one related to the prediction. For the case of fatigue, it is not of interest to compute states of operation that will not occur, or are unlikely to occur, in operation. In case of structural fatigue, the long-term cumulative character indicates that some operational states are more important than others. In particular the ones that occur very frequently may demand accurate predictions.

A weight function ($W(x)$) that relates to the predictive character of the physical problem is also required.

$$G^*(x) = G(x)W(x), \ \ \forall x \ \ and \ \ 0 < W(x) \leq 1 \tag{8.16}$$

In order to have a physical relation to D_T, $G(x)$ of D_t is weighted by $W(x)$, that is equal to $f(x)$, with $0 < f(x) \leq 1$.

For SN fatigue, as a problem of operational repetition, $W(x)$ is well approximated with the joint probability density function of the operational parameters ($f(x)$), Figure 8.9. The search in $G(x)$ will prioritize points that induce large D_t, but that are unlikely to occur, and therefore have negligible contribution to D_T. $G^*(x)$ can be seen as the direct function of the T damage weighted by occurrences. It has thus direct relation to D_T. The problem of SN fatigue is not one of identifying the most damaging states, but instead, the states that will contribute the most to decrease the fatigue survivability of the OWT component.

The results for the first six iterations using the EI improvement, the ψ and considering $G^*(x)$ are presented in Figure 8.10. First the individual EI search is computed, then the new criterion are progressively introduced in the computations. When only the EI is used almost all X_{n+1} points are picked in the same region of x. These points do not contribute significantly for accurate long-term SN structural fatigue predictions. These are points of low occurrence, but large D_t (large values of U and I). This was expected due to the character of the EI function that uses the minimum best known prediction.

Despite the expectation that the maximum values of damage are located near the maximum I, in offshore operational context, high I occurrence above the rated U (11.4 m/s) has very low probability of occurrence due to the negative correlation between U and I. As a result, these states of operation have very small contribution to the improvement the accuracy of the long-term fatigue estimations.

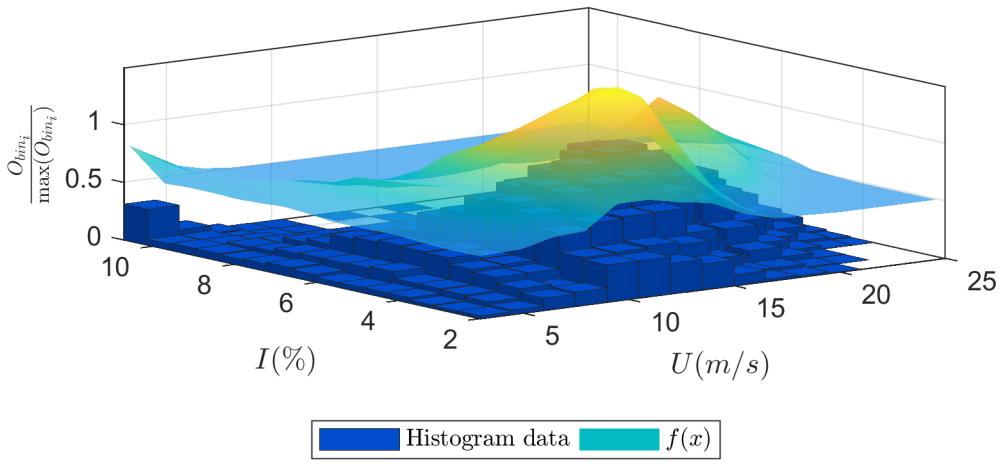


Figure 8.9: Normalized histogram of occurrences (O_{bin}) and joint probability density function of x. (for sake of representation $f(x)$ was plotted as $f(x) + 0.5$).

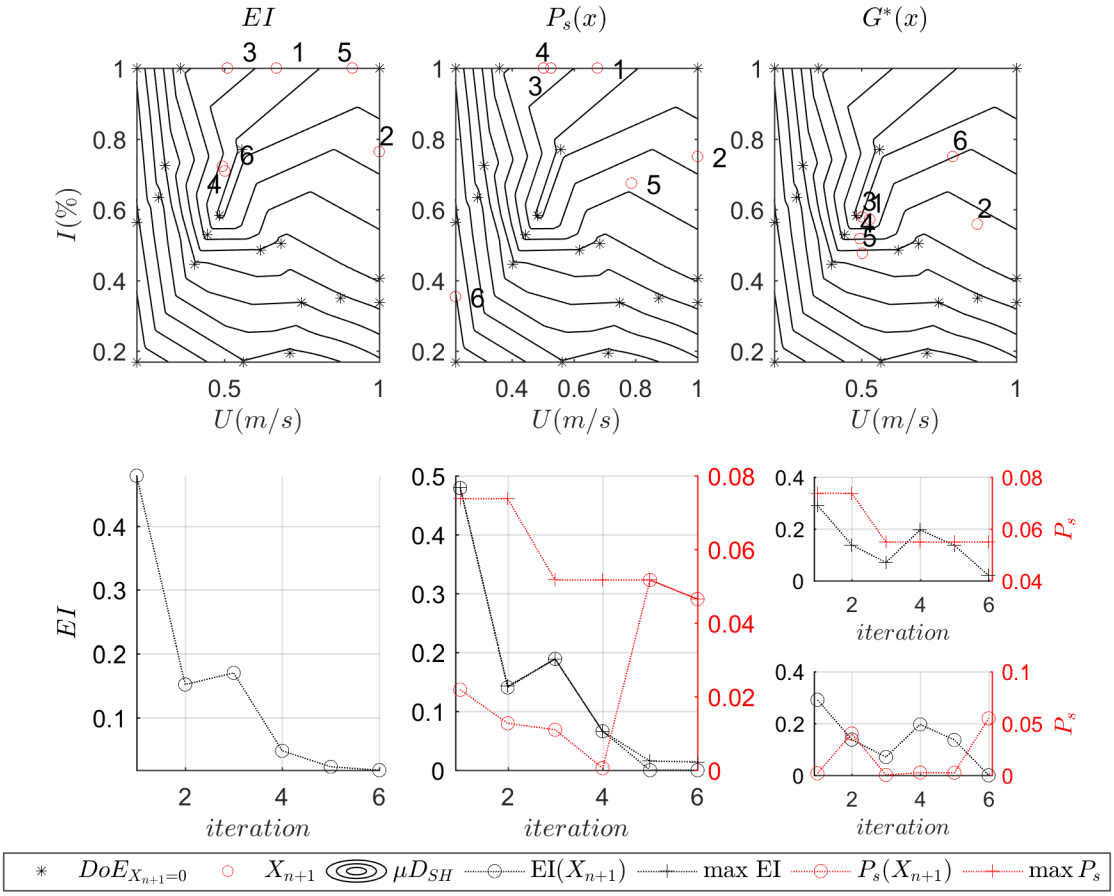


Figure 8.10: Comparison of the different x space search depending on the different criteria used. New criteria are added to the initial EI criteria progressively; first $EI + P_s(x)$ search, and finally weighted with $f(x)$.The order of the introduced points is also indicated.

The introduction of P_s results in a better distribution of points. In this case, the EI criterion dominates the selection of the next point and the first four iterations occur in a region of operation that has little interest to estimate D_T. After 4 iterations the $P_s(x)$ starts having relevance on the choice of X_{n+1}. This relevance can be identified in the positions of the 5^{th} and 6^{th} iterations.

With $f(x)$ weighting the search in $G(x)$, *i.e.* $G^*(x)$, the X_{n+1} prediction often occurs in regions of repetitive operation in x. The maximum damage region is prioritized by the ψ criteria, and $P_s(x)$ promotes global exploration at iteration 2 and 6.

The result is that the EI guarantees that the search in x does not miss important points, such as points 1 or 3, 4, 5, where operation is slightly above the rated power. It was seen that these operational points are important for fatigue calculations in Chapter 7. Moreover, Cheng [2002] showed that these are particularly loading states for pitch controlled turbines. And P_s enables a faster global search.

8.5.3 Fatigue assessment interpolating μD_t

Definition of a D_t surrogate surface allows to design for SN fatigue without the need to access a large number of environmental conditions x.

Results for the long-term prediction of the cumulated D_t ($\sum D_t$) using the meta-model as a surrogate of the computationally expensive OWT model were computed using data from [Teixeira et al., 2018a] to approximate the sampling of $T = 1$ years of operation. The mentioned data-set is presented in Section 8.5.1.

LHS approximation to $G(x)$ DoE.

The LHS has shown the potential to be a preferable method for random stratified sampling due to its capability of having an efficient probabilistic spread over the x space. It was shown that many authors previously applied the LHS in order to define the DoE of Gaussian process predictors when analysing OWT [Yang et al., 2015] or [Morató et al., 2016].

Predictions for a 1-year operational fatigue for different LHS DoEs are presented in Table 8.3. These are used as reference values of a surrogate approximation to the D_T prediction.

Table 8.3: Results for $G(x)$ approximation of D_T using LHS for generation of the DoE sample. Gaussian and Matérn correlations and a 1^{st} order polynomial approximation in $G(x)$ were used to compute the results. θ was converged in the interval $[0.05, 2]$.

			$\frac{\lvert\overline{D_{T=1yr}-D_{G(x)}}\rvert}{D_{T=1yr}}$	
	LHS	Sample size	Gaussian	Matérn
$m=3$	I	10 points; correlated	2.45%	10.9%
	II	10 points; uncorrelated	44.5%	53.9%
	III	15 points; correlated	17.4%	16.5%
	IV	20 points; uncorrelated	14.7%	12.2%
	V	20 points; correlated	19.3%	17.8%

Continued on next page

Continued from previous page

	VI	25 points; uncorrelated	17.9%	16.4%
	VII	25 points; correlated	9.1%	8.0%
	VIII	30 points; uncorrelated	1.8%	1.1%
	IX	30 points; correlated	6.52%	4.5%
	X	40 points; uncorrelated	55.4%	53.7%
	XI	40 points; correlated	14.5%	12.5%
$m = 5$	I	10 points; correlated	76.5%	69.1%
	II	10 points; uncorrelated	> 100 %	> 100%
	III	15 points; correlated	6.6%	4.2%
	IV	20 points; uncorrelated	23.9%	19.2%
	V	20 points; correlated	22.4%	22.5%
	VI	25 points; uncorrelated	> 100%	87.4%
	VII	25 points; correlated	10.7%	18.5%
	VIII	30 points; uncorrelated	15.9%	18.4%
	IX	30 points; correlated	13.8%	12.2%
	X	40 points; uncorrelated	53.1%	47.3%
	XI	40 points; correlated	34.7%	27.3%
Double m	I	10 points; correlated	67.1%	55.6%
	II	10 points; uncorrelated	> 100%	> 100%
	III	15 points; correlated	8.2%	4.3%
	IV	20 points; uncorrelated	15.5%	16.3%
	V	20 points; correlated	25.2%	25.8%
	VI	25 points; uncorrelated	> 100%	> 100%
	VII	25 points; correlated	11.4%	12.5%
	VIII	30 points; uncorrelated	11.3%	12.0%
	IX	30 points; correlated	12.7%	7.3%
	X	40 points; uncorrelated	81.6%	> 100%
	XI	40 points; correlated	29.6%	26.7%

LHS correlations were implemented using the method presented in Section 8.3, originally developed by Iman and Conover [1982]. The symbol ($^-$) represents the average value that accounts for variations in the θ convergence within the search interval (e.g. effect of θ different search algorithm).

All the points with I superior to 13% were computed at the border. Their occurrence is scarce (only common below $U = 11.4$m/s) and the additional D_t these induce is relatively low when compared with the points with the value of 13%. To define the search space,

a threshold (such as 99% or 95% of the damage density) can be used to limit the x space. This may allow to reduce the computational effort of evaluating fatigue using meta-modelling.

Considering an uncorrelated approximation of the LHS DoE sample, with the sample sizes studied, may result in significant errors when predicting D_T using $G(x)$. This is the case of the LHS for 25 points and $m = 5$ and double m. This might depend on the LHS methodology applied.

In the present scenario the LHS algorithm proposed by Viana et al. [2010] was used. When the LHS was generated by 25 points, the algorithm missed the region of operation that significantly contributes for the SN fatigue damage, shown in Figure 8.11. In the case of the component studied, the largest SN fatigue damage contribution to the one year SN damage is confined to a relatively small subset of x, which should not be missed by the sampling approach.

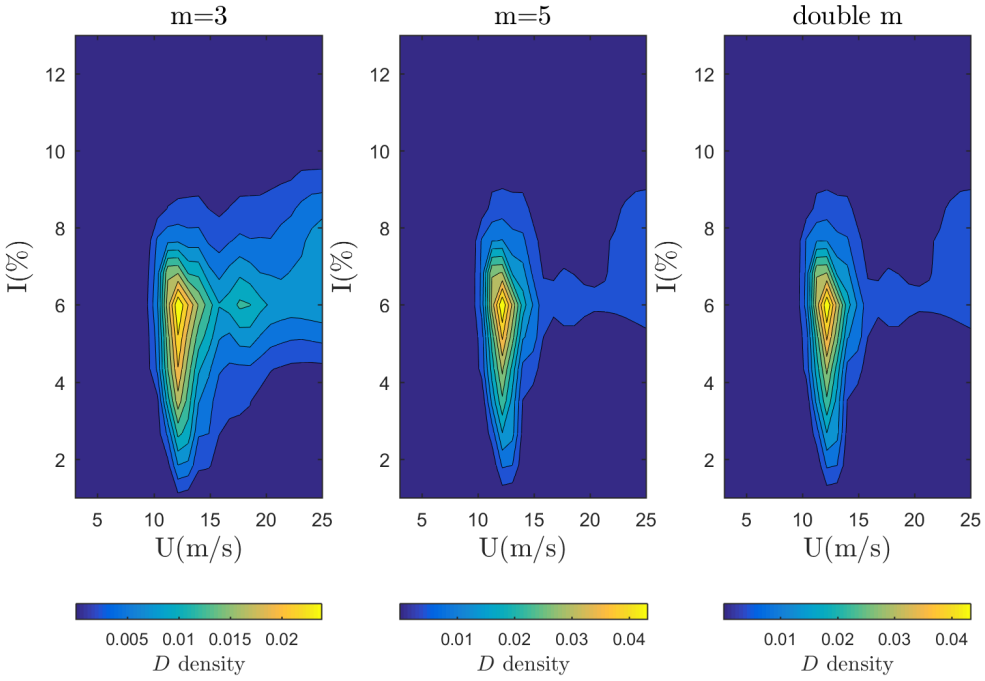


Figure 8.11: SN damage density dependence on x for 1-year data.

Correlated DoE generated with LHS resulted in a better approximation for the smallest sample sizes. As the DoE space is progressively in-filled with points, both LHS approximations (correlated and uncorrelated) present similar results. For very large sample sizes, no improvement in the approximation was identified. In fact it became less accurate. These are the cases of the LHS with 40 points. This may be connected to overfitting, which may generate high local deviations. Alternately, convergence of the model may be unstable for such a large density of points. Alternatively, as mentioned, the LHS algorithm may be influencing the accuracy of the results. In both cases, prediction over-estimated D_T.

As $G(x)$ interpolates results, large sample sizes does not necessarily mean better accuracy in approximating D_T. However, according to the results obtained, the larger the sample size the more likely it is for the approximation to D_T to be accurate up to the point where the predictions strongly diverge from the real value.

Small sample sizes may produce accurate surrogate predictions, but are expected to be less robust when replicating the results. Some points of x are highly relevant for the definition of an efficient surrogate model. If the original sample used to build the meta-model misses one of these points, or has very low resolution in its region, $G(x)$ predictions tends to be inaccurate. This is one of the justifications for the discrepancy between the uncorrelated and correlated LHS. As the correlated LHS considers $f(x)$, the LHS approximation is more likely to pick points with high probability of occurrence, but at the same time, guaranteeing that the selected points represent the entire space over x.

Nevertheless, further research on different possibilities for the decrease in accuracy for the 40 points LHS samples is required for more elaborate conclusions. The resolution attributed by the LHS to the region of x associated with high density of SN damage is crucial for the long-term predictions. The LHS with correlated samples is more likely to have points in the DoE that cover the high density region, resulting in a better approximation when few points are used to sample. Cross validation for the LHS sample size of 25 is presented in Figure 8.12. Ideally the cross validation results should be symmetric over the 45° (red) line. A small increasing trend in the predictions of large values of D_t can be identified in I. The result is the 11.8% error presented in Table 8.3. For II the prediction are considerably more inaccurate, which results in an error over 100%.

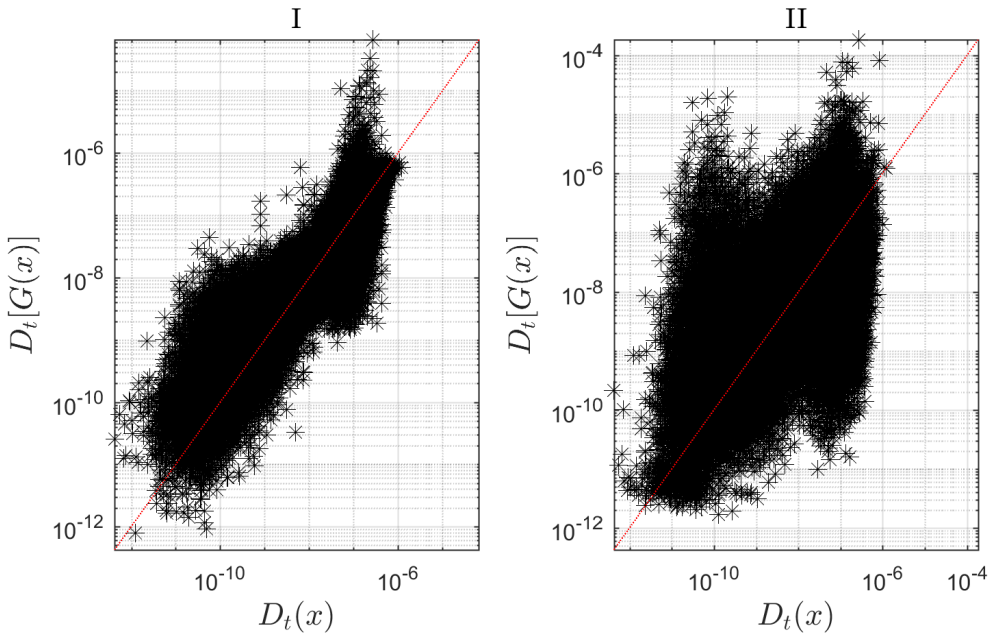


Figure 8.12: Cross validation results for the LHS sample of 25 points with a double m slope SN curve. I - with correlation. II - without correlation. Results for a Gaussian correlation function was implemented.

Both LHS samples, *i.e.*, correlated and uncorrelated, use a global definition of x (cover the whole x delimited by given boundaries), which is also relevant for accurate predictions of D_T. If there is no global character in $\boldsymbol{X}$ when fitting $G(x)$ the model might overestimate the predictions at the x states with low SN fatigue contribution. These, despite individually having low contribution to D_T, summed with other x states that also have mid to low contribution to D_T, impact the accuracy of the results. If no global information is available about the surface, *i.e.* only the damaging x states are known, $G(x)$ is expected

to over-predict D_t due to the correlation with known large D_t operational states.

