Wind Turbine Load Extrapolation

Uncertainty quantification of 50-year ultimate wind turbine load estimations

S.F. van Eijk

Aerospace Faculty, Section Wind Energy

TU Delft Delft University of Technology

Challenge the future

WIND TURBINE LOAD EXTRAPOLATION

UNCERTAINTY QUANTIFICATION OF 50-YEAR ULTIMATE WIND TURBINE LOAD ESTIMATIONS

by

S.F. van Eijk

For obtaining the degree of

Master of Science
in Sustainable Energy Technology at the Delft University of Technology,

at Delft University of Technology,
Saturday April 16, 2016

Wind Energy Research Group, Faculty of Aerospace Engineering, Delft University of Technology

Student number:	4242319	
Project duration:	October 1, 2014 – April 16, 2016	
Supervisor:	Ir. R. Bos	
Thesis committee:	Prof. dr. ir. G. J. van Bussel,	TU Delft
	Prof. dr. ir. P. van Gelder,	TU Delft
	Dr. ir. W. A. A. M. Bierbooms,	TU Delft

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

TU Delft Delft University of Technology

SUMMARY

Wind turbine design is subject to load calculation standards. These determine material use which in turn affects wind turbine costs. Wind turbine cost reductions are important as they contribute to the economic feasibility of wind turbines. Therefore careful evaluation of the load standards is paramount for wind turbine design. This thesis offers an evaluation of 50-year ultimate load estimates described in the IEC 61400-1 standard and a tool for wind turbine designers to manage the estimates.

The release of 96 years of on-shore wind turbine load simulations (Barone *et al.*, 2012a) enabled uncertainty quantification of the 50-year ultimate wind turbine load estimations. The method for calculating 50-year load estimates was evaluated in this thesis by means of discrete samples from a Gumbel distribution and with SANDIA wind turbine load simulation samples. Discrete Gumbel samples were used in order to understand the sheer calculation method without interference of wind turbine effects. Wind turbine load simulation samples were used in order to quantify the uncertainty in calculation of eleven types of 50-year wind turbine load and deflection estimations. In addition, a suggestion for improving the 50-year load estimate calculation method was made. Lastly, probabilities for wrong component and wind turbine concept selection were worked out.

Important findings from applying the 50-year load calculation method to discrete Gumbel samples is an initial upward bias at low sample sizes that can reach up to 30%, which was first noticed by White (1969). Furthermore the variation in eleven different plotting positions irrespective of extrapolation was researched. All plotting positions resided within of 7% of each other.

Uncertainty quantification of SANDIAs OoPDefl, IPDefl, RootMOoP, RootMFlp, RootMEdg, TwrBsMyt and TwrBsMzt 10-min load extreme samples showed a distinct *knee* shaped CDF curve. This indicates that multiple processes exist in wind turbine load simulation. The process of maximum thrust that is generated at rated wind speed is assumed to be one of these processes. The *knee* causes large amounts of uncertainty in calculating the 50-year load estimate. If a minimum of 300 minutes of load simulation is done as recommended by IEC 61400-1, 50-year load estimates of two times the mean and 40% of variation are observed. Appropriate measures to mitigate the upward bias and variation result in a 55% upward bias for the worst load types: IPDefl, LSSTipMya, LSSTipMza and TwrBsMzt. By raising the cut-in wind speed from 3 to 5,6 - 8,9 m/s depending on the load type, the variation was slightly reduced and the amount of simulations needed dropped between 16 and 45%. From load distribution analysis it was found that processes causing the *knee* where removed by raising the cut-in wind speed.

The limited computational resources of wind turbine designers forces them to make decisions about wind turbine simulations. The distribution of 50-year load estimates behaves as a GEV. This makes wind turbine concept and component selection susceptible for false negatives. Two days of simulation for a component that is twice as strong as the TwrBsMyt load, will produce a 12 to 30% probability for a false negative component selection. For 2 days of simulation for a component that is only 10% weaker than the TwrBsMyt load, the probability of obtaining a false positive lies between 25 and 40%. This indicates over-dimentionalisation of the loads and acts as an inherent safety factor. Hence, safety factors will only add to the over-dimentionalisation of loads, leading to increased material use.

The minimum simulation length of 300 min described in IEC 61400-1 appendix F, does not have any value in determining the 50-year load of a wind turbine concept due to the upward bias and the excessive amount of uncertainty. Furthermore the 50-year load estimate calculation method leads to over-dimentionalisation, omitting the need for safety factors.

PREFACE

This thesis was written in order to obtain my degree of master of science in Sustainable Energy Technology. Sustainable energy technology has not only taught me the importance and pressing need for clean and renewable energy but has also made me understand natural every day phenomena like wind, solar, tides, earth's heat and the living world around us. Renewable energy technologies elegantly utilise these phenomena and I feel privileged to have a deep understanding of these processes.

This thesis evaluates the standards that were set by humans. These standards are chosen with great care, but need to be updated from time to time to avoid conformation to old ways of thinking. Here an attempt is made to aid in this serious process by shining a new light over standards in the wind turbine design industry.

The completion of this paper concludes my life as a student and grants me a solid base from which I can start my professional career. For this I want to thank my mother and father who have supported me throughout my study each in their own way. They gave me the biggest gift in life: the opportunity for personal development. Furthermore I want to make a huge compliment to my daily supervisor René Bos who has the skills and had the patience to teach me the skills and knowledge needed to complete this thesis. Also I want to thank my siblings for supporting me unambiguously with every choice I make. In addition I want to thank my friends for making my life a fun place to live. I have experienced some tough moments during the small hours in the night and I want to thank Maltesers for making these moments a little bit less tough. Last but not least a special thanks to Dick Kramers for helping me study in a new environment and Niek Overgaauw for reviewing my work.

S.F. van Eijk
Amsterdam, April 2016

CONTENTS

List of Abbreviations and Nomenclature

Abbreviations

A-D	Anderson-Darling test
AWEA	American Wind Energy Association
CCGT	Combined Cycle Gas Turbine
CDF	Cumulative Density Function
CI	Confidence Interval
DLC	Design Load Case
EROI	Energy Return on energy invested
EV	Extreme Value Distribution
GEV	Generalised Extreme Value Distribution
HAWT	Horizontal Axis Wind Turbine
IEC	International Electrotechnical Commission
LCOE	Levelised Cost of Electricity
LSQfit	Least-square curve fit
PDF	Probability Density Function
RMSE	Root Mean Square Error
SANDIA	Sandia National Laboratories
VAWT	Vertical Axis Wind Turbine
VG	Vortex Generator

Nomenclature

$p_{95\%}$	95% confidence interval
A	swept area of the rotor
a	uniform distribution parameter a
A_r^2	Anderson-Darling coefficient
$\hat{a}$	estimated parameter a
b	uniform distribution parameter b
$\hat{b}$	estimated parameter b
β	Gumbel scale parameter with $\beta > 0$
$\hat{\beta}(N)$	estimated Gumbel scale parameter from N data points with $\beta > 0$
$\hat{c}$	estimated parameter c
c	chord of the blade
C_l	lift coefficient
C_p	power capacity factor
E_{year}	yearly energy yield
$\epsilon_F(N)$	normalisation of $E[\hat{x}_F(N)]$
$\varsigma_F(N)$	normalisation of $\hat{\sigma}_F(N)$
$E[\hat{x}_F(N)]$	estimate mean of r sets to F extrapolated values
γ	Euler's constant
Ω/ω	exitation of the rotor
x_{50y}	50 year value
$F(x;\mu,\beta)$	cumulative distribution of a Gumbel
F_c	centrifugal forces on the rotor
L	lift force
ω_n	natural frequency of the rotor
$f(U)$	wind regime

h	hub height
h_{ref}	reference height of 10 m
M_a	aerodynamic moment on the rotor
μ	Gumbel location parameter
$\hat{\mu}(N)$	estimated Gumbel location parameter from N data points
n	number of estimate evaluations
N	sample size
H_0	null hypothesis
M	number of iterations of design load cases
P	harmonics per rotation
Ω	rotational speed
P_r	power from the rotor
P_{el}	electric power output of the wind turbine
p-**value**	scalar parameter in range [0 1] for testing a statistical hypothesis
Q_r	torque on the rotor
r	the number of sets
R	rotor radius
r_h	rotor radius at the hub of the blade
ρ	air density
σ_r	rotor solidity
$\hat{\sigma}_F(N)$	estimated standard deviation
σ_A	aerodynamic stress on the rotor
σ_c	centrifugal stress on the rotor
σ_g	gravitational stress on the rotor
T_r	thrust on the rotor
λ	tip speed ratio
$N_{\boldsymbol{total}}$	total number of simulations available
U	wind speed
$U(h)_{ci}$	cut in wind speed
$U(h)_{co}$	cut out wind speed
$U(h)$	wind speed at hub hight
$U(h_{ref})$	wind speed at reference height
$\sigma_{\boldsymbol{A,c,g}}$	ultimate strength of wind turbine component
U_r	rated wind speed
W_r	weight of the rotor
$x_{i,j}$	sample value at set j and rank i
x_F	sample value of x at F
$\hat{x}_{F,j}(N)$	towards F extrapolated estimated value of set j
$x_{N,j}$	biggest sample value of set j being of rank N
z_0	local terrain roughness constant

List of Figures

LIST OF TABLES

1

INTRODUCTION

In this chapter a brief introduction to the wind turbine landscape is given. The first section covers the position of wind turbines in the sustainable energy technology market together with the fundamentals and a cost breakdown. In the problem definition the wind turbine load calculations in wind turbine standards is introduced. Thereupon the main research question and sub questions are based. The last paragraph describes the outline of the entire thesis briefly.

1.1. INTRODUCTION

With increasing droughts in desert areas, rising sea levels and increasing hurricanes in number and force, reducing greenhouse gas emissions is no longer optional but has become a necessity. Due to the increased awareness and acknowledgement of climate change and the willingness of world leaders to mitigate the effects of climate change, the number of renewable energy installations has grown exponentially in the last two decades.

Harnessing energy from natural airflows is one elegant way to produce sustainable electrical energy. There are encouraging properties in using wind turbines to produce sustainable electricity. For example the global feasible wind energy potential is estimated to exceed the world energy use one to six times (Wiser *et al.*, 2011). In addition, because wind turbines can be mass produced a large share of the worldwide wind energy potential can be utilised against ever decreasing costs. Furthermore the energy needed to construct a wind turbine, is generated between 5 and 33 times over the lifespan of that wind turbine (Kubiszewski *et al.*, 2010). The ratio of energy delivered to energy costs is called the energy return on energy invested (EROI). High EROI values yield significant economic and social benefits relative to other power generation systems (Kubiszewski *et al.*, 2010). Wind turbines show among the highest EROI values in figure 1.1, suggesting a favourable position for wind energy relative to other forms of renewable power generation (Kubiszewski *et al.*, 2010).

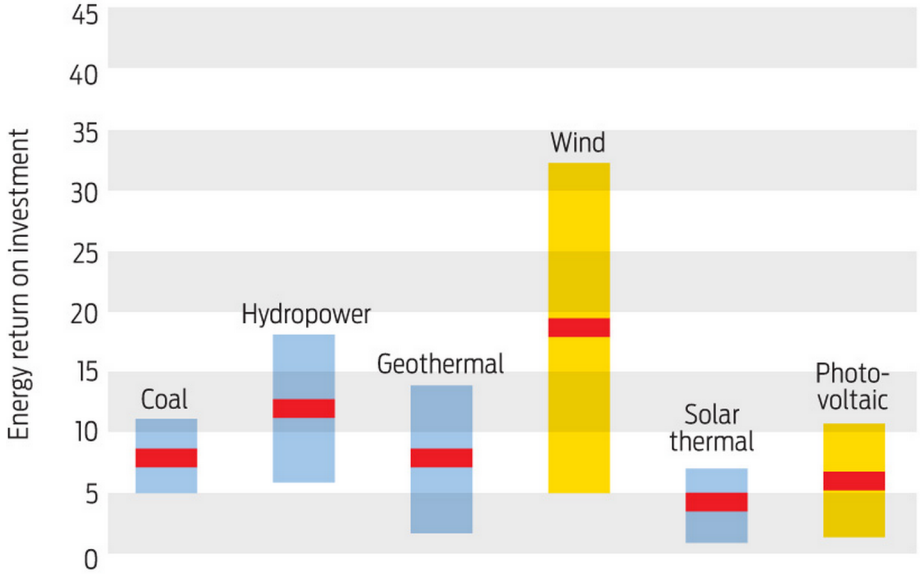


Figure 1.1: EROI and standard deviation of coal, hydropower, geothermal, wind, solar thermal and photovoltaic energy systems (Kubiszewski *et al.*, 2010)

However, there are still numerous challenges in the wind turbine industry. In most cases wind turbine systems do not produce electricity against grid parity yet which makes wind turbine projects dependent on government subsidies. Substantial research is done on wind turbine systems to achieve cost reduction in order to become cost competitive. The discounted cost price at which electricity is generated over the life-cycle of a wind turbine is called the Levelised Cost of Electricity (LCOE) and is further elaborated in paragraph 2.1. The LCOE allows for comprehensive comparison of electricity cost prices generated from different renewable sources. Estimated ranges of LCOEs are given in figure 1.2

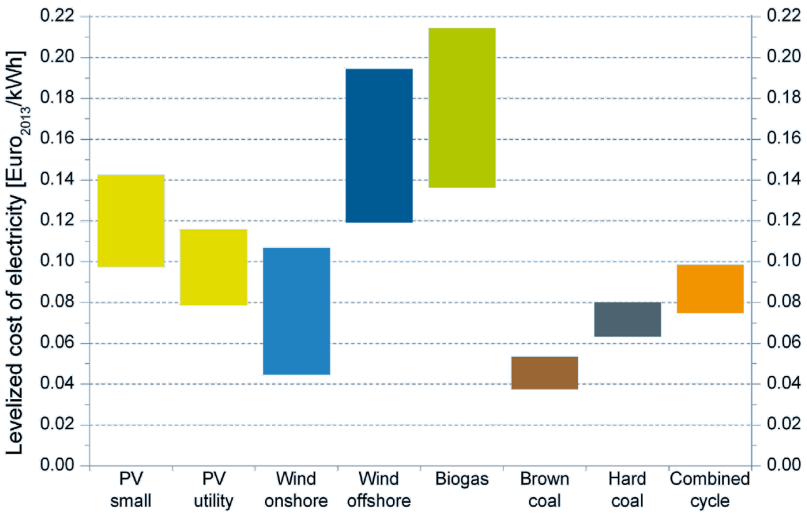


Figure 1.2: LCOE of Photovoltaic, Wind, Biomass, Brown coal, Hard coal and Combined Cycle Gas Turbine (CCGT) systems (Fraunhofer ISE, 2013)

Although wind turbines on land produce electricity cheaper than off-shore installations, off-shore wind farms are increasingly popular. Due to higher wind speeds at sea off-shore wind farms are able to produce relatively more electricity than on-shore wind farms. In addition, off-shore wind farms encounter less public opposition because of their low aesthetic impact.

Significant cost reductions in on-shore and especially off-shore wind turbine production and installation are

needed. A reduction of the LCOE of approximately 50% and 75% for on-shore and off-shore wind turbines respectively is required to achieve grid parity in 2030 (Fraunhofer ISE, 2013). The current trend to reduce the LCOE is to scale up wind turbines and this is confirmed by figure 1.3 which shows wind turbine designs since 1985.

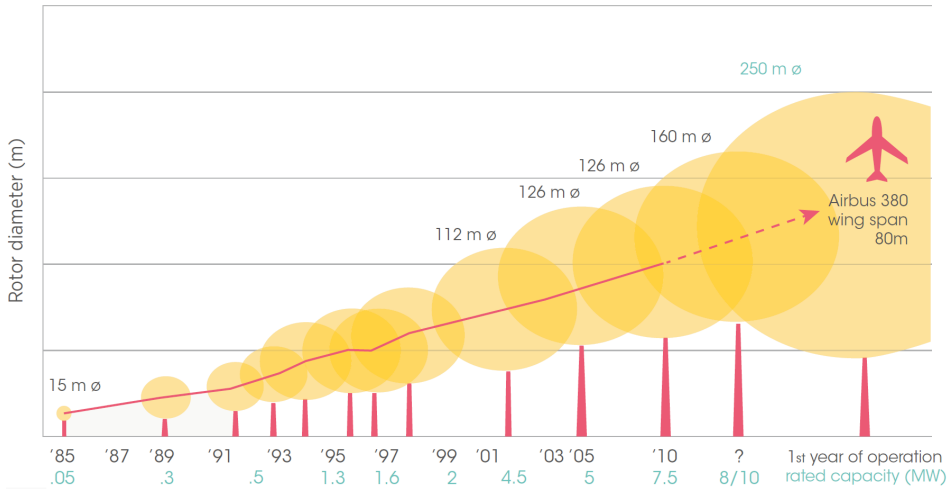


Figure 1.3: The trend of up scaling of wind turbines since 1985 (IRENA, 2012)

Key for the wind turbine up-scaling success lies in the the relative reduction of material use. Scaling-up wind turbines increases hub height(h) and rotor radius (R). This in turn causes wind speed (U) and the swept area of the rotor (swept area of the rotor (A)) to increase by relation (1.1) and (1.2) with wind speed reference height (h_{ref}), local terrain roughness constant (z_0) and wind speed at reference height ($U(h_{ref})$). A higher power production (P_r) following equation (1.3) can be achieved if a larger generating capacity is installed. Here C_p is defined as the wind turbine power coefficient and ρ the air density. Ultimately this will result in a higher energy yield (E_{year}) of the wind turbine in accordance with equation (1.4) where $U(h)_{co}$ and $U(h)_{ci}$ represent the cut-out and cut-in wind speed respectively. $f(U)$ is defined as the probability density of the wind speed, T as the number of hours in a year and P_{el} as the electric power output of the wind turbine.

Wind speed at hub hight:
$$U(h) = U(h_{ref}) \frac{\ln\left(\frac{h}{z_0}\right)}{\ln\left(\frac{h_{ref}}{z_0}\right)} \tag{1.1}$$

Swept area rotor:
$$A = \pi r^2 \tag{1.2}$$

Turbine power production:
$$P = \frac{1}{2} C_p \, \rho \, U(h)^3 \, A \tag{1.3}$$

Yearly energy yield:
$$E_{year} = T \int_{U_{ci}}^{U_{co}} P_{el}(U) \, f(U) dU \tag{1.4}$$

Thus energy yield increases with wind turbine size. However, this comes at the cost of turbine, blade, tower and foundation weight increase (Sieros *et al.*, 2012). Historically the increase in the weight of turbines has been limited by the utilisation of lighter materials and the optimisation of design, although it is not clear if this trend can continue (IRENA, 2012). The increase in energy yield ultimately slightly outweighs the offset material usage. In addition wind turbine scaling benefits wind farm scaling.

So the weight of the turbine appears to play a dominant role in wind turbine cost reductions. It can be deducted from the cost breakdown in figure 1.4 that average on-shore and off-shore wind turbine project costs

largely depend on the turbine. Cost reduction opportunities lie in material savings of the components for both off-shore and on-shore, although opportunities for cost reduction lie in multiple disciplines in off-shore wind energy. No single component dominates wind turbine costs, so savings are needed in all turbine components. The rotor is the highest cost item on most machines and requires the most reliability. Towers are normally the heaviest component and could benefit from weight reduction, but lightening the rotor or tower-top weight has a multiplier effect throughout the system including the foundation. The bulk materials in wind turbines are copper and steel and the cost of these commodities is therefore directly related to the cost of wind turbines. In off-shore projects steel and copper can account for as much as 20% to 40% of the total project cost (IRENA, 2012). So less use of material in the tower significantly reduces turbine cost.

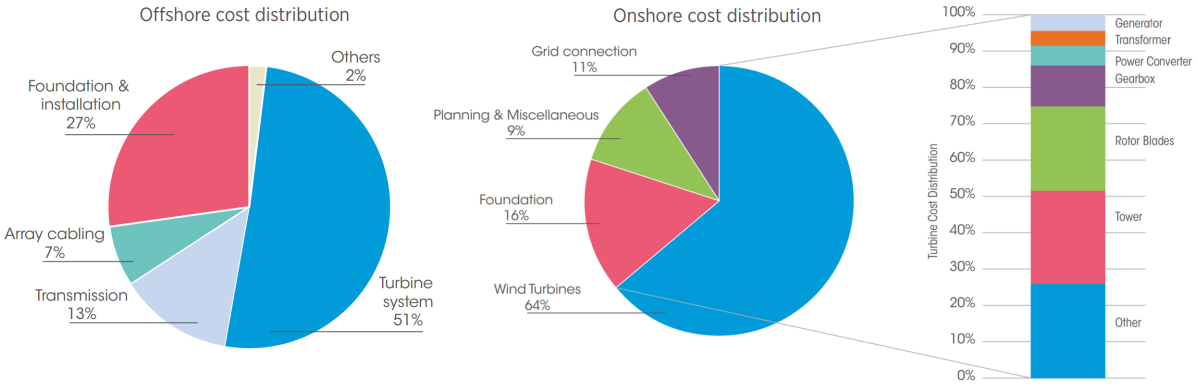


Figure 1.4: Capital cost breakdowns for typical on-shore and off-shore wind systems[1] (IRENA, 2012)

The tower and rotor material use is directly related to the forces that they must be able to resist. As it is costly to build concept wind turbines to analyse loads, wind turbulence and wind turbine load simulation models are used to evaluate the robustness of a concept. Because wind turbulence and wind turbine model calculation speed depend hugely on computational resources, a limited number of wind turbine load simulations can be performed (Fogle *et al.*, 2008). However, wind turbine load extremes are assumed to follow a particular trend. This trend is obtained from a limited amount of simulations and extrapolated to predict long term extreme load returns. These extreme values partly define the threshold force a wind turbine design must be able to withstand. Extreme load values for wind turbines are calculated for 1 and 50-year return values depending on the component and type of simulation. The ultimate load case calculations of 50-year return estimates possesses an unknown amount of uncertainty while being extremely relevant for it is the longest load simulation in the wind turbine design. Extrapolation techniques are allowed in calculating 50-year wind turbine load estimates and are specified in industry standards. European wind turbine standards are specified in the International Electrotechnical Commission (IEC) 61400-1 for on-shore IEC (2005) and in IEC 61400-3 IEC (2009) for off-shore wind turbine design. In the United States of America standards are specified by the American Wind Energy Association (AWEA). These standards are further elaborated in Chapter 2.

1.2. PROBLEM DEFINITION

Extrapolated load simulations influence wind turbine material use which in turn has a significant impact on wind turbine cost. The estimates of especially the 50-year extreme load returns is know to be subject to a significant amount of uncertainty. Inherent uncertainty exists in the calculation of wind turbine estimates. Safety factors are used to account for this. However, little is know about the behaviour of 50-year extreme wind turbine load estimates. In theory it could be possible that a cheap wind turbine design is falsely rejected because load simulations are extrapolated too high. Therefore identifying the exact uncertainty dependence can be of great value to wind turbine design. Based on an uncertainty inventory, safety factor use can be accurately weighted against known uncertainty in extreme load calculations. So the following advantages could be achieved by obtaining uncertainty quantification in wind turbine 50-year load estimates:

[1] The share of different cost components varies by country and project, depending on turbine costs, site requirements, the competitiveness of the local wind industry and the cost structure of the country where the project is being developed. On-shore turbine cost can vary from 64% to as much as 84% of the total project.

- Different stages of the wind turbine design process suit different levels of uncertainties, knowledge about 50-year load estimate uncertainty can help wind turbine designers better allocate computational resources during the design process.
- Material use in wind turbine design is based on the loads that the wind turbine must be able to withstand. The load calculation of 50-year estimates contains uncertainty which is compensated by means of heavy safety factors. Quantification of 50-year extreme wind turbine load extrapolation uncertainty can guide accurate allocation and intensity of safety factors. This will mitigate wind turbine designs that are able to withstand greater forces, moments and deflections than needed, leading to high material use and ultimately to a costly wind turbine design.

Furthermore, recently results of 96 years of on shore wind turbine load simulations have been published (Barone *et al.*, 2012a) as has data over 63 years of off-shore wind turbine load simulation (Barone *et al.*, 2012b). There exists great potential in this extensive data set to do research on the calculation of 50-year extreme load returns and the uncertainty associated with this calculation. The data set holds accurate 10-min extreme wind turbine load sample distributions and enables benchmarking of load estimates. Hence, this thesis will focus on identifying the uncertainty of 50-year extreme load calculation. This yields the following main research topic and sub questions:

How does the degree of uncertainty in 50-year extreme wind turbine load estimates depend on computational resources (e.g. the number of load simulations) and how does that impact wind turbine design?

1. What is the standardised procedure and safety factor use for 50-year extreme wind turbine load calculations?
2. What research has already been conducted on 50-year extreme wind turbine load estimate calculations?
3. What factors play a role in uncertainty of general extreme value estimate calculations?
4. What uncertainty exists in 50-year extreme wind turbine load estimates?
5. What factors play a role in 50-year extreme wind turbine load estimates and how can adverse behaviour be mitigated?
6. How does the uncertainty in 50-year extreme wind turbine load calculations influence the wind turbine design process?

1.3. THESIS OUTLINE

Sub questions one and two will be answered in chapter 2: Theory. This chapter gives a complete overview of the wind turbine design process and indicates where 50-year extreme load calculations is situated in the design process. This topic is worked out further to give the research context. Furthermore an extended literature study towards uncertainty in extrapolation and for wind turbines in particular is included. Topics such as: load calculation simulation procedures, the current status of 50-year extreme wind turbine load extrapolation and previous research on 50-year extreme wind turbine load extrapolation and estimate calculation will be covered here.

Chapter 3: Estimates by Extrapolation will deal with sub question three. This chapter will cover the fundamentals of uncertainty in extrapolation because remarkable behaviour was discovered during the research which warranted further investigation. It was found paramount to understand this behaviour before further data analysis could be preformed on wind turbine simulation data.

The SANDIA Data Set is covered in chapter 4 and forms the bases of the data analysis. Here, a detailed description and exploration of the data is performed and 50-year load estimation uncertainty is quantified. From this analysis it must be possible for wind turbine designers to deduct uncertainty levels for appropriate wind turbine design stages thereby answering sub question four. Furthermore an attempt is made to explain wind turbine load behaviour on which an additional new idea to mitigate adverse extrapolation behaviour is

based and presented which answers sub question five.

The practical consequences for the level of uncertainty present in wind turbine estimates is discussed in Chapter 5: Managing 50-year Wind Turbine Load Estimates. The discussion consists of two main topics: the dilemma that wind turbine designers face due to limited computational resources and how this can be managed by means of uncertainty quantification. Thereby answering sub question six. In addition the use safety factors is evaluated.

2

THEORY

In this chapter an attempt was made to accurately describe the wind turbine design process in order to understand the position of standardised procedure and safety factor use of 50-year extreme wind turbine load estimate calculations. In the first paragraph all phases of the wind turbine design process are covered with an emphases on wind turbine load calculation and standards. In the second paragraph the first sub question is answered by covering the IEC 61400-1 design standard is covered in. The third paragraph covers the ultimate stochastic wind turbine design load from the IEC 61400-1 design standard. The research done on this load case answers sub question two. The final section gives a brief conclusion of this chapter.

2.1. THE WIND TURBINE DESIGN PROCESS

A wind turbine consists of many different parts which have to work seamlessly together to harvest the power from the wind in a cost effective way. The wind turbine designer has to take many aspects into consideration where minimising the cost of energy is key. This can be achieved by using components consisting of inexpensive materials and minimising the material use. This often results in the use of of relatively small components which are susceptible to fatigue stress, especially due to the fluctuating nature of wind turbine loading for a long period typically being 20 years. This makes designing a strong reliable wind turbine whilst still keeping repair and operation costs at a minimum a challenge. Hence, the fundamental concern for wind turbine designers is to balance initial cost with the requirement of a long term fatigue-resistant wind turbine design. One of the goals of this thesis is to provide a helping hand for wind turbine designers to aid decision making in this dilemma by quantifying the uncertainty in load calculations.

As there are multiple approaches for wind turbine design, a general slightly altered design process after Manwell *et al.* (2009) is given below. The design process had been divided into three columns. The first column represents a phase in the design process that is characterised by major decisions of the project and is called the *conceptual design phase*. In this phase the broad scope and application of the wind turbine is defined which includes decisions and assumptions about the application, the type of wind turbine, the environment and an assessment about lessons learned from previous projects. With these decisions, input is created for a tentative wind turbine design. The first draft will by evaluated and refined in several steps. This iterative process comprises the second phase and here the wind turbine blueprint is formed. The second phase in the second column is called the *detailed design phase* and constitutes the core engine of the wind turbine design process. The third column represents a phase that is characterised by construction. Here a prototype is build and practical issues are resolved. Finally a wind turbine production and assembly plant is constructed and wind turbines are produced. This phase is called *prototyping & production*. All processes of the design stages are numbered within figure 2.1 and will be covered in short below.

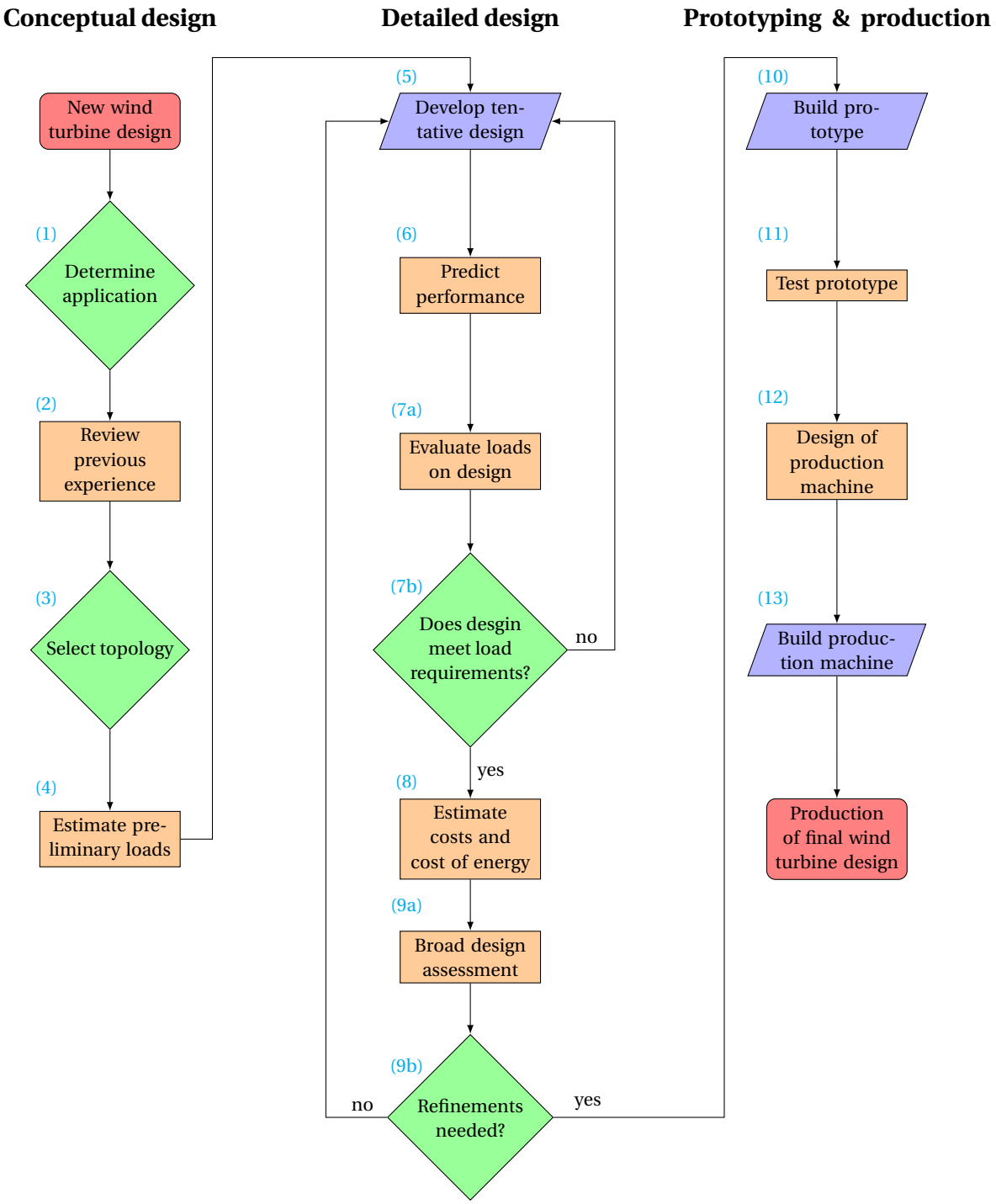


Figure 2.1: Wind turbine design process after Manwell *et al.* (2009)

(1) *Determine application:* the wind turbine design process begins with the demand for a new wind turbine design. First the the application of the wind turbine is defined. There are significant differences for wind turbines designed for stand alone off-grid systems than wind turbines that are developed for on-grid bulk electricity production. Wind turbines for bulk electricity production are generally bigger (1 to 5 MW) than typical off-grid remote communities (10 to 500 kW). Other design considerations are ease of installation, maintenance and simplicity in construction.

(2) *Review previous experience:* Since the repowering of wind turbines in 1995 a broad range of wind turbines have been conceptualised in multiple projects. Twenty years of experience can save costly mis-

takes and helps designers to narrow down options.

(3) *Select topology:* many options exist in defining the wind turbine model and most affect rotor operation. Ten key options for wind turbine designers are listed and elaborated below:

1. Rotor axis orientation: This is the most fundamental decision in wind turbine design. The Horizontal Axis Wind Turbine (HAWT) has its rotor orientation parallel or roughly to the ground. The Vertical Axis Wind Turbine (VAWT) has its rotor axis perpendicular to the ground. Examples of three main types of lift based designs are given in figure 2.2. Here two types of VAWT exist, the Giromill and the Darrieus models.

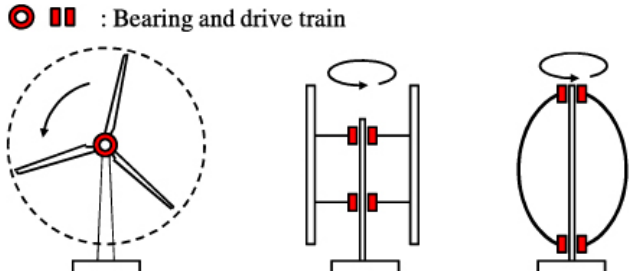


Figure 2.2: (a) HAWT (b) VAWT-Giromill (c) VAWT-Darrieus

Most modern wind farm projects make use of HAWT wind turbines. Three main reasons for this trend are:

+ The ratio of the total area of the rotor blades to the swept area of the rotor (A) and is called the rotor solidity (σ_r), here c is the chord of the blade and r_h is the rotor radius at the hub of the blade:

$$\sigma_r = \frac{1}{A}\int_{r_h}^{R} c\, dr$$

  The rotor solidity is lower for HAWT than for VAWT. This reduces weight of the blades and as explained in the introduction reduces the material cost through the wind turbine resulting in a lower cost of energy for HAWT types.
+ The swept area of the rotor is located at a higher average altitude for HAWT than for VAWT. At higher altitudes higher wind speeds reside (equation 1.1) and thus a higher powers and electricity production can be achieved.
+ The rated winds speed is relatively high in comparison to VAWT.
+ Because there is more experience with HAWTs the technology is more advanced than VAWT which makes it more attractive to investors.
- Difficult maintenance because there is no easy access to the drive train components up in the tower.
- High loads on the tower due to high center of gravity. This adds to the material cost.

The up and downside of VAWTs are:

+ Because the rotor blades form a round figure, the rotor can handle wind from any direction without yawing. So there is no need for a yawing system in VAWT systems.
+ In models such as depicted in figure 2.2(a/b/c) the chords do not need a chord twist because the whole blade experiences the same rotational speed, the angle of inflow remains constant over the blade. The constant chord makes for a simple and therefore cheap production process for the blades.
+ Except for the shaft, the whole drivetrain can be located on the ground. This makes installing and maintenance easy and cheap.
+ Because VAWTs have a low centre of gravity, they make a good solution for floating wind turbines.

- The connection points of the blades with the rotor are highly susceptible to fatigue damage.
- Incompatibility between the structure and aerodynamic control limits the power production above rated wind speeds. This causes drive train components to be larger than needed, adding costs.
- Since the best application for VAWT is to integrate them in high buildings, the noise level can be problematic.

Because the vast majority of modern wind turbines are HAWT and because the SANDIA data set (Barone *et al.*, 2012a) also assumes a HAWT, all further comments about wind turbines will assume a HAWT design if not stated otherwise. In the figure below the position and definition of the wind turbine components discussed in this thesis are given.

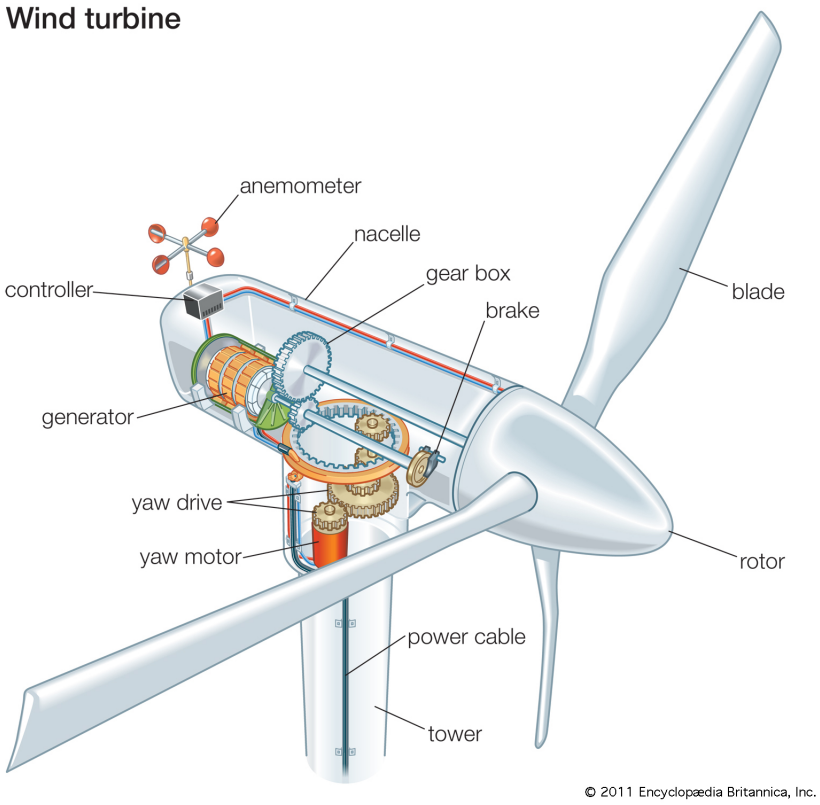


Figure 2.3: Wind turbine components (Britannica, 2015)

Loads on wind turbine design occur on various components. The key components have their own coordinate system to describe the load vectors. In this thesis six different coordinate systems: blade, chord, hub, rotor, yaw bearing and the tower coordinate system are considered. These coordinate systems are graphically represented and defined in appendix A.

2. Power control: Stall control is the result of reduced lift at large angles of attack. Low lift leads to low torque and lower power output. In a stall-controlled wind turbine maximum power is reached at high wind speeds leading to high torques on other parts. Also other breaking systems need to be in place in order to guarantee wind turbine shut down in extreme conditions. The blades can be controlled to rotate around their own axis in variable-pitch wind turbines. This complicates the attachment of the blades to the hub but will lead to more control options which results in achieving rated power at lower wind speeds than a stall controlled system leading ultimately to lower torques on the system. Another system that can manage lift are Vortex Generator (VG)s doing so at the blade roots. The last major power control option for wind turbines is yaw control. Yaw is the rotation towards or away from the wind direction. For HAWT this system is a must and thus the tower must be resistant against gyroscopic forces.

3. Rotor position: the rotor can be positioned downwind or upwind of the tower. Downwind rotors have the advantage that thrust is compensated by centrifugal forces. Therefore the blade root bending moments are relatively small. However, downwind rotor operates in the wake of the tower. The periodic passing of the blades induces periodic loading which increases the fatigue damage on the blades. In addition less power can be extracted from the wind and more noise is produced.

4. Yaw control: can be free or active. Downwind wind turbines follow the wind naturally like a weather vane and to help this movement the blades are coned some degrees to the downwind direction. Upwind turbines have to have active yaw control for the rotor to follow the wind direction. Gyroscopic forces are released on the gears and tower when the yaw system fixed and aligned with wind direction. Ultimately the tower must be able to resist these forces.

5. Rotor speed: knows two topologies: constant or variable. Optimum tip speed ratios can be achieved by using variable-speed rotors. At low wind speeds the tip speed ratio is optimised in order to optimise power output (to keep the power coefficient maximised), and at above-rated wind speeds the tip speed ratio is lowered to reduce loads on the drive train and keep power constant (keep power coefficient constant). For this reason modern wind turbines are equipped with variable-rotor speed option. Constant-speed rotor speeds are less complex and the rotational speed is determined by electrical generator and gearbox.

6. Design tip speed ratio and solidity: the design tip speed ratio (2.1) has a major impact on the design on the turbine:

$$\lambda_{design} = \frac{\Omega R}{U(h)} \tag{2.1}$$

The design tip speed ratio has a direct influence on the rotor solidity. A high design tip speed ratio rotor requires relatively less blade area (small chord) than a low tip speed ratio design with the same diameter in order to produce the same power, as visualised in figure 2.4. There is a structural limit to the thinness of a blades. If the design tip speed ratio is high, then the number of blades will decrease to provide for structural stability in the blades. High tip speed ratios are preferred for a number of reasons. Less air-foil surface is needed so lighter blades can be used. This decreases material of the blade and trough the turbine. Furthermore lower torques are accompanied at higher rotational speeds, reducing the load on the turbine. Also the balance of the drive train becomes lighter due to lesser conversion steps to high rotor RPM. However, high tip speed ratio wind turbine designs tend to be noisier.

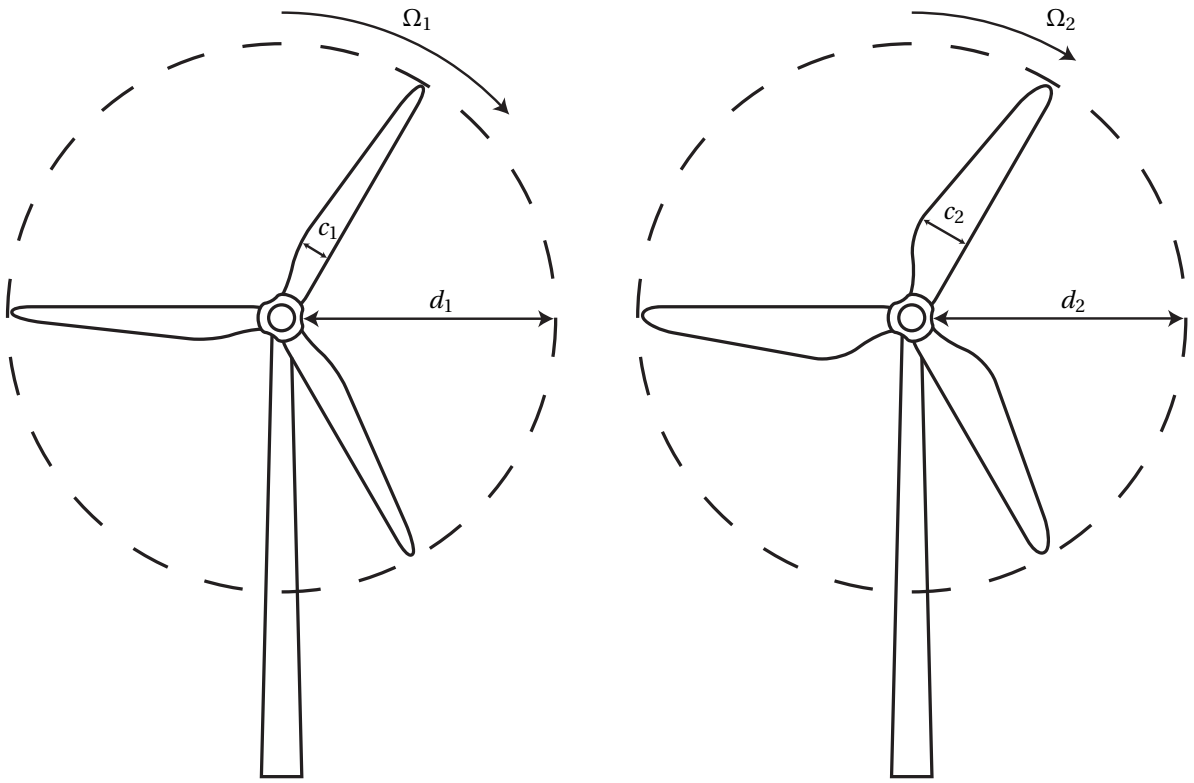


Figure 2.4: Relation of rotational speed (Ω) and chord length (c) with $d_1 = d_2$

The up-sizing of wind turbines has led to lower to more flexible structures and consequently to lower natural frequencies of the structure. High tip speed ratios lead to high frequencies. With bigger flexible structures it becomes increasingly difficult to prevent resonance of large off-shore wind turbines with the operating speed. Turbines with low tip speed ratios often possess rigid structures to prevent resonance.

7. Type of hub: the hub can assume rigid, teetering, hinged blades or gimballed;
   - *Rigid:*, here the blades can not move move in the flap wise and edgewise direction. Variable-pitch rotors also fall under this group of hubs.
   - *Teetering:* A part of the hub is installed with barrings to the rotor nacelle assembly. This enables the hub to teeter into and out of the plane of rotation. The advantage is of teetering is that the bending moments in the blades can be made very low due to upwind and downwind simultaneous movement of the blades.
   - *Hinged blades:* are used by some two-bladed wind turbines and allow independent in and out of plane movement of the blades. Special precautions must be made when the turbine is stopped.
   - *Gimballed hub:* uses a rigid hub but the whole rotor nacelle assembly is installed on horizontal bearings so that the machine can tilt up or down from horizontal. Imbalances in aerodynamic forces can be relieved this way.

8. Number of blades: Increasing the design tip speed ratio for a given solidity and all other wind turbine factors being constant, decreases the number of blades due to the structural limits of the blades. Especially the stresses increase with decreasing number of blades. The vast majority of the deployed wind turbines are three-bladed. An advantage of wind turbines with three blades or more is that the polar moment of inertia with respect to yawing is constant and thus independent of the position of the rotor. This makes for smooth yaw operation. Two-bladed wind turbines have a larger yaw moment of inertia when the blades are horizontal then when they are positioned vertically. This imbalance causes most two bladed turbines to use a teetering rotor.

9. Generator speed: has three forms: synchronous speed, multiple synchronous speeds or variable speed. The choice for generator type will determine to a large extend the configuration of the drive train. Low speed direct drive shafts need more power electronics such as a converter to produce the required 50 HZ grid output. Furthermore generators tend to be large due to the high number of poles (n) needed to boost output frequency:

$$f_{AC} = \frac{\Omega n}{4\pi} \tag{2.2}$$

Variable high speed shaft generators with variable pitch are the most common configuration and it needs a gearbox to keep the rotational to keep track of the rotational speed. Often the gearbox is a hybrid of a parallel-shaft gearbox and planetary gearbox.

10. Tower structure: height is chosen based on an economic trade-off based on energy gain at higher altitudes due to higher wind speeds and increased tower costs. This can be seen from wind turbine towers in urban areas where there is much turbulence at the ground. There the tower height and swept area ratio is large in comparison to off-shore wind turbines where relatively low turbulence and high wind speeds exists at the surface.

(4) *Estimate preliminary loads:* when the topology has been set, a global wind turbine design can be modelled. With this model load estimations can be calculated. In this stage of the wind turbine design process the calculations will be based on well known scaling laws. These laws are derived from physical laws such as equation 1.1 and must be interpreted as 'rule of thumb'. For the known scaling laws in table 2.1 to be valid, the following assumptions must be true:

- tip speed ratio (λ) remains constant;
- the number of blades, air foil and blade material are the same;
- geometric similarity is maintained to the extent possible

The most important scaling laws depend on the radius of the rotor (R) and are depicted in table 2.1.

