Structural Reliability of Wind Turbine Blades

Design Methods and Evaluation

Dimitrov, Nikolay Krasimirov; Staerdahl, Jesper; Friis-Hansen, Peter ; Berggreen, Christian

Publication date:
2013

Document Version
Publisher's PDF, also known as Version of record

Structural Reliability of Wind Turbine Blades: Design Methods and Evaluation

by

Nikolay Dimitrov

Department of Wind Energy
Technical University of Denmark

and

Siemens Wind Power A/S

February 2013

Preface

This dissertation is submitted as a partial fulfillment of the requirements for obtaining the degree Doctor of Philosophy. The work was carried out between September 2009 and February 2013 at Siemens Wind Power and at the Department of Wind Energy, the Technical University of Denmark. Supervisors of the study are Christian Berggreen, Associate Professor at the Department of Wind Energy, DTU, Peter Friis-Hansen, Program Director Energy at Det Norske Veritas, and Jesper Stærdahl, Team Leader at Siemens Wind Power A/S.

The financial support of Siemens Wind Power A/S and the Danish Ministry of Science and Innovation is gratefully appreciated.

I would like to thank all the people who helped me with knowledge and advice. Special thanks to Peter Friis-Hansen, who, despite his long and extremely busy workdays, never got tired of having discussions and getting new ideas. Thanks to Christian Berggreen for his supervision and especially for his realism, which surely prevented us from spending too much time on overcomplicated problems. Thanks to Jesper Stærdahl, whose practical approach based on business experience has contributed to the industry relevance of the study.

Special thanks to professor Armen Der Kiureghian. His exceptional qualities as mentor and researcher, and the interaction with his enthusiastic team of students helped my five-month stay at the University of California, Berkeley, to become one of the most memorable and inspirational periods of my study. Admittedly, along with the excellent academic atmosphere at UC Berkeley, the difference between the winter weather in California and in Denmark is also something I truly enjoyed.

Many thanks to all my colleagues at Siemens Wind Power for their help in solving tricky wind turbine-related problems and even trickier administrative tasks. The people I met at Siemens certainly prove that the perfect job and perfect colleagues do exist.

Finally, I would like to thank my wife Danuta and daughter Maja for their love and support, and I hope they will excuse me for the endless evenings and weekends I spent bent over the computer screen instead of spending the time with them.

Being given three years to work on a single scientific project is a unique opportunity. The amount of time and the freedom of taking decisions allows exploring a problem in a great depth while significantly enhancing various personal skills and experience. At the end of such a journey one feels enlightened and striving for new knowledge. The work on this dissertation has been very inspirational for me, and I truly hope that the reader will find it equally inspiring.

Nikolay Dimitrov
Lyngby, February 2013

Summary

In the past decade the use of wind energy has expanded significantly, transforming a niche market into a practically mainstream energy generation industry. With the advance of turbine technology the search for more efficient solutions has lead to increased focus on probabilistic modelling and design. Reliability-based analysis methods have the potential of being a valuable tool which can improve the state of knowledge by explaining the uncertainties, and form the probabilistic basis for calibration of deterministic design tools.

The present thesis focuses on reliability-based design of wind turbine blades. The main purpose is to draw a clear picture of how reliability-based design of wind turbines can be done in practice. The objectives of the thesis are to create methodologies for efficient reliability assessment of composite materials and composite wind turbine blades, and to map the uncertainties in the processes, materials and external conditions that have an effect on the health of a composite structure.

The study considers all stages in a reliability analysis, from defining models of structural components to obtaining the reliability index and calibration of partial safety factors. In a detailed demonstration of the process of estimating the reliability of a wind turbine blade and blade components, a number of probabilistic load and strength models are formulated, and the following scientific and practical questions are answered:

a) What material, load and uncertainty models need to be used

b) How can different failure modes be taken into account

c) What reliability methods are most suitable for the particular task

d) Are there any factors specific to wind turbines such as materials and operating conditions that need to be taken into account

e) Are there ways for improvement by developing new models and standards or carrying out tests

The following aspects are covered in detail:

- The probabilistic aspects of ultimate strength of composite laminates are addressed. Laminated plates are considered as a general structural reliability system where each layer in a laminate is a separate system component. Methods for solving the system reliability are discussed in an example problem.

- Probabilistic models for fatigue life of laminates and sandwich core are developed and calibrated against measurement data. A modified, nonlinear S-N relationship is formulated where the static strength of the material is included as a parameter. A Bayesian inference model predicting the fatigue resistance of face laminates based on the static and fatigue strength of individual lamina is developed. A series of tests of the fatigue life of balsa wood core material are carried out, and a probabilistic model for the fatigue strength of balsa core subjected to transverse shear loading is calibrated to the test data.

- A review study evaluates and compares several widely-used statistical extrapolation methods for their capability of modelling the short-term statistical distribution of blade loads and tip deflection. The best performing methods are selected, and several improvements are suggested, including a procedure for automatic determination of tail threshold level, which allows for efficient automated use of peaks-over-threshold methods.

- The problem of obtaining the long-term statistical distribution of load extremes is discussed by comparing the method of integrating extrapolated short-term statistical distributions against extrapolation of data directly sampled from the long-term distribution. The comparison is based on the long-term distribution of wind speed, turbulence, and wind shear, where a model of the wind shear distribution is specifically developed for the purpose.

- Uncertainties in load and material modelling are considered. A quantitative assessment of the influence of a number of uncertainties is done based on modelled and measured data.

- Example analyses demonstrate the process of estimating the reliability against several modes of failure in two different structures. This includes reliability against blade-tower collision, and the reliability against ultimate and fatigue failure of a sandwich panel.

The results from the reliability analyses are then used for calibrating partial safety factors against a target reliability level.

The main conclusions from the thesis are that a) the problem of estimating the reliability of wind turbine blades has been addressed in detail, and suitable methodologies for carrying efficient and robust reliability analysis have been identified, b) model uncertainties have a very high influence on the reliability estimate, and an effort to reduce model uncertainties can be rewarded with improved structural reliability or lower safety factors, and c) wind turbine design using present state-of-the-art models and standard safety factors results in sufficiently safe structures, and in some cases the actual reliability exceeds the assumed sufficient reliability levels significantly.

Aspects that need to be covered in further work are investigating the influence of defects, studying the fatigue properties of composites and defining physical models for fatigue damage accumulation, making risk-based decision analysis possible by including cost measures, quantification of uncertainties, and further improving the knowledge on turbine loads, including turbine controller behaviour, extreme operational events and frequency of faults, and assessment of the wind field across the entire rotor.

Resumé

Anvendelsen af vindenergi er udvidet betydeligt i løbet af det seneste årti. Fra at være et niche-marked har vindmølleindustrien udviklet sig til en betydelig aktør i energisektoren. Fremgangen i vindmølleindustrien kræver mere effektive løsninger, og det har skabt forøget fokus på probabilistisk modellering og design. Pålidelighedsbaserede design metoder har potentiale til at blive et effektivt værktøj som kan forbedre indsigten i teknologien, forklare usikkerhederne, og bruges som et probabilistisk grundlag for kalibrering af deterministiske designværktøjer.

Denne afhandling fokuserer på pålidelighedsbaseret design of vindmøllevinger. Hovedformålet er at tegne et klart billede af, hvordan pålidelighedsbaseret design af vindmøller skal udføres i praksis. Målsætningerne er at skabe metoder for effektiv pålidelighedsanalyse af kompositmaterialer og komposit-vindmøllevinger, samt kortlægge usikkerhederne i de processer, materialer og miljøfaktorer, som har en effekt på strukturens sikkerhed.

I studiet er taget hensyn til alle processer i en pålidelighedsanalyse, fra definering af strukturelle modeller af komponenter til estimering af pålidelighedsindeks og kalibrering af sikkerhedsfaktorer. Proceduren for beregning af svigtsandsynligheden af vindmøllevinger og vingekomponenter demonstreres i detaljer, der formuleres flere probabilistiske modeller for belastning og styrke, og der svares på følgende videnskabelige og praktiske spørgsmål:

a) Hvilke materiale-, last- og usikkerhedsmodeller skal bruges

b) Hvilke svigtmåder skal der tages hensyn til

c) Hvilke metoder for pålidelighedsanalyse er mest velegnede til de konkrete problemer

d) Skal der tages hensyn til faktorer som er specifikke for vindmøller, for eksempel materialer og driftsforhold

e) Er der mulighed for teknologiske fremskridt igennem udvikling af nye modeller, standarder, og udførelse af tests

Følgende områder diskuteres i detaljer:

- Probabilistiske aspekter af laminaternes maksimalstyrke diskuteres. Lamineredе plader betragtes som et generelt pålidelighedssystem, hvor hvert lag i laminatet er en separat systemkomponent. Løsning af et eksempelproblem viser hvordan pålideligheden af et sådant system kan beregnes.

- Probabilistiske modeller for udmattelse af laminater og sandwichkerner defineres og kalibreres til testmålinger. En modificeret, ulineær S-N kurve formuleres hvor materialets statiske styrke inkluderes som parameter. Udmattelseskapaciteten af laminater estimeres ved brug af en "Bayesian Inference"model, baseret på statisk og udmattelsesstyrke af individuelle lag af kompositmateriale. En serie af udmattelsestester er blevet udført, og en probabilistisk model for udmattelseskapaciteten af balsakernemateriale lastet med forskydningsspænding er kalibreret til testresultaterne.

- I et opsummeringsstudie sammenlignes adskillige statistiske ekstrapolationsmetoder for deres evne til modellering af den korttids-statistiske fordeling af vingelaster og tipudbøjning. De bedst fungerende metoder udpeges, og adskillige forbedringer foreslås, heriblandt en procedure for automatisk estimering af ekstremresponsens tærskelværdi, som giver mulighed for effektiv automatisk brug af "peaks-over-threshold"methoder.

- Problemet omkring evaluering af den langtids-statistiske fordeling af laster diskuteres ved hjælp af resultatsammenligning fra integration af ekstrapolerede korttids-statistiske fordelinger versus direkte sampling fra den langtids-statistisk fordeling. Sammenligningen er baseret på langtidsfordelingerne af vindhastighed, turbulens og vindprofil. En statistisk model for vindprofilet er blevet udviklet specifikt til formålet.

- Usikkerheder i last- og materialemodellering diskuteres. En kvantitativ evaluering af effekten af adskillige usikkerhedsfaktorer udføres, baseret på modeller og testresultater.

- Eksampelanalyser demonstrerer processen for estimering af pålideligheden i adskillige svigtmåder for to forskellige strukturer. Dette inkluderer pålideligheden mod vinge-tårn kollision, samt pålideligheden mod ekstrem- og udmattelsessvigt af et sandwichpanel. Resultaterne fra pålidelighedsanalyserne bruges til at kalibrere partielle sikkerhedsfaktorer til en fastsat pålidelighedsværdi.

Hovedkonklusionerne fra studiet er at: a) Problemet ved estimering af pålidelighed for vindmøllevinger er blevet undersøgt i detaljer, og passende metodologier til effektiv og robust

pålidelighedsanalyse er blevet identificeret, b) modelusikkerheder har en stor indvirkning på pålidelighedsværdierne, og en indsats for at reducere modelusikkerhederne kan belønnes med forbedret strukturpålidelighed og lavere sikkerhedsfaktorer, og c) vindmølledesign som bruger de nyeste modeller og standardværdier for sikkerhedsfaktorer resulterer i en tilstrækkelig pålidelighed, og i mange tilfælde er den aktuelle pålidelighed højere end hvad er antaget i standarderne.

Nogle områder som har behov for yderligere undersøgelse er analyse af defekters indvirkning, udmattelsesegenskaber af kompositter, definition af fysiske modeller til summering af udmattelsesskade, omkostningsstudier som kan danne grundlag for udførelse af risikobaserede beslutningsanalyser, kvantificering af usikkerhedsfaktorer, og videreudvikling af indsigten i vindmøllelaster, herunder vindmøllestyringen, ekstreme hændelser under drift, fejlhyppighed, og analyse af vindfeltet over rotoren.

Contents

Chapter 1

Introduction

1.1 Background

In the past decade wind turbine industry has experienced significant growth, fuelled by growing environmental concerns and increased government spendings. What was once a niche market has now practically become a mainstream energy generation technology, accounting for a significant share of new capacity installations worldwide.

As it can be expected, the expansion in turbine size and production quantities has also meant that design-engineers have to face new challenges. The fast growth in the size of new wind turbines has led to a lack of operational experience with structures of this size, complexity, and the short product roll-out times mean that adequate modelling is of critical importance, as in many cases decisions have to be taken solely on the basis of model predictions.

Detailed modelling of turbine behaviour requires a solid theoretical basis supported by design guidelines and standards. While the presently available design practices fulfill these conditions to a large extent, some uncertainties remain. One example is the actual reliability of a wind turbine structure achieved by designing in conformance with existing standards. Although the standards either state explicitly a target reliability level (EN 1990, [22]), or imply such a level by enforcing the use of certain safety factors(IEC61400-1 [37]), it is not guaranteed that the actual reliability of the structure will correspond to the intended reliability level. As stated in ISO 2394 [38], calibrated reliability values are related to a specific set of structural and probabilistic models, and using calibrated values in connection with other models could lead to unintentionally high or low levels of reliability. Keeping in mind that models used by designers will often differ from the models used for safety

factor tuning, that different turbine designs will have different behaviour and uncertainties associated with them, and that safety factors might be calibrated for different types of structures (as is the case with EN-1990 series of standards), it can be expected that there is some uncertainty in the actual reliability. This is supported by comparing studies of the reliability against fatigue failure of a wind turbine blade, where reliability indexes based on partial safety factors from IEC61400-1 vary between 2.7 (Sørensen et al. 2008 [77]), 3.1 (Toft & Sørensen, 2011 [86]), and 3.53 (Veldkamp, 2006, [88]).

Despite the uncertainties in the design methodology described above, deterministic design methods based on partial safety factors will for the time being remain the dominating approach to designing wind turbines, because they are simple, well established, and their shortcomings are known. However, reliability-based analysis is a valuable supplement which can improve the state of knowledge by explaining the uncertainties, and form the probabilistic basis for calibration of deterministic design tools.

Reliability of wind turbines as a whole is a broad field, requiring a multidisciplinary approach due to the combination of components with electrical, mechanical, and aerodynamic functions, and a corresponding variety of materials and load conditions involved. Therefore a project set within a limited time frame has to either narrow the scope of the study, or stay on a general level without considering too many details. For the present thesis the former approach is adopted, and the scope of the study is limited to the structural reliability of composite wind turbine blades.

1.2 Previous work

The interest towards structural reliability of wind turbine blades has been growing together with turbine sizes and rotor diameter, a parameter which increases the demands towards blade design. First comprehensive studies on the subject begin to emerge around the turn of the century, with works by Ronold et al., 1999 [68], and Ronold & Larsen, 2000 [69], which deal with reliability of wind turbine blades against ultimate and fatigue failure, respectively. Tarp-Johansen, Madsen & Frandsen, 2002, [79] attempt to calibrate partial safety factors for extreme load effects on large wind turbines, a study followed in 2003 by a similar analysis on fatigue loads by Tarp-Johansen [78]. Lekou & Philippidis, 2009 [43] combine a probabilistic strength model for a composite blade section and an aeroelastic simulation code, applying reliability analysis on the simulated time series. Sørensen et al., 2008 [77] investigate the reliability against fatigue failure of turbines situated in wind farms. Possibly the most

general study is a PhD thesis by Veldkamp, 2006 [88], where a large number of topics related to wind turbine reliability are discussed, including cost-benefit analysis and determining cost-optimal reliability levels. Another PhD study by Toft, 2011 [82], works related to which are also published in Toft et al., 2011 [85], Toft & Sørensen, 2011 [86], and Toft et al., 2012 [84], deals with several aspects of wind turbine blade reliability such as ultimate and fatigue strength of composite materials, extreme loads on blades, and uncertainty modelling.

Although the list above is not long, a number of additional references exist which do not specifically deal with wind turbine blade reliability, but the results from which are relevant to one or more of the stages in the reliability analysis, as for example probabilistic ultimate and fatigue strength of composites, statistical extrapolation of wind turbine loads, and others.

There are subjects however that deserve further scientific attention. One of them is system effects and system reliability analysis, which have not been discussed in detail. Other aspects where more detailed considerations would be beneficial are probabilistic modelling of material properties of layered composites including properties of sandwich panels. Further issues that need to be addressed are connected with the practical implementation of reliability-based methods by designers - these include choosing appropriate reliability methods for efficient solutions of different reliability tasks, finding efficient solutions for obtaining extreme load distributions, and applying reliability analysis to problems based on real-world wind turbine structure and behaviour. Finally, based on opinions in industry, more emphasis should be put on quantification and interpretation of the influence of random effects and uncertainties to the reliability, as this is an information which has high potential value.

1.3 Project basics

The present thesis focuses on reliability-based design of wind turbine blades. The main purpose is to draw a clear picture of how reliability-based design of wind turbines can be done in practice. The objectives are the following:

- Create methodologies for efficient reliability assessment of composite materials and composite wind turbine components.
- Map the uncertainties in processes, materials and external conditions that have an effect on the health and behaviour of a composite structure.

The study considers all stages in a reliability analysis, from defining models of structural components to obtaining the reliability index and calibration of partial safety factors. A

detailed demonstration of the process of estimating the reliability of a wind turbine blade and blade components is shown, with the aim of answering the following scientific and practical questions:

a) What material, load and uncertainty models need to be used

b) Which failure modes need to be taken into account

c) What reliability methods are most suitable for the particular task

d) Are there any factors specific to wind turbines such as materials and operating conditions that need to be taken into account

e) Are there ways for improvement by developing new models and standards or carrying out tests

1.4 Outline of the thesis

- In Chapter 2 the reader is introduced to some of the theoretical aspects of reliability analysis that are used in the thesis.

- In Chapter 3 the probabilistic aspects of ultimate and fatigue strength of composites are addressed. Laminated plates are considered as a general structural reliability system where each layer in a laminate is a separate system component. Methods for solving the system reliability are discussed in an example problem. Probabilistic models for fatigue life of laminates and sandwich core are developed and calibrated against measurement data. A modified, nonlinear S-N relationship is formulated where the static strength of the material is included as a parameter. A Bayesian inference model predicting the fatigue resistance of face laminates based on the static and fatigue strength of individual lamina is developed. A series of tests of the fatigue life of balsa wood core material are carried out, and a probabilistic model for the fatigue strength of balsa core subjected to transverse shear loading is calibrated to the test data.

- Chapter 4 discusses the statistical distribution of the loads acting on wind turbine blades. A review study evaluates and compares several widely used statistical extrapolation methods for their capability of modelling the short-term statistical distribution of blade loads and tip deflection. The best performing methods are selected, and several improvements are suggested, including a procedure for automatic determination

of tail threshold level, which allows for efficient automated use of peaks-over-threshold methods. The problem of obtaining the long-term statistical distribution of load extremes is discussed by comparing the method of integrating extrapolated short-term statistical distributions against direct extrapolation from data following the long-term distribution. The comparison is based on the long-term distribution of wind speed, turbulence, and wind shear, where a model of the wind shear distribution is specifically developed for the purpose.

- In Chapter 5 uncertainties in load and material modelling are considered. A quantitative assessment of the influence of a number of uncertainties is done based on modelled and measured data.

- Chapter 6 demonstrates the process of estimating the reliability against several modes of failure in two different structures. This includes reliability against blade-tower collision, and the reliability against ultimate and fatigue failure of a sandwich panel. The results from the reliability analyses are then used for calibrating partial safety factors against a target reliability level.

- Chapter 7 summarizes the most important conclusions from the study and suggests topics that need further attention.

Chapter 2

Structural Reliability

2.1 Introduction to structural reliability

Failure of components is a phenomenon that inevitably occurs in all machines using the present human technology. Regardless of how rigorously our devices are designed, there is always a possibility of unforeseen situations that can lead to failure. The frequency of occurrence of such faults, as well as the price for repairing them directly influences the profitability of a project. Higher failure rates or increased price per repair action will both lead to increasing the operation and maintenance costs. At the same time, an attempt to prevent failures by modifying the design can easily lead to increased costs too - because improving reliability can often mean adding costly materials and safety features. Reliability analysis gives the designer an opportunity to deal with these aspects. Performing a reliability analysis provides information about what is the expected frequency of failure, and it could identify the factors which influence the reliability. Using a reliability model adds a level of complexity to the analysis because variables are represented with their statistical distributions. Depending on the type and complexity of probabilistic modelling used, reliability analysis methods are usually categorized in four levels (DNV, [15]):

Level I methods are deterministic reliability methods that use a single, characteristic value to describe each uncertain variable. Level I methods correspond to standard deterministic design, where safety of the structure is usually ensured by introducing partial safety factors. Safety coefficients may be calibrated by means of a higher-level reliability method.

Level II methods use two values, the mean and the variance, as means of describing each stochastic variable in the model. This is supplemented by a measure of correlation

between variables, usually the covariance. The resulting safety measure is in the form of a reliability index, which can be interpreted geometrically as the distance in standard deviation units from the mean value to the design point.

Level III methods employ the joint probability distribution of all uncertain variables in the model. The basic reliability measure of level III methods is the probability of failure. A large number of methods such as simulation, FORM/SORM, and numerical integration fall within this category. Analysis using level III methods can be used to calibrate partial safety factors which are then implemented in level I design methods.

Level IV methods use economic measures as design optimality criteria. The cost-benefit analysis of construction, maintenance, repair, and consequences of failure results in the selection of an optimal reliability level from economical point of view. This cost-optimal reliability can be used as target reliability of a level III method.

Economic analysis of failure consequences can be a valuable precursor to reliability analysis. However, using cost-optimal target reliability is not necessarily applicable in all practical cases, because often the cost-optimal reliability will be lower than the minimum required by design and safety standards, where additional considerations not directly related to operating cost can be taken into account. An example of this is the study by Veldkamp, 2006 [88] where the optimal failure probability from economical point of view is about $3.5 \cdot 10^{-3}$ annually.

Cost analysis can be very useful even without defining target failure probabilities, as having information about cost of failures opens possibilities for carrying out risk analysis. In a risk analysis procedure the expected cost of different types of failure modes can be estimated. This information can be used to identify the causes of failure which are potentially most critical from risk point of view, and provide directions for eventual efforts towards improvement of the reliability.

The present thesis deals mainly with the process of applying level III reliability methods for determining the probability of failure of a wind turbine blade. Level I methods are also often involved in the discussion, with results from level III calculations used as means to calibrate safety factors which can be then used in level I methods. Figure 2.1 shows an overview of the main processes and inputs involved in the different levels of reliability analysis. The procedure of carrying out reliability analysis can be summarized as follows:

a) Establish a target reliability level, or decision model

b) Identify all significant modes of failure

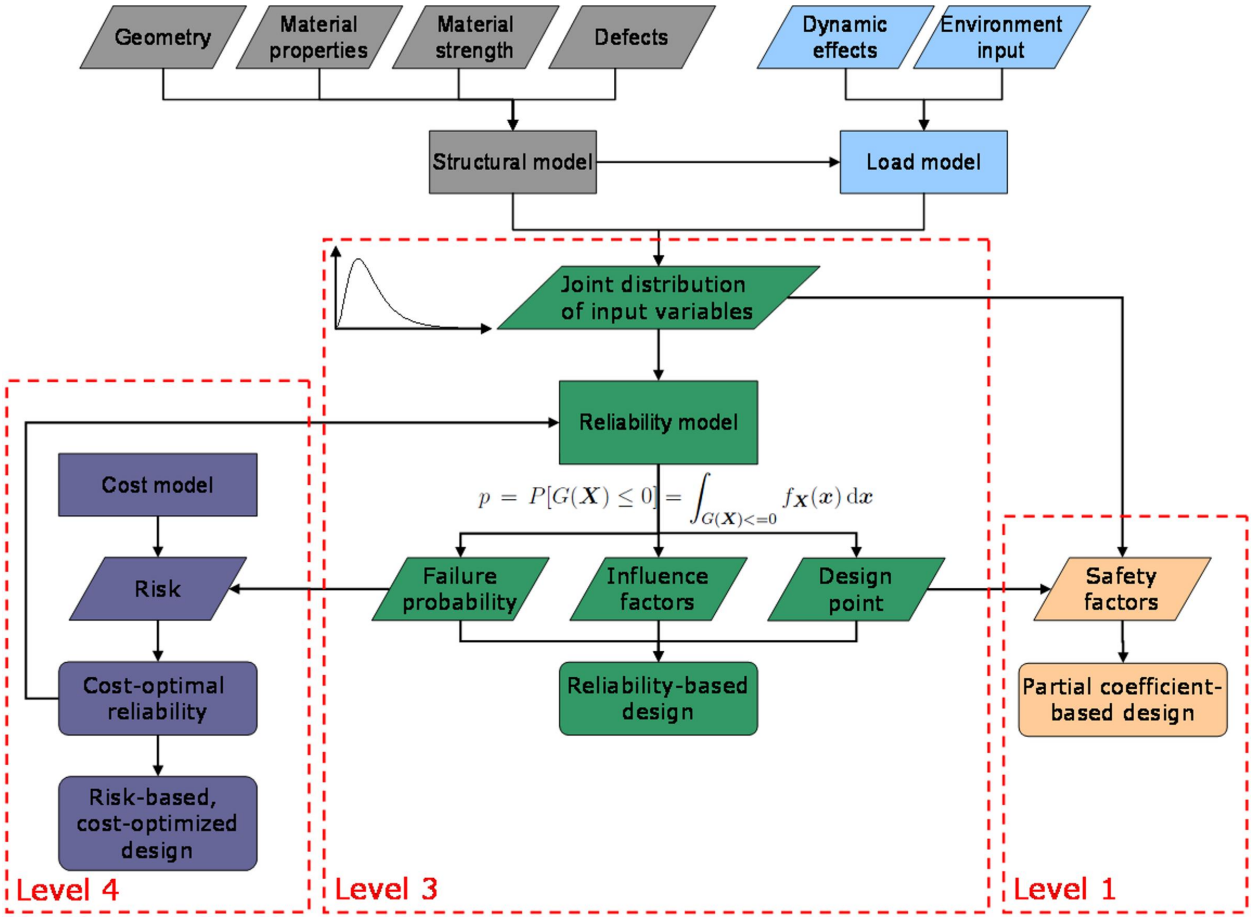

Figure 2.1: Reliability analysis summary.

c) Formulate limit state function for each failure mode

d) Identify stochastic variables and parameters included in the limit states, and define their statistical distributions

e) Carry out component reliability analysis and obtain the probability of failure associated with each failure mode

f) Obtain system probability of failure based on component reliabilities

g) Assess whether the resulting system reliability leads to a sufficiently safe design, and modify the design if necessary

h) Analyze parameter sensitivity results, and evaluate whether they provide any additional information that can improve the design

The remaining parts of this chapter discuss the theoretical aspects of some Level III reliability methods. The selection corresponds to the methods that are used later for calculating reliability of composite structures.

2.2 Structural reliability methods

2.2.1 Limit state equation

The limit state function $g(\mathbf{X})$ is used to describe the state of the structure given input $\mathbf{X}$. Usually $g(\mathbf{X}) = 0$ indicates the failure surface, and $g(\mathbf{X}) \leq 0$ describes the entire failure domain. Solving the reliability problem involves calculating the integral

$$P_f = \int \mathbb{1}_{g(\mathbf{X})\leq 0} f_{\mathbf{X}}(\mathbf{X}) \mathrm{d}\mathbf{X} \tag{2.1}$$

where $\mathbb{1}_{(.)}$ is the zero-one indicator function, and $f_{\mathbf{X}}(\mathbf{X})$ is the joint probability density of the vector of input variables $\mathbf{X}$.

The most commonly used implementations of the limit state equation divide $g(\mathbf{X})$ in two parts related to the capacity (resistance) of the structure, and the demand (usually the load) which it is subjected to. The capacity is denoted by R, and the demand - by S. Common expressions are

$$g(\mathbf{X}) = R(\mathbf{X}) - S(\mathbf{X}) \quad ; \quad g(\mathbf{X}) = \frac{R(\mathbf{X})}{S(\mathbf{X})} - 1 \quad ; \quad g(\mathbf{X}) = \log\left(\frac{R(\mathbf{X})}{S(\mathbf{X})}\right) \tag{2.2}$$

2.2.2 Transformation to standard normal uncorrelated space

Stochastic variables that are part of a reliability analysis can have different scales of magnitude and types of distribution. For practical reasons it can be difficult to operate with such inhomogeneous set of data. Therefore a transformation of the vector of input variables $\mathbf{X}$ into a set of i.i.d (independent and identically distributed), standard normal variables $\mathbf{U}$ is usually applied as part of the reliability analysis procedure:

$$U_i = \Phi^{-1}(F(X_i)) \tag{2.3}$$

where $F(X_i)$ is the cumulative distribution function of the i^{th} stochastic variable X_i, and Φ^{-1} is the inverse standard normal distribution. As the authors Hasofer & Lind (1974, [34]) show, this is an exact transformation, and the values of $\mathbf{X}$ can also be obtained from $\mathbf{U}$ by the inverse transformation:

$$X_i = F^{-1}(\Phi(U_i)) \tag{2.4}$$

Using equation 2.3 ensures that stochastic variables are identically distributed with mean of zero and standard deviation of 1, however it does not ensure independence. Transforming a vector of dependent stochastic variables $\mathbf{X}$ will result in a set of dependent variables $\mathbf{U}$. This vector can then be transformed into a vector of independent variables $\mathbf{Y}$ by decomposing the correlation matrix $\mathbf{R}$ into lower and upper triangular matrices $\mathbf{R} = \mathbf{L}\mathbf{L}^{\mathbf{T}}$, using for example Cholesky decomposition. Then, $\mathbf{Y}$ is obtained by

$$\mathbf{Y} = \mathbf{L}\mathbf{U} \tag{2.5}$$

Using the standard normal transformation makes it possible to define Hasofer-Lind's generalized reliability index as the shortest distance from the origin to the failure surface in U-space:

$$\beta = \min_{g(\mathbf{U})=0} \sqrt{\sum_{i=1}^{n} \mathbf{u}_i^2} = \mathbf{u}^{*T}\mathbf{u}^* \tag{2.6}$$

The point u^* on the failure surface at which the smallest distance to the origin is achieved is the design point. Being based on the standard U-space this definition of the reliability index is invariant with respect to the formulation of the limit state function and the actual distribution of input variables.