As the LHS sample size increases, the high density region is also populated for the LHS without correlation of inputs, and therefore, the results for D_T prediction improve, as seen, for instance, in case of 30 points.

It is surprising to note that the correlated LHS sample of 15 points produced a very accurate approximation of D_T for both curves that include a m slope of 5. The fact that the LHS of 10 and 20 points resulted in a decrease of the accuracy indicates that this occurs because of the efficient random balance between the LHS points for the LHS sample of 15 points. The x points that occurred out of the considered space where approximates by the "border" damage value at the same U. This improved the computation of $m = 5$.

The LHS sampling with correlation may be an efficient approach to build $G(x)$ for fatigue purposes. It is simple to implement and, if an efficient sample size is considered, it is likely to result in accurate prediction of D_T with $G(x)$. Moreover, for validation of $G(x)$, there is no need to always run a full one year assessment. A cross-validation exercise using an additional subset samples may suffice the validation exercise for prediction of D_T.

However, if a notion of improvement is considered when defining $G(x)$, prediction of D_T may be improved in terms of accuracy, computational efficiency and robustness. This is the motivation that supports the definition of the presented ψ search function. Results are discussed in the following Section.

Search function approximation to $G(x)$ DoE.

The ψ function considers an iterative approach to $\boldsymbol{X}$ definition where the current state of $G(X)$ is used to compute the new X_{n+1} iteration. It is built to search of maxima and coverage of the DoE space, allowing efficient characterization of the high density D_T region and, at the same time, sufficient spread of points over the x.

Convergence to D_T using $G(x)$ as a surrogate of D_t for an SN $m = 3$, $m = 4$, and double m is studied using the ψ function. A detailed scheme of the search code is presented in Annex, Figure B.1. Results of convergence, using a *Gaussian* correlation and a 1^{st} order polynomial function, to a full 1-year operation are presented in Table 8.4.

An initial DoE ($DoE_{X_{n+1}=0}$) size of 4 points was considered for the 3 cases. These 4 points comprised the corners (*i.e.*, extreme boundaries of the x space) of the search space. Based on previous iterations with the LHS, a maximum computational budget of 30 simulations was considered in order to infill the x space.

Table 8.4: Results for the application of the ψ search criteria. Gaussian correlation function and convergence intervals of $[0, 2]$ were applied to compute the results. i refers to the number of iterations, that is, $i = 1$ implies a DoE size of $4 + 1$ points.

	$m = 3$			$m = 5$			double m		
i	$\frac{\lvert D_{T=1yr} - D_{G(x)} \rvert}{D_{T=1yr}}$	EI	$P_s(x)$	$\frac{\lvert D_{T=1yr} - D_{G(x)} \rvert}{D_{T=1yr}}$	EI	$P_s(x)$	$\frac{\lvert D_{T=1yr} - D_{G(x)} \rvert}{D_{T=1yr}}$	EI	$P_s(x)$
1	0.603	4.166	0.078	0.842	5.49	0.072	0.813	6.088	0.137
2	0.67	$< 1e-3$	0.156	4.314	$< 1e-3$	0.165	0.233	$< 1e-3$	0.205
3	1.005	0.643	0.024	1.213	0.634	0.009	0.26	0.608	0.011
4	0.506	0.168	0.006	1.161	0.353	0.003	0.486	0.406	0.022
5	0.474	$< 1e-3$	0.119	0.231	0.146	0.032	2.551	0.272	0.005
6	0.263	0.084	0.026	0.231	$< 1e-3$	0.115	0.051	0.376	0.002
7	0.191	0.072	0.001	0.197	$< 1e-3$	0.077	0.878	$< 1e-3$	0.152

Continued on next page

Continued from previous page

8	0.171	$< 1e-3$	0.069	0.233	0.01	0.048	0.528	$< 1e-3$	0.126
9	0.164	$< 1e-3$	0.051	0.128	0.385	0.005	0.011	0.085	0.004
10	0.152	$< 1e-3$	0.049	0.106	$< 1e-3$	0.054	2.714	0.079	0.001
11	0.125	$< 1e-3$	0.039	0.052	0.001	0.051	2.491	0.224	0.002
12	0.12	$< 1e-3$	0.036	0.514	0.038	0.003	11.542	0.177	0.003
13	0.112	$< 1e-3$	0.032	0.499	0.004	0.035	1.555	0.168	0.001
14	0.095	$< 1e-3$	0.027	0.441	$< 1e-3$	0.034	1.296	0.082	0.021
15	0.036	$< 1e-3$	0.027	0.370	$< 1e-3$	0.032	1.165	0.089	0.001
16	0.065	0.028	$< 1e-3$	0.282	$< 1e-3$	0.027	1.045	$< 1e-3$	0.068
17	0.05	$< 1e-3$	0.026	0.232	$< 1e-3$	0.026	0.919	$< 1e-3$	0.067
18	0.045	$< 1e-3$	0.025	0.202	$< 1e-3$	0.025	0.506	0.003	0.051
19	0.05	$< 1e-3$	0.024	0.186	$< 1e-3$	0.024	0.413	0.001	0.051
20	0.045	$< 1e-3$	0.021	0.184	$< 1e-3$	0.021	0.579	$< 1e-3$	0.05
21	0.006	$< 1e-3$	0.020	0.148	$< 1e-3$	0.021	0.383	0.034	0.018
22	0.005	$< 1e-3$	0.019	0.151	$< 1e-3$	0.02	0.352	$< 1e-3$	0.046
23	0.013	$< 1e-3$	0.018	0.147	$< 1e-3$	0.019	0.253	$< 1e-3$	0.043
24	0.001	$< 1e-3$	0.018	0.146	$< 1e-3$	0.019	0.251	$< 1e-3$	0.032
25	0.001	$< 1e-3$	0.015	0.075	0.019	0.001	0.09	0.001	0.03
26	0.001	$< 1e-3$	0.015	0.063	$< 1e-3$	0.017	0.017	0.007	0.017
27	0.013	$< 1e-3$	0.015	0.057	$< 1e-3$	0.015	0.045	$< 1e-3$	0.023
28	0.011	$< 1e-3$	0.014	0.072	$< 1e-3$	0.015	0.016	$< 1e-3$	0.02
29	0.002	$< 1e-3$	0.014	0.069	$< 1e-3$	0.014	0.002	0.006	0.014
30	0.002	$< 1e-3$	0.014	0.078	$< 1e-3$	0.014	0.019	$< 1e-3$	0.018

For $m = 3$, convergence to D_T was efficient. After 15 iterations the error of approximation to D_T was below the 10%, and after 20 iterations, below 2%. For $m = 5$ and double m, 25 iterations were needed to guarantee a stable error inferior to 10%. In case of a double m SN curve (with m for high cycle of 5 and low cycle of 3), 25 iterations were sufficient to consistently guarantee an error inferior to 5%. For $m = 5$ the error remained below 8% after 25 iterations, but never lower than 5%.

Convergence was expected to occur more efficiently to $m = 3$. For an $m = 3$ the error in approaching μ_{D_t} (statistical mean of the D_t probability distribution) with 10 simulations is less likely to be large. On the other hand, for an $m = 5$ this error is more likely to be large as 10 simulations is a relatively low sample size for μ converge.

Even considering that $G(x)$ interpolates the region of x, the interpolation accuracy is still bounded by the $\boldsymbol{Y}$ used to build $G(x)$. Moreover, for the same material strength characteristics, a lower m results in larger damage. The damage density for $m = 3$ is spread over different x regimes facilitating the D_T characterization with a surface (less weight is given to inaccurate characterizations of the x regions).

It can be seen that the EI leads the initial search, while $P_s(x)$ dominates posterior iterations. The EI converges relatively fast to 0, but not without selecting first the points of high damage density. The $P_s(x)$ guarantees that the results are not highly overestimated by filling regions of the DoE with relatively low D_T damage density.

Example replications of the convergence for the $m = 5$ and the double m are presented in Figure 8.13.

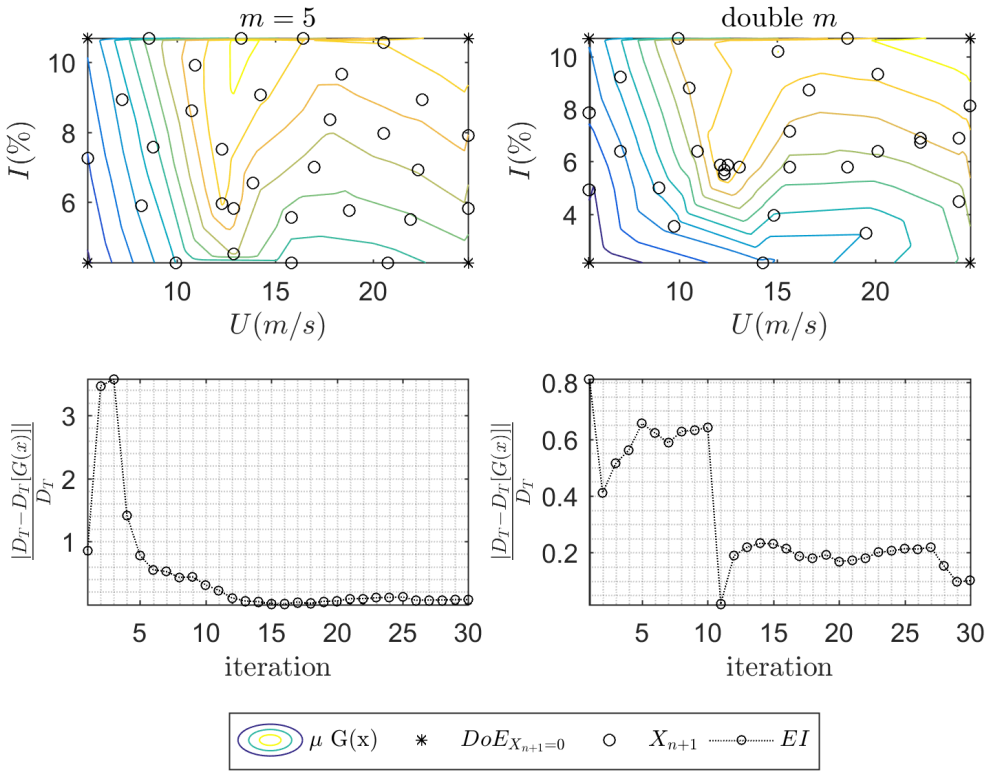


Figure 8.13: Convergence results for $m = 5$ and double m curves.

While in this case results for $m = 5$ were achieved after 15 iterations, the error after this iteration stabilized around values inferior to 5%, and only occasional surpassing this threshold. The convergence of the surface prediction to the real value depends on the convergence of θ, correlation function, polynomial approximation and also on the values of the μ_{D_t} used to define the predictor, and which have important influence on the EI search. This may lead to different search procedures using the same methodology, and achieving convergence with different number of iterations.

For all the cases studied with 4 initial points, 30 iterations were sufficient to converge the error. The replicated results for the double m SN curve are of interest to show the influence of sample convergence and search algorithm on the results. In this example the final iteration convergence error was of approximately 10%. The error in the approximation of D_t initially decreases, but after the 12^{th} iteration the error increases to around 20% and only a relatively small improvements are achieved after this iteration for the budget considered. In this case more search points were selected by the EI function. This can be identified in the agglomeration of sample points in some regions of operation. Lower spread of points in the space occurs for the same computational budget due to this. This is particularly notorious for the low I region. As the trend of the algorithm is to infill the space, it is expected for the convergence to eventually go to 0 if the error is not originated on sample convergence. A trend to an error of 10% may indicate deviation in the sample convergence. This can be also a potential inducer for lack of accurate convergence. In such a case, more iterations may not guarantee better convergence to D_T.

However, it is important to emphasize that the results computed and presented so far are highly demanding in terms of the search function. Only four corner points are given, and the ψ function needs to automatically search all the possibilities of x. While

for $m = 3$ the results were very accurate and efficient, if other information than the corner points of x is given, *e.g.* an initial larger sample to be improved, the results are expected to converge faster. The only recommendation for such approach is to indicate at least the boundaries, or corners, of the search space to the algorithm. Otherwise, areas of x may be undefined, generating wrong predictions by $G(x)$ (extrapolates outside of the x domain).

Results for an initial sample of 10 points were computed using the ψ criteria for an $m = 5$ and double m, Figure 8.14.

The prediction for D_T after 10 additional iterations gave errors of less than 10% for all the three cases. While for the double m curves the results rapidly approached D_T with errors of approximately 7% and 3% after 10 iterations (two more points were sampled in II to confirm the convergence), for $m = 5$ the error was only smaller than 10% in the last iteration. Five additional iterations show that it was stable below 3% after this iteration. This may indicate that $m = 5$ may need more iterations in order to ensure robust results. Double m SN curves are a mixed model between $m = 3$ and $m = 5$, and their fatigue performance was identified to be in between these two slopes.

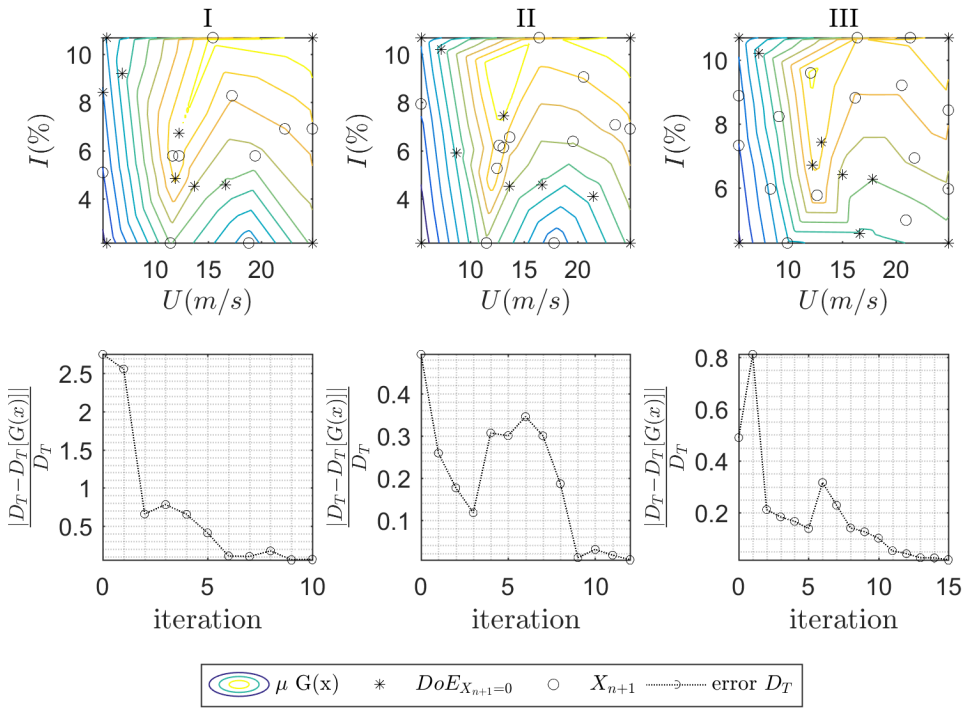


Figure 8.14: Results for the convergence of the Kriging cumulated sample when comparing with one year simulation of operation. I-III present 3 examples simulated considering different initial DoE ($DoE_{X_{n+1}=0}$) of 10 points. The initial set of points ($DoE_{X_{n+1}=0}$) considers the four corner points of the DoE space and additional six random points extracted from a LHS sample without replacement. I - II - SN curve with double m. III - $m = 5$.

To identify convergence using ψ is challenging. A computational budget may be selected in order to search the space guaranteeing that this minimum number results in convergence. Defining a computational budget is a common option in learning procedures as the one presented. In the current case, it was seen that 30 iterations were expected

to suffice convergence of the problem when $DoE_{X_{n+1}=0}$ had 4 points. And 10 when the initial DoE had 10 points.

The EI can be used and weighted with $f(x)$ as a stopping criteria, similarly to what is already used with the other applications of the EI. After several iterations with no influence of the EI, the algorithm may be stopped. In the current implementation the criterion of stability of ψ could be used, Figure 8.15.

A common trend which was identified, was the ratio of ψ value of consecutive iterations, $\frac{\psi_i}{\psi_{i-1}}$, to approach the value of 1 as new points are computed. A ratio in between 0.9 and 1 for consecutive iterations may be used as an indicator that the infill can be stopped. This occurs due to the exponential decrease of the expected improvement, and consequent individual dependence of the search on $P(x)$. This exponential decrease indicates that no further maxima damage density regions were identified. Moreover, when a threshold of $P(x)$ is reached, the space is expected to be reasonably covered. This threshold value of $\max P_s(x)$ can be calculated *a priori* for a specified x space.

Results may be accurate for both the ψ and the LHS generation of $\boldsymbol{X}$. In this context, the ψ function has the advantage of being robust in the calculations, and frequently, with a smaller sample size.

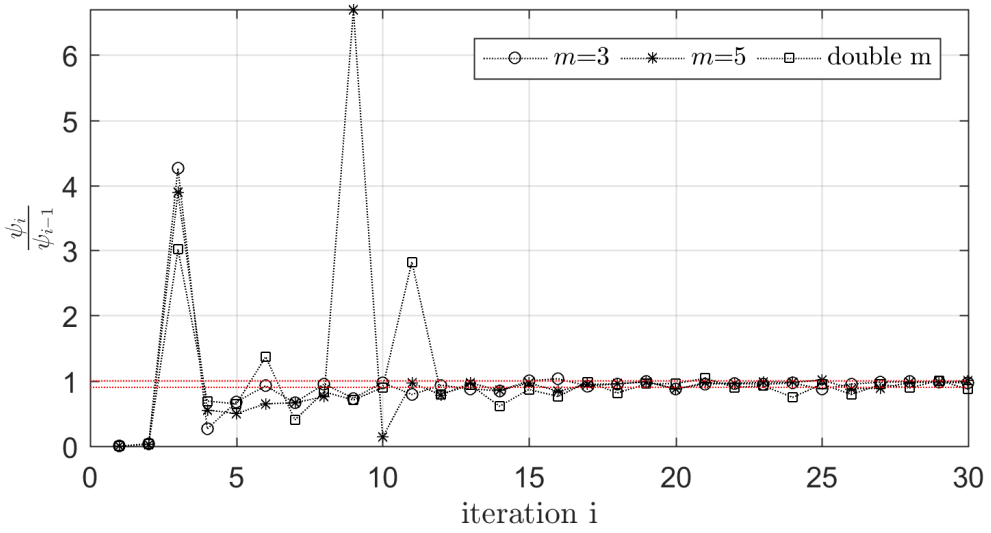


Figure 8.15: Convergence criteria example using the stability of ψ. Red band indicates the margins for the (ratio) range of 0.9 to 1.

Comparison to extrapolation of results using binned data

For reference, results for SN fatigue estimation are compared with the results of extrapolation using the methodology recommended by IEC [2005, 2009]. A loading spectra was defined for each environmental combinations of U and I and multiplied by their joint probability. Computations considered a s_0 of six simulations. IEC [2005, 2009] recommends the usage of six simulations for each combination of environmental variables in order to define the loading spectra. Results are presented in Figure 8.16.