Quantity	Symbol	Relation	Scale dependence
Power, forces and moments			
Power	P_r	$P_{r,1}/P_{r,2} = (R_1/R_2)^2$	$\sim R^2$
Torque	Q_r	$Q_{r,1}/Q_{r,2} = (R_1/R_2)^3$	$\sim R^3$
Thrust	T_r	$T_{r,1}/T_{r,2} = (R_1/R_2)^2$	$\sim R^2$
Rotational speed	Ω	$\Omega_1/\Omega_2 = (R_1/R_2)^1$	$\sim R^{-1}$
Weight	W_r	$W_{r,1}/W_{r,2} = (R_1/R_2)^2$	$\sim R^3$
Aerodynamic moments	M_a	$M_{A,1}/M_{A,2} = (R_1/R_2)^3$	$\sim R^3$
Centrifugal forces	F_c	$F_{c,1}/F_{c,2} = (R_1/R_2)^2$	$\sim R^2$
Stresses			
Gravitational	σ_g	$\sigma_{g,1}/\sigma_{g,2} = (R_1/R_2)^1$	$\sim R^1$
Aerodynamic	σ_A	$\sigma_{A,1}/\sigma_{A,2} = (R_1/R_2)^0 = 1$	$\sim R^0$
Centrifugal	σ_c	$\sigma_{c,1}/\sigma_{c,2} = (R_1/R_2)^0 = 1$	$\sim R^0$
Resonances			
Natural frequency	ω_n	$\omega_{n,1}/\omega_{n,2} = (R_1/R_2)^1$	$\sim R^{-1}$
Excitation	Ω/ω	$(\Omega_{r,1}/\omega_{n,1})/(\Omega_{r,2}/\omega_{r,2}) = (R_1/R_2) = 1$	$\sim R^0$

Table 2.1: Most important wind turbine scaling laws (Manwell *et al.*, 2009)

(5) *Develop tentative design:* In this stage a number of subsystems are designed that make up the overall design that was determined in the conceptual design. Common subsystems are: rotor, drive train, nacelle and main frame, yaw system and tower. After the design, performance, load, environment and economic evaluations this stage is used to reconsider the design in order to conform to the requirements of performance standards.

(6) *Predict performance:* translates into constructing the power curve for the wind turbine design. In this stage the rotor, gearbox, generator and control system are considered. Furthermore the calculation procedure differs for variable or fixed-speed wind turbines.

(7) *Evaluate loads on design:* after the main configurations for the design have been set and the theoretical power and torque generation are determined the loads of the design are evaluated. The techniques used to evaluate the resistance of the wind turbine design loading are specified in standards. Wind turbines have to comply to these standards in order for the design to be certified. Certification ensures that wind turbine requirements are met by computer modelling tests and prototype testing. A summary of wind turbine standards is given in table 2.2.

Source/Number	Title
IEC WT01	IEC System for Conformity Testing and Certification of Wind Turbines Rules and Procedures
IEC 61400-1	Wind Turbines – Part 1: Design Requirements, edition 2
IEC 61400-2	Wind Turbines – Part 2: Safety Requirements for Small Wind Turbines
IEC 61400-3	Wind Turbines – Part 3: Design Requirements for Off-shore Wind Turbines
ISO/IEC 81400-4	Wind Turbines – Part 4: Gearboxes for Turbines from 40 kW to 2 MW
IEC 61400-11 TS	Wind Turbines – Part 11: Acoustic Emission Measurement Techniques
IEC 61400-12	Wind Turbines – Part 12: Power Performance Measurements of Electricity Producing Wind Turbines
IEC 61400-13 TS	Wind Turbines – Part 13: Measurement of Mechanical Loads
IEC 61400-14	Wind Turbines – Part 14: Declaration of Apparent Sound Power Levels and Tonality Value of Wind Turbines
IEC 61400-21	Wind Turbines – Part 21: Power Quality Measurements
IEC 61400-22 TS	Wind Turbines – Part 22: Conformity Testing and Certification of Wind Turbines
IEC 61400-23 TS	Wind Turbines – Part 23: Full-scale Structural Testing of Rotor Blades
IEC 61400-24 TR	Wind Turbines – Part 24: Lightning Protection
IEC 61400-25	Wind Turbines – Part 25: Communications for Monitoring and Control of Wind Turbines

Table 2.2: Wind turbine load standards

Currently of the above standards, IEC 61400-1 (IEC, 2005) forms the key design load standard. There are many different loads acting on a wind turbine. In assessing these loads, two types are considered: fatigue and ultimate loads. Fatigue occurs when material is exposed to cyclic loading causing minuscule material cracks to be enlarged until the material fails. An ultimate load consists of a load limit which is the maximum load that a material or component can withstand for 3 seconds multiplied with a safety factor. The following classes of loads are defined for acting on a wind turbine: steady, cyclic, stochastic, transient and resonance-induced loads. These are elaborated in appendix B. Each load class contains a number of specific load situations and these are evaluated for fatigue and ultimate loads, where ultimate loads are comprised of an extreme event and an expected event at normal operation. This is depicted in figure 2.5. A specific load situation that is evaluated for a particular load event is called a load case.

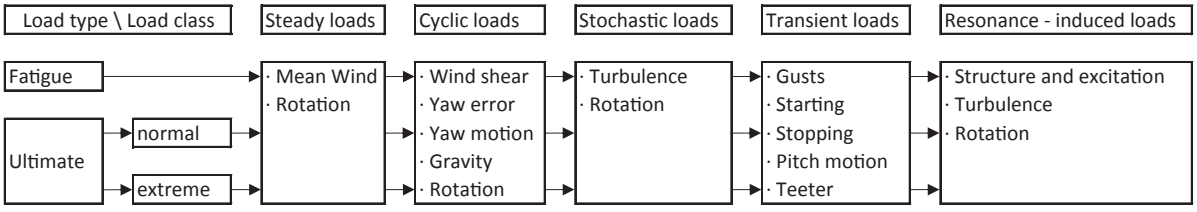


Figure 2.5: Load cases for different types of wind turbine loading

Experience with wind turbine loading has led to a selection of relevant load cases. These are codified in the IEC 61400-1 standard. Wind turbine designers need to comply to IEC 61400-1 in any country in which it is enforced. From figure 2.1 it can be noticed that the evaluation of the standards is included in two design loops. This causes the standards to be profoundly influential on the total wind turbine design. It is therefore that the research in this thesis focuses on the methods used to evaluate loads and the effect this has on the decision making in the design process. Because IEC 61400-1 is widely used, it is further elaborated in paragraph 2.2

(8) *Estimate costs and cost of energy:* when all relevant components are optimised to design standards the total cost and cost of energy of the wind turbine design is calculated. For the cost of the wind turbine the production, commercial components, custom components, installation and operational costs are considered. For the calculation of the cost of energy, the energy yield is calculated by wind turbine specifications of rotor diameter, height and turbine power together with site-specific weather conditions and effectively divided by the total costs. A widely used method is the formally discussed LCOE in which the total wind turbine cost is and the total energy production are calculated over the expected lifetime of the wind turbine (n), discounted and divided by each other according to the following formula:

$$LCOE = \frac{\sum_{t=1}^{n} \frac{I_t + M_t + F_t}{(1+r)^2}}{\sum_{t=1}^{n} \frac{E_t}{(1+r)^2}} \tag{2.3}$$

In which: I_t represents the investment expenditures, M_t are the operations and maintenance expenditures, F_t are the fuel expenditures which are 0 for renewable technologies, E_t represents the electricity generation and t indicates the year of the expenditure or electricity generation. Finally the discount rate is set with r. In the next design process step the cost and energy estimates are evaluated.

(9) *Broad design assessment:* in this design process stage all information obtained in the former design steps being the ability of the wind turbine concept to withstand loads, produce energy and to produce energy against a certain price are evaluated and benchmarked to predetermined boundaries. If the concept is found to be satisfactory the prototyping & production phase commences. If not, some design features in step 2.1(5) need to be re-evaluated and the whole detailed design phase is repeated.

(10) *Build prototype:* in this design process step an actual wind turbine prototype is build and installed on a site and is mainly used to verify assumptions in the wind turbine production and installation domain. Also small new concepts and adjustments are tested. The main goal of this stage is to ensure that the wind turbine can be fabricated, installed and operated as planned.

(11) *Test prototype:* here the prototype that has been build and installed and is subject to a wide variety of field tests. The main objective of this design process is to verify assumptions and calculation in the detailed design phase. So power curves of the prototypes are plotted and trough strain gauges the actual loads on the wind turbine are measured and compared to the predicted values.

(12) *Design of production machine:* taking the experience from the installation and production of the prototype the production line is designed. Also certain components that were adequate for the prototype might differentiate to the mass production version.

(13) *Build production machine:* during the production of the mass produced version, long term faults in the wind turbine design lead to slightly different production methods.

2.2. THE IEC 61400-1 DESIGN STANDARD

This thesis is concerned with wind turbine load calculations that wind turbine designs are tested against. As can be seen from figure 2.1 in the detailed design phase, step 7 evaluation of loads on design, is considered iteratively on two occasions. This makes for the standards that dictate the loads in this step to have a significant impact in the final design. This is the main reason why the IEC 61400-1 design standard is subject to significant research. This thesis focuses on Ultimate stochastic wind turbine design loads within the wind turbine design standard. This will be further covered in section 2.3.

This section is used to explain the overall design standard to get a good understanding of where the ultimate stochastic wind turbine design load fits into the IEC 61400-1 Design Standard procedure. The purpose of IEC

61400-1 is to: 'design requirements to ensure the engineering integrity of wind turbine and [...] to provide an appropriate level of protection against damage from all hazards during the planned lifetime.' This is done using the following procedure:

2.2.1. WIND PROFILES

There are three wind turbine design classes: I, II and III representing the most windy (50 m/s) to the least windy circumstance respectively in which the wind turbine normally operates. The wind speeds are generated using a type of distribution and are thereafter used to create a turbulence wind field. There are specific classes of wind turbulence which can be considered: A, B and C in decreasing turbulence intensity level order. For special cases a class S as been facilitated.

The wind speed and turbine intensity parameters can be combined in such a way that three types of wind condition groups can be identified Normal Wind Conditions, Extreme Wind Conditions and Rotationally Sampled Turbulence. Within these wind groups certain profiles or models exist. Depending of the location and purpose of the wind turbines applicable wind profiles and models are selected. These are the following:

- Normal Wind Profile (NWP)
- Normal Turbulence Model (NTM)
- Extreme Wind Speed (EMW)
- Extreme Operating Gust (EOG)
- Extreme Direction Change (EDC)
- Extreme Coherent Gust (ECG)
- Extreme Coherent Gust with Change in Direction (ECD)
- Extreme Wind Shear (EWS)

2.2.2. DESIGN LOAD CASES

The next step is defining the load cases. The load cases depend on the configuration of the wind turbine and the interaction with the wind field. In the standard there are eight relevant situations for which load cases exist:

1. Power production
2. Power production plus fault
3. Start-up
4. Normal shut-down
5. Emergency shut-down
6. Parked
7. Parked plus fault
8. Transport, assembly, maintenance and repair.

There can be multiple load cases in one situation. Many of the cases are ultimate load cases and are completed with fatigue load cases. The IEC 61400-1 load cases are presented in table 2.3.

Design situation	DLC	Wind Condition	Other conditions	Type of analysis	Partial safety factors
1) Power production	1.1	NTM $V_{in} < V_{hub} < V_{out}$	For extrapolation of extreme events	U	N
	1.2	NTM $V_{in} < V_{hub} < V_{out}$		F	*
	1.3	ETM $V_{in} < V_{hub} < V_{out}$		U	N
	1.4	ECD $V_{hub} = V_r - 2\text{m/s}$, $V_r, V_r + 2\text{m/s}$		U	N
	1.5	EWS $V_{in} < V_{hub} < V_{out}$		U	N
2) Power production plus occurrence of fault	2.1	NTM $V_{in} < V_{hub} < V_{out}$	Control system fault or loss of electrical network	U	N
	2.2	NTM $V_{in} < V_{hub} < V_{out}$	Protection system or preceding internal electrical fault	U	A
	2.3	EOG $V_{hub} = V_r \pm 2\text{m/s}$ and V_{out}	External or internal electrical fault including loss of electrical network	U	A
	2.4	NTM $V_{in} < V_{hub} < V_{out}$	Control, protection, or electrical system faults including loss of electrical network	F	*
3) Start up	3.1	NWP $V_{in} < V_{hub} < V_{out}$		F	*
	3.2	EOG $V_{hub} = V_r \pm 2\text{m/s}$ and V_{out}		U	N
	3.3	EDC $V_{hub} = V_r \pm 2\text{m/s}$ and V_{out}		U	N
4) Normal shut down	4.1	NWP $V_{in} < V_{hub} < V_{out}$		F	*
	4.2	EOG $V_{hub} = V_r \pm 2\text{m/s}$ and V_{out}		U	N
5) Emergency shut down	5.1	NTM $V_{hub} = V_r \pm 2\text{m/s}$ and V_{out}		U	N
6) Pared (standing still or idling)	6.1	EWM 50-year recurrence period		U	N
	6.2	EWM 50-year recurrence period	Loss of electrical network connection	U	A
	6.3	EWM 1-year recurrence period	Extreme yaw misalignment	U	N
	6.4	NTM $V_{hub} < 0,7V_{ref}$		F	*
7) Pared and fault conditions	7.1	EWM 1-year recurrence period		U	A
8) Transport, assembly, maintenance and repair	8.1	NTM V_{maint} to be stated by the manufacturer		U	T
	8.2	EWM 1-year recurrence period		U	A

Table 2.3: The IEC 61400-1 load cases

The types of analysis are ultimate strength (U) and fatigue (F). Safety factors depend on the situation of the load case and are denounced as: normal (N), abnormal (A) or transport and erection (T). The topic of safety factors will be further discussed in subsection 2.2.4.

2.2.3. APPLICATION OF DESIGN LOAD CASES

Four types of load analyses are done on the design load situation:

1. Analysis of maximum strength (ultimate strength).
2. Analysis of fatigue failure.

3. Stability analysis (e.g. buckling).
4. Deflection analysis (e.g. preventing blades from striking tower).

First the calculation of the expected loads values is done for a variety of operating conditions. From these estimates and dimensions of the components the stresses and deflections on the wind turbine are calculated. These stresses are then compared with the design stresses of the component. Consequently a change in component strength to withstand the calculated load stresses will influence the load stress. This iterative process continues until a safe and stable load / stress balance has been established.

2.2.4. METHOD OF PARTIAL SAFETY FACTORS

Because there are uncertainties in the load estimates and the characteristics of the material a partial safety factor is included in the load calculation. Specifying materials and sizing the various components requires a particular method:

- Determining design properties for materials by de-rating their characteristic (or published) properties.
- Selecting safety factors, which in effect increase estimates of the loads.

The requirement for ultimate loading is that the estimated *load function*, $S(F_d)$, multiplied by the *consequence of failure* safety factor, γ_n, must be equal to or less than *resistance function* $R(f_d)$. In the simplest case the load function is the highest value of the expected stress and the resistance function is the maximum allowable design value. The requirement follows:

$$\gamma_n S(F_d) \leq R(f_d)$$

where F_d = design values for loads and f_d = design values for materials. The design values for loads are found from the expected loads, F_k, by applying a partial safety factor for loads γ_f:

$$F_d = \gamma_f F_k$$

The design values for the materials are found from materials and stress tests. These materials always poses uncertainty in strength, f_k, is multiplied with safety factor for materials, γ_m

$$f_d = (1/\gamma_m) f_k$$

Partial safety factors are generally set greater than 1,0. Partial safety factors for loads range from 1,0 to 1,5. partial safety factors for materials are at least 1,1 and partial safety factors for consequence of failure are equal to at least 1,1.

2.3. DLC 1.1: THE ULTIMATE STOCHASTIC WIND TURBINE DESIGN LOAD

Every load case in table 2.3 has it purpose and use. DLC 1.1 contains with 50 years of simulation the longest period of operation of all DLCs. Because of the long simulation of real loads during operation, DLC 1.1 becomes highly relevant for load specifications. In chapter 1 it was covered that 50-year wind turbine load returns contain uncertainty due to calculation methods used. Wind turbine loads are described as stochastic variables. A stochastic process is defined as a process that ultimately contains a randomised input, as opposed to a deterministic process at which the whole process consists of a set of predetermined values. Any system or process that must be analysed using probability theory is stochastic at least in part. Ultimate wind turbine design loads are generated by a random input into a model that eventually yields a set of load values, of which the largest is the ultimate design load and is therefore stochastic. Depending on the length of simulation the data set increases and therefore also the likelihood for a higher ultimate wind turbine design load rises. The relation between this likelihood and increase in load values together with the model determines the uncertainty in the load returns. The high relevance of long load simulations and combined with the large degree of uncertainty present is problematic for the wind turbine industry. Wind turbine designers are obliged to calculate DLC 1.1 while outcomes are rarely used due to unpredictable nature of the outcome. This thesis therefore focuses on DLC .1.1. where it attempts to create more insight into the load calculation and uncertainty existing in DLC 1.1. and gives wind turbine designers a tool in managing wind turbine design load calculations.

Recommendations are done in appendix F of the IEC 61400-1 (IEC, 2005) standard on how to calculate 50-year load returns. There it is assumed that the largest load values occur at widely separated times and are thus statistically independent. The probability that the largest load F_{ext} exceeds a given load f in the observation time is then given by (Gumbel, 1958):

$$\text{Prob}(F_{ext} \geq F|V,T) = 1 - (F_{max}(F|V))^{E(n|V,T)}$$

where $F_{max}(F|V)$ is the short-term probability distribution function of the local maxima for the load process, and $E(n|V,T)$ is the expected number of local maxima in the observation time period. These statistical quantities are conditioned on the mean wind speed, V and where indicated also depend on the observation time period, T. The long-term exceedance probability is then given by integrating over the operating wind speeds,

$$\text{Prob}(F_{ext} \geq F|V,T) \equiv P_e(F,T) = \int_{V_{in}}^{V_{out}} \text{Prob}(F_{ext} \geq F|V,T)\, p(V)\, dV$$

where $p(V)$ is the probability density function for the hub hight wind speed defined by wind conditions. The acceptable probability of exceedance is reciprocal to the number of time intervals of length T in the recurrence period T_r associated with the characteristic load. The resulting characteristic load, F_k, is then given by solving the equation:

$$P_e(F_k,T) = \frac{T}{T_r}$$

Research and discussions about best practices on calculations for DLC 1.1 have been done extensively. In the next paragraphs a literature study will be performed on the research in order to position this thesis in the current research and to obtain previously acquired knowledge.

2.3.1. LITERATURE REVIEW

The guidelines for calculating DLC 1.1 above are vague at best. For example the number of simulations needed is not clearly specified. Furthermore no focus area is appointed to the most relevant load bins (for many load types around rated wind speed) although this might be desirable. In the past attempts have been made to improve the standard. In this section a literature review is given of that research.

2.3.2. GLOBAL AND BLOCK MAXIMA

In a paper called "Towards an Improved Understanding of Statistical Extrapolation of Wind Turbine Extreme Loads" of Fogle *et al.* (2008) global and block maxima are discussed. This research focuses on methods to determine the 10-min block load values. A global maximum load value is when only a single (highest) value of the 10-min block is taken. An other option is to take several time-separated (block) maxima. DLC 1.1 does not clearly describe whether extreme load statistics may be saved from each 10-min simulation in global or time separated block maxima illustrated in figure 2.6.

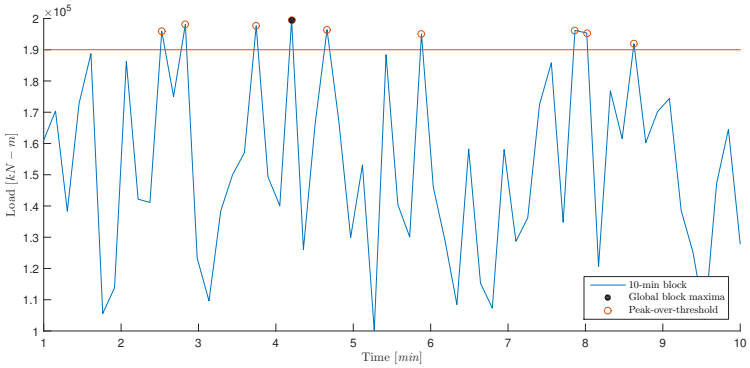


Figure 2.6: 10-min block maxima explained

A different load distribution appears with both techniques. Fogle *et al.* (2008) addresses the issue of independence in block sizes and statistical test for independence in which the predictions based on the use of global

an block maxima are covered. The analysis was done by means of LE[3] data set from a 5MW NREL wind turbine of 14.400 10-min wind speed bins simulations. It was found that block sizes of around 40-60 seconds for four load types (OopDefl, IPDefl, TwrBsMyt and LSSTipMya) led to independent block maxima when checked at the 1% significance level based on a well-accepted statistical test. Furthermore it was concluded that there is no advantage gained from single block maxima over global maxima when short-term loads are estimated.

2.3.3. CHOICE OF PROBABILITY DENSITY FUNCTION

As Fogle *et al.* (2008), Genz *et al.* (2006) investigates the peak-over-threshold by calculating an optimum. Furthermore the choice of probability density functions for fitting load simulation distribution is investigated. Four distributions are investigated:

- Gumbel
- 3-parameter Weibull
- Generalised Extreme Value
- 2-parameter Log-normal

The tests were based on goodness-of-fit data and visual comparison. Results indicate that the 3-parameter Weibull distribution with a central moment based estimation of parameters has the best fit. Also the threshold suggested in IEC-61400-1 is not robust for pitch regulated variable speed turbines and it may result in over estimation.

2.3.4. EXTRAPOLATION OF REAL LOAD VALUES VERSES EXTREME VALUE STATISTICS

In the paper of Peeringa (2009): "Comparison of extreme load extrapolations using measured and calculated loads of a MW wind turbine" 4 years of actual free stream and wake conditions resulting in loads of a 2,5 MW wind turbine are extrapolated and compared to extrapolated extremes produced by extreme value statistics. This is done for maximum blade root flap-wise bending moment (RootMFlp) [kN-m] and maximum tower base fore-aft bending moment (TwrBsMyt) [kN-M]. Furthermore attention is paid to the extreme value probability distribution and the amount of data needed. Global data approach is used in determining 10-min load block values yielding 150 measurements at 15,5 m/s. The empirical distributions are determined for $N = 35$ maxima.

Conclusions of the paper are that 50 maxima are sufficient to estimate the extreme load distribution conditional to the wind speed. Furthermore the GEV, Lognormal and Weibull fit both the RootMFlp and TwrBsMyt extreme load values well. The Gumbel overestimates the moments 4 to 11% and can only be used when it has the same shape as the GEV distribution. A comparison of empirical and simulated loads shows that the extrapolated calculated RootMFlp maxima are smaller than the measured maxima. For the TwrBsMyt the results are reversed. The order of differences is dependent on free, partial and full wake stream conditions but vary between 109% to 254%.

2.3.5. EARLY EXTRAPOLATION OF ULTIMATE EXTREME LOADS

In 2004 Moriarty *et al.* (2004) researched: "Extrapolation of Extreme and Fatigue Loads Using Probabilistic Methods". In the context of this thesis the focus is on the extreme ultimate loading using probabilistic methods. A general study was done on fundamental aspects of extrapolation of ultimate load extremes towards out of plane bending loads of a pitch regulated wind turbine. It was found that extrapolation methods over predict loads 20% or more. There it was noticed that models cannot follow the distribution due to loads near rated wind speed because there a change in behaviour of the loading takes place. Furthermore it was found that a Weibull model was most conservative, followed by a Gumbel, three parameter Weibull and finally a quadratic Weibull. The peak-over-threshold method was used to determine extremes and an attempt was made to optimise the threshold for most accurate load returns. It was found that the mean value plus 1,4 times the standard deviation of the original time series yielded the most accurate results. In the paper explicit emphasis is put on the need for alternative fitting.

2.3.6. SIMULATING THE ENTIRE LIFE OF AN OFF-SHORE WIND TURBINE

The original data for 96 years of wind turbine loads released by SANDIA (Barone *et al.*, 2012a) has triggered additional research like Barone *et al.* (2012b) for off-shore wind turbine loading. In the paper of Barone *et al.*

(2012b) 63 years of aero-hydro-elastic loads simulations where demonstrated for a 5 MW off-shore wind turbine. This allows for a new research area being the ability to benchmark 50-year loads estimates to extrapolation techniques. Emphasis in this paper is put on the uncertainty present in extrapolation techniques. Performing 2048 simulations gives a normalised interval width of approximately 10% for the extrapolated 1-year return load and 13% for the extrapolated 50-year return load of tower yaw moment. Explanation for this uncertainty is given by the linear regression model, which gives an inaccurate fit of especially the tail of the distribution. In addition the uncertainty is attributed to the fact that the highest extreme tower yaw loads occur at wind speeds near cut-out. In this thesis it is advised to sample separately for each mean wind speed bin and construct short-term load distributions for each bin and obtain 50-year value per bin.

2.4. CONCLUSION

Fundamental concern for wind turbine designers is to balance initial cost with the requirement of a long-term load-resistant wind turbine design. There are three design phases: conceptual design phase, detailed design phase and prototyping & production of which the detailed design phase has an iterative nature. This thesis focuses on the detailed design phase and in particular on the "predict performance" and "evaluate loads on design" design processes since these steps are being considered over multiple design loops. They therefore determine the material strength via load calculations and thereby the material use and total material cost of the wind turbine. Loading in these design processes are optimised towards the IEC 61400-1 design Standard making this design standard indirectly dictating wind turbine cost. In the current trend of cost reduction in the wind turbine design landscape, IEC 61400-1 is therefore highly relevant and evaluation of this standard is done extensively. This thesis focuses on DLC 1.1 since it is the design case that simulates the longest duration of loading.

In the literature review five research topics in ultimate wind turbine load calculations were found. The first topic constitutes the debate about the method in which the 10-min maxima is defined in order to create an extreme value distribution. The second topic consists of research on type of probability distribution that can be used to fit to the sample size. Thirdly, research is occupied with comparing real load and deflections to simulated and extrapolated load values. The fourth area comprises of research towards some fundamental aspects of obtaining 50-year load estimates by extrapolation of probabilistic methods. This lies closely to the topic of this thesis. However, the ability to benchmark the effects of wind turbine extrapolation against true 50-year load simulation values did not exist yet. With the release of the SANDIA data set Barone *et al.* (2012a) a new and last of the five research topics opened up. For now it becomes possible to quantify the load calculation uncertainty, enabling wind turbine designers to manage their number of simulations consciously. This will be further elaborated in chapter 5.

3

ESTIMATES BY EXTRAPOLATION

This chapter is used to evaluate the fundamentals of extrapolation answering sub question three. First a short explanation of the method used to evaluate uncertainty in extrapolation is given in section 3.1. Then the aspects of uncertainty in estimate calculations as itemised below are investigated in section 3.2:

3.2.1 Development of $\epsilon_F(N)$ over N
3.2.2 Development of $\varsigma_F(N)$ over N
3.2.3 Monte Carlo and Fit for $\varsigma_F(N)$ over N
3.2.4 Extreme Value Distribution Fit instead of a Normal Distribution Fit
3.2.5 Gumbel Plotting Positions

Lastly the conclusions of the research in this section will be given in section 3.3.

3.1. METHODOLOGY

As discussed in chapter 1, the aim is to quantify the distribution of an extrapolated Extreme Value Distribution (EV) estimates. The Extreme Value sample distribution used here is the Gumbel with parameters μ and β. This translates in obtaining estimated parameters $\hat{\mu}(N)$ and $\hat{\beta}(N)$ from a fixed known and sorted sample size N. These parameters are estimated using the Quantile Gumbel function and applying a linear fit. The Quantile Gumbel function is the inverse Gumbel CDF and the derivation is given in equation (3.3).

Gumbel CDF:
$$F(x;\mu,\beta) = e^{-e^{-\frac{(x-\mu)}{\beta}}} \tag{3.1}$$
$$-\ln\left[-\ln\left[F(x;\mu,\beta)\right]\right] = \frac{x-\mu}{\beta} \tag{3.2}$$
Quantile Gumbel function:
$$x(F;\mu,\beta) = \mu - \beta[\ln(-\ln(F))] \tag{3.3}$$

By using $\hat{\mu}(N)$ and $\hat{\beta}(N)$ the estimator $\hat{x}_{F,j}(N)$ can be calculated by extrapolation using equation (3.4) in tandem with the Weibull plotting position in equation (3.5). The role of plotting positions is further elaborated in section 3.1.1. The fit is done r times and thus r extrapolated estimators $\hat{x}_{F,j}(N)$ can be obtained.

Estimated extrapolated value:
$$\hat{x}_{F,j} = \hat{\mu}(N) - \hat{\beta}(N)[\ln(-\ln(F))] \tag{3.4}$$
with plot position:
$$F = \frac{i}{1+N} \tag{3.5}$$

To quantify the distribution of estimates by extrapolation, the mean of r $\hat{x}_{F,j}(N)$ values is taken according to equation (3.6) and the associated deviation is calculated by equation (3.7). The estimated mean and standard deviation are represented by the symbols $E[\hat{x}_F(N)]$ and $\hat{\sigma}_F(N)$ respectively.

$$E[\hat{x}_F(N)] = \frac{1}{r}\sum_{j=1}^{r}\hat{x}_{F,j} \tag{3.6}$$

$$\hat{\sigma}_F(N) = \sqrt{\frac{1}{r-1}\sum_{j=1}^{r}(E[\hat{x}_{F,j}(N)] - E[x_F])^2} \tag{3.7}$$

The matrix that is associated with this process is depicted in equation 3.8. There are r sets evaluated yielding r estimated extrapolations from a fixed amount of known values of rank N toward a fixed probability of rank F. This matrix is constructed for a n contiguous number of known data points. Furthermore the term (N) stresses the dependence on the number of known data points N from which the extrapolation took place. It should be noted that rank N must be $N \geq 2$, because no linear fit can be made for less than 2 known data points.

$$\begin{array}{c} \\ \text{set}\rightarrow \end{array}\begin{array}{c} \text{rank}\rightarrow \\ \begin{array}{ccccccccccc} x_{1,1} & < & x_{2,1} & < & \dots & < & x_{i,1} & \dots & \hat{x}_{N,1} & \rightarrow & \hat{x}_{F,1}(N) \\ x_{1,2} & < & x_{2,2} & < & \dots & < & x_{i,2} & \dots & \hat{x}_{N,2} & \rightarrow & \hat{x}_{F,2}(N) \\ \vdots & & \vdots & & & & \vdots & & \vdots & & \vdots \\ x_{1,j} & < & x_{2,j} & < & \dots & < & x_{i,j} & \dots & \hat{x}_{N,j} & \rightarrow & \hat{x}_{F,j}(N) \\ \vdots & & \vdots & & & & \vdots & & \vdots & & \vdots \\ x_{1,r} & < & x_{2,r} & < & \dots & < & x_{i,r} & \dots & \hat{x}_{N,r} & \rightarrow & \hat{x}_{F,r}(N) \\ & & & & & & & & & & \downarrow \\ & & & & & & & & & & E[\hat{x}_F(N)] \end{array} \end{array} \tag{3.8}$$

Finally, the $E[\hat{x}_F(N)]$ and $\hat{\sigma}_F(N)$ are normalised with respect to x_F trough equation (3.9) and (3.10) respectively.

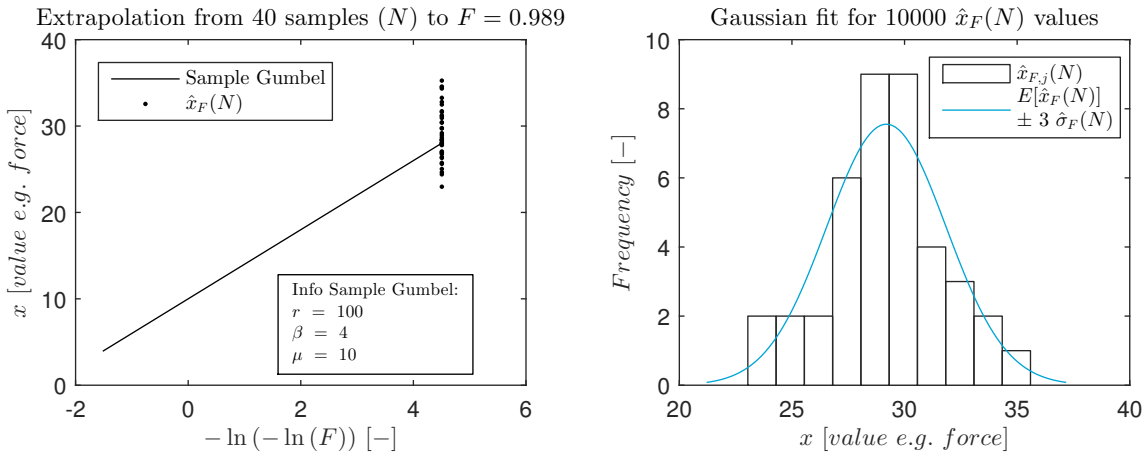
$$\epsilon_F(N) \equiv \frac{E[\hat{x}_F(N)] - x_F}{x_F} \tag{3.9}$$

$$\varsigma_F(N) \equiv \frac{\hat{\sigma} - x_F}{x_F} \tag{3.10}$$

3.1.1. EVALUATION PROGRAM

MATLAB® was used to program the above extrapolation process into the function `GumbelUnc`. `GumbelUnc` forms the core of the research done in the sections of this chapter with adjustments made fitting the research. For the sake of replicability and transparency of the research the code has been published in appendix C. Below a more detailed explanation is given to understand how `GumbelUnc` processes the above calculations.

In `GumbelUnc` the $x_{i,j}$ samples of a set are sorted and plotted verses a double natural logarithmic scale: $-\ln(-\ln(F))$. The double logarithmic scale presents the probability F by which a sample occurs and is situated on the x-axis. The sample values $x_{i,j}$ of a set are situated on the y-axis. Note that F can be defined as plotting position like equation (3.5) which is the Weibull (1939) plot position. There are many different plotting positions developed and all yield a slightly different Gumbel CDF curve. One of the research topics is the influence of different plotting positions on extrapolation. These will be researched in section 3.2.5. By using the linear relation in equation (3.3) and the known variables $x_{i,j}$ as well as the probability F, the parameters $\hat{\mu}(N)$ and $\hat{\beta}(N)$ can be obtained using the function `polyfit`. The estimated parameters form an estimated Gumbel which is extrapolated to F yielding $\hat{x}_{F,j}(N)$. $\hat{x}_{F,j}(N)$ values are stored and this process is repeated for r sets. The r $\hat{x}_{F,j}(N)$ values form a distribution upon which a Gaussian distribution fit is made from which $E[\hat{x}_F(N)]$ and $\hat{\sigma}_F(N)$ are obtained. The graphs in figure 3.1 illustrate this process. The line in figure 3.1(a) equals the sample Gumbel. Figure 3.1(b) represents the Gaussian fit to the $\hat{x}_{F,j}(N)$ distribution. In section 3.2.4 the use the Normal and EV distribution fits to the sets of $\hat{x}_{F,j}(N)$ are investigated.



(a) Multiple extrapolation with Weibull (1939) plot position $F = \frac{N}{N+1}$ yielding a $\hat{x}_{F,j}(N)$ distribution

(b) A normal distribution fit to $\hat{x}_{F,j}(N)$ yielding $E[\hat{x}_F(N)]$ and $\hat{\sigma}_F(N)$

Figure 3.1: The core process of evaluating the uncertainty in Gumbel extrapolation

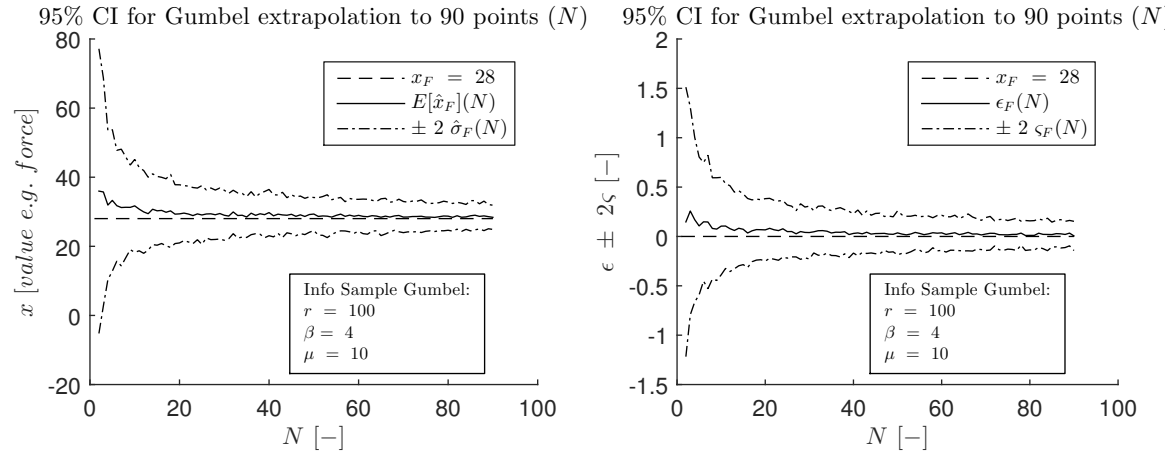
Uncertainty of extrapolation can now be constructed as a function of the number of known data points N. The process described in the matrix of equation 3.8 is thereon repeated for n times, storing all $E[\hat{x}_F(N)]$ and $\hat{\sigma}_F(N)$. In `GumbelUnc` defaults settings:

$$N \text{ takes on } 2,3,\ldots,N$$

and thus $n = N-1$ and Weibull (1939) extrapolation plot position becomes $F = \frac{N}{N+1}$. However, theoretically F could also be set to another value: e.g. for a 10-min event happening once every 50 years, the probability of non-exceedance F becomes:

$$F_{50y} = 1 - \frac{1}{50 * 365,25 * 24 * 6} \tag{3.11}$$

An output value in this form is called a 50-year value and is denoted x_{50y}. Furthermore n can be chosen on different intervals. The doubling sequence for $N = 16,32,64,\ldots,1024,2048$ is a much used interval in the field for it saves computational resources. The continuous linear series of `GumbelUnc` was chosen to produce a complete output. This returns the most graphical representation of uncertainty development over N and enables an accurate Monte Carlo test of the method which is covered in section 3.2.3. Multiple absolute $E[\hat{x}_F(N)]$ values and $\hat{\sigma}_F(N)$ deviation values from x_F over N are shown in figure 3.2(a). In figure 3.2(b) multiple $E[\hat{x}_F(N)]$ as $\hat{\sigma}_F(N)$ are normalised to x_F and denoted respectively $\epsilon_F(N)$ and $\varsigma_F(N)$ yielding factorised uncertainty.



(a) $E[\hat{x}_F(N)] \pm 2\ \hat{\sigma}_F(N)$ of absolute $E[\hat{x}_F(N)]$ values

(b) $\epsilon_F(N) \pm 2\ \varsigma_F(N)$ of normalised $\epsilon_F(N)$

Figure 3.2: The result of the function `GumbelUnc` which can show absolute or factorised uncertainty within the domain of $n = 2,3,\ldots,N$.

The operation that `GumbelUnc` preforms closely resembles a Monte Carlo simulation. The more samples drawn from a particular distribution, the more the sample distribution will resemble the distribution from which it is sampled. The sample variance in this process is defined in (3.12) and is known to have following general relation:

Unbiased estimate of variance:

$$var(\mu) \equiv \sigma_\mu^2 \tag{3.12}$$

$$var(\mu) = \frac{1}{N}\sum_{n=1}^{N} var(\hat{x}_n) \tag{3.13}$$

$$\begin{aligned} var(\mu) &= var\left(\frac{1}{N}\sum_{n=1}^{N}\hat{x}_n\right) \\ &= \frac{1}{N^2} var\left(\sum_{n=1}^{N} x_n\right) \\ &= \frac{1}{N^2}\sum_{n=1}^{N} var(x_n) \\ &= \frac{1}{N^2} N\, var(x_n) \\ var(\mu) &= \frac{1}{N} var(x_n) \end{aligned}$$

and

$$\sigma_\mu^2 = \frac{1}{N} var(x_n)$$

gives

$$\sigma_\mu = \frac{x_n}{\sqrt{N}} \tag{3.14}$$

$$\sigma_\mu \sim \frac{1}{\sqrt{N}} \tag{3.15}$$

First the dependence of N on the variance of μ is shown in (3.13). In equation (3.14) it can be seen that $\sqrt{var(x_n)}$ is a constant and thus it can be argued that the uncertainty deviation decreases with $\frac{1}{\sqrt{N}}$ shown by equation (3.15). Consequently `GumbelUnc` is expected to scale in the same way. This check is preformed in section 3.2.3.

3.2. RESEARCH

By generating samples from discrete sample generator, variations caused by external influences such as present in the SANDIA data set are kept to a minimum. Hence the effect of the calculation method in itself can be evaluated.

3.2.1. DEVELOPMENT OF $\epsilon_F(N)$ OVER N

In this section the behaviour of $\epsilon_F(N)$ over N is described. Graph 3.2(a) clearly shows a decreasing upward bias reaching up to 30% and decreasing over the entire length of the graph. After drawing a sample distribution a 100 times it might be expected that $E[\hat{x}_F(N)]$ values would average out on $E[x_F]$, but this is clearly not the case. This result is likely to occur in ultimate wind turbine load calculation since it uses the same method. This would indicate the stimulation of over-dimensionalisation of wind turbine loads and ultimately needlessly high wind turbine costs. This premise is covered in chapter 4 with actual wind turbine load data. The

results of this fundamental extrapolation experiment from `GumbelUnc` are confirmed by a study of White. Already in 1969 this phenomena was reported and researched by White in his paper: *The Moments of Log-Weibull Order Statistics*. There it is proved that function (3.16) describes the development of a normalised $E[\hat{x}_F(N)]$ mean over N which was formerly defined as $\epsilon_F(N)$.

$$E(i,n) = \sum_{j=0}^{j=i-1} \binom{n}{j} \Delta^j M(n-j) \tag{3.16}$$

Here $E(i,n)$ is the expected normalised deviation from any integrable function $g(X_{in})$ in $E\big(g(X_{in})\big)$. The symbols n is formerly defined as N in this thesis and i would be the equivalent of i in the domain of n:

$$N \text{ takes on } 2,3,\ldots,i,\ldots,N$$

Furthermore, the expression $\Delta^j M(n-j)$ describes the difference between the expected outcome x_F and the actual extrapolation outcome $E[\hat{x}_F(N)]$ while the function $M(n)$ represents the initial condition. In the paper *Simulating the Entire Life of an Off-shore Wind Turbine* of Barone *et al.* (2012b) one of the first attempts is made to evaluate uncertainty in extrapolation of wind turbine model generated data. In this paper the same upward bias is observed and is believed to have two causes:

1. *"due to the use of a linear regression fit to a distribution that has curvature present in the tail".*
2. *"due to the fact that the highest extreme tower yaw loads occur at wind speeds near cut-out"*

Although these causes may explain the upward bias, the role of a fundamental mechanism for extrapolation as discussed by White (1969) must also be considered. In the paper of White (1969) only a certain amount of functions for $g\left(X_{i,n}\right)$ are validated and no clear explanation is given on the upward bias. Here it is believed that it has most likely to do with the sorting of the data. With the help of a simple example in which the case of a complete data set is compared to a data set with a missing value. For example, six samples are drawn from a simple uniform Probability Density Function (PDF): $f(x) \sim U(a,b)$. By integration $\int f(x) \sim \int U(a,b)$ becomes the linear CDF $F(x) = \frac{x-a}{b-a}$, of which a= 2 and b= 8:

$$x = [3,63,\ 4,25,\ 4,81,\ 6,83,\ 7,20,\ 7,26]$$

These are sorted and plotted in figure 3.3. Due to the sorting, irrespectively of the missing sample, the next sample in line is always equal or bigger than the missing sample:

$$x_{j+1} \geq x_j$$
$$at$$
$$i_{j+1} = i_j + 1$$

Hence, the $\frac{dx(i)}{di}$ is always larger in the case of a sorted missing sample than in the complete case. Therefore a positive trend for the steepness of the slope in sorted data sets exists and could induce a higher $\hat{x}_{F,j}(N)$ than x_F. Furthermore the development of μ has also been investigated for highly unlikely events such as N equals 2000 in the Weibull plot position equivalencing the exeedance probability $1-F$ to 0,0005. Figure 3.4 shows the uncertainty in extrapolation from N= 500 to 2000.

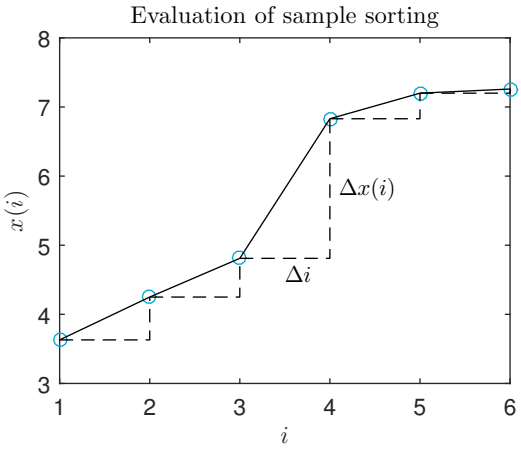


Figure 3.3: Slope of sorted data

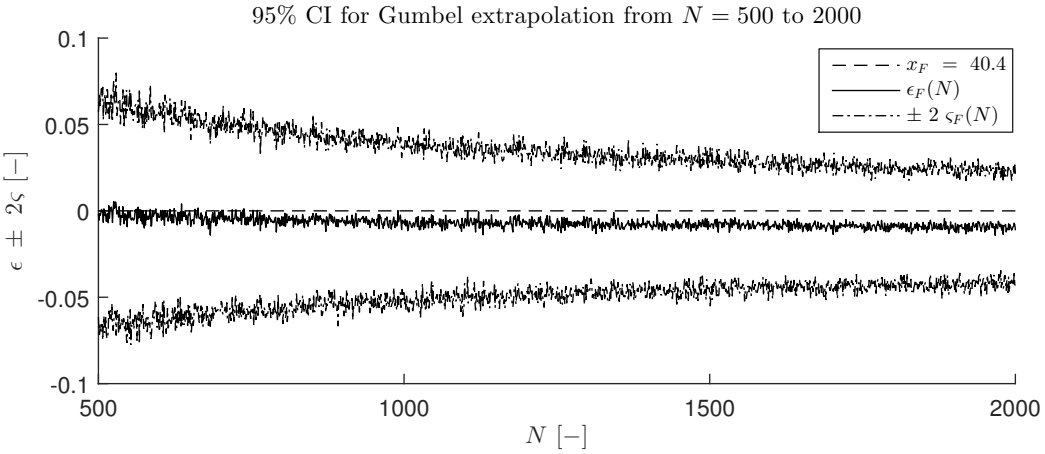


Figure 3.4: Extrapolation from 500 to a 2000 sample size

Contrary to the upward bias at small F, at large F a clear linear increasing downward bias can be observed from $N \gtrsim 500$. Because of limited computational resources in 1969 the paper of (White, 1969) only covers $N \rightarrow 20$ and has therefore not noted the downward bias. No literature could be found on this and neither a good mathematical explanation can be given to explain this phenomenon. The downward bias becomes −1% and thus significant for $N > 2000$ or $1-F < 0,0005$ and because $1-F_{50y} \ll 0,0005$ in the case of 10-min samples it is relevant for wind turbine load extrapolation. Therefore further mathematical and statistical research to understand this bias is recommended.

3.2.2. DEVELOPMENT OF $\varsigma_F(N)$ OVER N

In the graph 3.2 the dotted dashed line represents the Confidence Interval (CI) and has a value of $2\hat{\sigma}_F(N)$ there. Thus, by rule of thumb, 95% of all estimated extrapolations at a specific N fall within that boundary. In `GumbelUnc` it is possible to adjust other CI values values: 68,95% and 99,7%. When N is chosen large enough the confidence band is expected to narrow and this is confirmed by figure 3.4. However, in figure 3.4 at N=N still a small CI exists, while in a perfect case no uncertainty should be there. This shows that some inherent uncertainty permanently resides in within `GumbelUnc`.