2.2.3 Gradient-based methods

Approximate reliability methods such as FORM/SORM (first/second order reliability methods, see e.g., Ditlevsen & Madsen, 1996 [19] or Madsen et al., 1985 [48]) are usually very

fast and efficient as they involve relatively few evaluations of the limit state function. The main assumption behind this class of reliability methods is that the limit state surface can sufficiently be approximated by a first-order (FORM) or a second-order (SORM) polynomial surface. In case of the First-Order method (FORM) the surface parameters are defined in standard normal space by the unit normal vector

$$\boldsymbol{\alpha} = -\frac{\nabla g(\mathbf{U})}{||\nabla g(\mathbf{U})||}\bigg|_{\mathbf{U}=\mathbf{u}^*} \tag{2.7}$$

The linear approximation of the failure surface at a point $\mathbf{u}^*$ will be

$$M = \beta - \boldsymbol{\alpha}^{\boldsymbol{T}}\mathbf{U} \tag{2.8}$$

The probability mass bounded by the hyperplane surface can be assumed as an approximate estimate of the probability mass within the failure domain. Therefore, the FORM probability of failure is estimated by

$$P_f = P(g(\mathbf{U}) \leq 0) \approx P(\beta - \boldsymbol{\alpha}^{\boldsymbol{T}}\mathbf{U}) = \Phi(-\beta) \tag{2.9}$$

The vector $\boldsymbol{\alpha}$ also has a specific meaning. It is the gradient of the reliability index β at the design point $\mathbf{u}^*$, and the elements α_i are therefore measures of the relative influence of stochastic variables u_i to the reliability. The elements of $\boldsymbol{\alpha}$ are subsequently called importance factors.

Design points used in the context of FORM analysis are usually obtained iteratively by means of a gradient-based optimization algorithm which seeks the minimum of $\sqrt{\mathbf{U}^{\mathbf{T}}\mathbf{U}}$ subjected to the constraint $g(\mathbf{U}) = 0$. The results from FORM analysis can also be supplemented by a SORM calculation by evaluating the curvature of the limit state surface at the design point, Breitung, 1984, [7]. In most cases the linear approximation to the limit state surface that the FORM method assumes is reasonably accurate, because most of the probability mass of the failure domain is concentrated around the design point ([19]).

2.2.4 The Model Correction Factor method

Practically any reliability analysis procedure involves a large number of evaluations of the limit state function. This can either be in the form of gradient-based search techniques such as FORM/SORM and response surface (Bucher, 1990 [9]), or as a simulation with large number of random samples, i.e., crude Monte Carlo (Rubinstein, 1981 [71]), importance sampling (Augusti et al., 1984 [3], Bourgund & Bucher, 1986 [5]), asymptotic sampling (Bucher, 2009

[8]), and others. The large number of function calls will often result in prohibitively long computation times, because limit state functions can involve complex analysis routines such as for example nonlinear finite element modelling. The Model Correction Factor method (Ditlevsen & Arnbjerg-Nielsen, 1994, [18]) gives the possibility of alleviating this problem to a large extent by introducing a second, simplified model, which is tuned in a probabilistic sense to the advanced, "realistic" model. The "simplified" model typically requires far less computational resources in comparison with the "realistic" model, while still being able to capture some of the major physics that govern the realistic problem. Therefore it is possible to estimate the reliability of the structure by performing reliability analysis solely on the simplified model. The relationship established between the two models is the so-called model correction factor (MCF):

$$R^{REAL}(\mathbf{X}) = \nu(\mathbf{X}) \cdot R^{IDEAL}(\mathbf{X}) \tag{2.10}$$

Where $\nu(\mathbf{X})$ is the model correction factor; R^{REAL} is the resistance of the realistic model, and R^{IDEAL} is the resistance of the idealistic model.

The above equation 2.10 will always be valid. However, defining the actual function $\nu(\mathbf{X})$ explicitly might be very difficult or even impossible for the numerical models that are generally used. Therefore Ditlevsen & Arnbjerg-Nielsen (1994 [18]) suggest that the model correction factor is approximated by its Taylor-series expansion around the design point of the simplified model. The approach of using the simplified model for reliability analysis can be justified, because the simplified model must capture the main physical phenomena that govern the behaviour of the realistic model.

The reliability of the realistic model is determined in a sequence of iterations, where during each successive step a design point is obtained by reliability analysis of the simplified model. The procedure can be summarised by the following steps:

1. Perform an initial iteration, $i = 0$. Initialise $\nu(\mathbf{X})$ by solving eq. 2.10 at the mean values $E(\mathbf{X})$.

2. $i = i+1$. Obtain the design point $\mathbf{X}_i^*$ by performing reliability analysis of the idealised model.

3. Determine the realistic response $R^{REAL}(\mathbf{X}_i^*)$ at the newly obtained design point.

4. Solve (2.10) for the design point $\mathbf{X}_i^*$ and obtain an updated value of the model correction factor $\nu_i(\mathbf{X}_i^*)$.

5. Has the series of design points converged? If not, go back to step 2

6. If the series of design points has converged, the reliability of the realistic model can be approximated by the reliability found using the simplified model obtained at the final design point.

Using a zero- or first-order approximation around the design point for the idealized model $\mathbf{X}^*$ will in most cases give sufficiently precise approximations. It should be noted however that convergence to the correct design point is not assured.

An example application of the Model Correction Factor method to the reliability analysis of a composite blade section can be found in Dimitrov et al., 2013 [17].

2.2.5 Simulation-based methods

Crude Monte Carlo

Monte Carlo simulation technique (e.g., Rubinstein, 1981 [71]) is a very robust and straightforward method for determining probability of failure. The Monte Carlo (MC) method does not use any approximations of the limit state surface, and it is not limited to finding local extremes (as are some of the optimisation methods used together with FORM/SORM analysis). As a result, the MC technique is well suited for performing system reliability analysis, and it is one of the few reliability methods whose performance is not sensitive to the number of dimensions in the vector of random variables. Therefore Monte Carlo simulation is very often used as a method for validation of other reliability analysis procedures, or in cases when other reliability methods fail to achieve a solution.

Recalling equation 2.1:

$$P_f = \int \mathbb{1}_{g(\mathbf{X})\leq 0} f_{\mathbf{X}}(\mathbf{X}) \mathrm{d}\mathbf{X}$$

Equation 2.1 represents the first moment of $\mathbb{1}_{(\cdot)}$, so an unbiased estimator of equation 2.1 is:

$$P_f = \frac{1}{N} \sum_{j=1}^{N} \mathbb{1}_{g(\mathbf{x}_j)\leq 0} \tag{2.11}$$

Equation 2.11 represents the classical Monte Carlo method.

A limitation of the crude Monte Carlo simulation method is the large number of trials required, typically $200/P_f$ to achieve a 10% coefficient of variation of the failure probability P_f (Shooman, 1968 [73]), which implies that the method is very time consuming to apply in connection with the small failure probabilities which are usually associated with structural

design. However, a number of variance reduction techniques exist - for example directional simulation (Bjerager, 1988, [4]), adaptive VEGAS algorithm (Lepage, 1978, [45]), importance sampling (Augusti et al., 1984 [3]), adaptive (search-based) importance sampling (Melchers, 1990 [50]), asymptotic sampling (Bucher, 2009 [8]). All the listed methods have the common feature that the sampling density is modified in order to obtain more failure outcomes per number of trials. With the importance sampling technique in particular the sampling density is shifted from the mean point to a point closer to the failure domain.

Importance sampling

If Equation 2.1 is rewritten in the following way:

$$P_f = \int_{All\ \mathbf{x}} \mathbb{1}_{g(\mathbf{X})\leq 0}\ \frac{f_{\mathbf{X}}(\upsilon)}{h_{\mathbf{V}}(\upsilon)} h_{\mathbf{V}}(\upsilon)\mathrm{d}\upsilon \tag{2.12}$$

where $h_{\mathbf{V}}()$ is the importance sampling probability density function, then P_f can be estimated by:

$$P_f = \frac{1}{N}\sum_{j=1}^{N}\left(\mathbb{1}_{g(\mathbf{V}_j)\leq 0}\frac{f_{\mathbf{X}}(\mathbf{V}_j)}{h_{\mathbf{V}}(\mathbf{V}_j)}\right) \tag{2.13}$$

where $\mathbf{V}_j$ is a vector of sample values taken from the importance sampling distribution $h_{\mathbf{V}}(\cdot)$. With a proper choice of the sampling distribution the variance in J can be greatly reduced, a larger fraction of the trials will contribute to the probability sum, and as a result the solution will converge with fewer trials. An appropriate choice for $h_{\mathbf{V}}(\cdot)$ should aim at extracting as much information as possible from each sampling point $\mathbf{V}_j$, and this can be achieved by having samples with large probability density (close to the point with maximum likelihood) and falling within the failure domain ($[G(\mathbf{X}_j) \leq 0]$).

In practice it is very difficult to determine $h_{\mathbf{V}}(\cdot)$ such that all sample points fall within the failure domain. A robust and easily applicable approach of choosing the sampling point is to place the sample mean at the point from the failure domain which has maximum likelihood - i.e., the design point determined from another reliability analysis (Bourgund et al., 1988 [5]). This will ensure that approximately half of the samples will fall within the failure domain.

Asymptotic sampling

The asymptotic sampling method (Bucher, 2009 [8]) relies on the asymptotic behaviour of the failure probability in n-dimensional i.i.d standard normal space as the standard deviation σ of the variables and hence the failure probability P_f approaches zero. Introducing a scaling

variable $f = \frac{1}{\sigma}$, it is assumed that the reliability index β will behave asymptotically as $\beta \to \infty$, with the following functional dependence:

$$\beta = A \cdot f + \frac{B}{f} \tag{2.14}$$

The coefficients A and B are determined by generating several random samples using different values of f, and applying regression analysis to the obtained reliability estimates. For fitting purposes, equation 2.14 is rewritten in terms of a scaled reliability index $\frac{\beta}{f}$ as

$$\frac{\beta}{f} = A + \frac{B}{f^2} \tag{2.15}$$

This formulation ensures equal weight to all support points for the regression analysis. The procedure of carrying out Asymptotic Monte Carlo analysis typically starts with choosing an initial value for $f < 1$ such that enough points fall within the failure domain as to obtain a sufficiently accurate estimate for the scaled reliability index $\frac{\beta}{f}$. Subsequent simulations are then carried out with further stepwise reduction of the scaling factor f, until enough support points are obtained.

Importance factors from Monte Carlo analysis

When using FORM, importance factors are directly obtained as their calculation is part of the reliability analysis procedure. For simulation-based methods such as crude Monte Carlo however importance factors are not a direct outcome of the analysis. It is useful to have a measure of the influence of the stochastic variables also when using Monte Carlo, and for that purpose the approximate technique by Melchers & Ahammed, 2004 [50] can be applied, where importance factors are estimated by fitting a hyperplane to the observations of the limit-state function values which fall close to the limit state surface (i.e., the sample points for which $g(\mathbf{X}) \approx 0$). The linearized version of the limit state function is given by

$$g_L(\mathbf{X}) = a_0 + \sum_{i=1}^{n} a_i X_i = 0 \tag{2.16}$$

where (a_0, a_i), $i = 1, \ldots, n+1$ are the coefficients of the hyperplane fit to the outcomes of the original limit state function, $g(\hat{\mathbf{x}})$. Since the amount of samples for which $g(\hat{\mathbf{x}}) \approx 0$ might be very limited, a less accurate but more practical approach is to use all points in the failure domain, $g(\hat{\mathbf{x}}) \leq 0$ for the fit. The mean and standard deviation of g_L are given by

$$\begin{aligned} \mu_g &= a_0 + a_1\mu_{X1} + a_2\mu_{X2} + \ldots \\ \sigma_g &= \left[\sum_{i=1}^{n} \left(a_i \sigma_{X_i}\right)^2 \right]^{1/2} \end{aligned} \tag{2.17}$$

When the variable space $\mathbf{X}$ is mapped to the standard normal space $\mathbf{U}$, the gradients of the reliability index will follow directly from FORM theory:

$$\left.\frac{\partial \beta}{\partial U_i}\right|_{\mathbf{u}^*} = \alpha_i = \frac{a_i \sigma_{X_i}}{\sigma_g} \tag{2.18}$$

2.2.6 Calibration of partial safety factors

In most practical cases design requirements are based on deterministic analysis, and the uncertainties in model predictions are accounted for by using safety factors. Reliability analysis is rarely used directly as a design tool, however it also gives a possibility for tuning the partial safety factors that can subsequently be used in a classical rule-based design process.

The standard approach at using safety factors is dividing the overall safety to several partial safety factors related to the loads, the material strength, consequences of failure, and other. For a limit state equation consisting of two parts, the resistance R and the load S, a corresponding design equation can be defined

$$g = \frac{1}{\gamma_m \gamma_n} R_c - \gamma_f S_c \tag{2.19}$$

where γ_m, γ_f and γ_n are partial safety factors for material strength, load effects and consequences of failure respectively, R_c is the characteristic (typically 5%) quantile value of the resistance of the structure, and S_c is the characteristic (98% quantile) value of the loads. If by means of reliability analysis a design point $\mathbf{X}^*$ is obtained such that $G(\mathbf{X}^*) = R^* - S^* = 0$, and knowing the characteristic values of R and S, the partial safety factors can be estimated by

$$\gamma_m \gamma_n = \frac{R_c}{R^*} \quad , \quad \gamma_f = \frac{S^*}{S_c} \tag{2.20}$$

By definition at the design point from reliability analysis $R^* = S^*$. Therefore if one is interested simply in the combined safety factor $\gamma = \gamma_m \gamma_f \gamma_n$, it can be found by $\gamma = R_c / S_c$.

In some cases it might be desired to calibrate partial safety factors against a target reliability level. The target reliability is achieved by adjusting the limit state equation by introducing an additional design parameter z. If for example the partial safety factor for structural resistance γ_m needs to be calibrated against a given reliability while the other partial safety factors are kept unchanged, the modified limit state equation will read

$$g(\mathbf{X}) = R(z\mathbf{X}_R) - S(\mathbf{X}_S) \tag{2.21}$$

Above, the vector of stochastic variables $\mathbf{X}$ is divided into components $\mathbf{X}_R$ and $\mathbf{X}_S$ representing the parts of the stochastic input that have influence on the structural resistance and on the loads, respectively. Applying the design parameter directly on the random variables assures that the results will be invariant to the formulation of the limit state function. When the design parameter z is determined from equation 2.21 for the given target reliability level, the calibrated partial safety factor is obtained by solving the design equation for γ_m:

$$g = R\left(z\frac{1}{\gamma_m}\gamma_n\mathbf{X}_{R,c}\right) - S\left(\gamma_f\mathbf{X}_{S,c}\right) = 0 \tag{2.22}$$

2.2.7 System reliability analysis

When multiple failure modes are present in a structure it is considered to be a reliability system. The state of a system is related to the state of its components through logical unions and intersections. The simplest types of system are the parallel and the series systems. In a series system the failure probability is the union of the component failure probabilities, i.e., failure of any component will lead to overall system failure. In parallel systems the failure probability is given by the intersection of component failure probabilities, meaning that the overall failure of the system is reached when all components have failed. There are systems that can neither be characterized as parallel nor as series systems. In these so-called general systems failure may occur when several, but not all of the system components have failed.

The limit state surface corresponding to a reliability system is represented as multiple safety margins. The overall failure surface is also non-linear and non-continuous at the intersection points (see e.g., Ditlevsen, [19] for details). FORM/SORM reliability methods are based on the assumption of a single failure boundary which is approximately linear, and are therefore not able to directly predict the reliability of a system with multiple failure modes. Instead, each of the system components needs to be treated separately in a component reliability analysis.

Simulation-based methods such as crude Monte Carlo or Asymptotic Sampling that use the mean of the stochastic variables as a sampling center point are not affected by presence of multiple safety margins and can be therefore applied directly in a system reliability analysis. Importance sampling techniques however rely on a modified sampling center point, which in the case of multiple safety margins may be placed close to only one of the limit states, while entirely omitting the other ones. It is therefore not guaranteed that an importance sampling calculation will be able to correctly estimate the system failure probability.

Series and parallel systems

The probability of failure for a series or a parallel system can be obtained by applying the multivariate normal distribution to the component reliability indexes obtained by considering each component (or failure mode) separately. For series and parallel systems the expressions for system failure probability are

$$P_f = 1 - \Phi_k(\boldsymbol{\beta}, \mathbf{R}) \qquad \text{(Series system)} \tag{2.23}$$

$$P_f = \Phi_k(-\boldsymbol{\beta}, \mathbf{R}) \qquad \text{(Parallel system)} \tag{2.24}$$

where $\boldsymbol{\beta}$ is the vector of component reliability indexes, $\mathbf{R}$ is the matrix containing the correlation coefficients between component reliabilities, and $\Phi_k(\cdot)$ is the k-dimensional multivariate standard normal distribution. The correlation coefficient between components i and j can be estimated from the importance factors resulting from the component reliability analysis, $R_{ij} = \boldsymbol{\alpha}_i^T \boldsymbol{\alpha}_j$. There is no closed-form expression for the multivariate normal density, and it has to be evaluated either by direct numerical integration (e.g., Drezner, 1992 [20]), or by some approximate method as for example Gollwitzer & Rackwitz, 1988 [28], and Pandey, 1998 [59]. In the present thesis the efficient asymptotic approximation described by Pandey, 1998 [59] is used.

General systems

While the reliability of parallel and series systems can be approximated in a relatively straightforward way, solving general systems requires (except when using simulation) more advanced considerations. The solution methods which the author has reviewed can in principal be divided into four groups, described below.

Simulation methods Simulation-based methods are well suited for analyzing large systems of any type as their validity and efficiency is not influenced by the number of system components and relations between them, as it is not influenced by the number of stochastic variables. Applicable methods in this class include crude Monte Carlo, adaptive (search-based) importance sampling (Melchers, 1990 [50]), asymptotic sampling (Bucher, 2009 [8]).

Cut-set and tie-set representations Every general system can in principle be represented either as a series system of parallel subsystems or a parallel system of series subsystems. The former representation involves the definition of so-called minimal cut-sets, and

the latter corresponds to defining minimal tie-sets. Methods involving cut-set or tie-set definitions are e.g., Sequential Compounding (Kand & Song, 2010 [42]), linear programming (Song & Der Kiureghian, 2003 [76]), or direct multivariate integration of the cut set / tie set when possible.

Order statistic By using order statistic each of the correlated component events can be transformed into independent events, and their joint probability distributions can be computed (Friis-Hansen, 1994 [25]). This approach can be very useful for systems where the reliability of each component is dependent on a single stochastic variable. Using more than one variable per component makes it necessary to introduce two limit states, where one describes the component failure based on the input stochastic variables, and the other describes the system failure, based on the values of the component limit states.

Bounds by direct integration of the safe set One of the features of a general system is that some of the component failures will not lead to system failure. Using this property, reliability bounds for a general system can also be determined by evaluation of all component failure sequences leading to a point in the safe domain. This, combined with the probability of no component failures, will result in an evaluation of the probability of the safe event set (Ditlevsen & Madsen, 1996 [19])

Chapter 3

Modelling Composite Materials Behaviour

3.1 Structural modelling of composites

The word "composite" originates from Latin, and literally means "made up of distinct parts". This is indeed the most distinct feature of composite materials, which constitute of at least two different material phases and as a result have a number of unique properties. The combination of two material phases (fibers and matrix), arranged in multiple layers, results in a material with non-isotropic, non-homogeneous elastic strength properties, which can fail in a number of different failure modes.

3.1.1 Length scales

The causes of mechanical behavior of composite structures can be traced all the way to the microscopic level where interaction between the fiber surfaces and the surrounding matrix material takes place. As sketched on Figure 3.1, aligning the individual fibers along a common direction in a thin layer results in the so-called lamina, which due to the unidirectional fiber alignment, has different mechanical properties in different directions. A lamina may fail in at least two different failure modes - when subjected to load parallel to the fiber direction, the lamina will fail in a mode dominated by the fiber breakage strength, while for a load transverse to the fiber direction the failure will be dominated by the matrix strength. Moving up to larger structural scales introduces new possible failure mechanisms: an ordered stack of laminas, called a laminate (see Figure 3.1), may fail when any of the constituent laminas fails, or if the interface material, connecting adjacent lamina, fails too.

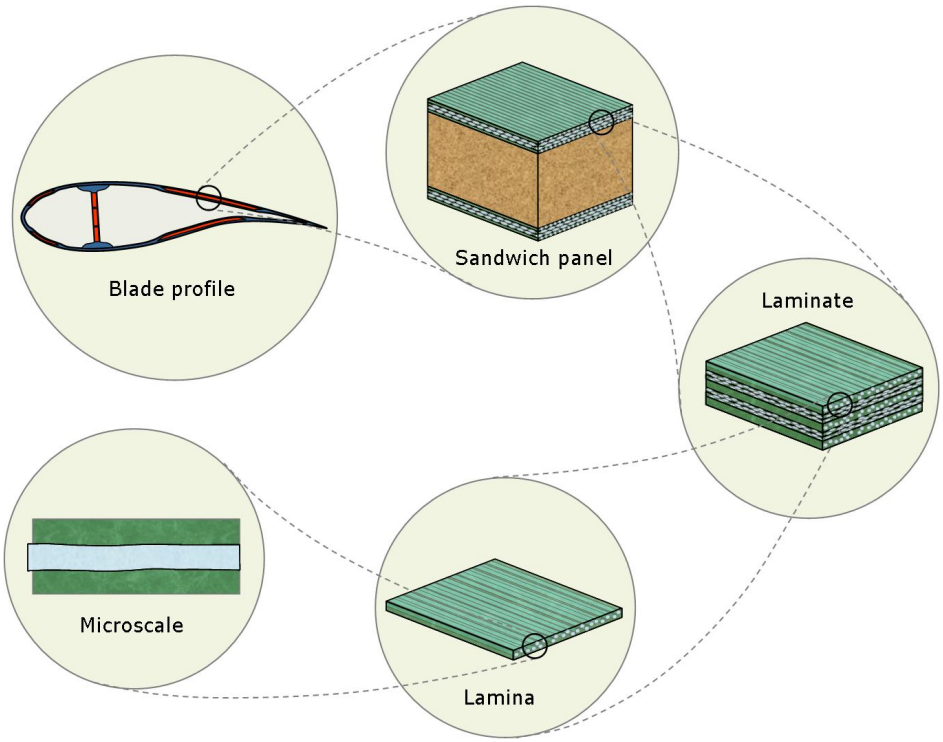

Figure 3.1: Length scales in composite materials and structures.

A sandwich panel, where two composite laminate faces are separated by a lightweight core in order to increase the flexural rigidity of the structure, will fail if either of the outer faces fails in tension, compression or local buckling (called wrinkling). Core fracture (typically in shear) is another possible mode of sandwich failure. Finally, sandwich panels as well as monolithic laminated plates are the components used to build full-scale composite structures and sub-structures such as wind turbine blades. On the structural level, again a number of failure mechanisms are possible, at different locations and different loading conditions. For a wind turbine blade for example a typical set of common ultimate failure modes is: extreme material failure, fatigue, buckling, and extreme deflection (leading to a collision of the blade with the wind turbine tower, possibly resulting in the destruction of the whole turbine). In order to avoid excessive complexity, an analysis of a composite structure needs to be limited to a reasonable span of structural size scales. Including the microscopic scale levels will mean that micro-mechanics must be involved in calculating structural characteristics, along with the mechanics for larger size scales based on Hooke's law. On the other hand, considering the entire structure as a detailed reliability system with its possibly unlimited number of different failure modes is not practically possible. The problem can instead be limited to

smaller parts of the structure such as single components or hot-spots, which will have far less structural complexity and will be subject to similar loading conditions throughout the modeling domain. If necessary, the results from such analyses can be further compounded into a global reliability estimate by the use of additional system reliability analysis. In the present thesis most of the modeling is limited to three scale levels - individual lamina, laminated panels, and sandwich panels. Taking into account individual unidirectional lamina allows modeling the non-isotropic mechanical properties of composites, while avoiding the use of micro-mechanics. Including laminates and sandwich panels into the analysis allows modeling the reliability against failure of a given location within a structure - for example, a so-called "hot-spot" which is known or expected to be critically decisive for the overall safety of the structure.

3.1.2 Modelling of laminated composite plates

When detailed and accurate representation of the properties of composite structures is needed, the most commonly used solution is advanced finite-element modelling. There are cases however when simplicity is more important than the level of detail, and faster, approximate approaches are preferred. In the present thesis, depending on the problem that needs to be solved, both finite-element and simplified analytical models are used for modelling composite structures. Here, a short outline of the analytical methods used to model composite laminates and sandwich panels is given.

One of the most widely used and accepted methods for modelling the global laminate behaviour is the so-called Classical Lamination Theory (see e.g., Reisner, 1961 [66], or more recently Jones, 1998 [39]). This theory idealizes the composite laminate as a thin plate with homogeneous properties, which are derived from the underlying lamina properties and the orientations of their axes. The underlying assumptions and kinematic relations are analogous to the ones used in Kirchhoff-Love thin plate theory (Love, 1888 [47], also described in Reddy, 2007 [65]): (a) Straight lines normal to the mid-surface remain straight and normal to the mid-surface after deformation; (b) The thickness of the plate does not change during deformation; and (c) The thickness of the plate is much smaller than its other dimensions. In order to account for the directionality in properties, specific non-isotropic stress-strain relations must be introduced. For that purpose, two additional assumptions for the lamina are introduced: (d) The lamina are predominantly subjected to a plane state of stress; and (e) The lamina properties in the two non-primary material directions are identical (so-called transverse isotropy assumption). This means that the elastic properties of a lamina are

represented by 4 independent elastic constants. The strength properties are given by 5 constants, making a total of 9 material constants characterizing each lamina.

Elastic properties of composite laminates

Under the above assumptions, the stress-strain relations for the laminate are derived by integrating individual layer properties through the thickness of the laminate, where the properties of each lamina are subjected to a coordinate transformation corresponding to the angle of orientation of the lamina. The global stress-strain relation takes the form

$$\left[\begin{array}{c} \mathbf{N} \\ -- \\ \mathbf{M} \end{array}\right] = \left[\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ -- & -- \\ \mathbf{B}^T & \mathbf{D} \end{array}\right] \left[\begin{array}{c} \boldsymbol{\varepsilon^0} \\ -- \\ \boldsymbol{\kappa} \end{array}\right] \tag{3.1}$$

where $\mathbf{N} = [N_x, N_y\ N_{xy}]^T$ are forces per unit length, $\mathbf{M} = [M_x,\ M_y,\ M_{xy}]^T$ are moments per unit length, $\boldsymbol{\varepsilon^0} = [\varepsilon_x^0,\ \varepsilon_y^0,\ \varepsilon_{xy}^0]^T$ are strains, and $\boldsymbol{\kappa} = [\kappa_x,\ \kappa_y,\ \kappa_{xy}]^T$ are rotation rates. $\mathbf{N}$ and $\mathbf{M}$ represent global force measures derived by integrating layerwise stresses through the thickness. The 3×3 sub-matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{D}$ store the elastic coefficients and are derived by layerwise integration of the in-plane stiffness matrices for the individual lamina. $\mathbf{A}$ represents the membrane stiffness of the laminate, $\mathbf{D}$ describes the bending (flexural) stiffness, and $\mathbf{B}$ represents the coupling between bending and membrane stresses. Once the sub-matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{D}$ are defined, the stress-strain relations described above are used to determine the stresses and strains in the principal material directions for each lamina, thus making it possible to assess the structural integrity of the laminate.

Modelling of strains and displacements in laminated composite plates

Laminated composite plates are one of the most common applications of the Classical Lamination Theory. The modelling of such plates is described in details in Zenkert & Batley, 2006 [93]. Here, a short overview of some of the basic equations is given. The solutions are based on the assumptions from Kirchoff-Love plate theory discussed above. The mid-surface strains are given in terms of the displacements:

$$\left[\begin{array}{c} \varepsilon_{x0} \\ \varepsilon_{y0} \\ \varepsilon_{xy0} \end{array}\right] = \left[\begin{array}{cc} \frac{\partial}{\partial x} & 0 \\ 0 & \frac{\partial}{\partial y} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} \end{array}\right] \left[\begin{array}{c} u_0 \\ v_0 \end{array}\right] \quad ; \quad \left[\begin{array}{c} \kappa_x \\ \kappa_y \\ \kappa_{xy} \end{array}\right] = \left[\begin{array}{c} -\frac{\partial^2}{\partial x^2} \\ -\frac{\partial^2}{\partial y^2} \\ -2\frac{\partial^2}{\partial x \partial y} \end{array}\right] w \tag{3.2}$$

The above can be written out in short matrix form as

$$\boldsymbol{\varepsilon}_0 = \mathbf{Ku} \quad ; \quad \boldsymbol{\kappa} = \mathbf{Lw} \tag{3.3}$$

where $\mathbf{u} = [u_0, v_0]^T$.

Using the force-strain relations defined by the **ABD**-matrix from equation 3.1, the force equilibrium equations for a composite laminate plate are

$$\begin{aligned} \mathbf{K}^T\mathbf{AKu} + \mathbf{K}^T\mathbf{BL}w &= 0 \\ \mathbf{L}^T\mathbf{BKu} + \mathbf{L}^T\mathbf{DL}w &= q^* \end{aligned} \tag{3.4}$$

where

$$q^* = q + N_x\frac{\partial^2 w}{\partial x^2} + N_y\frac{\partial^2 w}{\partial y^2} + 2N_{xy}\frac{\partial^2 w}{\partial x \partial y}, \tag{3.5}$$

and q is a distributed transverse load. Solving the equilibrium equations gives the possibility to model the in-plane deformations, out-of-plane bending, buckling, and free vibration of a laminated composite plate. The solution can be rather complicated in cases when coupling terms are present in the **ABD** matrix, as this causes coupling between the out-of-plane displacement w and the in-plane displacements u, v. Often simplifications are possible which would eliminate some of the terms in the equations, for example using a symmetric laminate where all terms in the **B** matrix are equal to 0, or a specially orthotropic laminate, where bending-extension couplings (matrix **B**), shear-extension couplings (elements A_{16} and A_{26} from the **A** matrix), and bend-twist couplings (elements $\mathbf{D_{16}}$ and $\mathbf{D_{26}}$ from matrix **D**) are all eliminated.