A similar procedure was implemented in Moriarty et al. [2004], where the authors performed more than 4000 simulations for a precise implementation (9 simulations per environmental condition were used). In the present case, as the I space is smaller due to the offshore description, less combinations of U and I were needed resulting in a maximum

of 1584 simulations for a precise implementation using binned x data at values of 1. It can be seen that, for the cases where bins of value 1 were considered, $(T/t \sum D_t)$ approached $\sum D_T$ with an error of less than 1% for the three materials.

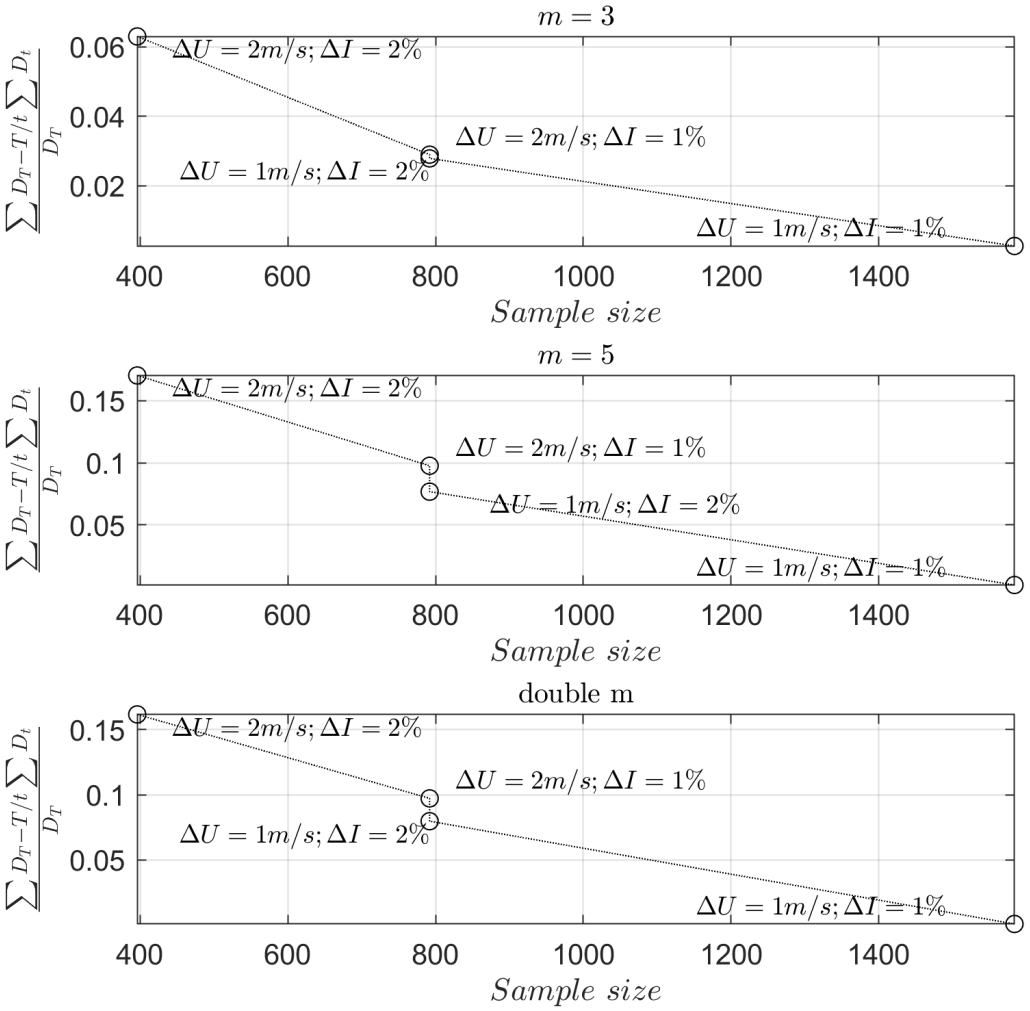


Figure 8.16: Convergence of the $\sum D_t$ depending on the approach to bin x data and consequent number of simulations. Four scenarios where considered to study the environmental variables; bins of $2m/s$ and 2% for U and I respectively; bins of $1m/s$ and 1% and alternating between the values of 1 and 2 for both U and I.

Using bins of 1 and 2 or only 2 the computational cost was cut by 2 and 4 respectively, however this resulted in a relative error of approximately 9% and 17% for $m = 5$ and double m curves, and 6% and 3% for $m = 3$. It was seen that within a bin of $2m/s$ or 2% some variability may occur in the D_t.

Errors are higher in the cases of $m = 5$ and double m SN curves. The double m SN fatigue in the present case is mostly dominated by high cycle loading accounted in the $m = 5$ slope. For the same s_0 the error in estimating μD_t may be larger for these curves when comparing with $m = 3$.

The effort reduction benefit of using a Gaussian predictor as a surrogate model is significant. Depending on the accuracy intended for the approximation, the computational cost can be cut by a factor of 4 to 10 for $m = 3$ and 4 to 8, in $m = 5$. This results in an approximate reduction of the computational time needed to assess SN fatigue design by up to 85% without significantly compromising accuracy.

In case of the calculations for $m = 3$, bins of value 2 produce relatively accurate predictions of D_T. However, with approximately a half to a third of the number of simulations, $G(x)$ produced even more accurate results. If comparing with the most accurate case, reduction of computational time may go up to a factor of 10. This value may be further reduced if a $s_0 = 6$ is considered to calculate $G(x)$, or *vice-versa* for the extrapolation (considering a $s_0 = 10$ for the binned data)

In case of double m and $m = 5$ the converged value of D_T in 30 simulations was less precise when comparing to m=3. Results improved when an initial DoE of 10 points was considered. Sample size s_0 has larger influence in this case. Bins of value 1 resulted in a level of accuracy not attained by the initial $G(x)$ predictions. Results of $G(x)$ may be comparable to the values given by alternate bins of 1 and 2, however with approximately 25% of the computational time. In these cases where 10 initial points were given for the search in $G(x)$ the results are comparable with the ones using a bin of 1. However, imposing a reduction of computational time by a factor of approximately 8.

Bins of value 2 recurrently presented less accurate results than $G(x)$. This may be connected to the important interpolating capability of $G(x)$ in the x space. Size s_0 may be an additional source of uncertainty for the less accurate precision given by the binned predictions, when comparing with the $G(x)$ approximation.

In the case of $m = 3$ it would be interesting to analyse the influence of slightly increasing or decreasing the number of samples to converge μ_{D_t}. Results from binned data indicate that, instead of 10 simulations, 6 simulations may be used to define each new X_{n+1}. A reduction from 10 to 6 samples, as recommended for the Design Load Case 1.2 [IEC, 2009], could provide further reduction of the computational time required to build $G(x)$ by 40%. Nevertheless, for larger m materials, Chapter 7 showed that 6 simulations may not be sufficient to compute D_t. Nonetheless, for optimization procedures s_0 can be significantly reduced and validated at the last, or reference, iterations.

The same methodology can be implemented using any material and component. Note the fact that if a larger m is applied, extrapolation and large sample sizes may be needed to produce accurate approximations. In such a scenario, the same methodology can be implemented considering that calculations of D_t to define $\boldsymbol{Y}$ need to comprise the tail region influence. The main benefit of using $G(x)$ is the possibility of covering a wide range of x with a small number of points. Challenges related to the accuracy of s_0 are not addressed by the approach. Nevertheless, $G(x)$ properties may be used to interpolate confidence intervals in the indicator approached with the meta-model. This possibility will be discussed later on Section 8.5.5.

For a comparative reference, D_T was computed using only $f(\boldsymbol{\beta}; x)$ and ignoring $Z(x)$, Figure 8.17. The analysis was then extended to a comparative analysis of $G(x)$ with the polynomial approximations of μ and μ_{LN} of D_t, Figure 8.17 (b1-b2) and (c1-c2).

The analysis of polynomial fit was developed using cases of double m and $m = 3$. Results for double m and $m = 3$ are similar. Disregarding the $Z(x)$ in $G(x)$ under-predicts D_T. Using a polynomial approximation of μ_{LN} may strongly overestimate D_T. The values of local D_t are small (significantly smaller than 1) and their fit is challenging. A 3^{rd} polynomial approximation to the $m = 3$ surface results in a error only slightly larger than 5%. However, other polynomial approximations show that the polynomial over-

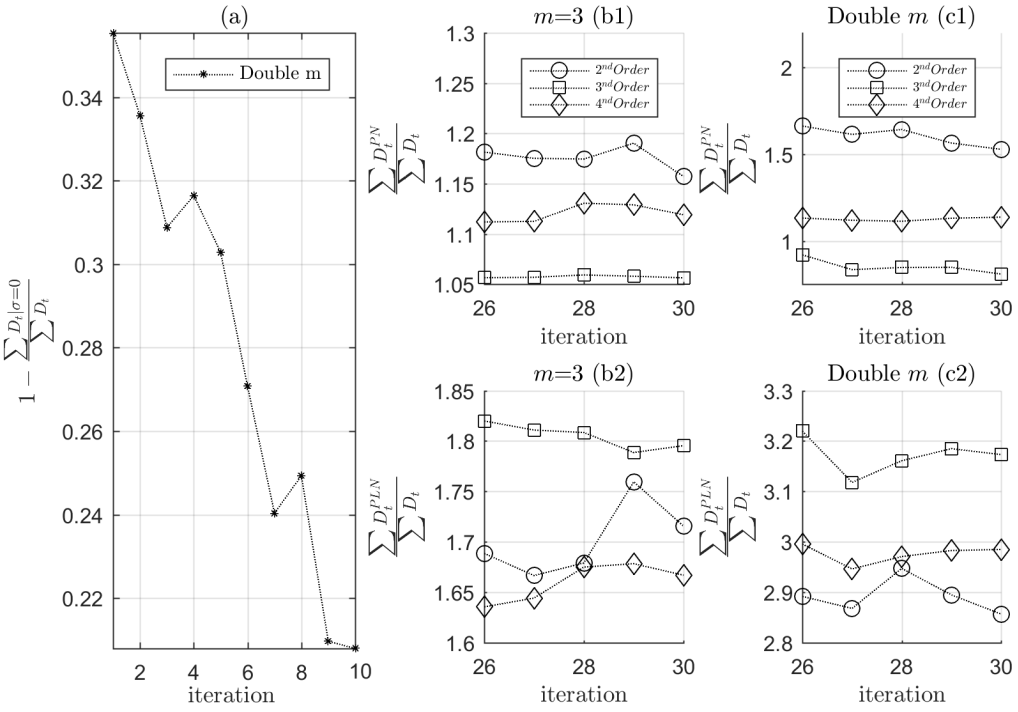


Figure 8.17: Results for the analysis of the influence of using the Kriging model as a surrogate of D_t. (a) presents the difference in $\sum D_t$ if only $f(\beta; X)$ is used in the prediction for the case of double m. (b1), (b2), (c1), (c2) present the results for polynomial fits of μ and μ_{LN} to the X of $m = 3$ and double m five iterations respectively. 2^{nd}, 3^{rd} and 4^{th} order polynomial functions are used. $\sum D_t^{PN}$ and $\sum D_t^{PLN}$ represent the one year sum of the surrogate of μ and μ_{LN} of D_t respectively.

predict the D_T estimation. Large degree polynomial fits are susceptible to overfitting. Under-prediction caused when $Z(x)$ is neglected, indicating that the Gaussian behaviour may be of interest to replicate $\sum D_t$ due to the lognormality of the long-term mean. This occurs because the fit is implemented using a logarithmic transformation (in order to avoid the challenging fit of the lognormally distributed D_t variable). It is noted that further comparative research should be conducted with methodologies specified for polynomial approximations. Results presented were computed for $G(x)$, and may not be totally adequate for comparison.

The methodology introduced in the current chapter that uses a learning search can be implemented in any work that fulfils two conditions; (i) a quantified indicator of fatigue damage can be defined; and (ii) it follows a normal or lognormal distribution. In the case this indicator does not follow a lognormal distribution, multiple surface may be sampled from $G(x)$ and an interval for D_T defined. This approach would represent replicating the designing process as many times as the number of surfaces sampled, under the condition that the $\boldsymbol{X}$ points are a true approximation of $D_t(x)$ at x. This same approach can be implemented as an alternative to infer on the uncertainty associated with the lognormal or normal estimations. Section 8.5.5 researches on this potential application using a noisy model.

It is important to highlight that, if the number of repetitions to be extracted from the surrogate model is relatively low, or if these are not cumulative as in the case of SN fatigue, then σ_G may be relevant. In such scenarios, the usage of noisy Kriging to approach probabilistic fields is of interest.

Finally, it is noted that low frequency loading wind effects were not considered. Veldkamp [2008] showed that these have a negligible influence in the long-term fatigue damage of low slope S-N materials. Sutherland [1999] also remarked that ground cycles have negligible influence on D_t calculations.

8.5.4 Long-term statistical convergence of the SN fatigue.

Figure 8.18 presents the evolution of the $\sum D_t$ statistical behaviour for each n_i cumulated sample up to the full lifetime (T) of operation.

With $G(x)$ as the predictor of μD_t and σD_t (the second approximated using a noisy component) the process of estimating D_T is repeated 100 times. Using the noise to approximate the interpolation of σ, $G(x)$ should be understood as a model that compiles the statistical distribution of D_t for each x. $\pm 20\sigma$ predictions were deliberately considered to account for potential errors in the noisy approximation of σ. For each x a new D_t prediction is taken from $G(x)$ and added to the cumulated damage at $nD_T - 1$.

To infer on the statistical behaviour of $\sum D_t$, 100 cumulative samples were used to characterize it at each cumulative D_t point.

In the figure it is possible to infer that σ only influences the $\sum D_t$ statistical variability if the number of cumulated samples is relatively small. For the case of D_t with 1^{th} polynomial function and *Gaussian* correlation, in approximately 36×10^4 independent samples, 99.7% of $\sum D_t$ is confined to a $\pm 1\%$ variation of $\mu\left(\sum D_t\right)$. Considering a $+20\%$ increase in σ ($\Delta\sigma$) this number increases to 55×10^4. Despite the increase, the relative influence of $\Delta\sigma$ is still negligible. For a sample of 36×10^4 independent points, using a $+20\%\Delta\sigma$, 99.7% of the PDF of $\sum D_t$ is confined to a $\pm 1.2\%$ variation in $\mu\left(\sum D_t\right)$. While for $-20\%\Delta\sigma$, $\pm 0.8\%$ variation of $\sum D_t$ is expected within 99.7% of the PDF.

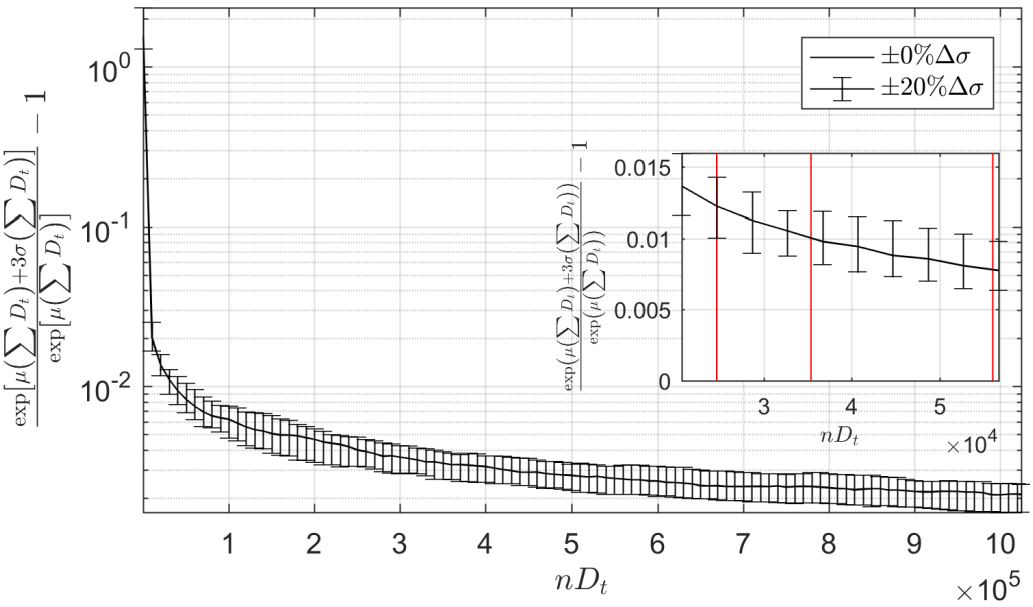


Figure 8.18: Statistical evolution of the interval where 99.7% of $\sum Dt$ is confined with the number of cumulated samples. Note: $\frac{\exp(\mu+3\sigma)}{\exp(\mu)}$ represents the variation to which 3 standard deviations are confined, therefore 99.7% data.

The long-term variability of the $\sum D_t$ is low. In 20 years of operation the $\sum D_t$ has 99.9% of its occurrences confined to a deviation of less than 1% of the mean value.While the variability of the results (given by the ±99.7 statistical variability) is high when nD_t is small, it converges to a value very close to 0 as the cumulated sample increases. Local uncertainty of D_t was identified to be significant (200% variations were identified between maximum and minimum in 50 simulations), however, D_t lognormality balances the cumulated damage over time. The expected reason for such a behaviour derivates from the fact that $\sum D_t$ is calculated considering a sum of independent samples from a D_t distribution. After many consecutive repetitions the probability of consecutive occurrences far from the mean (and therefore with low probability) is increasingly less. As the number of observations is very large, according to the law of large numbers, the average of the random D_t for each x will prevail in the approximation of D_T.

Considering the well-established case of summing Gaussian distributions, μ grows in comparison with σ by a ratio of $\frac{N}{\sqrt{N}}$, with N being the number of summed samples. A similar growing trend occurs in the present case, confirming the fact that D_T prediction is a problem of mean approximation to $D_t(x)$.

These results corroborate the previously discussed idea that the challenge of the long-term D_T definition is not one of statistical characterization, but instead, one of how much is not known about the true value of $\mu_{D_t}(x)$ $\forall x$

8.5.5 Interpolation of the sampling mean variance.

An alternative approach to use the Gaussian process predictor models to interpolate D_t is to interpolate the mean uncertainty through the variance ($\sigma^2_{\mu_{D_t}}$) of the estimated mean value. This variance may be used to estimate confidence intervals (CI) in the mean.

The design procedure to structural fatigue requires multiple x points to be assessed.

Each estimate at x, depending on s_0, has an associated uncertainty. The surrogate can account for this local uncertainty using $\sigma_{\mu_{D_t}}$. Therefore, instead of sampling operational points from $G(x)$, a full surface may be sampled at a time and then used to predict D_T Teixeira et al. [2019d]. This surface will comprise a range of mean values that are possible realisations of the true population mean.

To note that CIs are not an estimate of variability of the sample mean, but it is reasonable to understand these as an estimate of plausible values for the population mean.