3.2.3. MONTE CARLO AND FIT FOR $\varsigma_F(N)$ OVER N

As discussed in the 3.1, it is expected that if the sample size quadruple, the uncertainty deviation decreases by a factor of 2 according to equation (3.15). This is tested by fitting the general power curve as in equation (3.17):

$$f(x) = a * x^b + c \tag{3.17}$$

to the upper CI of exactly one $\varsigma_F(N)$. Of the estimated parameters $\hat{a}$, $\hat{b}$ and $\hat{c}$, the parameter $\hat{b}$ is of most interest. It should assume the value of $\approx -0,5$ for $\frac{1}{\sqrt{N}} = N^{-0,5}$ to verify the Monte Carlo principle. Since some initial upward bias has been observed, two cases are present:

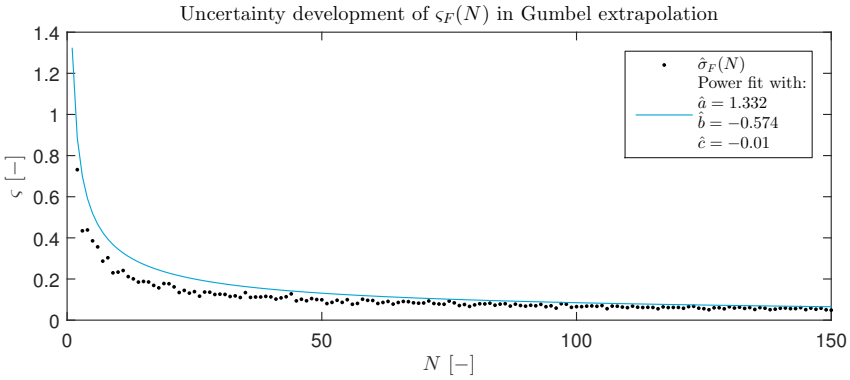
(a) $\varsigma_F(N)$
(b) $\epsilon_F(N)+\varsigma_F(N)$

translating equation (3.17) to:

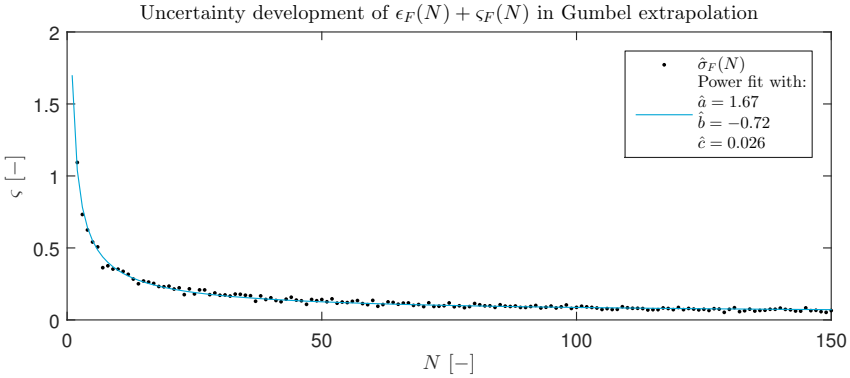
$$\varsigma_F(N) = \hat{a}N^{\hat{b}} + \hat{c} \tag{3.18a}$$

$$\epsilon_F(N) + \varsigma_F(N) = \hat{a}N^{\hat{b}} + \hat{c} \tag{3.18b}$$

In both cases the fits are statistically evaluated by R^2, adjusted-R^2 and Root Mean Square Error (RMSE) in figure 3.5:



(a) A power fit to $\varsigma_F(N)$; with R^2, adjusted R^2 and RMSE assuming: 0,9917, 0,9916 and 0,0076 respectively



(b) A power fit to $\epsilon_F(N)+\varsigma_F(N)$; with R^2, adjusted R^2 and RMSE assuming: 0,9781, 0,9778 and 0,0204 respectively

Figure 3.5: Evaluation of power fit curve of $\varsigma_F(N)$ to Monte Carlo principle, in order to evaluate the `GumbelUnc` process

It can be seen from the figures and the statistics that in both cases a good fit exist. Also both fits show for $\hat{b}$ a value close to $-0,5$. Case (b) has a slightly better fit than (a). However, it is noteworthy that after performing this test a dozen times, parameter $\hat{b}$ consequently approximates $-0,5$ more closely in case (b) than it does in case (a). Both cases systematically assume a higher value for $\hat{b}$. Therefore in case (a), b assumes systematically higher values than in case (b) which can be explained by the initial upward bias included in $\epsilon_F(N)$ at (b), A smaller negative root in case (b) is needed to adjust for the increased curve caused by $\epsilon_F(N)$. `GumbelUnc`'s $\varsigma_F(N)$ preforms an good fit and a good approximation of parameter $\hat{b}$ with slight deviations which is likely due to the initial bias. However, in the Monte Carlo test the value of $\hat{b}$ is found to be in a close enough proximity of -0,5 to be considered a viable calculation method.

3.2.4. EXTREME VALUE DISTRIBUTION FIT INSTEAD OF A NORMAL DISTRIBUTION FIT

Because of the of the initial upward bias discussed in section 3.2.1 another type of fit to obtain $E[\hat{x}_F(N)]$ and $\hat{\sigma}_F(N)$ than the Gaussian fit was researched. In essence the estimates by extrapolation, so sets of $\hat{x}_{F,j}(N)$, are tested to be of a Normal or s EV distributive nature. An EV distribution was chosen to be researched as alternative due to the shape of the upward bias, for it might yield a better result in $E[\hat{x}_F(N)]$ becoming closer to x_F. First the nature of the $\hat{x}_{F,j}(N)$ distribution is tested and secondly the result of $E[\hat{x}_F(N)]$ deviation to x_F is researched.

There are many statistical methods for evaluating and quantifying distribution fits. Here the Anderson-Darling test (A-D) (Anderson and Darling, 1952) in concurrence with p-value (Shorack and Wellner, 1986) is used. The A-D test gives us the Anderson-Darling coefficient (A_r^2) by:

$$A_r^2 = -r - \sum_{i=1}^{r} \frac{2i-1}{r} \left[\ln\left(F\left(X_i\right)\right) + \ln\left(1 - F\left(X_{r+1-i}\right)\right)\right] \tag{3.19}$$

The value for A_r^2 is compared to a predetermined significance level which depends on: the known parameters, the type of distribution and significance. Here a significance level of 5% is taken. If A_r^2 exceeds the significance level the null hypothesis is rejected, otherwise the null hypothesis is accepted. Some more explanation

about null hypothesis can be found in subsection 5.1.1. MATLAB® functions were used to calculated the exceedance. Exceedance of the significance level results in a `true`, meaning that the null hypothesis is to be rejected and thereby indicating a correlation with the distribution. The two null hypothesis that are tested are:

1. Is the extrapolated estimates data set normally distributed?
2. Is the extrapolated estimates data set EV wise distributed?

Furthermore the goodness of the fit is quantified by the p-value being a scalar parameter in range [0 1]. The higher the p-value the better the fit for the given distribution. As researched by Anderson and Darling (1952) and Shorack and Wellner (1986) the A-D test needs for Normal distribution a minimum data set: $r > 5$ and for a EV distribution: $r > 8$. Here r is taken 10.000 to ensure a large enough data set. As shown in section 3.2.1 due to the bias a distinction can be made between the initial phase and long therm phase. Therefore the initial: $N = 1,2,\ldots,20$ and long term: $N = 1024, 1536, 2048$ extrapolated estimates are evaluated. Note that for $N = 1,2,\ldots,20$, $F = \frac{N}{20+1}$ and for $N = 1024, 1536, 2048$, $F = \frac{N}{2048+1}$ like the continues linear series in `GumbelUnc`. Quantification of the goodness of distribution fits by the A-D test and the p-value is summarised in table 3.1. Comparison for which distribution fits best is summarised in the last column: Fit result.

N	$A_r^2 > 5\%$		p-value		Fit result
	Normal	Extreme Value	Normal	Extreme Value	
2	`false`	`true`	0,2613	0,0341	Normal
3	`false`	`true`	0,2347	0,0011	Normal
4	`false`	`false`	0,5405	0,6034	Extreme Value
5	`false`	`true`	0,1779	0,0009	Normal
6	`false`	`false`	0,6449	0,1366	Normal
7	`false`	`false`	0,9320	0,3160	Normal
8	`false`	`false`	0,8428	0,8713	Extreme Value
9	`false`	`false`	0,5933	0,5022	Normal
10	`false`	`true`	0,4101	0,0093	Normal
11	`false`	`false`	0,1119	0,0970	Normal
12	`false`	`false`	0,6251	0,9591	Normal
13	`false`	`true`	0,1543	0,0011	Normal
14	`false`	`false`	0,7060	0,4196	Normal
15	`false`	`false`	0,9300	0,1039	Normal
16	`false`	`false`	0,4253	0,0743	Normal
17	`false`	`false`	0,3407	0,0691	Normal
18	`false`	`true`	0,3823	0,0203	Normal
19	`false`	`false`	0,4170	0,3455	Normal
20	`false`	`false`	0,5553	0,2202	Normal
1024	`false`	`true`	0,0750	0,0005	Normal
1536	`false`	`true`	0,3336	0,0005	Normal
2048	`false`	`true`	0,8296	0,0005	Normal

Table 3.1: Results for $A_r^2 > 5\%$ and p-value at $N = 1,2,\ldots,20$ extrapolation estimates for $F = \frac{N}{20+1}$ and $N = 1024, 1536, 2048$ extrapolation estimates for $F = \frac{N}{2048+1}$. For all $r = 10.000$

Although the two occasions where the EV distribution fits slightly better than a Normal distribution, it can be concluded that the data is predominately normally distributed. The two occasions are believed to be caused by the heavy uncertainty being present for low Ns. This is also confirmed by the very low p-value at very high Ns. So from this chapter it can be concluded that the data sets of extrapolated estimates are generally normally distributed.

3.2.5. GUMBEL PLOTTING POSITIONS

In estimating an extreme value trough extrapolating the EV, the plotting position used to assign a probability to the sorted $x_{i,j}$ has influence on the curve of the fitted extreme value distribution. Extrapolating this could in turn have a magnifying effect on the uncertainty. For this reason estimates of different fitted plotting

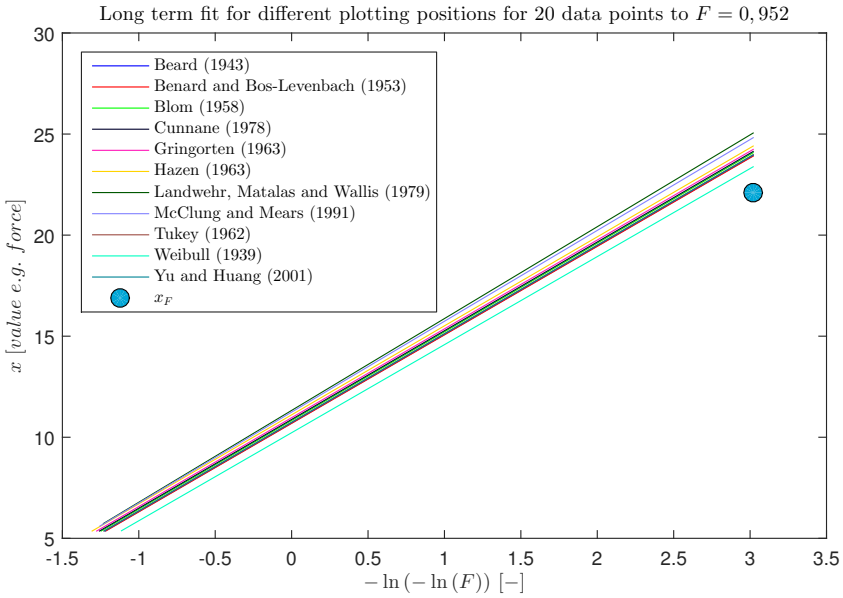
positions is investigated. Over the years many plotting positions have been suggested mainly in the form of equation (3.20). Equations proposed in literature are summarised in table 3.2.

$$F = \frac{j-a}{n+1-2a} \tag{3.20}$$

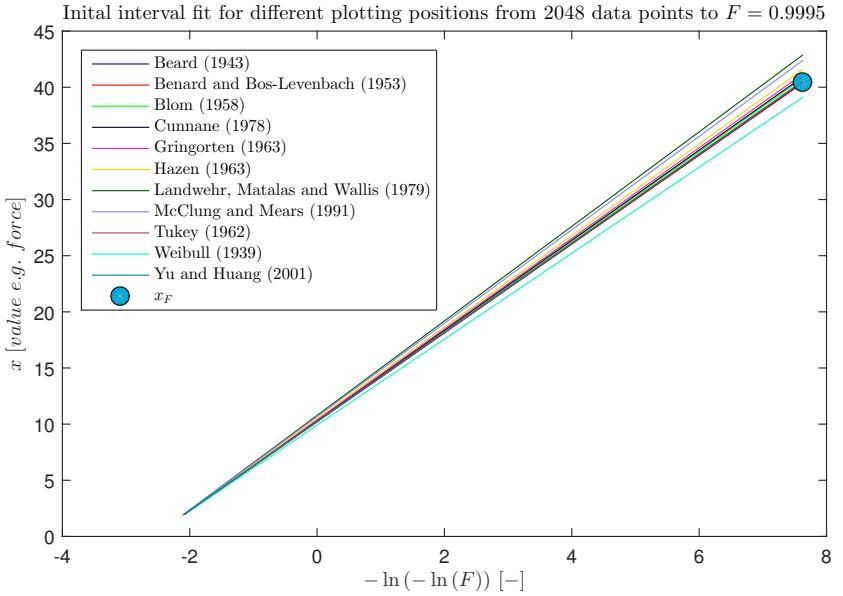
Author	F	a
Beard (1943)	$(j-0,31)/(n+0,38)$	0,31
Benard and Bos-Levenbach (1953)	$(j-0,3)/(n+0,4)$	0,3
Blom (1958)	$(j-\frac{3}{8}/(n+\frac{1}{4})$	$\frac{3}{8}$
Garcia (1979)	$(j-0,4)/(n+0,2)$	0,4
Gringorten (1963)	$(j-0,44)/(n+0,12)$	0,44
Hazen (1914)	$(j-\frac{1}{2})/n$	$\frac{1}{2}$
Landwehr *et al.* (1979)	$(j-0,35)/n$	n/a
McClung and Mears (1991)	$(j-0,4)/n$	n/a
Tukey (1962)	$(3j-1)/(3n+1)$	$\frac{1}{3}$
Weibull (1939)	$j/(n+1)$	0
Yu and Huang (2001)	$(j-0,326)/(n+0,348$	n/a

Table 3.2: Plotting Positions from literature

It is interesting to look at how much the fits differ with different plotting positions. Since an initial and long term phase was previously defined, both cases will be examined. The fits for the initial and long term have the following number of discrete Gumbel sampled samples to fit from: N= 20 and N= 2048. In this case no extrapolation takes place in order to determine the minimum difference between plotting positions. Further research could investigate the influence of extrapolation by using discrete Gumbel samples and actual wind turbine load data set. Note that the plotting positions have a different definition of the probability of exceedance F. To be able to compare all plotting positions equally, the probability F is taken as the Weibull for both terms: So all plotting positions are fitted with the same amount of N and plotted toward $\frac{20}{20+1}$ and $\frac{2048}{2048+1}$ for short and long term respectively. The deviation of the x_F value is due to the inherent uncertainty in sampling. The results are visualised in figure 3.6.



(a) A Gumbel fit for different plotting positions at $N = 20$



(b) A Gumbel fit for different plotting positions at $N = 2048$

Figure 3.6: Gumbel fits for different plotting positions for the short and long term

From the above graphs it can be concluded that the fit of the plotting positions results in plot position estimates residing within 7% proximity of each other. In this case the difference of the fit for plotting positions is given. The influence of extrapolation on different plotting positions of wind turbine load data can produce leads for an optimal plotting position that yields the most accurate wind turbine load estimate.

3.3. CONCLUSION

Extreme value estimate calculation in this chapter was based on drawing samples form a Gumbel distribution to understand factors that influence the uncertainty other than influences from wind turbine phenomena. Calculation of estimate uncertainty of this distribution is executed by preforming extrapolation on a set of discrete generated and sorted samples followed by the calculation of the distribution around the true load value. Because the wind turbine load samples originate from a purely mathematical construct, this research

construct is not influenced by natural variations enabling an evaluation of the estimate calculation method in itself.

Research on this mathematical construct has yielded five insights. The first remarkable insight is that the mean of the extreme extrapolated sample distribution does not average out onto the mean of the original sample distribution. It shows a clear upward bias especially at low number of samples. Upward biases reach from 30% at low sample sizes to -1% at bigger sample sizes. An explanation of the upward bias phenomena is given by the fact that if samples are sorted, every following sample is equal or higher than the next one. A model for predicting this behaviour has been developed by White (1969). Secondly, inherent uncertainty of the calculation technique exits. Still 1-2% uncertainty was found in extrapolating to an equal amount of samples drawn as the probability of exceedance to which the extrapolation takes place. Thirdly, by means of the confidence interval, the method of sampling was tested against the Monte Carlo principle. The parameter $\hat{b}$ and was found close to -0,5 which indicates a correct decrease of uncertainty over sample size.

Fourthly, ways of minimising the estimate distribution bias were explored by comparing the extreme value distribution fit with an Gaussian distribution fit. This was done by means of the A-D and p-value tests. For 22 different sample sets two small sample sizes indicated that an extreme value fit was more suitable. A possible explanation for this is that the less sample values are available, the more the distribution resembles an own extreme value distribution because it is sampled from the same source. Though a normal distribution gives the overall best fit. Lastly, the effect of fitting different plotting positions was investigated. Eleven definitions of plotting positions where executed for small (20) and larger (2048) sample sizes. All plotting positions fitted within 7% of each other. The influence of extrapolation on different plotting positions of wind turbine load data can produce leads for an optimal plotting position that yields the most accurate wind turbine load estimate.

Many aspects of the cause of the perceived upward bias is not yet clear and must be understood better because it stimulates over-dimensionalisation of loads and ultimately needlessly high wind turbine costs. So further mathematical and statistical research to understand this bias is recommended.

4

THE SANDIA DATA SET

Evaluating calculation methods for wind turbine design standards is paramount in the development of wind turbines. The IEC 61400-1 DLC 1.1 is a notoriously uncertain calculation method for it relies heavily on extrapolation. One of the initiatives to understand this load case is the release of a data set containing 96 years of wind turbine operation simulation by SANDIA [1] (Barone *et al.*, 2012a). This data set has the potential to map the uncertainty in DLC 1.1 depending on the number of load simulations and may increase the understanding of the calculations method in this load case. In this chapter an attempt is made to quantify the uncertainty in 50-year extreme wind turbine load estimates thereby answering sub question four.

Before the data can be used to analyse the load case, the data set itself will be presented. This chapter introduces the SANDIA data set where it is explained and elaborated into detail. The first two paragraphs of this chapter contains an explanation and basic plots of the data set. In in the final paragraph estimate uncertainty is quantified and an attempt is made to identify wind turbine specific factors that play a role in the uncertainty calculation results. Lastly, existing ways for mitigating uncertainty and a new method to mitigate uncertainty of wind turbine estimates is suggested, thereby answering sub question five.

4.1. SANDIA CASE STUDY EXPLAINED

Normally wind turbine designers have a limited amount of computational resources which are used to simulate wind turbine behaviour. These values are then allowed to be extrapolated to obtain the 50-year load value. However, the SANDIA has removed the computational restriction by using a high performance-computer to calculate the loads over a simulated time period of 96 years. This exercise has produced over 5 million aero-elastic simulations resulting in 10-min extreme loads as well as fatigue cycles on various turbine components. The aero-elastic on-shore 5 MW NREL reference turbine (Jonkman *et al.*, 2009) is used in the simulation. The following specifications apply:

[1] Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

- Blades: 3
- Rotor: Upwind
- Rotor diameter: 126 m
- Hub height: 90 m
- Turbine control:
  - Variable speed operation in region 2
  - Variable collective blade pitch in region 3
  - No active yaw control
- Cut-in ($U(h)_{ci}$): 3 m/s
- Rated (U_r): 11,4 m/s
- Cut-out ($U(h)_{co}$): 25 m/s

Furthermore the wind turbine model assumes that the commanded yaw position is held constant at zero relative to the normal wind direction, while allowing for small yaw deflections subject to flexibility and damping in the yaw drive. This wind turbine was simulated using FAST version 7.00.01.a-bjj. The turbulent wind fields are generated by the NREL TurbSim code version 1.50. The turbulence intensity and average wind speed in IEC 61400-1 are chosen to be class I-B, corresponding with 10 m/s average wind speed at hub height[2]. From this a Rayleigh distribution was used to sample 10-min wind speeds. Wind speeds that fell out of the cut-in and cut-out wind speed range where automatically resampled until they fell within this threshold. This has probably caused the noise in the data set. This noise has been filtered out using MATLAB ® reducing the data set to 5.020.189 values, representing 95,45 years of loads and deflections. Further information regarding seed management is beyond the scope of this thesis and can be found in the paper of Barone *et al.* (2012a).

Eleven different types of stresses and deflections on the wind turbine where calculated using FAST. These are listed below with their designator in the data set. All stresses and deflections have their corresponding wind speeds delivered with them, making it easy to perform vectorised calculations and visualisations of the data set.

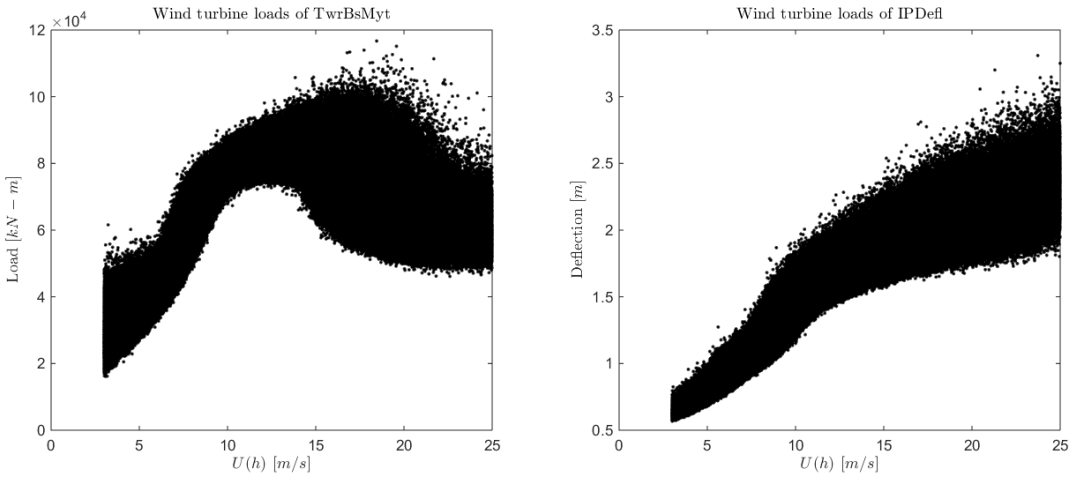
- OopDefl Maximum out-of-plane blade tip deflection, taken over all 3 blades (m)
- IPDefl Maximum in-plane blade tip deflection, taken over all 3 blades (m)
- RootMOop Maximum blade root out-of-plane bending moment, taken over all 3 blades (kN-m)
- RootMIP Maximum blade root in-plane bending moment, taken over all 3 blades (kN-m)
- RootMFlp Maximum blade root flap-wise bending moment, taken over all 3 blades (kN-m)
- RootMEdg Maximum blade root edge-wise bending moment, taken over all 3 blades (kN-m)
- LSSTipMya Maximum rotating low-speed shaft bending moment about the y_a-axis (*see* FAST user manual) (kN-m)
- LSSTipMza Maximum rotating low-speed shaft bending moment about the z_a-axis (*see* FAST user manual) (kN-m)
- TwrBsMxt Maximum tower base side-to-side bending moment (kN-m)
- TwrBsMyt Maximum tower base fore-aft bending moment (kN-m)
- TwrBsMzt Maximum tower base yaw moment (kN-m)

The TwrBsMyt moment shall serve as the example throughout this chapter for two reasons. Firstly, as discussed in chapter 1 the tower forms a huge part of the wind turbine cost. According to figure 1.4 16% of the cost of the total wind turbine is made up by the tower. Secondly, weight, based on weight which mainly depends on load calculations, determines for 16% the cost of the foundation and installation especially in the off-shore wind turbine industry.

4.2. PLOTTING THE SANDIA DATA SET

In this section an attempt is made to plot the SANDIA data in a practical way that enables new insights in the data. First, a simple plot of the load versus the wind speed is made. Then, a begin is made with plotting probability density functions of the loads. And finally, the probability density function is plotted per wind speed bin to get further insight into wind turbine loading (*see figure* 4.1).

[2]The turbulence class and average wind speed have been revised from B-II and 8,5 m/s to B-I and 10 m/s in later on-line publications: http://energy.sandia.gov/energy/renewable-energy/wind-power/wind-resources/wind-software-downloads/sandia-wind-turbine-loads-database/



(a) TwrBsMyt SANDIA data

(b) IPDefl SANDIA data

Figure 4.1: Wind speed verses wind turbine load extremes derived van SANDIA data set

The remainder of loads and deflections can be found in Appendix D. Figure 4.1 shows that different areas and moments on the wind turbine can have a different behaviour. To get more insight in the effect of this behaviour in wind turbine design, a Cumulative Density Function (CDF) must be constructed. In a CDF load data is visualised using probabilities of exeedance in 10-min time intervals. These loads are sorted from the lowest to the highest value. The lowest value is exceeded every 10-min and the highest value is exceeded ever 96 years in this data set. The probability that is associated with these numbers is called the probability of non-exceedance (F) and is calculated trough:

$$F = \frac{i}{(1 + \text{number of 10-min intervals in the time period})}$$

In this case, a daily return period corresponds with a probability of $1 - \frac{1}{24*6} = 0,993$ and a 50-year return period with $1 - \frac{1}{50*356,25*24*6} = 0,9999996$. If these probabilities are plotted against the load values a Cumulative Distribution Function (CDF) (for example a Gumbel distribution) appears with a typical S-shape. In figure 4.2 an example has been drawn to illustrate that the 50-year return period lies well far in the tail of the S.

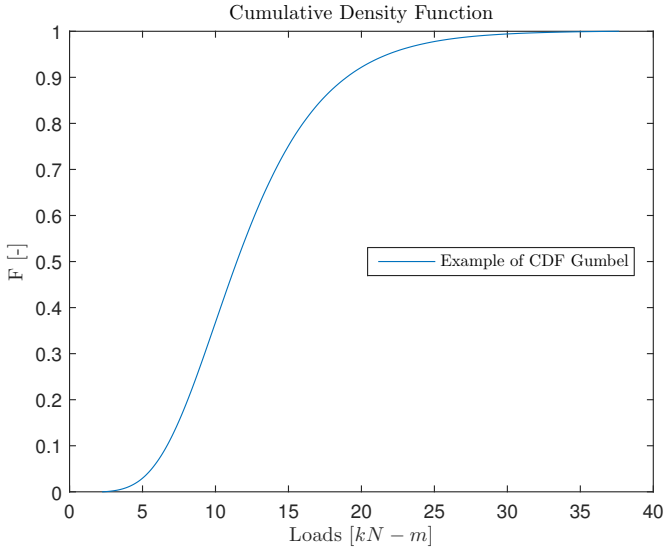
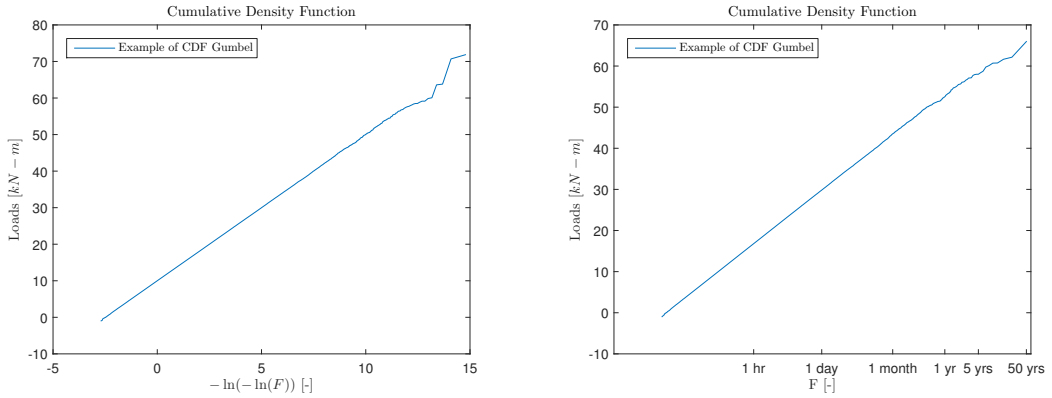


Figure 4.2: Example of Cumulative Density Function

Since the 50-year value and the tail is of most interest a log-log plot (on Gumbel paper) is preferred over the

classical S-shape. The Gumbel chart illustrates the same data in a different way. The probability of occurrence is then situated on the x-axis and the loads on the y-axis. In addition the probability of occurrence is scaled to a $-\ln(-\ln(F))$ scale, turning the S-shape to a linear graph, facilitating easier analysis. Statistical fundamentals and calculations that prove these statements are discussed in chapter 3. In figure 4.3(a) an example is given of how 50-year load data is visualised.

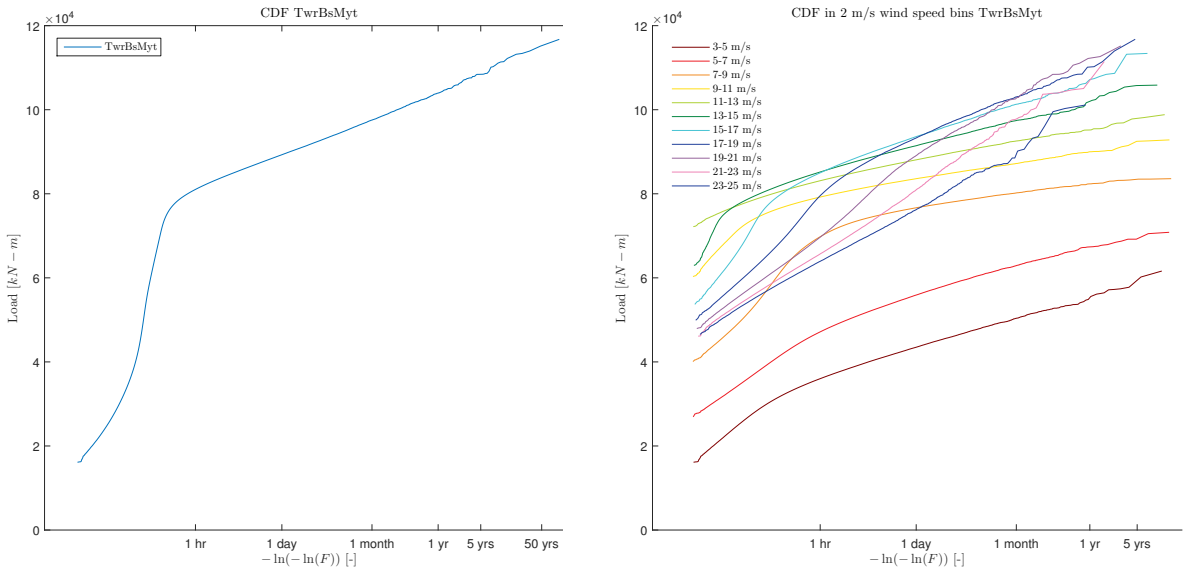


(a) Example cumulative probability density with $-\ln(-\ln(F))$ and swapped axis

(b) Example cumulative probability density translated to time scale

Figure 4.3: Examples of graphical presentation probability density function curve

Figure 4.3(a) and (b) are equal, except for that the probability of occurrence in figure 4.3(b) is translated into human time scales. This is done in order to make interpretation of the figures more intuitive. Note that the data is now plotted on a $-\ln(-\ln(\ldots))$ scale, causing the data to concentrate at the low probabilities of occurrences and dilute towards the high probability of occurrences. This can also be seen from the smoothness of the graph in the initial fragment and the wrinkles at the last segment of the graph. In figure 4.4 the TwrBsMyt moment graph is plotted in the exact same way. The other SANDIA loads and deflections can be obtained in appendix E.



(a) CDF of TwrBsMyt SANDIA data

(b) CDFs of TwrBsMyt in 2 m/s wind speed bins

Figure 4.4: Total and binned extreme load distribution

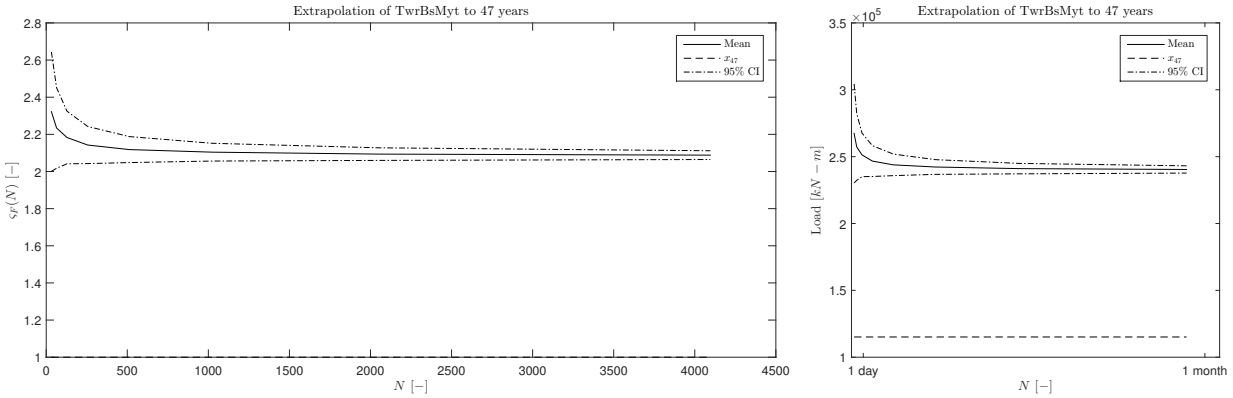
Remarkably, many of the extreme load distributions show a curved graph instead of the expected linear graph. In the case of OoPDefl, IPDefl, RootMOoP, RootMFlp, RootMEdg, TwrBsMyt and TwrBsMzt a clear and peculiar *'knee'* can been observed in the graph. This peculiar shape seems to indicate a dynamic situation for wind turbine load and deflections distributions. The involvement of intertwined distributions from other processes are generally at the root of structural extreme value distribution deviations. In order to understand which other processes are causing this behaviour it can be insightful to observe the building blocks of the *knee* individually. The SANDIA data set contains wind speed information attached to the loads and deflections, which enables an analysis of load behaviour at different wind speeds. The data of a particular load or deflection is divided into eleven 2 m/s wind speed bins. The binned data is plotted in one graph and presented next to the general load or deflection CDF. The total data set is built via the earlier stated Rayleigh distribution. Therefore not all wind speed bins have equal amounts of loads which causes some wind speed binned CDFs to stretch further than others.

4.3. SPECIFICS OF THE SANDIA DATA SET

The results presented in section 4.2 are analysed and discussed in this section. It is particular interesting to understand the behaviour and effect that the *knee* has on calculating the 50-year ultimate stochastic wind turbine load.

In chapter 2 and 3 the calculation method for ultimate stochastic wind turbine design loads is covered. The estimate distribution that is fitted to the samples assumes a certain behaviour, normally a linear graph. However, as shown in the figures 4.4 in section 4.2 entirely different curves with a *knee* like curve can be observed. The use of extrapolation on the bases of models that do not fit the data well can significantly influence the outcome leading to a large uncertainty in the 50-year loads.

To illustrate the enormous amount of inaccuracy in the calculation of the 50-year loads or deflections the estimates are calculated for different simulation sample subsets: $N = 32, 64, 128, \ldots, 4096$. Each subset has been sampled 100 times (r), and from that set a normal average and standard deviation were taken. Again, only the TwrBsMyt is presented here in figure 4.5 and the rest of the loads and deflections can be found in Appendix F. Chapter 3.1 covers the extrapolation process into detail. The emphasis on the graph below is on the enormous deviation that the extrapolation outcome has relative to the actual 50-year value. Notice that the graph works with a 47-year load/deflection. This is because the 47-year load/deflection is an actual point in the data set, making the calculation easier and more accurate while still close enough to 50-year to illustrate the point of large uncertainty using basic calculation techniques.



(a) Normalised uncertainty in TwrBsMyt extrapolation

(b) Uncertainty per number of simulations in TwrBsMyt

Figure 4.5: Uncertainty in estimates from basic sample data

The most striking feature is that all graphs show an elevated average as opposed to the true 47-year extreme value. The explanation for this over-dimensionalization is two fold:

1. The linear fit will tend to overshoot because the slope of the initial *knee* fragment will cause an upward

linear fit slope.

2. Secondly, due to the concentrated data at the initial fragment of the graph, the much steeper slope before the *knee* gets a significant larger preference for the model to fit on. In other words, the steep fragment weighs much larger in the fit than the correct final slope of the graph. In some cases as much as 90% of the data lies before the *knee* curve.

Another noticeable feature occurring in all figures is an initial upward bias at low probabilities of occurrences. This phenomenon is complicated as found out in chapter 3.2.1 and cannot be fully explained. In practice often small amounts of load simulation are executed due to a limitation of the computational resources and simulations are regularly done for 300 minutes to 3 days, yielding 30 to 432 10-min extreme values. From the graphs in figure 4.5 it can be observed that in the best case of a 3 day load simulation the mean still deviates twice the value of the actual 47-year load. Taking this finding into consideration and looking at the IEC 61400-1 standard in appendix F (IEC, 2005), it appears that the recommended minimum of 300 minutes of simulation (equal to thirty 10-min block maxima) exposes calculations to a dangerous bias of up to two times the mean and large variations up to 40%.

4.3.1. THE SANDIA DATA CDF CURVE

The findings in 4.3 show that the *knee* in the CDF curve has a significant effect in wind turbine load extrapolation. Looking at other disciplines using the same techniques can provide valuable insights. Load extrapolation is otherwise done in civil engineering. But when looked at load extrapolation of CDFs in for example bridge deflections, a much more linear CDF can be observed in figure 4.6.

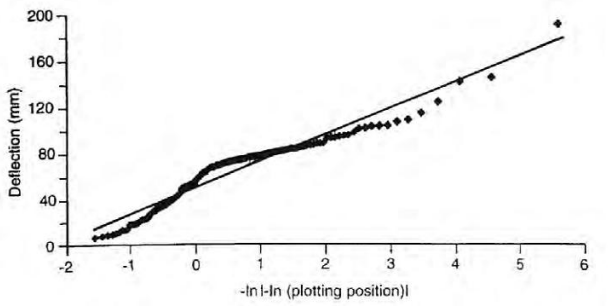


Figure 4.6: Gumbel extrapolation for the Foyle bridge, Reproduced from O'Brien *et al.* (1995)

Although figure 4.6 does not show a perfect linear curve, the shape is much less curved than the wind turbine loads plotted in figure 4.4. Also in other extreme value disciplines the general trend is a linear one. The finding that in other disciplines the method of extrapolation works, confirms the suspicion that another process is influencing the wind turbine load distribution. The purpose of this chapter is not to discover the mathematical fundamentals of the wind turbine load CDF. Managing the data in such a way that it yields an acceptable outcome together with the consequences of the accompanied uncertainty is rather the goal. The method to obtain acceptable outcomes for uncertainty analysis is covered in section 4.3.2. The uncertainty in that method of extrapolation and the consequences thereof are covered in chapter 5.

A first attempt is made below to investigate the origins of the peculiar CDF shape of wind turbine load simulations. Apparently another process plays a significant role in the creation of extreme wind turbine loading. When figure 4.5 is analysed it can be seen that in the load types where the curve of the *knee* is particular strong, loads from multiple wind speed bins run trough each other. This can be seen in for example TwrBsMyt. It might be expected that increasing wind speeds are accompanied with increasing loads. However, in the case of TwrBsMyt the wind speed bins 9-11, 11-13, and 13-15 m/s all start with a higher value than the extremest wind speed bin of 23-25 m/s. This indicates that there are more high loads produced at these wind speed intervals. Looking at general wind turbine loading it becomes evident that around rated wind speed the largest forces on the rotor are produced. This is because the turbine is then still pitched in such a way to produce maximum power and the rotor operates at maximum rotational speed. Equation 4.1 and 4.2 describe the thrust and power against wind speed.

Thrust: $$T \equiv \frac{1}{2}\rho U(h)^2 \pi R^2 C_t \tag{4.1}$$

Power $$P \equiv \frac{1}{2}\rho U(h)^3 \pi R^2 C_p \tag{4.2}$$

What stands out the formulas is the difference in wind speed dependence by one power. When these curves are plotted, a sharp peak of thrust loading can be observed just before and after U_r caused by the difference in wind speed dependence. To cope with these high moment of forces on the wind turbine the wind turbine pitch regulating system sets the pitch somewhat earlier to less rotational speed. This releases the wind turbine of the highest forces and is also called 'peak shaving'. This principle is depicted in figure 4.7.

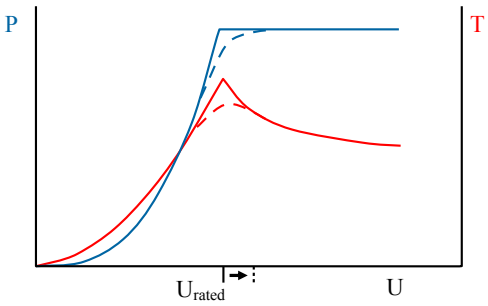


Figure 4.7: Peak shaving of thrust

When the peculiar wind speed binned loads from 9 trough 15 m/s are analysed, it can be seen that all bins around the U_r of 11,4 m/s and with the principles of 4.7 in mind, it might be so that the pitch regulating system has a large influence on load production and could be on of the causes that is responsible for the *knee* in the wind turbine CDF.

4.3.2. MANAGING THE SANDIA DATA CDF CURVE

There are generally two distinct ways by which the prediction of ultimate loads can be optimised:

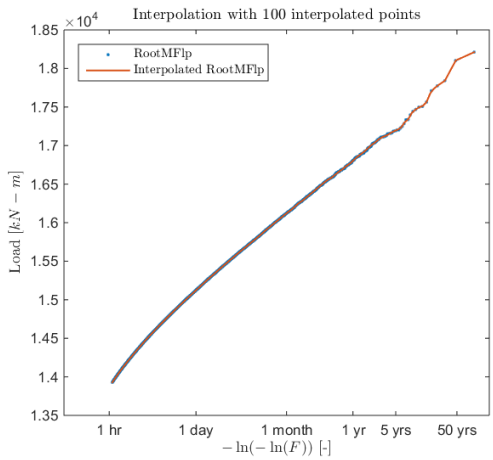
- Devise models that can allow for inclusion of the *knee*.
- Manipulate the data in such a way that current models can suffice and deliver accurate outcomes.

The second manner is often used in practice by companies, while the first is mostly the domain of research institutions and universities. This thesis focuses on the effect of extrapolation methods used by wind turbine designers and therefore a practical manner will be devised to calculate 50-year ultimate loads. The following practical method is used to calculate the uncertainty in 50-year return estimates out of a set of 96 years of wind turbine load simulations:

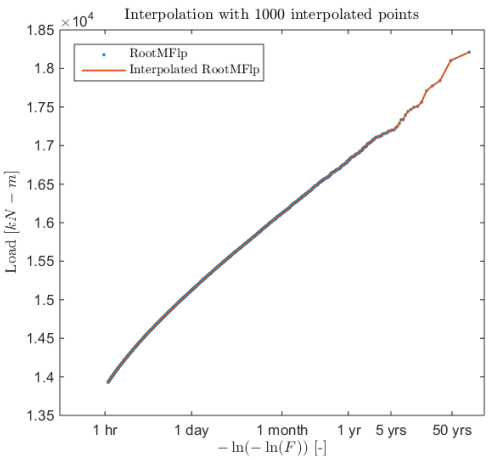
1. Particular subsets of samples (N) are drawn from the SANDIA data set called ranks (*see figure* 3.8). The ranks that where calculated are: $N = 32, 64, 128, \ldots, 4096$.

2. The first 85% of the sample data is cut away to eliminate the *knee* from the fit and ensure an approximate linear graph in the $-\ln(-\ln(F))$ plane.

3. To account for the influence that the ascending data density has on the fit, the data is interpolated. A representation of 100 equally spaced points is used to represent the final 15% of sample data. More interpolation points do not add to the accuracy of the 50-year load extrapolation. From figure 4.8 it can be seen that the interpolated graph hardly differs from the original 15% graph. Furthermore it can be seen from 4.8(b) the graph is not represented significantly better when more interpolated data points are used.



(a) Interpolation using 50 interpolation points (RootMFlp)



(b) Interpolation using 50 interpolation points (RootMFlp)

Figure 4.8: Effect of interpolation on tail of CDF, here RootMFlp is used as example

4. There are roughly 4 accepted distributions that fit the tail of the data. These are the: Generalised Extreme Value (GEV), Gumbel, 3 parameter Weibull, and Lognormal distribution:

Generalised Extreme Value:

$$F(x) = \begin{cases} e^{\left(-\left(1-\frac{\beta(x-\gamma)}{\alpha}\right)^{1/\beta}\right)} & \text{, for } \beta \neq 0 \\ e^{\left(-e^{\left(-\frac{x-\gamma}{\alpha}\right)}\right)} & \text{, for } \beta = 0 \end{cases} \tag{4.3}$$

Gumbel:

$$F(x) = e^{-e^{-\frac{(x-\mu)}{\beta}}} \tag{4.4}$$

3 parameter Weibull:

$$F(x) = 1 - e^{\left(-\left(\frac{x-\gamma}{\alpha}\right)^{\beta}\right)} \tag{4.5}$$

Lognormal:

$$F(x) = 0.5 + 0.5 erf\left(\frac{\ln(x) - \mu}{\sigma}\right) \tag{4.6}$$

$$erf(x) = \frac{2}{\sqrt{\pi}} \int_0^x \mathrm{e}^{-t^2}\, \mathrm{d}t$$

From literature (Genz *et al.* (2006) and Peeringa (2009)) it can be concluded that the best fitting method is the 3 parameter Weibull distribution. By means of a Least-square curve fit (LSQfit) the 3 parameter Weibull distribution is represented and a fit is made to the final 15% interpolated data. The LSQfit must be used in order to account for the position of the curve, this is not possible with a fitting method that relies on maximum likelihood estimation.

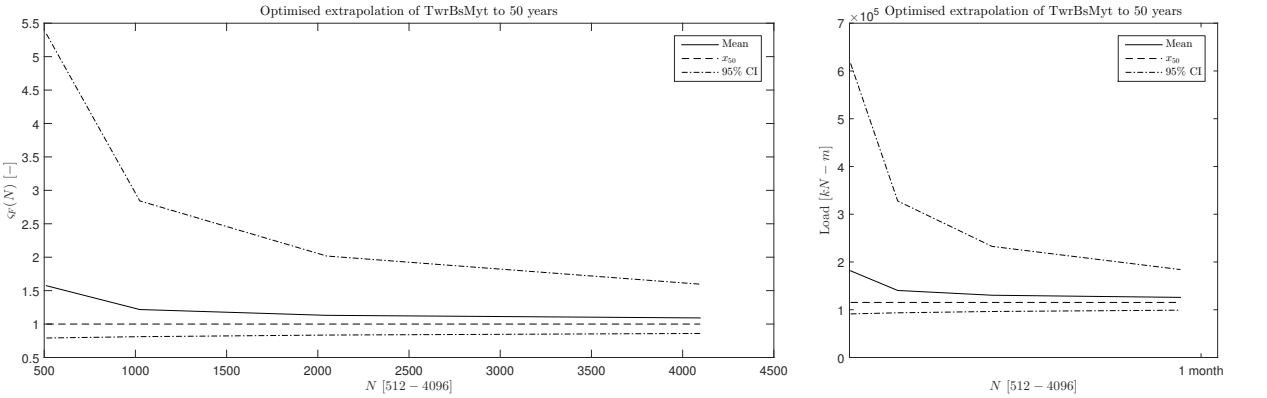
5. The 50-year value is extracted from the fit. This process is repeated for a number of sets (*see figure* 3.8). A big enough set of 50-year load estimates is used in order to extract a reliable mean and confidence bound of the estimates. Here a 1000 sets are used. In contrast to what was found in section 3.2.4 the distribution of the estimates does not adhere to a Gaussian fit. This is further investigated in section 4.3.3. A solution for representing the 95% confidence interval ($p_{95\%}$) is found by selecting the 25$^{\text{th}}$ and 975$^{\text{th}}$ value of the sorted set. The calculation of the mean and uncertainty in the data can be mathematically described in the following way:

$$\textit{Confidence Interval:} \qquad p_{95\%} = \hat{x}_{F,1}(N) < \cdots < \hat{x}_{F,r} \rightarrow \begin{cases} \hat{x}_{F,3}(N) & = \text{lower } 2.5\% \\ \hat{x}_{F,98}(N) & = \text{upper } 97.5\% \end{cases} \tag{4.7}$$

$$\textit{Mean} \qquad E[\hat{x}_F(N)] = \frac{1}{r}\sum_{j=1}^{r} \hat{x}_{F,j} \tag{4.8}$$

6. From the set of 50-year values a estimated mean is extracted as well as a normalised standard deviation. These are plotted against the number of sampled 10-min load extremes.

Using the above optimisation method and tools 50-year load estimates are calculated. Below only TwrBsMyt is plotted in figure 4.9, while the graphs of other loads and deflections can be found in Appendix G.



(a) Optimised normalised uncertainty in TwrBsMyt extrapolation

(b) Optimised uncertainty per number of simulations in TwrBsMyt

Figure 4.9: Optimised uncertainty estimation of extrapolation

For display purposes, the graphs of different load cases are plotted from subset ranks of 256 or 512 toward 4096. Noteworthy is that all graphs still show an upward bias of the mean estimates from the actual 50-year extreme load value (x_{50y}). This implicates that the calculation of the 50-year value is structurally over-dimensionalised causing a constant elevated 50-year return value. It must be emphasised that this inherent over-dimensionalisation depends on the number of simulations (rank) and type of load or deflection. The earlier stated 300 minuets of simulation cannot be insightfully presented due to the immense uncertainty. Three and a half day worth of wind turbine simulations are executed which accounts for roughly 500 10-min load extremes. For this duration at IPDefl, LSSTipMya, LSSTipMza and TwrBsMzt still a $\sim$ 55% upward bias of the estimated mean can be observed. So for this simulation length and these load types, this translates into an over-estimation of material use needed based on $\sim$ 55% higher loads. Because of this significant impact this phenomenon is further investigated in chapter 5.