In the case of bending of a rectangular plate, the deflections are expressed by infinite sums of sinusoidal half-wave deflection shapes (see Zenkert & Batley, 2006 [93]):

$$\begin{aligned} w &= \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} w_{mn} \sin\frac{m\pi x}{a}\sin\frac{n\pi y}{b} \\ u &= \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} u_{mn} \cos\frac{m\pi x}{a}\sin\frac{n\pi y}{b} \\ v &= \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} v_{mn} \sin\frac{m\pi x}{a}\cos\frac{n\pi y}{b} \end{aligned} \tag{3.6}$$

In the above, a and b are respectively the length and width of the plate, m and n represent the number of half-waves present in the deflection shape, and w_{mn}, u_{mn} and v_{mn} are coefficients

dependent on the load distribution and the elastic constants of the plate. The out-of plane loading $q(x, y)$ is expressed in a similar fashion:

$$q(x,y) = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} q_{mn} \sin\frac{m\pi x}{a} \sin\frac{n\pi y}{b} \tag{3.7}$$

where q_{mn} is a coefficient dependent on the load magnitude and location.

3.1.3 Modelling of composite sandwich panels

In sandwich panels, the thickness of the core is much larger than the face thickness. This means that the variation of normal stresses in the faces will be relatively small, and under certain conditions it can be disregarded, assuming that the normal stresses are constant across the face (see Zenkert, 1995 [92]). Under this assumption, the sandwich model can be simplified by replacing the layer-wise description of the face laminate with equivalent, homogeneous orthotropic layer.

An analytical, linear elastic solution for a rectangular, orthotropic, simply supported sandwich plate is used to approximate composite sandwich properties, Zenkert, 1995 [92]. The model is capable of estimating the stiffness, the stress and deflection distribution, and the buckling capacity of the plate. Similarly to the laminated plate model, the out-of-plane deflection and force distributions are obtained by infinite sums of sinusoidal half-wave deflection shapes, and the membrane and transverse stresses are superimposed with the assumption that they are independent of each other. All derivations are based on the assumption that the faces are thin compared to the core and the stress within them can be considered uniform, while the core is weaker than the faces and does not carry any membrane stresses. It is considered that the assumptions for thin faces and weak core are valid when the following conditions are fulfilled:

$$\frac{d}{t_f} > 5.77(\text{thin faces}) \qquad ; \qquad \frac{6E_f t_f d^2}{E_c t_c^3} > 100(\text{weak core}) \tag{3.8}$$

where t_f and t_c are the face and core thickness respectively, $d = t_c + t_f$ is the reference thickness of the sandwich, E_f is the apparent elastic modulus of the face, and E_c is the elastic modulus of the core.

Elastic properties

The properties of the plate are fully described by 7 constants - the flexural rigidities D_x and D_y, the twist stiffness D_{xy}, the Poisson ratios ν_x and ν_y, and the shear stiffnesses S_x and S_y.

These plate constants are given by the following:

$$D_x = \frac{E_{x1}t_1E_{x2}t_2d^2}{E_{x1}t_1+E_{x2}t_2} \quad , \quad D_y = \frac{E_{y1}t_1E_{y2}t_2d^2}{E_{y1}t_1+E_{y2}t_2}$$

$$D_{xy} = \frac{2G_{xy1}t_1G_{xy2}t_2d^2}{2G_{xy1}t_1+G_{xy2}t_2} \quad , \quad \frac{\nu_{xy}}{D_x} = \frac{\nu_{yx}}{D_y} \tag{3.9}$$

$$S_x = \frac{G_{cx}d^2}{t_c} \quad , \quad S_y = \frac{G_{cy}d^2}{t_c}$$

where subscripts 1, 2 and c refer to the lower face, upper face, and core, respectively. E_x and E_y are the apparent elastic moduli of the faces in x and y direction, D_x and D_y are flexural rigidities, M_x and M_y are bending moments, N_x and N_y are membrane forces, T_x and T_y are shear forces, t is the face thickness, and $d = t_c + (t_1 + t_2)/2$ is the sandwich reference thickness.

Stress distribution

The stress quantities most often used in sandwich plate analysis are the in-plane stresses in the faces, and transverse shear stresses in the core. They are given by the following expressions:

$$\sigma_{fx1} \approx \frac{N_xE_{x1}}{E_{x1}t_1+E_{xc}t_c+E_{x2}t_2} - \frac{M_xE_{x1}E_{x2}t_2d}{D_x(E_{x1}t_1+E_{x2}t_2)} \quad , \quad \sigma_{fx2} \approx \frac{N_xE_{x2}}{E_{x1}t_1+E_{xc}t_c+E_{x2}t_2} + \frac{M_xE_{x1}E_{x2}t_1d}{D_x(E_{x1}t_1+E_{x2}t_2)}$$

$$\sigma_{fy1} \approx \frac{N_yE_{y1}}{E_{y1}t_1+E_{yc}t_c+E_{y2}t_2} - \frac{M_yE_{y1}E_{y2}t_2d}{D_y(E_{y1}t_1+E_{y2}t_2)} \quad , \quad \sigma_{fy2} \approx \frac{N_yE_{y2}}{E_{y1}t_1+E_{yc}t_c+E_{y2}t_2} + \frac{M_yE_{y1}E_{y2}t_1d}{D_y(E_{y1}t_1+E_{y2}t_2)}$$

$$\tau_{cxz} = \frac{T_x}{d} \quad , \quad \tau_{cyz} = \frac{T_y}{d} \tag{3.10}$$

The stress values obtained using equation 3.10 are approximate, because they are derived based on the assumption of thin faces and weak core, i.e., that the faces are thin and the normal stresses in them can be considered uniform while the shear stresses can be neglected, and that the core is compliant (weak) enough so that the normal stresses in it can be neglected.

Out-of-plane deflection

The out-of-plane deflection is approximated by double Fourier sine series similar to the ones in equation 3.6:

$$w = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} -\frac{W_{mn}q_{mn}}{Z_{mn}} \sin\frac{m\pi x}{a} \sin\frac{n\pi y}{b} \tag{3.11}$$

where W_{mn} and Z_{mn} are a set of coefficients, dependent on the load distribution and plate elastic constants, and q_{mn} is a function of the transverse loads acting on the plate.

Buckling capacity

Determining the analytical buckling capacity is done by finding the stability limit, which is reached when the deflection tends to infinity, $w \to \infty$. In order this condition to be fulfilled, the coefficient Z_{mn} from equation 3.11 has to become 0:

$$Z_{mn} = -N_x \left(\frac{m\pi}{a}\right)^2 W_{mn} + \left(\frac{m\pi}{a}\right) X_{mn} - \left(\frac{n\pi}{b}\right) Y_{mn} = 0 \tag{3.12}$$

where X_{mn} and Y_{mn} are again coefficients dependent on the loads, plate dimensions and plate elastic properties. Solving equation 3.12 for N_x and for every value of m and n results in obtaining a set of buckling loads as function of the number of sine half-waves which the buckling mode shape represents.

3.1.4 Failure of composite laminates

Identifying failure events. Damage evolution algorithms

Due to the presence of a number of layers, the failure of a laminated composite will often happen as a gradual event, with different layers failing in sequence. Between successive layer failures the load previously carried by the failed lamina is redistributed to the layers remaining intact. Consequently, the definition of an ultimate failure event for a composite structure can to a high extent be a matter of subjective judgement. Probably the simplest approach to defining failure is identifying the so-called "first-ply failure" event, where failure of any of the laminate components (i.e., the first matrix or fiber failure occurring under gradually increasing load) will be considered as a total structural failure. Such an approach can however lead to very inefficient designs, because typically the first failures are associated with matrix cracking, which does not have a significant effect on the residual stiffness and strength of the structure (see Figure 3.2). There is also some degree of redundancy in composite structures, meaning that the structure will often be able to withstand loads higher than the load at which the first fiber failure has occurred.

A progressive failure analysis procedure (Figure 3.3) can be used to find the maximum load which the structure can withstand while maintaining static equilibrium. This is the approach which most fully describes the failure process in the panel, however it might not be very useful for design purposes, because large amounts of irreversible damage can occur

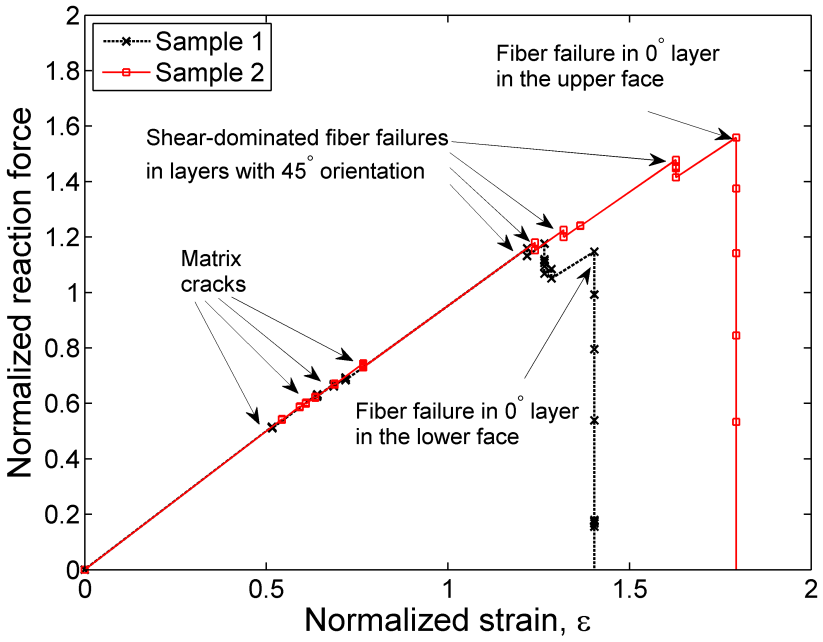

Figure 3.2: Example load-response curve of a sandwich panel subjected to compressive axial loading, simulated with a progressive failure algorithm.

at load levels below the indicated ultimate strength. Finally, the approach adopted in this thesis is to use a progressive failure analysis procedure to find the load level at which the first fiber breaching occurs. The use of progressive failure analysis means that the development of matrix-related failure events will be followed until the first fiber failure occurs. Although not describing the ultimate failure, this approach might be more realistic for design purposes as the failure event is at the point when the first significant damage occurs.

Static failure criteria

In order to assess what is the load at which damage is initiated in a certain lamina, a failure criterion needs to be applied. Typically failure criteria evaluate the stresses or strains in the principal material directions for each individual lamina, and compare these to a limit value determined by a certain physical or (more often) empirical law.

One of the widely used failure criteria is the so-called Hashin criterion [32]. The Hashin failure criterion is a stress-based criterion, featuring shear interaction as well as distinguishing between matrix and fiber failure modes. The ability of the Hashin criterion to distinguish between failure modes offers an advantage over uni-modal criteria such as the Tsai-Wu

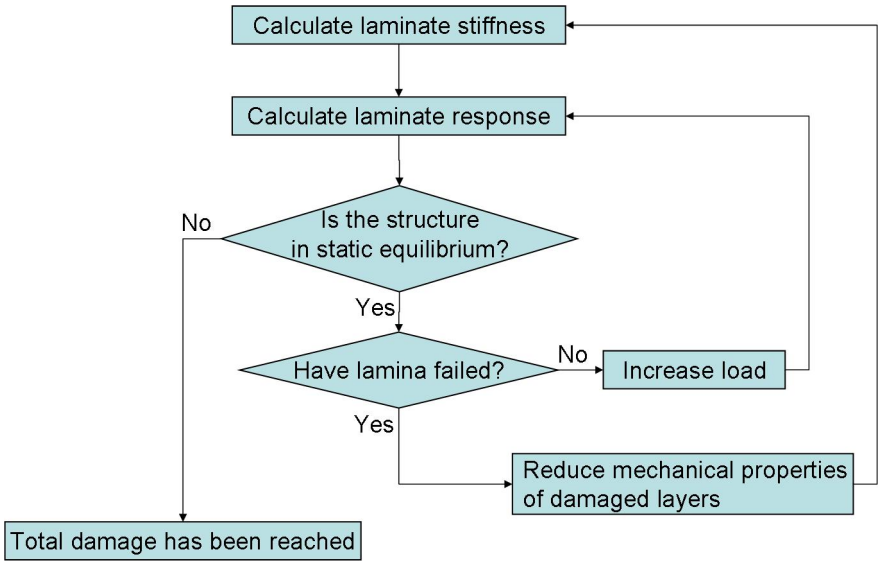

Figure 3.3: Flow chart of a progressive failure algorithm.

criterion [87] for applications involving progressive failure analysis of composites, because it enables the user to modify the stiffness of a failed lamina in the particular material direction where the failure has occurred. The Hashin criterion involves the computation of four failure indices describing the material conditions in fiber tension and compression, and matrix tension and compression (see equation (3.13):

$$\begin{aligned}
F_f^T &= \left(\tfrac{\sigma_{11}}{X^T}\right)^2 + \alpha\left(\tfrac{\tau_{12}}{S^L}\right)^2 \qquad (Fiber\ tension, \sigma_{11} > 0) \\
F_f^C &= \left(\tfrac{\sigma_{11}}{X^C}\right)^2 \qquad (Fiber\ compression, \sigma_{11} \leq 0) \\
F_m^T &= \left(\tfrac{\sigma_{22}}{Y^T}\right)^2 + \left(\tfrac{\tau_{12}}{S^L}\right)^2 \qquad (Matrix\ tension, \sigma_{22} > 0) \\
F_m^C &= \left(\tfrac{\sigma_{22}}{2S^T}\right)^2 + \left[\left(\tfrac{Y^C}{2S^T}\right)^2 - 1\right]\tfrac{\sigma_{22}}{Y^C} + \left(\tfrac{\tau_{12}}{S^L}\right)^2 \qquad (Matrix\ compression, \sigma_{22} \leq 0)
\end{aligned} \tag{3.13}$$

where $\sigma_{11}, \sigma_{22}, \tau_{12}$ are the components of the stress tensor, X^T and X^C are the longitudinal tensile and compressive strengths, Y^T and Y^C are the tensile and compressive strengths in the transverse direction, and S^L and S^T are the longitudinal and transverse shear strengths respectively. The term α is a shear interaction coefficient, which can vary between 0 and 1. Hashin [32] suggests that the default value should be $\alpha = 1$.

3.1.5 Ultimate failure of large panels subject to compression

When large panels are loaded in compression two possible failure mechanisms come into action: compressive material failure and buckling. The two phenomena are closely interconnected - loss of stability results in out-of-plane displacements which will influence the stress distribution, while material failures will change the stiffness and influence the stability of the panel. As a result the response of a panel loaded in compression will be highly nonlinear around the failure point. The ultimate failure under these conditions can neither be characterized as purely material nor as a buckling failure. It may instead be characterized as buckling-driven material fracture: the first material failures will usually occur at the locations where significant bending stresses due to buckling-related out-of-plane displacement appear.

Hayman et al. [35] suggest that the failure in composite sandwich panels subjected to compression is modelled by employing a geometrically nonlinear progressive failure analysis in a finite-element model with applied initial imperfection. The progressive failure algorithm results in gradual reduction of stiffness which ultimately leads to the collapse of the panel. The initial imperfection is needed in order to initiate buckling in the model, which may otherwise fail to produce a buckling response. The imperfection may be included as part of the model geometry, or it may be induced by applying a small transverse force leading to out-of-plane displacement. The latter approach is relevant for wind turbine blade modelling as the outer surface of the blades is subject to aerodynamic pressure.

The approach described above is implemented using the commercial finite-element modelling program ANSYS. Figure 3.4 shows an example of the load-displacement history for a sandwich panel with glass-fiber reinforced faces. On the Figure the result from the nonlinear progressive failure analysis is compared to the strength predicted by linear eigenvalue analysis. The nonlinear model predicts higher strength, however the effect of buckling is also present, and the panel response becomes nonlinear when the load becomes close to the eigenvalue buckling limit. In many cases including the simulation shown on Figure 3.4 the ultimate failure will happen shortly after the first fiber breakage, because the failure of load-carrying fibers leads to stress redistribution and cascade of failures in other layers.

3.2 Probabilistic modelling of ultimate strength

To the reliability engineer, the presence of distinct failure modes means that composites will exhibit system reliability behavior. This system behavior will appear on several levels,

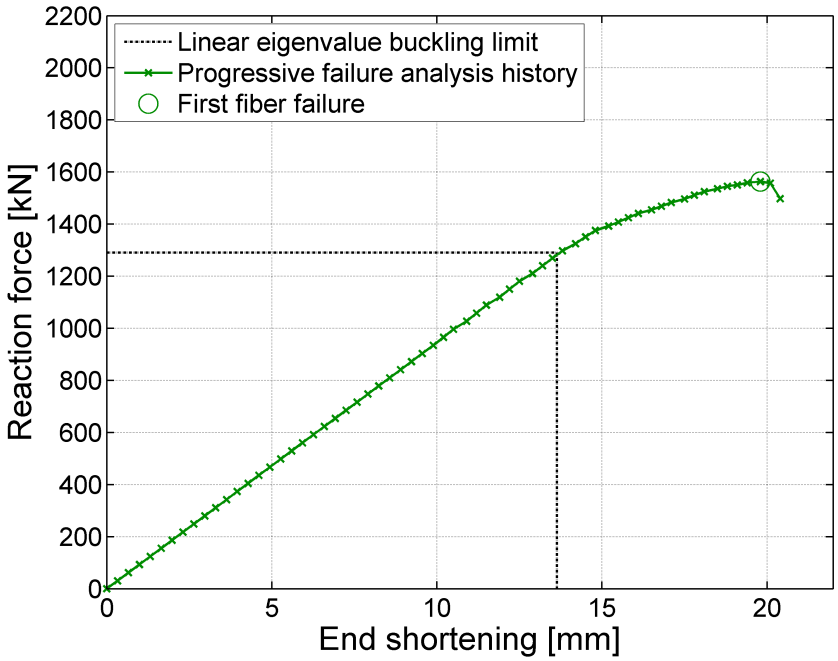

Figure 3.4: Failure of a composite sandwich panel subjected to compression.

corresponding to the length scales at which the respective failure modes occur - starting from micro-scale (fibers, matrix and interface between them), through lamina (individual layers with unidirectional fibers), structural components such as laminate panels (stacks of individual laminas) and sandwich panels, and up to whole structures such as an entire wind turbine blade.

3.2.1 Reliability system characterization

As discussed in the preceding sections, progressive failure of composite laminates is characterized by redistribution of loads following each of the successive failure events. Individual lamina will have random strength, as well as random stiffness distributions, meaning that the sequence of layer failures might be different for different realizations of the stochastic quantities. This was illustrated on Figure 3.2, where for two different realizations of the random lamina strength properties the sequence of failure events is different. Loading is not uniform either - a laminate can carry loads in a number of different directions, including both forces and bending moments, meaning that there will be a varying stress field inside the laminate.

A system with the features as described (varying stiffness, strength, loading, and load redistribution following failures) could in principal be modeled as a parallel system, as long as the overall failure event occurred when the last system component has failed. However, according to the definitions of laminate failure given above, ultimate failure occurs when the maximum load bearing capacity of the structure has been exceeded, or significant damage has occurred. These events might happen when there are still a large number of intact layers, with the failure event triggering a cascade failure of the remaining intact layers. This shows that a composite laminate can only be characterized as a general system, unless a first-ply failure approach is adopted, which would allow a series system representation.

Daniels, 1945 [11] (see Gollwitzer & Rackwitz, 1990 [29] for more modern discussion) has described the properties and derived reliability bounds for a similar type of system, consisting of equally loaded wires with random strength. The behavior of composite panels bears similarities with a Daniel's system, with failure being similarly characterized as the situation when the load exceeds the remaining bearing capacity of the structure. A difference with the classical Daniel's system is the fact that the individual components in the presently discussed problem have different stiffnesses, and are possibly loaded differently.

3.2.2 Solution of the system reliability problem

Due to the presence of multiple layers and multiple failure modes per lamina (in the example given below there are two modes per lamina) for a typical composite panel the total number of system components becomes fairly large. For example, a sandwich panel with two identical faces, each of them having 8 fiber-reinforced layers with 2 failure modes per layer, will have a total of 33 failure modes (if the core is considered to have a single failure mode). Analyzing a general system with 33 components in its entirety is difficult using most of the methods given above. Using Ditlevsen's method for reliability bounds of a general system (Ditlevsen & Madsen 1996 [19]) would require excessive computational efforts, because all possible safe sequences need be identified, making it necessary to investigate up to $n! \approx 8.6 \cdot 10^{36}$ different possible sequences. If it is possible to identify part of the system components as being most critical and thus governing the system probability of failure, these components can be arranged in an ordered set according to the failure probability of each component. Using order statistic, the correlated component events are transformed into independent events, and their joint probability distribution can be computed (Friis-Hansen, 1994 [25]). If only one stochastic variable per layer (i.e., the compressive fiber strength) is used, this computation is straightforward and fast. However, as discussed in Section 2.2.7, using more

than one variable per component makes it necessary to introduce two limit states, where one describes the component failure based on the input stochastic variables, and the other describes the system failure, based on the values of the component limit states. Some adaptive simulation techniques such as search-based importance sampling will also struggle with the problem due to the high number of dimensions, which makes it very difficult for the search algorithm to find the region in the failure domain with highest probability density. Given the above, three most suitable approaches to solving the present reliability problem are identified:

- Use a technique not suffering from problem dimensionality (e.g., Monte Carlo, Asymptotic Sampling)

- Simplify the problem by identifying possible critical components and focusing the analysis on those components. Consider each of the critical components as a cut-set, i.e., its failure results in an overall system failure

- Use FORM/SORM analysis to find a possible design point, and correct the reliability estimate for any nonlinearity in the failure surface by applying importance sampling at the design point. This analysis can serve as an indicator to the degree of system behaviour present in the structure - under milder conditions this technique should be able to determine the correct probability of failure, while if there are significant system effects present with multiple design points far from each other, the reliability estimate will not be correct.

Representing laminated faces as single equivalent layers

Using an equivalent single-layer representation of the faces in a sandwich panel will reduce the structural system to three components - upper face, lower face, and core. If the strength distributions of faces and core are available, it is straightforward to determine the system reliability by making a three-component series system analysis.

Although very useful in deterministic stiffness and strength analysis, it is possible that using the equivalent single-layer representation of the sandwich faces can result in some deviations in the reliability estimate, because not all system effects and load redistribution effects under a progressive failure process will be accounted for. The extent to which a homogeneous-layer representation of the faces can be used in reliability analysis is tested by comparing a reliability model based on homogeneous faces with the reliability estimates found from a full layer-wise model.

Table 3.1: Layup description of analyzed sandwich panel.

Component	Layer material	Orientation	Thickness (mm)
Lower Face	Glass-epoxy	$-45°$	0.3
	Glass-epoxy	$45°$	0.3
	Glass-epoxy	$0°$	0.3
	Glass-epoxy	$0°$	0.3
	Glass-epoxy	$0°$	0.3
	Glass-epoxy	$0°$	0.3
	Glass-epoxy	$45°$	0.3
	Glass-epoxy	$-45°$	0.3
Core	Balsa wood	$0°$	40
Upper Face	Glass-epoxy	$-45°$	0.3
	Glass-epoxy	$45°$	0.3
	Glass-epoxy	$0°$	0.3
	Glass-epoxy	$0°$	0.3
	Glass-epoxy	$0°$	0.3
	Glass-epoxy	$0°$	0.3
	Glass-epoxy	$45°$	0.3
	Glass-epoxy	$-45°$	0.3

3.2.3 Example system reliability analysis of a composite sandwich beam

In the following, the discussion continues with the help of an example reliability problem, where the reliability against ultimate failure of a sandwich composite beam is considered.

Description of structure

The structure under consideration is a simply supported, unit-width sandwich beam with faces made of glass-fiber reinforced fibers, and a balsa wood core. The two faces are identical laminates, consisting of eight layers each. Table 3.1 lists the thickness, angle orientation, and material for each of the layers within the structure.

The layup sequence of the faces, $\boldsymbol{\theta} = [-45^\circ \ +45^\circ \ 0^\circ \ 0^\circ \ 0^\circ \ 0^\circ \ +45^\circ \ -45^\circ]$, is simple, but in principle very similar to the layup sequences typically found on wind turbine blades,

where the vast majority of layers have orientation of 0 and ±45 degrees. The zero-degree layers have the highest stiffness, and therefore are typically arranged along the predominant load direction, which in the wind turbine blade is along the main span of the blade.

The sandwich beam is subject to a downward distributed load of $Qz = -7kN/m$, and a compressive axial force $Fx = -1400kN$. This resembles a loading condition which can be typically found on the downwind surfaces of wind turbine blades. Due to the distributed load, the sandwich beam will be loaded by an internal bending moment, highest at the middle of the beam, and a transverse shear force, which will be highest at the supports (and with opposite signs at opposite supports). Due to its larger magnitude the axial force is dominant and both faces will be loaded in compression. However, due to the presence of out-of-plane bending moment, the normal stresses will be larger in the upper face, where the bending moment contribution results in compression. In the lower face the contribution of the bending moment results in tension, thus counteracting the axial compressive force. In general, both faces will be loaded in compression, with higher loads the upper face, and the core will be loaded mainly in shear.

Modeling of stochastic material properties

In order to allow for random variations in the material properties within the laminate, each layer should be represented by a separate set of stochastic variables describing the material properties. For the 17-layer laminate this would mean at least 149 random variables (9 for each of the face layers, plus 5 for the core, where the transverse shear strength is the only strength property considered). In order to reduce the number of stochastic variables and to simplify the problem, it is decided to represent the elastic properties of materials with their mean values, with the argument that elastic properties have less variation and have in general less influence on the load bearing capacity of the structure compared to the strength properties. This results in a total of 81 stochastic variables representing the material properties.

Table 3.2 lists the assumed statistical distributions of the five strength parameters for composite lamina used in the present study: fiber tensile strength X_t, fiber compressive strength X_c, matrix tensile strength Y_t, matrix compressive strength Y_c, and shear strength S, as well as the shear strength of balsa core material, S_c. The mean values of properties of the face materials are determined from material test data of unidirectional lamina given in the OptiDat public database (Nijssen 2006 [56]). The definition of correlation between the material variables has to represent both the spatial variation of a given mechanical property,

Table 3.2: Statistical distributions of strength parameters.

Parameter	Designation	Mean	Cov	Distribution
Fiber tensile strength	X_t	$780MPa$	0.06	Lognormal
Fiber compressive strength	X_c	$528MPa$	0.19	Lognormal
Matrix tensile strength	Y_t	$54MPa$	0.08	Lognormal
Matrix compressive strength	Y_c	$165MPa$	0.14	Lognormal
Lamina shear strength	S	$82MPa$	0.15	Lognormal
Balsa core shear strength	Sc	$2.2MPa$	0.1	Lognormal

as well as correlation between different properties (e.g., tensile strength and shear strength). Thus the correlation coefficient between any two stochastic variables is taken as the product of two correlation coefficients,

$$\rho_{ij} = \rho_{Li,Lj} \cdot \rho_{Mi,Mj} \tag{3.14}$$

where

- i, j refers to the variable number
- $L(i), L(j)$ are the layer numbers which variables i and j refer to
- $M(i), M(j)$ are the two material properties which the variables i and j refer to, respectively.
- $\rho_{Li,Lj}$ is the correlation coefficient between the same property for different layers. It is taken as a fixed value, with $0 \leq \rho_{Li,Lj} \leq 1$ for $i \neq j$, and $\rho_{Li,Lj} = 1$ for $i = j$.
- $\rho_{Mi,Mj}$ is the correlation coefficient between the different material properties which variables i and j represent, regardless of layer number. The values of $\rho_{Mi,Mj}$ are taken from the study by Toft (2010) [82], based on micromechanics laws, and are listed in Table 3.3.

It is considered that the core strength properties are not correlated with face properties, and that properties in different faces are not correlated either.

The value of the correlation between the strength of different layers $\rho_{Li,Lj}$ is not known with certainty, as it is in principle very difficult to determine experimentally. It is nevertheless expected that some degree of correlation exists, because all material layers at a certain location within the structure have been subjected to similar conditions during the casting and curing process. In order to investigate the influence of inter-layer correlations, reliability

Table 3.3: Correlation coefficients between material strength variables.

	X_t	X_c	Y_t	Y_c	S
X_t	1	0.8	0	0.2	0.2
X_c	0.8	1	0	0.2	0.2
Y_t	0	0	1	0.8	0.8
Y_c	0.2	0.2	0.8	1	0.8
S	0.2	0.2	0.8	0.8	1

analyses are carried out with a number of different correlation levels, $\rho_{Li,Lj} = [0; 0.3; 0.7; 1]$. Correlation level of $\rho_{Li,Lj} = 1$ will mean that all layers in a single face will have the same strength properties.

Evaluation of ultimate capacity of the sandwich model

The stress distribution within the panel is calculated using the assumption that the core is loaded primarily in shear, with negligible normal stresses, while the faces carry the normal stresses with negligible transverse shear loading (see Zenkert, 1995 [92]). The layer-wise stress distribution is calculated using Classical Lamination Theory. In order to determine the failure load, a step-by-step analysis is performed, where at each step the layer which will fail under the smallest load is identified. The mechanical properties of the failing layer are downgraded, the structure is updated, and a new analysis step is performed on the remaining layers. The procedure is interrupted when a fiber-mode failure is observed, which will typically happen after a number of matrix-mode failures have already occurred, which is also visible on Figure 3.2.

In principle, if matrix degradation is disregarded, the first-fiber failure approach makes the structure equivalent to a series system. Including the influence of matrix failures means that there might be some path-dependency, where the matrix degradation influences the failure load and the location where the first fiber failure occurs. Therefore, it is considered that the laminate will behave as a general system even when using a first-fiber failure assumption. However due to the low stiffness of the matrix its influence will be limited, and a series system of components representing the fiber strength might be a good approximation.

If the analysis is continued until all layers have failed, the analysis step where the failure has occurred at the highest load will mark the ultimate load bearing capacity of the panel.

For the structure and loading conditions used for the present example, the ultimate bearing capacity will in many cases actually correspond exactly to the point of first fiber failure. This is because the load is carried mainly by just 8 layers (4 in each of the two faces) with 0-degree fiber orientation, and if one of them fails, in most cases the remaining 7 layers will not have enough remaining capacity to withstand the stress increase due to load redistribution. If there is more redundancy in the laminate, for example a thicker panel with, say, 10 layers in the 0° orientation in each face, then following a fiber failure event it is very possible that the laminate will still be able to carry the load, and as a result the first-fiber failure and the ultimate-load failure events will happen at different load values.

The load at which a lamina fails is identified using the previously discussed Hashin composite failure criterion (Hashin, 1980 [32]). This failure criterion assumes that a lamina can fail in two distinct modes: fiber failure or shear-dominated matrix failure. Both matrix and fiber failures can be tensile or compressive, depending on the stresses acting on the lamina.

System reliability

The reliability of the structure described above is estimated using the three different approaches suggested in section 3.2.2, as well as using a model where laminate faces are represented as single equivalent layers. A comparison of the performance of these four approaches is given in the following.