This process can be repeated up to a very large number of extracted surfaces (n_{surf}). Each surface, which is a deterministic prediction from $G(x)$, accounts for potential values of μ_{D_t} within the CI estimations. As a result, these surfaces are expected to enclose an interval that has plausible values of the population mean value within the CI boundaries chosen. It can then be used to research on the variability of the D_T estimation using $G(x)$. The procedure of sampling multiple surfaces is equivalent to repeating the design process multiple times assuming that the surface estimation is the true realisation of the population parameters for each surface. In reality, there is only a true surface. This allows the definition of a measure of uncertainty for the results attained with $G(x)$.

The idea of interpolating σ and consequently the CI can be implemented using the noisy approximation with:

$$\tau^2 = \sigma^2_{D_t} \tag{8.17}$$

The sample mean variance is given by

$$\sigma^2_{\mu^2_{D_t}} = \frac{\sigma^2_{D_t}}{N} \tag{8.18}$$

which can be used to calculate the CI according to the Cox's parametric formulation for the lognormal CI,

$$CI = \mu_{D_t} + \frac{\sigma^2_{D_t}}{2} \pm Z_{1-\alpha}\sqrt{\frac{\sigma^2_{D_t}}{N} + \frac{\sigma^4_{D_t}}{2(N-1)}} \tag{8.19}$$

N is the sample size used to converge μ_{D_t} and σ_{D_t}.

As in the present implementation 10 simulations with distinct seeds were performed for each DoE point, the CI are calculated based on this value. Due to sample size, the t-student distribution for N-1 degrees of freedom (DOF) is required to estimate the sample mean uncertainty interval. The critical value for the 95% CI of a t-distribution with 10-1 DOF is 2.262. According to Olsson [2005] this is one of the alternative procedures to estimate the confidence intervals when analysing a lognormal distribution. The most commonly applied approach to define the lognormal CI is the so-called "naive" method. Teixeira et al. [2018b] use this method to implement a similar procedure. The "naive" method involves less computations but may wrongly estimate the CI for the population mean. This occurs because it disregards the contribution from the σ in the lognormal mean.

Figure 8.19 shows the results for a noisy Gaussian predictor that interpolates CI. The idea of using the local DoE σ is to interpolate locally the CI, attending to its heterogeneity over x. It is highlighted that lack of points in some regions of x may significantly overestimate the CI prediction in that region.

Nonetheless, this interesting capability of the Gaussian process predictor may be of interest to account for the local uncertainty when applying the methodology proposed to calculate SN fatigue. It allows the consideration of potential convergence errors without any additional computational efforts. Moreover, the interest of having an estimate of

uncertainty increases when computational time is highly relevant. In such scenarios, errors due to sample size and interpolation are more likely to occur as the computational budget is limited.

Results for Figure 8.19 were computed using a noisy $G(x)$ for the 10^{th} iteration of the cases I-III from Figure 8.14. In all cases the error in the estimation of D_T was under 10 % (respectively 6.2%, 3.5% and 9.8%).

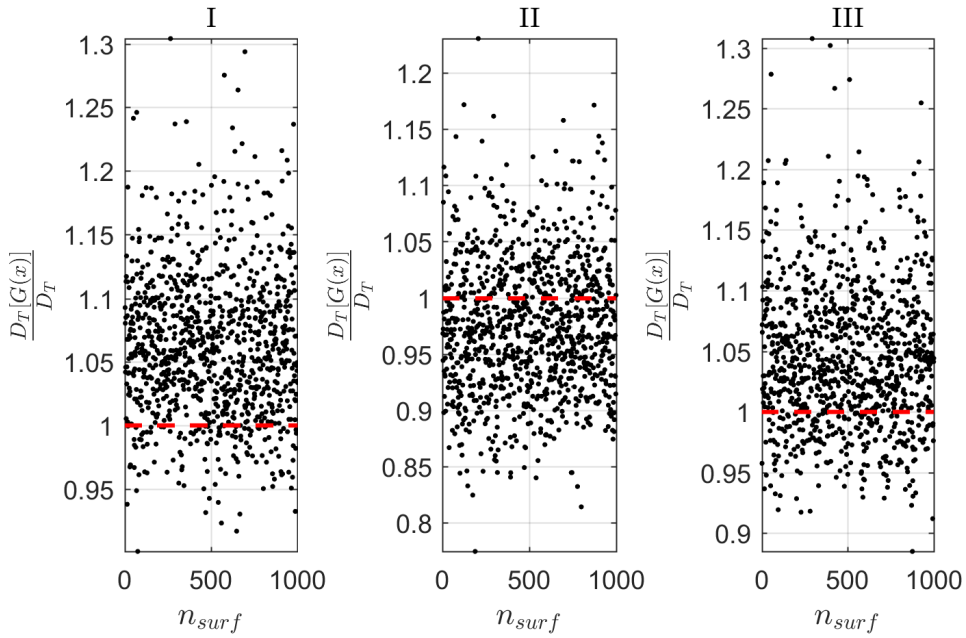


Figure 8.19: Variability of D_T estimation using a noisy $G(x)$ model to interpolate the 95% CI intervals in x. 1000 surfaces were sampled from $G(x)$. Trimmed (red) line represents the calculated value for D_T. Results were computed using the 10^{th} iteration from Figure 8.14. In I-III every point corresponds to the D_T predicted by a surface extracted from $G(x)$.

For I to III plotting multiple surfaces encloses the true value of D_T even when uncertainty is assessed using a small n_{surf}. In all the cases the considered iteration of $G(x)$ gave a relatively small error in the D_T prediction. When this is not the case, plotting the CI is not sufficient to have an estimation of the interval that may enclose the D_T. This approach is of more relevance as a complementary assurance of uncertainty in case the design procedure underestimates the SN design fatigue.

An alternative to sampling multiple surface is to use the enveloping surface of $G(x)$. However, note that CI estimated with such an approach may lead to a very conservative design interval. It is unlikely for D_t to be on the limit of the CI for all the 95% cases where the CI will enclose the mean.

8.6 Conclusions

The present chapter discussed a methodology for the assessment of SN fatigue using metamodelling. SN fatigue of the tower component of an offshore wind turbine installed on a monopile foundation was researched.

A meta-model was implemented as a surrogate of the short-term SN damage rates calculated as per recommendation of IEC61400. Using its interpolation capabilities it is possible to interpolate the SN fatigue predictions without the need to run the computational expensive simulations that define it. Only a limited number of points needs to be assessed to achieve accurate prediction of the long-term fatigue. The ultimate goal of using such models is to significantly reduce the amount of computational effort devoted to the fatigue calculation in the design phase, facilitating optimization procedures. It was seen in previous chapters that the amount of effort needed to assess fatigue is a major concern in the design of wind turbines.

According to the IEC standard, the SN method is recommended to design OWT to fatigue. Fatigue is expected to be assessed using time-domain simulations of complex codes at different operational conditions. The calculations for fatigue use then a counting algorithm, a stress-cycle curve and the well-established Palmgren-Miner rule. The meta-model is defined as surrogate of the local damage values calculated using this procedure.

Gaussian process predictors, also known as Kriging models, were applied as a meta-model to surrogate the SN fatigue surface. These already showed high flexibility to surrogate complex engineering problems. Moreover, they comprise a statistical measure of uncertainty in the approximation that was identified to be of relevance to achieve accurate predictions.

When applying a meta-model as a surrogate of a complex model the first approach is to understand the problem in-hand, and preferentially focus on its most relevant variables. For the tower component, mean wind speed and turbulence intensity were identified to be the most relevant variables in regard of the tower's SN damage rates variability and uncertainty. A modified indicator used as a measure of statistical distance was proposed in order to measure sensitivity, enclosing uncertainty.

Using these two wind variables, meta-modelling of SN damage rates was researched. Results obtained using the traditional approach that bins data to design for fatigue were compared with the results from meta-modelling. LHS and an innovative, powerful infill criterion were compared as alternatives to select the DoE of the Gaussian process predictor. The infill criterion tackles some of the limitations identified when applying LHS. It considers a notion of improvement to select new points that are expected to improve the SN fatigue surrogate prediction. All the results were validated using a comparative, computationally expensive, full 1-year loading assessment.

The main conclusions to highlight from the presented work are remarked in the following bullet points.

- The proposed modified sensitivity indicator is of relevance to assess sensitivity when considering the statistical characteristics of the design of experiments. In some problems, such as the case of OWT, and due to the intrinsic non-linear behaviour of the system, the same set of input characteristics generates a statistically characterized output. It is built on the Kullback-Leibler discrimination and accounts for uncertainty in the analysed output. Results showed that this indicator successfully characterizes sensitivity while accounting for uncertainty. Nevertheless, its potential interest significantly increases for research of sensitivity in reliability problems.

- The LHS has been widely applied before to assess surrogate modelling using Gaussian predictors. Depending on the problem, it may be of interest to use a correlated LHS sample to define the DoE. In average, for fatigue calculations, the convergence of the prediction when compared to a full year assessment was better for the correlated LHS. Uncorrelated LHS were identified to potentially induce large errors in the

estimation of the SN fatigue life. Due to the lack of relationship between variables, uncorrelated LHS may be highly dependent on the LHS generation algorithm. As a result, no relation to the physical problem of fatigue exists in uncorrelated LHS. For SN fatigue design, some operational points have significant more contribution to decrement fatigue life than others. If these are not considered in the DoE, or too much weight is given on the fitting to these, fatigue design may comprise significant errors.

- The proposed infill criterion predicts accurately long-term fatigue using a reduced amount of computational time. When comparing with the traditional approach that bins environmental conditions, reduction of computational time by a factor of 4 to 10 without compromising accuracy was achieved using this infill criterion. It produces robust results, mainly due to the multi-step definition of the surrogate. It held accurate prediction for the challenging case where a minimum amount of information was given for the search procedure. Convergence and (computational) efficiency increased when more information was conceded to the search procedure in the initial step.

- When applying the Gaussian process predictor to approximate complex problems, it is of relevance to have some notion of improvement in the meta-model definition. The definition of the meta-model should be related to the physical problem being analysed. Every iteration should be selected in order to improve the approximation of the meta-model predictions to the real model. This implies establishing indicators from which a search function to learn. In the current application these were related to, first, the requirement to cover high SN damaging areas, and second, the need to have enough spread of points to define the surrogate.

- The surrogate definition using a learning infill criterion (it learns on the current surface), or a LHS scheme, varies in accuracy due to convergence of the model hyperparameters, sample convergence and learning calculations (calculation of the EI). Due to this fact, convergence to the true value of SN damage may occur within different number of iterations. Therefore, it is important, when applying these models to a particular problem, to develop validation techniques so that a minimum sample size that guarantees accurate predictions can be defined.

- Alternatively, full advantage can be taken from the Gaussian process predictors for SN fatigue calculations if a noise component is used to interpolate confidence intervals in the mean. This allows the considerations of the uncertainty in the estimated SN fatigue life while not increasing the computational cost of the assessment.

- Future developments should research on the influence of convergence thresholds used to estimate the meta-modelling hyperparameters. Different weights depending on the influence of the variable may contribute to improve convergence in both the learning method, as well, in the LHS case.

- Finally, the methodology proposed to assess SN fatigue has universal character. It can be implemented to calculate the fatigue of any component using the SN fatigue damage method. If implemented for large m materials, attention should be given to the definition of the SN damage indicator used to build the surface. Furthermore, it can be applied for fatigue assessment using other methodologies and account for their statistical behaviour, *e.g.* crack growth.

Chapter 9

Conclusions

Contents

9.1 Conclusions

In order to enable a better understanding to the reader, the different conclusions from the work developed are discussed in the context of the proposed goals for the presented PhD thesis. Three major goals were highlighted *a priori*, each one of them relating to a distinct topic.

9.1.1 Conclusions with respect to extrapolation of physical quantities

Significant attention in the developed work was given to the necessity of dealing, in a design framework, with the statistical extrapolation of physical quantities. Statistical extrapolation is a procedure highly relevant for engineering design.

The following points highlight the major conclusion from the analysis developed.

- For extrapolation of significant wave heights, results of the Goodness-of-Fit analysis and extrapolation showed that there is no evidence to reject the Generalised Pareto distribution over the two-parameter Weibull and the Exponential distributions. Moreover, the support of the Generalised Pareto distribution and its bounded character is expected to contribute to a physically more realistic statistical characterization. Nevertheless, attention should be given to the selection and fitting process when applying Peak-over-threshold to model *exceedance* data. Results showed that when using this approach the user should be careful during the selection of the threshold level that will be used to truncate the data.

- The threshold level is the variable of major influence in the outcome results of the extrapolation with the peak-over-threshold. The method used to ensure independence

and the number of field points do not show the same decisive influence in the results as the threshold. A contribution to the choice of threshold was added through the introduction of a new methodology to define the threshold level in a straightforward way. It consists in settling the threshold value by analysing the three derivatives of the cumulative density function.

- For a studied data set, it may be impossible to generalize findings from which statistical distribution is going to produce a better fit. The process should be evaluated in conjunction with a sensitivity analysis of the parameters of the study; threshold level, decoupling time, and fitting methodology; so that a definitive conclusion can be drawn. No indicators to reject any of the distributions researched were found. However, strong evidence was found that backs the fact that the statistical support of the theoretical models analysed are highly important in the results.

- Using a logarithmic transformation of the data when applying the Generalised Pareto distribution to model a physical limited phenomena, such as significant wave heights, was adequate for the present case, and therefore must be considered for future studies of wave data and other variable.

- If the interest is to characterize the statistical tail region, limited interest exists in using the full data-set available in order to characterize the tail region. The "bulk" of the data is usually locate around the mean region, and its quantity of points is likely to bias the tail region. Even truncating the data at high threshold values, strong overestimation of the results may occur. As a example, truncation of the fatigue loading data at the quantile of 95%, which is recommended by IEC61400, frequently produces an heavy tail (estimated tail has more probability density than the real tail) that strongly overestimates the SN fatigue damage.

- Finally, the interest of the presented probabilistic-based methodologies to extrapolate physical variables highlights the relevance of the role of the different parameters in the extrapolation process. The methodologies implemented can be used to produce robust results on a design basis approach.

9.1.2 Conclusions with respect to uncertainty based on design procedures

A significant part of the presented work addressed uncertainty originated by the statistical procedures. Despite being highly related with the previous topic, some additional conclusions are highlighted.

- The individual application of probability extrapolation may not be sufficient to produce accurate prediction results. An adequate sample size is of large relevance to have accurate statistical extrapolation and characterization of the tail region. This is of particular importance when applying rules of thumb, such as the quantile thresholds, that have limited theoretical background Embrechts et al. [1999] and that may lead to erroneous extrapolated values. Modelling considerations, such as the support or number of free-parameters to fit, are of large relevance to produce accurate and realistic predictions in the tail, that are of major interest when dealing with physical variables, such as the loading. The influence of these is not frequently fully addressed.

- In the case of SN fatigue, long-term accurate estimations were seen to depend mainly on the short-term damage estimation mean value. If an accurate short-term damage rate mean is defined for a certain operational state, so it is the long-term estimation of the SN fatigue damage for that same operational state. This is true regardless of how large the long-term is, whether it is 1 year or 100 years. As a result, no matter what statistical procedures are applied, if a representative sample is not established, the long-term SN fatigue damage estimations are not robust. Significant errors in this regard may be committed for large SN slope materials. For these, a SN fatigue representative sample needs to be larger. When considering the design standard, the recommended sample size of a minimum of six seeded simulations, does not produce consistent results.

- For small SN slopes, limited interest was found in the application of complicated statistical fitting techniques. In a design procedure, fitting techniques for such materials consistently introduce uncertainty in the SN fatigue estimations. This occurs because the statistical theoretical models are used to approximate a spectra, instead of being used to quantify statistically the variability of the loading variable. Still, their use may be of interest to perform parametric analysis. However, the uncertainty resulting from the fitting needs to be addressed (in particular in the tail region, where, even for low SN slope, important SN damage density exist).

- On the other hand, for large SN slope values (*e.g.*, $m = 10$) these are of interest, and should be applied (*e.g.*, considering an adequate threshold value) so that an accurate approximation of the long-term fatigue is achieved. For large SN slope components, most of the fatigue life is depleted by the large load ranges that are described by the tail of the loading spectra. Therefore, an adequate approximation of large load ranges is of major relevance for these. Nonetheless, probability extrapolation, for reference, is on the safe side of the calculations as it increases the current predictions for the short-term damage (for the same number of cycles in the tail region).

- It was shown how sample size limitation could be tackled by using non-parametric bootstrapping techniques. These allow the estimation confidence intervals of the true mean, and to mitigate errors from sample size limitations. The asymptotic values for the confidence intervals obtained using the percentile method were accurate for low SN slope materials. For high SN slope materials the bootstrapping methodologies slightly under-predicted the confidence level. These asymptotic values were identified to be stable regardless of the sample size. This is of interest in a design procedure. For small samples the range of the confidence intervals may be too large for appropriate consideration on a design basis. This limitation can be mitigated by simply increasing the sample size. The increase of the sample size from 6 to 25 allowed the reduction of the confidence interval by approximately 50% while maintaining the confidence level estimate.

- Block bootstrapping techniques were also implemented in order to infer on the effect of the autocorrelation of the time series. However, block bootstrapping applications to fatigue time series are a challenging. While the time series have an autocorrelation function with several significant lags, the RC load ranges, despite being related to the time series, are purely random. Previous works that applied block bootstrapping for fatigue purposes were identified, however, in the current work limited interest was found in their application when comparing with the empirical bootstrapping.

Only a slightly improvement of the uncertainty estimation was achieved. Furthermore, if not applied correctly, when applying the RC algorithm, important load ranges may be eliminated from the analysis. Fatigue estimations, in such a case, are underestimated.

- Finally, it is important to highlight that little interest has been given to the potential uncertainty generated by the heterogeneity of design procedures, such as consideration of small sample sizes. Such uncertainty is relevant when calibrating OWT safety factors, in particular when statistical convergence is inherently imposed by standard recommendations.

9.1.3 Conclusions with respect to effort reduction for design procedures

Designing effort was identified to be one of the major drawbacks of OWT design, and in particular for some concrete cases, such as fatigue design. OWT design is strongly supported by computational methods. Therefore, strong focus was given in the current work to merge the current design practices with some new techniques that have been recently applied to cut the computational cost required to analyse complex codes. Meta-modelling was researched as an alternative to interpolate the currently recommended fatigue design procedure. A meta-model that surrogates the fatigue design is of relevance because it allows the prediction of the operational damage for points not assessed with the OWT model.

The main conclusions of the implementation performed are the following:

- Global sensitivity analysis are a powerful tool to decrease the cost needed to solve an engineering problem. Within a complex problem, it is common for only a limited number of variables to comprise for almost all the influence in the problem's output. By conducting a sensitivity analysis, it is possible to better comprehend the problem in-hands, and more efficiently analyse it. In the present implementation, sensitivity analysis was performed with the ultimate goal of reducing the dimension of a meta-model. Nevertheless, relevant informations about the problem in-hands, which would then later be used to build an improvement criterion, were output from the global sensitivity analysis procedure.

- A transformation was proposed to the Kullback-Leibler discrimination measure in order to make it more suitable for application in a probabilistic global sensitivity analysis, and in a reliability or uncertainty characterization context. A similar transformation can be applied to the Kullback-Leibler divergence, which is a statistical distance measure. The proposed transformation makes the Kullback-Leibler discrimination completely symmetric and monotonic in different domains of interest in order to reflect symmetrically the sensitivity between the distributions. This measure is based on a relative comparison between statistical distributions, and complements the information given by the former.