Furthermore different decreasing 95% confidence bounds over the number of simulations can be observed. Some loads and deflections lose there uncertainty relatively fast like RootMOop, RootMFlp and TwrBsMxt. However, this still indicates that for these loads 20480 minutes of simulations (fourteen days) need to be executed before reasonable amounts of variation are avoided. For most other load and deflection cases the 95% confidence interval still decreases significantly after 20480 minutes of simulation. Therefore by simulating for $\sim$ the 300-min standard minimum, a excessive probability exists that the load is estimated too high. This causes a valid design to be discarded needlessly. This phenomenon is better known as a false negative. As clearly shown in chapter 2, load calculation can have a significant impact on the wind turbine material and ultimate wind turbine design cost. The high probability of *false negative*[3] outcomes in 50-year load calculations can therefore seriously influence wind turbine cost. This justifies more research into the consequences of the findings in this chapter and forms the bases of chapter 5.

[3]The principle of false negative is extensively researched in section 5.1.1

4.3.3. DISTRIBUTION OF 50-YEAR LOAD ESTIMATES

From the previous research done with the SANDIA data set it became clear that the 50-year wind turbine load estimates shows no clear normal distribution, even though a Gaussian distribution was fitted to the extrapolated set. Figure 4.10 below shows five box plots summarised in one figure of the 50-year load estimates at different amounts of 10-min extreme load value samples (N). The red dots are outliers defined by 1,5 points[4] and for reasons of visualisation the box plot is only made for $N = 256, 512, \ldots, 4096$. The set consists of a 1000 50-year load estimates. If the data is normally distributed, the four areas between the five markers in the box plot will be equally spaced. Note that the distribution of the 50-year load estimates are calculated according to the optimisation strategy described in section 4.3.2 for the TwrBsMyt moment.

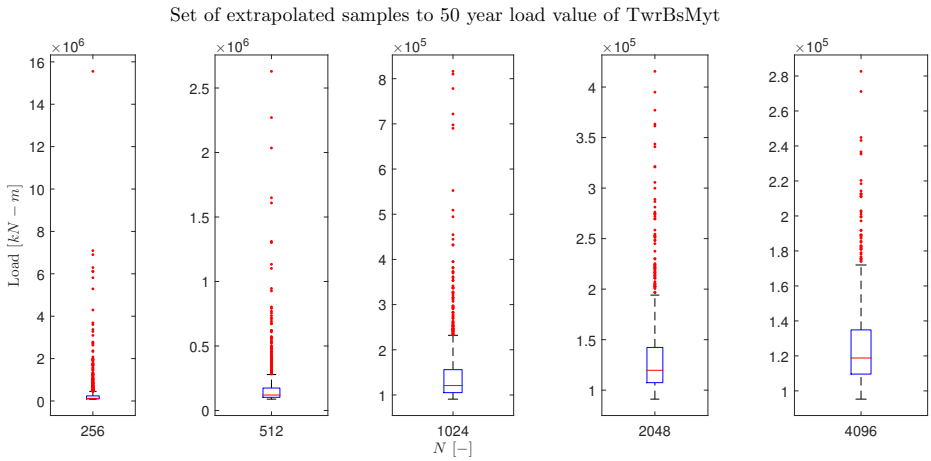
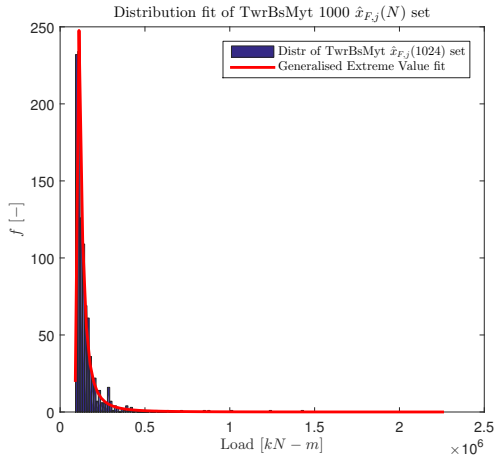
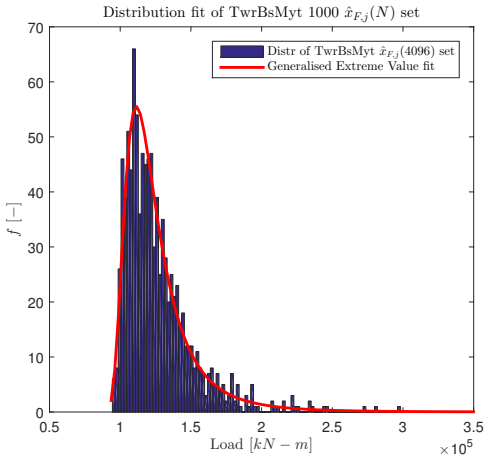


Figure 4.10: Box plots with $N = 256, 512, \ldots, 4096$ of a set of 50-year load estimates in TwrBsMyt

It can be seen that none of the box plots satisfy equally spaced regions. Instead a concentrated amount of data in the low regions can be observed with a long tail with many outliers. This indicates an other distribution rather than a normal distribution. Therefore the results are further investigated by plotting the data as probability density histogram. The histogram is drawn in figure 4.11 for the 10-min extreme sample sizes of $N = 1024$ and 4096 in the 50-year load estimates of TwrBsMyt.



(a) GEV fit for extrapolated extreme values fitted from N = 1024 TwrBsMyt values



(b) GEV fit for extrapolated extreme values fitted from N = 4096 TwrBsMyt values

Figure 4.11: Histogram and GEV fit of estimate distribution

[4] $w = 1,5$ points and outliers are drawn if they are larger than $q3 + w(q3–q1)$ or smaller than $q1–w(q3–q1)$, where $q1$ and $q3$ are the 25^{th} and 75^{th} percentiles, respectively. 1,5 points corresponds to approximately $\sim 2,7\sigma$ and 99.3 coverage if the data is normally distributed.

The Normal, Generalised Extreme Value (GEV), Extreme Value, Weibull, Rayleigh and Lognormal distributions where fitted to the 50-year load set. The distribution of $\hat{x}_{F,j}(N)$ values in figure 4.10 changes as N progresses. The distributions at low sample sizes seem to attain extreme shapes and decreases to an interpretable distribution from $N = 512 - 1054$ onwards. For this reason the highest sample set available ($N = 4096$) was chosen to evaluate the best fit. It was found that GEV gives the best fit. The red curve in figure 4.11 indicates the GEV fit. The results of the different fits at $N = 4096$ can be found in appendix H. The fact that a GEV gives the best fit indicates that as suspected in section 3.2.4 the estimate distribution takes on a type of extreme value distribution. An explanation for a better Gaussian fit in section 3.2.4 might be sought in the difference of the probability of non-exceedance or in that the SANDIA data set possesses other properties which causes other distributions to fit better. Further research into this topic might explain the behaviour of estimates better.

From figure 4.10 it can be seen that especially the low sample sizes sometimes produce ridiculous 50-year load estimations that can reach in the most extreme case up to 180 times the actual 50-year load value (x_{50y}). Thus, the automated algorithm sometimes produces unrealistic fits that would be discarded by wind turbine designers based on experience and expectations. If the wind turbine designer obtains a fit that runs exponentially upwards yielding an unrealistic 50-year load estimate, the designer might decide to for example use another distribution to optimise the fit and obtain a realistic 50-year load. Here an attempt was made to take this customisation process into account by means of a check on each calculated 50-year load $\hat{x}_{F,j}(N)$. The check resamples, fits and extrapolates a new distribution until the $\hat{x}_{F,j}(N)$ falls within the range of the true 50-year value (x_{50y}) as described in equation 4.9:

$$x_{50y}0.5 \geq \hat{x}_{F,j}(N) \geq x_{50y}2,5 \tag{4.9}$$

If the $\hat{x}_{F,j}(N)$ is calculated obeying the range in equation 4.9 indeed a more confined picture of the distribution of the $\hat{x}_{F,j}(N)$ set at low N emerges. The results are given in figure 4.12.

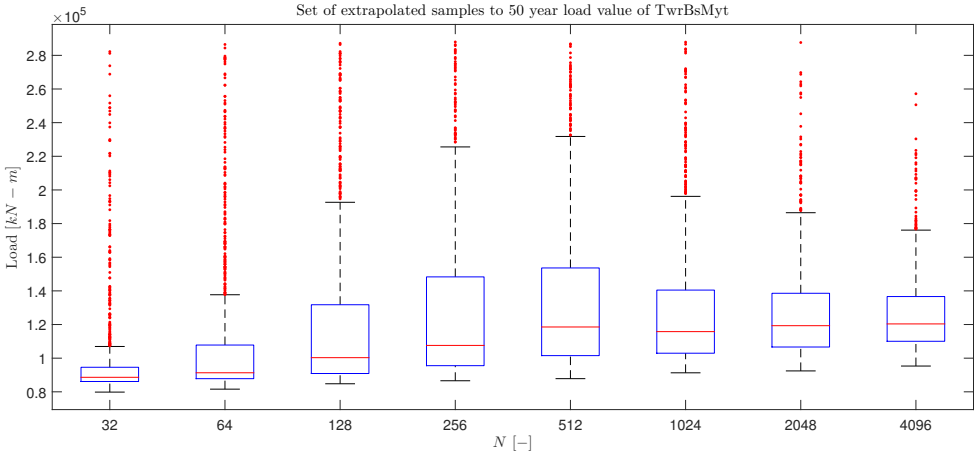


Figure 4.12: Box plots with $N = 32, 64, \ldots, 4096$ of a set of estimated 50-year load estimates re-sampled within bounds of equation 4.9

From the results in figure 4.12 a very peculiar behaviour of the sets emerges. Below 128 sample sizes the median is located *below* the 50-year value. The median climbs from 32 samples to a maximum at 512 sample sizes where it varies as sample sizes increase. In chapter 3 a research indicated that it is expected of extrapolated sets to have an upward bias. In addition and maybe more importantly, it was noted earlier that the data sets for sample sizes lower than 512 show severe uncertainty. Because this behaviour of the data can not be explained or interpreted this resampling technique is not further used.

In addition a wind turbine designer has many tools to optimise the fit to obtain a realistic fit. Besides changing the number of simulations a wind turbine designer can choose to cut the data later or earlier after the *knee*, use a different distribution for the fit and the manner of interpolation can differ to optimise $\hat{x}_{F,j}(N)$. These variations in estimate calculations can significantly influence the 50-year wind turbine loads estimation distribution. These irregularities make it difficult to draw conclusions about the calculations performed.

The results in figure 4.12 show an unexplainable behaviour from a sample size smaller than 128 samples. Together with the trend of exponential increase of variation with decrease in sample size as seen in figure 4.10 it is decided that 50-year loads estimates from 128 samples or smaller are ignored from now on. If new modelling techniques are developed to mimic these variables in a more realistic way, more accurate uncertainty quantification can be obtained.

It must be noted that this research (the distribution of $\hat{x}_{F,j}(N)$ set) was conducted only on TwrBsMyt. Therefore the conclusion that less than 128 10-min wind turbine load extremes yield an unacceptable amount of uncertainty only applies to TwrBsMyt. When looking at other wind turbine loads and deflections, it can be seen that some loads and deflections are less prone to the uncertainty than TwrBsMyt. In plotting the graphs for other loads and defections in appendix G it was noted that all wind turbine loads and deflections begin giving reasonable amounts of uncertainty starting between 256 to 512 simulations.

4.3.4. OPTIMISING WIND TURBINE SAMPLING

The possible origins and consequences of the *knee* in CDF of extreme load estimate distributions has been covered. An already used method for dealing with this problem is to cut the CDF just after the *knee* in order to obtain an approximate linear graph in the $-\ln(-\ln(\ldots))$ plane which can then be fitted (Barone *et al.*, 2012b). From looking at the plots in figure E.1 in appendix E it can be seen that all *knees* have been passed when $-\ln(-\ln(F)) = 1,8$. This can be worked out to a 0,15 probability level, hence only 15% of the data constitutes the linear tail section. This finding indicates that wind turbine designers are only able to use 15% of their wind turbine simulations. In other words, 85% of the strenuously obtained load simulations are thrown away.

With the background theory covered in section 4.3.1 and with a further look into the data it might be possible to aid wind turbine designers in mitigating the 10-min extreme load values present *knee*. Here an attempt is made to manage the valuable 10-min load extremes more efficient for 50-year load estimation calculations. What stands out is the enormous amount of data present in the *knee* that is later removed. It might be possible to exclude this data from the simulations if a large portion of this data shares a common property thereby saving valuable computational resources. This is called importance sampling. As covered in chapter 2, wind turbine loads are calculated using a wind field that is calculated using three inputs:

1. A Wind speed sample from a particular wind speed distribution
2. The level of turbulence intensity
3. The level of wind shear

The first parameter is given in the SANDIA data set with the corresponding loads. The seeds that are used to calculate the other parameters where also given, but were rounded and therefore could not be reproduced. The properties of the SANDIA data set thus only allow for comparison regarding wind speeds. Hence, importance sampling is done by means of the wind speed property. Figure 4.13 is used to illustrate the importance sampling process. In this figure two graphs are drawn in which the loads and wind speeds are plotted separately. The loads are plotted as a CDF as in figure 4.4 and the corresponding wind speeds are plotted on a different axis, adhering to the CDF.

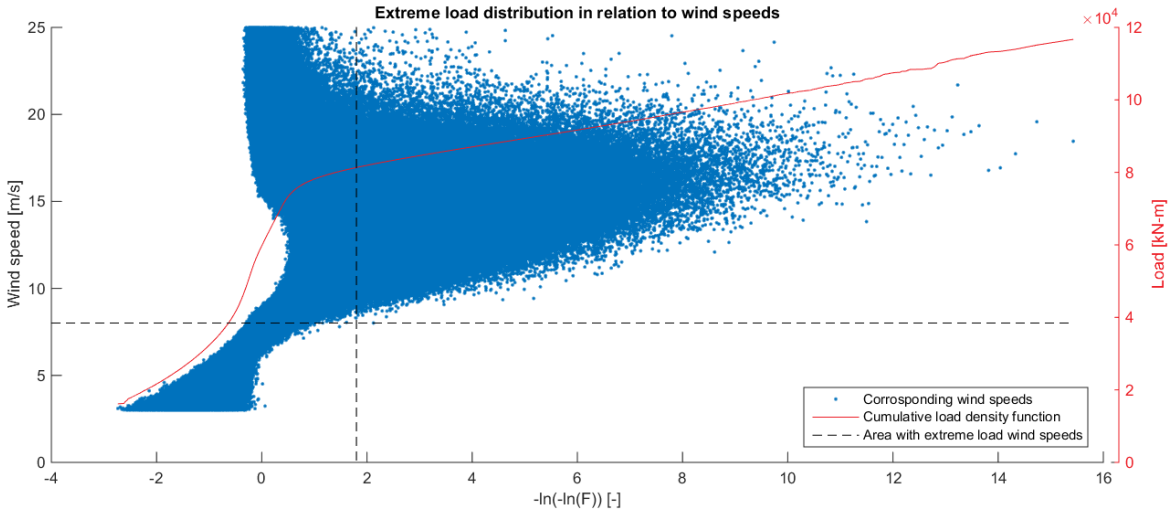


Figure 4.13: Extreme load distribution with corresponding wind speeds and new cut-in wind speed threshold

From the graph in figure 4.13 it was calculated that the loads beyond the *knee* are caused by wind speeds not lower than 8,0099 m/s for the TwrBsMyt moments. For this reason it might be interesting to investigate the possibility of omitting calculation of wind turbine simulation below the 8,0099 m/s threshold. So the importance sampling of simulations consist of effectively raising the cut-in wind speed in order avoid calculating wind turbine loads causing the *knee*. The Rayleigh distribution from which the wind speeds are drawn, is *only* modified for the cut-in wind speed, keeping cut-out and average wind speed constant. This is illustrated in figure 4.14.

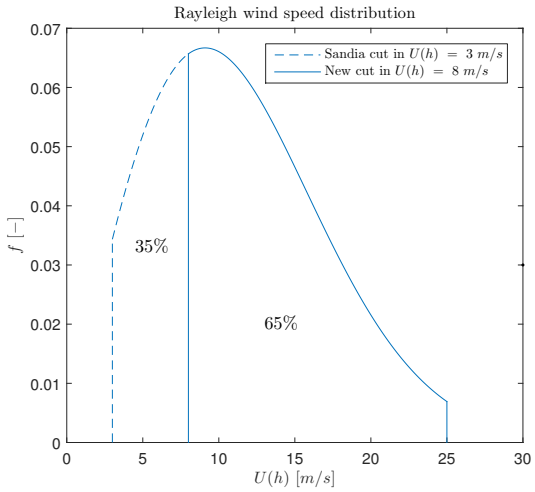


Figure 4.14: Newly defined Rayleigh with $\bar{U}(h) = 11,4$ m/s, $U(h)_{ci} = 8$ and $U(h)_{co} = 25$ m/s for TwrBsMyt

Figure 4.13 indicates that the loads produced by wind speeds below 8,0099 m/s are omitted anyhow during the cutting of the *knee* of the TwrBsMyt moment. When this wind speed is translated to a percentage of simulations saved to figure 4.14 accordingly, it turns out that this corresponds to 35% of the data. So calculating 50-year load estimates from 10-min load extremes originating from a 8 m/s cut-in wind speed Rayleigh distribution will in theory discard 85% − 35% = 50% of the data against 85% in simulating from the 3 m/s cut-in wind speed Rayleigh distribution. In other words, an equal amount of more relevant simulations yields a more accurate 50-year load extrapolation. This hypothesis is tested in two ways:

1. by sampling 65% of the 10-min load extremes (N) from an elevated cut-in wind speed Rayleigh distribution (−35% simulations) of the TwrBsMyt load data set and only discard 50% of the simulations. It is expected that this should result in the same level of uncertainty as doing simulations from a Rayleigh distribution with 3 m/s cut-in wind speed

2. by sampling the full amount of 10-min load extremes (N) form an elevated cut-in wind speed Rayleigh distribution (−35% simulations) of the TwrBsMyt load data and cut away 50% of the simulations. It is expected that this should result in a reduced amount of uncertainty as doing simulations from a Rayleigh distribution with 3 m/s cut-in wind speed.

The differences of the elevated cut-in wind speed Rayleigh distribution uncertainty with the original uncertainty over number of load simulations is given for $N = 512, 1024, 2048, 4096$. Apart from the number of simulations and the reduced load sample distribution, the same assumptions and method as defined in section 4.3.3 apply.

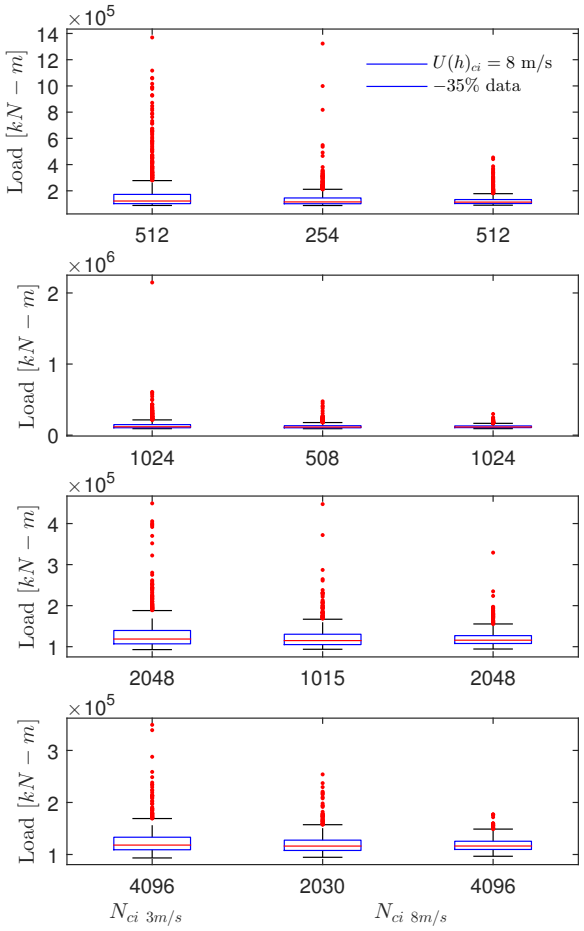


Figure 4.15: TwrBsMyt sampled with different cut-in wind speeds: 3 and 8 m/s ≤ $U(h)$ ≤ 25 m/s ($\bar{U}(h)$ =11,4 m/s)

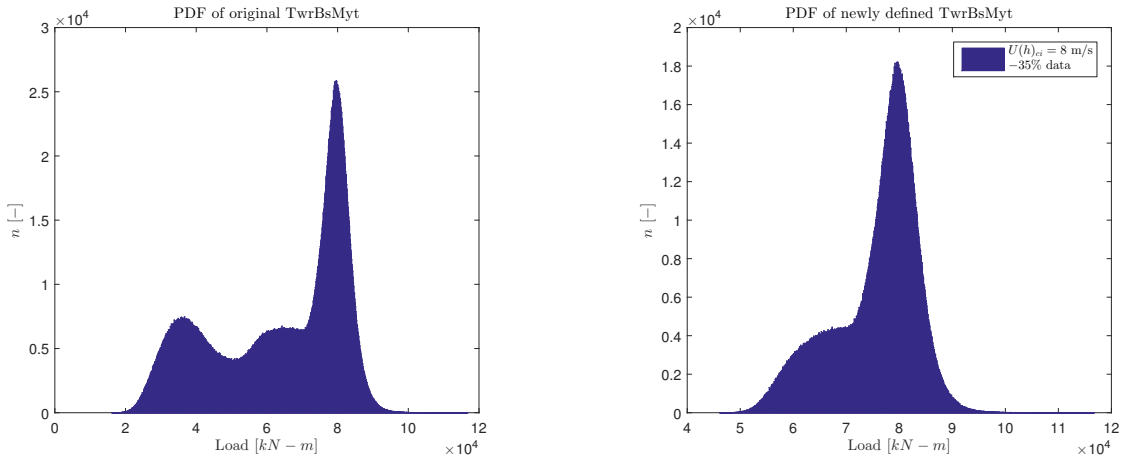
The results for TwrBsMyt are plotted in figure 4.15. The first column contains the 50-year load estimates from sampling from the original Rayleigh wind speed distribution with cut-in wind speed of 3 m/s. The second and third columns contain the results of extrapolating from sampling from 8 m/s cut-in Rayleigh wind speed distribution, with 35% less and equal number of 10-min load extremes respectively. The results indicate that for the same amount or 35% less wind turbine load simulations following a higher cut-in wind speed Rayleigh distribution, significantly less uncertainty exists in extrapolating towards the 50-year load estimate.

The above procedure was done for all loads and deflections present in the SANDIA data set. In all cases, the lowest wind speed causing the loads beyond the 0,85 probability of exeedance was taken as cut-in wind speed in the same way as showed in figure 4.13. The reason for this uniform threshold for the probability of exeedance is that all *knees* exists before this threshold. Note that the calculation of different cut-in wind speeds for different loads and deflections are accompanied with different amounts of load simulation savings

that correspond with the elevated cut-in wind speed Rayleigh distribution. The results for the other loads and deflections can be found in appendix I.

Using these results the wind turbine designer has the opportunity to look at the Rayleigh that is specified for the wind turbine concept and estimate under which wind speed it is safe to omit values using the wind speeds in the graph of appendix I. That way, the probability associated with the area under the Rayleigh from the original cut-in wind speed to the chosen cut-in wind speed may be counted towards the omitted values in the extrapolation process. If there is doubt that lower wind speeds induce high wind turbine loads (for example with a radical new concept), then it is advised to simulate from the normal cut-in wind speed.

Looking at the results of raising the cut-in wind speed, many surprising improvements can be observed. It has been found that an explanation might lie in the alternation of the PDF load distribution. By effectively cancelling loads originating from a wind speed threshold a newly defined load distribution is obtained. Comparison of the original load and deflection PDF and the important sampled load and deflection PDF might yield more insight into load behaviour. The histograms of the two PDFs of TwrBsMyt are drawn in figure 4.16.



(a) TwrBsMyt sampled from 3m/s ≤ $U(h)$ ≤ 25 m/s and $\bar{U}(h)$ = 11,4 m/s

(b) TwrBsMyt sampled from 8m/s ≤ $U(h)$ ≤ 25 m/s and $\bar{U}(h)$ = 11,4 m/s

Figure 4.16: Load sampled histogram from SANDIA data set and load importance sampled histogram by winds speed

In graph 4.16(a) three distinctive peaks can be observed, indicating three processes at which concentrated amount of loads occur. This corresponds with the results discussed from figure 4.4 that multiple processes are present in the SANDIA load and deflection distributions. With effectively removing the loads from the PDF that where produced at wind speeds below 8 m/s in figure 4.13(b) the results show one peak less than the original data. This might indicate that loads produced below 8 m/s constitute a separate process with its own distribution which influences the extreme load distribution in such a way that it ultimately distorts the calculation of the 50-year load value. The same trend can be observed for many loads and deflections for which the results are shown in appendix J. In some cases the the PDF strongly resembles a normal or even an extreme value distribution which indicates that there is no *knee* present at all. For a wind turbine designer it might appear to be detrimental to run load simulations at higher cut-in wind speeds than a lower cut-in wind speed because relatively higher loads are produced. However, for the sake of calculating 50-year loads estimates it is best to do so because the 10-min extremes resulting from the simulations have more relevance.

4.4. CONCLUSION

Eleven different stresses and deflections where simulated over a period of 95,45 years using the FAST wind turbine model released by SANDIA (Barone *et al.*, 2012a). The data set further contains the associated wind speeds. From plotting the wind turbine load simulations verses the wind speed, it can be seen that different deflections and loads show different behaviours. OoPDefl, IPDefl, RootMOoP, RootMFlp, RootMEdg, TwrBsMyt and TwrBsMzt show a distinct *knee* when their CDF is plotted. When the CDF was plotted in multiple

wind speed bins it becomes partly apparent that the *knee* in the distribution is mainly caused at rated wind speeds. The non-linear fashion of the CDF indicates that the distribution is influenced by other distributions from other processes. One process that is believed to play a role in influencing the distribution is that maximum thrust is generated at rated wind speeds.

The *knee* is at least in part responsible for the large uncertainty resulting from calculating 50-year load estimates as shown by figure 4.5. If a basic fit is done for the recommended minimum of 300 minutes of simulation, the estimates are subject to a dangerous upward bias of 2 times the mean and large variations up to 40%. Two explanations are given for the over-dimensionalisation: the upward slope of the *knee* and the sample concentration of the *knee* lies on that part of the slope that progresses steeply upward. The initial upward bias researched in chapter 3 is present in all graphs.

There are three measures that can be undertaken to minimise the effect of the *knee*:

1. Cut away the data before the *knee*
2. Interpolate the remaining linear data
3. The 3 parameter Weibull gives the best fit

Even with these measures the 300 minuets of simulation gives immense uncertainty. The 3 and a half day simulation case for IPDefl, LSSTipMya, LSSTipMza and TwrBsMzt still results in a $\sim$ 55% upward bias of the estimated mean, this translates into an over-estimation of material use needed based on $\sim$ 55% higher loads. Because of this significant impact the relation of this phenomenon is further investigated in chapter 5. The figures in appendix G constitute the answer on sub question four regarding uncertainty quantification in 50-year load estimates. Furthermore it was found that the distribution of the estimates fits a GEV best. The discrepancy between this result and the findings in 3.2.4 have to be researched further in order to understand the behaviour of estimates better.

In addition to the standard measures named above, research was done to devise an additional method. From plotting the CDF of the loads with the corresponding wind speeds it was determined that depending on the wind turbine load type an elevated cut-in wind speed could be set. The loads from these wind speeds would have been discarded by the cut off measure anyhow. The level with which the wind speed can be elevated is load specific and varies between 8,9 and 5,6 m/s. The simulations that could be saved using this method varies between 16 and 45% depending on the type of load. Using higher cut-in wind speeds not only reduced the simulations needed but also improved estimate distributions because less outliers are present.

A link is presumed between the positive results of an elevated cut-in wind speed and the idea that other natural process play a role in distorting the load distribution. From the load probability density graphs multiple Gaussian peaks can be distinguished. The origins of these peaks can be further investigated but by elevating the cut-in wind speed, many of these peaks are cut away from the probability density functions. This indicates a reduced influence from other processes giving models a better fit and more accurate estimates. With these insights and findings sub question five on discovering and mitigating factors that influence 50-year load estimates has been answered.

5

MANAGING 50-YEAR WIND TURBINE LOAD ESTIMATES

This chapter is concerned with the consequences of the uncertainty in 50-year load estimates and the impact it has on wind turbine design as described in sub question six. In this chapter, first the importance of the 50-year load case in wind turbine design will be covered. Reasoning further on the research from chapter 4 the effects of the uncertainty in extrapolation will be further elaborated. This is done trough working out a specific case. TwrBsMyt is chosen to represent the results for all loads in this chapter for the reasons mentioned in section 4.1 and because TwrBsMyt shows a clear *knee* in the data as described in chapter 4. Using such a load includes the effects of the dynamic load distribution.

The same analysis performed on TwrBsMyt are also performed on the remaining ten loads in the SANDIA data set. These results can be found in the appendix. The aim of this case is to translate the uncertainty into terms of probabilities that can be easily calculated to cost savings. In the final section the effect of these concrete numbers on policy and wind turbine design in general are covered:

5.1. APPLICATION OF CALCULATED UNCERTAINTY IN DLC 1.1

In practice, the limiting factor for calculating wind turbine loads is the budget appointed for wind turbine load calculations. In chapter 2 a clear overview is given of the wind turbine design process. The detail design phase has been covered extensively there. Wind turbine designers have to make decisions about certain design features in the detailed design phase based on the DLCs. As explained in chapter 4 a certain amount of uncertainty exists in the 50-year load extrapolation and can cause incorrect selection of wind turbine designs or components. The concept of testing a statement on the bases of observing a process that is modelled via a set of random variables falls under the discipline of *statistical hypothesis testing*.

5.1.1. STATISTICAL HYPOTHESIS TESTING

Statistical hypothesis testing makes use of the null hypothesis (H_0) which refers to a general statement or default position that there is no relationship between two measured phenomena, or no difference among groups. Rejection of the null hypothesis therefore implies correlation between the phenomena but does not prove it. Due to uncertainty in the modelled process, a probability exists that the outcome of the modelled process falls in false spectrum and can therefore be wrongly interpreted as true or false. These wrong rejections of the null hypothesis are known as *errors*:

Type I error	The null hypothesis is wrongly rejected i.e. a condition is judged as present, while it actually is not. This is also known as a false positive.
Type II error	The null hypothesis is wrongly accepted i.e. a condition is judged as not present, while it actually is. This is also known as a false negative.

Statistical tests always involve a trade-off between a acceptable level of false positives and false negatives. The more restrictive the test is, the more rejection of true positive is risked. While a more sensitive test risks accepting false positives. A summary of the scenarios is given in table 5.1.

Error types in Null Hypothesises (H_0)		H_0 is	
		Valid/True	**Invalid/Flase**
Judgement of H_0	**Reject**	*Type I error (false positive)*	Correct inference (true positive)
	Fail to reject	Correct inference (true negative)	*Type II error (false negative)*

Table 5.1: Scenarios for accepting and rejecting null hypothesis

In section 4.3.2 it was found that the calculation method of extrapolation is conservative due to the over-dimentionalisation and can therefore be labelled as a form of restrictive testing. The selection of a wrong wind turbine concept or component is therefore susceptible for the *Type II error* thus also know as *false negative*.

5.2. DILEMMA IN WIND TURBINE DESIGN

The detailed design phase was covered in detail in chapter 2. Furthermore the consequences of the iterative nature of this process was highlighted. With repetition of restrictive testing the likelihood for false negatives increases with each iteration. In the design process wind turbine designers have to balance the wind turbine load simulation budget in the amount of load simulations (10-min load extremes) and the amount of concept iterations. The computational resource budget effectively represents the total number of simulations available (N_{total}) in the wind turbine design load calculation. These simulations must be divided over a number of iterations of design load cases (M), amounting to equation 5.1 below.

$$N_{total} = M \times N \tag{5.1}$$

The wind turbine designers dilemma is that the uncertainty in the load case is unknown and therefore no conscious decision could be made regarding the allocation of simulations to DLC iterations (M) or the number of simulations per load case (N). This dilemma that wind turbine designers face is visualised in figure 5.1. The diamond represents a decision moment if the load is equal or smaller than the component ultimate strength ($\sigma_{A,c,g}$).

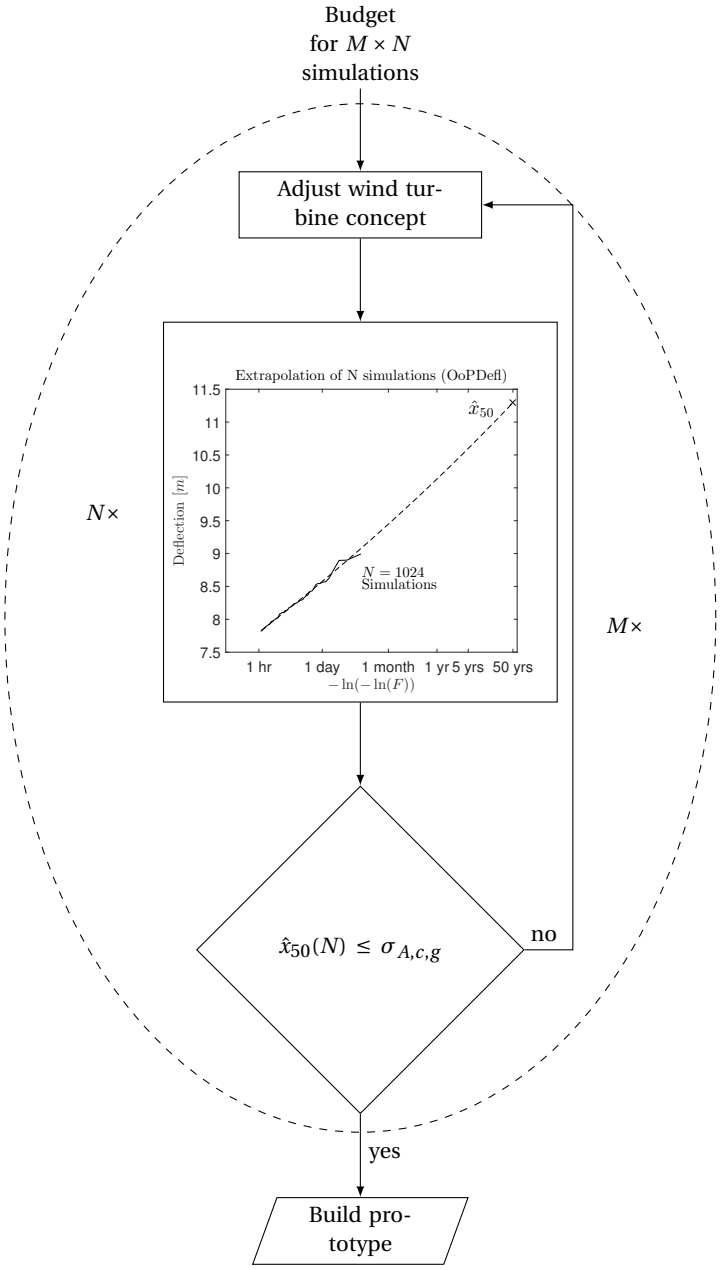


Figure 5.1: The dilemma for wind turbine designers in allocating wind turbine load simulations

In subsection 5.1.1 it was explained that the wind turbine ultimate load calculation suffers predominately from type II errors. In chapter 4 the uncertainty in the 50-year loads and deflections estimates was calculated and therefore the probability of an estimate being a type II error can be calculated by using N simulations (10-min load extremes). This chapter aims to produce a tool for wind turbine designers in which a predetermined acceptable level for a type II error can be set trough which the tool will indicate the number of load simulations needed for that level. This way the allocation of load simulations can be more consciously managed.

5.3. INCORRECT LOAD CALCULATION

With the distribution of the 50-year load estimates ($\hat{x}_{F,j}(N)$) uncertainty depending on the number of load simulations exists. Wind turbine designers have to make decisions concerning wind turbine designs and components based an acceptable amount of 50-year load estimate uncertainty. The inherent uncertainty in

this selection process brings about the principles of statistical hypothesis testing. In this thesis two types of decision cases for wind turbine designers are defined:

- decisions about wind turbine components
- decisions about wind turbine designs or concepts

Wind turbine components are characterised by a known fixed ultimate stress. Wind turbine designs or concepts are defined as having different loads but with the same 50-year estimate distribution. These concepts are further elaborated in sections 5.3.1 and 5.3.2.

5.3.1. DECISIONS ABOUT WIND TURBINE COMPONENTS

Here it is assumed that for the decision about a component the ultimate strength ($\sigma_{A,c,g}$) of that component is known. In this context the four scenarios from statistical hypothesis testing apply. A load can be calculated too high while the component is actually strong enough. This error is defined as a false negative, since the component is incorrectly rejected. On the other hand a load can be calculated too small leading to believe that the component is strong enough to resist while in reality it is not. Because the component is wrongly accepted this type of error is referred to as a false positive. The false positive scenario is illustrated in figure 5.2.

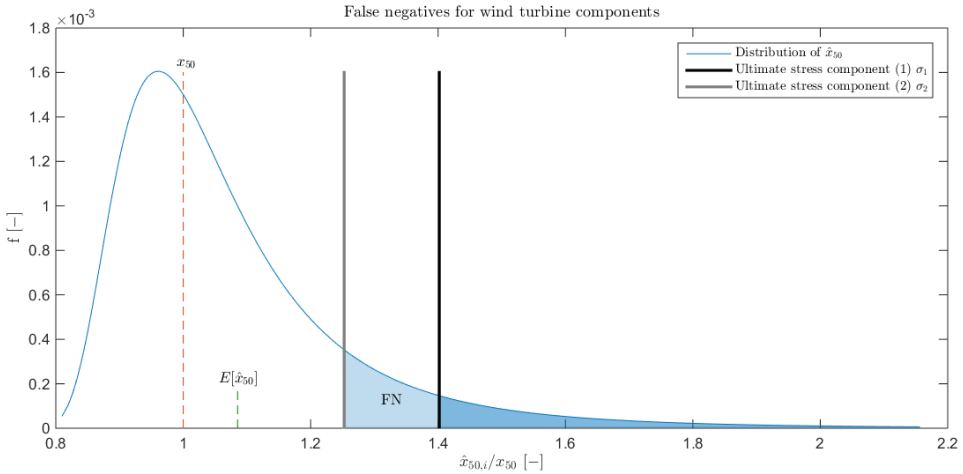


Figure 5.2: False negative wind turbine component calculation with component examples

The example graph in figure 5.2 represents an $\hat{x}_{F,j}(N)$ distribution which was taken from the TwrBsMyt load type. The estimate distribution is derived from N = 4096 which corresponds with a 28-day simulation. Notice that $E[x_F] > x_{50y}$, indicating a larger probability for high load estimates. Analysing figure 5.2 further, firstly the dark blue area covered to the right of component (1) resembles the probability of a sample that indicates that the component is too weak (or the force is too large). However, when looked at the position of the true 50-year load (x_{50y}) it can be seen that it is actually strong enough. Hence the dark blue area represents the probability for false negatives. When looked at component (2), which is situated more closely to x_{50y}, the probability of obtaining a false positive increases. Thus for every point to the right of component (2), including the probability for component (1), the probability for a false negative exists equal to the area to the right of component (2). Therefore the total area to the right of a component which is situated to the right of x_{50y} can be given a cumulative probability. Because the probability of a false negative increases with decreasing component strength, an inverse proportionally exists between the probability for a false negative and the position of a component stronger than the 50-year:

$$P(\sigma_{A,c,g} > x_{50}) = 1 - F(x_{50})$$

Furthermore for the example in figure 5.2, if the true 50-year load value is equal to the component ultimate strength, a higher probability for a too large load for the components ultimate strength exists. This is due to the load over-dimensionalisation. In figure 5.3 multiple cumulative probability graphs at $N = 256, 512, \ldots, 4096$ number of 10-min load extremes are drawn against the probability of a false negative. So

in this graph the probability for a false negative of a given component strength relative to the real 50-year wind turbine load value (x_{50y}) can be found. The green line represents the probability curve discussed in the example of figure 5.2.

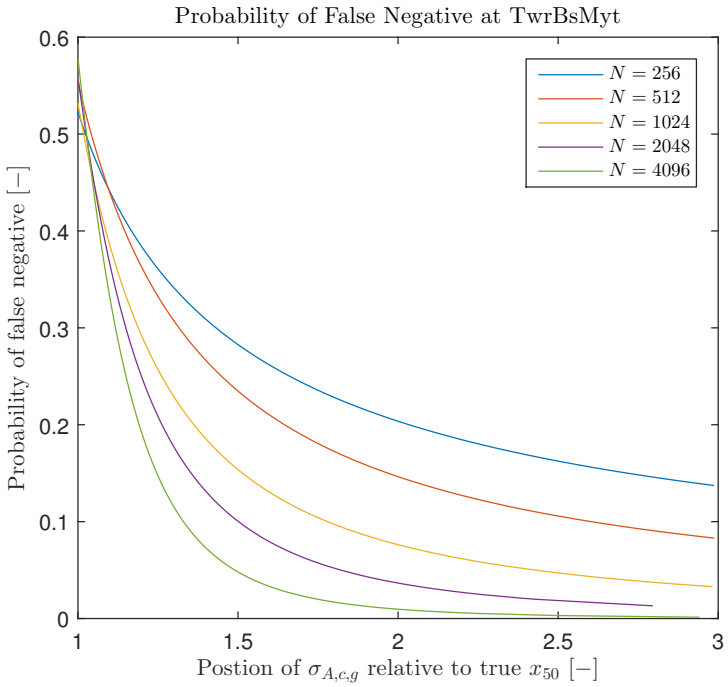


Figure 5.3: Probability of false negative TwrBsMyt for 3 times the distance between the component strength and true 50-year load

As predicted the graph in figure 5.3 for N = 4096 starts above the 50% threshold. Similar graphs as figure 5.3 were constructed for all other load and deflection types. These can be found in appendix K. For 256 10-min load extremes representing roughly 2 days of simulation and a twice as strong material the probability for a false negative still ranges between 12 and 30% looking at all load types. In the case of of TwrBsMyt for 256 extreme values that when the material is 3 times stronger than the actual load, still a 15% probability exist of rejecting the component. Following the trend in figure 5.3 of the number of simulation curves upward, the minimum of 300 minutes of simulations corresponding with 30 simulations advised by IEC 61400-1 in appendix F IEC (2005), can be regarded as having no relevance.

In figure 5.4 the opposite case, that of a false positives, is depicted. Here component (3) lies to the left of the true 50-year load value. So every estimate that indicates a load lower than the component strength presents a false positive error. This situation is depicted with the dark blue area in figure 5.4.

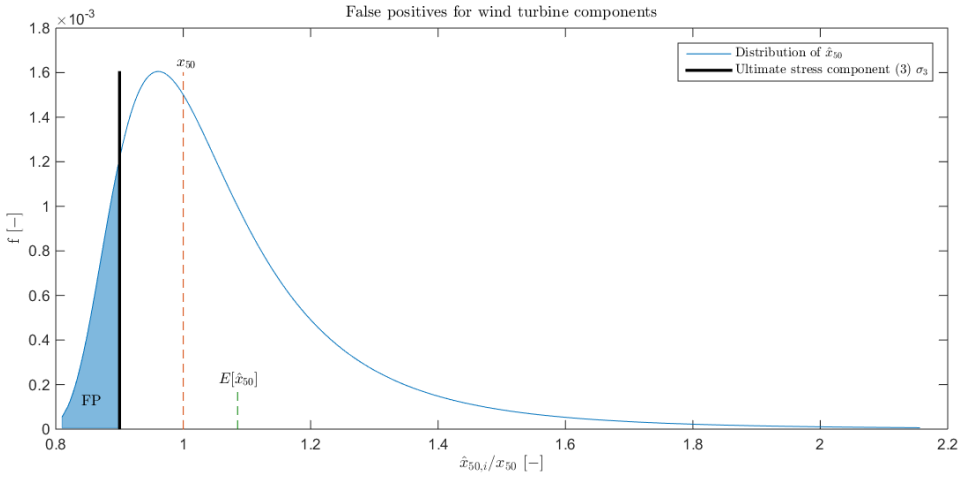


Figure 5.4: False positive wind turbine component calculation with component example

Cumulative false probabilities can be described similar to the theory of false negatives. However, because the probability of false positives increases with increasing ultimate strength the following expression applies:

$$P(\sigma_{A,c,g} < x_{50}) = F(x_{50})$$

For each of the load types the above expression was executed for false positives at $N = 256, 512, \ldots, 4096$ number of simulations. In figure 5.5 the results for TwrBsMyt are plotted and the green line indicates the graph of the example in figure 5.4.

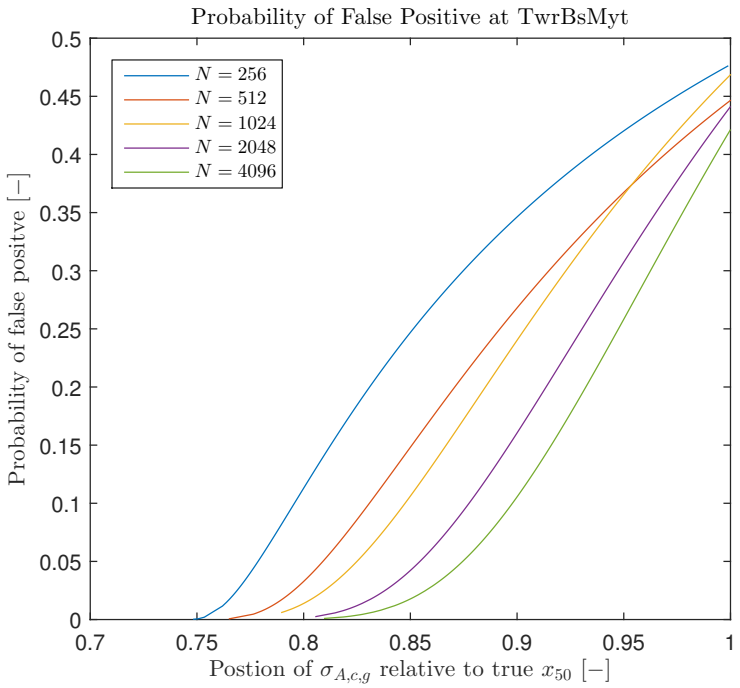


Figure 5.5: Probability of false positive TwrBsMyt for 0,7 times the distance between the component strength and true 50-year load

Notice in this graph that it starts below the 50% threshold in accordance with the graph for false negatives in figure 5.2. Due to the shape of the estimate distribution a much lesser probability for false positives is expected. For 256 extreme values representing roughly 2 days of simulation and a 10% weaker material than the load the probability for a false positive still ranges between 25 and 40% depending on the load. It can be

seen from figure 5.5 that for a material 25% weaker than the TwrBsMyt load, 0% probability exists of obtaining a false positive when doing 256 (2 days) simulation. In this case the right decision will always be made: discard the component. Calculation of probability of false positives at the position of a component relative to the true 50-year load value of all load and deflection types have been put in appendix K.

5.3.2. DECISIONS ABOUT WIND TURBINE DESIGNS

Different wind turbine designs with different 50-year load and deflection values are presented to wind turbine designers. Choices have to be made between these designs. In this case two wind turbine designs with the same 50-year load estimate distribution but with different true 50-year load values exist. Choosing the right design in this case constitutes of the design with the lowest load. So there exists a probability of selecting a wrong wind turbine concept in the area where the distributions overlap each other. The area depends on how far the location parameters, or true 50-year load values, are separated from each other. Thus the probability for the wrong selection is directly proportional to the distance between the location of the load of distributions. This conclusion makes sense, since the probability for a wrong selection decreases when the difference between the actual load calculations lie further apart. This concept is illustrated in figure 5.6.

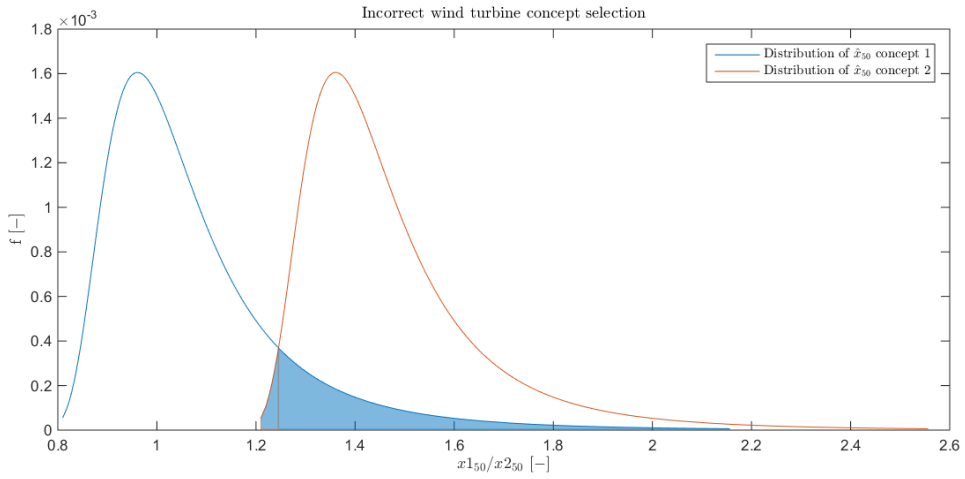


Figure 5.6: Incorrect wind turbine concept selection

In figure 5.6 the blue shaded area represents the probability for a wrongly selected wind turbine design. In this case no distinction can be made for a false positive or false negative, since the wrongly selected wind turbine is always the one with the higher load. Through brute force calculation ($n = 10^6$) the probability of the a wrong concept selection was calculated at different sample sizes over the distance between the x_{50y} of both concepts following: $x_{50y} * (1+i)$ with $i = 0,\ 0{,}05\ 0{,}10,\ \ldots,\ 2$. This is worked out in the graph in figure 5.7.