Crude Monte Carlo Given that enough failure events are observed, crude Monte Carlo (MC) simulations are probably the most robust method for determining reliability, and are therefore used as a reference value here. All simulations used in the present study are run until at least 400 failure events have been observed, which according to the formula by Shooman (1968) [73] corresponds to a 95% confidence that the error in the reliability estimate is less than 10%.

The results from the Monte Carlo simulations have identified that the layers in the face laminate oriented at 0 degrees are the most critical for the integrity of the structure (in most of the samples, indicating failure, the structure loses static equilibrium and collapses in a cascade of failures, following failure of one of the aforementioned 0-degree layers). In an attempt to simplify the problem and make use of more efficient reliability methods, this observation can be used as means to restrict the system analysis to the most critical parts of the structure.

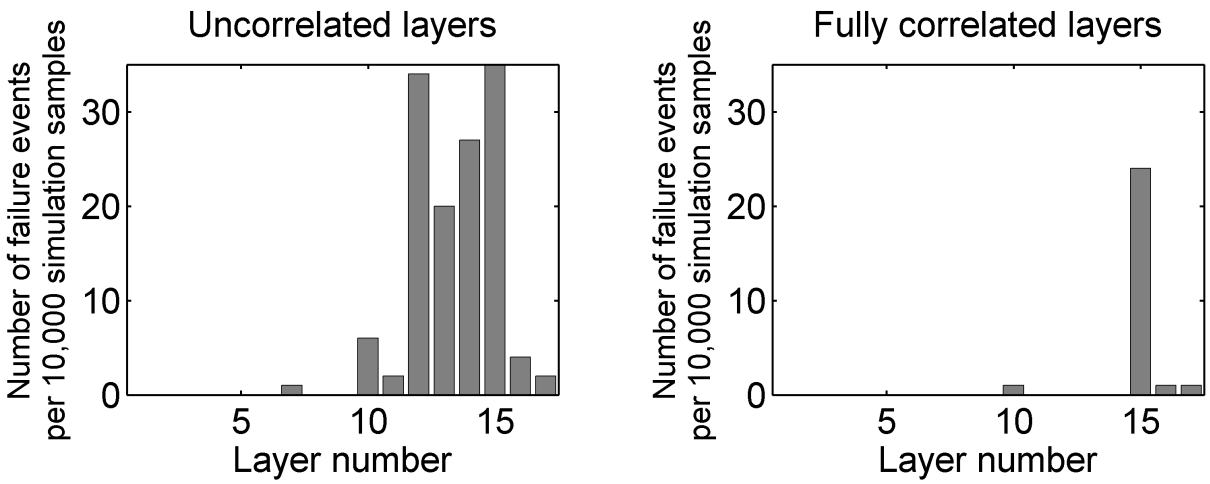

Figure 3.5: Distribution of failure events from Monte Carlo simulation: a) Layer strengths uncorrelated, b) Layer strengths fully correlated.

The effects of correlation between properties of different lamina are also clearly visible from the MC simulations. As Figures 3.5a) and 3.5b) illustrate, in the case of uncorrelated material strengths in different lamina, the ultimate failure event is located at various layers, mostly the four layers with 0-degree orientation of the upper face. In the case of fully correlated layer strengths within the face laminates, almost all failure events are concentrated at a single layer, which is the highest-loaded layer with 0° fiber angle. Such a behavior is expected, as when all the lamina in a given laminate have the same properties, the one failing first will simply be the one subject to the highest stress. Results from Monte Carlo simulations, along with results from other used methods, are shown on Figure 3.6.

Series system of critical components The results from the Monte Carlo simulations showed that some of the layers (the ones with 0-degree orientation) are the most critical to the integrity of the structure. As the failure definition adopted is first fiber breaking, the failure in the fiber material in any of these layers will be considered as an ultimate failure of the panel. Then, if the effect of matrix degradation is disregarded, each of the layers can be considered as a cut-set consisting of a single component representing the fiber strength. This approach makes the structure equivalent to a series system. The reliability of each of the cut-sets is equal to the component reliability, which, using single-component FORM analysis, is found to be $\boldsymbol{\beta}_{component} = [2.586, 2.574, 2.562, 2.550]$ for the four 0-degree layers loaded in compression. System failure probabilities can then be estimated as given by Equation 2.23,

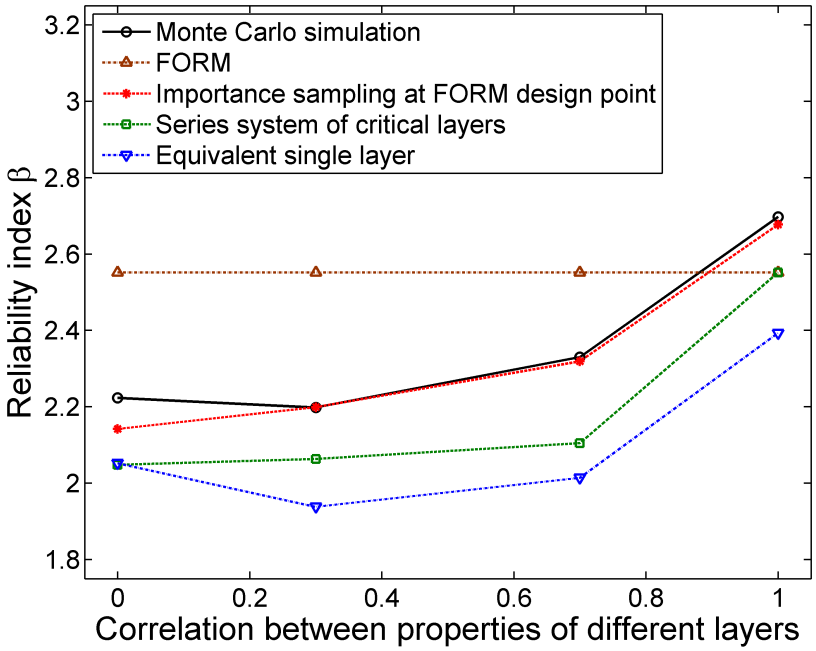

Figure 3.6: Results from system reliability analysis.

where $P_{f,system} = 1 - \Phi_k(\boldsymbol{\beta}_{component}, \mathbf{R}_{component})$.

Equivalent layer model The stochastic distributions of the equivalent strengths of the laminate faces used in the present problem are determined using simulation. For small failure probabilities, the tail of the strength distributions will have the greatest importance, and in order to ensure a good estimate some tail approximation technique has to be employed, as for example the ACER method by Naess (2009) [55]. For the present problem the probabilities of failure are relatively large, and a simple fit of a lognormal distribution to the data set from the MC simulation results gives excellent results too. The total probability of failure is then the complement of the product of the failure probability complements for the three components:

$$P_{f,system} = 1 - \prod_{i=1}^{N} (1 - P_i) \tag{3.15}$$

Comparison of results

System reliability estimates obtained by all the methods discussed above are shown on Figure 3.6. The presence of system behavior is evident when looking at the performance of FORM

method, which fails to predict the correct reliability index, and converges to the same design point, regardless of the degree of correlation between layers. This design point corresponds to failure of the 0-degree layer which is subjected to highest load (layer number 15 from Table 3.1). However, due to the randomness in material properties, failure modes corresponding to layers with lower stress are also present.

Importance sampling runs, centered at the design point obtained from FORM, converge to the correct probability of failure in most of the cases considered; however the convergence is very slow, because the sampling center points correspond to a single failure mode. Other failure modes occur relatively rarely, because their probability densities are far apart from the sampling center point. As a result, the importance sampling procedure described above has efficiency similar to that of a crude MC simulation, because the importance sampling density maximizes the chances of occurrence of just one of the several failure modes present.

The two simplified approaches for solving the system problem (using a smaller part of the system, or representing the faces as homogeneous) both capture the main trends in the system behavior, however the reliability estimates are conservative, sometimes with significant difference from the reference MC simulation results. These differences are explained by the facts that with such simplified representations the redistribution of stresses between the two sandwich faces following matrix-mode failures cannot be captured. The accuracy of the homogeneous layer representation is further limited by the underlying assumption that the stresses are constant throughout the laminate face. Although being relatively inaccurate, the single homogeneous layer calculation yields an approximate, conservative estimation of the system reliability, and can be useful in some reliability calculations where model complexity will not allow the use of full layerwise representation of the sandwich faces.

3.3 Probabilistic models for fatigue life of composites

Fatigue of materials is associated with a high degree of uncertainty due to random scatter in the material properties and defects. Therefore, in reliability analysis of structures with fatigue limit states, it is necessary to take the random variation of fatigue life properties into account. The probabilistic nature of fatigue in composites and the uncertainties associated with it have been part of the discussion in a number of scientific publications, as for example Passipoularidis, 2009 [62], Liu, 2005 [46], Naderi, 2011 [53].

The phenomenon of fatigue is usually modelled as a gradual progression of microscopic damage in the material. Most fatigue models, such as the classical $S - N$ curve and Paris

law [61], are phenomenological in nature and describe empirical relationships, which are determined based on observations. In the deterministic case, the most common empirical relationship is a curve with a slope describing the amount of incremental damage resulting from a load cycle of certain amplitude. A straightforward way to take the random nature of fatigue into account is to consider the parameters of the empirical fatigue law as random variables. The distribution of these random variables is determined by statistical analysis of test data. The present thesis adopts such an approach.

Laminated composites have a layered structure, where the properties of individual layers determine the global properties of the structure. The structural part of the proposed model is based on individual layer properties, which makes it possible to apply the model to a wide range of laminate configurations, as long as the layer properties are the same as in the data set used for assessing the model.

3.3.1 Fatigue life model using static material strength as parameter

Fatigue properties of structures are commonly described in terms of a damage accumulation law. A commonly used representation, chosen for the present problem, is the $S-N$ curve or W"ohler diagram (W"ohler, 1870 [91]), where the number of cycles to failure is given as a function of the amplitude of the stress cycle. This representation is convenient because it provides a good analogy to typical fatigue life tests, which commonly subject the material specimen to a constant-amplitude cyclic loading. In its typical form, the $S-N$ fatigue life model is described by the relation

$$NS^m = K \tag{3.16}$$

where N is the number of stress cycles of constant amplitude S to reach failure and K and m are material constants. Taking logarithm of both sides and rearranging terms, one obtains

$$\log N = \log K - m \log S \tag{3.17}$$

The above equation describes a linear relation between $\log N$ and $\log S$, with $\log K$ as the intercept and m as the magnitude of the negative slope. This relation has been found to well describe the fatigue life of materials over a wide range of stress amplitudes, except when the stress amplitude approaches the static strength of the specimen, or the fatigue threshold level (the endurance limit). An illustration of this is in Figure 3.7, which shows test data for a laminate from the public OptiDat database [56] together with the fitted model (solid line) with $\log(K) = 27$ and $m = 9.27$.

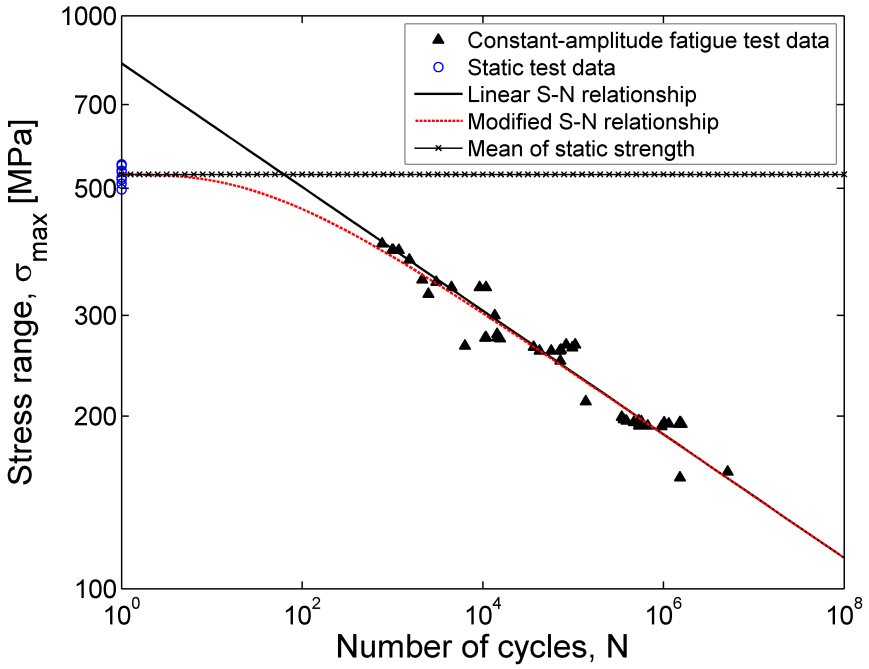

Figure 3.7: Linear and non-linear $S - N$ curves.

If equation (3.17) is solved for $\log N = 0$, i.e., for $N = 1$, the resulting stress estimate based on the fitted line in Figure 3.7 overestimates the static strength by about 50 percent. The classical $S - N$ curve fails to predict the static strength of the laminate, because it assumes a linear relationship between stress amplitude and damage accumulation per cycle. Under low-cycle, high-stress conditions, there is significant nonlinearity in the stress-strain relationship of the material due to plasticity effects. This results in nonlinear damage accumulation and leads to the observed discrepancy between the actual static strength and the static strength predicted by the classical $S - N$ relationship. It is clear that the relation between $\log N$ and $\log S$ should be nonlinear, at least for small values of N.

Physical phenomena that govern fatigue and ultimate capacities of a material have significant similarities. For example, it can be conjectured that variations in both the fatigue life and the ultimate capacity of a material are governed by the presence of flaws, such as microcracks, voids, and fiber misalignments. Therefore, one can expect a significant correlation between the static and fatigue properties. For a given lamina/laminate configuration, typically data on static properties are much more readily available than data on fatigue life due to the inherent difficulties associated with fatigue testing of composites, especially when determining the properties in shear. Therefore, it is beneficial to use the static properties

as an additional source of information for determining the fatigue life. Another argument for using the static strength as a variable in the fatigue law is that the original parameters (K and m) are only regression coefficients and have no direct physical meaning. Introducing the static strength as an additional variable allows to relate the model to the underlying physical phenomena.

One way to include the effect of the static strength in the fatigue law is to express the material constant K (see equations (3.16) and (3.17)) as a function of the static strength. This is achieved by evaluating equation (3.17) for $N = 1$, yielding

$$\log K = m \log S_1 \tag{3.18}$$

where S_1 denotes the stress amplitude for which the $S - N$ law predicts a fatigue life of one cycle. S_1 is expressed in terms of the static strength, denoted S_u, as

$$S_1 = S_u q \tag{3.19}$$

where q is a material constant. Using the above relation in equation (3.18), we have $K = S_u^m q^m$. By substituting this relation in equation (3.17) the modified $S - N$ law is obtained, where now three material quantities are present - m, q, and S_u,

$$N S^m = S_u^m q^m \tag{3.20}$$

The useful feature of this expression is that it introduces a direct relation between the static material strength and the fatigue properties. Such a relation is particularly useful when developing a probabilistic fatigue model. Using equation (3.20), some of the variation in the fatigue properties can be ascribed to the variation in the static strength. As already mentioned, there are far more data available for static properties of composites compared to their fatigue properties. If the static and fatigue properties of a material were fully correlated, the parameter q would be a deterministic constant. In a probabilistic analysis, q is considered an uncertain parameter with a statistical distribution, which accounts for the possible lack of perfect correlation between the static and fatigue properties.

Now a modification to the above model is proposed so that it accounts for the low-cycle fatigue region, where, as already seen in Figure 3.7, the original $S - N$ curve fails to correctly predict the fatigue life. The modification amounts to replacing the linear $S - N$ curve with a nonlinear curve, which intercepts the y-axis at $S = S_u$ with a zero slope, while remaining coincident with the original linear curve at $N >> 1$. Noting that $\frac{dS}{dN} = -\frac{1}{m}$ represents the slope of the logarithmic $S-N$ curve, the constant slope has to be replaced with

$\lim_{N\to 1}(\frac{dS}{dN}) = 0$ and $\lim_{N\to\infty}(\frac{dS}{dN}) = -\frac{1}{m}$. Furthermore, since we want $S = S_u$ at $N = 1$, we must have $\lim_{N\to 1}(S(N)) = S_u$ and $\lim_{N\to\infty}(S(N)) = S_u q N^{-\frac{1}{m}}$. These conditions can be satisfied by multiplying the coefficient $\log q$ in the logarithmic $S - N$ equation with a function of N that varies between 0 at $N = 1$ and 1 at $N \to \infty$:

$$\log S = \log S_u + f(N) \cdot \log q - \frac{1}{m} \log N \tag{3.21}$$

Consider a parameterized version of the logistic function

$$f(\log N) = 2 \cdot \left[\frac{1}{1 + e^{-t \log N}} - 0.5 \right] \tag{3.22}$$

The function $f(\log N)$ smoothly and monotonically varies between -1 at $N \to -\infty$ and 1 as $N \to \infty$, and equals 0 at $\log N = 0$. The parameter t is used to adjust the slope of the function at $\log N = 0$ which equals $f'(0) = 0.5t$. The value of t that satisfies the requirements for asymptotic behaviour given above, is

$$t = \frac{1}{0.5m \log q} \tag{3.23}$$

The end result is the modified $S - N$ curve

$$\log S = \log S_u + 2 \cdot \left[\frac{1}{1 + e^{-\frac{\log N}{0.5m \log q}}} - 0.5 \right] \cdot \log q - \frac{1}{m} \log N \tag{3.24}$$

Figure 3.7 shows the modified $S - N$ law (dashed line) compared with the classical $S - N$ law for the measured data. For this particular fitting, $S_u = 528MPa$, $q = 1.56$, $m = 9.27$ and $t = 1.1$.

For some materials, when the stress amplitude does not exceed a given threshold value (the so-called endurance limit), the amount of accumulated fatigue damage is negligible, and the number of cycles to failure will asymptotically approach infinity. The $S - N$ curve defined in equation 3.24 does not account for an eventual endurance limit, i.e., the only stress amplitude for which the number of cycles to failure equals infinity is at $S = 0$. However, it is possible to include an endurance limit in the suggested fatigue law, by imposing the condition $\lim_{N\to\infty}(S(N)) = S_{endurance}$, where $S_{endurance}$ is the endurance limit, on equation 3.24.

3.3.2 Description of fatigue load cycles

During a fatigue test, the specimen is subjected to repetitive load cycles. The cycles typically have a sinusoidal shape, which means that the properties of a given load cycle are entirely

defined by any two of the following four quantities: mean, amplitude, maximum, and minimum load. For constant-amplitude fatigue tests the standard load definition consists of the load amplitude (or the range, i.e., twice the amplitude), and the R-ratio, which is the ratio between the algebraic minimum and maximum load values:

$$R = \frac{\sigma_{min}}{\sigma_{max}} \tag{3.25}$$

The load amplitude is the variable which is most often used for describing the value of S in an $S - N$ curve. However, this representation is not very convenient for comparing the fatigue life with the static strength of the material, because the cyclic load amplitude and the load at static failure will only be consistent with each other for $R \to \infty$ and $R \to 0$ (i.e., $\sigma_{max} = 0$ or $\sigma_{min} = 0$). An approach which allows better comparison between static and fatigue properties is using the maximum load level in a cycle, σ_{max}, as load reference. In the present thesis, this approach is used in cases when it is necessary to consider the static strength, as for example in figures showing the nonlinear $S - N$ curve defined in Section 3.3.1, which includes the static strength as a parameter.

The definition of load by mean, maximum, amplitude, or R-ratio can not be used for the case of variable-amplitude (spectrum) loading either, because all the quantities mentioned will vary from cycle to cycle. Instead, the load is defined in terms of a spectrum and a cycle sequence (as for example the WISPER and WISPERX standardized sequences for testing of composites, Ten Have, 1992 [81], [80]). When constructing an $S - N$ curve based on variable-amplitude test data, the largest single load peak from the spectrum is used as a reference for the load magnitude, and the number of spectrum repetitions (passes) to failure are used instead of the number of cycles.

3.3.3 Fatigue failure criteria for laminates

Different orientations of the material coordinate systems of the individual lamina within an FRP composite laminate result in varying stress tensors from lamina to lamina. This means that a typical unidirectional loading of the laminate will result in multidirectional stresses within the individual lamina. Therefore, in order to assess the fatigue life of a laminate based on individual lamina properties, it is necessary to make assumptions regarding the manner in which individual stress components contribute to the fatigue damage within a lamina. One assumption often used in the literature (see e.g., Hashin [33], Favaz [23], Liu [46], Philippidis [63], Papanikos et.al [60]) is that the contributions of cyclic stress components to the fatigue damage are analogous to the contributions of static stress components to the static failure

stress. This assumption implies that failure criteria similar to those developed for the static strength of composites can be used to model the fatigue life. The failure criterion considered here is the fatigue analog to the static Hashin failure criterion ([32]). Two failure modes are considered - a fiber failure mode, and a matrix failure mode:

$$\begin{aligned} \left(\frac{\sigma_{11}}{\sigma_A(R,N)}\right)^2 + \left(\frac{\tau_{12}}{\tau_A(R,N)}\right)^2 = 1 \qquad (fiber\ mode) \\ \left(\frac{\sigma_{22}}{\sigma_T(R,N)}\right)^2 + \left(\frac{\tau_{12}}{\tau_A(R,N)}\right)^2 = 1 \qquad (matrix\ mode) \end{aligned} \tag{3.26}$$

In the above, σ_{11} is the cyclic stress amplitude in the primary material direction (fiber direction); σ_{22} is the cyclic stress amplitude in the transverse material direction (matrix direction); τ_{12} is the cyclic shear stress amplitude; $\sigma_A(R,N)$ is the failure stress amplitude in the primary material direction as a function of the number of cycles N and the ratio R between the (algebraic) minimum and maximum stresses in the stress cycles; $\sigma_T(R,N)$ is the failure stress amplitude in the transverse material direction as a function of N and R; and $\tau_A(R,N)$ is the failure stress amplitude in shear as a function of N and R. A lamina subjected to cyclic loading fails in the mode that corresponds to a lower fatigue life when solving equation 3.26 for N. In case of unidirectional loading, the dominant failure mode will depend on the orientation of the lamina with respect to the load direction: the fiber failure mode occurs for lamina orientations $0° \leq \theta \leq \theta_{cr}$, and the matrix failure mode occurs for lamina orientations $\theta_{cr} \leq \theta \leq 90°$, where θ denotes the angle between the fiber direction and the loading direction. At the transition angle θ_{cr} both failure modes should be satisfied, which defines the transition angle as:

$$\theta_{cr} = \tan^{-1}\sqrt{\sigma_T(R,N)/\sigma_A(R,N)} \tag{3.27}$$

3.3.4 Laminate fatigue life model based on lamina data and quasi-static failure criteria

Exploiting the analogy between static and fatigue failure criteria, a fatigue life algorithm is laid out in a way similar to progressive failure algorithms used for determining the ultimate static capacity of a laminated plate (see e.g., Jones [39]). However, several changes to the static algorithm are necessary, first because the model should give the number of cycles to failure rather than a critical failure stress, and second because the fatigue properties depend on the cyclic load ratio R. The resulting algorithm for estimating the fatigue life of a laminate based on lamina-level data is shown in figure 3.8.

If full correlation exists between the properties of individual lamina, the input to the fatigue model is the following set of lamina-level variables:

- In-plane elastic constants $[E_1,\ E_2,\ G_{12},\ \nu_{12}]$,
- The in-plane lamina strengths $[X_t,\ X_c,\ Y_t,\ Y_c,\ S_{12}]$. In the fatigue law these constants are given as $[S_{u,A},\ S_{u,T},\ S_{u,S}]$, where

$$\begin{aligned} S_{u,A} &= X_t \ \text{ for } \sigma_x \geq 0 \\ &= X_c \ \text{ for } \sigma_x < 0 \\ S_{u,T} &= Y_t \ \text{ for } \sigma_y \geq 0 \\ &= Y_c \ \text{ for } \sigma_y < 0 \\ S_{u,S} &= S_{12} \end{aligned} \tag{3.28}$$

where the subscripts A, T and S refer to axial, transverse and shear loading, while t, c and 12 refer to tension, compression and shear loading.

Other necessary inputs to the model are the lamina-level fatigue life parameters m, q and t. Their values depend on the load ratio R to which each lamina is subjected. A different load ratio activates a different failure mechanism, hence altering the fatigue behaviour of the material. For example, for a load ratio of $R = 0.1$ the material is always loaded in tension, whereas for $R = 10$ the material is loaded in compression. Clearly, tensile and compressive loadings of a composite lamina would lead to different failure mechanisms.

For a single load ratio, and assuming full correlation between individual lamina properties, the laminate fatigue model requires a total of 18 input quantities - 9 static material constants and 9 material fatigue parameters:

$$\begin{aligned} \text{Static constants}: \quad & E_1, E_2, G_{12}, \nu_{12}, X_t, X_c, Y_t, Y_c, S_{12} \\ \mathbf{q} = \quad & [q_A(R), q_T(R), q_S(R)] \\ \mathbf{m} = \quad & [m_A(R), m_T(R), m_S(R)] \\ \mathbf{t} = \quad & [t_A(R), t_T(R), t_S(R)] \end{aligned} \tag{3.29}$$

where the subscripts A, T and S again refer to material properties with respect to axial, transverse and shear loading. If the assumption of full correlation between laminas is not valid and individual lamina properties are randomly drawn from a population, the total number of inputs will be 18 multiplied by the number of layers.

The above fatigue model can be generalized to take into account multiple load ratios, if the available data are for a mixture of load ratios. For example, if the model is to take into account the three most-commonly used load ratios, $R = [-1; 0.1; 10]$, a total of nine values

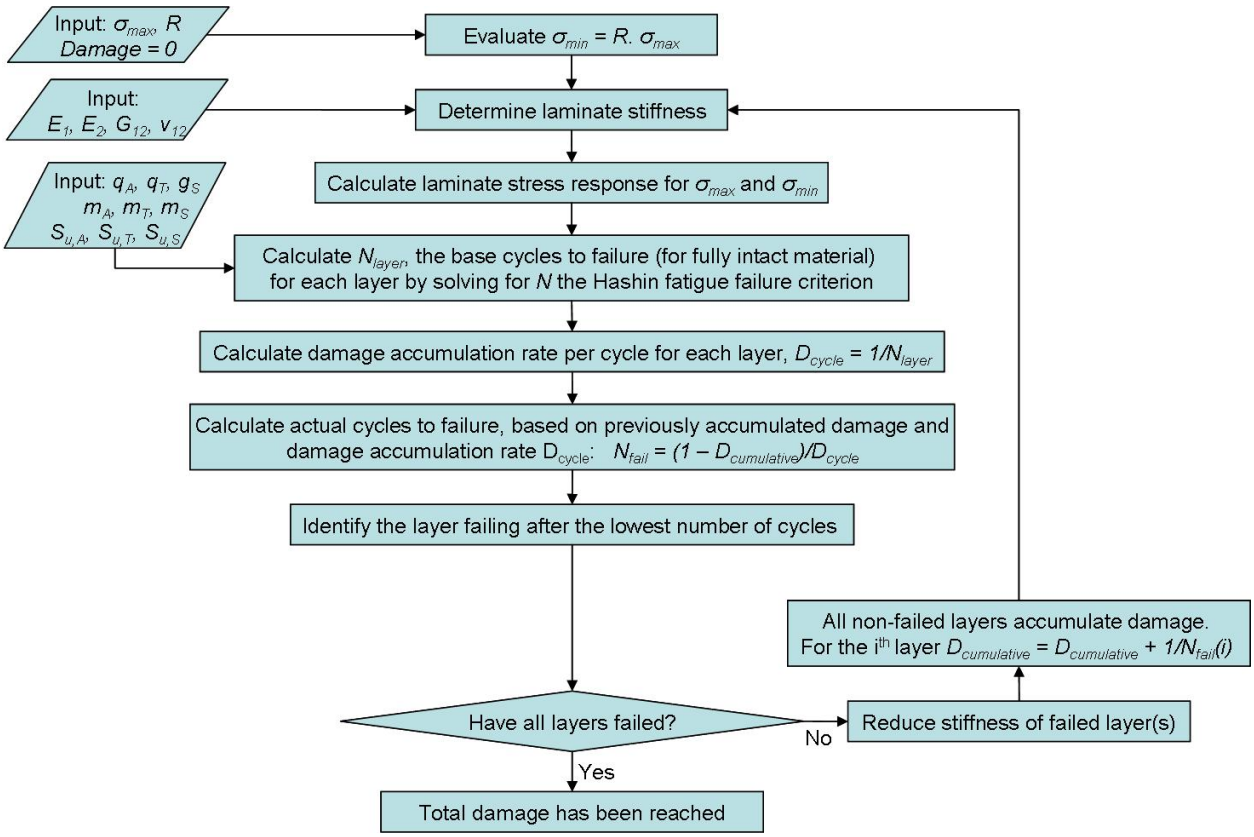

Figure 3.8: Algorithm for calculating the fatigue life of a laminate.

will be needed for each of the parameter q, m and t, i.e., for the three stress components and three load ratios:

$$\begin{aligned}
\mathbf{q} = &\ [q_A(-1), q_A(0.1), q_A(10), q_T(-1), q_T(0.1), q_T(10), q_S(-1), q_S(0.1), q_S(10)] \\
\mathbf{m} = &\ [m_A(-1), m_A(0.1), m_A(10), m_T(-1), m_T(0.1), m_T(10), m_S(-1), m_S(0.1), m_S(10)] \\
\mathbf{t} = &\ [t_A(-1), t_A(0.1), t_A(10), t_T(-1), t_T(0.1), t_T(10), t_S(-1), t_S(0.1), t_S(10)]
\end{aligned} \tag{3.30}$$

As can be seen, each additional load ratio adds nine parameters to the model. With this, the total number of material constants and model parameters required as input for three load ratios is 36 (9 material static constants + 27 fatigue parameters).

3.3.5 Fatigue damage accumulation law

The $S-N$ curves discussed in the previous section give a relation between the cyclic stress amplitude and the cycles to failure, however such a relation is only valid for constant-amplitude loading cycles. Varying-amplitude loads are usually described in terms of a load spectrum, where the number of cycles are a function of the stress amplitude, and in case of a 2-dimensional spectrum, a function of the stress amplitude and mean values. Therefore, a relation between the constant-amplitude-based $S-N$ curve and the damage caused by a load spectrum is necessary. The simplest and one of the most widely used damage accumulation laws is the Palmgren-Miner's rule ([58], [51]), which is based on the assumption of linear damage accumulation. According to Palmgren-Miner's rule, the following fundamental condition is satisfied at failure:

$$\sum_i \frac{n_i}{N(S_i)} = 1 \tag{3.31}$$

where i is the i^{th} stress amplitude bin from the load spectrum, S_i is the value of the stress amplitude in that bin, n_i is the number of cycles with this amplitude, and $N(S_i)$ is the number of cycles to failure for a constant-amplitude loading S_i, as estimated by the material $S-N$ curve.