- A meta-modelling technique was successfully implemented in order to reduce the computational time of the fatigue design for OWT analysis. It uses a Gaussian process predictor to interpolate the SN fatigue damage from different operational states. It was applied to the tower component, but the same methodology can be extended to any other component.

- Different approximation to the design variables for the meta-model were researched. A "notion" of improvement with regard to the pursued approximation may enable efficient application of surrogate procedures. Such "notion" can be attained by using an infill criterion that picks the next points in the design of experiments and that relates to the specificity of the analysis being implemented. In the case of fatigue, this was achieved by relating the local damage rates with the joint-probability density function.
- The new search procedure successfully implemented to meta-model fatigue was identified to reduce the computational cost of the SN fatigue procedure, without significantly compromising uncertainty, by up to 80%. Nevertheless, for the same accuracy the reduction in computational time was never less than 60%.
- As the last finding of relevance, it is highlighted that the intrinsic characteristics of the Gaussian process predictor allow the estimation of the deviation in the design procedure. This is achieved by sampling multiple deterministic surrogates from a single Gaussian process predictor, which characterizes potential deviations in the design procedure. If the Gaussian is an accurate predictor of the short-term SN damage rates, so it is also, of the associated uncertainty. In such a scenario, the Gaussian predictor should be built as an uncertainty interpolator of the short-term damage rate mean value.

9.2 Future works

Due to the innovative character some of the work developed, many new lines of research emerged. The main are highlighted in the following points:

- It would be of interest to study how the new methodology used to define the threshold for the Peak-over-threshold may compare with other alternative techniques. The procedure proposed to estimate the threshold of the peak-over-threshold method uses the distribution function shape to establish the threshold level. This is of relevance to match the *exceedances* distribution and the tail region. However, such a method may not hold for the causes where the data do not have the bell shape characteristic of many distribution functions. It was seen to produce accurate results for significant wave heights. Moreover, it reduces the analyst bias. It would be of interest to study its robustness within the different methods used to define the threshold value.
- The analysis of significant wave heights and the associated return periods showed that special attention should be given to the model properties when fitting the tail with a multi-parameter distribution. Return levels may be underestimated or overestimated due to them. The present work is representative of a data set of the Irish coast, which in turn is representative of the north Atlantic and the Irish Sea. It is of interest to extended the presented analysis to other regions.
- The uncertainty quantification methodology presented that uses bootstrap results failed to asymptotically approach the confidence levels for the confidence intervals estimated when the SN slope was large ($m = 10$). It would be of interest to study on how the bootstrap performance could be improved for the case of large SN slopes. Uncertainty quantification is highly relevant for these. Studying statistical extrapolation uncertainty using bootstrap is one of the alternatives. Correcting the bootstrap results due to bias is another example. There is, hence, an extensive set of

methodologies that may be researched in order to improve the performance of the bootstrap for large SN slope materials.

- Statistical uncertainty was addressed using the tower component. It would be of interest to expand it to the blade component and monopile foundation critical components in order to understand the robustness of the methodology. This was partially addressed in one of the publications listed in Section 9.3.

- Further research on meta-modelling should research the accuracy of the long-term fatigue estimation when reducing the sample size convergence. A direct sample size reduction from 10 to 6 simulations may contribute to a further decrease of computational time by 40%. However, the effects on the interpolation accuracy should be addressed.

- The ultimate goal of the methodology presented that reduces the computational cost of the SN fatigue design is to implement an optimization procedure. The natural next step for the meta-modelling proposed is then to further exploit optimization. It is of relevance to identify what the gains of the implemented meta-modelling are in an optimization procedure. Different prediction errors can be assumed at different iteration steps. An optimum combination of computational convergence (computational budget and size of samples per iteration) within the meta-modelling is expected to exist, and should be identified so that the effort reduction methodology presented can be exploited to its full potential.

- The infill methodology researched for meta-modelling of fatigue showed remarkable accuracy and convergence. Nevertheless, better convergence stopping criteria may allow further increase of the efficiency of the method. A computational budget sufficed the iterations performed. However it was identified that a minimum of 4 iterations could be disregarded and the meta-model prediction would still be accurate. In some cases, up to half the computational budget could be reduced if the infill methodology was able to identify the first new iteration that guaranteed accurate predictions.

9.3 List of publications

The results of the present PhD thesis have been presented and disclosed in the following journal and conference papers:

Journal Papers

- *Teixeira, R., Nogal, M., O'Connor, A.* (2018) On the suitability of the generalized Pareto to model extreme waves. Journal of Hydraulic Research, 56(6): 755-770. (doi.org/10.1080/00221686.2017.1402829)

- *Teixeira, R., O'Connor, A., Nogal, M.* (2019) Probabilistic Sensitivity Analysis of OWT using a transformed Kullback-Leibler discrimination. Structural Safety, 81: 101860 (doi.org/10.1016/j.strusafe.2019.03.007)

- *Teixeira, R., Nogal, M., O'Connor, A.* (2019) Analysis of wind turbine operational fatigue design for low SN slope materials. Wind Energy; 1-18 (doi.org/10.1002/we.2389)

- *Teixeira, R., Nogal, M., O'Connor, A., Nichols, J., Dumas, A.* (2019). Stress-cycle fatigue design with Kriging applied to offshore wind turbines. International Journal of Fatigue, 125: 454-467 (doi.org/10.1016/j.ijfatigue.2019.04.012)

Conference Papers

- *Teixeira, R., O'Connor, A., Nogal, M.* (2016). Comparative Analysis of the Probabilistic Methods to Estimate the Probability of Failure of Offshore Wind Turbine Towers. CERI2016. Galway, Ireland.
- *Teixeira, R., O'Connor, A., Nogal, M., Nichols, J. Spring, M.* (2017). Structural probabilistic assessment of Offshore Wind Turbine operation fatigue based on Kriging interpolation. ESREL2017. Portoroz, Slovenia.
- *Teixeira, R., O'Connor, A., Nogal, M., Nichols, J. Krishnan, N.* (2017). Analysis of the design of experiments of offshore wind turbine fatigue reliability design with Kriging surfaces." Procedia Structural Integrity 5 (2017): 951-958.
- *Teixeira, R.,O'Connor, A., Nogal, M.* Interpolation of confidence intervals for fatigue design using a surrogate model. IFED2018, Lake Louise, Canada
- *Teixeira, R.,O'Connor, A., Nogal, M.,* (2018). Application of Gaussian process regression for structural analysis. CERI2018. Dublin, Ireland
- *Teixeira, R., Nogal, M., O'Connor, A.* (2018). On the calculation of offshore wind turbine load spectra for fatigue design. Porto, Portugal.
- *Teixeira, R., O'Connor, A., Nogal, M.* (2018). Convergence analysis for Offshore Wind Turbine operational fatigue reliability analysis. IALCCE2018. Ghent, Belgium.

Part IV

Appendix

Appendix A

This appendix presents complementary results and discussion on the comparative analysis of procedures for SN fatigue calculations presented in Section 7.3.6. 25 LHS operational points were computed in order to complement the analysis performed with the NTM.

It was seen that the methodology presented by IEC for extrapolation of fatigue calculations divides the estimation procedure in above and below u. This methodology is presented as one of the alternatives for the calculation of D_T. When truncating the data for fatigue calculations, the accuracy of the results highly depends, in addition to the loads, on the number of *exceedance* cycles, e.g., case 8 in Table A.1. As a result, the problem of extrapolation for fatigue is not only a problem of accurate prediction of T loads, which is itself challenging, but also a problem of accurate definition of the expected *exceedance* cycles in the tail region (M) at the time T.

The estimation of the T loading for fatigue calculations using a $Q99$ as value of u, with exception of simulation 4, 6 and 14, with respective errors of extrapolated L of approximately, 30%, 20% and 10%, did not produce significant errors in the extrapolated L when using the Weibull distribution. And with exception of the cases 4, 7, 11, 14 when using the GP distribution. Nevertheless, much larger estimation errors of SN damage can be identified derivated from its dependence also on the cycle counts for the tail.

If the columns relative to the extrapolated number of cycles are compared with the error in damage, an association can be found between the error in the SN fatigue estimation with the error of extrapolation in L and M. In case 4, despite the good approximation of the number of cycles in the tail region, the results of the tail prediction are underestimated due to the underestimation of the extrapolated L. In case 8, the extrapolated L is relatively accurate (error of 6.5% for the GP and 3% for the Weibull extrapolation), however the error in the expected number of cycles in the tail region is significant (approximately 50%), resulting in an under-prediction of the SN fatigue damage in the tail region by approximately 40%.

Comparison of the expected *exceedance* cycles in Tables A.1 and A.2 shows that an $s_0 = 6$ may not be adequate to produce accurate results for M. As s_0 increases, with the usage of a fixed percentile to set the value of u, the statistical region above u becomes increasingly more populated with occurrences. The u values becomes more stable and the approximation to the real number of *exceedance* cycles becomes more accurate.

Table A.3 presents the results of bootstrapping for different operational conditions relative to the analysis of uncertainty discussed in the same chapter.

Table A.1: Parameters of fatigue design for 25 LHS points, comparing 100 estimations of D_t with extrapolating from a $s_0 = 6$. E refers to the exponential distribution that was considered for complementarity. α is the rate parameter. K are the cycles below the u value, and M above. The index R refers to the real values (e.g. L_R is the real load).

Point	**Stat. Parameters**					**Number of cycles**				max L **(MNm)**			
	ζ_{GP}	σ_{GP}	ζ_W	σ_W	α_E	K	M	K_R	M_R	L_{GP}	L_W	L_E	L_R
$u = Q95$													
1	-0.11	1316	0.91	1140	1187	7875	308	7750	453	10.84	12.94	11.99	10.56
2	-0.22	2217	0.83	1670	1803	7954	296	7972	284	11.73	18.11	14.77	11.25
3	0.00	3544	0.84	3263	3551	8271	308	8251	333	26.91	32.82	26.83	19.39
4	-0.56	7079	1.17	4718	4504	8508	300	8614	314	18.69	27.55	32.27	28.21
5	-0.73	9184	1.07	4839	4714	7612	283	7676	275	17.78	29.73	31.98	20.61
6	-0.22	7420	0.95	5863	5998	7854	300	7922	297	29.66	42.5	39.89	25.98
7	-0.68	11203	1.35	6791	6205	7817	279	7886	294	23.04	31.26	41.85	27.19
8	-0.45	9662	0.94	6230	6387	8071	279	7945	327	30.27	49.85	46.54	30.35
9	-1.09	13454	1.41	6469	5944	6833	233	6682	255	24.39	33.66	44.53	25.91
10	-0.27	4108	1.08	3300	3212	5679	183	5720	158	34	37.95	39.43	32.94
11	-1.08	9512	1.13	4059	3923	6158	204	6039	272	31.01	39.94	43.08	35.21
12	-0.64	7249	1.23	4478	4205	6358	225	6281	204	38.31	44.97	50.15	39.95
13	-0.25	6022	1.07	4904	4786	4929	154	4972	192	30.74	36.41	38.48	31.2
14	-0.64	12528	1.1	7248	7001	7262	279	6855	359	36.34	52.23	56.8	40.09
15	-0.27	3907	1.34	3364	3092	6800	258	6757	298	20.33	21.16	26.29	19.72
16	-0.38	4081	1.22	3129	2932	6492	233	6679	250	20.74	23.84	27.26	22
17	-0.73	6782	1.29	4081	3801	7562	275	7471	291	21.36	27.81	33.53	22.63
18	-0.29	4988	1.08	3936	3825	7596	283	7602	306	28.19	33.81	35.83	28.71
19	-0.37	2936	1.18	2240	2119	7079	233	7168	269	15.8	18.32	20.49	15.95
20	-0.2	2537	0.91	2008	2087	7279	292	7546	250	20.12	25.15	23.45	19.27
21	-0.1	2119	1.02	1947	1928	7717	292	7790	251	20.6	21.96	22.28	18.66
22	-0.44	4202	1.33	3155	2926	8008	288	8031	275	22.37	25.2	30.18	24.28
23	-0.41	3001	1.23	2243	2105	7292	254	7311	299	15.58	18.03	20.68	15.11
24	-0.58	3300	1.22	2190	2076	7746	258	7582	316	15.78	19.24	21.87	17.67
25	-0.35	3227	1.1	2421	2331	7808	296	7902	278	20.15	23.85	25.43	20.59
$u = Q99$													
1	-1.29	2318	1.08	957	933	8146	38	8154	49	9.24	10.6	10.84	10.56
2	0.27	642	0.85	777	849	8212	38	8223	33	11.87	11.51	11.04	11.25
3	-1.21	7018	2.06	3631	3228	8537	42	8548	36	18.74	19.86	25.03	19.39
4	-0.42	2724	1.26	2023	1883	8762	46	8881	47	18.65	19.33	20.68	28.21
5	-1.6	7244	2.41	3248	2866	7854	42	7918	32	17.54	18.62	23.7	20.61
6	-1.7	9257	0.96	3323	3359	8117	38	8194	24	24.15	31.35	30.89	25.98
7	-1.34	6264	1.22	2732	2621	8058	38	8147	32	22.56	25.82	27.44	27.19
8	-1.25	9062	1.97	5251	4722	8312	38	8199	73	28.38	31.3	38.33	30.35
9	-1.54	4440	2.38	2119	1889	7029	38	6918	19	24.41	25.17	28.37	25.91
10	-1.12	5929	1.34	2644	2413	5837	25	5864	14	32.84	34.36	36.09	32.94
11	-1.3	4928	1.45	2700	2511	6333	29	6246	65	31	33.67	36.06	35.21
12	-1.28	5925	3.41	3392	3041	6554	29	6440	44	38.05	38.37	44.13	39.95
13	0.03	2889	0.99	2984	2991	5063	21	5145	19	32.28	32.04	32	31.2
14	-1.13	4231	1.19	1622	1530	7508	33	7178	36	35.8	36.8	37.49	40.09
15	-0.42	3496	1.37	2647	2392	7025	33	7018	37	19.9	20.06	21.85	19.72
16	-1.27	5050	1.21	1954	1834	6692	33	6895	34	20.09	21.7	22.61	22
17	-1.1	4258	1.75	1972	1743	7800	38	7729	33	21.22	21.54	23.73	22.63
18	-1.36	8662	1.88	4055	3588	7842	38	7855	53	25.91	27.64	32.59	28.71
19	-1.1	3320	1.67	1457	1287	7279	33	7414	24	15.36	15.49	16.92	15.95
20	-0.4	1820	0.85	1160	1249	7533	38	7764	32	18.94	20.71	19.96	19.27
21	-1.09	3354	0.97	1325	1340	7971	38	8002	39	17.89	19.9	19.73	18.66
22	-0.26	1868	1.18	1558	1466	8258	38	8268	38	22.34	22.58	23.24	24.28
23	-1.16	3707	1.12	1353	1294	7512	33	7576	34	14.95	15.94	16.34	15.11
24	-1.15	3050	1.55	1242	1110	7962	42	7842	56	15.66	15.94	17.18	17.67
25	-1.16	3786	1.8	1763	1562	8067	38	8150	31	19.12	19.5	21.57	20.59

Table A.2: Comparative results for 25 LHS sampled points, comparing estimations of D_T resulting from 50 simulations using a $s_0 = 6$. d refers to the contribution to damage of the loads under u, and D to above u. The index R refers to the "real" or reference value in T.