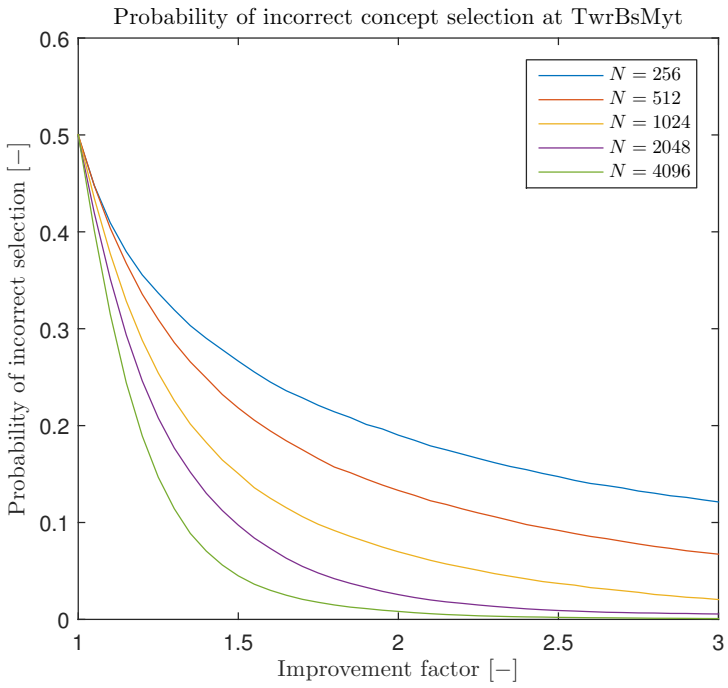


Figure 5.7: Probability of incorrect concept selection TwrBsMyt at various N

The of figure 5.7 can be illustrated by figure 5.6. The blue shaded area in figure 5.6 represents two distributions that are obtained with N = 4096 samples and the means are 1,40 times separated from each other. When looked at figure 5.7 it can be seen from the green graph that the probability represented by this blue area is then $\sim$ 0,09. Hence, the probability of selecting the wrong concept when 4096 10-min load extremes were obtained for both concepts and the x_{50y} differ 1,4 times is 9%. Again, looking at the trend of simulation lengths, the minimum of 30 10-min load extremes implies that very little can be said about a comparison between the concepts at these simulation lengths. Calculation of the probabilities for wrong concept selection of different wind turbine load and deflection types have been put in appendix L. This wind turbine concept selection evaluation has been done on the premisses of that both concepts contain an equal amount of simulations. It might be interesting to work out concept comparison between concepts with different numbers of simulations in further research.

5.3.3. SAFETY FACTORS

It has become apparent that the 300 minutes of simulation minimum set by the IEC 61400-1 IEC (2005) offers little support for wind turbine designers to accurately assess their designs. Furthermore it was covered in chapter 2.3 that safety factors reach between 1 and 1,5. Safety factors have no influence on concept selection but do on component selection. If the most severe safety factor of 1,5 is used, the x-axis in figure 5.5 effectively moves 0,5 points to the left, situating 1,5 where previously 1 was located and 2 where previously 1,5 was located. The probability for false positives thus increase even more to dramatic proportions. As shown in both chapter 3 and 4 an inherent upward bias is present in the calculation technique which acts already as an inherent safety factor. It might even be worth thinking about a reversed safety factor. Hypothetically, a safety factor of 0,5 would then induce a $\sim$ 23% probability for a false negative selections for a 2-day simulation at 1,5 times the component strength of the true TwrBsMyt load.

5.4. CONCLUSION

Wind turbine designers have limited computational resources and are therefore forced to balance the number of simulations and the number of concept iterations. Wind turbine load calculation uncertainty needs to be managed correctly to avoid unnecessary material cost. For determining the probability of a false outcome of a wind turbine load estimation a distinction was made between wind turbine components and wind

turbine concepts or designs. Components compose of a hard load threshold and wind turbine concepts are composed of a load distribution on another load location. The probability for the selection of a wrong component or concept depending on their position, has been worked out in appendices K and L. Wind turbine concept selection evaluation has been done on the premisses of that both concepts contain an equal amount of simulations. It might be interesting to work out concept comparison between concepts with different numbers of simulations.

For 256 10-min load extremes, representing roughly 2 days of simulation, and a twice as strong material than the load, the probability of obtaining a false negative ranges between 12 and 30% regarding the load type. For 256 extreme values and only 25% weaker component the probability of obtaining a false positive is 0%. Hence, as concluded in chapter 4 the probability of obtaining a false positive from the GEV distributed estimates is bigger than under-dimentianolisation. Furthermore, in the light of the IEC 614000-1 appendix F, even when 256 10-min load extremes are taken of TwrBsMyt representing 2 days of simulation at 3 times the component strength than the actual load, still a 15% probability exist of rejecting the component. In IEC 61400-1 appendix F, the minimum recommended simulation time is 300 minutes of simulation. This corresponds with only 30 10-min load extremes and has little value for either concept or component selection.

The use of safety factors escalates the probability of false negative component selection due to the GEV distributed 50-year estimates. The calculation method of 50-year estimates contains an inherent safety factor and a reversed safety factors might even be considered. Sub question six refers to the influence of uncertainty in the wind turbine design process. With the results presented in this thesis, wind turbine designers can manager their safety factors and number of simulations (use the graphs in appendices K and L) in order to obtain manageable results.

6

CONCLUSIONS AND RECOMMENDATIONS

6.1. CONCLUSIONS

Challenges in the wind turbine industry are mainly cost-related since the LCOE of wind energy has in many cases not yet reached grid parity. Hence, the field of wind turbine research focuses mainly on cost reduction. Load calculation standards significantly influence wind turbine costs. The conclusions and implications for the research on uncertainty of 50 year wind turbine load calculations are summed up below.

Uncertainty in extrapolation techniques

Because the wind turbine load samples originate from a purely mathematical construct, it is not influenced by natural variations, thereby enabling evaluation of the extrapolation method in itself. Five insights were extracted by doing mathematical research on the 50-year load estimate calculation method:

- The estimate mean distribution does not average on the mean of the sample distribution over various sample quantities. An initial upward bias was noticed at low sample numbers towards +30% until $N <$ 500. Larger sample sizes at $500 \leq N \leq 2000$ a negative bias of approximately -1% of the mean can be observed. This phenomenon was already discovered by White (1969).
- As the amount of drawn samples corresponds with the amount of extreme values, still uncertainty exists. This means that there is an inherent uncertainty in the research method in which the level of uncertainty is related to the sample method from the sample distribution.
- A Monte Carlo test was performed which was found the $\hat{b}$ parameter to be within bounds around the target value of $-0,5$. This indicates a correct calculation method.
- In this case the best fit for the estimate distribution was found to be the Gaussian distribution. It is believed that at low amounts of samples the estimate distribution resembles an extreme value distribution and thus extreme value distributions fit better at lower sample sizes.
- The fit of plotting positions without the influence of extrapolation was researched. For small (20) and larger (2048) sample sizes the plotting positions fitted within 7% each other.

Two important conclusions can be drawn from these points. Firstly, the calculation method applied appears to be mathematical correct. Secondly from this sound mathematical construct a significant upward bias can be detected, which is thus solely caused by the 50-year wind turbine load calculation method. Hence, when evaluating wind turbine load simulations this effect must be taken into consideration.

Wind turbine load simulations

- Plotting the wind speed verses the eleven loads and deflections reveals different behaviour for different loads and deflections. The CDF of OoPDefl, IPDefl, RootMOoP, RootMFlp, RootMEdg, TwrBsMyt and TwrBsMzt show a distinct *knee*. The *knee* is mainly caused at rated wind speeds. The non-linear fashion indicates that the distribution is influenced by other distributions belonging to other processes. One

process that is believed to play a role in influencing the distribution is the maximum thrust that is generated at rated wind speed. Furthermore it was found that the distribution of the estimates fits a GEV model best. The discrepancy between this result and the findings in 3.2.4 have to be researched further in order to understand the behaviour of estimates better.

- The *knee* is at least partly responsible for the large uncertainty resulting from calculating 50-year load estimates as shown by figure 4.5. If a basic fit is done for the recommended minimum of 300 minutes of simulation, the estimates are subject to a dangerous upward bias of 2 times the mean and large variations up to 40%. Two explanations for the over-dimensionalisation are given: the upward slope of the *knee* and that the weight of the curve lies on that part of the slope that progresses steeply upward. There are three existing methods to manage the *knee*: cutting off the steep initial slope, interpolating the tail and apply the 3 parameter Weibull regression model for the fit. With these measures in place three days of simulation for IPDefl, LSSTipMya, LSSTipMza and TwrBsMzt results in a $\sim$ 55% upward bias of the estimated mean. This immense upward load bias ultimately leads to high material use and wind turbine cost.

- Research was done to devise an additional method to mitigate uncertainty in 50-year estimates. From plotting the CDF of the loads with the corresponding wind speeds it was determined that elevated cut-in wind speeds could be set depending on the wind turbine load type. The wind speed elevation is load-specific dependent and varies between 8,9 and 5,6 m/s. The amount of simulations that could be saved using this method therefore varies between 16 and 45% depending on the type of load. Using higher cut-in wind speeds not only reduced the simulation time needed but also improved estimate distributions because less outliers were present. A relation is believed to exist between the positive results of an elevated cut-in wind speed and the idea that other natural process play a role in distorting the load distribution. From the load probability density graphs multiple Gaussian peaks can be distinguished. By elevating the cut-in wind speed, many of these peaks are removed from the probability density functions, indicating a reduced influence of other processes giving models a better fit and extrapolation estimate. The origins of these peaks can be further investigated to identify the other processes involved.

Implications uncertainty in wind turbine design

- Wind turbine designers have limited computational resources and are therefore forced to balance the number of simulations and the number of concept iterations. A tool was designed that attempts to enable wind turbine designers to mange the number of simulations. In determining the probability of a false decision of a wind turbine designer, a distinction was made between wind turbine components and wind turbine concepts. For 256 extreme values representing roughly 2 days of simulation and a material twice as strong than the load, the probability that results in a false negative component ranges between 12 and 30%. For 256 extreme values and an only 25% weaker component the probability of obtaining a false positive is 0%. Hence the probability of obtaining a false positive from the GEV distributed estimates is bigger than false negatives. This makes wind turbine concept and component selection susceptible to over-dimentionalisation leading to high material use and cost.

- The standard set by IEC 61400-1 to conduct extrapolation on 30 extreme values contains such a significant amount of uncertainty that it can be regarded as having no relevance for 50-year loads. When current safety factors of 1,5 are added to the result, the probability of obtaining a false negative will dramatically increase the amount of false rejections. The calculation method of 50-year estimates contains an inherent safety factor and reversed safety factors might even be considered.

6.2. RECOMMENDATIONS

The upward initial bias was first described by White (1969). No clear explanation as yet been given for this behaviour. In this thesis the sorting step is suggested to cause the upward bias but a deeper mathematical understanding of the origin of the bias might lead to new insights for improving the accuracy for 50-year load calculations. Furthermore the influence of extrapolation on different plotting positions of wind turbine load data can produce leads for an optimal plotting position that yields the most accurate wind turbine load estimate.

The discrepancy between the estimate distribution fits originating from a Gumbel distribution (section 3.2.4) and the SANDIA data set (section 3.2.4) stands out. In the case of samples originating from the Gumbel a Gaussian distribution appears to fit the estimate distribution, while samples originating from the SANDIA data set assume a GEV distribution. An explanation might be sought in the difference of the probability of non-exceedance or in that the SANDIA data set possesses other properties which causes other distributions to fit better. Further research into this topic might explain the behaviour of estimates better.

The resampling method for unrealistic 50-year estimates yielded inexplicable behaviour and was therefore ignored in this thesis. However, if the understanding of this behaviour is increased, a better method for managing unrealistic estimates might be devised. This can lead to a better description of the uncertainty in 50-year load estimates in practice.

Two research results presented in this thesis indicate that multiple processes play a role in wind turbine loading. The irregular wind speed binned CDF (*see section* 4.2) and the load probability density distribution with multiple Gaussian peaks (*see section* 4.3.4) both indicate that multiple load processes are at hand. One process that is believed to play a role in influencing the distribution is that maximum thrust is generated at rated wind speeds. Other processes need yet to be identified. If the origin of these load processes is better understood, the process causing 50-year load estimates might be derived and by isolating the extremes of this process a far more accurate 50-year load estimate calculation can be performed.

The proposition of raising the cut-in wind speed for wind turbines simulations appears to be valid on the SANDIA data set (Barone *et al.*, 2012a). However, validation of this method is needed on other data sets. The SANDIA data set (Barone *et al.*, 2012b) might yield other results partly because it is a different data set and partly because this wind turbine is situated in an off-shore environment.

The wind turbine concept selection evaluation that is presented in section 5.3.2, was performed on the premisses that both concepts contain an equal amount of simulations. It may benefit wind turbine designers to further work out concept selection evaluations between concepts with different numbers of simulations.

With the use of safety factors the probability of obtaining a false negative will increase, thereby yielding an increased amount of false rejections. The calculation method of 50-year estimates contains an inherent safety factor. Future research might shed light on the effect of using reversed safety factors to counterbalance the inherent safety factor present in the 50-year load estimate calculations.

BIBLIOGRAPHY

M. Barone, J. Paquette, B. Resor, and L. Manuel, *Decades of Wind Turbine Load Simulation,* 50th AIAA Aerospace . . . , 1 (2012a).

J. S. White, *The Moments of Log-Weibull Order Statistics,* Technometrics **11**, 373 (1969).

R. Wiser, Z. Yang, M. Hand, O. Hohmeyer, D. Infield, P. H. Jensen, V. Nikolaev, M. O'Malley, G. Sinden, and A. Zervos, *IPCC Special Report on Renewable Energy Sources and Climate Change Mitigation*, Tech. Rep. (2011).

I. Kubiszewski, C. J. Cleveland, and P. K. Endres, *Meta-analysis of net energy return for wind power systems,* Renewable Energy **35**, 218 (2010).

Fraunhofer ISE, *LEVELIZED COST OF ELECTRICITY RENEWABLE ENERGY TECHNOLOGIES Levelized Cost of Electricity Renewable Energy Technologies*, Tech. Rep. November (2013).

IRENA, *Renewable energy technologies: cost analysis series Wind Power*, Tech. Rep. 5/5 (2012).

G. Sieros, P. Chaviaropoulos, J. D. Sø rensen, B. H. Bulder, and P. Jamieson, *Upscaling wind turbines: theoretical and practical aspects and their impact on the cost of energy,* Wind Energy **15**, 3 (2012).

J. Fogle, P. Agarwal, and L. Manuel, *Towards an improved understanding of statistical extrapolation for wind turbine extreme loads,* Wind Energy **11**, 613 (2008).

IEC, *International standard IEC 61400-1,* (2005).

IEC, *International Standard IEC 61400-3,* (2009).

M. Barone, J. Paquette, and B. Resor, *Simulating the Entire Life of an Offshore Wind Turbine,* (2012b).

J. Manwell, J. McGowan, and A. Rogers, *Wind Engineering*, 2nd ed. (2009) pp. vii, 577 p.

I. Britannica, *Wind Turbine,* (2015).

E. Gumbel, *Statistics of Extremes,* Columbia University Press, New York (1958).

R. Genz, N. Kristian Bendix, and P. Hauge Madsen, *An investigation of load extrapolation according to Abstract :,* Ewec (2006).

J. Peeringa, *Comparison of extreme load extrapolations using measured and calculated loads of a MW wind turbine,* Wind Energy , 16 (2009).

P. J. Moriarty, W. E. Holley, and S. P. Butterfield, *Extrapolation of Extreme and Fatigue Loads Using Probabilistic Methods,* NREL (2004).

W. Weibull, *A statistical theory of the strength of materials,* (1939).

T. W. Anderson and D. A. Darling, *Asymptotic Theory of Certain "Goodness of Fit" Criteria Based on Stochastic Processes,* (1952).

G. R. Shorack and J. A. Wellner, *Empirical Processes with Applications to Statistics,* in *Empirical Processes with Applications to Statistics* (SIAM, 1986) 2009th ed., p. 239.

L. R. Beard, *Statistical analysis in hydrology,* Transactions of the American Society of Civil Engineers **108**, 1110 (1943).

A. Benard and E. C. Bos-Levenbach, *Het uitzetten van waarnemingen op waarschijnlijkheids-papier,* Statistica Neerlandica **7**, 163 (1953).

G. Blom, *Statistical estimates and transformed beta variables* (NY:Wiley, New York, 1958).

A. Garcia, *Unbiased plotting positions — A review — Comments,* (1979).

I. I. Gringorten, *A plotting rule for extreme probability paper,* (1963).

A. Hazen, *Storage to be provided in impounding reservoirs for municipal water supply.* Transactions of the American Association of Civil Engineers 77 , 1539 (1914).

J. M. Landwehr, N. C. Matalas, and J. R. Wallis, *Probability weighted moments compared with some traditional techniques in estimating Gumbel Parameters and quantiles,* (1979).

D. McClung and A. Mears, *Extreme value prediction of snow avalanche runout,* (1991).

J. W. Tukey, *The Future of Data Analysis,* (1962).

G. H. Yu and C. C. Huang, *A distribution free plotting position,* Stochastic Environmental Research and Risk Assessment **15**, 462 (2001).

J. Jonkman, S. Butterfield, W. Musial, and G. Scott, *Definition of a 5-MW reference wind turbine for offshore system development,* Contract , 1 (2009).

E. O'Brien, D. Sloan, K. Bulter, and K. J., *Traffic load 'fingerprinting' of bridges for assessment purposes.* The Structural Engineer **73**, 320 (1995).

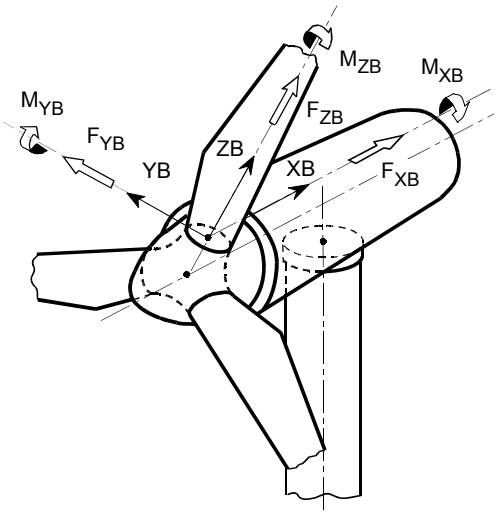
Germanischer Lloyd, *Rules and Guidelines Industrial Services: Guideline for the Certification of Offshore Wind Turbines,* , 155 (2005).

J. Connell, *The spectrum of wind speed fluctuations encountered by a rotating blade of a wind energy conversion system,* (1982).

G. Lloyd, *Wind Energy*, Tech. Rep. (Germanischer Lloyd, Hamburg, 2005).

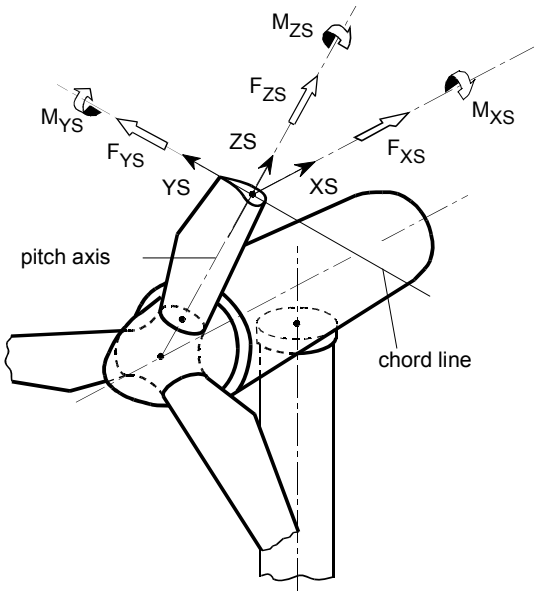
A

WIND TURBINE COORDINATE SYSTEM



XB in direction of the rotor axis
ZB radially
YB so that XB, YB, ZB rotate clockwise

Figure A.1: Blade bearing coordinate system after Germanischer Lloyd (2005)

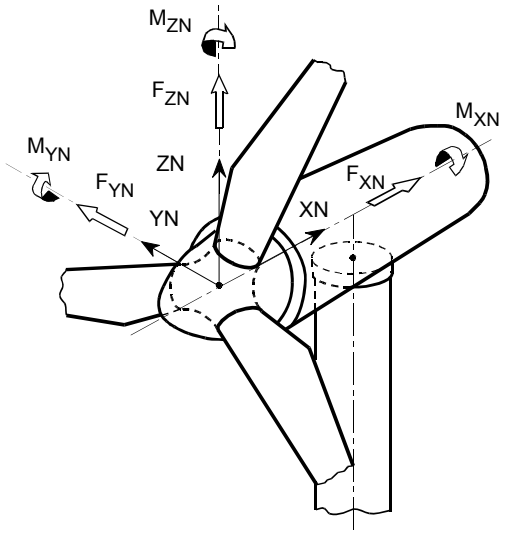


YS in direction of the chord, orientated to blade trailing edge

ZS in direction of the blade pitch axis

XS perpendicular to the chord, so that XS, YS, ZS rotate clockwise

Figure A.2: Chord bearing coordinate system after Germanischer Lloyd (2005)

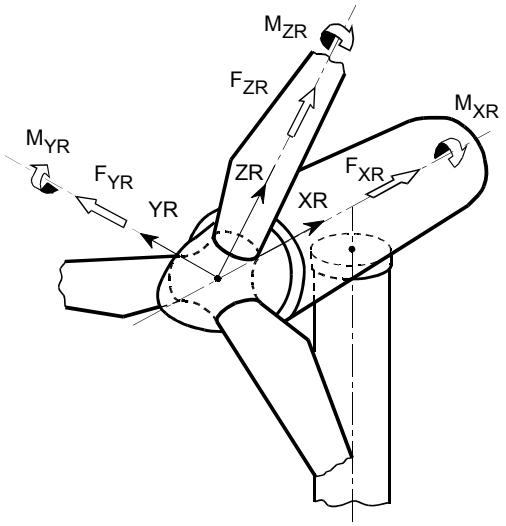


XN in direction of the rotor axis

ZN upwards perpendicular to XN

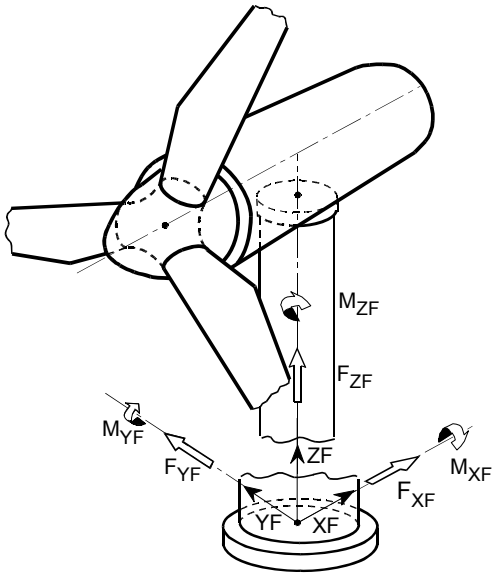
YN horizontally sideways, so that XN, YN, ZN rotate clockwise

Figure A.3: Hub bearing coordinate system after Germanischer Lloyd (2005)



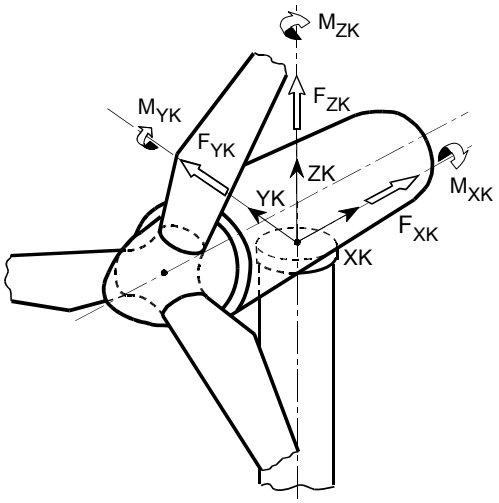
XR in direction of the rotor axis
ZR radially, orientated to rotor blade 1 and perpendicular to XR
YR perpendicular to XR, so that XR, YR, ZR rotate clockwise

Figure A.4: Rotor bearing coordinate system after Germanischer Lloyd (2005)



XF horizontal
ZF vertically upwards in direction of the tower axis
YF horizontally sideways, so that XF, YF, ZF rotate clockwise

Figure A.5: Tower bearing coordinate system after Germanischer Lloyd (2005)



XK horizontal in direction of the rotor axis, fixed to nacelle
ZK vertically upwards
YK horizontally sideways, so that XK, YK, ZK rotate clockwise

Figure A.6: Yaw bearing coordinate system after Germanischer Lloyd (2005)

B

LOAD CLASSES EXPLAINED

Steady loads: are loads that act continuously on the wind turbine. Examples are the weight of the turbine on the tower, mean wind speeds and centrifugal forces due to the rotation of the rotor.

Cyclic loads: arise from rotor rotation and the most profound forces (of HAWT) are located at the blade roots due to gravity. Other periodic loads originate from wind shear, crosswind (yaw error), vertical wind, yaw velocity and tower shadow.

Stochastic loads: in general are used to describe a load process that is unpredictable due to the influence of a random variable. A load process where at least one random variable is introduced to account for the unpredictability in a variable is at least in part a stochastic process. A pure stochastic process is a completely random process, opposite of that of a pure deterministic process in which the process is completely defined. Naturally many of the origins of wind turbine loads are induced by wind turbulence. Because of the unpredictable nature of wind turbulence, it is defined as a stochastic load process. Wind turbulence can vary in time, space and across the rotor which causes rapidly changing forces across the blades. As the blades passes trough the turbulence of each other wake, the turbulence is also affected by the rotation itself Connell (1982).

Transient loads: are loads witch occur on an irregular short term bases. Starting and stopping loads are frequent transient loads for wind turbines. Other important transient loads are due to sudden wind gusts, changes in wind direction, blade pitching motions and teetering. Wind gust analysis is identified as a key load case

Resonance-induced loads: emerge if a structure is exited to its natural frequency. Although designers try to avoid this, turbulence is often responsible for some excitation to a natural frequency. Structure frequency resonance dependence on rotor speed can be evaluated using a Campbell diagram such as in figure B.1.

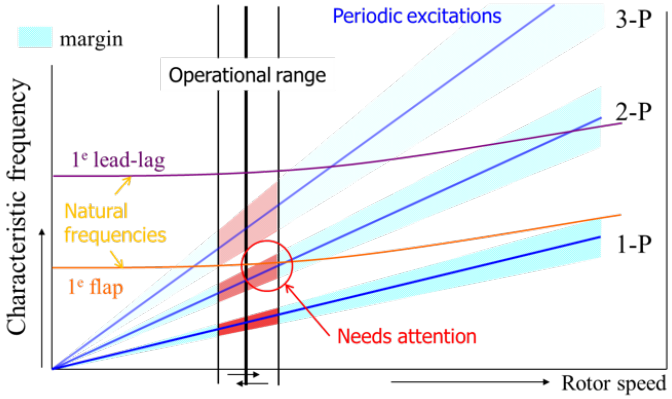


Figure B.1: Campbell diagram

The abbreviation P are the harmonics at which the component is exited per rotation of the rotor. Components of the wind turbine have a different natural frequency over a range of rotational speeds. It should be noted that the harmonics of the blades are dependent on the number of blades. For tower shadow induced periodic load the harmonics of the number of blades add, while the remaining harmonics cancel. So a three blades wind turbine only experiences $3P$, $6P$, $9P$ ... harmonics while two bladed wind turbine blades experience, $2P$, $4P$, $6P$ ... harmonics. The Campbell diagram shows problem areas where the natural frequencies of components meet a similar turbulence induced integrated frequency that can excite the component thereby enhancing crack formation. So the natural frequencies of the components cannot intersect with the harmonics within the operational range. This can be avoided by ether changing the natural frequency by stiffening the components or change the operational range. More standards and guidelines on this area is given by Lloyd (2005).

C

FUNCTION GUMBELUNC

```
%% GumbelUnc
% Quantifictaion of Gumbel extrapolation unceirtainty
%% Parameters
% # |N| = Available data points
% # |r| = Fits per |N|
% # |CI| = Confidence Interval of extrapolated values: |68| , |95| or
% |99.7| %
% # |normalise| = normalises the axis with: |'off'| or |'yon'|
% # Optional: |mu| = mu of the Sample Gumbel |(default = 10)|
% # Optional: |beta| = beta of the Sample Gumbel |(default = 4)|
%% Syntax
% |[Returnlvlplot] = GumbelUnc(N,r,CI,normalise,mu,beta)|
%
% |[Returnlvlplot,musigma] = GumbelUnc(N,r,CI,normalise,mu,beta)|
%
% |[Returnlvlplot,musigma,SampleGumbel] =
% GumbelUnc(N,r,CI,normalise,mu,beta)|
%% Description
% *|GumbelUnc(N,r,CI,normalise,mu,beta)|* quantifies in a
% normalised way the uncertainty of extrapolation from 2 towards |N|
% available data points. It plots the return level plot and saves under
% GumbelUnc. Both axis can be normalised with |normalise|. Furthermore the
% sample Gumbel can be adjusted entering for |mu| and |beta|.
%%
% *|[Returnlvlplot] = GumbelUnc(N,r,CI,normalise,mu,beta)|* gives same
% output as |GumbelUnc(N,r,CI,normalise,mu,beta)|
%%
% *|[Returnlvlplot,musigma] = GumbelUnc(N,r,CI,normalise,mu,beta)|* stores
% the mu and sigmas of |N| normal distributions in the workspace in |musigma|
%%
% *|[Returnlvlplot,musigma,SampleGumbel] =
% GumbelUnc(N,r,CI,normalise,mu,beta)|*
% Gives a figure of the Gumbel from which N samples are drawn for |r| times
% per N. Negative values of the Sample Gumbel or a Sample Gumbel that is
% elevated too high may distort return level plot outcomes.

%% Function
function [Returnlvlplot,musigma,SampleGumbel] = ...
    GumbelUnc(N,r,CI,normalise,mu,beta)

%some in- and output checks
if nargout > 3
    error('GumbelUnc:Too many outputs, give a maximum of 3 outputs');
end


if nargin > 6
    error('GumbelUnc: Too many inputs, give maximum of 6 inputs');
elseif nargin < 4
```

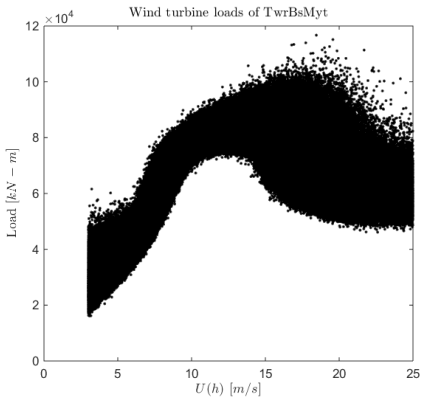
```
50      error('GumbelUnc: Too fieuw inputs, give a minimum of 4 inputs');
51  end
52
53  switch nargin
54      case 4
55          mu = 10;
56          beta = 4;
57      case 5
58          error('Specify both mu and beta for sample Gumbel');
59  end
60
61  if mu > N;
62      error('mu cannot be greater than N.')
63  end
64
65  if CI == 68;
66      Nsigma = 1;
67  elseif CI == 95;
68      Nsigma = 2;
69  elseif CI == 99.7;
70      Nsigma = 3;
71  else
72      error('Confidence Interval (CI) must be equal to 68%, 95% or 99.7% without ...
            percent sign.')
73  end
74
75  %%
76  % Construction Sample Gumbel
77  %
78  % The confidence interval from which is Sampled is between
79  % 0.1% and 99.9%.
80  %confidence interval
81  Rb = mu - (beta * log(-log(0.999)));
82  Lb = mu - (beta * log(-log(0.001)));
83
84  %creating values that are drawn (X) acording to the Sample Gumbel (f)
85  X = linspace(Lb,Rb,N);
86  f = gevpdf(X,0,beta,mu);
87
88  %some warnings for Sample Gumbel to prevent distortion in return level plot
89  if f(N)/max(f) > 0.05;
90      warning('Sample Gumbel does not represent extreme values well because of limited ...
            use of probabilities. Change mu, beta, Duration0 or V.')
91  end
92
93  if f(find(X ≥ 0,1))/max(f) ≥ 0.05
94      warning('Sample Gumbel does not represent extreme values well because of ...
            significant negative values probabilities.')
95  end
96
97  %%
98  % Unceirtanty of extrapolation toward |N|
99  %
100 % Here de data is generated using |polyfit| as fitting tool. Due to the
101 % logaritmic axis a linear relation exists. The |r| fitted lineair graphs
102 % are extrapolated towards |FN| which is the |N| th |X| value. A normal
103 % distribution of these values is fitted and described by mu and sigma.
104 % both are stored in the parameter |Normal|. |N| mu's and sigma's are
105 % iteratively stored in |Allmusigma|.
106
107 FN = -log(-log(N/(1+N)));   %the probability of the extremest value
108
109 Allmusigma = zeros(N,2);
110 for i = 2:N;                %minimal of 2 datapoints for fit
111    betamu = zeros(r,2);
112    Q = zeros(r,1);
113    for j = 1:r;             %r extreme value points
114        x = sort(X(discretesample(f,i)));        %Sort values
115        P = -log(-log((1:i)/(1+i)));             %Plot positions
116        betamu(j,:) = polyfit(P,x,1);            %Param. of lineair graph
117        Q(j,:) = betamu(j,2) + (betamu(j,1)*FN); %All extrapolated values
```

```
    end
  Normal = fitdist(Q,'Normal');                  %Fit a Normal distribution
  Allmusigma(i,:)= [Normal.mu Normal.sigma];     %Store the mu and sigma's
end

%%
% Plot normal distributions

%normalisation towards Xtrue
Xtrue = ones(N,1)*(mu+(beta*FN));

NormaliseN = ((1:N)./N);
NormaliseXmu = ((Allmusigma(:,1) - Xtrue)./Xtrue);
NormaliseXup = (((Allmusigma(:,1)+(Nsigma*Allmusigma(:,2)))...
    - Xtrue)./Xtrue);
NormaliseXdown = (((Allmusigma(:,1)-(Nsigma*Allmusigma(:,2)))...
    - Xtrue)./Xtrue);

%Sample Gumbel that represents the data
SampleGumbel = figure;
plot(X,f,'k')
xlabel('$X$','Interpreter','LaTeX')
ylabel('$f$','Interpreter','LaTeX')
title({'Sample Gumbel'},'Interpreter','LaTeX')

set(gcf, 'Color', 'none')
set(gcf, 'Units', 'centimeters')
set(gcf, 'OuterPosition', [5, 5, 9.3, 10])

%manage output arguments
switch nargout
    case 0
        close figure 1
    case 1
        close figure 1
    case 2
        close figure 1
end

%check axis for normalisation
normalis1 ='off';
normalis2 ='yon';

if strcmp(normalis1,normalise)==1;
    Returnlvlplot = figure;
    hold on
    plot(1:N,Xtrue,'--k')
    plot(2:N,Allmusigma(2:N,1),'k')
    plot(2:N,Allmusigma(2:N,1)+(Nsigma*Allmusigma(2:N,2)),'-.k')
    plot(2:N,Allmusigma(2:N,1)-(Nsigma*Allmusigma(2:N,2)),'-.k')
    xlabel('$N¬[-]$','Interpreter','LaTeX')
    ylabel('$x¬[value¬e.g.¬force]$','Interpreter','LaTeX')
    legend({['$x_F≠¬' num2str(round(Xtrue(1),1)) '$'],'$E[\hat{x}_{F}](N)$',['$\pm¬' ...
        num2str(Nsigma) '¬\hat{\sigma}_{F}(N)$']},...
        'Interpreter','LaTeX')
    title([ num2str(CI) '\% CI for Gumbel extrapolation to $' num2str(N) '$ points ...
        $(N)$'],...
        'Interpreter','LaTeX')
    h = annotation('textbox',[0.491,0.21,0.31,0.15],...
        'String',{'Info Sample Gumbel:',...
        ['$r≠¬' num2str(r) '$'],...
        ['$\beta=¬' num2str(beta) '$'],...
        ['$\mu≠¬' num2str(mu) '$']},...
        'VerticalAlignment','middle','HorizontalAlignment','center',...
        'FitBoxToText',...
        'on','LineWidth',0.1,'Interpreter','LaTeX');
    h.FontSize = 8;
    hold off

elseif strcmp(normalis2,normalise)==1;
    Returnlvlplot = figure;
```

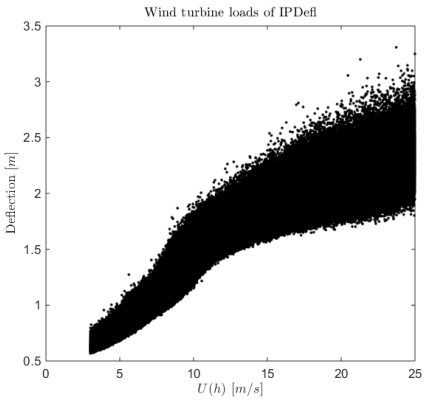
```
    hold on
    plot(1:N,zeros(N,1),'--k')
    plot(2:N,NormaliseXmu(2:N),'k')
    plot(2:N,NormaliseXup(2:N),'-.k')
    plot(2:N,NormaliseXdown(2:N),'-.k')
    xlabel('$N¬[-]$','Interpreter','LaTeX')
    ylabel(['$\epsilon¬\pm¬' num2str(Nsigma) '\varsigma¬[-]$'],'Interpreter','LaTeX')
    legend({['$x_F≠¬' num2str(round(Xtrue(1),1)) '$'],'$\epsilon_F(N)$',['$\pm¬' ...
        num2str(Nsigma) '¬\varsigma_F(N)$']},...
        'Interpreter','LaTeX')
    title([ num2str(CI) '\% CI for Gumbel extrapolation to $' num2str(N) '$ points ...
        $(N)$'],...
        'Interpreter','LaTeX')
    h = annotation('textbox',[0.491,0.21,0.31,0.15],'String',...
        {'Info Sample Gumbel:',...
        ['$r≠¬' num2str(r) '$'],...
        ['$\beta≠¬' num2str(beta) '$'],...
        ['$\mu≠¬' num2str(mu) '$']},...
        'VerticalAlignment','middle','HorizontalAlignment','center',...
        'FitBoxToText','on','LineWidth',0.1,'Interpreter','LaTeX');
    h.FontSize = 8;
    hold off
end


%manage output arguments
musigma = Allmusigma;

switch nargout
    case 0
        clear musigma
    case 1
        clear musigma
end

set(gcf, 'Color', 'w')
set(gcf, 'Units', 'centimeters')
set(gcf, 'OuterPosition', [5, 5, 9.5, 9.3])
export_fig(['C:...\Returnlvlplot_standard_N' num2str(N) '_r' num2str(r) '_CI' ...
    num2str(CI) '_' num2str(normalise) '_mu' num2str(mu) '_beta' num2str(beta) ...
    '.pdf'], '-nocrop')
end

%% Note
% This function uses *discretesample.m* and thus before use this must
% be present in the same directory
```

D

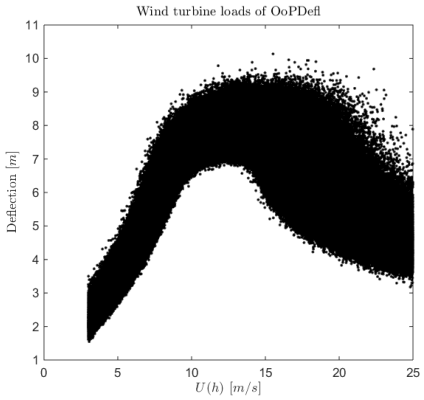
WIND SPEED OVER WIND TURBINE LOAD SIMULATIONS



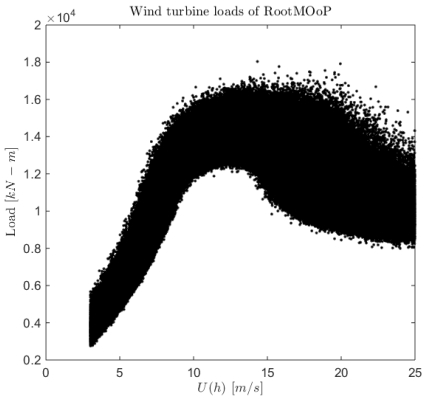
(a) TwrBsMyt SANDIA data



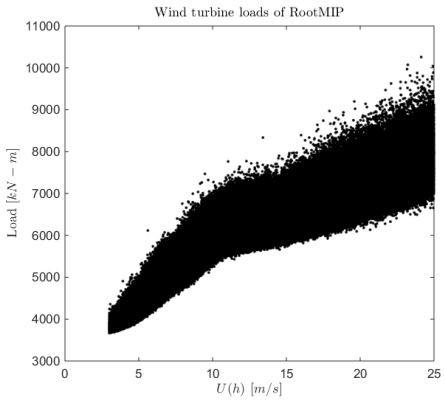
(b) IPDefl SANDIA data



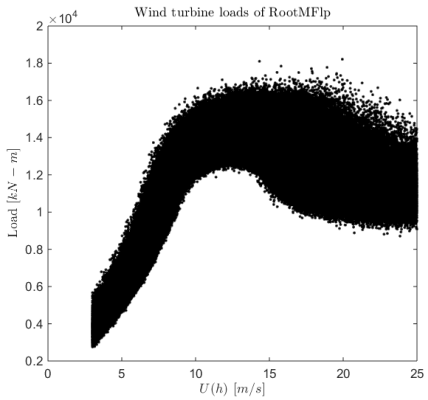
(c) OoPDefl SANDIA data



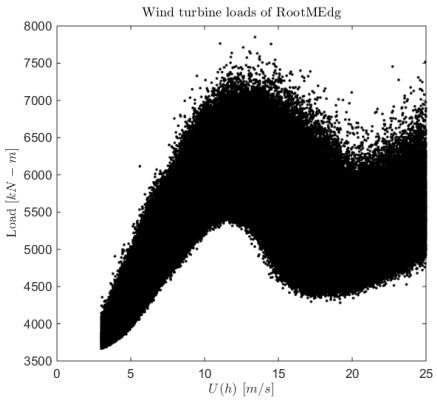
(d) RootMOoP SANDIA data



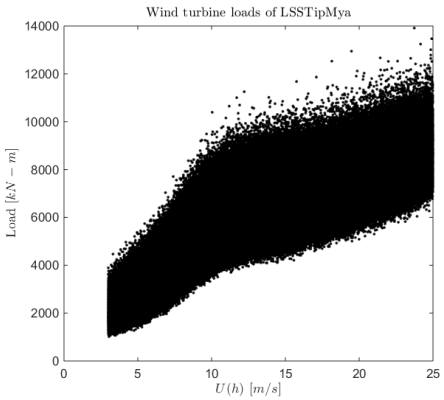
(e) RootMIP SANDIA data



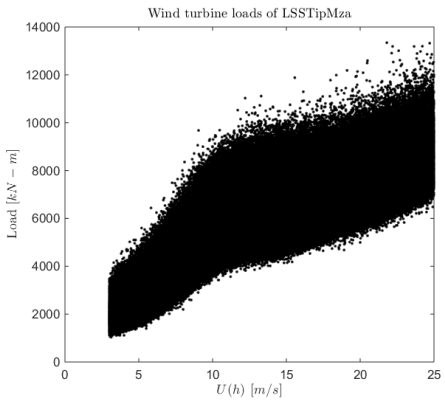
(f) RootMFlp SANDIA data



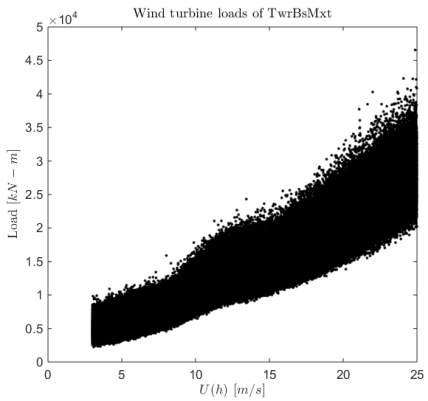
(g) RootMEdg SANDIA data



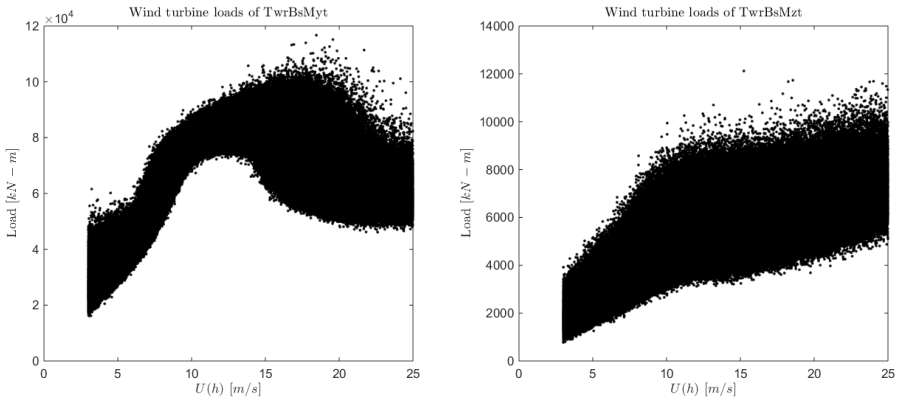
(h) LSSTipMya SANDIA data



(i) LSSTipMza SANDIA data



(j) TwrBsMxt SANDIA data



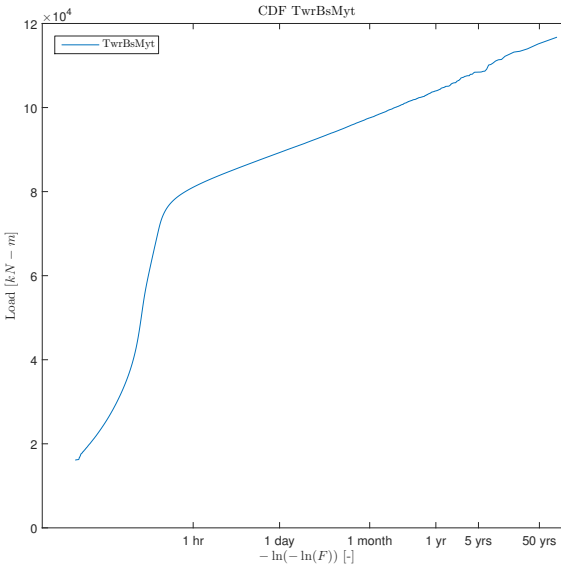
(k) TwrBsMzt SANDIA data

(l) TwrBsMzt SANDIA data

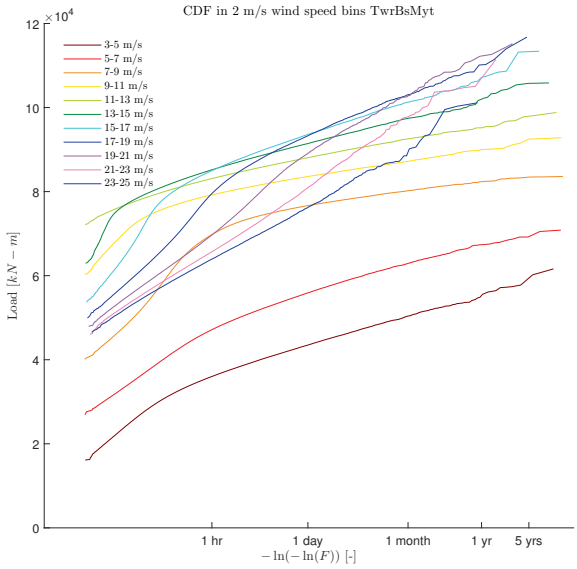
Figure D.1: Wind speed verses wind turbine load simulations derived from SANDIA data set