Fatigue damage accumulation laws exist, aiming at representing the fatigue phenomenon more accurately than the linear damage accumulation hypothesis. However, due to the very large scatter of data typically observed in fatigue tests of composites, the gain of accuracy when using other models is not very significant (Nijssen, 2006 [57]), despite the added complexity. It is therefore decided that Palmgren-Miner's rule is the best choice in the present thesis, due to its simplicity. The consequences of using Palmgren-Miner's rule are further discussed in Chapter 5, where the accuracy in model predictions of fatigue life of composite

laminates subjected to variable amplitude loading is analyzed.

3.3.6 Bayesian inference model for fatigue life of laminated composites

Introduction to Bayesian inference

The input quantities of a mathematical model are either variables or parameters. Measurable or observable quantities, which with the help of mathematical expressions (i.e., the model) produce the desired mathematical output, are the variables, here denoted by the vector $\mathbf{x}$. The mathematical expressions within the model are normally formulated with the help of a set of parameters, here denoted by the vector $\boldsymbol{\theta}$. The parameters may or may not have physical meanings and are usually not observable. A predictive model with this distinction between observable variables and unobservable parameters is written as

$$y = g(\mathbf{x}, \boldsymbol{\theta}) \tag{3.32}$$

where y is the variable to be predicted and $g()$ is the model function. Given a set of observations of $\mathbf{x}$ and the corresponding y, the task is to estimate the model parameters $\boldsymbol{\theta}$. In Bayesian inference [6], this is accomplished by use of the Bayesian updating formula

$$f(\boldsymbol{\theta}) = \kappa \mathcal{L}(\boldsymbol{\theta}) p(\boldsymbol{\theta}) \tag{3.33}$$

where $f(\boldsymbol{\theta})$ is the posterior distribution representing the updated knowledge about the model parameters $\boldsymbol{\theta}$, $p(\boldsymbol{\theta})$ is the prior distribution representing the initial knowledge about$\boldsymbol{\theta}$, $\mathcal{L}(\boldsymbol{\theta}) \propto P(y|\mathbf{x}, \boldsymbol{\theta})$ is the likelihood, which is a function proportional to the conditional probability of observing the values of y for the given observations of $\mathbf{x}$ and current values of the parameters $\boldsymbol{\theta}$, and $\kappa = \left[\int \mathcal{L}(\boldsymbol{\theta}) p(\boldsymbol{\theta}) \mathrm{d}\boldsymbol{\theta}\right]^{-1}$ is a normalization constant. In order to fit a Bayesian model to a set of data, two functions must be specified: The likelihood function, $\mathcal{L}(\boldsymbol{\theta})$, and the prior distribution, $p(\boldsymbol{\theta})$. Formulations of these functions are presented below. But first a model must be selected.

Choice of the model

The mathematical model $g(\mathbf{x}, \boldsymbol{\theta})$ symbolically represents the algorithm shown in Figure 3.8. For a given stress cycle, the variables $\mathbf{x}$ represent the elastic and strength properties $(E_1, E_2, G_{12}, \nu_{12}, X_t, X_c, Y_t, Y_c, S_{12})$ and the parameters $\boldsymbol{\theta}$ represent the material constants $(\mathbf{q}, \mathbf{m}, \mathbf{t})$. Given $\mathbf{x}$ and $\boldsymbol{\theta}$, $g(\mathbf{x}, \boldsymbol{\theta})$ produces the cyclic stress amplitude (or its logarithm, see

below) that will result in failure after a given number of cycles. However, mathematical models of physical phenomena are invariably imperfect, and the proposed fatigue model is no exception. To account for this error, equation (3.32) is rewritten in the form

$$y = \hat{g}(\mathbf{x}, \boldsymbol{\theta}) + \epsilon_m \tag{3.34}$$

where the superposed hat on $\hat{g}$ indicates the approximate nature of the model and ϵ_m denotes the model error. With the objective of developing an unbiased model, i.e., a model that, on the average, predicts the correct mean of the output variable, the mean of ϵ_m is set to zero. Its standard deviation, σ_{ϵ_m}, is a measure of the model error and an additional parameter to be estimated.

To facilitate statistical inference, now two fundamental assumptions are made. Firstly, homoscedasticity is assumed, i.e., σ_{ϵ_m} is independent of $\mathbf{x}$. In other words, the model error does not depend on the set of input variables. Secondly, the distribution of ϵ_m and, therefore, that of y for a given $\mathbf{x}$ is assumed to be normal. To achieve these conditions, at least approximately, the variance-stabilizing logarithmic transformation is employed such that $x = \log N$ and $y = \log S$, where N is the number of cycles to failure and S is the cyclic stress magnitude. Observation of the scatter of fatigue data, such as that in Figure 3.7, supports the validity of this transformation.

Formulating the likelihood function

Given a set of observations $(\mathbf{x}_i, y_i)$, $i = 1, ..., r$, the task is to formulate the likelihood function as the conditional probability of making the above observations given the parameter values $\boldsymbol{\theta}$ and σ_{ϵ_m}. This task is further complicated by the inevitable presence of measurement errors in the observations. First the error in the measurement of y is considered. Let $\hat{y}_i$ represent the ith measured value so that the true value is $y_i = \hat{y}_i + \epsilon_{y,i}$, where $\epsilon_{y,i}$ denotes the measurement error. The values of $\epsilon_{y,i}$ are assumed to be independent outcomes of a normal random variable ϵ_y with zero mean (unbiased measurement error) and standard deviation σ_{ϵ_y}. For the ith observation, equation (3.34) can be rearranged to read $\epsilon_m - \epsilon_{y,i} = \hat{y}_i - \hat{g}(\mathbf{x}_i, \boldsymbol{\theta})$. It is clear that the right hand side is the outcome of a normal random variable with zero mean and variance

$$\sigma^2 = \sigma^2_{\epsilon_m} + \sigma^2_{\epsilon_y} \tag{3.35}$$

Thus, assuming the observations are statistically independent, the likelihood function takes the form of the product of normal probability densities with zero mean, standard deviation σ and outcomes $\hat{y}_i - \hat{g}(\mathbf{x}_i, \boldsymbol{\theta})$, i.e.,

$$\begin{aligned}\mathcal{L}(\boldsymbol{\theta}, \sigma^2) &\propto \prod_{i=1}^{r} \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(\hat{y}_i - g(\mathbf{x}_i, \boldsymbol{\theta}))^2}{2\sigma^2}\right] \\ &\propto \frac{1}{\sigma^r} \exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^{r} \left(\hat{y}_i - g(\mathbf{x}_i, \boldsymbol{\theta})\right)^2\right]\end{aligned} \tag{3.36}$$

Often the variance of the measurement error, $\sigma^2_{\epsilon_y}$, is known from calibration of the measuring devices or methods. In that case, the model variance $\sigma^2_{\epsilon_m}$ is computed from equation (3.35) after estimating σ^2. If the variance of the measurement error is not known, then the total variance σ^2 is all that can be estimated.

The likelihood function in equation 3.36 is general for any inexact model $\hat{g}(\mathbf{x}, \boldsymbol{\theta})$ that satisfies the homoscedasticity and normality assumptions, provided the observations $\mathbf{x}_i$ have no measurement error. This is often not the case, and so now the attention is turned to formulating the likelihood when the observations of the input variables contain measurement errors.

The set of variables for the laminate fatigue life model described in section 3.3.4, excluding precisely known variables such as the number of layers and load ratio, is

$$\mathbf{x} = (S, E_1, E_2, G_{12}, \nu_{12}, X_t, X_c, Y_t, Y_c, S_{12}) \tag{3.37}$$

Suppose the observed values $\hat{\mathbf{x}}_i$, $i = 1, \ldots, r$ of $\mathbf{x}$ have the random error vector $\boldsymbol{\epsilon_x}$ such that $\mathbf{x}_i = \hat{\mathbf{x}}_i + \boldsymbol{\epsilon_x}$. It is assumed that the elements of $\boldsymbol{\epsilon_x}$ have the joint normal distribution with zero means (unbiased measurements). To make this assumption valid, the elements of $\mathbf{x}$ may be appropriately transformed, e.g., $\log x$ may be used in place of x, if the latter has the lognormal distribution. In some cases there are no measured values of $\mathbf{x}$, but it is known that its observations $\mathbf{x}_i$ (or their transformed values) are drawn from a joint normal distribution. In that case $\hat{\mathbf{x}}_i$ is the mean of the distribution and $\boldsymbol{\epsilon_x}$ denotes the variability around the mean. If an observation x is precisely known, then the corresponding ϵ_x is identical to zero. For the ith observation, replacing y_i with $\hat{y}_i + \epsilon_{y,i}$ and $\mathbf{x}_i$ with $\hat{\mathbf{x}}_i + \boldsymbol{\epsilon_x}$ in equation (3.34) results in

$$\hat{y}_i = \hat{g}(\hat{\mathbf{x}}_i + \boldsymbol{\epsilon}_\mathbf{x}, \boldsymbol{\theta}) + \epsilon_m - \epsilon_{y,i} \tag{3.38}$$

For a general function $\hat{g}$, the distribution of the right hand side is difficult to obtain. Hence, assuming the variabilities in $\boldsymbol{\epsilon}_\mathbf{x}$ are small relative to $\hat{\mathbf{x}}_i$, a first-order approximation of the function $\hat{g}$ relative to $\boldsymbol{\epsilon}_\mathbf{x}$ around zero is employed. This gives

$$\hat{g}(\hat{\mathbf{x}}_i + \boldsymbol{\epsilon_x}, \boldsymbol{\theta}) \approx \hat{g}(\hat{\mathbf{x}}_i, \boldsymbol{\theta}) + \nabla_\mathbf{x} \hat{g}(\hat{\mathbf{x}}_i, \boldsymbol{\theta}) \boldsymbol{\epsilon}_\mathbf{x} \tag{3.39}$$

where $\nabla_{\mathbf{x}}\hat{g}(\hat{\mathbf{x}}_i, \boldsymbol{\theta})$ is the gradient row vector of $\hat{g}$ with respect to $\mathbf{x}$ computed at $\hat{\mathbf{x}}_i$. Substituting this approximation in equation (3.38), we have

$$\hat{y}_i = \hat{g}(\hat{\mathbf{x}}_i, \boldsymbol{\theta}) + \nabla_{\mathbf{x}}\hat{g}(\hat{\mathbf{x}}_i, \boldsymbol{\theta})\boldsymbol{\epsilon}_{\mathbf{x}} + \epsilon_m - \epsilon_{y,i} \tag{3.40}$$

The right-hand side of this expression is now a linear function of jointly normal random variables so that $\hat{y}_i - \hat{g}(\hat{\mathbf{x}}_i, \boldsymbol{\theta})$ is the outcome of a normal random variable with zero mean and variance

$$\sigma_i^2 = \nabla_{\mathbf{x}}\hat{g}(\hat{\mathbf{x}}_i, \boldsymbol{\theta})\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}_{\mathbf{x}}\boldsymbol{\epsilon}_{\mathbf{x}}}\nabla_{\mathbf{x}}\hat{g}(\hat{\mathbf{x}}_i, \boldsymbol{\theta})^T + \sigma^2 \tag{3.41}$$

where σ^2 is as given in equation 3.35. Thus, assuming statistically independent observations, the likelihood function takes the form

$$\begin{aligned} \mathcal{L}(\boldsymbol{\theta}, \sigma^2) &\propto \prod_{i=1}^{r} \tfrac{1}{\sigma_i\sqrt{2\pi}} \exp\left[-\tfrac{(\hat{y}_i - \hat{g}(\hat{\mathbf{x}}_i, \boldsymbol{\theta}))^2}{2\sigma_i^2}\right] \\ &\propto \exp\left[-\tfrac{1}{2}\sum_{i=1}^{r}\left(\tfrac{\hat{y}_i - \hat{g}(\hat{\mathbf{x}}_i, \boldsymbol{\theta}}{\sigma_i}\right)^2\right] \end{aligned} \tag{3.42}$$

Note that the σ^2 term is hidden in each of the terms σ_i in the right-hand side of equation (3.35).

Weighting of the likelihood function

Sometimes model fitting is done on a combination of several data sets. If the data sets contain different numbers of data points, the larger sets will have higher influence on the likelihood values. If equal importance should be given to each data set under consideration, weighting functions are applied to the likelihood expression as suggested by Cetin & Der Kiureghian [10]:

$$\mathcal{L}(\hat{\mathbf{y}}|\boldsymbol{\theta}) \propto \prod_{Set1} \phi(\hat{y}_i|\mu_y(\boldsymbol{\theta}), \sigma_y(\boldsymbol{\theta}))^{w_{Set1}} \times \prod_{Set2} \phi(\hat{y}_i|\mu_y(\boldsymbol{\theta}), \sigma_y(\boldsymbol{\theta}))^{w_{Set2}} \tag{3.43}$$

The weight functions for two data sets would be given by

$$w_{Set1} = 1 \qquad ; \qquad w_{Set2} = \tfrac{N_1}{N_2} \tag{3.44}$$

where N_1 and N_2 are the number of data points in sets 1 and 2, respectively.

Prior distribution of parameters

Choosing an appropriate prior distribution $p(\boldsymbol{\theta})$ is a delicate task, because this choice can influence the posterior distribution. The significance of this choice decreases with increasing data, but it is nevertheless possible that an inappropriate choice leads to distortion of

the posterior distribution. Therefore, when little or no prior information is available, it is appropriate to choose a non-informative prior [6], which minimally influences the posterior. In practice, approximately non-informative prior distributions may be selected according to the following rule:

$$\begin{aligned} P(\theta_i) &\propto \tfrac{1}{\theta_i}, \qquad \text{for } 0 \leq \theta_i < \infty \\ p(\theta_i) &\propto \text{const, for } -\infty < \theta_i < \infty \end{aligned} \tag{3.45}$$

In the present analysis, non-informative priors are used. Since all the parameters $\boldsymbol{\theta} = (m_A, m_T, m_S, q_A, q_T, q_S, t_A, t_T, t_S, \sigma_{\epsilon_m})$ are by definition non-negative, the prior distribution in equation (3.45) is appropriate. Thus, assuming the parameters are a priori statistically independent, the prior distribution takes the form $p(\boldsymbol{\theta}) \propto \frac{1}{\boldsymbol{\theta}}$.

Estimating posterior distribution of model parameters

For complex models, such as the fatigue model under consideration, analytical solution of the posterior distribution of the parameters is infeasible. Hence, a numerical method to approximately evaluate the posterior distribution and its second moments is employed. For this purpose, an importance sampling method suggested by Gardoni and Der Kiureghian [27] is used.

Approximating posterior statistics by importance sampling Posterior analysis with Bayes' formula in equation 3.33 requires computing integrals of the form

$$I = \int B(\boldsymbol{\theta}) \mathrm{d}\boldsymbol{\theta} \tag{3.46}$$

where $B(\boldsymbol{\theta}) = \boldsymbol{\omega}(\boldsymbol{\theta})\mathcal{L}(\boldsymbol{\theta})p(\boldsymbol{\theta})$ is the Bayesian integrand. (For the sake of brevity of the notation, $\boldsymbol{\theta}$ here is meant to include the unknown parameter σ^2). Depending on the choice of the vector quantity $\boldsymbol{\omega}(\boldsymbol{\theta})$, the integral I represents different measures of the posterior density. Selecting $\boldsymbol{\omega}(\boldsymbol{\theta}) \equiv 1$ results in $I = 1/\kappa$, the inverse of the Bayesian normalization constant κ. Using this normalization factor in equation 3.33, the posterior joint density is obtained. If $\boldsymbol{\omega}(\boldsymbol{\theta}) = \kappa\boldsymbol{\theta}$ is selected, the integral yields the posterior mean vector of the parameters, $\boldsymbol{M_\theta}$. And if $\boldsymbol{\omega}(\boldsymbol{\theta}) = \kappa\boldsymbol{\theta}\boldsymbol{\theta}^T$ is selected, the result is $I = E[\boldsymbol{\theta}\boldsymbol{\theta}^T]$, the posterior mean square matrix, from which the posterior covariance matrix is computed as $\boldsymbol{\Sigma_{\theta\theta}} = E[\boldsymbol{\theta}\boldsymbol{\theta}^T] - \boldsymbol{M_\theta}\boldsymbol{M_\theta}^T$. For any selection of $\boldsymbol{\omega}(\boldsymbol{\theta})$, the integral(s) in equation 3.46 is evaluated by importance sampling (see Ditlevsen & Madsen, [19]). Defining an importance sampling density $S(\boldsymbol{\theta})$, the Bayesian integral takes the form

$$I = \int \left[\frac{B(\boldsymbol{\theta})}{S(\boldsymbol{\theta})}\right] S(\boldsymbol{\theta}) \mathrm{d}\boldsymbol{\theta} \tag{3.47}$$

It is seen that the value of the integral is equivalent to the expectation of the ratio $B(\boldsymbol{\theta})/S(\boldsymbol{\theta})$ relative to the probability density $S(\boldsymbol{\theta})$. Thus, for a sample of N random realisations drawn from the sampling density $S(\boldsymbol{\theta})$, an estimate of the integral is

$$\bar{I} = \frac{1}{N} \sum_{i=1}^{N} \frac{B(\boldsymbol{\theta}_i)}{S(\boldsymbol{\theta}_i)} \tag{3.48}$$

Choice of the importance sampling density The speed of convergence and the quality of the estimate in the Bayesian importance sampling procedure is tightly connected to the choice of the sampling distribution. An improper choice can lead to a large fraction of the samples yielding a zero or extremely small likelihood, thus not adding information to the analysis and reducing the speed of convergence.

One possible approach towards defining a suitable sampling density is to localize the mode of the posterior distribution. Under certain mild conditions (Richards, 1961 [67]), the difference between the value of $\boldsymbol{\theta}$ that maximizes the likelihood, i.e., the maximum-likelihood estimate (MLE), and the posterior mean $\boldsymbol{M}_{\boldsymbol{\theta}}$ asymptotically vanishes as the number of observations grows. Furthermore, the negative of the inverse of the Hessian, taken at the same point, is an estimate of the posterior covariance: $\boldsymbol{\Sigma}_{\boldsymbol{\theta}\boldsymbol{\theta}} \approx -\left[\nabla\nabla \ln\left(\mathcal{L}(\boldsymbol{\theta})\right)\right]^{-1}$. These values can be taken as first estimates of the posterior mean and covariance. However, finding the value of $\boldsymbol{\theta}$ that maximizes the likelihood is a challenging task for the present problem. The function describing the laminate fatigue life does not have explicitly formulated derivatives, is based on integer inputs of the number of cycles, and is not necessarily well-behaved. This makes gradient-based optimisation algorithms potentially unreliable and time consuming. A modification that makes using an optimisation procedure possible is allowing non-integer values of the cycles to failure, which makes the model function continuous and differentiable.

The above procedure describes the process of determining the mean and covariance of the initial sampling distribution. A choice of the distribution is also needed. A suitable choice for representing the non-negative quantities of the type contained in $\boldsymbol{\theta}$ is the joint lognormal distribution. However, it is more convenient to work with the normal distribution. Hence, in implementation of the algorithm, the model parameters $\boldsymbol{\theta}$ are represented by their natural logarithms, for which the joint normal distribution is appropriate. Since the logarithmically transformed parameters are unbounded, the second form of the prior distribution in equation 3.45 is used. Once a suitable initial sampling density has been identified, a preliminary importance sampling run is carried out and the posterior mean vector and covariance matrix are estimated. These are then used to define a more suitable sampling distribution.

This iterative adaptive process is continued a few cycles until convergence in the estimated posterior mean vector and covariance matrix is achieved.

Model reduction

Including many parameters in a model might result in a loss of precision in the parameter estimates, and possibly in over-fitting. Therefore, it is desirable to include as few parameters as practicable without losing accuracy and generality of the model. Once the posterior means and variance/covariances of the model parameters are estimated, parameters that do not add useful information to the model due to their high variances may be left out of the model, and parameters with strong correlations can be replaced by linear functions of their correlated pairs. For example, Gardoni & Der Kiureghian [26] suggest that when a pair of parameters θ_i and θ_j has correlation coefficient $|\rho_{\theta_i\theta_j}| \geq 0.7$, one of the parameters, say θ_i, be replaced by its best linear estimator,

$$\hat{\theta}_i = \mu_{\theta_i} + \rho_{\theta_i\theta_j} \frac{\sigma_{\theta_i}}{\sigma_{\theta_j}} \left(\theta_j - \mu_{\theta_j}\right) \tag{3.49}$$

The same approach is adopted in the present study with the threshold of the correlation coefficient set at 0.8.

3.3.7 Example with application to OptiDat database

In this section the Bayesian parameter estimation method is applied to fit a laminate fatigue life model to a set of measured fatigue strength data. Four different data sets are represented, which include three laminates made of the same glass-epoxy material, but with different layup sequences and subjected to different load ratios. One of the laminates is subjected to load ratios of $R = 0.1$ and $R = 10$, while the other two laminates are only subjected to load ratio $R = 0.1$. The particular data sets are chosen to have different types of mechanical behaviour, which will allow to infer more parameters from the list given in equation 3.30.

Input description

The test data used to fit the fatigue life model are taken from the open-access OptiDat database ([56]). Tests on three laminates are used - the “UD3”, “MD2” and “MD3” laminates, as described in Table 3.4. The UD3 laminate which consists of 7 identically-oriented layers is loaded transversely to the fiber orientation, therefore the layer orientation is given as 90°. All laminates are constructed of the same type of glass-epoxy UD lamina, thus the

Table 3.4: Laminates from OptiDat database.

Laminate type	Number of layers	Thickness	Layup	Number of test data
UD3	7	$6.30mm$	$[90]_7$	22
MD3	10	$3.05mm$	$[\pm 45]_5$	17
MD2	14	$6.57mm$	$[[\pm 45\ 0]_4\ \pm 45]$	28 ($R = 0.1$) 30 ($R = 10$)

individual layers have similar mechanical properties. In the multi-directional laminates MD2 and MD3 the average thickness of the 0-degree layers is $1.63mm$, several times more than the thickness of the layers with $+45°$ or $-45°$ orientation, which is $0.305mm$. These are average thicknesses, since the actual dimensions of the test specimens vary randomly as a consequence of variations in the production process. In order to eliminate the specimen dimensions as source of uncertainty, each specimen is precisely measured before testing and the applied loads are divided by the specimen actual area, resulting in the loads being expressed as stresses.

As discussed in Section 3.3.6, the Bayesian inference model takes into account the possible pertubations in some of the input variables. In the present example, it is considered that these pertubations are caused by the random natural variations in the material properties. The statistics of the mechanical properties are determined based on data from static tests from the OptiDat database, and are listed in Table 3.5. Table 3.6 gives the correlations between the variables. The correlation coefficients are determined based on laws of micromechanics, and are taken from a study by Toft [82].

Model fitting

The laminate fatigue life model defined in section 3.3.4 is used as the core algorithm for calculating fatigue life of the examples given above. The non-linear $S - N$ curve given by (3.24) is used, although in the present case there is no practical importance whether a linear or nonlinear $S - N$ curve is assumed because the available test data are limited to the high-cycle fatigue region, and they do not provide enough information for the material behaviour in the non-linear part of the $S - N$ curve. The model parameters and their probability distributions are determined using Bayesian inference, as described in Section 3.3.6.

A look at the four test data sets (Figure 3.9 shows that there are differences between

Table 3.5: Model variables with random pertubations.

Name	Desc.	Unit	Mean	C.o.v	Based on
E_1	Axial modulus	Pa	$39 \cdot 10^9$	0.02	UD, 0°
E_2	Transverse modulus	Pa	$14 \cdot 10^9$	0.02	UD, 90°
ν_{12}	Poisson ratio	-	0.29	0.09	UD, 0°
G_{12}	Shear modulus	Pa	$4.2 \cdot 10^9$	0.03	MD3, ±45°
X_t	Axial strength, tension	Pa	$777 \cdot 10^6$	0.05	UD, 0°
X_c	Axial strength, compression	Pa	$522 \cdot 10^6$	0.03	UD, 0°
Y_t	Transverse strength, tension	Pa	$54 \cdot 10^6$	0.05	UD, 90°
Y_c	Transverse strength, compression	Pa	$165 \cdot 10^6$	0.03	UD, 0°
S_{12}	Shear strength	Pa	$82 \cdot 10^6$	0.02	MD3, ±45°

Table 3.6: Correlations between static material properties.

	E_1	E_2	ν_{12}	G_{12}	X_t	X_c	Y_t	Y_c	S_{12}
E_1	1	0.8	0.0	0.8	0.6	0.6	0.0	0.2	0.2
E_2		1	0.0	0.8	0.6	0.6	0.0	0.2	0.2
ν_{12}			1	0.0	0.0	0.0	0.0	0.0	0.0
G_{12}				1	0.6	0.6	0.0	0.2	0.2
X_t					1	0.8	0.0	0.2	0.2
X_c						1	0.0	0.2	0.2
Y_t							1	0.8	0.8
Y_c								1	0.8
S_{12}									1

Table 3.7: Variables and parameters included in the laminate fatigue model examples.

Quantity	**UD3**	**MD3**	**MD2**	
	R0.1	**R0.1**	**R0.1**	**R10**
Stress, S	•	•	•	•
Number of layers, N	•	•	•	•
Layup angles, $\theta_1, \theta_2, \ldots, \theta_N$	•	•	•	•
Layer thicknesses, $t_1, t_2, \ldots, t_N$	•	•	•	•
UD lamina elastic constants, $E_1, E_2, G_{12}, \nu_{12}$	•	•	•	•
Fiber tensile strength, X_t	-	•	•	•
Fiber compressive strength, X_c	-	-	-	•
Matrix tensile strength, Y_t	•	•	•	-
Matrix compressive strength, Y_c	-	-	-	•
Shear strength, S_{12}	-	•	•	•
Lamina axial fatigue constants, m_A, q_A, for $R = 0.1$	-	-	△	-
Lamina axial fatigue constants, m_A, q_A, for $R = 10$	-	-	-	△
Lamina transverse fatigue constants, m_T, q_T, for $R = 0.1$	△	•	-	-
Lamina fatigue constants in shear, m_S, q_S, for $R = 0.1$	-	△	-	-
Model error	△	△	△	△
• : quantity is used as variable in the model △ : quantity is used as parameter in the model - : quantity is not used in the model				

them. The test data for laminate MD3 are characterized with much lower variance compared to the MD2 and UD2 laminates. The observed difference can possibly be attributed to the fact that the specimens in the data sets with larger scatter are tested at various labs and come from up to 15 different composite plates, while the more coherent data set comes from a single lab, and from specimens originating from just two composite plates. Another possibility that cannot be ruled out is errors in the measurements or the test set-up. Due to the differences in sample data sets, at first the data sets are considered separately and a Bayesian inference model is fit to each of them independently. In this way a more clear understanding of the data is obtained and the variance in each of the samples can be properly assessed.

The different load ratios used in different data sets mean that the material is subjected to either purely tensile ($R = 0.1$) or purely compressive ($R = 10$) loads. Therefore, not all of the variables will be active in each of the problems, where variables describing the compressive strength are irrelevant under purely tensile loading, and vice versa. Similarly, not all model parameters will be active in all models, because the strength of the MD2 laminate is dominated by the 0-degree layers and hence the axial lamina properties, while for the MD3 and UD2 laminates the shear and transverse properties are dominant, respectively. A list of all the active variables and parameters is given in Table 3.7. One should note that most of the variables (except the stress level) vary only from laminate type to laminate type, not from specimen to specimen, as one laminate type will always have the same number and orientation of layers. For the present problem the possibility that there are random variations in the angle orientations and thicknesses of individual layers is not taken into account, and these are considered deterministic.

A first estimate of model parameter values is obtained by maximizing the likelihood function and taking the inverse Hessian as described in Section 3.3.6. Then a series of importance sampling (IS) runs are carried out. The first IS run is centered at the likelihood mode, and subsequent runs are using the updated sampling density from the preceding IS run. 5000 samples are used in each pass, and the calculations are carried out until convergence. The convergence criterion is fulfilled when the estimates of the mean values do not change with more than 2% between two IS runs. The maximum likelihood for the MD3 laminate is not unique, because for this particular layup both the transverse and the shear properties of individual lamina have influence on the fatigue life of the laminate. As a result, there is an infinite number of combinations between pairs of properties (q_S, m_S) and (q_T, m_T) that result in the same fatigue life estimate. In order to solve that, first the model for the UD3 laminate is calibrated where shear stresses are zero, and the parameters q_T and m_T can be estimated. Subsequently q_T and m_T are included in the model for the MD3 laminate as variables with the distributions obtained from the UD3 model calibration.

A comparison between model predictions and measurement data for all four data sets is shown on Figure 3.9. The 95% confidence intervals on the figure are estimated by adding and subtracting 1.96 times the standard deviation of the total model error σ. Calibrated parameter values are given in Table 3.8.

The model error estimated for Model 2 (laminate MD3) is practically zero. This result is in agreement with what has been discussed previously: the data set for laminate MD3 is characterized with unusually low dispersion. Practically all the variance in the data is

Table 3.8: Parameter values: 1) initial sampling density from likelihood maximization, 2) results from first importance sampling run, 3) Converged results.

Parameter name	Initial values		First IS run		Final IS run	
	Mean	*c.o.v*	Mean	*c.o.v*	Mean	*c.o.v*
Model 1: Laminate UD3 90°, $R = 0.1$						
$q_{T,R=0.1}$	1.531	*0.0863*	1.499	*0.0922*	1.502	*0.0983*
$m_{T,R=0.1}$	10.70	*0.0888*	11.08	*0.0932*	11.08	*0.0992*
$\sigma_{\varepsilon,MODEL1}$	0.0426	*0.1687*	0.0471	*0.162*	0.0476	*0.168*
Model 2: Laminate MD3, $R = 0.1$						
$q_{S,R=0.1}$	1.235	*0.0339*	1.226	*0.0320*	1.235	*0.0295*
$m_{S,R=0.1}$	11.43	*0.0360*	11.54	*0.0347*	11.45	*0.0313*
$\sigma_{\varepsilon,MODEL2}$	≈ 0	-	≈ 0	-	≈ 0	-
Model 3: Laminate MD2, $R = 0.1$						
$q_{A,R=0.1}$	1.380	*0.036*	1.373	*0.0443*	1.365	*0.0476*
$m_{A,R=0.1}$	10.48	*0.0354*	10.54	*0.0447*	10.61	*0.0431*
$\sigma_{\varepsilon,MODEL3}$	0.0147	*0.255*	0.0159	*0.288*	0.0158	*0.297*
Model 4: Laminate MD2, $R = 10$						
$q_{A,R=10}$	1.145	*0.0400*	1.131	*0.0411*	1.126	*0.0473*
$m_{A,R=10}$	35.45	*0.124*	37.91	*0.128*	38.59	*0.142*
$\sigma_{\varepsilon,MODEL4}$	0.0153	*0.237*	0.0160	*0.266*	0.0163	*0.258*

explained by the error in input variables, in this case mainly the shear strength S_{12} and the matrix tensile strength Y_t.