	$m=3$						$m=5$						*double m*					
Point	d	D_{GP}	D_W	D_E	d_R	D_R	d	D_{GP}	D_W	D_E	d_R	D_R	d	D_{GP}	D_W	D_E	d_R	D_R
									$Q95$									
1	0.87	0.866	0.95	0.903	1	1	0.796	1.064	1.42	1.219	1	1	0.796	1.064	1.42	1.219	1	1
2	1.045	1.555	2.25	1.791	1	1	1.06	2.038	5.779	3.148	1	1	1.06	2.038	5.779	3.148	1	1
3	1.065	1.998	2.574	1.994	1	1	1.094	4.195	7.964	4.167	1	1	1.094	4.195	7.964	4.167	1	1
4	0.912	1.383	1.761	2.14	1	1	0.872	1.35	2.755	4.532	1	1	0.872	1.35	2.755	4.532	1	1
5	0.975	2.022	2.692	2.979	1	1	0.967	2.482	6.537	8.363	1	1	0.967	2.482	6.537	8.363	1	1
6	0.964	1.715	2.552	2.313	1	1	0.938	2.624	7.157	5.715	1	1	0.938	2.624	7.036	5.701	1	1
7	0.947	1.552	1.696	2.46	1	1	0.944	1.755	2.6	6.637	1	1	0.944	1.755	2.6	6.561	1	1
8	1.025	0.995	1.464	1.34	1	1	1.043	1.124	3.434	2.748	1	1	1.043	1.124	3.119	2.608	1	1
9	1.205	1.432	1.522	1.996	1	1	1.326	1.85	2.481	5.274	1	1	1.326	1.85	2.481	5.103	1	1
10	1.002	1.399	1.437	1.46	1	1	1.004	1.605	1.743	1.83	1	1	1.004	1.605	1.743	1.829	1	1
11	1.052	0.898	0.899	0.937	1	1	1.082	0.999	1.096	1.236	1	1	1.082	0.999	1.095	1.216	1	1
12	1.007	1.219	1.241	1.300	1	1	1.004	1.294	1.39	1.594	1	1	1.004	1.295	1.337	1.434	1	1
13	1.034	1.158	1.264	1.327	1	1	1.065	1.529	2.021	2.33	1	1	1.065	1.529	2.021	2.33	1	1
14	0.994	1.036	1.17	1.264	1	1	1.007	1.153	1.841	2.293	1	1	1.007	1.155	1.564	1.775	1	1
15	1.06	1.358	1.348	1.582	1	1	1.115	1.962	1.924	3.103	1	1	1.115	1.962	1.924	3.103	1	1
16	1.083	1.166	1.21	1.317	1	1	1.14	1.362	1.541	1.999	1	1	1.14	1.362	1.541	1.999	1	1
17	1.156	1.376	1.473	1.685	1	1	1.234	1.729	2.193	3.27	1	1	1.234	1.729	2.193	3.27	1	1
18	1.006	1.226	1.311	1.363	1	1	1.037	1.519	1.892	2.131	1	1	1.037	1.519	1.892	2.131	1	1
19	1.108	1.144	1.192	1.274	1	1	1.13	1.421	1.625	1.994	1	1	1.13	1.421	1.625	1.994	1	1
20	1.012	1.359	1.486	1.428	1	1	1.019	1.546	2.04	1.812	1	1	1.019	1.546	2.04	1.812	1	1
21	1.062	1.342	1.368	1.377	1	1	1.1	1.553	1.652	1.686	1	1	1.1	1.553	1.652	1.686	1	1
22	1.055	1.301	1.329	1.466	1	1	1.087	1.513	1.617	2.158	1	1	1.087	1.513	1.617	2.158	1	1
23	0.98	1.182	1.223	1.326	1	1	0.98	1.525	1.709	2.178	1	1	0.98	1.525	1.709	2.178	1	1
24	0.977	1.02	1.07	1.148	1	1	0.978	1.172	1.359	1.671	1	1	0.978	1.172	1.359	1.671	1	1
25	0.997	1.368	1.423	1.468	1	1	1.015	1.653	1.89	2.084	1	1	1.015	1.653	1.89	2.084	1	1
									$Q99$									
1	0.805	0.834	0.831	0.838	1	1	0.713	0.816	0.843	0.862	1	1	0.713	0.816	0.843	0.862	1	1
2	1.056	1.188	1.178	1.166	1	1	1.114	1.207	1.175	1.134	1	1	1.114	1.207	1.175	1.134	1	1
3	1.003	1.564	1.579	1.751	1	1	0.95	1.818	1.848	2.594	1	1	0.95	1.818	1.848	2.594	1	1
4	1.024	0.83	0.828	0.847	1	1	1.251	0.647	0.65	0.698	1	1	1.251	0.647	0.65	0.698	1	1
5	1.017	1.507	1.538	1.681	1	1	1.09	1.553	1.604	2.164	1	1	1.09	1.553	1.604	2.164	1	1
6	0.92	1.906	2.03	2.014	1	1	0.905	2.04	2.538	2.476	1	1	0.905	2.04	2.538	2.476	1	1
7	0.897	1.209	1.221	1.253	1	1	0.796	1.145	1.206	1.3	1	1	0.796	1.145	1.206	1.3	1	1
8	0.969	0.601	0.654	0.711	1	1	1.001	0.627	0.73	0.96	1	1	1.001	0.627	0.73	0.96	1	1
9	1.002	2.182	2.223	2.268	1	1	1.024	2.214	2.282	2.428	1	1	1.024	2.214	2.282	2.428	1	1
10	1.03	2.264	2.19	2.198	1	1	1.039	2.41	2.29	2.331	1	1	1.039	2.41	2.29	2.331	1	1
11	0.97	0.499	0.518	0.526	1	1	0.959	0.497	0.533	0.557	1	1	0.959	0.497	0.533	0.557	1	1
12	1.081	0.768	0.797	0.807	1	1	1.105	0.795	0.842	0.887	1	1	1.105	0.796	0.843	0.847	1	1
13	0.903	1.238	1.237	1.237	1	1	0.841	1.249	1.245	1.244	1	1	0.841	1.249	1.245	1.244	1	1
14	0.818	0.943	0.909	0.911	1	1	0.738	0.885	0.834	0.839	1	1	0.738	0.889	0.838	0.842	1	1
15	0.998	1.197	1.182	1.21	1	1	0.992	1.386	1.353	1.485	1	1	0.992	1.386	1.353	1.485	1	1
16	1.046	1.277	1.22	1.229	1	1	1.078	1.316	1.237	1.272	1	1	1.078	1.316	1.237	1.272	1	1
17	1.077	1.266	1.216	1.236	1	1	1.069	1.275	1.189	1.258	1	1	1.069	1.275	1.189	1.258	1	1
18	1.019	0.936	0.938	0.981	1	1	1.073	0.998	1.006	1.19	1	1	1.073	0.998	1.006	1.19	1	1
19	1.013	1.738	1.639	1.66	1	1	0.961	1.888	1.706	1.794	1	1	0.961	1.888	1.706	1.794	1	1
20	1.051	1.373	1.381	1.37	1	1	1.073	1.367	1.411	1.375	1	1	1.073	1.367	1.411	1.375	1	1
21	1.091	1.233	1.199	1.196	1	1	1.128	1.286	1.255	1.247	1	1	1.128	1.286	1.255	1.247	1	1
22	1.063	1.029	1.027	1.031	1	1	1.09	0.998	0.995	1.01	1	1	1.09	0.998	0.995	1.01	1	1
23	0.946	1.272	1.173	1.179	1	1	0.933	1.422	1.264	1.286	1	1	0.933	1.422	1.264	1.286	1	1
24	0.949	0.859	0.809	0.819	1	1	0.944	0.895	0.81	0.844	1	1	0.944	0.895	0.81	0.844	1	1
25	1.001	1.452	1.41	1.436	1	1	1.005	1.539	1.464	1.553	1	1	1.005	1.539	1.464	1.553	1	1

Table A.3: Results for the BB and MBB calculated fatigue damage confidence intervals for simulations 1-11 and for a $b = 120s$. Three estimations with random s_0 of six simulations were used to compute the results. C represents the upper limit, and c the lower.

		BB												MBB									
		$m=3$			$m=5$			$m=10$			*double m*			$m=3$			$m=5$			$m=10$			*double m*		
iteration		1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
$S1$	CD_T	0.97	1.11	1.19	1.06	1.55	1.33	2.29	1.86	1.50	1.06	1.55	1.33	0.95	1.07	1.2	0.89	1.41	1.48	2.22	1.4	2.43	0.89	1.41	1.48
	cD_T	0.59	0.85	0.91	0.49	0.8	0.82	0.39	0.39	0.54	0.49	0.8	0.82	0.62	0.85	0.93	0.5	0.75	0.86	0.53	0.17	0.62	0.5	0.75	0.86
$S2$	CD_T	1.07	1.25	1.05	1.28	1.75	1.14	1.60	2.93	1.26	1.28	1.75	1.14	1.04	1.24	1.05	1.18	1.64	1.09	5.36	2.59	1.47	1.18	1.64	1.09
	cD_T	0.77	0.88	0.84	0.56	0.82	0.65	0.48	0.54	0.28	0.56	0.82	0.65	0.81	0.87	0.82	0.63	0.81	0.61	1.42	0.29	0.3	0.63	0.81	0.61
$S3$	CD_T	1.26	1.36	1.28	1.56	1.71	1.49	0.98	1.46	0.54	1.49	1.72	1.56	1.22	1.31	1.17	1.47	1.67	1.18	1.04	1.49	0.66	1.45	1.68	1.26
	cD_T	0.84	0.74	0.81	0.69	0.59	0.62	0.21	0.33	0.12	0.68	0.61	0.65	0.75	0.77	0.75	0.57	0.6	0.48	0.26	0.35	0.16	0.6	0.63	0.51
$S4$	CD_T	1.06	1.02	1.21	1.18	1.25	1.49	1.36	1.09	2.38	1.14	1.1	1.35	1.15	1.18	1.23	1.20	1.42	1.52	3.09	1.22	2.22	1.21	1.32	1.35
	cD_T	0.71	0.67	0.76	0.53	0.56	0.61	0.32	0.37	0.51	0.59	0.57	0.66	0.81	0.71	0.76	0.68	0.66	0.66	0.71	0.37	0.67	0.75	0.67	0.7
$S5$	CD_T	1.34	1.1	1.1	1.57	1.08	1.08	1.11	1.19	2.15	1.41	1.09	1.11	1.22	1.08	1.04	1.36	1.05	1.00	1.14	1.1	2.27	1.33	1.07	1.03
	cD_T	0.93	0.84	0.87	0.92	0.75	0.77	0.61	0.63	0.91	0.94	0.76	0.81	0.87	0.81	0.83	0.86	0.71	0.72	0.7	0.61	0.95	0.87	0.74	0.76
$S6$	CD_T	1.1	1.17	1.26	1.19	1.24	1.38	1.21	1.52	1.14	1.15	1.23	1.35	1.15	1.26	1.27	1.31	1.36	1.41	1.18	1.55	1.09	1.22	1.34	1.38
	cD_T	0.88	0.9	0.92	0.86	0.86	0.87	0.46	0.71	0.57	0.86	0.88	0.88	0.88	0.95	0.96	0.86	0.94	0.93	0.48	0.67	0.53	0.87	0.96	0.95
$S7$	CD_T	1.21	1.08	1.26	1.42	1.14	1.56	1.92	0.55	0.58	1.31	1.12	1.4	1.13	1.09	1.31	1.21	1.13	1.49	2.3	0.47	0.48	1.19	1.12	1.46
	cD_T	0.87	0.84	0.83	0.72	0.66	0.71	0.61	0.21	0.19	0.8	0.74	0.78	0.84	0.83	0.9	0.65	0.69	0.81	0.81	0.16	0.16	0.75	0.75	0.86
$S8$	CD_T	1	1.03	1.09	0.94	1.12	1.3	0.66	1.33	1.94	1.02	1.08	1.17	0.98	1.01	1.07	0.89	1.06	1.29	0.79	1.31	2.03	1.02	1.01	1.13
	cD_T	0.76	0.8	0.8	0.54	0.61	0.67	0.18	0.26	0.59	0.62	0.69	0.71	0.74	0.78	0.79	0.51	0.59	0.62	0.21	0.1	0.56	0.6	0.65	0.66
$S9$	CD_T	1.08	1.13	1.07	1.15	1.31	1.15	0.84	2.4	0.47	1.15	1.23	1.09	1.09	1.15	1.08	1.17	1.38	1.19	0.72	2.99	0.53	1.17	1.29	1.13
	cD_T	0.86	0.88	0.9	0.72	0.77	0.72	0.28	0.25	0.18	0.78	0.82	0.8	0.87	0.89	0.9	0.74	0.82	0.75	0.29	0.29	0.22	0.8	0.83	0.82
$S10$	CD_T	1.03	1.19	1.02	1.21	1.37	1.11	3.82	1.46	0.77	1.09	1.32	1.09	1.06	1.17	1.08	1.28	1.38	1.2	3.77	1.61	0.86	1.14	1.31	1.17
	cD_T	0.89	0.97	0.86	0.83	0.9	0.77	0.53	0.74	0.46	0.84	0.93	0.79	0.89	1.00	0.9	0.8	0.99	0.84	0.5	0.79	0.48	0.81	1.02	0.85
$S11$	CD_T	1.08	1.04	1.03	1.22	1.11	1.19	3.54	1.07	1.43	1.13	1.03	1.08	1.1	1.08	1.07	1.25	1.2	1.27	3.38	1.21	1.46	1.16	1.1	1.15
	cD_T	0.92	0.89	0.88	0.83	0.76	0.82	1.19	0.46	0.6	0.84	0.79	0.81	0.92	0.92	0.9	0.82	0.8	0.86	0.93	0.43	0.61	0.84	0.84	0.85

Appendix B

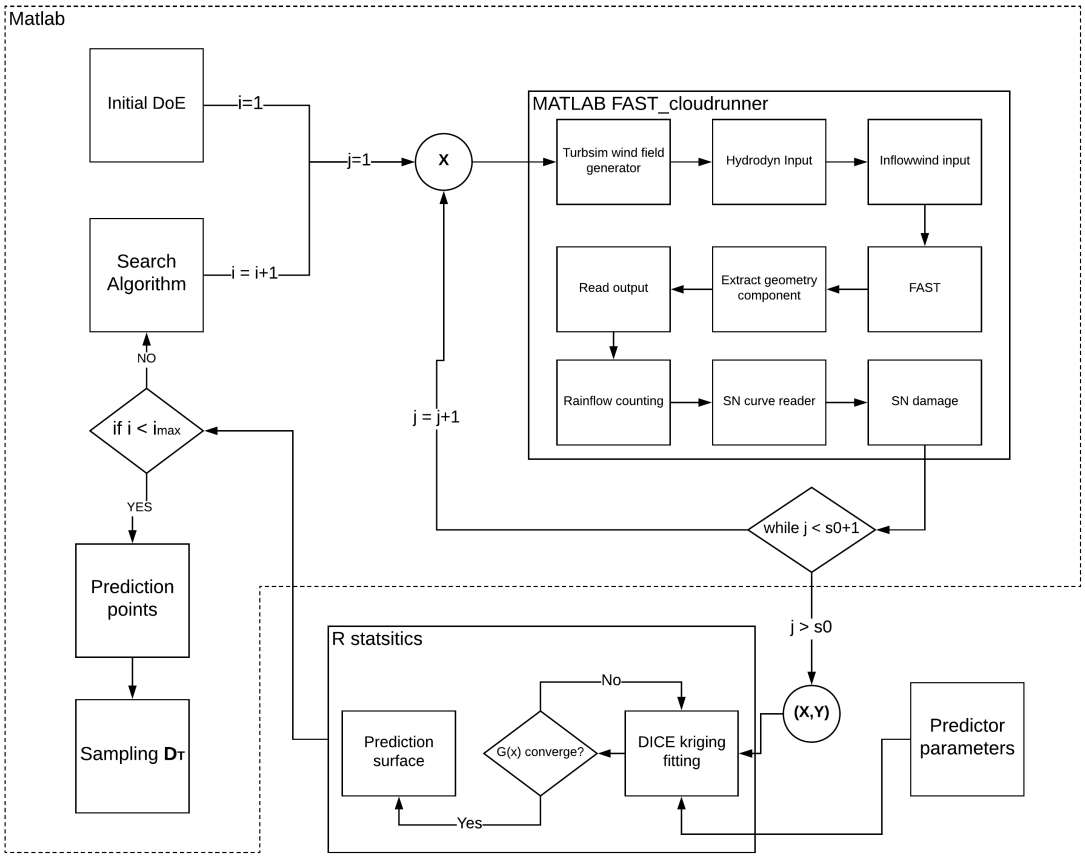


Figure B.1: Configuration of Kriging Gaussian process predictor code using an active learning search.

Bibliography

P. Agarwal. *Structural reliability of offshore wind turbines*. The University of Texas at Austin, 2008.

M. A. Aizerman. Theoretical foundations of the potential function method in pattern recognition learning. *Automation and remote control*, 25:821–837, 1964.

H. Arabian-Hoseynabadi, H. Oraee, and P. J. Tavner. Failure modes and effects analysis (FMEA) for wind turbines. *International Journal of Electrical Power & Energy Systems*, 32(7):817–824, 2010.

C. Bak, F. Zahle, R. Bitsche, T. Kim, A. Yde, L. C. Henriksen, M. H. Hansen, J. P. A. A. Blasques, M. Gaunaa, and A. Natarajan. The DTU 10-MW reference wind turbine. In *Danish Wind Power Research 2013*, 2013.

M. Balesdent, J. Morio, and J. Marzat. Kriging-based adaptive importance sampling algorithms for rare event estimation. *Structural Safety*, 44:1–10, 2013.

R. J. Barthelmie, K. Hansen, S. T. Frandsen, O. Rathmann, J. G. Schepers, W. Schlez, J. Phillips, K. Rados, A. Zervos, E. S. Politis, et al. Modelling and measuring flow and wind turbine wakes in large wind farms offshore. *Wind Energy: An International Journal for Progress and Applications in Wind Power Conversion Technology*, 12(5): 431–444, 2009.

J. A. Battjes. Long-term wave height distributions at seven stations around the british isles. *Deutsche Hydrografische Zeitschrift*, 25(4):179–189, 1972.

Y. Bazilevs, A. Korobenko, X. Deng, and J. Yan. Novel structural modeling and mesh moving techniques for advanced fluid–structure interaction simulation of wind turbines. *International Journal for Numerical Methods in Engineering*, 102(3-4):766–783, 2015.

P. Bernardara, F. Mazas, X. Kergadallan, and L. Hamm. A two-step framework for over-threshold modelling of environmental extremes. *Natural Hazards and Earth System Sciences*, 14(3):635–647, 2014.

B. J. Bichon, M. S. Eldred, L. P. Swiler, S. Mahadevan, and J. M. McFarland. Efficient global reliability analysis for nonlinear implicit performance functions. *AIAA journal*, 46(10):2459–2468, 2008.

M. Bigerelle, D. Najjar, B. Fournier, N. Rupin, and A. Iost. Application of lambda distributions and bootstrap analysis to the prediction of fatigue lifetime and confidence intervals. *International Journal of Fatigue*, 28(3):223 – 236, 2006.

E. Bommier. Peaks-Over-Threshold Modelling of Environmental Data. Technical report, U.U.D.M. Project Report 2014:33, 2014.

E. Borgonovo. A new uncertainty importance measure. *Reliability Engineering & System Safety*, 92(6):771–784, 2007.

T. Burton, N. Jenkins, D. Sharpe, and E. Bossanyi. *Wind energy handbook*. John Wiley & Sons, 2011.

BVG. Uk offshore wind supply chain: capabilities and opportunities. Technical report, BVG Associates, 2014.

B. Cañellas, A. Orfila, F. J. Méndez, M. Menéndez, L. Gómez-Pujol, and J. Tintoré. Application of a POT model to estimate the extreme significant wave height levels around the Balearic Sea (Western Mediterranean). *Journal of Coastal Research*, 50: 329–333, 2007.

Y. Caniou. *Global sensitivity analysis for nested and multiscale modelling*. PhD thesis, Université Blaise Pascal-Clermont-Ferrand II, 2012.

E. Carlstein. The use of subseries values for estimating the variance of a general statistic from a stationary sequence. *The Annals of Statistics*, pages 1171–1179, 1986.

W. Carswell, J. Johansson, F. Løvholt, S. R. Arwade, C. Madshus, D. J. DeGroot, and A. T. Myers. Foundation damping and the dynamics of offshore wind turbine monopiles. *Renewable energy*, 80:724–736, 2015.

E. Castillo and A. S. Hadi. Fitting the Generalized Pareto distribution to data. *Journal of the American Statistical Association*, 92(440):1609–1620, 1997.

E. Castillo, A. S. Hadi, N. Balakrishnan, and J.-M. Sarabia. *Extreme value and related models with applications in Engineering and Science*. Wiley Hoboken, NJ, 2005.

E. Castillo, A. O'Connor, M. Nogal, and A. Calvino. On the physical and probabilistic consistency of some engineering random models. *Structural Safety*, 51:1–12, 2014.

CEN. Eurocode 3: Design of steel structures - part 1-9: Fatigue. Technical Report EN1993, European Committee for Standardization, 2005.

CEN. Eurocode 9: Design of aluminium structures - part 1-3: Fatigue. Technical Report EN1993, European Committee for Standardization, 2011.

P. W. Cheng. *A reliability based design methodology for extreme responses of offshore wind turbines*. 2002.

P. W. Cheng, G. J. W. van Bussel, G. A. M. van Kuik, and J. H. Vugts. Reliability-based design methods to determine the extreme response distribution of offshore wind turbines. *Wind Energy*, 6(1):1–22, 2003.

V. Choulakian and M. A. Stephens. Goodness-of-fit tests for the Generalized Pareto distribution. *Technometrics*, 43(4):478–484, 2001.

M. K. Chryssanthopoulos and T. D. Righiniotis. Fatigue reliability of welded steel structures. *Journal of Constructional Steel Research*, 62(11):1199–1209, 2006.

G. Cortina, V. Sharma, and M. Calaf. Investigation of the incoming wind vector for improved wind turbine yaw-adjustment under different atmospheric and wind farm conditions. *Renewable Energy*, 101:376 – 386, 2017.

N. Cosack. *Fatigue load monitoring with standard wind turbine signals.* PhD thesis, University of Stuttgart, 2010.

D. D. Cox and S. John. Sdo: A statistical method for global optimization. *Multidisciplinary design optimization: state of the art*, pages 315–329, 1997.

W. Cui. A state-of-the-art review on fatigue life prediction methods for metal structures. *Journal of marine science and technology*, 7(1):43–56, 2002.