E

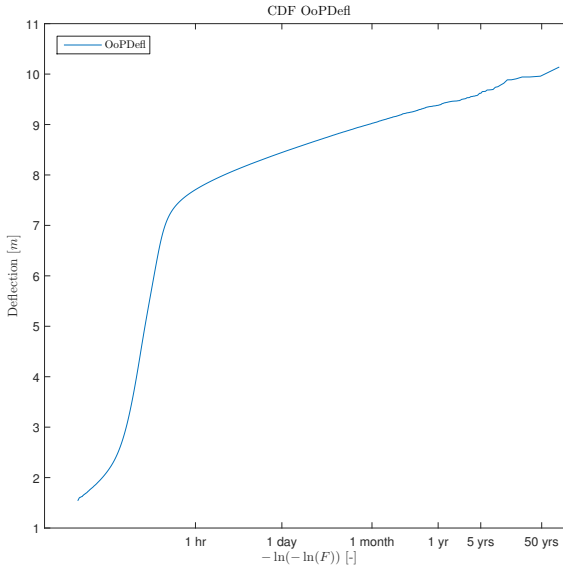
TOTAL AND BINNED SANDIA CDF PLOTS



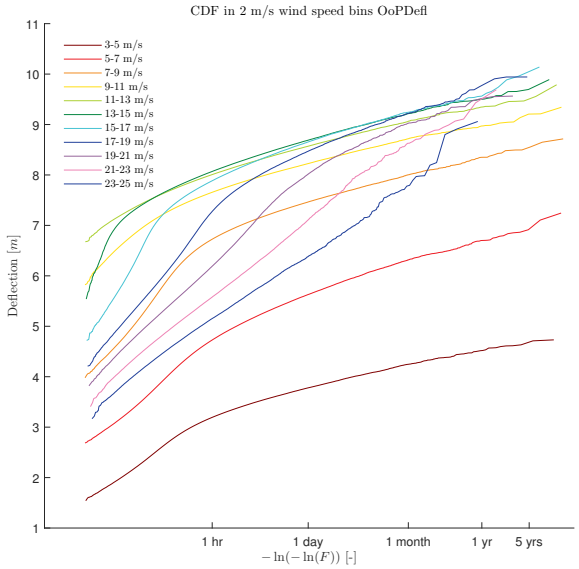
(a) CDF of TwrBsMyt SANDIA data



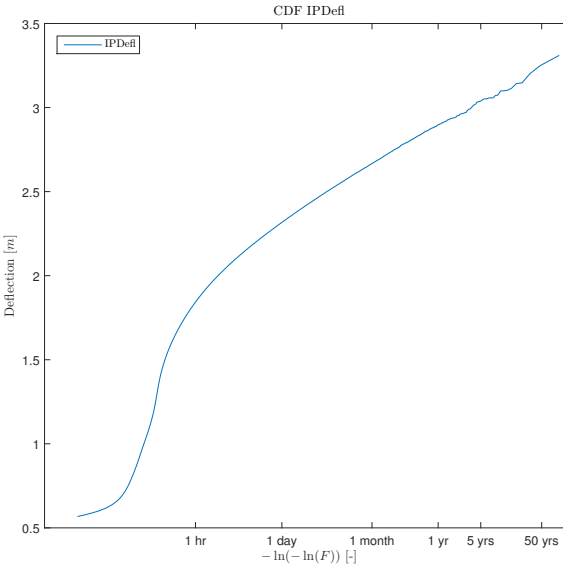
(b) CDF's of TwrBsMyt in 2 m/s wind speed bins



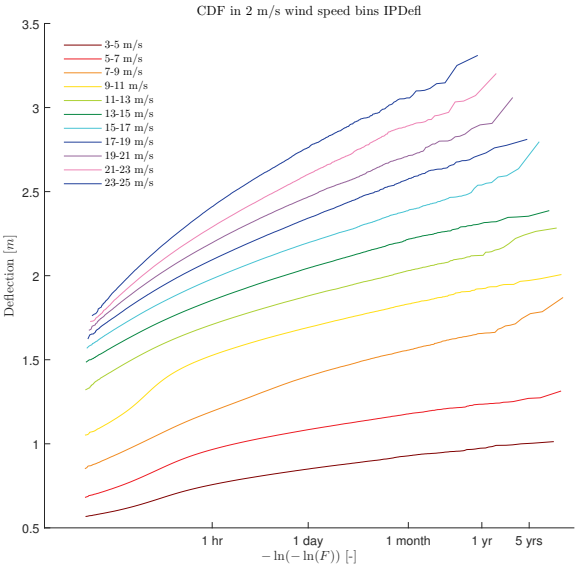
(c) CDF of OoPDefl SANDIA data



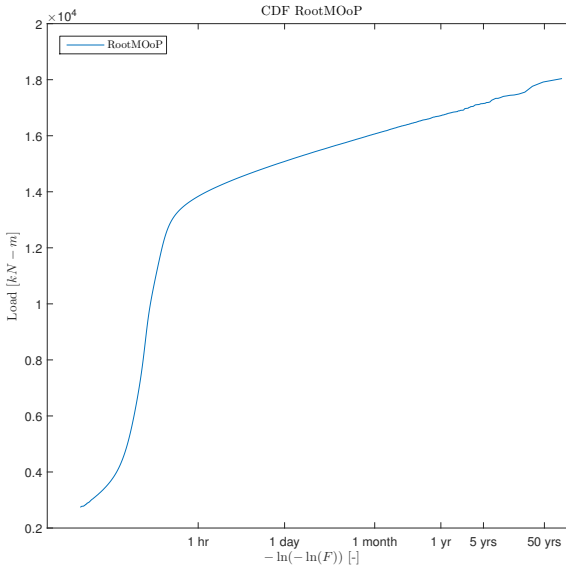
(d) CDF's of OoPDefl in 2 m/s wind speed bins



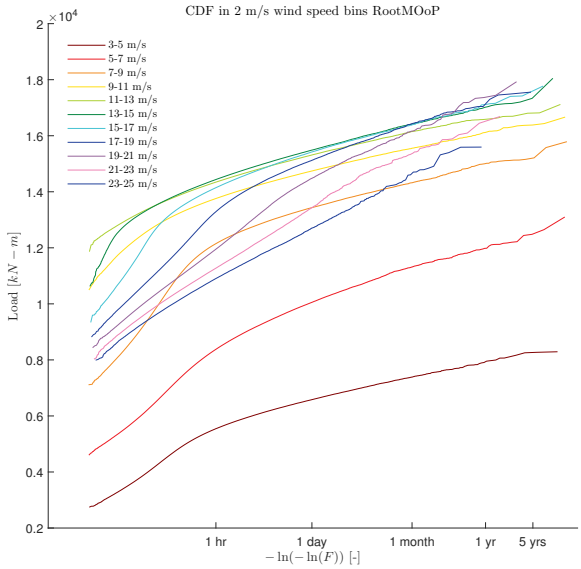
(e) CDF of IPDefl SANDIA data



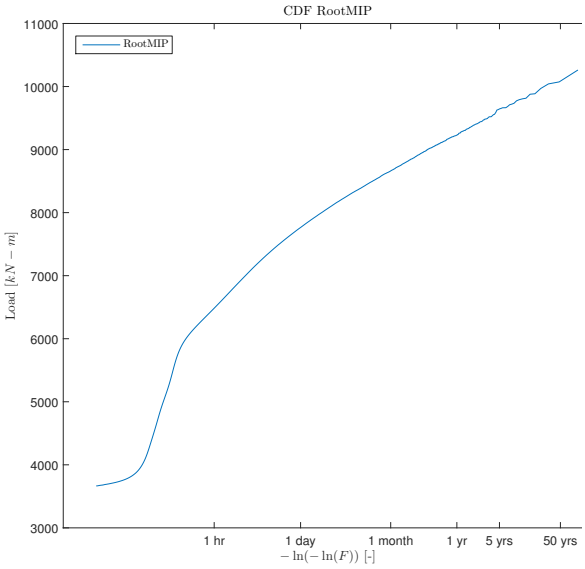
(f) CDF's of IPDefl in 2 m/s wind speed bins



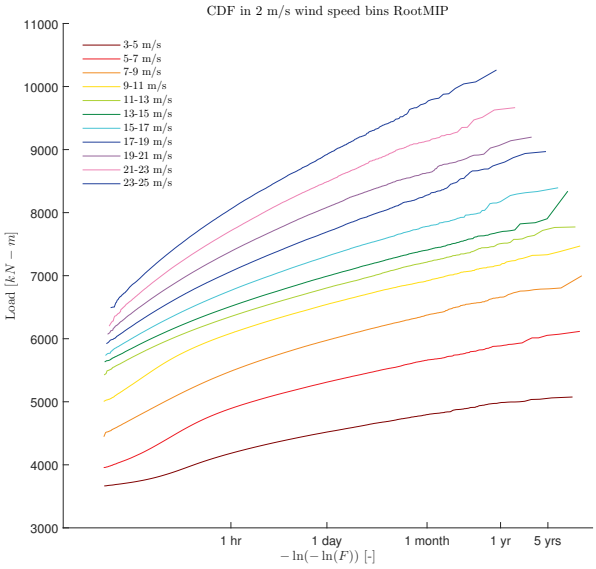
(g) CDF of RootMOop SANDIA data



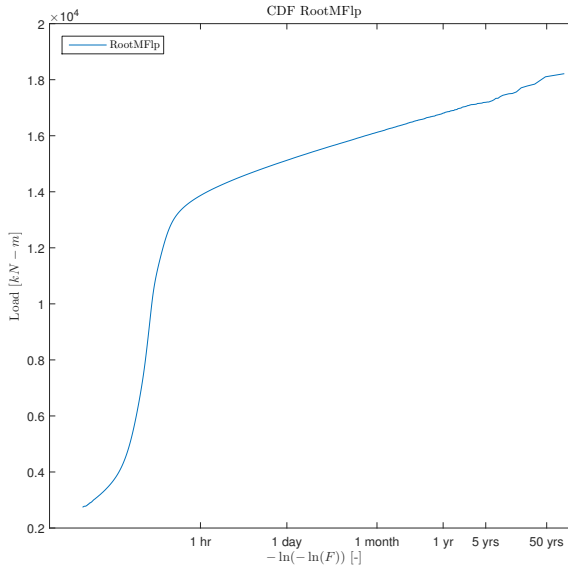
(h) CDF's of RootMOop in 2 m/s wind speed bins



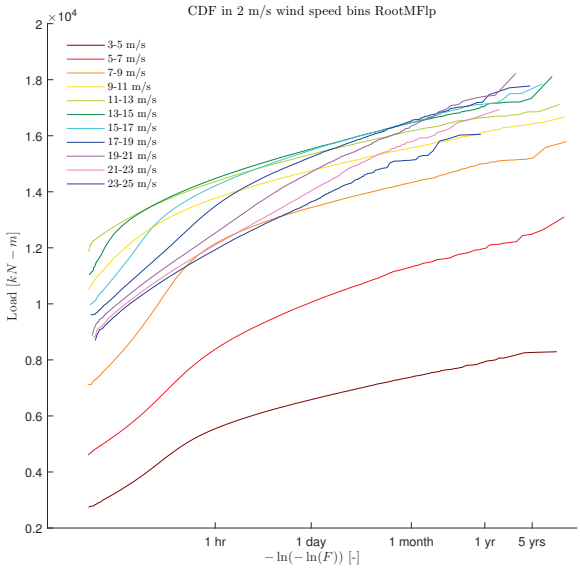
(i) CDF of RootMIP SANDIA data



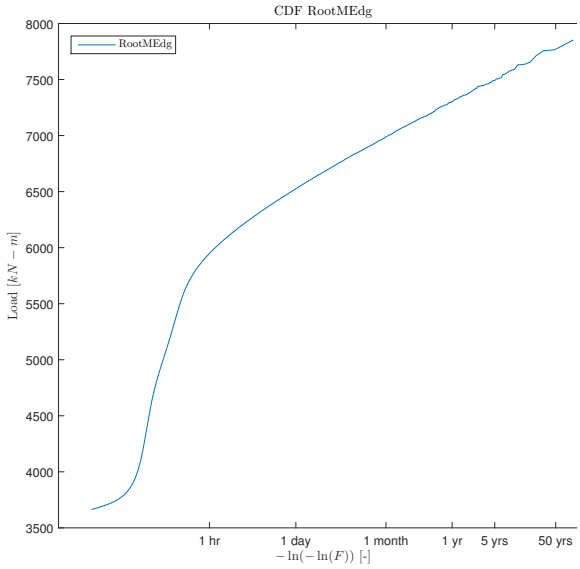
(j) CDF's of RootMIP in 2 m/s wind speed bins



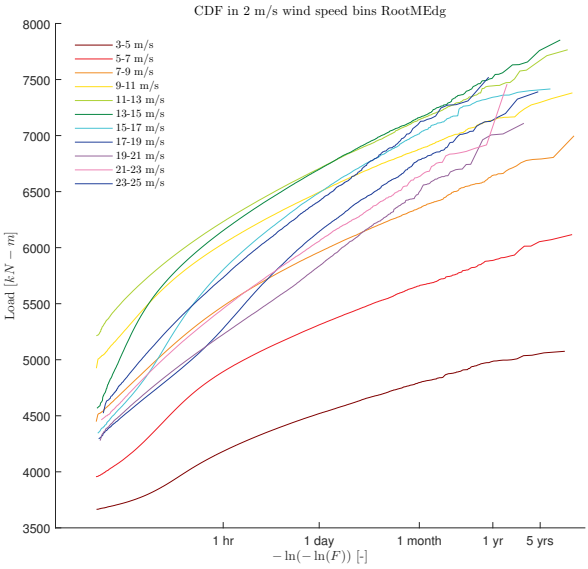
(k) CDF of RootMFlp SANDIA data



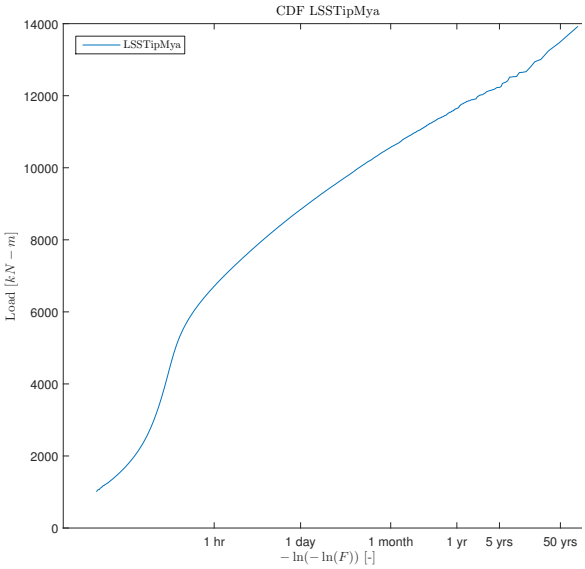
(l) CDF's of RootMFlp in 2 m/s wind speed bins



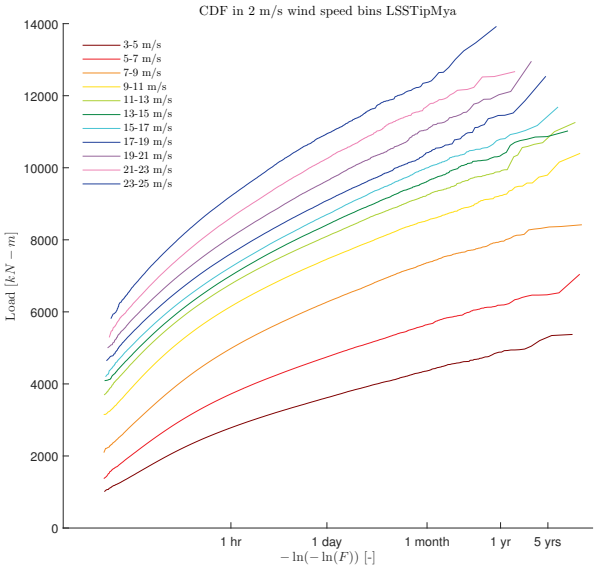
(m) CDF of RootMEdg SANDIA data



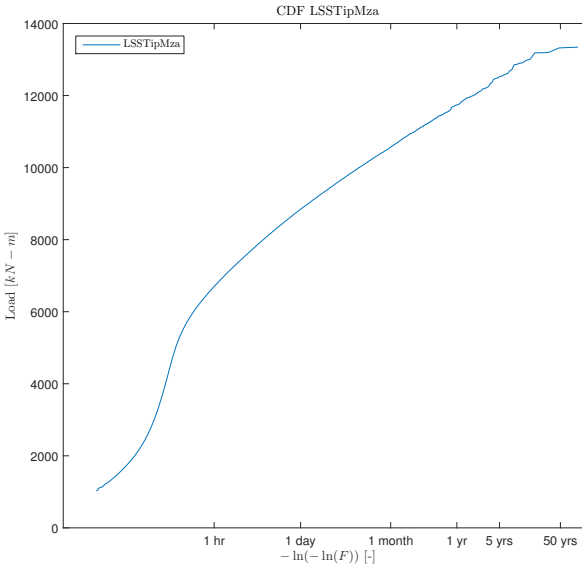
(n) CDF's of RootMEdg in 2 m/s wind speed bins



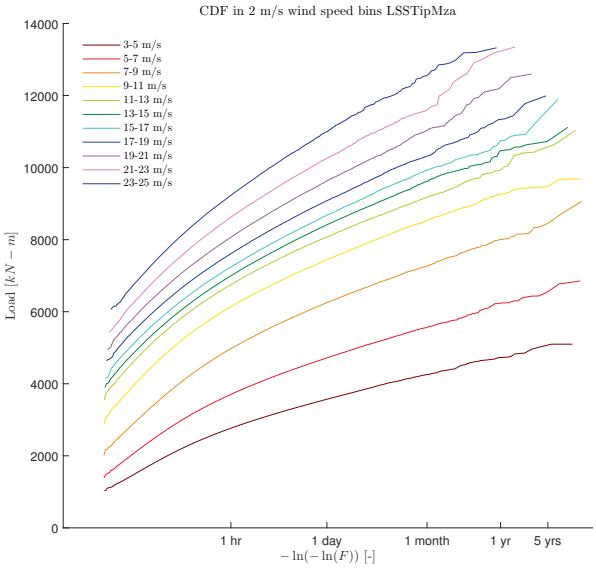
(o) CDF of LSSTipMya SANDIA data



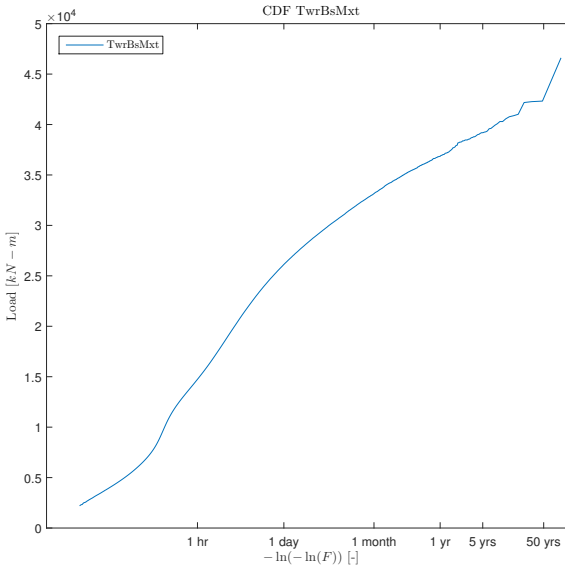
(p) CDF's of LSSTipMya in 2 m/s wind speed bins



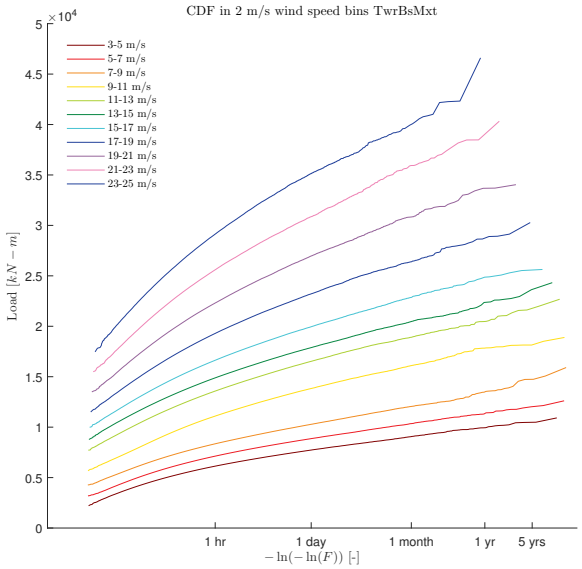
(q) CDF of LSSTipMza SANDIA data



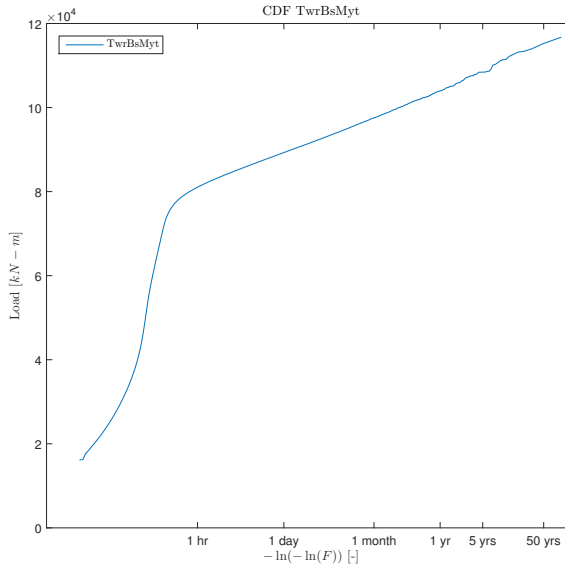
(r) CDF's of LSSTipMza in 2 m/s wind speed bins



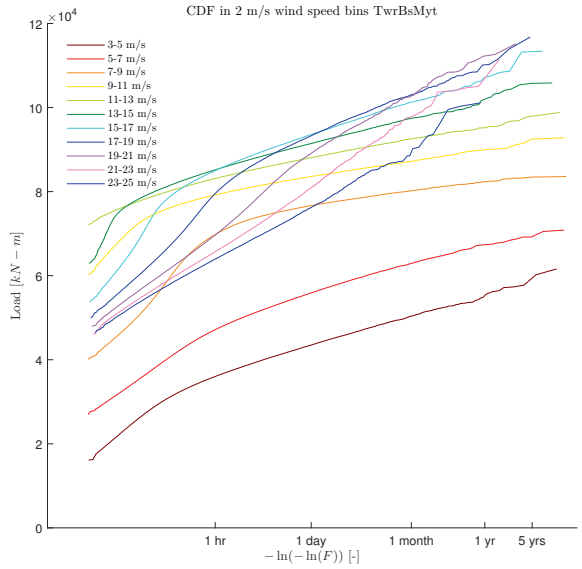
(s) CDF of TwrBsMxt SANDIA data



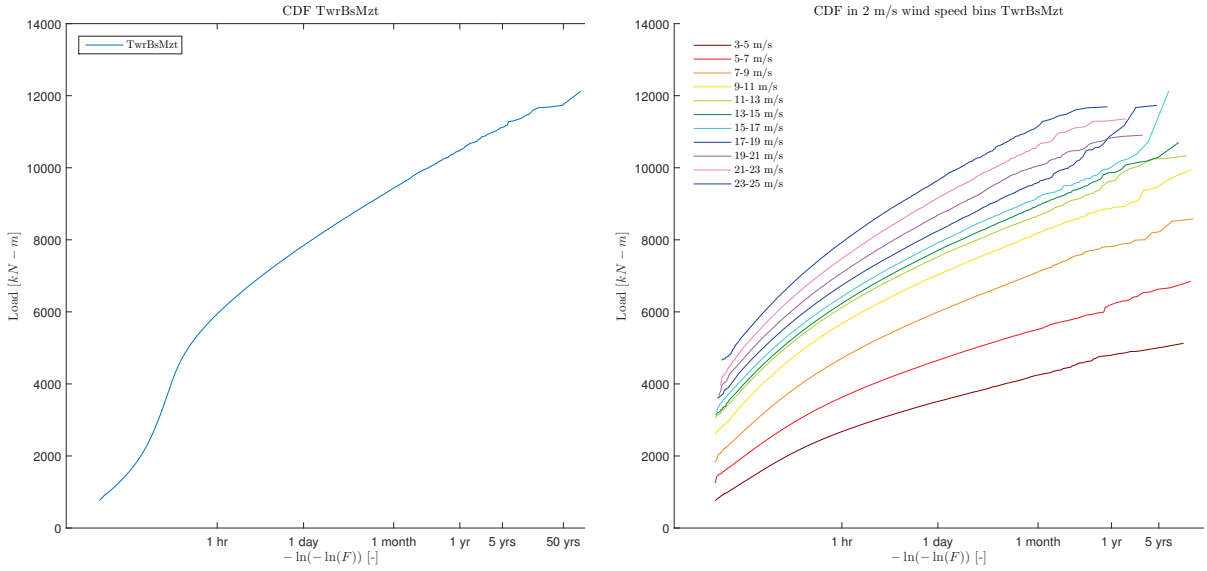
(t) CDF's of TwrBsMxt in 2 m/s wind speed bins



(u) CDF of TwrBsMyt SANDIA data



(v) CDF's of TwrBsMyt in 2 m/s wind speed bins



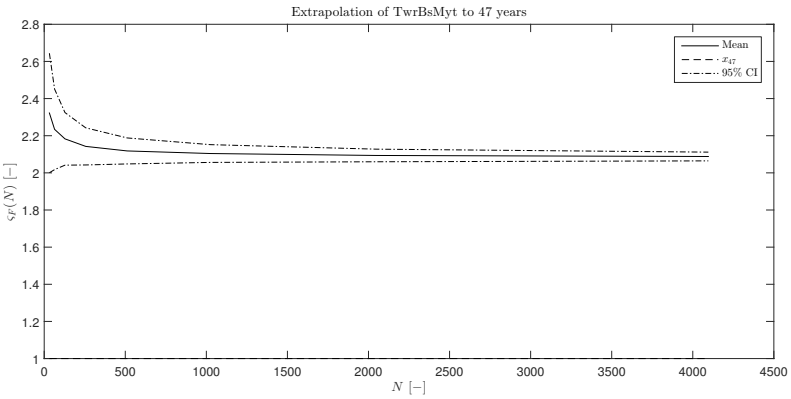
(w) CDF of TwrBsMzt SANDIA data

(x) CDF's of TwrBsMzt in 2 m/s wind speed bins

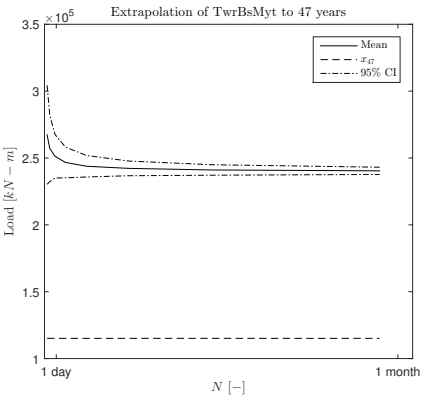
Figure E.1: CDF and 2 m/s binned CDF's of wind turbine load simulation from SANDIA data

F

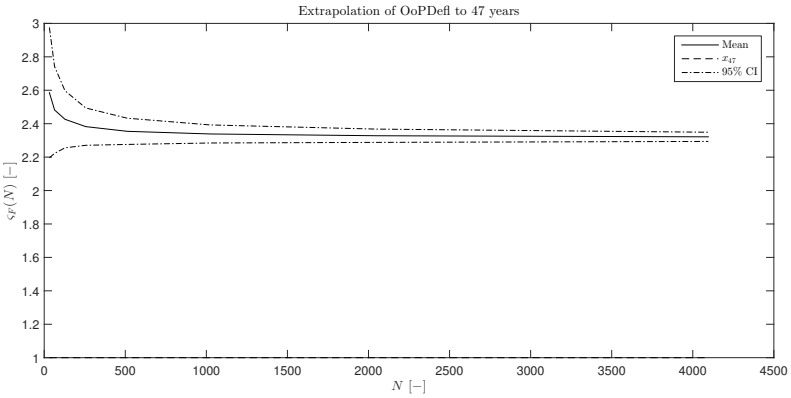
Uncertainty in Basic Wind Turbine Load Estimations



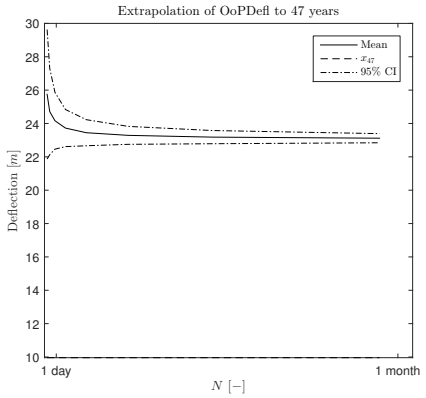
(a) Normalised uncertainty in TwrBsMyt estimates



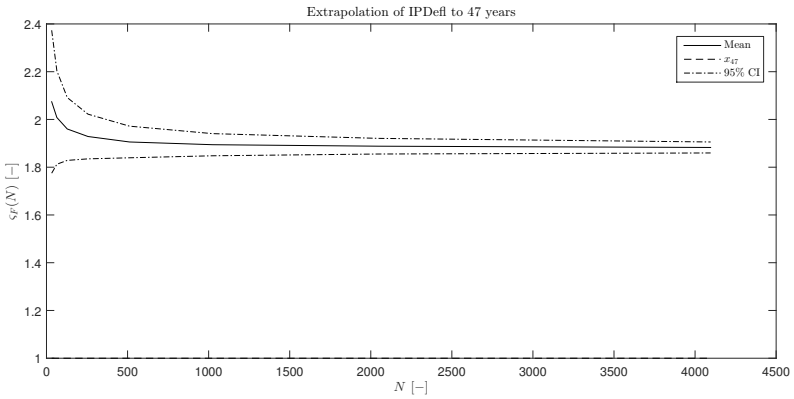
(b) Uncertainty over simulation length TwrBsMyt



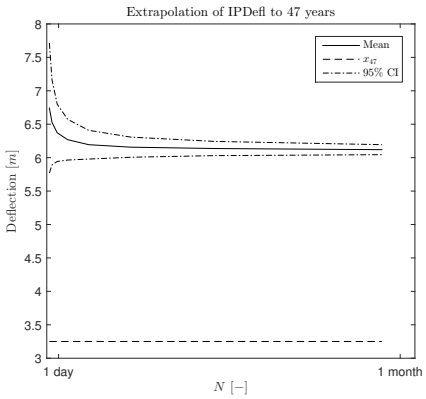
(c) Normalised uncertainty in OoPDefl estimates



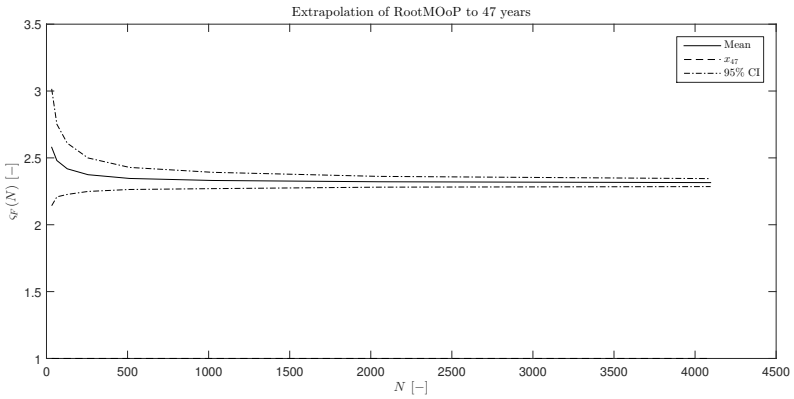
(d) Uncertainty over simulation length OoPDefl



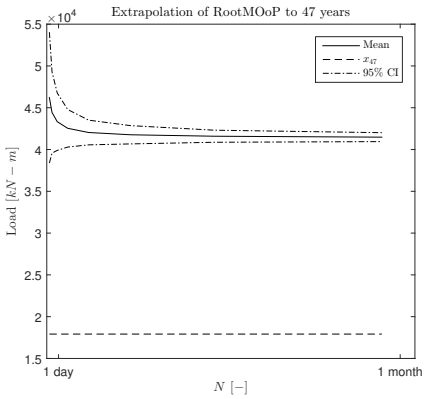
(e) Normalised uncertainty in IPDefl estimates



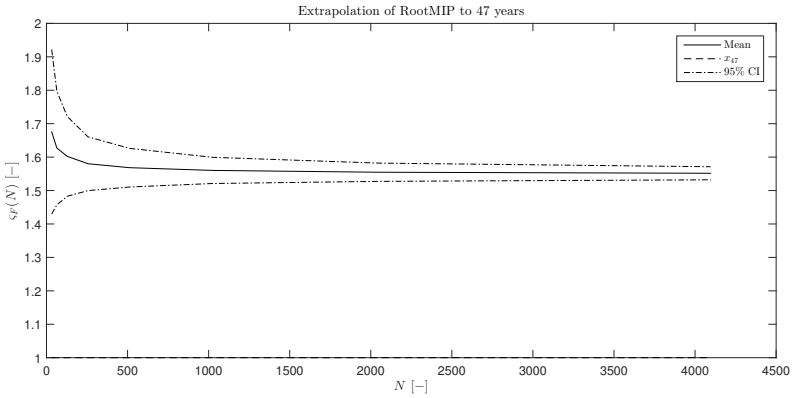
(f) Uncertainty over simulation length IPDefl



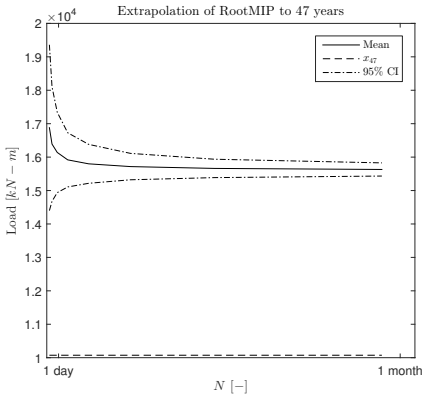
(g) Normalised uncertainty in RootMOop estimates



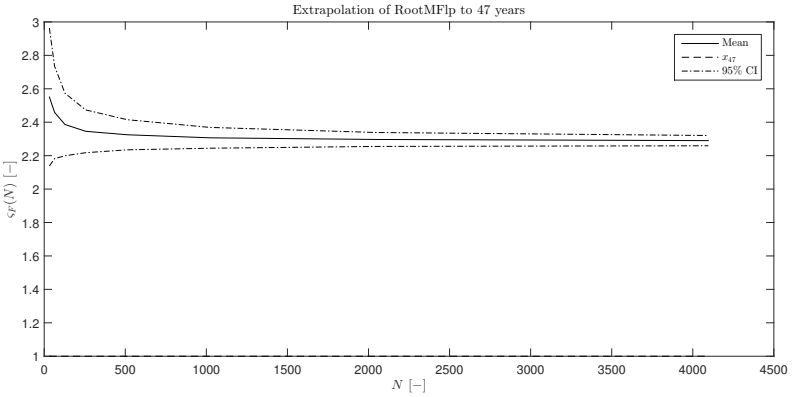
(h) Uncertainty over simulation length RootMOop



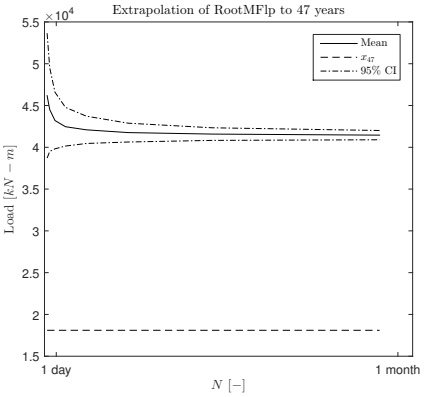
(i) Normalised uncertainty in RootMIP estimates



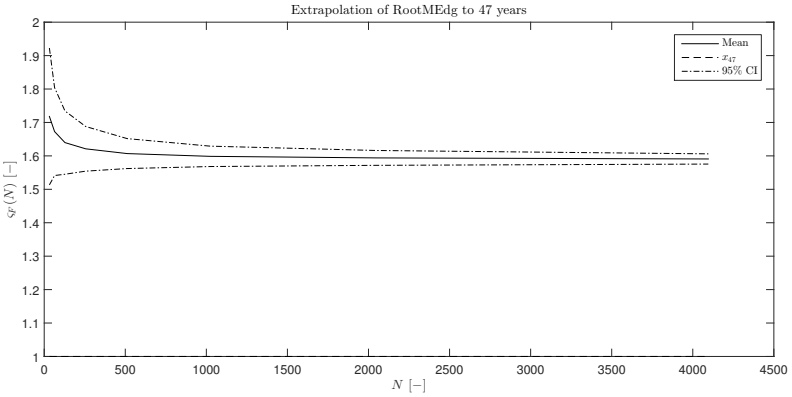
(j) Uncertainty over simulation length RootMIP



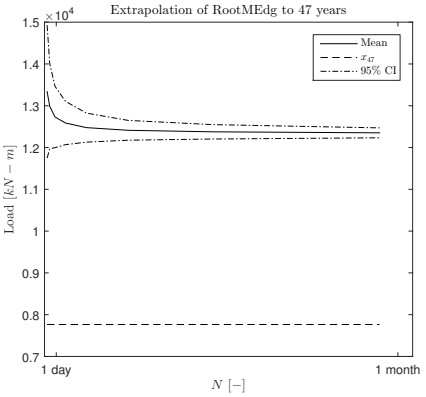
(k) Normalised uncertainty in RootMFlp estimates



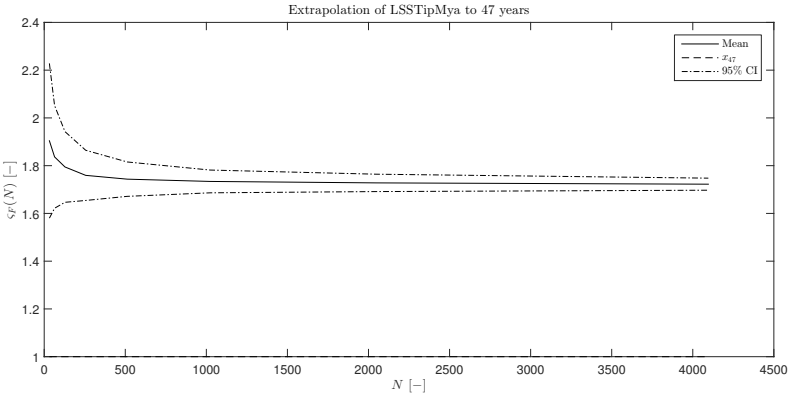
(l) Uncertainty over simulation length RootMFlp



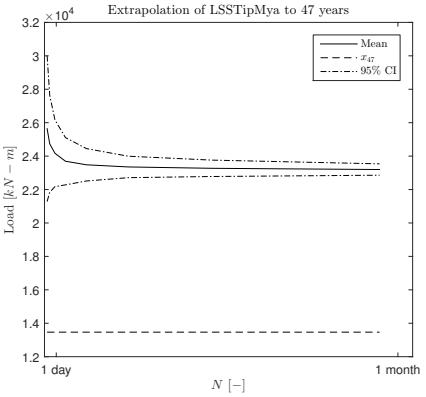
(m) Normalised uncertainty in RootMEdg estimates



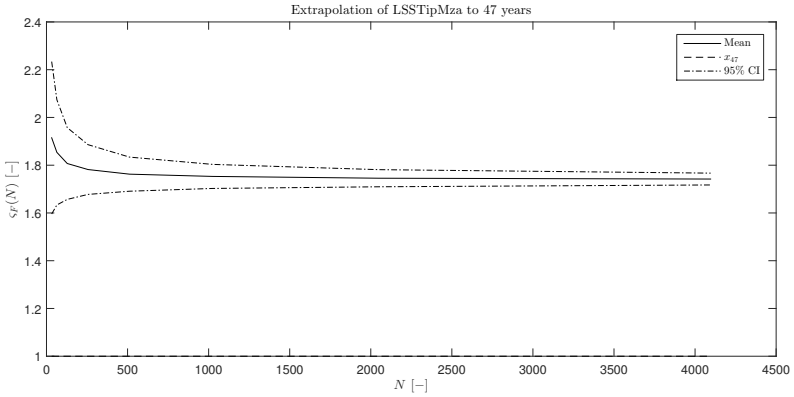
(n) Uncertainty over simulation length RootMEdg



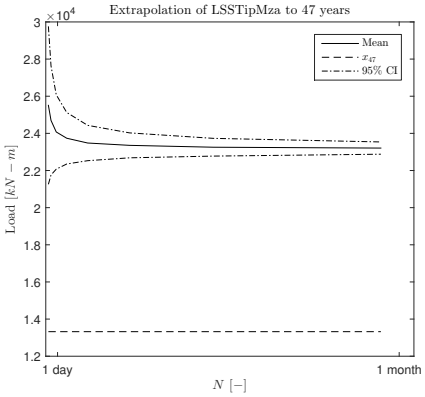
(o) Normalised uncertainty in LSSTipMya estimates



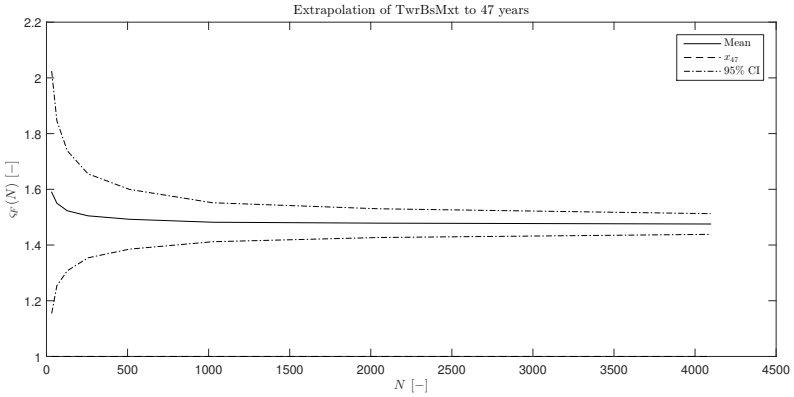
(p) Uncertainty over simulation length LSSTipMya



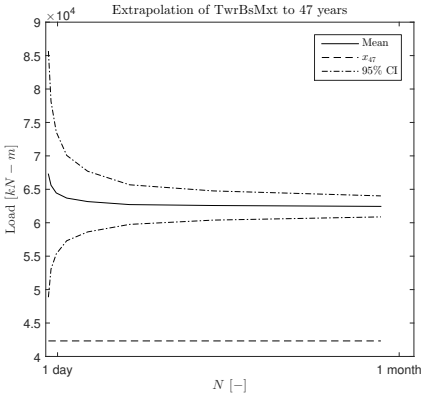
(q) Normalised uncertainty in LSSTipMza estimates



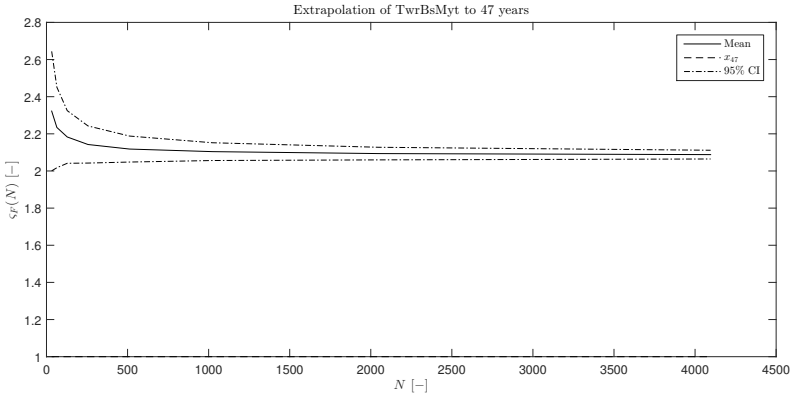
(r) Uncertainty over simulation length LSSTipMza



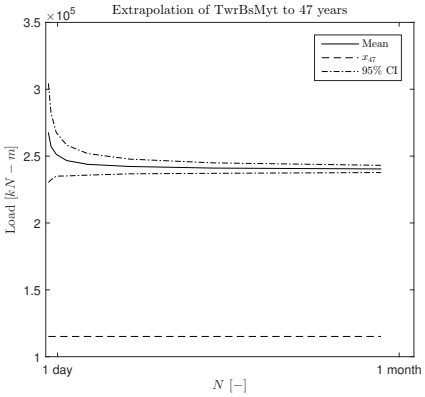
(s) Normalised uncertainty in TwrBsMxt estimates



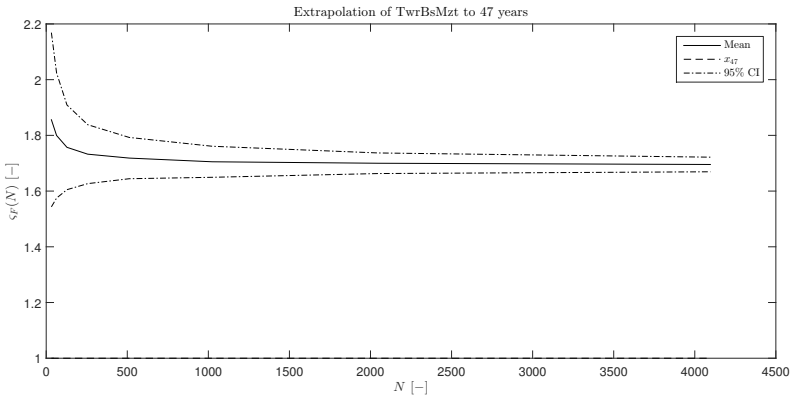
(t) Uncertainty over simulation length TwrBsMxt



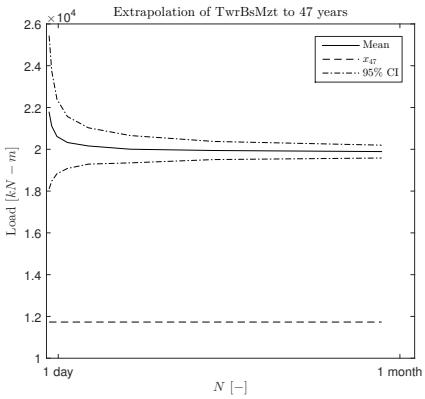
(u) Normalised uncertainty in TwrBsMyt estimates



(v) Uncertainty over simulation length in TwrBsMyt



(w) Normalised uncertainty in TwrBsMzt estimates

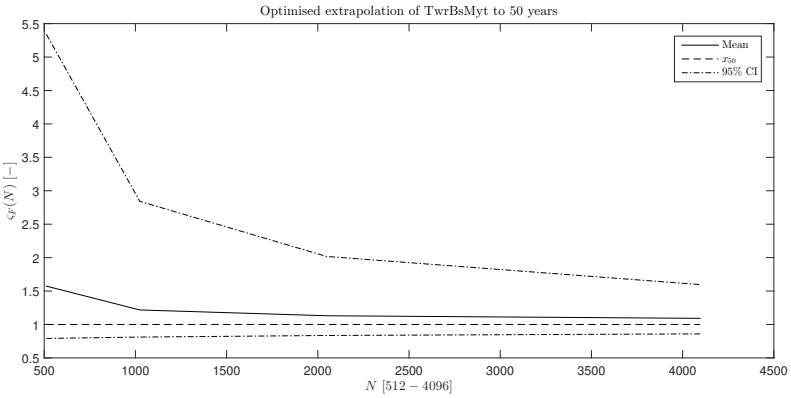


(x) Uncertainty over simulation length in TwrBsMzt

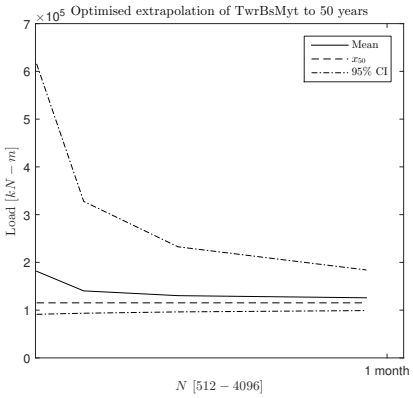
Figure F.1: Uncertainty based on a number of sub set simulations: $N = 32, 64, 128, \ldots, 4096$

G

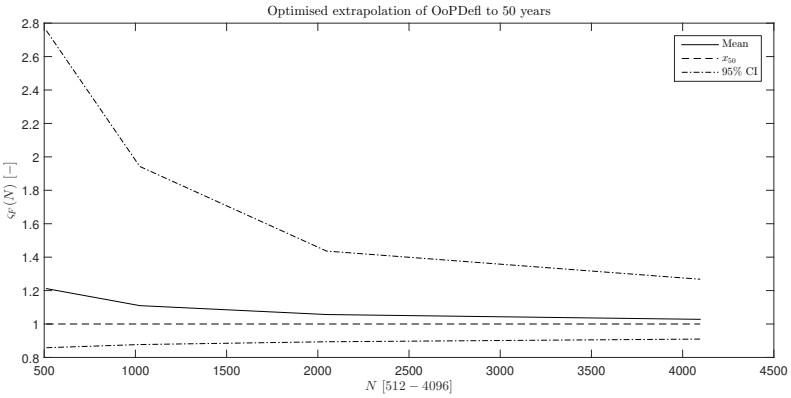
Uncertainty in Optimised Wind Turbine Load Estimations