A common outcome from the calibration of all models is that the pair of parameters q and m (q_A vs. m_A, q_S vs. m_S, q_T vs. m_T) have very high degree of correlation, with correlation coefficients $-1 < \rho \leq -0.96$. As a result, one of the parameters (q) can be removed from the model by approximating its value as a function of the other parameter in the correlated pair (see Section 3.3.6).

Finally, a general model is fit by using all available data sets simultaneously. The only parameter common to all models from single data sets is the standard deviation of the model error, σ_ϵ, and the value of this parameter is the only difference between the single-data-set models and the general model. The resulting general model error standard deviation

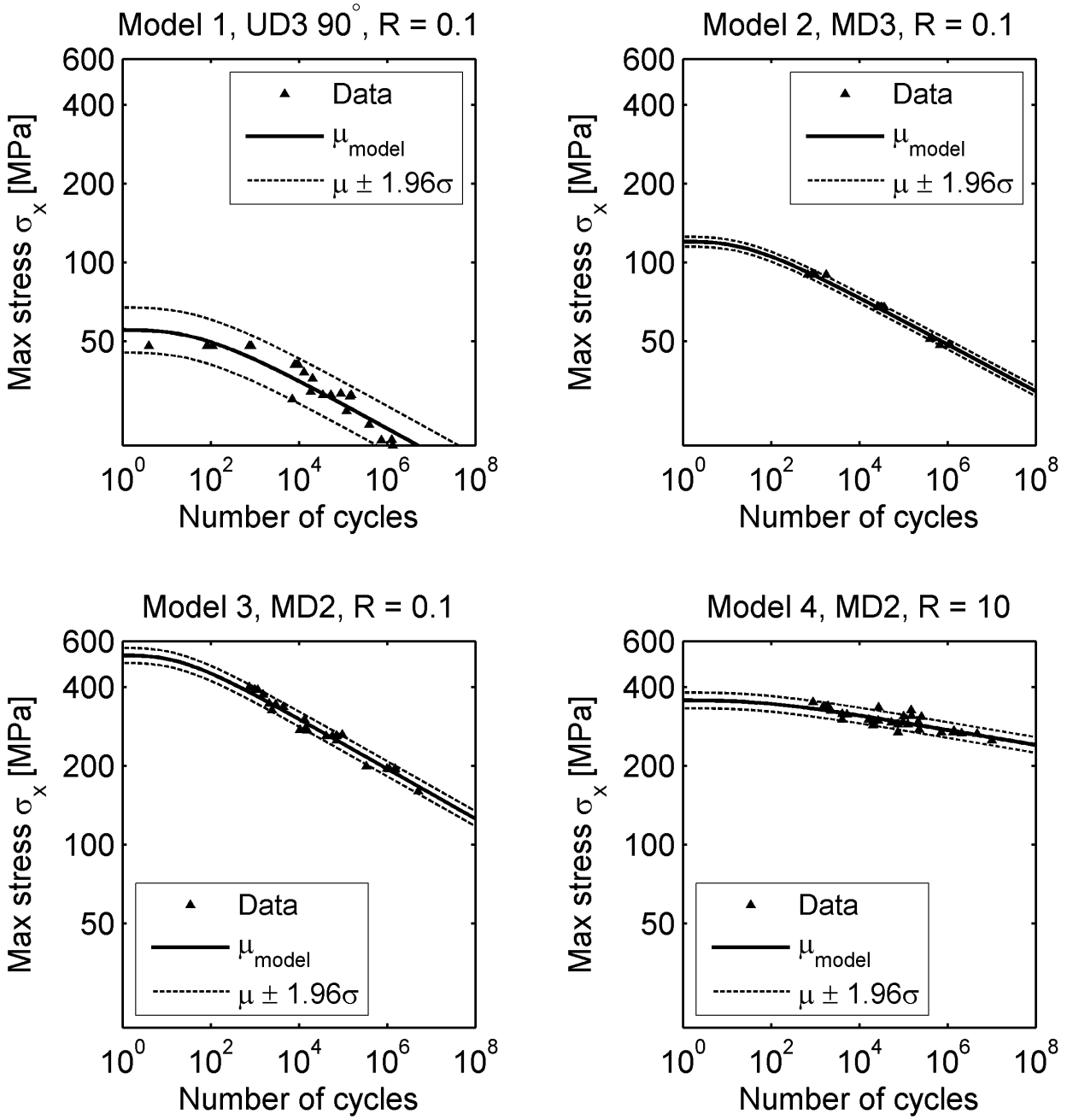

Figure 3.9: Fatigue life models calibrated to individual data sets.

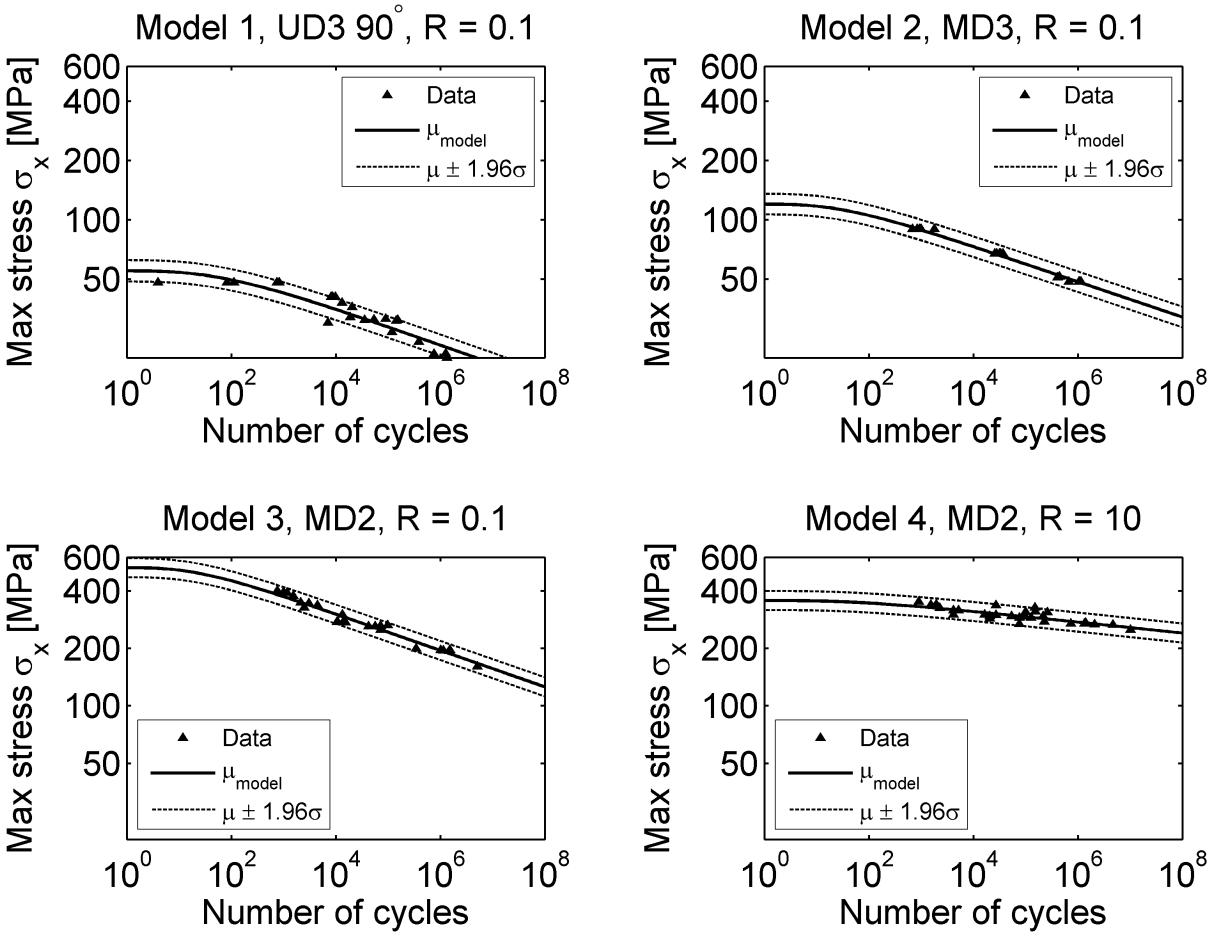

Figure 3.10: General fatigue life model calibrated to four data sets simultaneously.

is $\sigma_\epsilon = 0.025$. This is smaller than the model error for Model 1 (UD3 laminate), but larger than the errors for all remaining models. This illustrates the trade-off between model generality and model accuracy. The "general" model predicts the fatigue life of a wider range of laminates, but this comes at the expense of not being as accurate in representing the variance. This can also be seen on Figure 3.10, where the general model is compared to test data.

Discussion on probabilistic fatigue life modelling of laminates

When choosing a model for estimating the fatigue capacity of a laminated composite panel, a compromise has to be done between the generality of the model, and its complexity and accuracy. For a particular laminate, the simplest fatigue model would be to directly fit an $S - N$ relationship, valid for the laminate as a whole, by using data from tests of the

entire laminate. An example implementation for this type of model can be found in Chapter 5 of Madsen, Krenk & Lind [48]. Such a model will be accurate and simple to develop, because it is relatively easy to do fatigue tests on entire laminates. However, this model will only be valid for the particular laminate, for which it has been developed. Changing the layup (number of layers, layer orientations and thicknesses) or the material properties will make this model invalid. On the other hand, the model that includes numerous lamina-level parameters such as elastic constants, static strengths, and lamina fatigue parameters, will be far more general, but it will possibly produce larger deviations, and will be harder to develop due to the inherent difficulties in testing some of the lamina properties, especially the ones in shear.

Despite the greater difficulties in building up a lamina-level model, developing such a model is justified by the advantages it gives to its user:

- The model is applicable to more than one laminate with different layup sequences
- It can be used as a forecasting model, i.e., the fatigue strength of a non-tested or non-existing laminate can be estimated based on data from other tests, thus giving opportunity for parametric studies
- This lamina-level model takes into account the material stiffness and static strength properties and the uncertainties associated with them. This gives a physical anchor point to the otherwise fully parametric description of fatigue life laws.

3.3.8 Fatigue life of core material

Introduction

Balsa wood used as core material in sandwich composites is highly appraised for its qualities such as relatively high stiffness-to-weight and strength-to-weight ratio. The precise material properties are however often unknown, or just a single reference value is used, which gives no information about the uncertainty in the properties. This is understandable because balsa is the product of a naturally occurring growth process, which cannot be monitored and controlled to the same extent as industrial manufacturing processes. Balsa core sheets are usually assembled out of blocks a few cm in size which are placed randomly within the sheet. The properties of balsa wood are to a great extent determined by its density (Soden & McLeish, 1976, [75], Da Silva & Kyriakides, 2007, [74]), and therefore balsa core producers typically list the density of an assembled sheet of core material as its main characteristic

property. However, the density reported is just the average density of the sheet, and due to the large natural scatter in balsa properties the densities of individual blocks can vary.

It can be expected that the resulting balsa core material properties have a relatively high variance and low uniformity which can result in material properties markedly different from what the usually quoted values might suggest. Quantifying the uncertainties in material properties and their influence on the safety of the structure is therefore of high importance for ensuring the quality of the engineering design of balsa-cored sandwich structures. To the knowledge of the author, extensive statistical data for fatigue properties of balsa and studies on the importance of these properties for structural design are not available. Therefore, as part of the present study, constant-amplitude fatigue tests are carried out in order to construct a probabilistic model for fatigue life of balsa core, which is then used as basis for reliability analysis.

Constant amplitude fatigue tests

When a model is calibrated to the results from test data, the size of the data sample has a strong influence to the uncertainty in the model parameters. To the knowledge of the author, data sufficiently numerous to represent the statistical properties of the fatigue life of balsa core material subjected to transverse shear are not publicly available and such a study is not mentioned in literature. Therefore, in order to obtain an extensive data set, a series of constant-amplitude cyclic 3-point bending tests were carried out. The following test setup was used:

- Specimen dimensions: width 90mm; length 400mm (360mm support length)
- Core material: Baltek SP100 end-grain balsa core manufactured by 3A composites [1], nominal density $153 kg/m^3$, sheet thickness 18mm
- Face configuration: GFRP laminate, layup $[\pm 45, 0_4]$, thickness 2.25mm
- Roller diameter: test roller $\phi 40mm$, support rollers $\phi 40mm$
- Subject to 3-point bending with force control
- R-ratio is 0.1
- All specimens experience shear failure in core
- 47 constant-amplitude tests were carried out, 10 static tests were run as reference

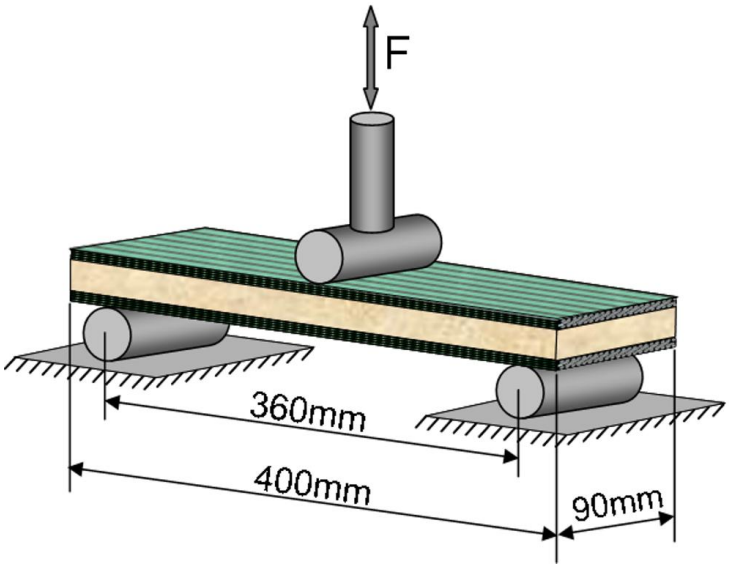

Figure 3.11: 3-point bending test setup scheme.

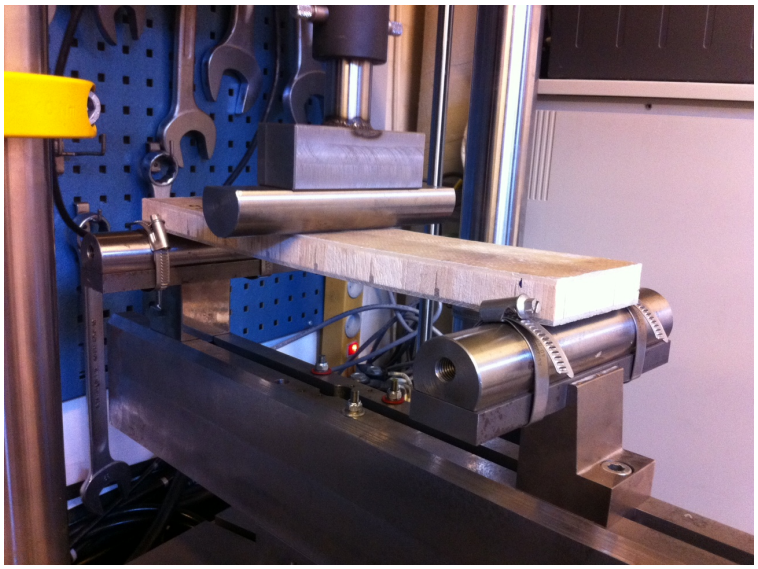

Figure 3.12: Actual setup for 3-point bending test.

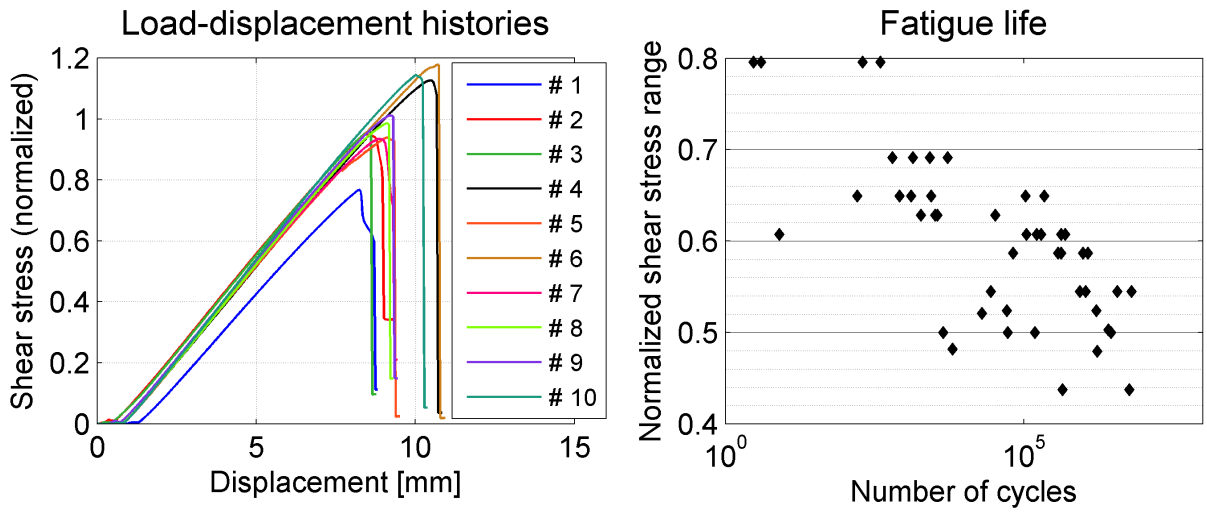

Figure 3.13: Results from static and constant-amplitude fatigue tests of balsa core.

The test setup is shown schematically on Figure 3.11. Figure 3.12 shows the actual test setup, and Figure 3.13 shows test results plotted on a $\log S$ - $\log N$ paper. The significant scatter in properties seen in Figure 3.13 underlines the need for using a probabilistic model for describing the fatigue life of balsa core material.

Probabilistic fatigue life model for balsa core

The fact that the mechanical properties of balsa are strongly dependent on the wood density is used as the basis for the assumption that the large observed scatter in fatigue life is due to variations in the density of the individual balsa blocks, as well as the presence of natural, randomly occurring imperfections. Failure under fatigue loads will occur first in the block which has lowest mechanical properties. Such a behaviour is consistent with Weibull's weakest link theory [90], meaning that the scatter in fatigue life can be modelled by a Weibull distribution. However, one should bear in mind that due to the relatively large size of the individual blocks it is expected that size effects will be present. This is the main reason for choosing a relatively large specimen width for the fatigue tests described above.

The model equation describing the probabilistic relationship between cyclic stress amplitude and number of cycles to failure for balsa core is given in terms of the sum of a mean value represented by an $S - N$ relationship with a random, Weibull-distributed deviation term:

$$\log S = g(\log N, \boldsymbol{\theta}) - \epsilon_m \tag{3.50}$$

where $\log S$ is the logarithm of the cyclic stress amplitude, $g(\log N, \boldsymbol{\theta})$ is the model function (in this case the $S-N$ relationship given in equation 3.20), $\boldsymbol{\theta}$ are the model parameters, and ϵ_m is the random deviation term.

The presently used definition of the scatter in fatigue life as deviation from a reference value requires the use of the three-parameter Weibull distribution as it allows both positive and negative values of the variate, in contrast with the two-parameter Weibull distribution which is only defined for positive values. The distribution of the deviation term will be skewed towards the positive side (as defined in equation 3.50 a positive error term will mean reduced fatigue resistance).

The model for balsa fatigue life when using the above definitions requires a total of six parameters:

$$\boldsymbol{\theta} = [S_u, q, m, A_{\mathcal{W}}, k_{\mathcal{W}}, \gamma_{\mathcal{W}}]$$

where S_u is the ultimate static strength, q is the over-prediction parameter as described in equation 3.20, m is the slope of the $S-N$ curve, and $A_{\mathcal{W}}, k_{\mathcal{W}}, \gamma_{\mathcal{W}}$ are the parameters of the Weibull distribution describing the scatter term. Parameter values are determined by maximizing the likelihood function

$$\mathcal{L}(\boldsymbol{\theta}) = \prod_{j=1}^{n} f_{\mathcal{W}}\left(\hat{\epsilon}_j | \boldsymbol{\theta}\right) \tag{3.51}$$

where $f_{\mathcal{W}}$ is the Weibull probability density of the vector of observed deviation terms $\hat{\boldsymbol{\epsilon}}$:

$$f_{\mathcal{W}}(\hat{\epsilon}_j | \boldsymbol{\theta}) = f_{\mathcal{W}}\left(\left\{ g(\log N_j, S_u, q, m) - \log S_j \right\} \middle| A_{\mathcal{W}}, k_{\mathcal{W}}, \gamma_{\mathcal{W}} \right) \tag{3.52}$$

Mean values of model parameters are determined by maximizing the likelihood function in equation 3.51. The covariance matrix of the parameters is estimated using the property (Richards, 1961 [67]) that the negative of the inverse Hessian of the log-likelihood function, taken at the likelihood-maximizing point, is an estimate of the parameter covariance matrix, $\boldsymbol{\Sigma}_{\boldsymbol{\theta}\boldsymbol{\theta}} = -[\nabla\nabla \log(\mathcal{L})]^{-1}$. Table 3.9 lists the distributions of the model parameters determined by maximization of the likelihood function. The only significant correlation is between m and S_u, and it is $\rho = 0.71$ (Table 3.10).

Figure 3.14 illustrates the probability distribution of the error terms evaluated using the likelihood-maximizing values of $\boldsymbol{\theta}$. Using the basis $S-N$ relationship and the Weibull-distributed random error, the probabilistic fatigue life model for balsa core is constructed. It represents a family of $S-N-P$ curves, where the $S-N$ curve is shifted corresponding to the probability level $P = \text{Prob}\left[S \geq g(\log N, \boldsymbol{\theta}) - \epsilon\right]$. Figure 3.15 shows a graphical

Table 3.9: Parameters of probabilistic model for balsa fatigue life.

Parameter	Designation	Mean	c.o.v.	Unit
Ultimate strength	S_u	$5.185 \cdot 10^6$	0.013	Pa
Overprediction of S_u by linear model	q	1.000	0.0	-
Fatigue slope	m	25.330	0.037	-
Weibull scale parameter	A	0.0879	0.0	-
Weibull shape parameter	κ	1.583	0.0	-
Weibull location parameter	γ	-0.114	0.0	-

Table 3.10: Correlations between parameters of the probabilistic model for balsa fatigue life.

Parameter	S_u	m
S_u	1	0.71
m	0.71	1

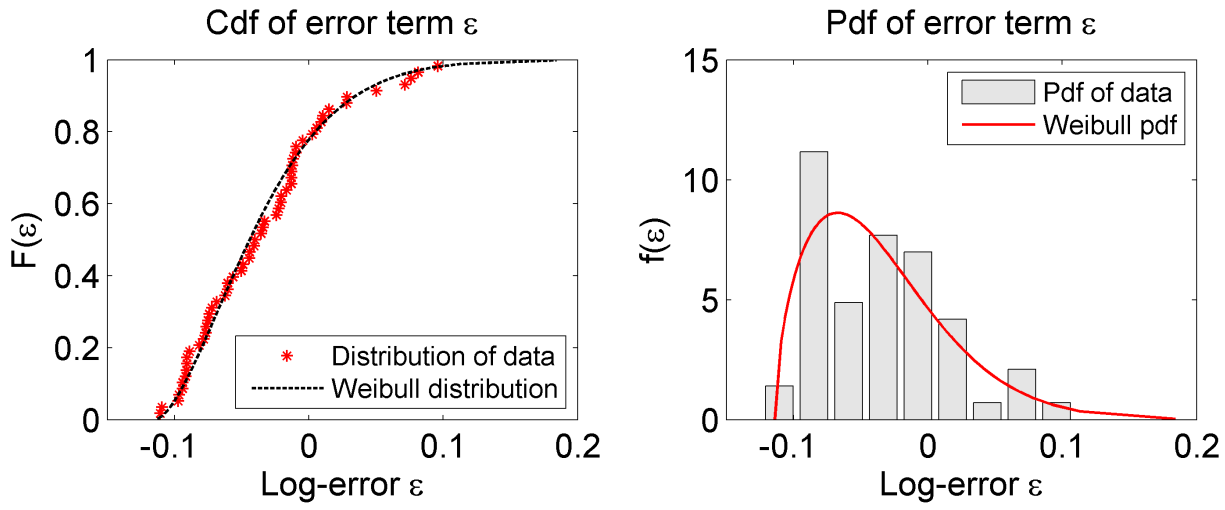

Figure 3.14: Probability density of model deviation term.

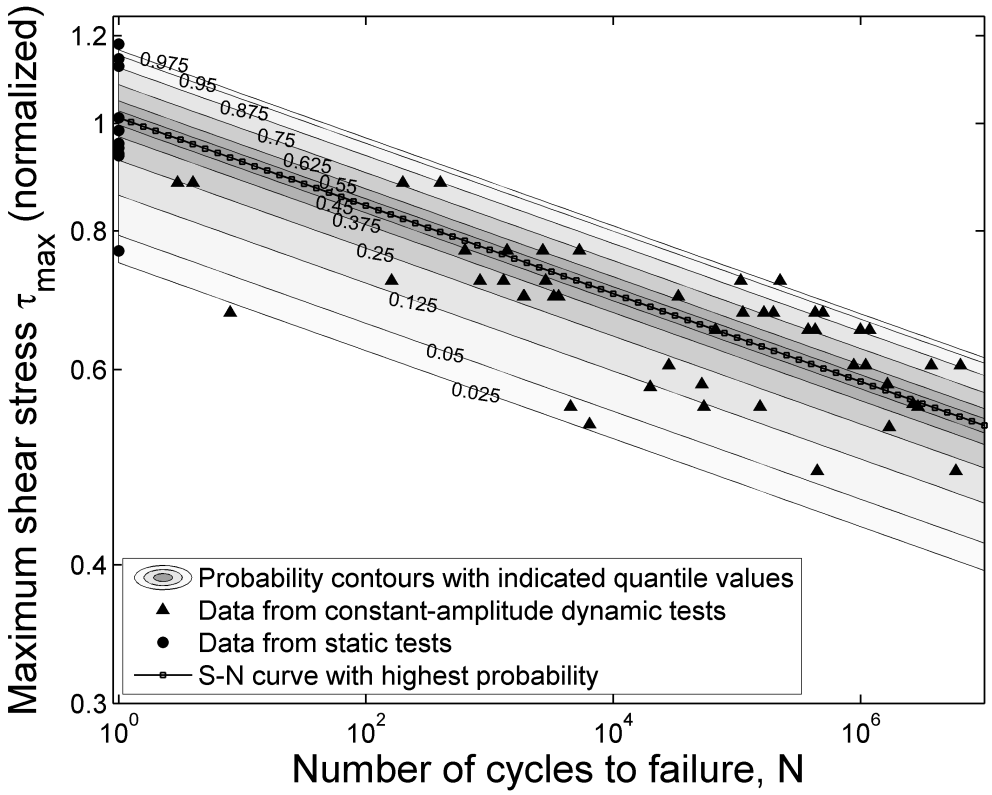

Figure 3.15: Probabilistic model for balsa fatigue life shown in an $S - N$ diagram.

representation of the model with probabilities given in terms of confidence bounds. The $S - N$ relationship is given in terms of the maximum stress τ_{max}, which allows for comparison between the static and fatigue properties of the material, as discussed in Section 3.3.2.

The optimal value of the parameter q for balsa core equals 1 as shown in table 3.9. This result suggests that the nonlinearity in the low-cycle fatigue of balsa core is very limited, because the $S - N$ relationship remains linear for loads as high as the static strength of the material. The standard deviation of the Weibull-distributed error term ϵ_m is 16% of the mean static strength of balsa. This is at least three times higher than the corresponding standard deviation for the face laminates, which does not exceed about 5% of their mean static strength.

3.4 Discussion on probabilistic modelling of composite material properties

The present chapter treated the behaviour of composites under ultimate and fatigue limit states. Laminates and face laminates in sandwich composites show significant system behaviour in ultimate limit state, and have to be modelled as a reliability system. Due to the significant load redistribution following fiber failures, using a progressive failure approach requires treating the laminate as a general system. Matrix degradation does not play a large role in the ultimate strength estimate, and therefore the problem may be approximated with a simpler, series-system representation by assuming failure at the point when first fiber failure occurs, while still allowing matrix degradation prior to fiber failures. As seen in the results on Figure 3.6, the series system representation is able to capture the main physics of the behaviour of the structure, however it is slightly conservative.

The lamina static properties are used as variables in the fatigue life model, and it can be therefore expected that the scatter in fatigue life is similar to the scatter in ultimate strength. This is the case for most of the models where the tensile fatigue properties are modelled, however as seen on Figure 3.9, d), under compressive loads the scatter in fatigue properties can be much larger. The high variance of fatigue life under compressive loading will have an influence on the structural reliability, and can possibly result in lower reliability, or in a requirement for higher safety factors.

The variance in balsa properties, both ultimate and fatigue, is several times higher than the variance in face laminate properties. This can be attributed to the fact that balsa wood is a natural product. The loads on the sandwich cores in a wind turbine blade structure are not very high, because the bending stresses are significantly larger than the shear stresses, and the out-of-plane loads in terms of aerodynamic pressure are relatively small. Therefore it is not expected that core material properties will have a decisive role in the reliability of the blade, however the high scatter in balsa properties has to be kept in mind when designing other structures where the core strength is as critical as the strength of the faces.