R. C. da Costa and L. V. S. Sagrilo. Statistical uncertainty analysis in time-domain fatigue assessment of steel risers. *Journal of Offshore Mechanics and Arctic Engineering*, 140 (3):031701, 2018.

N. Dalili, A. Edrisy, and R. Carriveau. A review of surface engineering issues critical to wind turbine performance. *Renewable and Sustainable Energy Reviews*, 13(2):428 – 438, 2009.

R. Damiani, J. Jonkman, and G. Hayman. Subdyn user's guide and theory manual. Technical report, National Renewable Energy Lab.(NREL), Golden, CO (United States), 2015.

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

C. Desmond, J. Murphy, L. Blonk, and W. Haans. Description of an 8 MW reference wind turbine. In *Journal of Physics: Conference Series*, volume 753, No. 9, page 092013. IOP Publishing, 2016.

T. J. DiCiccio and B. Efron. Bootstrap confidence intervals. *Statistical science*, pages 189–212, 1996.

N. Dimitrov, A. Natarajan, and J. Mann. Effects of normal and extreme turbulence spectral parameters on wind turbine loads. *Renewable Energy*, 101:1180–1193, 2017.

ENISO DIN. 19902: 2008-07: Petroleum and natural gas industries–fixed steel offshore structures (ISO 19902: 2007). *English version EN ISO*, 19902:182–202, 2007.

DNV. Fatigue design of offshore steel structures. Recommended Practice DNV-RP-C2013, DNV GL AS, 2011.

DNV. Design of offshore wind turbine structures. Offshore Standard DNV-OS-J101, DNV GL AS, 2014a.

DNV. Environmental condition and environmental loads. Recommended Practice DNV-RP-C205, DNV GL AS, 2014b.

S. D. Downing and D. F. Socie. Simple rainflow counting algorithms. *International journal of fatigue*, 4(1):31–40, 1982.

L. Draper. Derivation of adesign wavefrom instrumental records of sea waves. *Proceedings of the Institution of Civil Engineers*, 26(2):291–304, 1963.

V. Dubourg. *Adaptive surrogate models for reliability analysis and reliability-based design optimization.* PhD thesis, Université Blaise Pascal-Clermont-Ferrand II, 2011.

B. Echard, N. Gayton, and M. Lemaire. Ak-mcs: an active learning reliability method combining kriging and monte carlo simulation. *Structural Safety*, 33(2):145–154, 2011.

B. Echard, N. Gayton, M. Lemaire, and N. Relun. A combined importance sampling and kriging reliability method for small failure probabilities with time-demanding numerical models. *Reliability Engineering & System Safety*, 111:232–240, 2013.

B. Efron and C. Stein. The jackknife estimate of variance. *The Annals of Statistics*, pages 586–596, 1981.

B. Efron et al. Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, 7(1):1–26, 1979.

P. Embrechts, S. I. Resnick, and G. Samorodnitsky. Extreme value theory as a risk management tool. *North American Actuarial Journal*, 3(2):30–41, 1999.

A. Fatemi and L. Yang. Cumulative fatigue damage and life prediction theories: a survey of the state of the art for homogeneous materials. *International journal of fatigue*, 20 (1):9–34, 1998.

J. A. Ferreira and C. Guedes Soares. An application of the peaks over threshold method to predict extremes of significant wave height. *Journal of Offshore Mechanics and Arctic Engineering*, 120(3):165–176, 1998.

G. R. Fischer, T. Kipouros, and A. M. Savill. Multi-objective optimisation of horizontal axis wind turbine structure and energy production using aerofoil and blade properties as design variables. *Renewable Energy*, 62:506–515, 2014.

T. Fischer, W. E. De Vries, and B. Schmidt. *UpWind design basis (WP4: Offshore foundations and support structures)*. Upwind, 2010.

P. A. Fleming, A. Ning, P. M. O. Gebraad, and K. Dykes. Wind plant system engineering through optimization of layout and yaw control. *Wind Energy*, 19(2):329–344, 2016.

J. Fogle, P. Agarwal, and L. Manuel. Towards an improved understanding of statistical extrapolation for wind turbine extreme loads. *Wind Energy*, 11(6):613–635, 2008.

R. G. Forman. Study of fatigue crack initiation from flaws using fracture mechanics theory. *Engineering Fracture Mechanics*, 4(2):333–345, 1972.

A. I. J. Forrester, A. J. Keane, and N. W. Bressloff. Design and analysis of "noisy" computer experiments. *AIAA journal*, 44(10):2331, 2006.

G. Z. Forristall. On the statistical distribution of wave heights in a storm. *Journal of Geophysical Research: Oceans*, 83(C5):2353–2358, 1978.

G. Freebury and W. Musial. Determining equivalent damage loading for full-scale wind turbine blade fatigue tests. In *ASME Wind Energy Symposium*, page 50, 2000.

P. Garrett. Life cycle assessment of electricity production from a v90-2.0 mw gridstreamer wind plant. 2011.

B. Gaspar, A. P. Teixeira, and C. Guedes Soares. Assessment of the efficiency of kriging surrogate models for structural reliability analysis. *Probabilistic Engineering Mechanics*, 37:24–34, 2014.

Germanischer Lloyd WindEnergie GmbH. Certificate: Fast and adams incl. aerodyn module, 2005.

Y. Goda. On the methodology of selecting design wave height. In *Coastal Engineering 1988*, pages 899–913. 1989.

Y. Gōda. *Random seas and design of maritime structures*. World scientific, 2010.

Y. Goda, M. Kudaka, and H. Kawai. Incorporation of Weibull distribution in L-moments method for regional frequency of peak-over-threshold. *Coastal Engineering Proceedings*, 1(32), 2011.

G. Greegar and C. S. Manohar. Global response sensitivity analysis of uncertain structures. *Structural Safety*, 58:94–104, 2016.

J. A. Greenwood and M. M. Sandomire. Sample size required for estimating the standard deviation as a per cent of its true value. *Journal of the American Statistical Association*, 45(250):257–260, 1950.

J. A. Greenwood, J. M. Landwehr, N. C. Matalas, and J. R. Wallis. Probability weighted moments: definition and relation to parameters of several distributions expressable in inverse form. *Water Resources Research*, 15(5):1049–1054, 1979.

N. Gupta. Probabilistic load flow with detailed wind generator models considering correlated wind generation and correlated loads. *Renewable Energy*, 94:96–105, 2016.

T. R. Gurney. *Fatigue of welded structures*. CUP Archive, 1979.

Lorenz Haid, Gordon Stewart, Jason Jonkman, Amy Robertson, Matthew Lackner, and Denis Matha. Simulation-length requirements in the loads analysis of offshore floating wind turbines. In *ASME 2013 32nd International Conference on Ocean, Offshore and Arctic Engineering*, pages V008T09A091–V008T09A091. American Society of Mechanical Engineers, 2013.

P. J. Hawkes, D. Gonzalez-Marco, A. Sánchez-Arcilla, and P. Prinos. Best practice for the estimation of extremes: A review. *Journal of Hydraulic Research*, 46(S2):324–332, 2008.

G. J. Hayman and M. Buhl Jr. Mlife users guide for version 1.00. Technical report, National Renewable Energy Lab.(NREL), Golden, CO (United States), 2012.

A. Heege, J. Betran, and Y Radovcic. Fatigue load computation of wind turbine gearboxes by coupled finite element, multi-body system and aerodynamic analysis. *Wind Energy: An International Journal for Progress and Applications in Wind Power Conversion Technology*, 10(5):395–413, 2007.

J. C. Helton and F. J. Davis. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliability Engineering & System Safety*, 81(1):23–69, 2003.

J. C. Helton, F. J. Davis, and J. D. Johnson. A comparison of uncertainty and sensitivity analysis results obtained with random and latin hypercube sampling. *Reliability Engineering & System Safety*, 89(3):305–330, 2005.

L. H. Holthuijsen. *Waves in oceanic and coastal waters*. Cambridge university press, 2010.

M. C. Holtslag, W. A. A. M. Bierbooms, and G. J. W. van Bussel. Wind turbine fatigue loads as a function of atmospheric conditions offshore. *Wind Energy*, 19(10):1917–1932, 10 2016.

M. C. Holtslag, W. Bierbooms, and G. van Bussel. Extending the diabatic surface layer wind shear profile for offshore wind energy. *Renewable Energy*, 101:96–110, 2017.

S. M. Hoseyni, F. Di Maio, M. Vagnoli, E. Zio, and M. Pourgol-Mohammad. A bayesian ensemble of sensitivity measures for severe accident modeling. *Nuclear Engineering and Design*, 295:182–191, 2015.

HSE. Environmental considerations. Technical report, Health & Safety Executive, 2002.

M. C. Hsu, I. Akkerman, and Y. Bazilevs. Finite element simulation of wind turbine aerodynamics: validation study using NREL phase VI experiment. *Wind Energy*, 17 (3):461–481, 2014.

D. Huang, T. T. Allen, W. I. Notz, and N. Zeng. Global optimization of stochastic black-box systems via sequential kriging meta-models. *Journal of global optimization*, 34(3): 441–466, 2006.

IEC. Wind turbines– Part 1: Design requirements. Technical Report 61400-1, International Electrotechnical Commission, Geneva, Switzerland, 2005.

IEC. Wind turbines– Part 3: Design requirements for offshore wind turbines. Technical Report 61400-3, International Electrotechnical Commission, Geneva, Switzerland, 2009.

R. L. Iman and W.-J. Conover. A distribution-free approach to inducing rank correlation among input variables. *Communications in Statistics-Simulation and Computation*, 11 (3):311–334, 1982.

R. L. Iman and J. C. Helton. Comparison of uncertainty and sensitivity analysis techniques for computer models. Technical report, Sandia National Labs., Albuquerque, NM (USA), 1985.

IRENA. Renewable energy technologies: Cost analysis series. 2012.

IRENA. The power to change: Solar and wind cost reduction potential to 2025. Technical report, International Renewable Energy Agency, 2016.

IRENA. Renewable power generation costs in 2017. Technical report, International Renewable Energy Agency, 2017.

ISO. General principles on reliability for structures. Technical Report 2394, International Organization for Standardization, 2015.

P. Iván, R. Tom, and M. Ariola. Offshore wind in europe: Key trends and statsitics 2017. Technical report, WindEurope, 2017.

H. O. Jahns, J. D. Wheeler, et al. Long-term wave probabilities based on hindcasting of severe storms. In *Offshore Technology Conference*. Offshore Technology Conference, 1972.

Y. M. Ji and K. S. Han. Fracture mechanics approach for failure of adhesive joints in wind turbine blades. *Renewable energy*, 65:23–28, 2014.

Z. Jiang, W. Hu, W. Dong, Z. Gao, and Z. Ren. Structural reliability analysis of wind turbines: A review. *Energies*, 10(12):2099, 2017.

D. R. Jones. A taxonomy of global optimization methods based on response surfaces. *Journal of global optimization*, 21(4):345–383, 2001.

D. R. Jones, M. Schonlau, and W. J. Welch. Efficient global optimization of expensive black-box functions. *Journal of Global optimization*, 13(4):455–492, 1998.

J. Jonkman. The new modularization framework for the fast wind turbine cae tool. In *51st AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition*, page 202, 2013.

J. Jonkman and M. L. Buhl Jr. Fast user's guide-updated august 2005. Technical report, National Renewable Energy Laboratory (NREL), Golden, CO., 2005.

J. Jonkman and W. Musial. Offshore code comparison collaboration (OC3) for IEA task 23 offshore wind technology and deployment. *Contract*, 303:275–3000, 2010.

J. Jonkman, S. Butterfield, W. Musial, and G. Scott. Definition of a 5-MW reference wind turbine for offshore system development. *National Renewable Energy Laboratory, Golden, CO, Technical Report No. NREL/TP-500-38060*, 2009.

J. Jonkman, A. Robertson, and G. Hayman. Hydrodyn user's guide and theory manual. Technical report, National Renewable Energy Lab., Golden, CO (US), 2014.

J. Jonkman, G. Hayman, B. Jonkman, and R. Damiani. Aerodyn v15 user's guide and theory manual. Technical report, National Renewable Energy Laboratory (NREL), Golden, CO., 2016.

J. B. Jonkman and L. Kilcher. Turbsim user's guide: Version 1.06. 00, national renewable energy laboratory. Technical report, Technical Report. Draft available at: https://wind. nrel. gov/designcodes/preprocessors/turbsim/TurbSim. pdf, 2012.

J. C. Kaimal, J. C. J. Wyngaard, Y. Izumi, and O. R. Coté. Spectral characteristics of surface-layer turbulence. *Quarterly Journal of the Royal Meteorological Society*, 98 (417):563–589, 1972.

M. Karimirad. Modeling aspects of a floating wind turbine for coupled wave–wind-induced dynamic analyses. *Renewable Energy*, 53:299–305, 2013.

T. Kashef and S. Winterstein. Relating turbulence to wind turbine blade loads: Parametric study with multiple regression analysis. *Journal of solar energy engineering*, 121(3): 156–161, 1999.

K. M. Kecskemety and J. J. McNamara. Influence of wake dynamics on the performance and aeroelasticity of wind turbines. *Renewable Energy*, 88:333–345, 2016.

N. Kelley and H. J. Sutherland. Damage estimates from long-term structural analysis of a wind turbine in a US wind farm environment. In *35th Aerospace Sciences Meeting and Exhibit*, page 950, 1997.

D. G. Krige. A statistical approach to some basic mine valuation problems on the witwatersrand. *Journal of the Southern African Institute of Mining and Metallurgy*, 52(6): 119–139, 1951.

M. J. Kühn. *Dynamics and design optimisation of offshore wind energy conversion systems*. PhD thesis, DUWIND, Delft University Wind Energy Research Institute, 2001.

S. Kullback. *Information theory and statistics*. Courier Corporation, 1997.

S. Kullback and R. A. Leibler. On information and sufficiency. *The annals of mathematical statistics*, 22(1):79–86, 1951.

H. R. Kunsch. The jackknife and the bootstrap for general stationary observations. *The annals of Statistics*, pages 1217–1241, 1989.

H. J. Kushner. A new method of locating the maximum point of an arbitrary multipeak curve in the presence of noise. *Journal of Basic Engineering*, 86(1):97–106, 1964.

C. Lalanne. *Mechanical vibration and shock analysis, fatigue damage*, volume 4. John Wiley & Sons, 2010.

M. Lang, T. B. M. J. Ouarda, and B. Bobée. Towards operational guidelines for over-threshold modeling. *Journal of hydrology*, 225(3):103–117, 1999.

C. H. Lange. Probabilistic fatigue methodology and wind turbine reliability. Technical report, Sandia National Labs., Albuquerque, NM (United States), 1996.

H. W. Lilliefors. On the Kolmogorov-Smirnov test for the exponential distribution with mean unknown. *Journal of the American Statistical Association*, 64(325):387–389, 1969.

H. Liu, W. Chen, and A. Sudjianto. Relative entropy based method for probabilistic sensitivity analysis in engineering design. *Journal of Mechanical Design*, 128(2):326–336, 2006.

J. Liu, Z. H. Han, and W. P. Song. Comparison of infill sampling criteria in kriging-based aerodynamic optimization. In *28th Congress of the International Council of the Aeronautical Sciences*, pages 23–28, 2012.

R. A. Lockhart and M. A. Stephens. Estimation and tests-of-fit for the three parameter Weibull Distribution. Technical report, Office of Naval Research. Technical Report No. 476, 1993.

K. Luo, S. Zhang, Z. Gao, J. Wang, L. Zhang, R. Yuan, J. Fan, and K. Cen. Large-eddy simulation and wind-tunnel measurement of aerodynamics and aeroacoustics of a horizontal-axis wind turbine. *Renewable Energy*, 77:351–362, 2015.

K. Maki, R. Sbragio, and N. Vlahopoulos. System design of a wind turbine using a multi-level optimization approach. *Renewable Energy*, 43:101–110, 2012.

G. Manache and C. S. Melching. Identification of reliable regression-and correlation-based sensitivity measures for importance ranking of water-quality model parameters. *Environmental Modelling & Software*, 23(5):549–562, 2008.

J. F. Mandell and D. D. Samborsky. Doe/msu composite material fatigue database: test methods, materials, and analysis. Technical report, Sandia National Labs., Albuquerque, NM (United States), 1997.

L. Manuel, P. S. Veers, and S. R. Winterstein. Parametric models for estimating wind turbine fatigue loads for design. *Journal of solar energy engineering*, 123(4):346–355, 2001.

S. Marelli and B. Sudret. Uqlab: A framework for uncertainty in matlab. In *Proceedings 2nd Int. Conf. on Vulnerability Risk Analysis and Management (ICVRAM)*, 2014.

J. C. Marin, A. Barroso, F. Paris, and J. Canas. Study of fatigue damage in wind turbine blades. *Engineering failure analysis*, 16(2):656–668, 2009.

E. Marino, A. Giusti, and L. Manuel. Offshore wind turbine fatigue loads: The influence of alternative wave modeling for different turbulent and mean winds. *Renewable Energy*, 102:157–169, 2017.

S. Márquez-Domínguez and J. D. Sørensen. Fatigue reliability and calibration of fatigue design factors for offshore wind turbines. *Energies*, 5(6):1816–1834, 2012.

L. A. Martínez-Tossas, M. J. Churchfield, and S. Leonardi. Large eddy simulations of the flow past wind turbines: actuator line and disk modeling. *Wind Energy*, 18(6): 1047–1060, 2015.

M Mathiesen, and Hawkes P J Goda, Y, E. Mansard, M. J. Martín, E. Peltier, E. F. Thompson, and G. Van Vledder. Recommended practice for extreme wave analysis. *Journal of hydraulic Research*, 32(6):803–814, 1994.

A. J. McEvily and J. Groeger. On the threshold for fatigue crack growth. In *ICF4, Waterloo (Canada) 1977*, 1977.

F. J. Méndez, M. Menéndez, A. Luceño, and I. J. Losada. Estimation of the long-term variability of extreme significant wave height using a time-dependent peak over threshold (POT) model. *Journal of Geophysical Research: Oceans*, 111(C7), 2006.

M. Menéndez, F. J. Méndez, C. Izaguirre, A. Luceño, and I. J. Losada. The influence of seasonality on estimating return values of significant wave height. *Coastal Engineering*, 56(3):211–219, 2009.

G. Mesmacque, S. Garcia, A. Amrouche, and C. Rubio-Gonzalez. Sequential law in multiaxial fatigue, a new damage indicator. *International Journal of Fatigue*, 27(4):461–467, 2005.

M. A. Miner et al. Cumulative fatigue damage. *Journal of applied mechanics*, 12(3): A159–A164, 1945.

A. Morató, S. Sriramula, and N. Krishnan. Reliability analysis of offshore wind turbine support structures using kriging models. In *Proc. ESREL Conference, Glasgow, September*, 2016.

A. Morató, S. Sriramula, N. Krishnan, and J. Nichols. Ultimate loads and response analysis of a monopile supported offshore wind turbine using fully coupled simulation. *Renewable Energy*, 101:126–143, 2017.