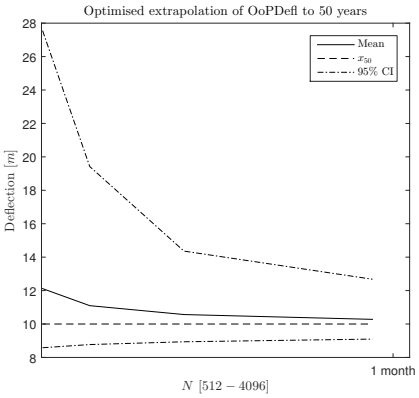
(a) Optimised normalised uncertainty in TwrBsMyt estimates



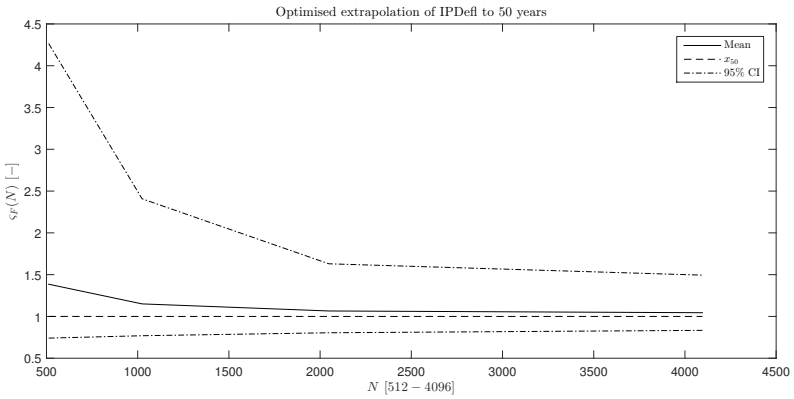
(b) Optimised uncertainty over simulation length in TwrBsMyt



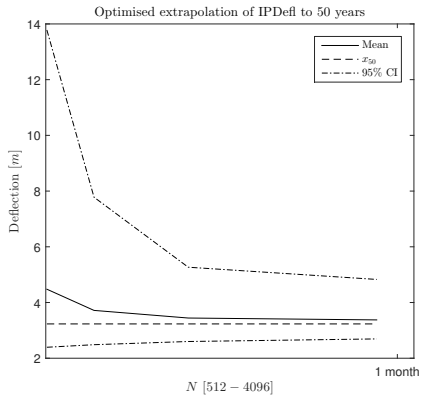
(c) Optimised normalised uncertainty in OoPDefl estimates



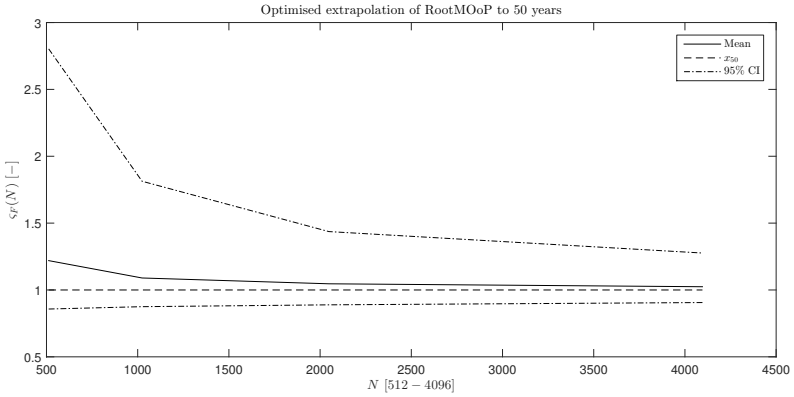
(d) Optimised uncertainty over simulation length in OoPDefl



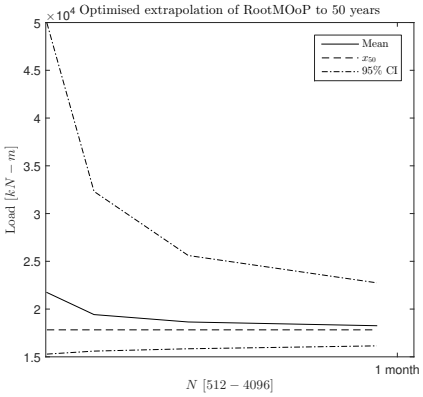
(e) Optimised normalised uncertainty in IPDefl estimates



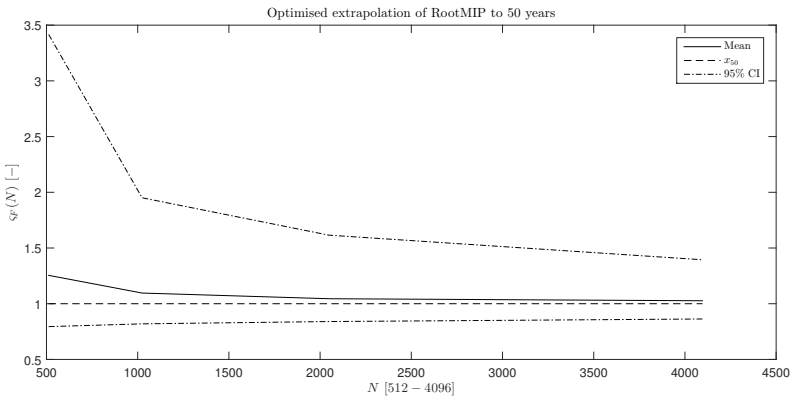
(f) Optimised uncertainty over simulation length in IPDefl



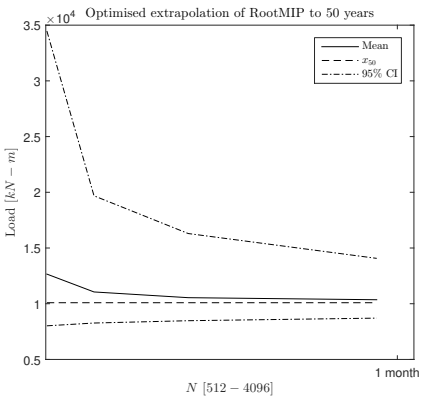
(g) Optimised normalised uncertainty in RootMOop estimates



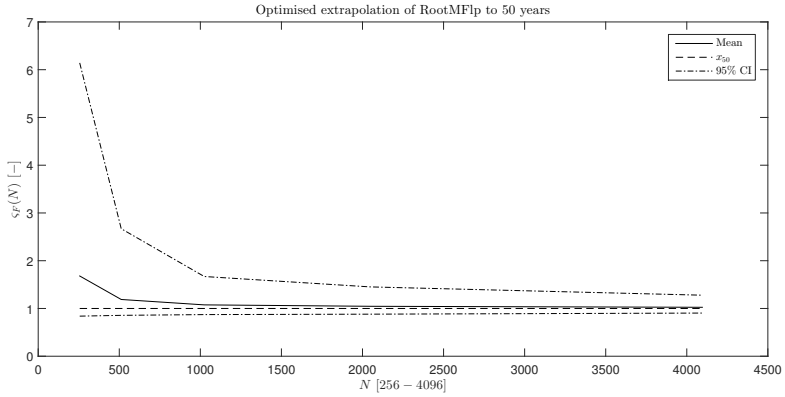
(h) Optimised uncertainty over simulation length in RootMOop



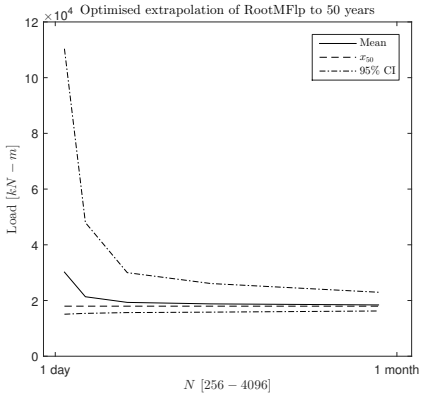
(i) Optimised normalised uncertainty in RootMIP estimates



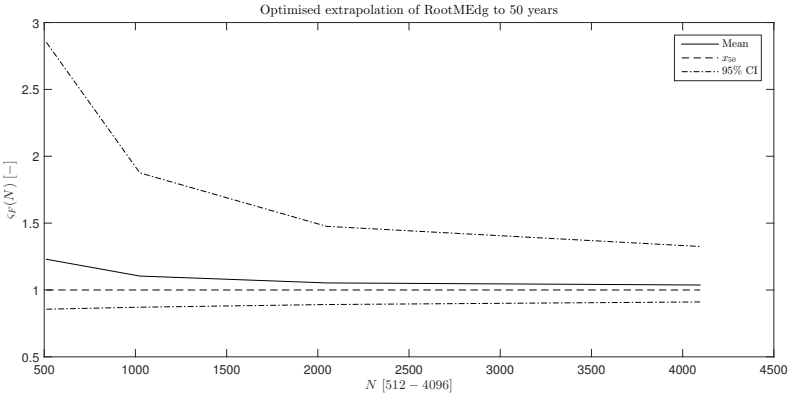
(j) Optimised uncertainty over simulation length in RootMIP



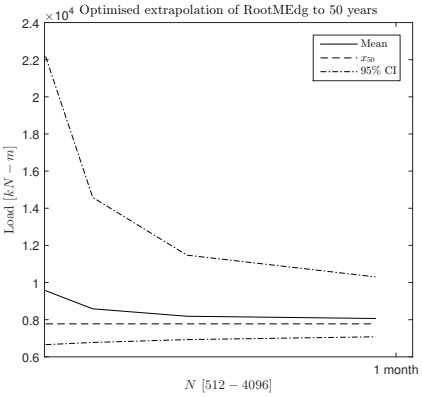
(k) Optimised normalised uncertainty in RootMFlp estimates



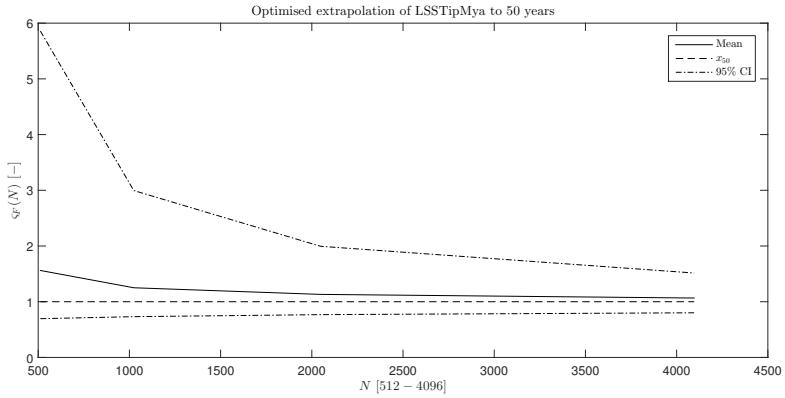
(l) Optimised uncertainty over simulation length in RootMFlp



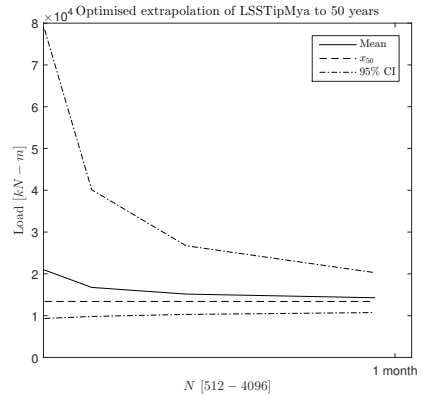
(m) Optimised normalised uncertainty in RootMEdg estimates



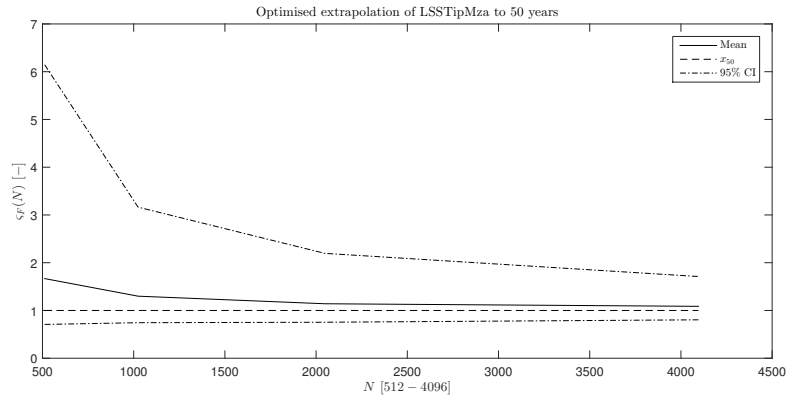
(n) Optimised uncertainty over simulation length in RootMEdg



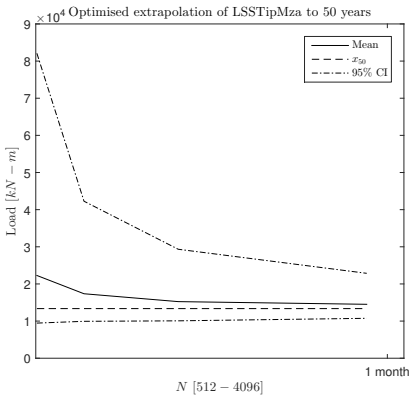
(o) Optimised normalised uncertainty in LSSTipMya estimates



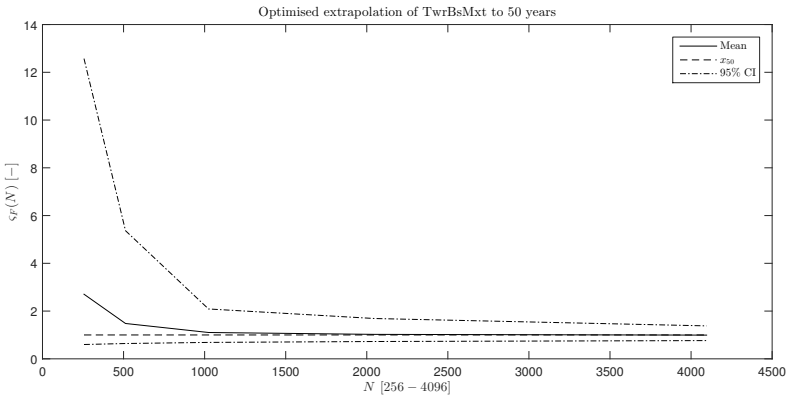
(p) Optimised uncertainty over simulation length in LSSTipMya



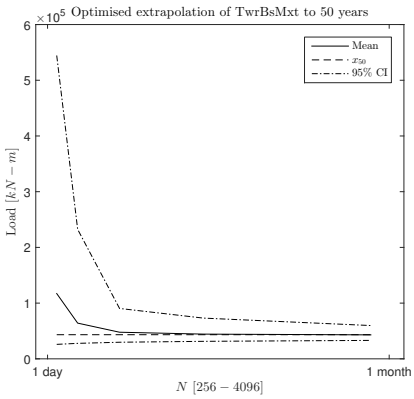
(q) Optimised normalised uncertainty in LSSTipMza estimates



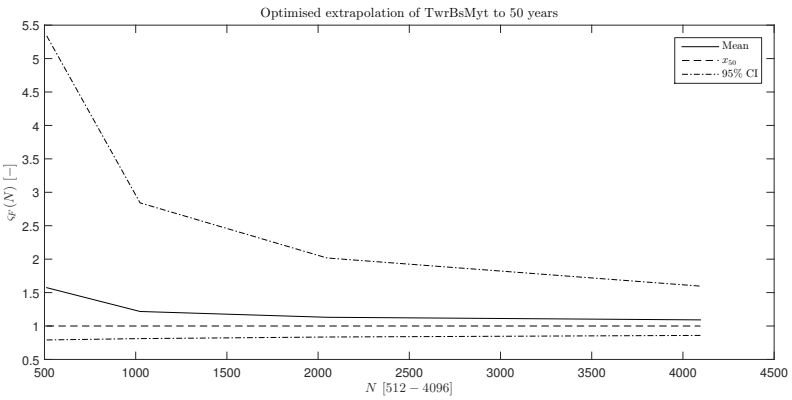
(r) Optimised uncertainty over simulation length in LSSTipMza



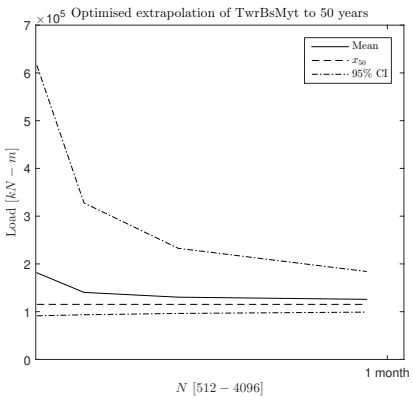
(s) Optimised normalised uncertainty in TwrBsMxt estimates



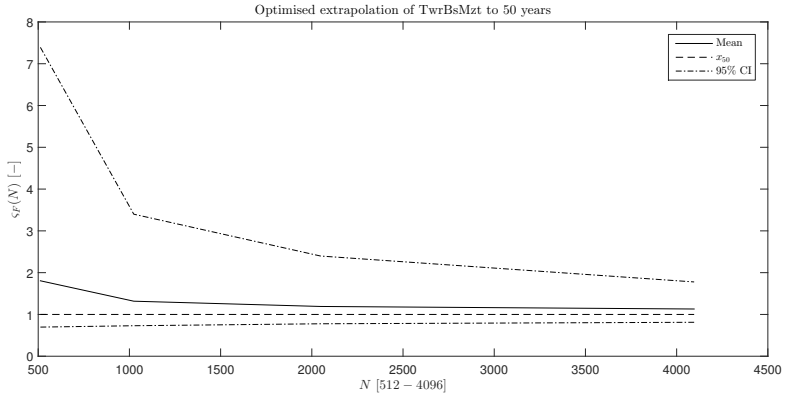
(t) Optimised uncertainty over simulation length in TwrBsMxt



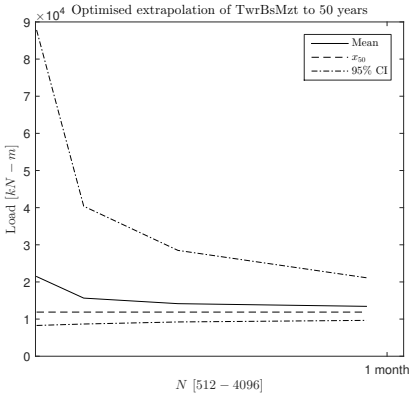
(u) Optimised normalised uncertainty in TwrBsMxt estimates



(v) Optimised uncertainty over simulation length in TwrBsMyt



(w) Optimised normalised uncertainty in TwrBsMzt estimates

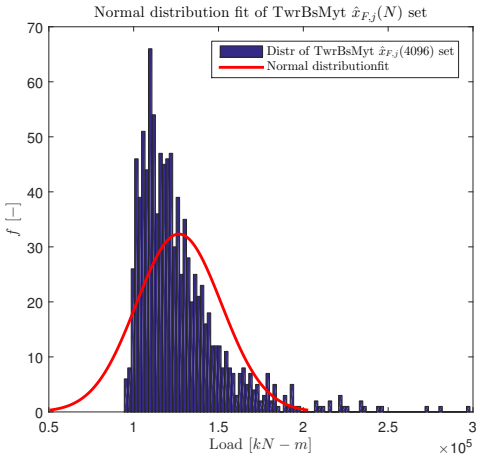


(x) Optimised uncertainty over simulation length in TwrBsMzt

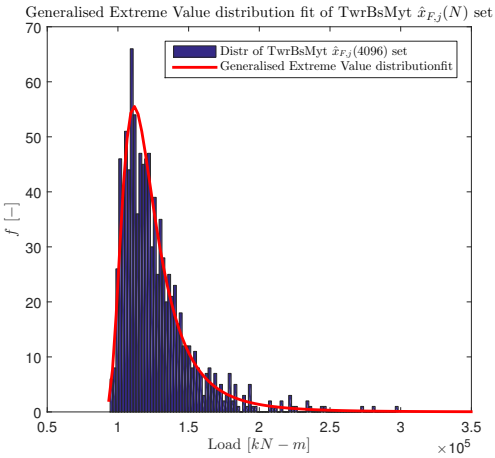
Figure G.1: Optimised uncertainty based on a number of sub set simulations: $N = 256, 512, \ldots, 4096$

H

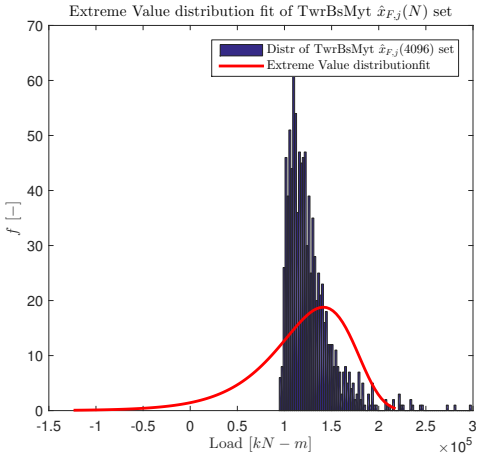
Fits of 50 year Load Estimate Distribution



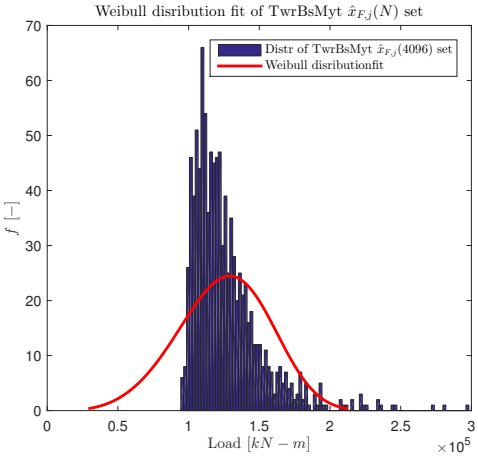
(a) Normal distribution fit for extreme value estimates fitted from N = 4096 TwrBsMyt values



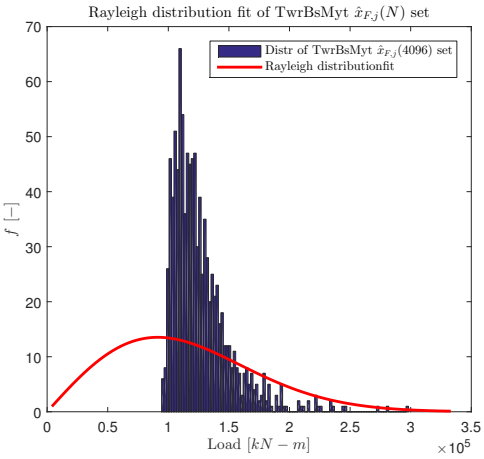
(b) GEV fit for extreme value estimates fitted from N = 4096 TwrBsMyt values



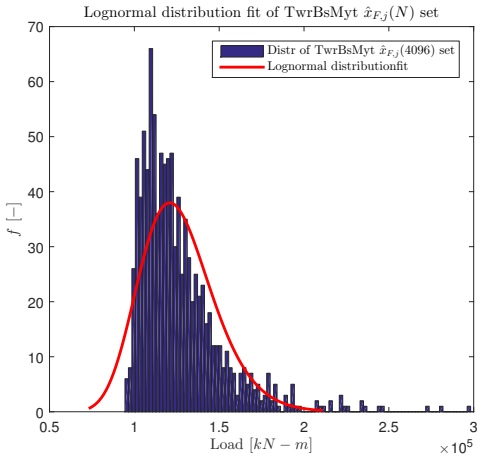
(c) Extreme value distribution fit for extreme value estimates fitted from N = 4096 TwrBsMyt values



(d) Weibull distribution fit for extreme value estimates fitted from N = 4096 TwrBsMyt values



(e) Rayleigh distribution fit for extreme value estimates fitted from N = 4096 TwrBsMyt values

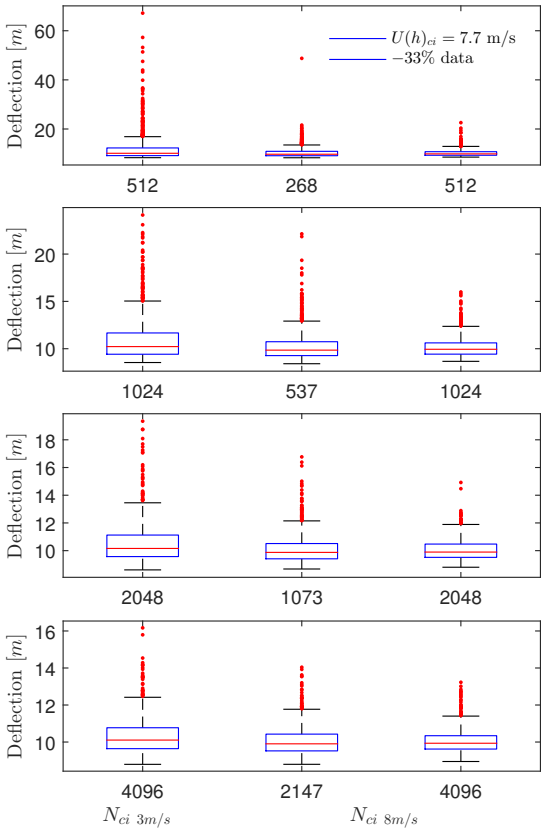


(f) Lognormal distribution fit for extreme value estimates fitted from N = 4096 TwrBsMyt values

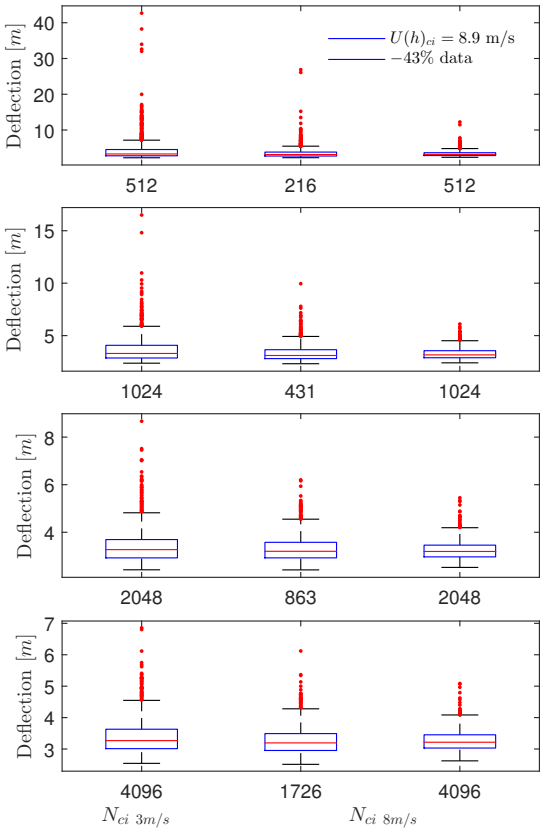
Figure H.1: Estimate distributions fits from the extreme simulation set of $N = 4096$

I

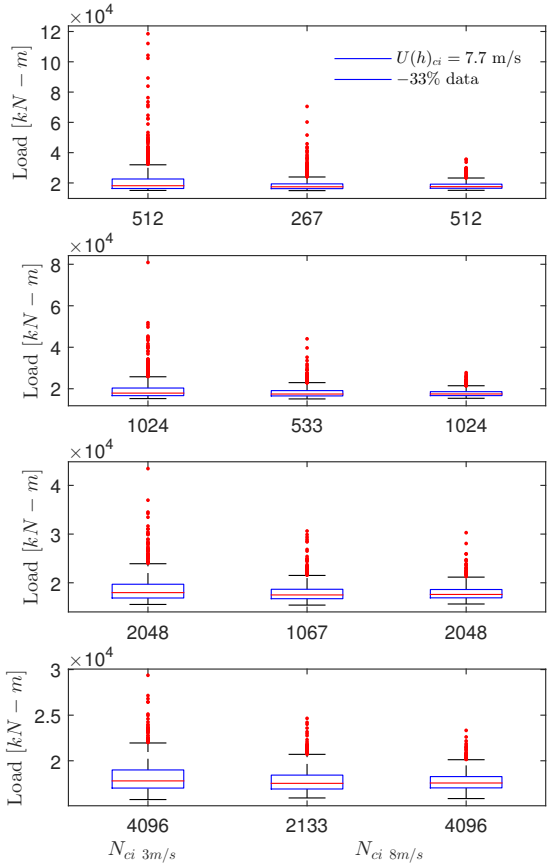
Uncertainty in 50-year Load Estimate Distribution Sampled by Elevated Cut-in Wind Speed



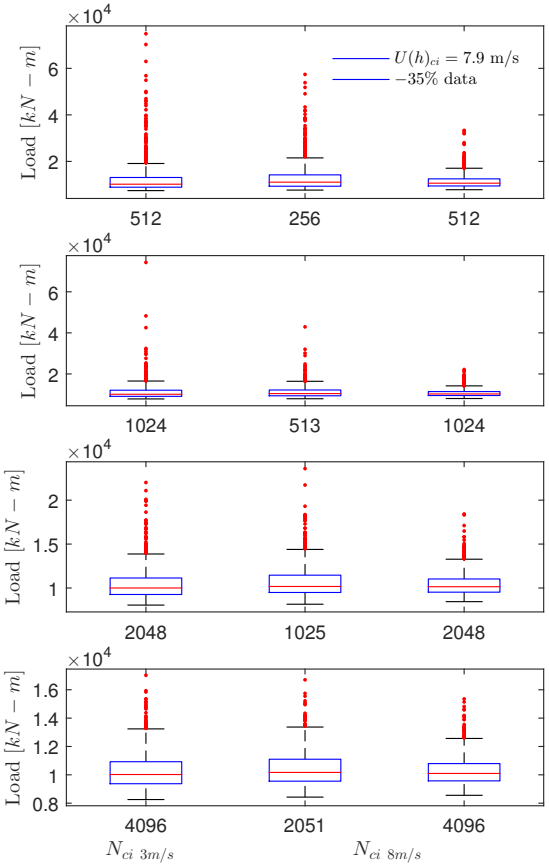
(a) Distribution of 50 year extreme OoPDefl values originating from N = 512, 1024, 2048 and 4069 compared to the estimate distribution from elevated cut-in wind speeds with equal and less number of simulations accompanied with 33% of saved data calculations



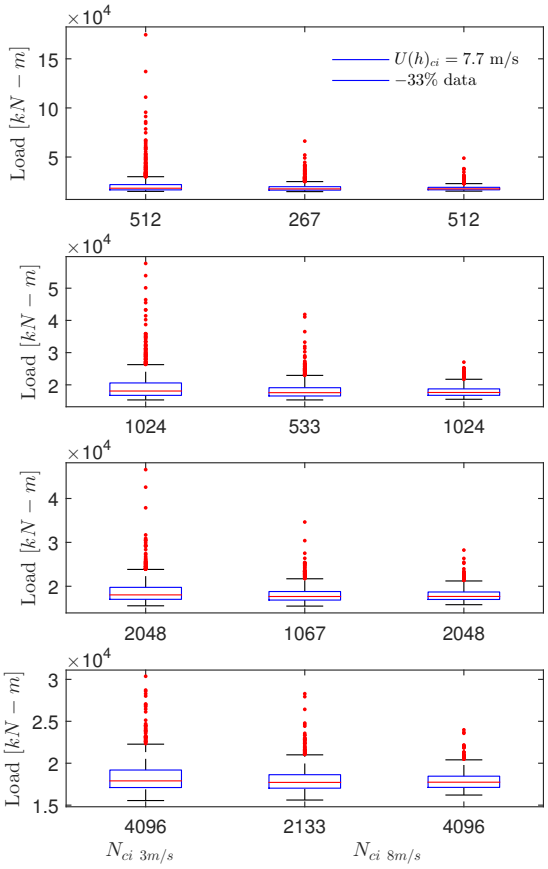
(b) Distribution of 50 year extreme IPDefl values originating from N = 512, 1024, 2048 and 4069 compared to the estimate distribution from elevated cut-in wind speeds with equal and less number of simulations accompanied with 43% of saved data calculations



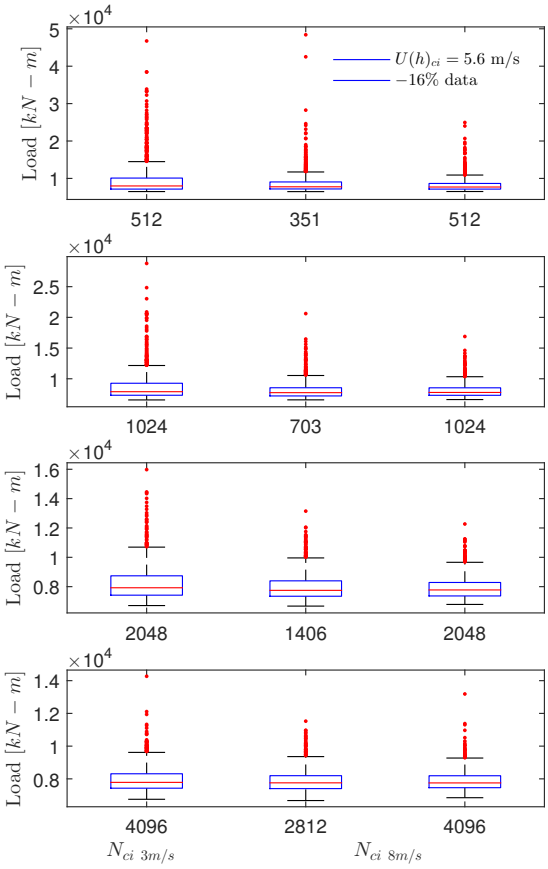
(c) Distribution of 50 year extreme RootMOoP values originating from N = 512, 1024, 2048 and 4069 compared to the estimate distribution from elevated cut-in wind speeds with equal and less number of simulations accompanied with 33% of saved data calculations



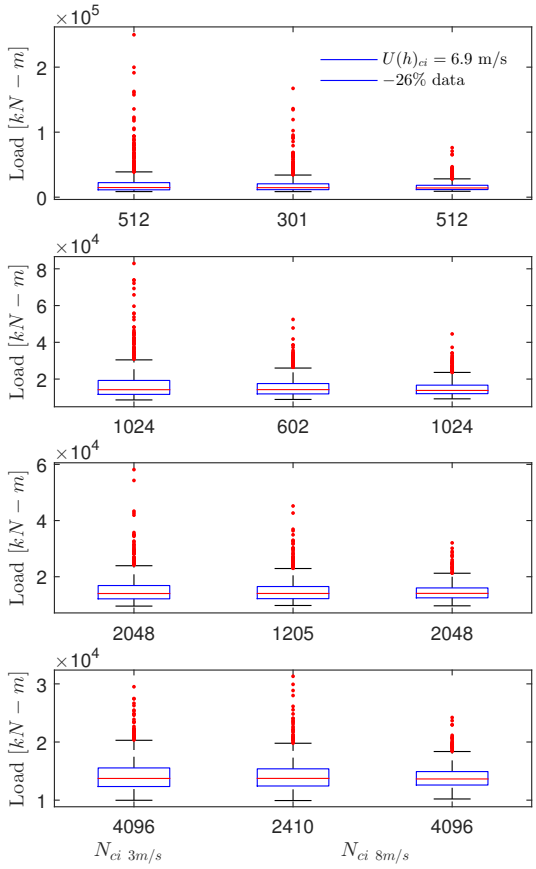
(d) Distribution of 50 year extreme RootMIP values originating from N = 512, 1024, 2048 and 4069 compared to the distribution from elevated cut-in wind speeds with equal and less number of simulations accompanied with 35% of saved data calculations



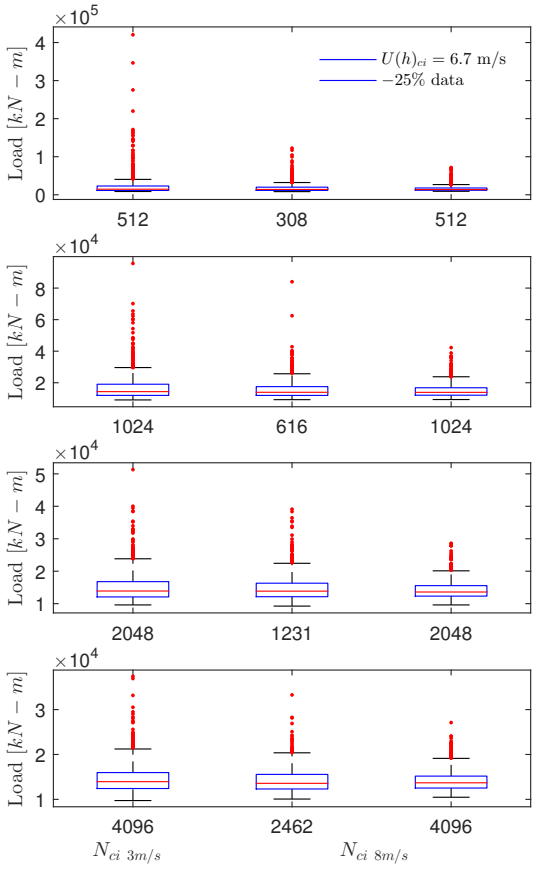
(e) Distribution of 50 year extreme RootMFlp values originating from N = 512, 1024, 2048 and 4069 compared to the estimate distribution from elevated cut-in wind speeds with equal and less number of simulations accompanied with 33% of saved data calculations



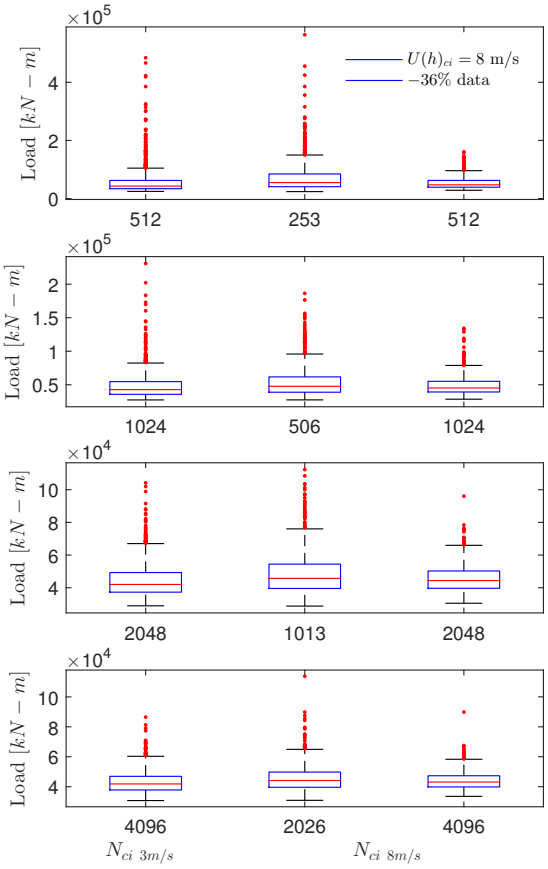
(f) Distribution of 50 year extreme RootMEdg values originating from N = 512, 1024, 2048 and 4069 compared to the estimate distribution from elevated cut-in wind speeds with equal and less number of simulations accompanied with 16% of saved data calculations



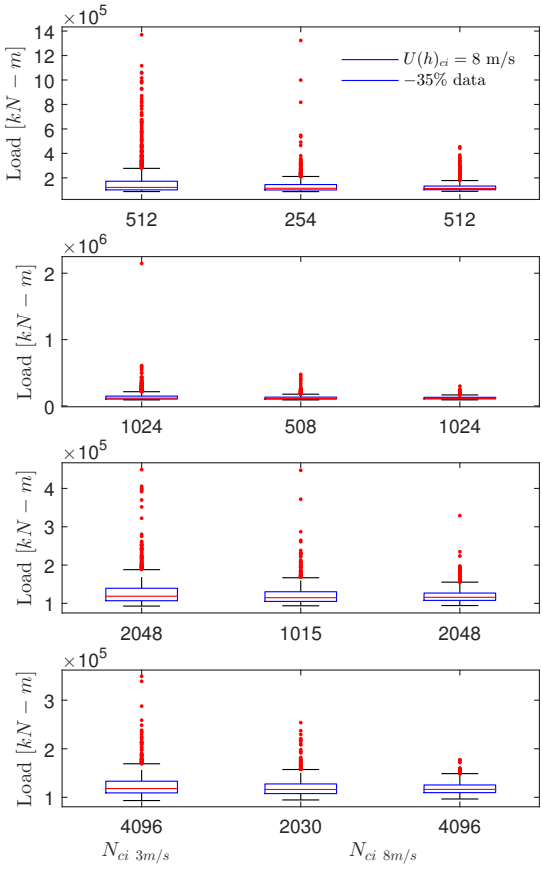
(g) Distribution of 50 year extreme LSSTipMya values originating from N = 512, 1024, 2048 and 4069 compared to the estimate distribution from elevated cut-in wind speeds with equal and less number of simulations accompanied with 26% of saved data calculations



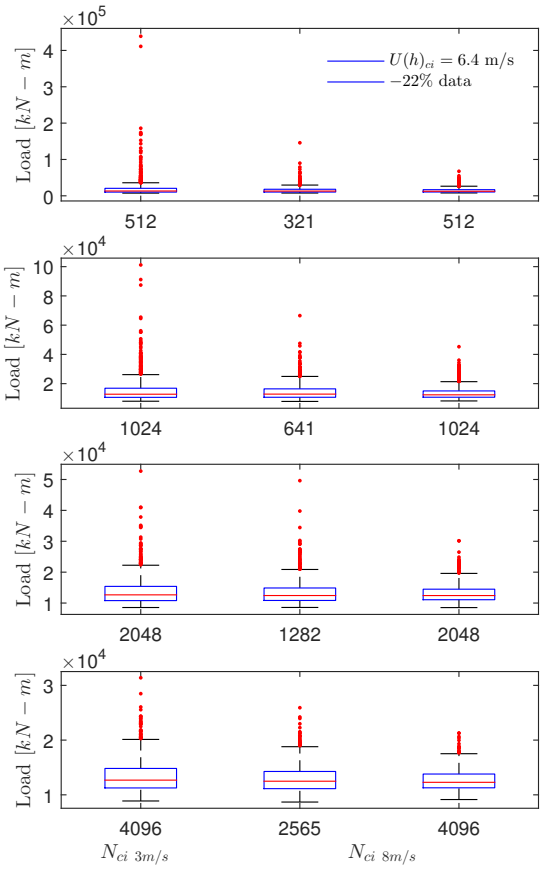
(h) Distribution of 50 year extreme LSSTipMza values originating from N = 512, 1024, 2048 and 4069 compared to the estimate distribution from elevated cut-in wind speeds with equal and less number of simulations accompanied with 25% of saved data calculations



(i) Distribution of 50 year extreme TwrBsMxt values originating from N = 512, 1024, 2048 and 4069 compared to the estimate distribution from elevated cut-in wind speeds with equal and less number of simulations accompanied with 36% of saved data calculations



(j) Distribution of 50 year extreme TwrBsMyt values originating from N = 512, 1024, 2048 and 4069 compared to the estimate distribution from elevated cut-in wind speeds with equal and less number of simulations accompanied with 35% of saved data calculations

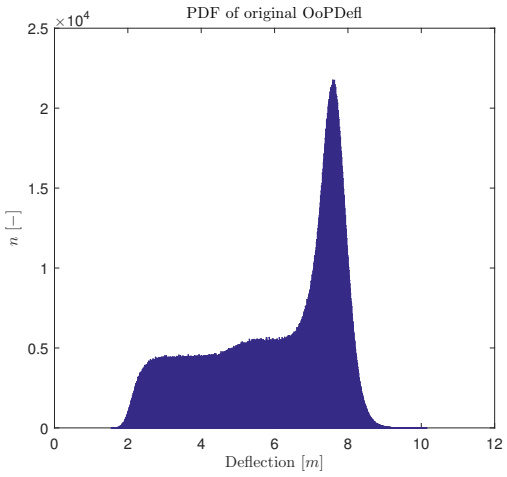


(k) Distribution of 50 year extreme TwrBsMzt values originating from N = 512, 1024, 2048 and 4069 compared to the estimate distribution from elevated cut-in wind speeds with equal and less number of simulations accompanied with 22% of saved data calculations

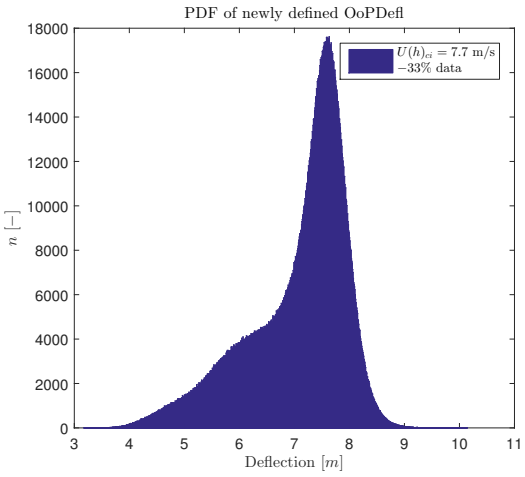
Figure I.1: Uncertainty in elevated Rayleigh distributed (8 m/s) cut in wind speed with extreme sub set simulations: $N = 256, 512, \ldots, 4096$

J

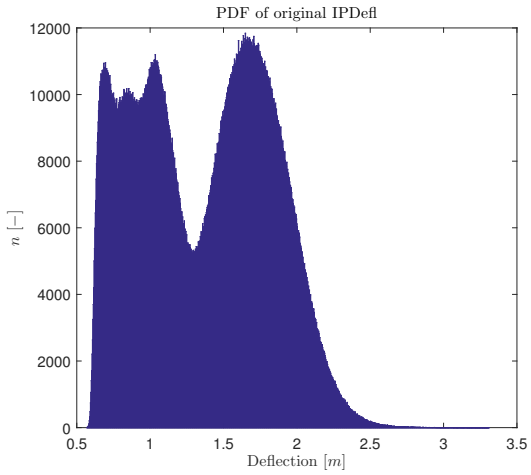
Distribution of Loads at Elevated cut-in Wind Speed Rayleigh Distribution



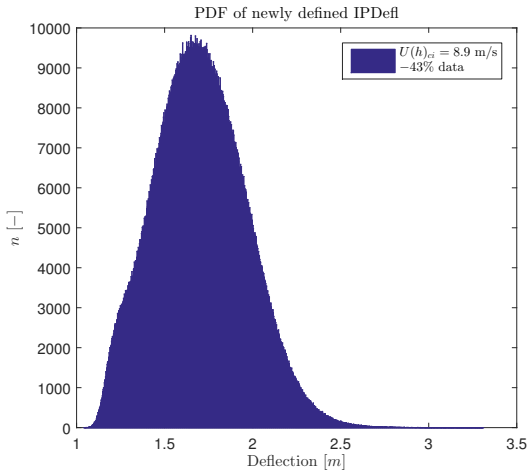
(a) Original load distribution OoPDefl



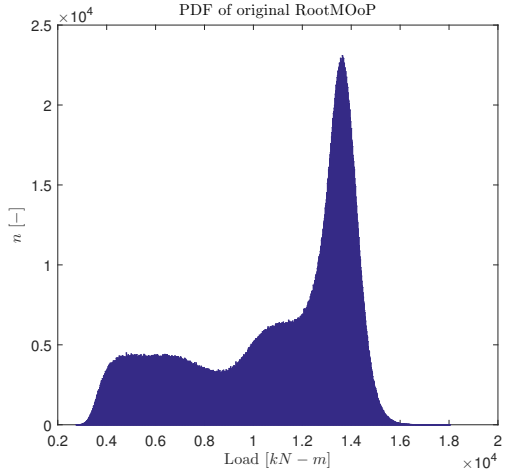
(b) Load distribution OoPDefl with cut in wind speed: 7,7 m/s



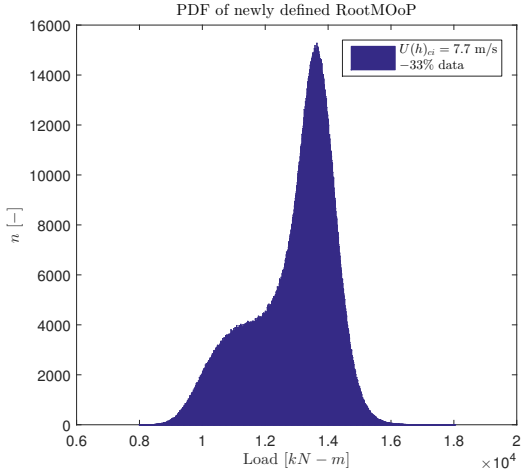
(c) Original load distribution IPDefl



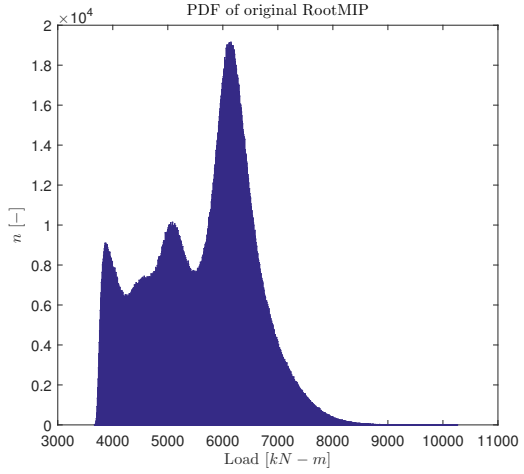
(d) Load distribution IPDefl with cut in wind speed: 8,9 m/s