Chapter 4

Load statistics

Loads are a major source of uncertainty in the design of many engineering structures. For wind turbines which are exposed to a turbulent wind field and are subject to dynamic effects the importance of load characteristics is even more obvious. Wind turbine design loads are typically determined using dynamic simulations of the turbine response, and despite the constantly improving computational capabilities of computers, the number of simulations that can be used is nevertheless limited to a few hours of time-series per simulation condition. Design load levels are usually obtained by employing one or both of the following two approaches: (a) Simulate events with rare occurrence that cause high load levels and are thus likely to be the design drivers, and (b) Simulate turbine response under normal operating conditions, and extrapolate the response by fitting a probability distribution to the response extremes. The IEC61400-1 standard [37] requires the use of both approaches in combination for obtaining extreme design loads, while the number of load cycles relevant for fatigue limit state is governed mainly by the simulations under normal operating conditions. In the present thesis loads under normal operating conditions are considered, which gives the possibility to analyze the behaviour of both extreme and fatigue loads, without the need to involve possibly lengthy discussions about frequency and severity of extreme events and malfunctions.

4.1 Defining short-term distribution of load extremes

The IEC61400-1 standard [37], which is one of the most widely used guidelines for structural design of wind turbines, states that a minimum of 15 simulations per wind speed (each simulation having a duration of 10 minutes) are required for calculating ultimate limit state

loads under normal production conditions. Extreme load events with a period of recurrence of 50 years have to be evaluated based on information for extreme events with a return period of a few hours or days. In other words, a load level which is several orders of magnitude less probable than the largest observed load level has to be estimated.

The present state-of-the-art statistical extrapolation methods perform very well at moderate degrees of extrapolation, as for example the assessment of an environmental condition (wind speed, wave elevation) with a 100-year return period based on up to 20 years of measured data (see for example Naess & Gaidai, 2009 [55]). However, when the amount of data are limited to much shorter time periods, the uncertainty is much higher and accurate prediction is not assured. It is therefore important to assess what is the minimum amount of data required for robust and accurate prediction of extreme load events using statistical load extrapolation. The present chapter shows a study on the performance of several load extrapolation methods, and gives a set of recommendations for ensuring high accuracy of the extreme load predictions. The first part of the study on extreme loads has the objectives of assessing the performance of several load extrapolation techniques currently used by industry, and describing an optimal procedure for estimating the short-term statistical distribution of extreme load events. Then it is demonstrated how the short-term load statistics are integrated in order to obtain the long-term probability distribution of extreme loads.

A necessary feature of design methods that are used by industry is their practicality, which requires that the process is easy to use and needs minimal intervention. However, the load extrapolation process often involves the selection of a peak threshold level, indicating the load range above which the recorded events can be considered as extremes. In practice either a fixed threshold level, e.g., a certain number of standard deviations above the mean value of the process, or a manually selected threshold are used. Here a method for automatic estimation of the threshold level is also defined, which allows for fully automated load extrapolation process.

The present study evaluates the performance of three families of asymptotic load analysis techniques - the block maxima method (Gumbel, 1958 [30]), the peaks-over-threshold method (see Dekkers et al., 1989 [13], Davison & Smith, 1990 [12]), and the Average Conditional Exceedance Rates (ACER) method, Naess & Gaidai, 2009 [55]. The latter is a relatively new method based on a Markov-chain-type cascade of conditional approximations. It is included due to its apparent robustness, the simple way in which peak correlation is accounted for, and due to its positive reception in literature, e.g., Toft et al., 2011 [85].

Comparative assessment of the different load analysis methods is done by applying all

the methods to the same set of problems and evaluating their performance according to the following criteria:

- Accuracy of the estimate (precise estimate and narrow confidence bands, including the cases of non-Gaussian response and autocorrelated time-series)
- Minimal deterioration in performance for very large extrapolations
- Amount of data necessary for obtaining a sufficiently reliable estimate

The load extrapolation procedure will be considered as a process consisting of two stages: first, identifying independent peak events that belong to the extreme tail of the response (peak extraction), and second, fitting an extrapolation function, usually in the form of a statistical distribution, to the extracted response extremes.

4.1.1 Analysis methods

Peak extraction

The extremes of a random process follow a distribution different than the distribution which characterizes the points close to the mean value of the process. In order to assess the statistical distribution of the extremes, peak events have to be identified. In addition to selecting process values above a certain threshold, in the case of an auto-correlated process the extracted peaks should be declustered in order to make sure that each peak corresponds to an independent event.

Block maxima The simplest way to obtain a set of independent extremes is to make sure that the peaks selected are sufficiently far apart. This is done by dividing the response history into a set of time periods (blocks) of equal length, and extracting a single peak (the absolute extreme) out of each block in the set. If the length of the time period is sufficiently large, or if the blocks represent independent realizations of the random process, the peaks are uncorrelated. If the data points belong to the same realization of the random process, an additional requirement will be that peaks extracted from adjacent blocks are not too close apart, i.e., the time separation of the peaks is greater then the autocorrelation length. While this approach in general ensures that the sampled data are non-correlated and they clearly belong to the tail of the response distribution, it has a drawback - only a very limited part of the data are used, making the estimate susceptible to noise and bias effects.

Peaks over threshold The method of peaks over threshold (POT) aims at eliminating the problem of limited utilisation of the available information inherent to the block maxima method. In the peaks-over-threshold method all independent response peaks which exceed a certain high threshold level are included in the analysis. This requires that there are means of extracting independent peaks from the time series by identifying uncorrelated events. Here, independence of peak events is assumed when a given peak event is sufficiently separated in time from other peak events, i.e., if it is not preceded or succeeded by a larger peak for a certain number of time steps.

Average Conditional Exceedance Rates method Naess & Gaidai, 2009 [55] have recently developed a method for determining exceedance probabilities on the basis of cascade of conditioning approximations. As in the other methods discussed, the goal is to accurately determine the distribution function of the extreme value $M_N = \max[X_j, j = 1 \dots N]$. Specifically, the value of $P(\eta) = P(M \leq \eta)$ has to be estimated accurately for large values of η.

From the definition of $P(\eta)$ it follows that

$$\begin{aligned} P(\eta) = P\{X_1 \leq \eta, \dots, X_N \leq \eta\} = \\ = P\{X_N \leq \eta | X_1 \leq \eta, \dots, X_{N-1} \leq \eta\} \cdot P\{X_1 \leq \eta, \dots, X_{N-1} \leq \eta\} = \\ = \prod_{j=2}^{N} P\{X_j \leq \eta | X_1 \leq \eta, \dots, X_{j-1} \leq \eta\} \cdot P(X_j \leq \eta) \end{aligned} \tag{4.1}$$

Equation 4.1 takes such a form due to the fact that variables X_j are generally dependent. For completely independent variables equation 4.1 will be reduced to the much simpler expression

$$P(\eta) \approx \prod_{j=1}^{N} P(X_j \leq \eta) \tag{4.2}$$

A way to account for the dependency between X_j to some extent is the assumption of a Markov-like, one-step memory cascade of conditional approximations:

$$P(X_j \leq \eta | X_1 \leq \eta, \dots, X_{j-1} \leq \eta) \approx P(X_j \leq \eta | X_{j-1} \leq \eta) \tag{4.3}$$

for $2 \leq j \leq N$. This can be extended by considering three steps instead of two:

$$P(X_j \leq \eta | X_1 \leq \eta, \dots, X_{j-1} \leq \eta) \approx P(X_j \leq \eta | X_{j-2} \leq \eta, X_{j-1} \leq \eta) \tag{4.4}$$

for $3 \leq j \leq N$, and so on. Equations 4.3 and 4.4 represent series with increasing amount of dependence between adjacent points taken into account (equation 4.4 accounts for the

correlation between successive time steps in a higher degree compared to equation 4.3). Including even more points would further enhance the correlation assumption, although in most cases including 2 or 3 point series (equations 4.3 or 4.4) would be sufficient.

Introducing the quantities α_{kj} as the exceedance probabilities based on $k-1$ previous non-exceedances:

$$\alpha_{kj}(\eta) = P(X_j > \eta | X_{j-k+1} \leq \eta, \ldots, X_{j-1} \leq \eta) \tag{4.5}$$

Then, the concept of average conditional exceedance rates (AER) $\bar{\varepsilon}_k$ is also introduced:

$$\bar{\varepsilon}_k(\eta) = \frac{1}{N-k+1} \sum_{j=k}^{N} \alpha_{kj}(\eta) \tag{4.6}$$

The empirical estimation of the conditional AER proceeds by counting the total number of favourable events (exceedance of the given threshold, given the required number of previous non-exceedances), and then dividing by $N-k+1$. The numerical implementation of this can be easily done by introducing the following two quantities:

$$A_{kj}(\eta) = \mathbf{1}\{X_j > \eta, X_{j-1} \leq \eta, \ldots, X_{j-k+1} \leq \eta\}, \quad j = k, \ldots, N, \quad k = 2, 3, \ldots \tag{4.7}$$

$$B_{kj}(\eta) = \mathbf{1}\{X_{j-1} \leq \eta, \ldots, X_{j-k+1} \leq \eta\}, \quad j = k, \ldots, N, \quad k = 2, 3, \ldots \tag{4.8}$$

where $\mathbf{1}(\mathcal{A})$ denotes the indicator function of an event $\mathcal{A}$. Then

$$\alpha_{kj}(\eta) = \frac{E[A_{kj}(\eta)]}{E[B_{kj}(\eta)]} \tag{4.9}$$

where $E[.]$ denotes the expectation operator. Assuming an ergodic process, $\bar{\varepsilon}_k(\eta) = \alpha_{kk}(\eta) = \ldots = \alpha_{kN}(\eta)$. Then, for finite time-series

$$\bar{\varepsilon}_k(\eta) = \lim_{N \to \infty} \frac{\sum_{j=k}^{N} A_{kj}(\eta)}{\sum_{j=k}^{N} B_{kj}(\eta)} \tag{4.10}$$

For both stationary and non-stationary time-series the sample estimate of $\bar{\varepsilon}_k(\eta)$ would be

$$\hat{\varepsilon}_k(\eta) = \frac{1}{R} \sum_{r=1}^{R} \hat{\varepsilon}_k^{(r)} \tag{4.11}$$

where R is the number of realisations (samples), and

$$\hat{\varepsilon}_k^{(r)} = \frac{\sum_{j=k}^{N} A_{kj}^{(r)}(\eta)}{\sum_{j=k}^{N} B_{kj}^{(r)}(\eta)} \tag{4.12}$$

where the index (r) refers to realisation number r. Further simplifications to equations 4.8 - 4.12 can be done by assuming that

$$\lim_{\eta\to\infty}\sum_{j=k}^{N} B_{kj}(\eta) = N-k+1 \approx N \tag{4.13}$$

This assumption will lead to

$$\hat{\varepsilon}_k^{(r)} = \frac{\sum_{j=k}^{N} A_{kj}^{(r)}(\eta)}{N-k+1} \tag{4.14}$$

Since equation 4.14 is derived under the assumption that $\eta \to \infty$, it is expected that it will be valid in the extreme part of the response, and can thus be used instead of equation 4.12 without significant loss of accuracy.

A predictive function for the extreme response levels can now be defined based on the assumption that the far tail of the response will asymptotically follow a Gumbel distribution. The function describing the asymptotic behaviour is assumed to be of the form $\exp\{-a(\eta - b)^c\}, \quad (\eta \geq \eta_1 \geq b)$, where a, b and c are constants, and η_1 is a chosen tail threshold level. Using this function, the average conditional exceedance rates determined by equation 4.12 or 4.14 can be represented in the following form:

$$\tilde{\varepsilon}_k(\eta) \approx q_k(\eta)\exp\{-a_k(\eta - b_k)^{c_k}\} \tag{4.15}$$

The value $c_k = 1$ corresponds to the asymptotic case when the data follows exactly the Gumbel distribution. Introducing c_k as a parameter greatly increases the flexibility of the extrapolation function, allowing it do describe a larger family of distributions, a subset of which is the Gumbel distribution. By assuming that the function $q_k(\eta)$ is slowly varying with respect to η, it can be replaced by a constant, q_k. Now the extrapolation function is defined by four constants, a_k, b_k, c_k, and q_k. Their values are determined using constrained optimisation on the weighted residual function

$$\sum_{j=1}^{N} w_j \left| \log\hat{\nu}^+(\eta_j) - \log q + a(\eta_j - b)^c \right|^2 \tag{4.16}$$

where $w_j = (\log CI^+(\eta_j) - \log CI^-(\eta_j))^{-2}$ denotes a weight factor putting more emphasis to more reliable data points. If necessary, changing the exponent value from -2 to a larger value, say 1, will put more emphasis on the less certain data points.

Statistical distribution of extremes

In the present study all extrapolation methods considered are based on asymptotic analysis, meaning the basic assumption is that the process extremes will asymptotically follow a certain type of a generalized extreme value distribution.

Gumbel distriution Under the assumption that the epochal extremes of a random process follow asymptotically a Gumbel (type I extreme value) distribution, the probability of the process being below a certain level can be defined as

$$P(X \leq x) = F_G(x, \alpha, \beta) = e^{-e^{\alpha(x-\beta)}} \tag{4.17}$$

The doubly-exponentiated expression $\alpha(x-\beta)$ is the equation of a straight line with a slope of α and offset $\alpha\beta$, meaning that a plot of x vs. $\ln(-\ln(F(x)))$ will represent a straight line. Therefore, under the assumption that the data follow a Gumbel distribution, the distribution parameters can be estimated by making a least-squares line-fit to the double logarithm of the cumulative distribution function $F(x)$.

Weibull distribution The three-parameter Weibull distribution offers greater flexibility due to the presence of an extra parameter in comparison to the Gumbel distribution. Distribution parameters A, κ and γ are obtained based on a least-squares optimization, similar to what is used for fitting the ACER function:

$$r = \sum_{j=1}^{N} w_j \Big[\log\left\{\hat{P}_j\right\} - \log\left\{1 - F_{\mathcal{W}}(\hat{x}_j, A, \kappa, \gamma)\right\} \Big]^2 \tag{4.18}$$

where $F_{\mathcal{W}}(x, A, \kappa, \gamma)$ is the Weibull density function, $\hat{x}_i$, $i = 1 \ldots N$ is the set of observations, and $\hat{P}_j = P(X \geq \hat{x}_j)$ is the empirically observed probability of exceeding response level $\hat{x}_j$.

Pareto distribution The underlying assumption for using the Pareto distribution is that the independent response peaks which exceed a certain high threshold level will follow a generalized Pareto distribution, if and only if the parent distribution belongs to the domain of attraction of one of the extreme value distributions (Pickands, 1975 [64]). The distribution function of the generalized Pareto distribution is the following:

$$P(X \geq y) = F_{GP}(y, a, c) = 1 - \left(1 + c\frac{y}{a}\right)_+^{-1/c} \tag{4.19}$$

where $a > 0$ is a scale parameter, $c(-\infty < c < \infty)$ is the shape parameter of the distribution. $(z)_+ = \max(0, z)$.

Equation (4.19) can be used to represent the conditional cumulative distribution function of the excess $Y = X - u$ of the observed variate X, where u is a sufficiently large threshold [64]. The cases of $c > 0$, $c = 0$ and $c < 0$ correspond to type II (Fréchet), type I (Gumbel) and type III (reverse Weibull) domains of attraction, respectively. In the case when $c = 0$ Equation (4.19) is considered in a limiting sense as $P_{GP}(y, a) = 1 - \exp(-y/a)$.

While Equation (4.19) gives the possibility to calculate the probabilities associated with a certain load level, it is sometimes more convenient to do the inverse action - estimating the load level which is associated with a given return period, R. The return period R of a given response level is defined as the inverse of the probability that this level will be exceeded during the time frame of interest. If λ denotes the mean crossing rate of a certain threshold u for a time period $T = 1$, then the return period R of the value of X corresponding to the level $x_R = u + y$ will be given by

$$R = \frac{1}{\lambda \cdot P(X > x_R)} = \frac{1}{\lambda \cdot P(Y > y)}. \tag{4.20}$$

It follows that

$$P(Y \geq y) = 1 - \frac{1}{\lambda R} \tag{4.21}$$

Substituting $P(Y \geq y)$ from Equation 4.19 into Equation 4.21 yields the following expression for x_R:

$$x_R = u - a\left[1 - (\lambda R)^c\right]/c. \tag{4.22}$$

If the parameters a and c can be estimated from an available data set, then equations (4.19) or (4.22) can be used to assess the probabilities and response levels associated with extreme events. A widely used approximation to a and c is given by the so-called de Haan estimators (Dekkers et al., 1989 [13]). These estimators are based on the following quantities:

$$H_{k,n} = \frac{1}{k}\sum_{i=0}^{k-1}\left\{\ln(X^*_{n-i}) - \ln(X^*_{n-k})\right\} \tag{4.23}$$

$$H^{(2)}_{k,n} = \frac{1}{k}\sum_{i=0}^{k-1}\left\{\ln(X^*_{n-i}) - \ln(X^*_{n-k})\right\}^2 \tag{4.24}$$

In the equations above, k is the number of observations above the threshold, u. X^*_i represent the variates, ordered accodring to their values, where $X^*_n, X^*_{n-1}, \ldots, X^*_{n-(k+1)}$ are the highest, second highest,..., k^th highest variate, respectively. An estimate for the mean crossing rate λ is $\hat{\lambda} = k/n_T$, where n_T is the total number of time units T that are covered by the data set. Finally, estimators for a and c are given by:

$$\hat{a} = \rho X^*_{n-k} H_{k,n} = \rho u H_{k,n} \tag{4.25}$$

and

$$\hat{c} = H_{k,n} + 1 - \frac{1}{2}\left\{1 - \frac{(H_{k,n})^2}{H^{(2)}_{k,n}}\right\}^{-1} \tag{4.26}$$

where $\rho = 1$ if $\hat{c} \geq 0$, while $\rho = 1$ if $\hat{c} < 0$.

Another set of estimators (the so-called moment estimators) give estimates of a and c in terms of the mean value $E(Y)$ and the standard deviation $\sigma(Y)$:

$$a = \frac{1}{2}E(Y)\left(1 + [E(Y)/\sigma(Y)]^2\right) \tag{4.27}$$

and

$$c = \frac{1}{2}\left(1 - [E(Y)/\sigma(Y)]^2\right) \tag{4.28}$$

Although the de Haan estimators are widely used, there are indications that the moment-based estimators might produce more accurate results, see Naess & Clausen, 2001 [54].

Exponential function asymptotically approximating the extreme value distribution The extrapolation function used in connection with the ACER method (equation 4.15) represents a class of double-exponential functions, a subclass of which is the asymptotic Gumbel distribution for $c = 1$. Using $c \neq 1$ results in greater flexibility as it allows the representation of a much larger class of asymptotic functions. This increased flexibility can be useful and possibly improve the performance of extrapolation methods compared to when using a more restricted class of distributions such as the Gumbel distribution. Therefore, here the exponential function given by equation 4.15 is also used for peak data that are extracted with methods other than the ACER method. To avoid confusion, when using the ACER extrapolation function outside the context of the ACER method, it will be referred to as "Asymptotic Extreme Value (AEV)" function.

Estimation of confidence intervals

Confidence intervals play an important role in assessing whether the obtained estimates for response levels and return periods are acceptable and can be trusted. The simplest way to determine a confidence interval is to assume that the error in predictions will follow the Normal distribution. Then, if a number of predictions are available, the confidence intervals at confidence level $1 - q$ can be determined from the variance of the available predictions:

$$CI_{\pm}(X) = Y(X) \pm \Phi^{-1}(1 - q/2)\sigma_Y(X) \tag{4.29}$$

For such an approach, about 20-30 sample sets would give a good confidence estimate.

When only a single data set is available it is not possible to directly estimate the variance of the estimate. In such cases the confidence intervals are determined by bootstrapping (Efron, 1979 [21]). Here the non-parametric bootstrapping technique is described, where the distribution of the response variable $Y = f(\mathbf{X})$ is determined purely empirically by creating

a number of realisations of an input sample $\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_l}$. The bootstrap samples are drawn from the original input sample the following way: first, all the data points from the input sample of size N are given an integer index number $[1, N]$. Then a new bootstrap sample of size N is created by drawing N random integers from the interval $[1, N]$. The new sample is constructed from the datapoints from the original sample with indexes corresponding to the generated set of integers. Since the drawing procedure allows number repetitions, the resulting sample will in most cases differ from the original sample. Confidence intervals can be obtained based on the variance of the bootstrap sample by equation 4.29. If one wants to avoid the assumption of normality of the bootstrap sample, the empirical density function of the bootstrap sample can be obtained by ranking the sampled response realisations $y_i = f(x_i)$ according to their values. If l number of bootstrap samples have been realised, and the set of obtained responses, ranked by value is designated as R^*, with the i^{th} ranked value being R_i^*, then the lower and upper bounds with a confidence level of $1 - q$ are determined by:

$$CI_- = R^*_{[ql/2]} \qquad , \qquad CI_+ = R^*_{[(1-q/2)/l]} \tag{4.30}$$

where the square brackets $[a]$ mean the integer part of a. This technique for estimating confidence intervals is as simple to apply as the variance based method, but there are no assumptions made about the distribution of the error. However this comes at the expense of having to generate much larger sample. Typically 1000 or more samples will be necessary for obtaining a good estimate using the empirical cdf of the bootstrap sample and equation 4.30.

Discarding uncertain data

Despite that the main idea of peak extrapolation methods is to model the extreme tail of the response, the highest peaks are sometimes so scattered that their presence deteriorates the quality of the calculation (see Figure 4.1). Peaks with too high uncertainty are excluded by considering the width of confidence intervals associated with them. For the ACER method, where the data series are split in a number of samples R, the 95% confidence bounds of the exceedance probabilities ε at each threshold level η are determined based on sample variance. Any threshold level for which the relative confidence exceeds a certain quantity δ is excluded from the calculations:

$$\frac{1.96\hat{s}_k(\xi)/\sqrt{(R)}}{\tilde{\varepsilon}(\eta)} > \delta \tag{4.31}$$

In the present study, a value of $\delta = 0.9$ is used. Naess & Gaidai, 2009 [55] recommend values between 0.5 and 1.

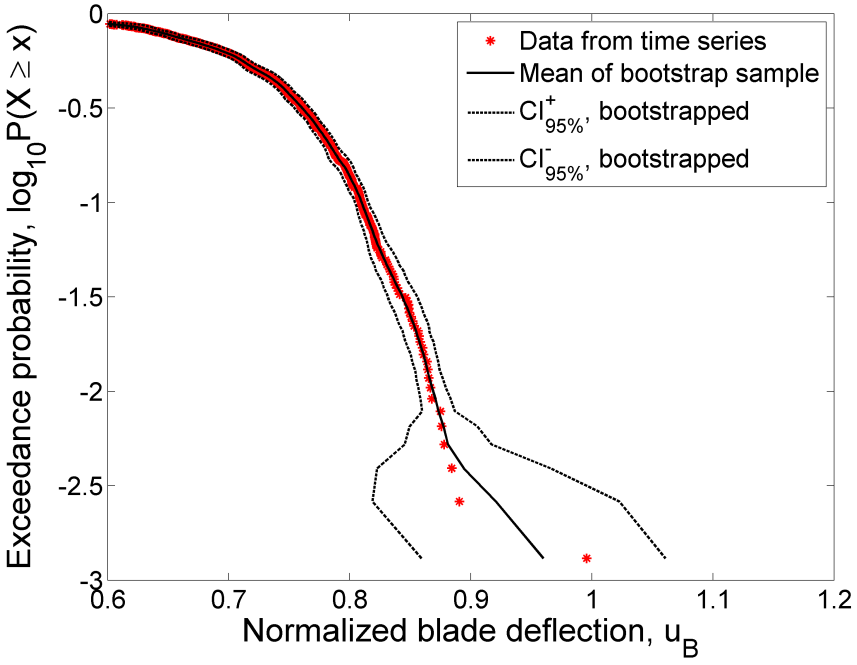

Figure 4.1: Example of a single peak event with high uncertainty.

When using the peaks-over-threshold method the empirical exceedance data are not necessarily averaged as in the ACER method, but instead all data points may be used as a single data set. Since in that case the variance between samples cannot be calculated, confidence intervals are estimated using bootstrapping. Then all data points for which the width of the 95% confidence interval is greater than 15% of the response level, are excluded. This corresponds to data points for which the coefficient of variation of the bootstrap sample is greater than 3.75%.

Convergence criteria

Quantifying the amount of data necessary for obtaning a quality extrapolation requires defining convergence criteria. Here, the width of the estimated confidence intervals at a target extrapolated response level will be used as an objective measure of convergence. IEC61400-1 ([37]) suggests that the width of the 90% confidence interval at the 84% quantile of the extrapolated load distribution function should be less than 15% of the quantile value. It can be expected however that even when this convergence criterion is satisfied the uncertainty in the extrapolated load levels at higher quantiles might be unacceptably high. Therefore, a stricter criterion is also defined for the present study, stating that convergence is achieved

when the width of the 95% confidence interval at the target extrapolation level should be not more than 20% of the estimated response value for that probability level. Under the assumption that the error is approximately normally distributed, this means that convergence is achieved when the coefficient of variation of the response estimate is less than ≈ 0.05. The importance of convergence criteria is assessed later by comparing the uncertainty in the estimates achieved under the two convergence conditions specified above.

Determining the threshold level

Using an appropriate peak threshold level is of significant importance for all extrapolation methods which are being discussed. When determining the peak distribution parameters it has to be made sure that all data points used for the fit belong to the extreme tail of the response, while still enough data need to be included in order to make a quality fit. While a manual selection of the threshold by inspection of the data allows the highest degree of control and is thus probably the "safest" approach for avoiding errors, it is not very practical. Therefore a method for automatic determining of a threshold level is proposed, based on the assumption that process extremes will asymptotically follow the Gumbel distribution.

When considering the tail as asymptotically following the Gumbel distribution, we are invariably making the assumption that at sub-asymptotic levels the response will follow a different distribution. Therefore, at the transition between regular response levels and extreme response levels one should expect a gradual change in behaviour of the threshold exceedance probabilities. If such a change in behaviour can be observed, then it is possible to determine a threshold level above which asymptotic tail behaviour is dominant. For any Gumbel-distributed data set, a plot of x vs. $\ln\left(-\ln\left(F(x)\right)\right)$ represents a straight line with a slope of $-\alpha$ and intercept of $\alpha\beta$, where α and β are the parameters of the Gumbel distribution. Therefore, if a straight line is fit to the highest extremes of an asymptotically Gumbel-distributed process, the residual error should be relatively small. On the other hand, if part of the sub-asymptotic region is included in the fit, it is expected that the relative error will increase as the lower response levels might not follow the Gumbel distribution (i.e., they will not represent a straight line in double-log space). In order to determine where the asymptotic tail behaviour becomes dominant, a series of line fits is created starting from the highest possible extremes, with each subsequent fit including more data by lowering the threshold level. For each of the fits, an error estimate corrected for sample size is obtained

using the root mean square error:

$$\varepsilon_{RMS} = \sqrt{\sum_{n=1}^{N} \frac{(\hat{F}_i - F_{GB}(\hat{x}_i))^2}{N}} \tag{4.32}$$

where $\hat{F}_i$ are the empirically observed probabilities of non-exceedance of a response level $\hat{x}_i$, and $F_{GB}(\hat{x}_i)$ is the Gumbel cumulative distribution function evaluated at $\hat{x}_i$. Figure 4.2 shows a typical example of the behaviour of the error as function of threshold level. The interpretation of the results is the following: it is expected that for high threshold levels (the right side of the graph on Figure 4.2) the error will vary somewhat erratically, but will be relatively low. For lower threshold levels, the behaviour will be more regular due to the increased sample size, but at the same time the error will start to grow as more data from the sub-asymptotic region are included, which do not follow the Gumbel distribution. This property of the fitting error can be used to define a criterion for choosing a tail threshold level. Experience from a number of tests has showed that the point where ε_{RMS} stops being monotonically decreasing is a suitable choice for tail threshold level. In order to avoid too many exceptions and false estimations due to noise, some adjustments to the input data can be made - e.g., excluding the highest few data points which have the highest uncertainty, and establishing some minimum required number of points over which to perform the fit.

Synthetic data following the extreme value distribution A set of synthetic data drawn from the extreme value distribution has been used by Naess & Gaidai, 2009 [55] as means to test the performance of the ACER method. Here, data with the same distribution will be used as a test to the performance of the automatic procedure for threshold selection. Replicating the results in [55] will also serve as means for verifying the present implementation of the ACER method.

First, it is assumed that the underlying stochastic process follows a Gaussian distribution with mean of zero and standard deviation of one. Then, using the Poisson assumption, the distribution of the extremes will be given by the following relation:

$$F^{1yr}(\eta) = \exp\{-\nu^+(\eta)T\} \tag{4.33}$$

where ν^+ is the upcrossing ratio of the process, and T is the time period.

If the period of interest is defined as $T = 1$ year, and assuming that the product of the mean zero upcrossing rate $\nu^+(0)$ and the time period $\nu^+(0)T = 1000$, the modelled process will resemble wind speed data. The distribution of the yearly maxima will then be

$$F^{1yr}(\eta) = \exp\left\{-10^3 \exp(-\frac{\eta^2}{2})\right\} \tag{4.34}$$

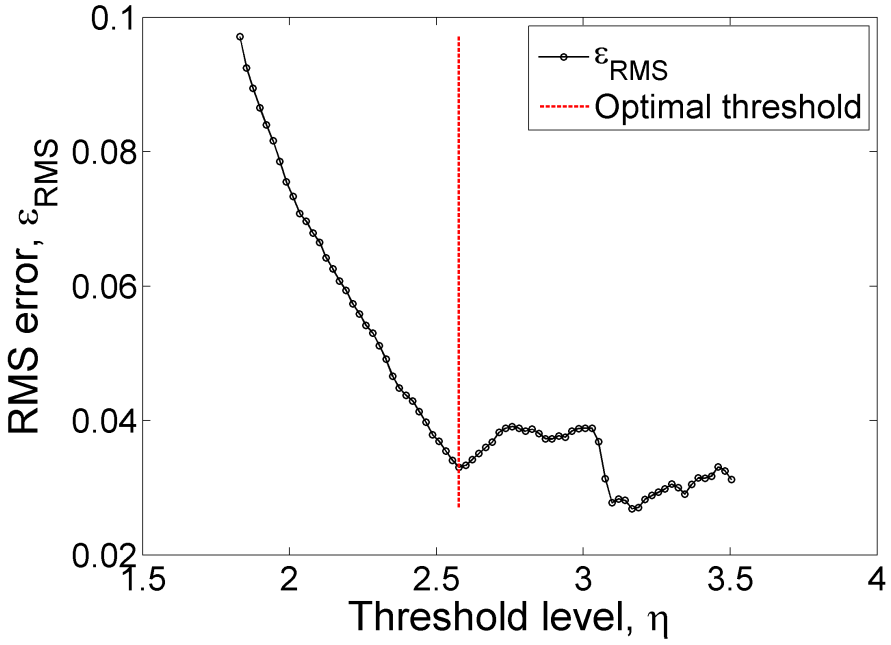

Figure 4.2: Example of RMS error from Gumbel fit as function of threshold level.