P. J. Moriarty. Database for validation of design load extrapolation techniques. *Wind Energy*, 11(6):559–576, 2008.

P. J. Moriarty and A. C. Hansen. Aerodyn theory manual. Technical report, National Renewable Energy Lab., Golden, CO (US), 2005.

P. J. Moriarty, W. Holley, and S. P. Butterfield. Extrapolation of extreme and fatigue loads using probabilistic methods. Technical report, National Renewable Energy Lab., Golden, CO (US), 2004.

M. D. Morris and T. J. Mitchell. Exploratory designs for computational experiments. *Journal of statistical planning and inference*, 43(3):381–402, 1995.

L. R. Muir and A. H. El-Shaarawi. On the calculation of extreme wave heights: a review. *Ocean Engineering*, 13(1):93–118, 1986.

J. P. Murcia, P. E. Réthoré, N. Dimitrov, A. Natarajan, J. D. Sørensen, P. Graf, and T. Kim. Uncertainty propagation through an aeroelastic wind turbine model using polynomial surrogates. *Renewable Energy*, 119:910–922, 2018.

A. Natarajan, W. E. Holley, R. Penmatsa, and Brahmanapalli B. C. Identification of contemporaneous component loading for extrapolated primary loads in wind turbines. *Wind Energy*, 11(6):577–587, 2008.

A. R. Nejad, Z. Gao, and T. Moan. On long-term fatigue damage and reliability analysis of gears under wind loads in offshore wind turbine drivetrains. *International Journal of Fatigue*, 61:116–128, 2014.

D. Nezamolmolki and A. Shooshtari. Investigation of nonlinear dynamic behavior of lattice structure wind turbines. *Renewable Energy*, 97:33–46, 2016.

M. Nogal, E. Castillo, A. Calvino, and A. O'Connor. Coherent and compatible statistical models in structural analysis. *International Journal of Computational Methods*, 13(02): 1640008, 2016.

N. Nordenstrøm. Long-term distribution of wave heights and periods. Technical Report 69.21 S, Det Norske Veritas, 1969.

A. Olsson, G. Sandberg, and O. Dahlblom. On latin hypercube sampling for structural reliability analysis. *Structural safety*, 25(1):47–68, 2003.

U. Olsson. Confidence intervals for the mean of a log-normal distribution. *Journal of Statistics Education*, 13(1), 2005.

M. I. Ortego, R. Tolosana-Delgado, J. Gibergans-Báguena, J. J. Egozcue, and A. Sánchez-Arcilla. Assessing wavestorm hazard evolution in the nw mediterranean with hindcast and buoy data. *Climatic Change*, 113(3):713–731, 2012.

A. Palmgren. The life span of ball bearings. *Zeitschrift der Vereines Detscher Ingenieure*, 68, n. 14:339–341, 1924.

P. Paris and F. Erdogan. A critical analysis of crack propagation laws. *Journal of basic engineering*, 85(4):528–533, 1963.

C. K. Park and K.-Il Ahn. A new approach for measuring uncertainty importance and distributional sensitivity in probabilistic safety assessment. *Reliability Engineering & System Safety*, 46(3):253–261, 1994.

P. Passon, M. Kühn, S. Butterfield, J. Jonkman, T. Camp, and T. J. Larsen. OC3 benchmark exercise of aero-elastic offshore wind turbine codes. In *Journal of Physics: Conference Series*, volume 75, No. 1, page 012071. IOP Publishing, 2007.

V. Petrovic and C. L. Bottasso. Wind turbine envelope protection control over the full wind speed range. *Renewable Energy*, 111:836 – 848, 2017.

C. Petruaskas, P. M. Aagaard, et al. Etrapolation of historical storm data for estimating design wave heights. In *Offshore Technology Conference*. Offshore Technology Conference, 1970.

V. Picheny, D. Ginsbourger, and Y. Richet. Noisy expected improvement and on-line computation time allocation for the optimization of simulators with tunable fidelity. 2010.

V. Picheny, T. Wagner, and D. Ginsbourger. A benchmark of kriging-based infill criteria for noisy optimization. *Structural and Multidisciplinary Optimization*, 48(3):607–626, 2013.

J. Pickands III. Statistical inference using extreme order statistics. *The Annals of Statistics*, pages 119–131, 1975.

W. Popko, F. Vorpahl, A. Zuga, M. Kohlmeier, J. Jonkman, A. Robertson, T. J. Larsen, A. Yde, K. Saetertro, K. M. Okstad, J. Nichols, et al. Offshore code comparison collaboration continuation (OC4), phase i-results of coupled simulations of an offshore wind turbine with jacket support structure: Preprint. Technical report, National Renewable Energy Lab.(NREL), Golden, CO (United States), 2012.

P. Ragan and L. Manuel. Comparing estimates of wind turbine fatigue loads using time-domain and spectral methods. *Wind engineering*, 31(2):83–99, 2007.

C. E. Rasmussen. Gaussian processes in machine learning. In *Advanced lectures on machine learning*, pages 63–71. Springer, 2004.

M. Ribatet. A user's guide to the pot package (version 1.4). *September*, 2006.

A. L. Richard, F. O'Reilly, and M. A. Stephens. Tests for the extreme value and Weibull distributiuons based on normalized spacings. Technical report, Office of Naval Research, 1985.

A. N. Robertson, F. Wendt, J. Jonkman, W. Popko, H. Dagher, S. Gueydon, J. Qvist, F. Vittori, J. Azcona, E. Uzunoglu, et al. OC5 project phase ii: validation of global loads of the deepcwind floating semisubmersible wind turbine. *Energy Procedia*, 137: 38–57, 2017.

K. O. Ronold and C. J. Christensen. Optimization of a design code for wind-turbine rotor blades in fatigue. *Engineering Structures*, 23(8):993–1004, 2001.

K. O. Ronold, J. Wedel-Heinen, and C. J. Christensen. Reliability-based fatigue design of wind-turbine rotor blades. *Engineering structures*, 21(12):1101–1114, 1999.

O. Roustant, D. Ginsbourger, and Y. Deville. Dicekriging, diceoptim: Two r packages for the analysis of computer experiments by kriging-based metamodeling and optimization. *Journal of Statistical Software*, 2012.

T. Sakai. Review and prospects for current studies on very high cycle fatigue of metallic materials for machine structural use. *Journal of solid mechanics and materials engineering*, 3(3):425–439, 2009.

A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola. *Global sensitivity analysis: the primer*. John Wiley & Sons, 2008.

A. Sánchez-Arcilla, J. Gomez Aguar, J. J. Egozcue, M. I. Ortego, P. Galiatsatou, and P. Prinos. Extremes from scarce data: The role of bayesian and scaling techniques in reducing uncertainty. *Journal of Hydraulic Research*, 46(S2):224–234, 2008.

C. Scarrott and A. MacDonald. A review of extreme value threshold estimation and uncertainty quantification. *REVSTAT–Statistical Journal*, 10(1):33–60, 2012.

W. Scott, P. Frazier, and W. Powell. The correlated knowledge gradient for simulation optimization of continuous parameters using gaussian process regression. *SIAM Journal on Optimization*, 21(3):996–1026, 2011.

N. Sedaghatizadeh, M. Arjomandi, R. Kelso, B. Cazzolato, and M. H. Ghayesh. Modelling of wind turbine wake using large eddy simulation. *Renewable Energy*, 115:1166–1176, 2018.

D. J. Sheskin. *Handbook of Parametric and nonparametric statistical procedures*. CRC Press LLC, 2003.

M. Shokrieh and R. Rafiee. Simulation of fatigue failure in a full composite wind turbine blade. *Composite structures*, 74(3):332–342, 2006.

J. O. Smith. The effect of range of stress on the fatigue strength of metals. Technical report, University of Illinois at Urbana Champaign, College of Engineering. Engineering Experiment Station., 1942.

A. J. Smola and B. Schölkopf. A tutorial on support vector regression. *Statistics and computing*, 14(3):199–222, 2004.

I. M. Sobol. Sensitivity estimates for nonlinear mathematical models. *Mathematical Modelling and Computational Experiments*, 1(4):407–414, 1993.

J. D. Sørensen, S. Frandsen, and N. J. Tarp-Johansen. Effective turbulence models and fatigue reliability in wind farms. *Probabilistic Engineering Mechanics*, 23(4):531–538, 2008.

J. D. Sørensen et al. Reliability-based calibration of fatigue safety factors for offshore wind turbines. In *The Twenty-first International Offshore and Polar Engineering Conference*. International Society of Offshore and Polar Engineers, 2011.

J. J. Spineili and M. A. Stephens. Tests for exponentiality when origin and scale parameters are unknown. *Technometrics*, 29(4):471–476, 1987.

S. Stanzl-Tschegg. Very high cycle fatigue measuring techniques. *International Journal of Fatigue*, 60:2–17, 2014.

M. Stein. Large sample properties of simulations using latin hypercube sampling. *Technometrics*, 29(2):143–151, 1987.

M. A. Stephens. Use of the Kolmogorov-Smirnov, Cramér-von Mises and related statistics without extensive tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 115–122, 1970.

M. A. Stephens. Asymptotic results for goodness-of-fit statistics with unknown parameters. *The Annals of Statistics*, pages 357–369, 1976.

M. A. Stephens. Goodness of fit for the extreme value distribution. *Biometrika*, 64(3): 583–588, 1977.

M. A. Stephens. The Anderson-Darling Statistic. Technical report, U.S. Army Research Office, 1979.

S. A. Stephens and R. M. Gorman. Extreme wave predictions around new zealand from hindcast data. *New Zealand Journal of Marine and Freshwater Research*, 40(3):399–411, 2006.

R. J. A. M. Stevens. Dependence of optimal wind turbine spacing on wind farm length. *Wind energy*, 19(4):651–663, 2016.

R. J. A. M. Stevens, L. A. Martínez-Tossas, and C. Meneveau. Comparison of wind farm large eddy simulations using actuator disk and actuator line models with wind tunnel experiments. *Renewable Energy*, 116:470–478, 2018.

H. J. Sutherland. Fatigue case study and loading spectra for wind turbines. Technical report, Sandia National Labs., Albuquerque, NM (United States), 1994.

H. J. Sutherland. Frequency domain analysis of the fatigue loads on typical wind turbine blades. *Journal of solar energy engineering*, 118(4):204–211, 1996.

H. J. Sutherland. On the fatigue analysis of wind turbines. Technical report, Sandia National Labs., Albuquerque, NM (US); Sandia National Labs., Livermore, CA (US), 1999.

H. J. Sutherland. A summary of the fatigue properties of wind turbine materials. *Wind Energy*, 3(1):1–34, 1 2000a.

H. J. Sutherland. A summary of the fatigue properties of wind turbine materials. *Wind Energy: An International Journal for Progress and Applications in Wind Power Conversion Technology*, 3(1):1–34, 2000b.

H. J. Sutherland and C. Butterfield. A review of the workshop on fatigue life methodologies for wind turbines. In *Presented at the 13th Energy-Sources Technology Conference and Exhibition (ETCE) on Wind Energy, New Orleans, LA, 23-26 Jan. 1994*, 1993.

H. J. Sutherland and J. F. Mandell. The effect of mean stress on damage predictions for spectral loading of fibreglass composite coupons. *Wind Energy: An International Journal for Progress and Applications in Wind Power Conversion Technology*, 8(1): 93–108, 2005.

H. J. Sutherland and P. S. Veers. Effects of cyclic stress distribution models on fatigue life predictions. *Wind Energy*, 16:83–90, 1995.

N. J. Tarp-Johansen. Examples of fatigue lifetime and reliability evaluation of larger wind turbine components. *Report No.: Risø*, 2003.

N. V. Teena, V. S. Kumar, K. Sudheesh, and R. Sajeev. Statistical analysis on extreme wave height. *Natural hazards*, 64(1):223–236, 2012.

R. Teixeira. Application of gaussian process regression for structural analysis. In *Training in Reducing Uncertainty in Structural Safety (TRUSS) Workshop*, page 24. TRUSS ITN, 2018.

R. Teixeira, A. O'Connor, and M. Nogal. Comparative analysis of the probabilistic methods to estimate the probability of failure of offshore wind turbine towers. In *Proceedings of the Civil Engineering Research in Ireland Conference (CERI 2016)*.

R. Teixeira, A. O'Connor, M. Nogal, N. Krishnan, and J. Nichols. Analysis of the design of experiments of offshore wind turbine fatigue reliability design with kriging surfaces. *Procedia Structural Integrity*, 5:951–958, 2017a.

R. Teixeira, A. O'Connor, M. Nogal, J. Nichols, and M. Spring. Structural probabilistic assessment of offshore wind turbine operation fatigue based on kriging interpolation. In *Proc. ESREL Conference, Portoroz*, 2017b.

R. Teixeira, M. Nogal, and A. O'Connor. On the suitability of the generalized pareto to model extreme waves. *Journal of Hydraulic Research*, 56(6):755–770, 2018a.

R. Teixeira, A. O'Connor, and M. Nogal. Interpolation of confidence intervals for fatigue design using a surrogate model. In *Internation forum on Engineering Decision Making*, 2018b.

R. Teixeira, M. Nogal, and A. O'Connor. On the calculation of offshore wind turbine load spectra for fatigue design. In *Mechanical Fatigue of Metals*, pages 383–390. Springer, 2019a.

R. Teixeira, M. Nogal, and A. O'Connor. Analysis of long-term loading characterization for stress-cycle fatigue design. *Wind Energy*, pages 1–18, 2019b.

R. Teixeira, M. Nogal, A. O'Connor, J. Nichols, and A. Dumas. Stress-cycle fatigue design with kriging applied to offshore wind turbines. *International Journal of Fatigue*, 125: 454–467, 2019c.

R. Teixeira, A. O'Connor, and M. Nogal. Fatigue reliability using a multiple surface approach. In *International Conference on Applications of Statistics and Probability in Civil Engineering (ICASP13)*, 2019d.

R. Teixeira, A. O'Connor, and M. Nogal. Probabilistic sensitivity analysis of offshore wind turbines using a transformed kullback-leibler divergence. *Structural Safety*, 81:101860, 2019e.

W. Tian, Z. Mao, B. Zhang, and Y. Li. Shape optimization of a savonius wind rotor with different convex and concave sides. *Renewable Energy*, 117:287–299, 2018.

C. Tibaldi, C. Bak, and L. C. Henriksen. *Concurrent Aeroservoelastic Design and Optimization of Wind Turbines*. PhD thesis, DTU Wind Energy, 2015.

C. Tibaldi, L. C. Henriksen, M. H. Hansen, and C. Bak. Wind turbine fatigue damage evaluation based on a linear model and a spectral method. *Wind Energy*, 19(7):1289–1306, 2016.

H. S. Toft, L. Svenningsen, J. D. Sørensen, W. Moser, and M. L. Thøgersen. Uncertainty in wind climate parameters and their influence on wind turbine fatigue loads. *Renewable Energy*, 90:352–361, 2016.

I. Troen and E. L. Petersen. European wind atlas. risø national laboratory, roskilde. *Weibull W.(1951). A statistical distribution function of wide applicability. J. Appl. mech*, 18:293–297, 1989.

M. Türk and S. Emeis. The dependence of offshore turbulence intensity on wind speed. *Journal of Wind Engineering and Industrial Aerodynamics*, 98(8):466–471, 2010.

T. van der Zee, M. J. de Ruiter, and I. Wieling. The c-tower project–a composite tower for offshore wind turbines. *Energy Procedia*, 137:401–405, 2017.

P. H. Van Gelder. Statistical methods for the risk-based design of civil structures. *Communications on Hydraulic and Geotechnical Engineering Report (Netherlands)*, 2000.

P. H. Van Gelder and C. V. Mai. Distribution functions of extreme sea waves and river discharges. *Journal of Hydraulic Research*, 46(S2):280–291, 2008.

H. J. Van Grol and B. H. Bulder. *Procedures to determine the fatigue life for large size wind turbines*. Netherlands Energy Research Foundation ECN, 1993.

E. Vanem, A. B. Huseby, and B. Natvig. Modelling ocean wave climate with a bayesian hierarchical space–time model and a log-transform of the data. *Ocean dynamics*, 62(3): 355–375, 2012.

P. S. Veers and S. Winterstein. Application of measured loads to wind turbine fatigue and reliability analysis. *Journal of Solar Energy Engineering*, 120(4):233–239, 1998.

D. Veldkamp. A probabilistic evaluation of wind turbine fatigue design rules. *Wind Energy*, 11(6):655, 2008.

H. F. Veldkamp. *Chances in wind energy: a probalistic approach to wind turbine fatigue design*. PhD thesis, TU Delft, Delft University of Technology, 2006.

F. A. C. Viana, G. Venter, and V. Balabanov. An algorithm for fast optimal latin hypercube design of experiments. *International journal for numerical methods in engineering*, 82(2):135–156, 2010.

J. K. Vrijling and P. H. A. J. M. Van Gelder. Uncertainty analysis of water levels on lake ijssel in the netherlands: A decision-making analysis. In *Proceedings of the 9th Annual Conference Risk Analysis: Facing the New Millennium, Rotterdam, The Netherlands*, 1999.

L. Wang, X. Liu, and A. Kolios. State of the art in the aeroelasticity of wind turbine blades: Aeroelastic modelling. *Renewable and Sustainable Energy Reviews*, 64:195–210, 2016.

A. G. Watson and R. J. Barnes. Infill sampling criteria to locate extremes. *Mathematical Geology*, 27(5):589–608, 1995.

W. Weibull. A statistical distribution function of wide applicability. *Journal of applied mechanics*, 103:293–297, 1951.

S. R. Winterstein, L. M. Fitzwater, L. Manuel, and P. S. Veers. Moment-based fatigue load models for wind energy systems. *Sandia National Laboratories US: Swets & Zeitlinger*, 2001.

P. H. Wirsching and Y.-N. Chen. Considerations of probability-based fatigue design for marine structures. *Marine Structures*, 1(1):23–45, 1988.

W. F. Wu, H. Y. Liou, and H. C. Tse. Estimation of fatigue damage and fatigue life of components under random loading. *International journal of pressure vessels and piping*, 72(3):243–249, 1997.

H. Yang and A. Wang. Fatigue reliability based design optimization of bending stiffener. *Journal of ship Research*, 56(2):120–128, 2012.

H. Yang, Y. Zhu, Q. Lu, and J. Zhang. Dynamic reliability based design optimization of the tripod sub-structure of offshore wind turbines. *Renewable Energy*, 78:16–25, 2015.

I. R. Young, J. Vinoth, S. Zieger, and A. V. Babanin. Investigation of trends in extreme value wave height and wind speed. *Journal of Geophysical Research: Oceans*, 117(C11), 2012.

S. Zergane, A. Smaili, and C. Masson. Optimization of wind turbine placement in a wind farm using a new pseudo-random number generation method. *Renewable Energy*, 125: 166 – 171, 2018.

L. Zhang, Z. Lu, and P. Wang. Efficient structural reliability analysis method based on advanced kriging model. *Applied Mathematical Modelling*, 39(2):781–793, 2015.

D. Zwick and M. Muskulus. Simplified fatigue load assessment in offshore wind turbine structural analysis. *Wind Energy*, 19(2):265–278, 2016.