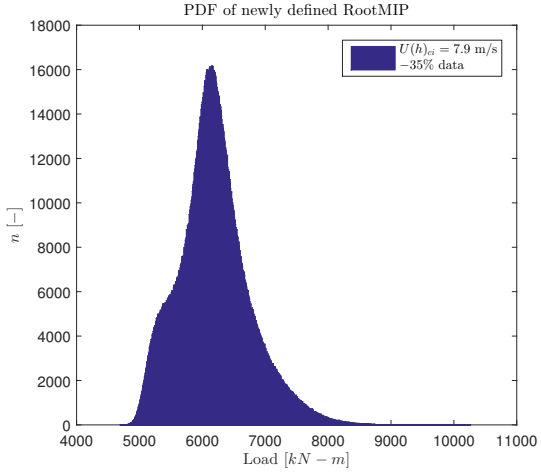
(e) Original load distribution RootMOop



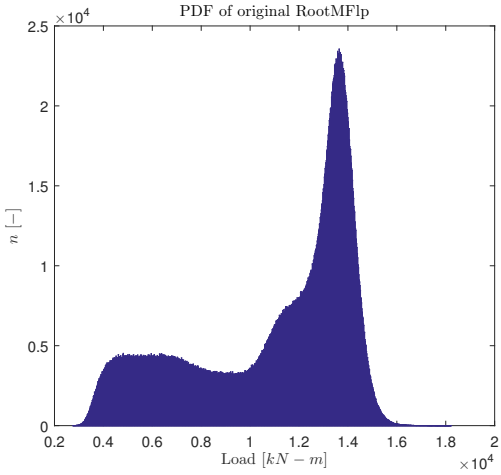
(f) Load distribution RootMOop with cut in wind speed: 7,7 m/s



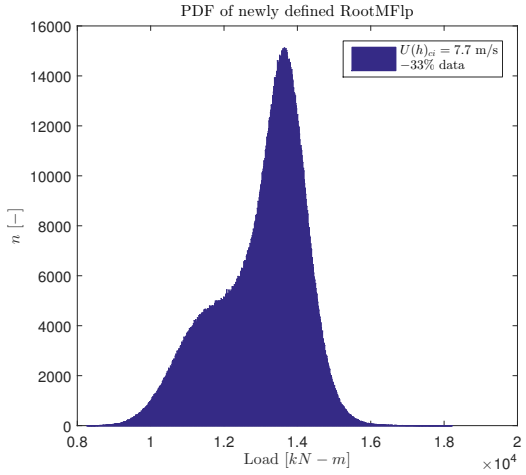
(g) Original load distribution RootMIP



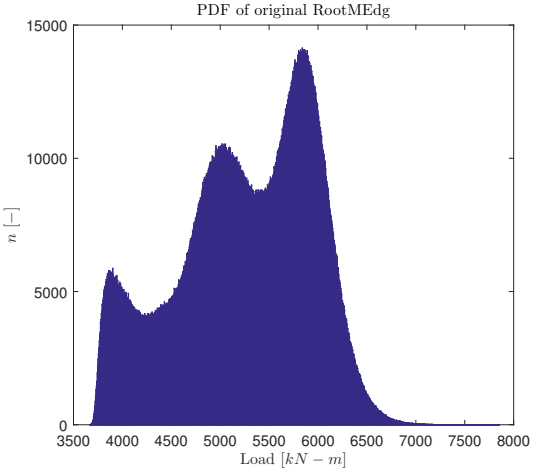
(h) Load distribution RootMIP with cut in wind speed: 7,9 m/s



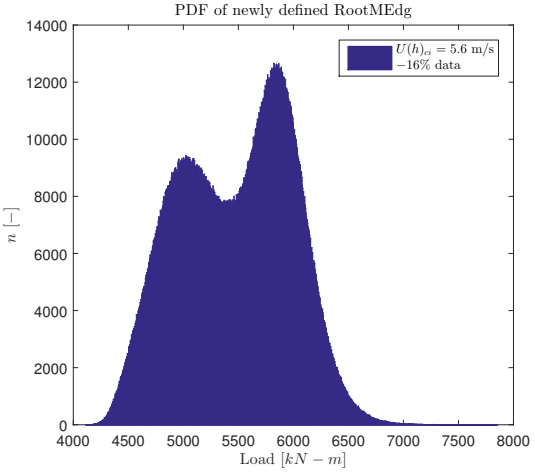
(i) Original load distribution RootMFlp



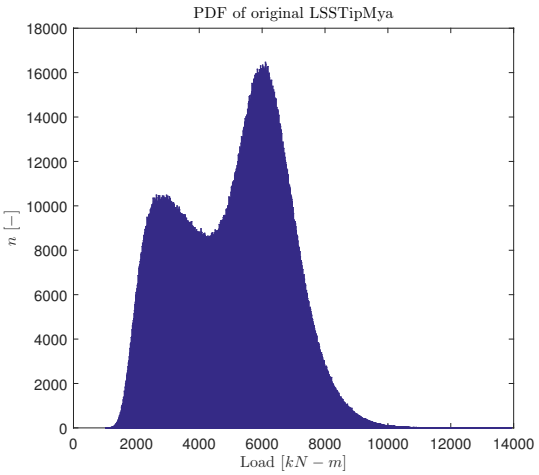
(j) Load distribution RootMFlp with cut in wind speed: 7,7 m/s



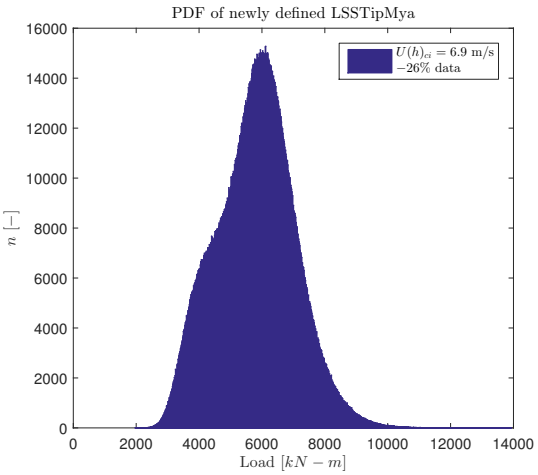
(k) Original load distribution RootMEdg



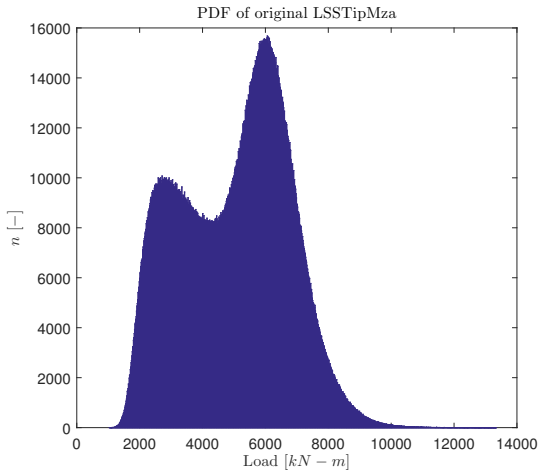
(l) Load distribution RootMEdg with cut in wind speed: 5,6 m/s



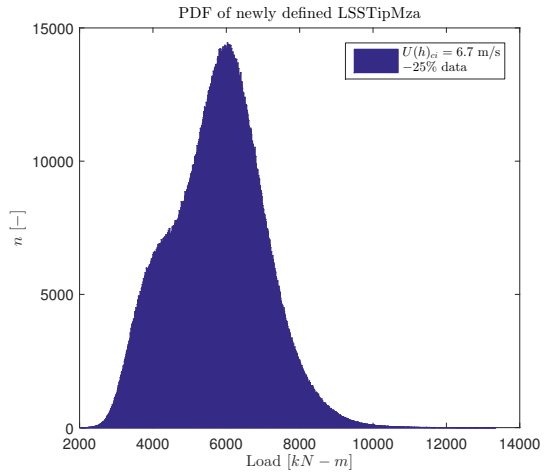
(m) Original load distribution LSSTipMya



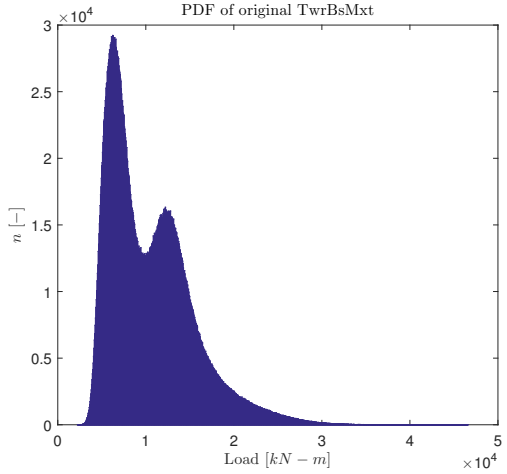
(n) Load distribution LSSTipMya with cut in wind speed: 6,9 m/s



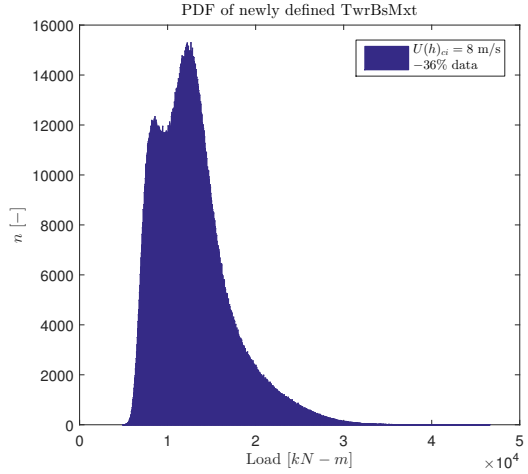
(o) Original load distribution LSSTipMza



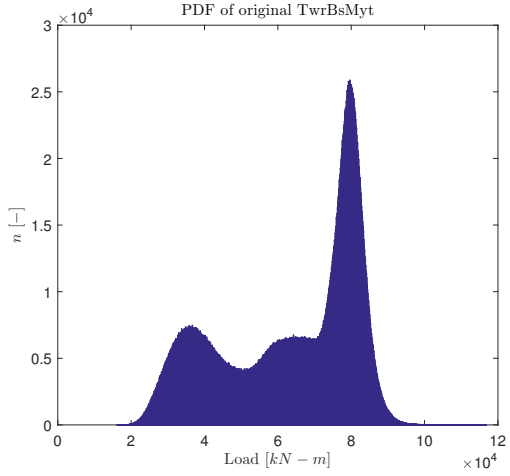
(p) Load distribution LSSTipMza with cut in wind speed: 6,7 m/s



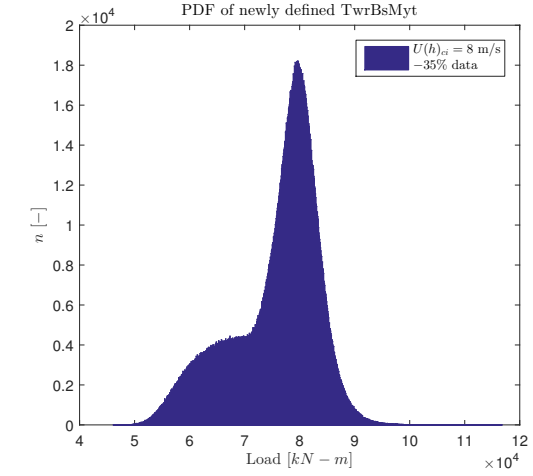
(q) Original load distribution TwrBsMxt



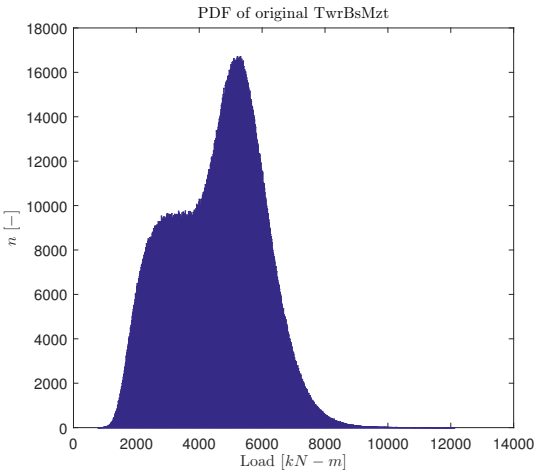
(r) Load distribution TwrBsMxt with cut in wind speed: 8 m/s



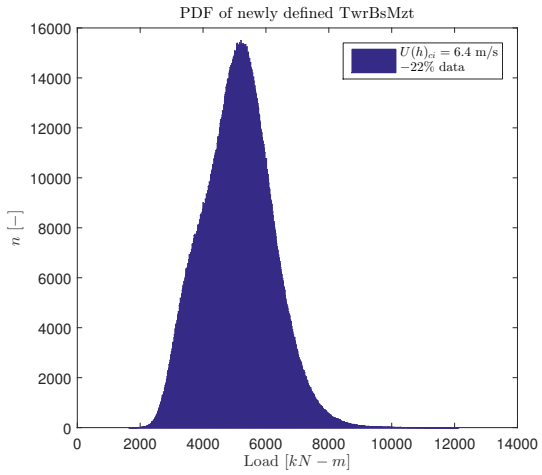
(s) Original load distribution TwrBsMyt



(t) Load distribution TwrBsMyt with cut in wind speed: 8 m/s



(u) Original load distribution TwrBsMzt

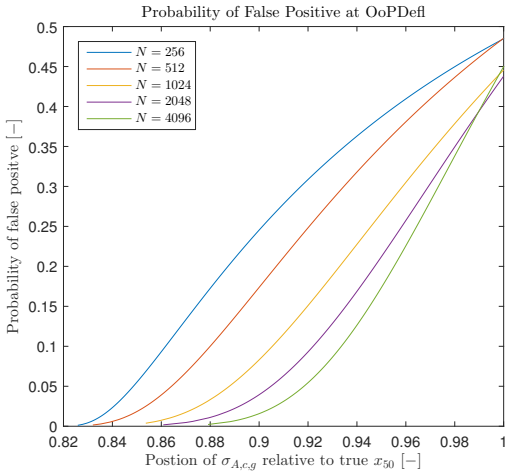


(v) Load distribution TwrBsMzt with cut in wind speed: 6,4 m/s

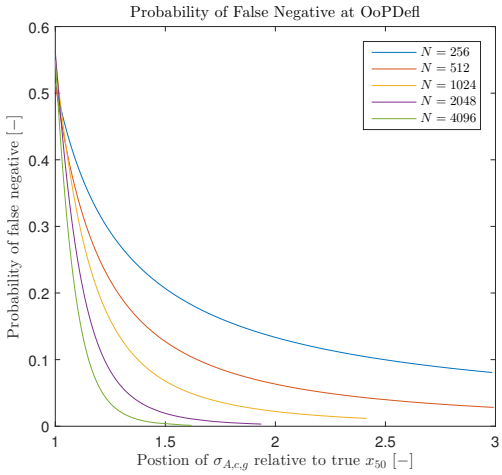
Figure J.1: Distributions of loads and deflections for elevated cut in wind speeds, calculated from the lowest wind speed in the lowest 35% loads region

K

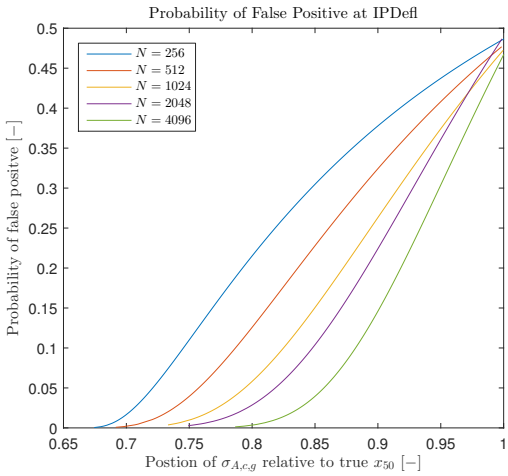
Probability of False Component Selection



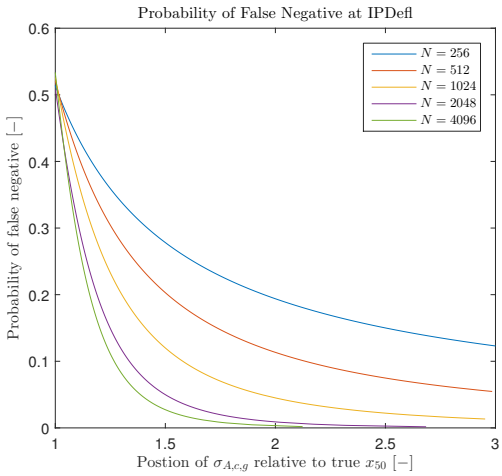
(a) False Positive component selection OoPDefl



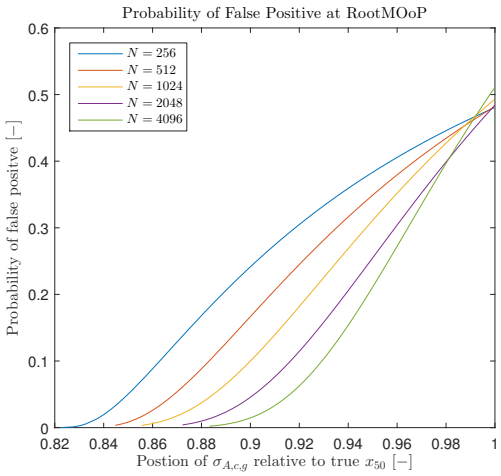
(b) False Negative component selection OoPDefl



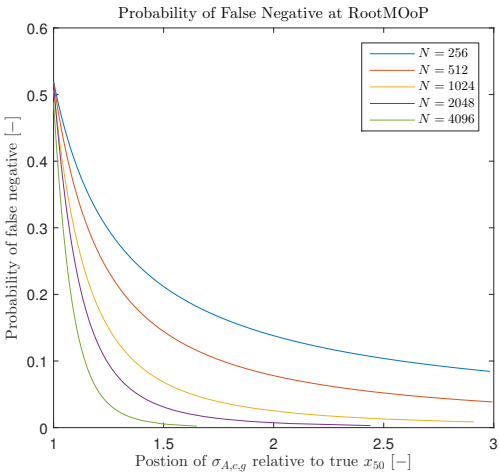
(c) False Positive component selection IPDefl



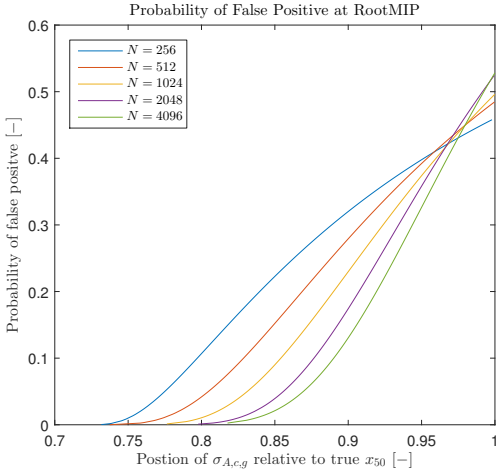
(d) False Negative component selection IPDefl



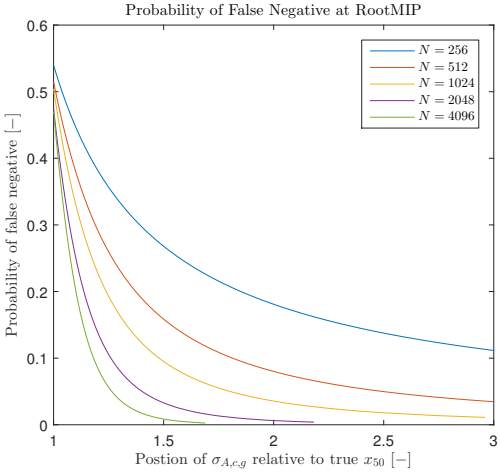
(e) False Positive component selection RootMOop



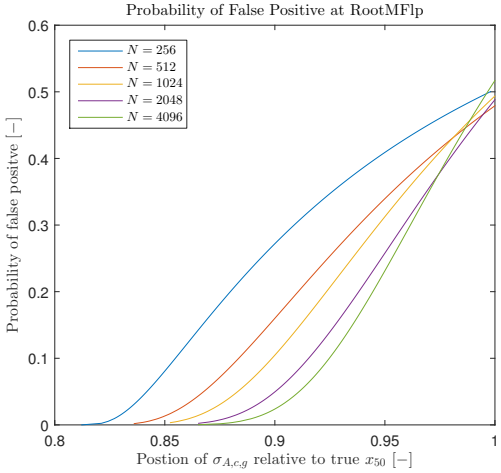
(f) False Negative component selection RootMOop



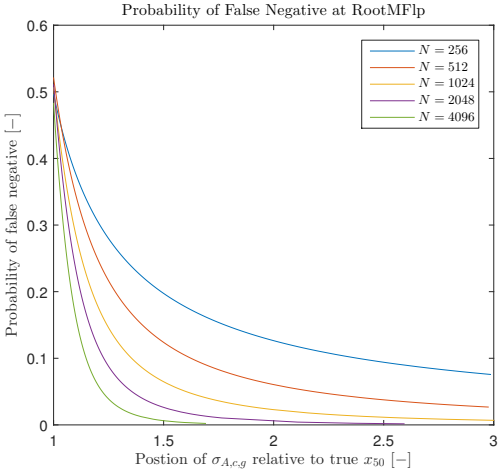
(g) False Positive component selection RootMIP



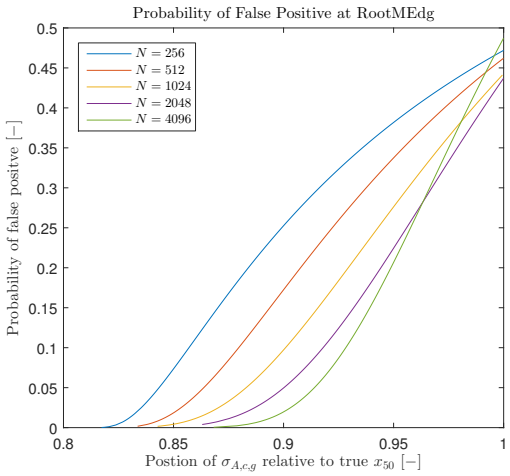
(h) False Negative component selection RootMIP



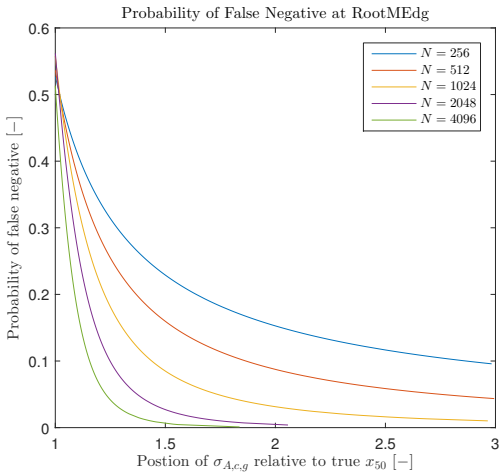
(i) False Positive component selection RootMFlp



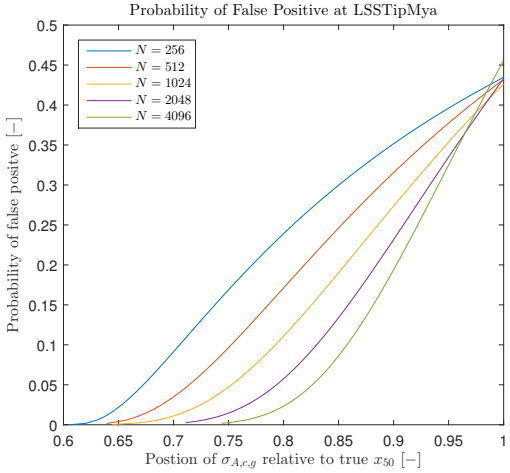
(j) False Negative component selection RootMFlp



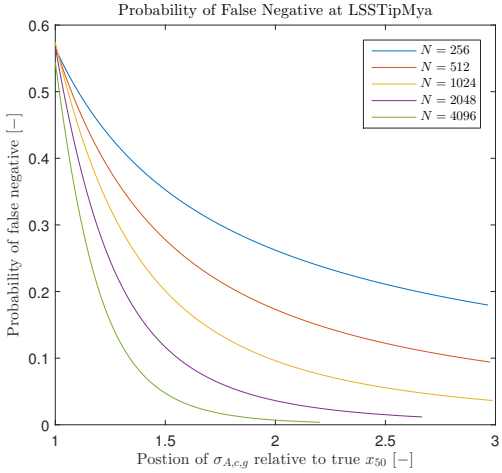
(k) False Positive component selection RootMEdg



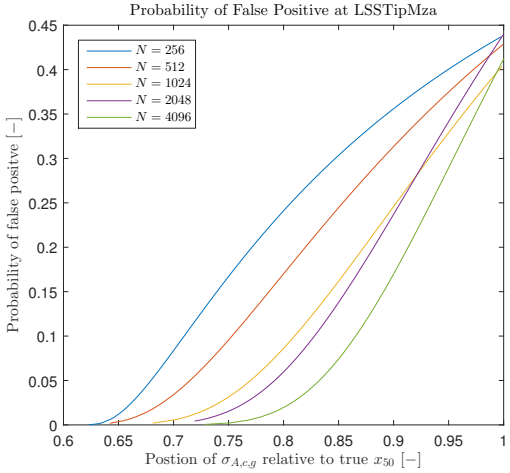
(l) False Negative component selection RootMEdg



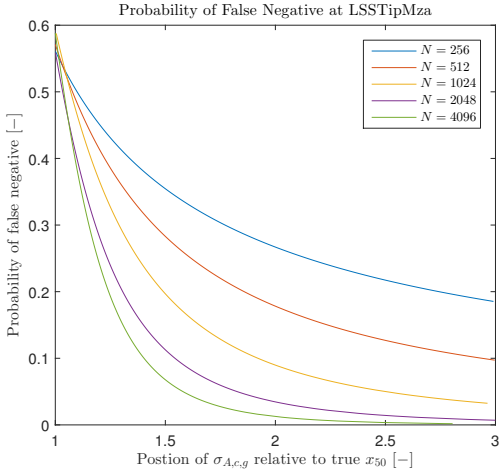
(m) False Positive component selection LSSTipMya



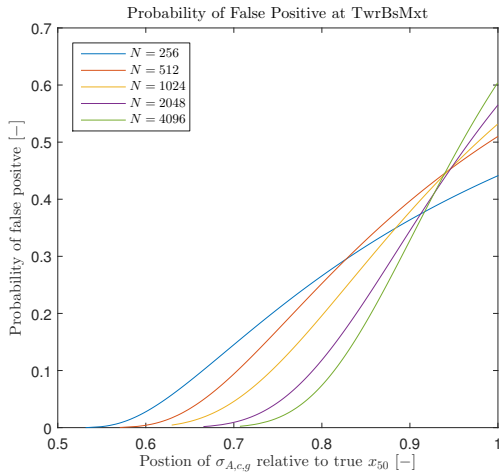
(n) False Negative component selection LSSTipMya



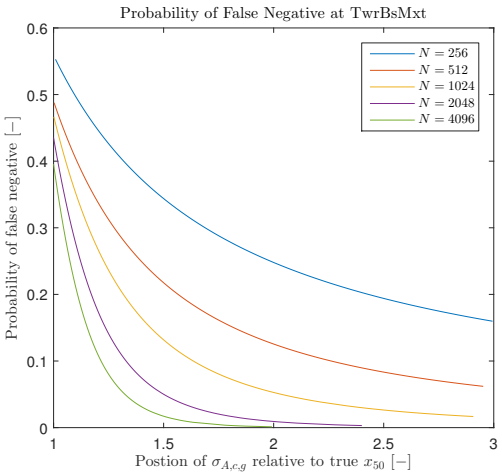
(o) False Positive component selection LSSTipMza



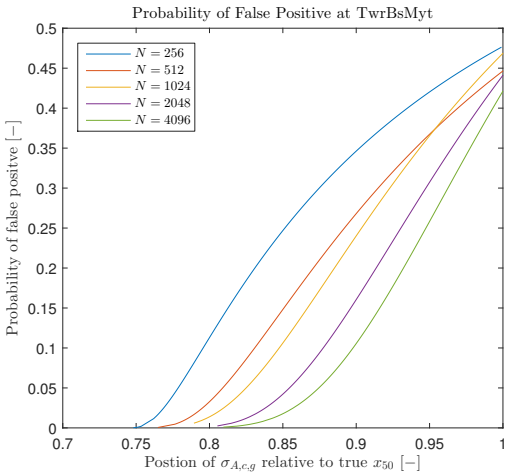
(p) False Negative component selection LSSTipMza



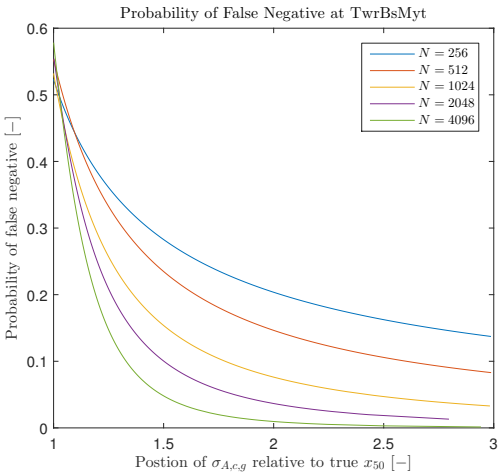
(q) False Positive component selection TwrBsMxt



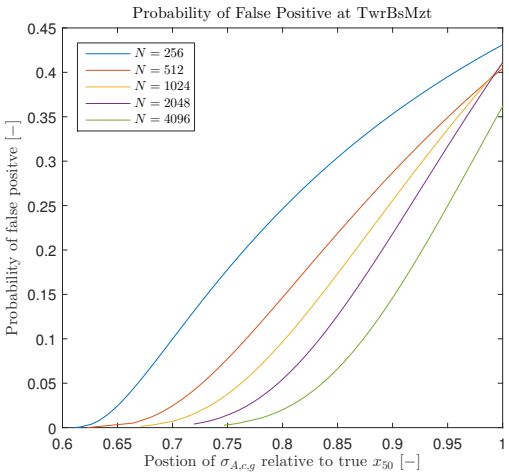
(r) False Negative component selection TwrBsMxt



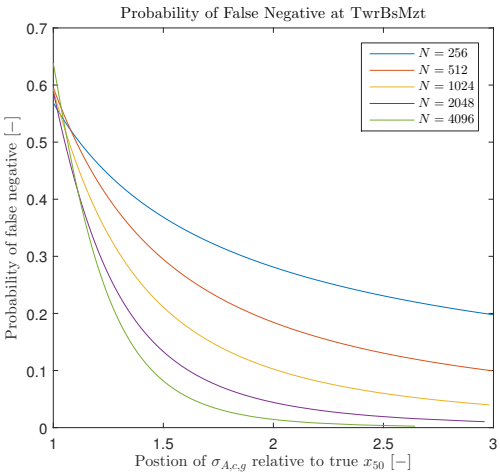
(s) False Positive component selection TwrBsMyt



(t) False Negative component selection TwrBsMyt



(u) False Positive component selection TwrBsMzt

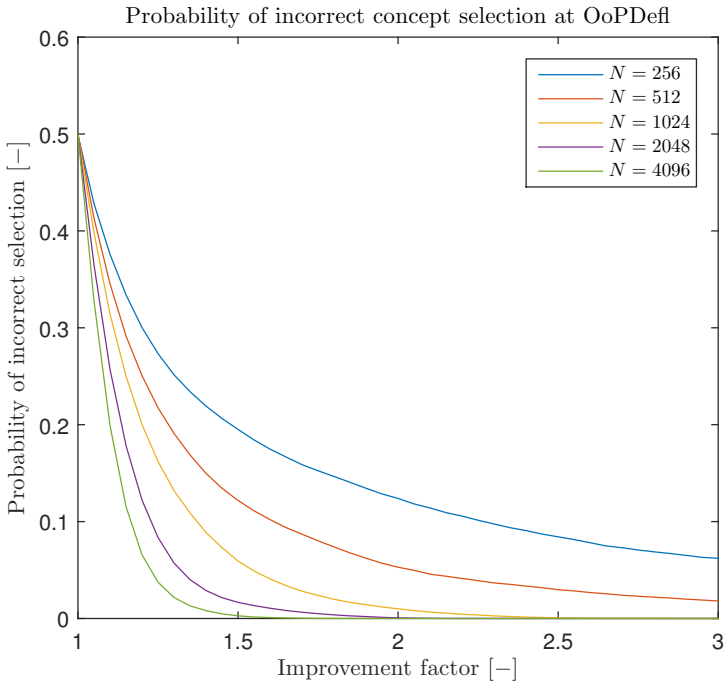


(v) False Negative component selection TwrBsMzt

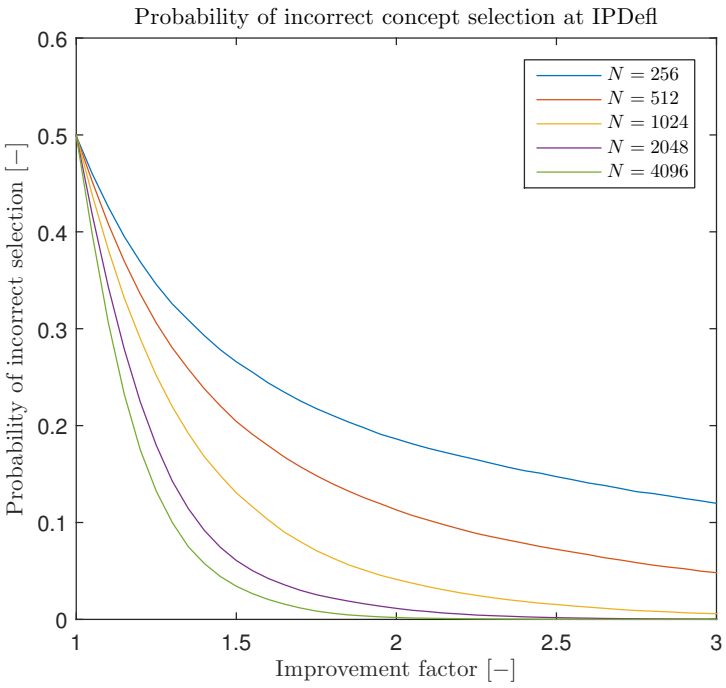
Figure K.1: False positive and false negative component selection of eleven different wind turbine loads and deflections at various N

L

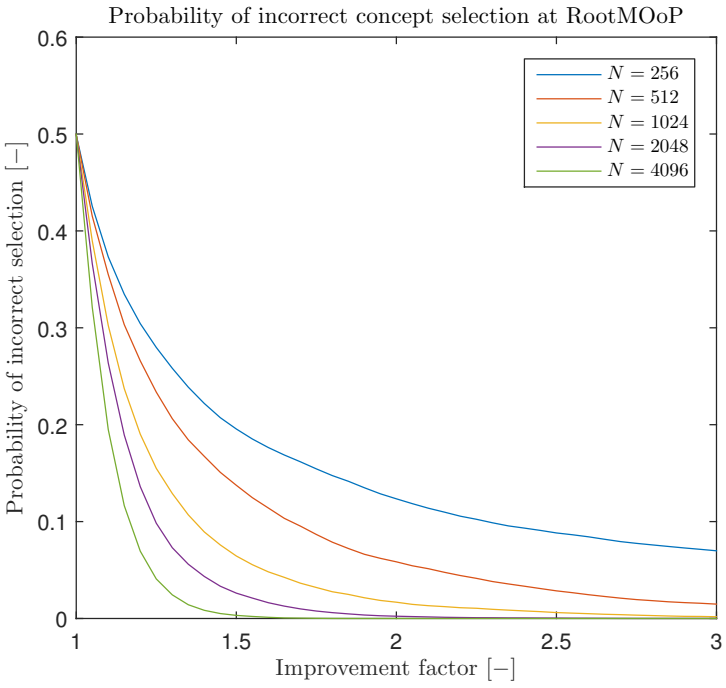
Probability of False Concept Selection



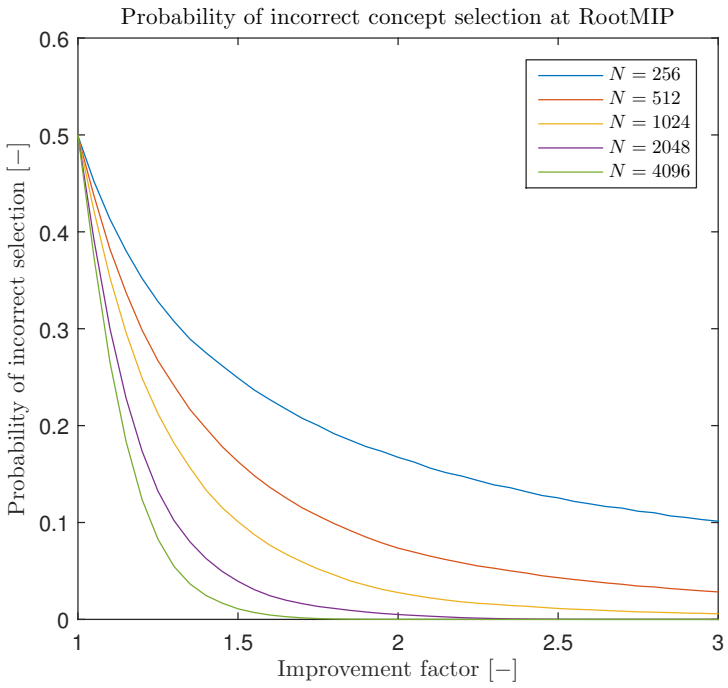
(a) False concept selection of OoPDefl at various N



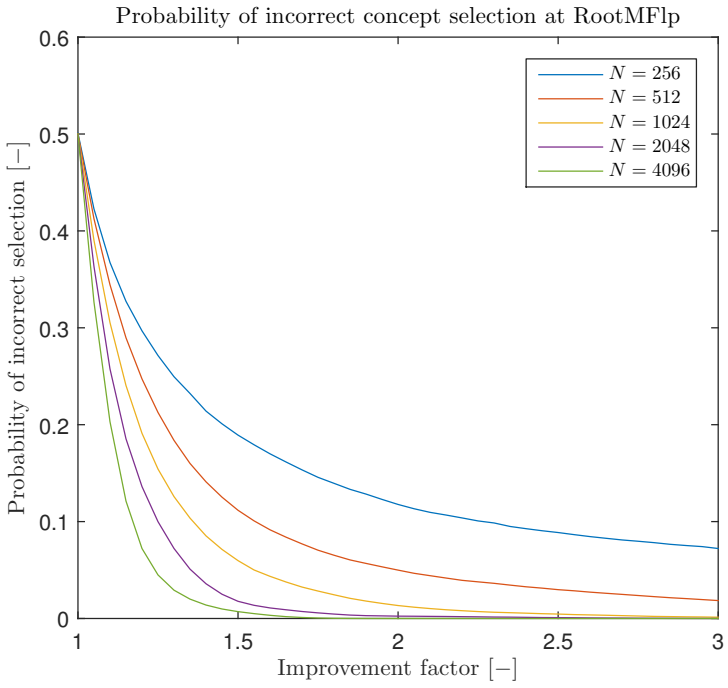
(b) False concept selection of IPDefl at various N



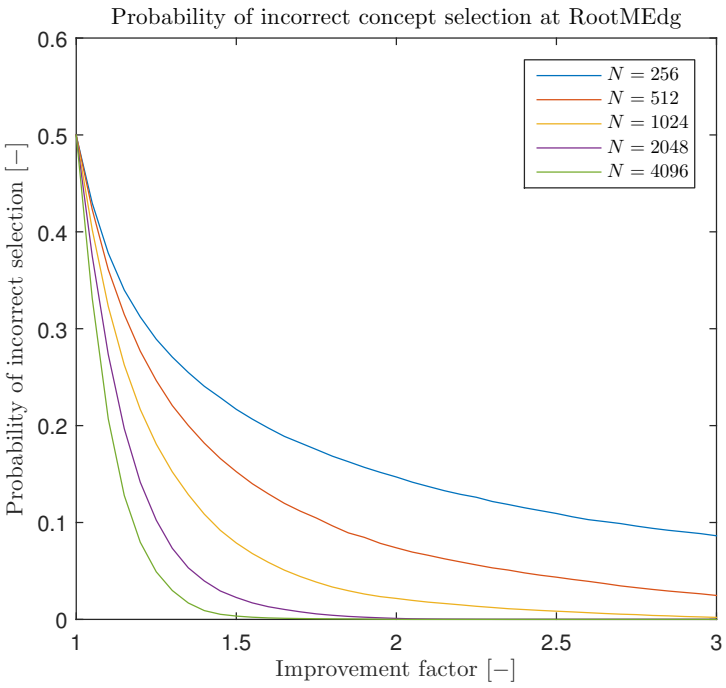
(c) False concept selection of RootMOoP at various N



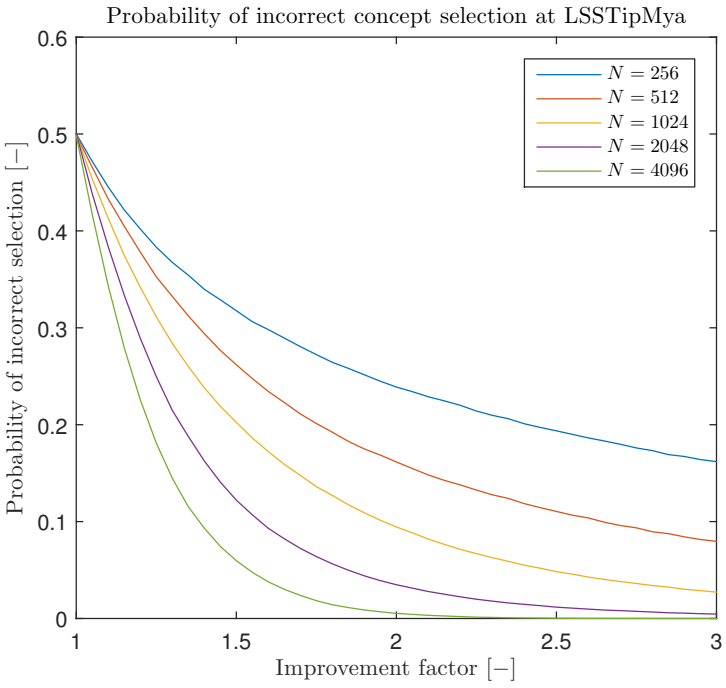
(d) False concept selection of RootMIP at various N



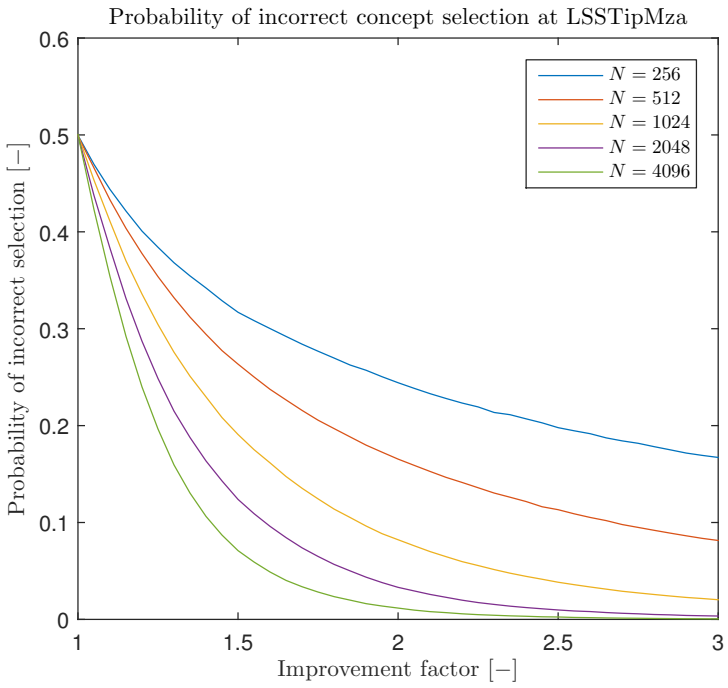
(e) False concept selection of RootMFlp at various N



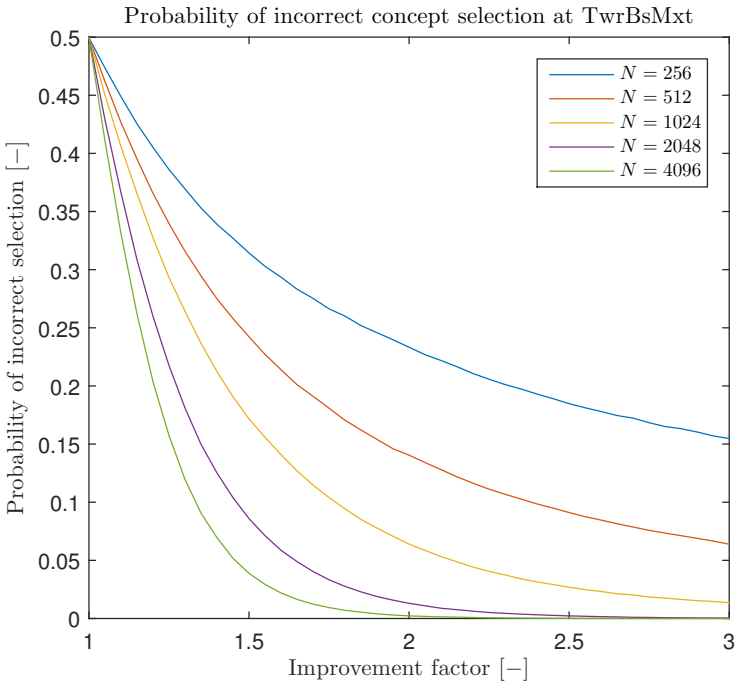
(f) False concept selection of RootMEdg at various N



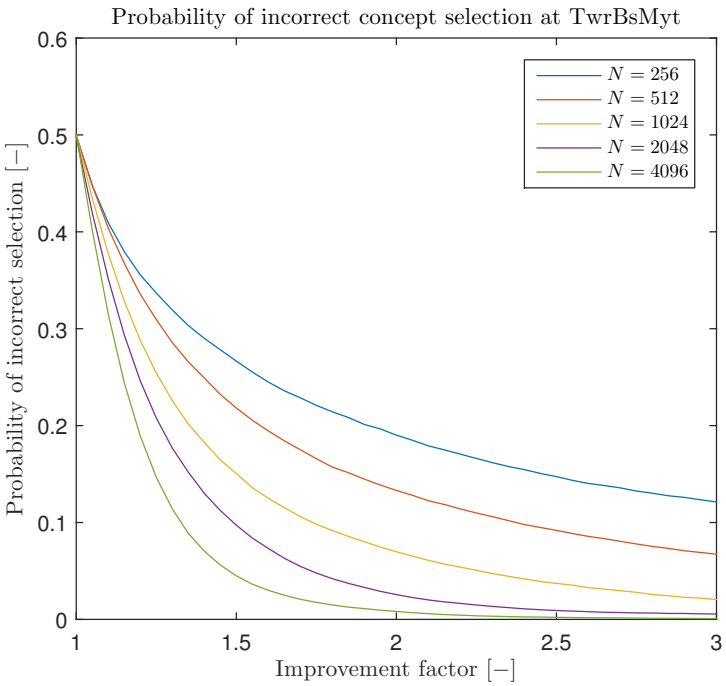
(g) False concept selection of LSSTipMya at various N



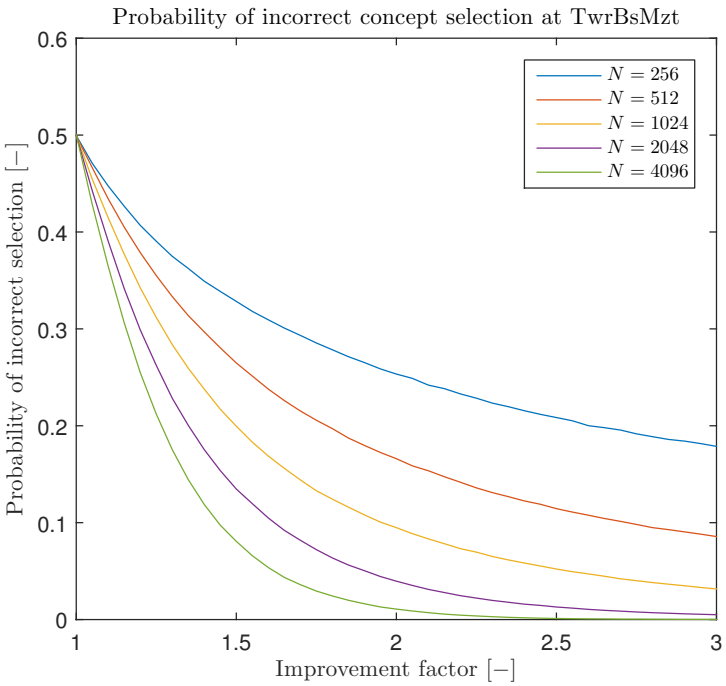
(h) False concept selection of LSSTipMza at various N



(i) False concept selection of TwrBsMxt at various N



(j) False concept selection of TwrBsMyt at various N



(k) False concept selection of TwrBsMzt at various N

Figure L.1: False concept selection of eleven wind turbine loads and deflections at various N