The 100-year load value will be found from the relation $F^{1yr}(\eta^{100yr}) = 1 - 1/100$, and it is $\eta^{100yr} = 4.80$. Such wind speed data will typically have peaks separated by several days, say 3-4 days, and therefore peak events are generated from the extreme value distribution:

$$F^{3d}(\eta) = \exp\left\{-q\exp(-\frac{\eta^2}{2})\right\} \tag{4.35}$$

where $q = \nu^+(0)T = 10$, which corresponds to peak separation of approximately 3.65 days, which, as mentioned above, corresponds well with the expected period of separation between independent peaks of wind speed.

Since the peaks generated by equation 4.35 are independent, there is no need to account for the correlation between peaks. Subsequently, extrapolation is applied directly to all data above the tail threshold level, and peak extraction is not carried out. For the ACER method this means that the parameter k (defining the length of the chain of conditional approximations) is set to $k = 1$.

Equation 4.35 is used to generate 30 sets consisting of 2000 points each (i.e., 30 sets of 20 years of data). Extreme values with 100-year return period are estimated using the ACER method, and with applying three different approaches to selecting a threshold level:

- A fixed threshold of $\eta = 2.5$;

Table 4.1: Mean and standard deviation of 100-year wind speed estimate η_{100yr} obtained by the ACER method.

Threshold method	Mean of 30 estimates	Standard deviation
Fixed threshold $\eta = 2.5$	4.838	0.225
Manual threshold	4.897	0.300
Automatic threshold	4.847	0.221

- A manually determined threshold level from visual inspection of time series;

- Automatically determined threshold using the method described above.

Table 4.1 lists the statistics of the calculated 100-year estimates, and Figure 4.3 shows a sample-by-sample comparison of η_{100yr} and the threshold levels used. There is high degree of correlation between the 100-year extreme wind speeds estimated with different threshold levels. This shows that the ACER method is not too sensitive to the threshold level used. It can be seen however that the automatic threshold selection procedure outperforms the manual one because it results in smaller variance of the estimates.

4.1.2 Short-term statistical distribution of wind turbine response

The main application of statistical extrapolation methods in wind turbine design is the estimation of short-term statistical distributions of wind turbine response. The standard extrapolation procedure from IEC61400-1 [37] involves the use of 15 time series per simulation case, which is a very limited amount of data compared to the recurrence periods under consideration. It is therefore not guaranteed that the correct load levels will be predicted indeed. In the following, it is investigated what amount of data is necessary to use with different extrapolation methods in order to obtain extrapolations with sufficiently low levels of uncertainty.

Simulated wind turbine response

A data set of 15000 time series with 10-minute duration is used as a basis for the analysis. The time series represent the simulated response of a 2.3MW Siemens wind turbine under turbulent wind conditions with average horizontal wind velocity of $U = 11m/s$, and turbulence intensity of $I = 0.15$. Simulations are performed using Siemens' proprietary aeroelastic

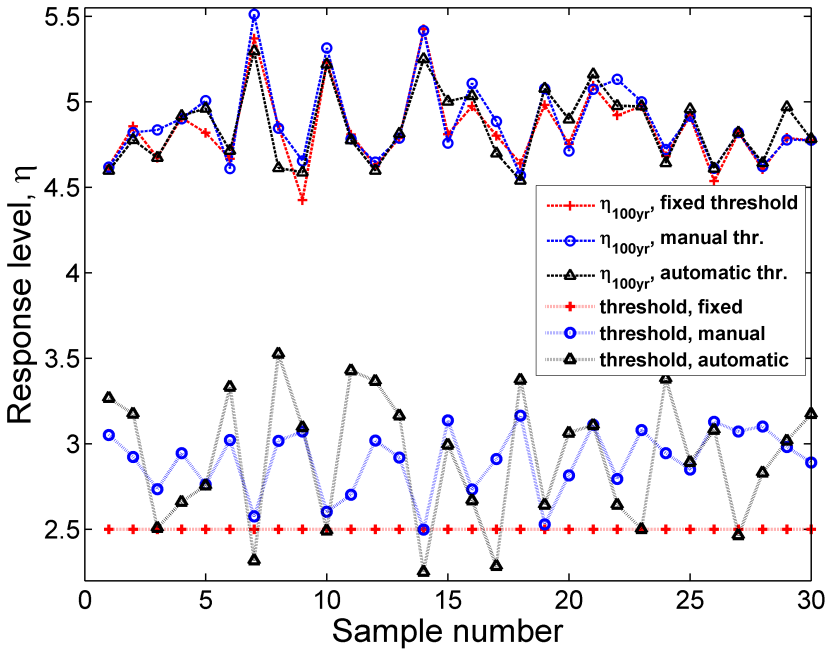

Figure 4.3: Comparison of performance of methods for identifying tail threshold levels.

code, BHawC. The total duration of the time series is $15000 \cdot (1/6) = 2500h \approx 104$ days, or slightly more than 3 months. The response quantities in focus are blade tip deflection u_B, and blade root bending moment in flapwise (along-wind) direction, M_y. For blade deflection in particular the extrapolation procedure is aimed at obtaining the response extremes which will allow for evaluating the possibility of blade-tower collision, therefore only the deflection at the moment of blade-tower passage is considered. This is done by extracting the peaks out of time history showing the distance between blade tip and tower surface (see Figure 4.4). For the sake of keeping the numbers more intuitive, the sign of the blade tip-tower distance is changed, and it is added to the value of blade-tower clearance at rest:

$$u_B = -d(\theta) + d_0 \tag{4.36}$$

where $d(\theta)$ is the blade tip-tower distance at any given rotor azimuth position θ, while d_0 is the blade-tower clearance for zero wind speed and stationary turbine. If the distance $d(\theta)$ is evaluated at the moment of blade-tower passage (i.e., $\theta = 180°$), the resulting value u_B gives the actual blade deflection at that moment. When the blade deflection equals the blade-tower clearance, $u_B = d_0$, a blade-tower collision event will occur. Using the representation in equation 4.36 will also allow for consequent comparisons between blade deflection during

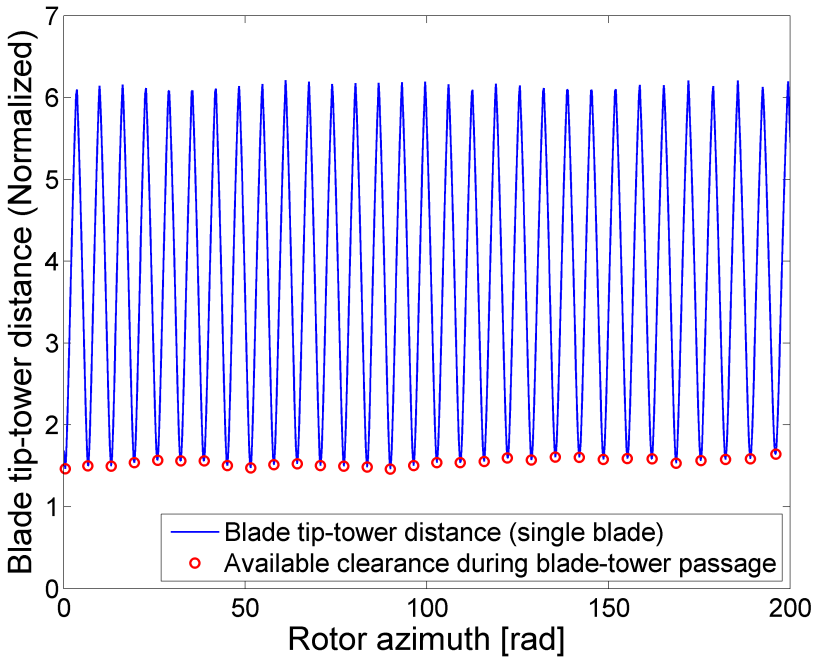

Figure 4.4: Blade-tower distance and blade-tower passage events.

blade-tower passage and the overall blade deflection measured regardless of rotor asimuth position.

The set of 15000 simulations is used to estimate extreme response values with 1-month recurrence periods, $u_{B,1month}$ and $M_{y,1month}$. First, the empirical cumulative distribution function of the response is evaluated from the time series by data ranking:

$$F_k = \frac{k}{n+1} \tag{4.37}$$

where F_k is cumulative distribution function of a point with rank k, and n is the total number of data points. In order to exclude correlation between peaks, and to avoid excessive data amounts, only the highest value from each of the 15000 time series is taken into account when determining the empirical exceedance probabilities. Confidence intervals are obtained by bootstrapping with 10000 bootstrap samples. Figure 4.5 shows a plot of the empirical exceedance probabilities obtained from data, together with the bootstrapped 1-month estimate and confidence intervals for u_B and M_y. In the following analysis, this response level with 1-month recurrence period is used as target reference for evaluating the performance of extrapolation methods. In order to avoid disclosure of company proprietary information, turbine responses have been normalized so that the values of u_B and M_y with estimated

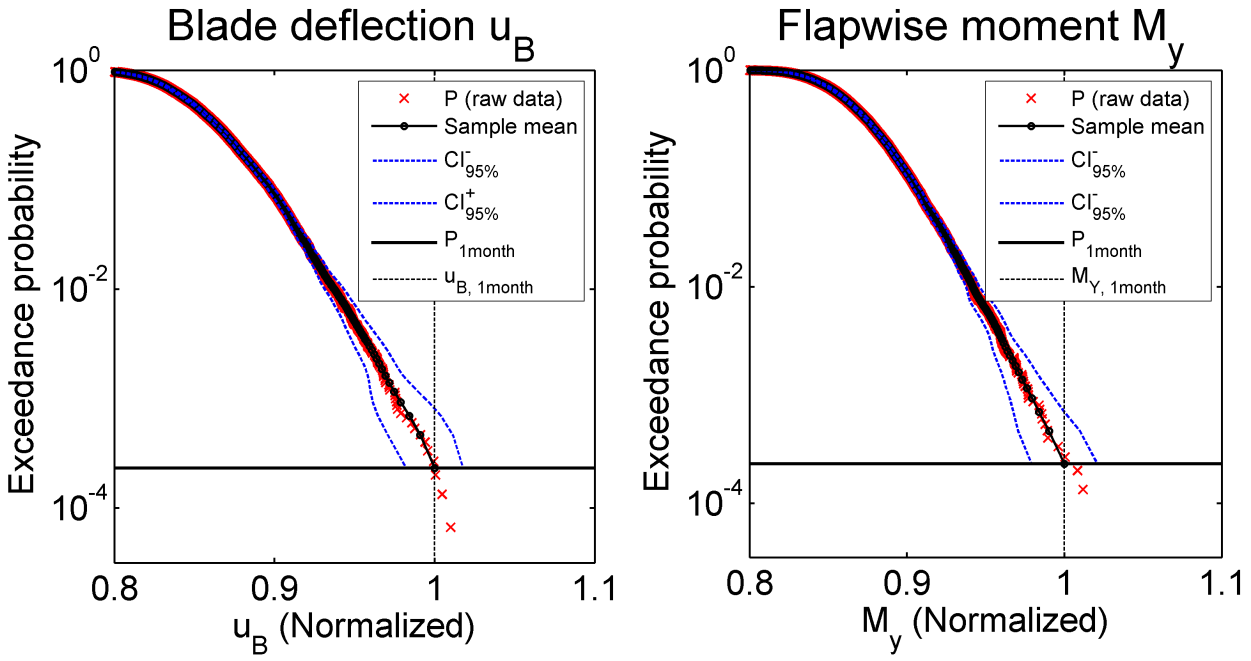

Figure 4.5: Exceedance probabilities for blade deflection and flapwise bending moment with condifence intervals determined by bootstrapping.

1-month recurrence period equal 1. Figures 4.6 and 4.7 show an example of an extrapolation based on 100 time series: the former shows a comparison of the distributions of independent peaks extracted using the three methods considered in this study, and the latter shows the extrapolation functions based on four different statistical distributions. It is visible that for large response values the three peak identification methods ACER, peaks-over-threshold, and block maxima result in very similar empirical distributions of peaks. This is a logical result as each of these three methods should serve the same purpose of identifying uncorrelated extreme events in a response history. The extrapolation functions, however, show a different picture, where the extrapolated response levels differ significantly depending on the choice of statistical distribution. This, as it will be seen later, has a visible effect on the accuracy of response predictions.

Results from estimation of 1-month extreme response levels

The variance of extrapolated response values is determined by carrying out a number of independent sample runs, with up to 100 time series included in each sample. For each of the sample runs, the estimated extreme blade deflection u_B and blade root flapwise moment M_y with 1-month recurrence period are determined as functions of the number of time series

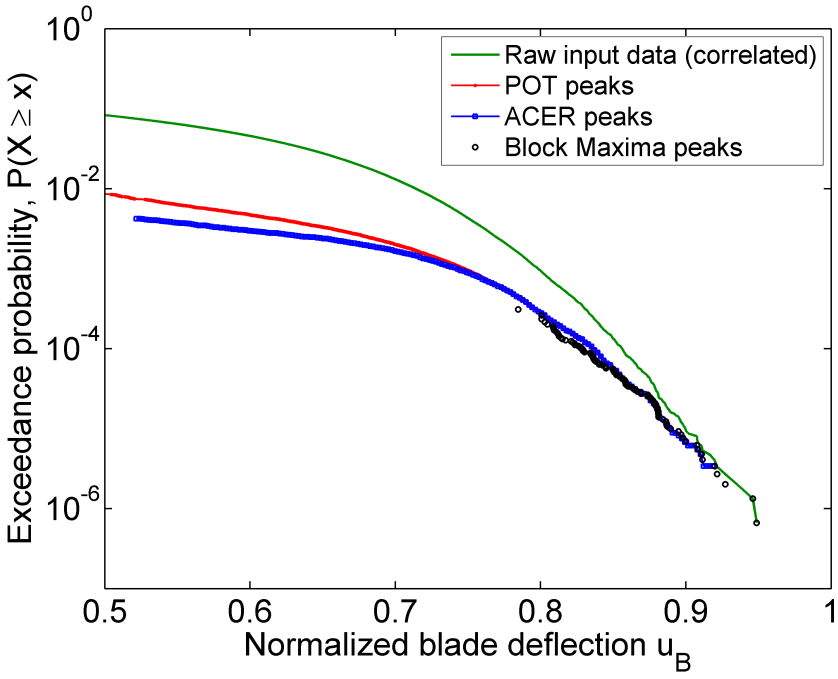

Figure 4.6: Exceedance probabilities for peak response events as determined by different peak extraction methods.

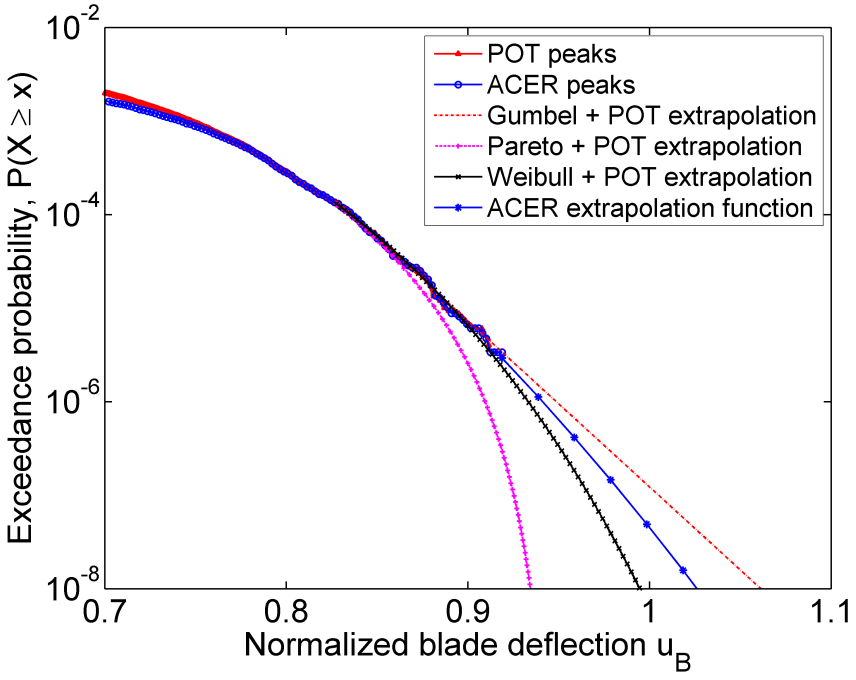

Figure 4.7: Comparison of extrapolation functions based on different statistical distributions.

used, starting with a minimum of 4 series and up to a maximum of 100 series. Three different peak extraction procedures are tested, and a number of statistical distributions is fit to the extremes obtained by each of the three methods:

- Average Conditional Exceedance Rates method (ACER) in its standard implementation with asymptotic double-exponential function fit to the response tail;
- Peaks Over Threshold method, with automatically defined threshold. Peaks are declustered by assuming a minimum separation between successive peaks of 7 data points. Note that for blade-tower passage events, which happen 3 times during a rotor revolution, this time separation roughly corresponds to two rotor revolutions. The following peak distributions are assumed:
  - Gumbel distribution;
  - Pareto distribution;
  - AEV function;
  - 3-parameter Weibull distribution;
- Block Maxima method with block length of 600s. The block length is equal to the length of time series, meaning that only one peak (global maximum) per series is extracted. All extracted peaks are thus independent. The following distributions are used:
  - Gumbel distribution;
  - Pareto distribution;
  - AEV function;
  - 3-parameter Weibull distribution;

Extreme blade deflection during blade-tower passage Figures 4.8 and 4.9 show respectively the mean and coefficient of variation of $u_{B,1month}$, obtained from 30 independent sample runs, and table 4.2 gives an overview of the performance of the procedures being assessed. The results for 1-month extreme blade deflection suggest that the most important factor influencing the accuracy of extrapolation is the choice of distribution. The means of the estimates obtained using the 4-parameter ACER function and the 3-parameter Weibull distribution are close to the reference value, while the Gumbel distribution tends to slightly overpredict, and the Pareto distribution to underpredict the response. Except for the case

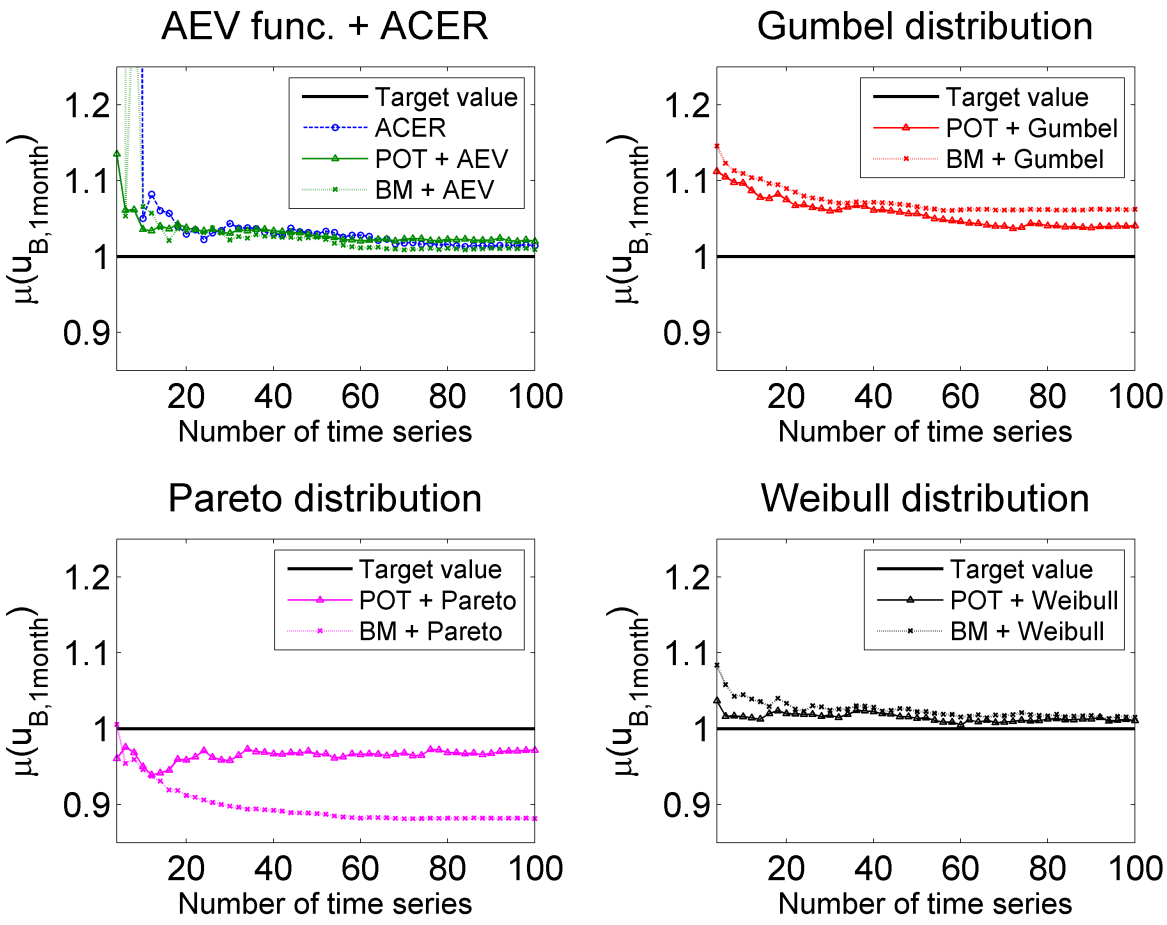

Figure 4.8: Comparison of accuracy of extrapolation methods in predicting extreme blade deflection.

Table 4.2: Estimates of extreme blade deflection with 1-month recurrence period.

Method	Average estimate (mean of 30 runs)	95% CI	Time series necessary for convergence
ACER	1.028	(0.939, 1.118)	42
POT, AEV func.	1.034	(0.944, 1.124)	12
POT, Gumbel dist.	1.068	(0.973, 1.164)	24
POT, Pareto dist.	0.970	(0.885, 1.055)	36
POT, Weibull dist.	1.020	(0.930, 1.109)	20
Block Max., AEV func.	1.026	(0.929, 1.124)	32
Block Max., Gumbel dist.	1.104	(1.019, 1.188)	12
Block Max., Pareto dist.	0.912	(0.824, 1.001)	20
Block Max., Weibull dist.	1.033	(0.950, 1.117)	20

when using Pareto distribution, the choice of peak extraction method seems to have less influence on the mean values of the estimates, however it influences the variance and hence the amount of data needed for convergence.

Results from estimation of 1-month extreme blade root bending moment Figures 4.10, 4.11, and table 4.3 show results from extrapolating the extreme blade root flapwise moments. The information that these calculations reveal is very similar to what was seen from extrapolating the blade deflection, with the difference that the flapwise response converges faster. This is expected because the blade deflection is considered only at the points where it is relevant for blade-tower collision, meaning that calculations with u_B are based on a far more limited data set compared to calculations with M_y. This is of course valid only for methods that use all data points, and not for the block maxima method, where only one data point per time series is used regardless of the type of response. The optimization procedure for fitting the AEV function to peaks extracted by POT and BM method showed numerical instabilities due to arithmetic overflow caused by the exponential function $q\exp(a(\eta - b)^c)$. Therefore calculations of blade flapwise moment with the AEV function were performed using normalized response time series, while all other calculations used the original time series, and normalization was performed directly on the obtained extrapolated estimates. Normalization reduced the values of the response with several orders of magnitude, which resulted in significantly improved stability of the optimization.

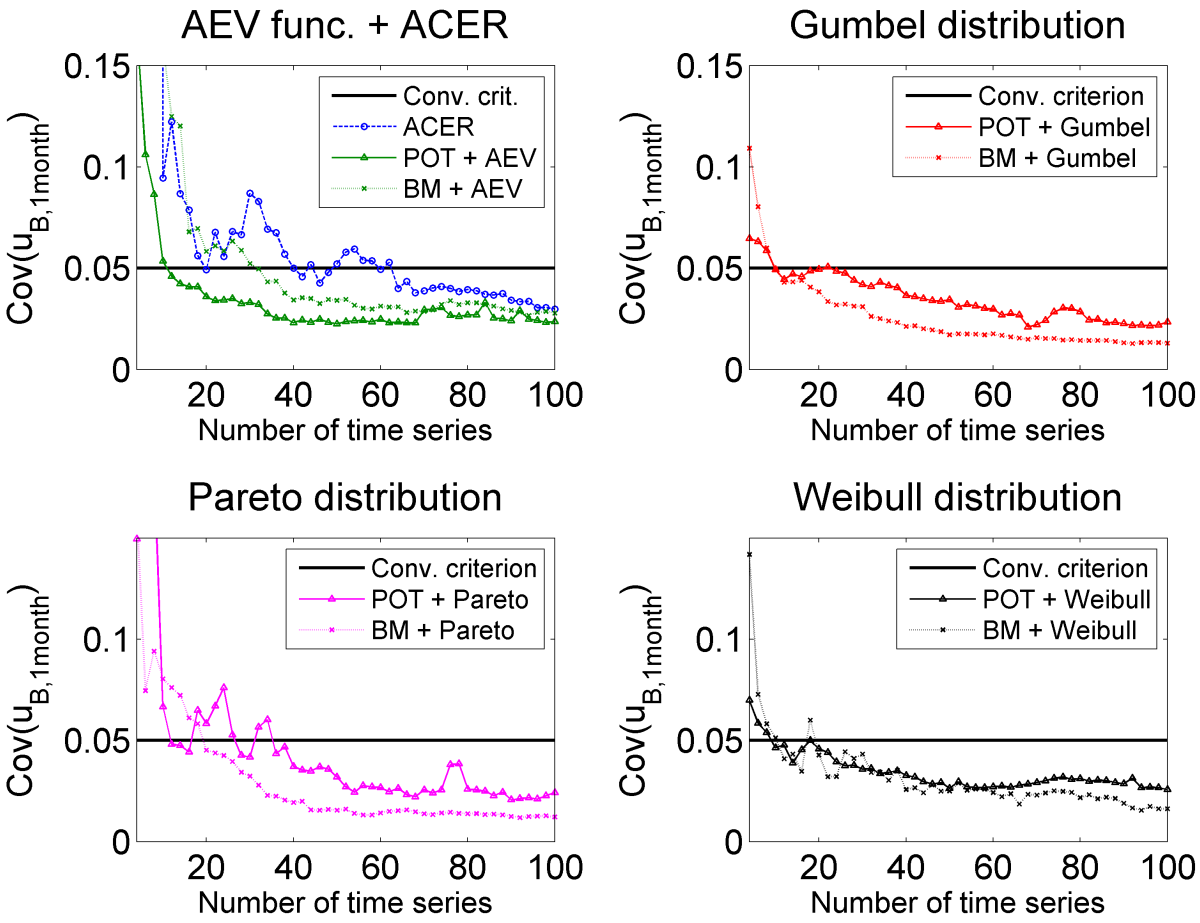

Figure 4.9: Rates of convergence of 1-month extreme blade deflection estimate for different extrapolation methods.

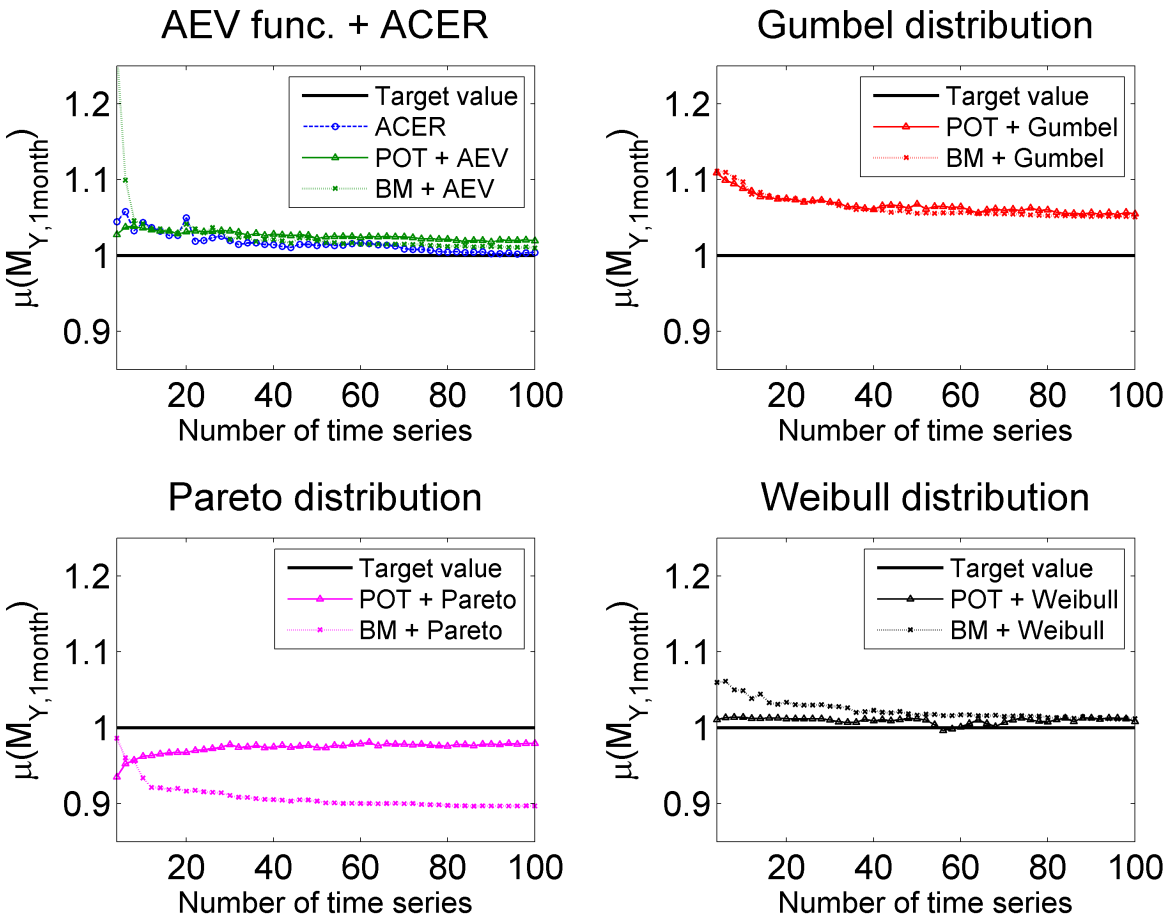

Figure 4.10: Comparison of accuracy of extrapolation methods in predicting extreme blade root flapwise moment.

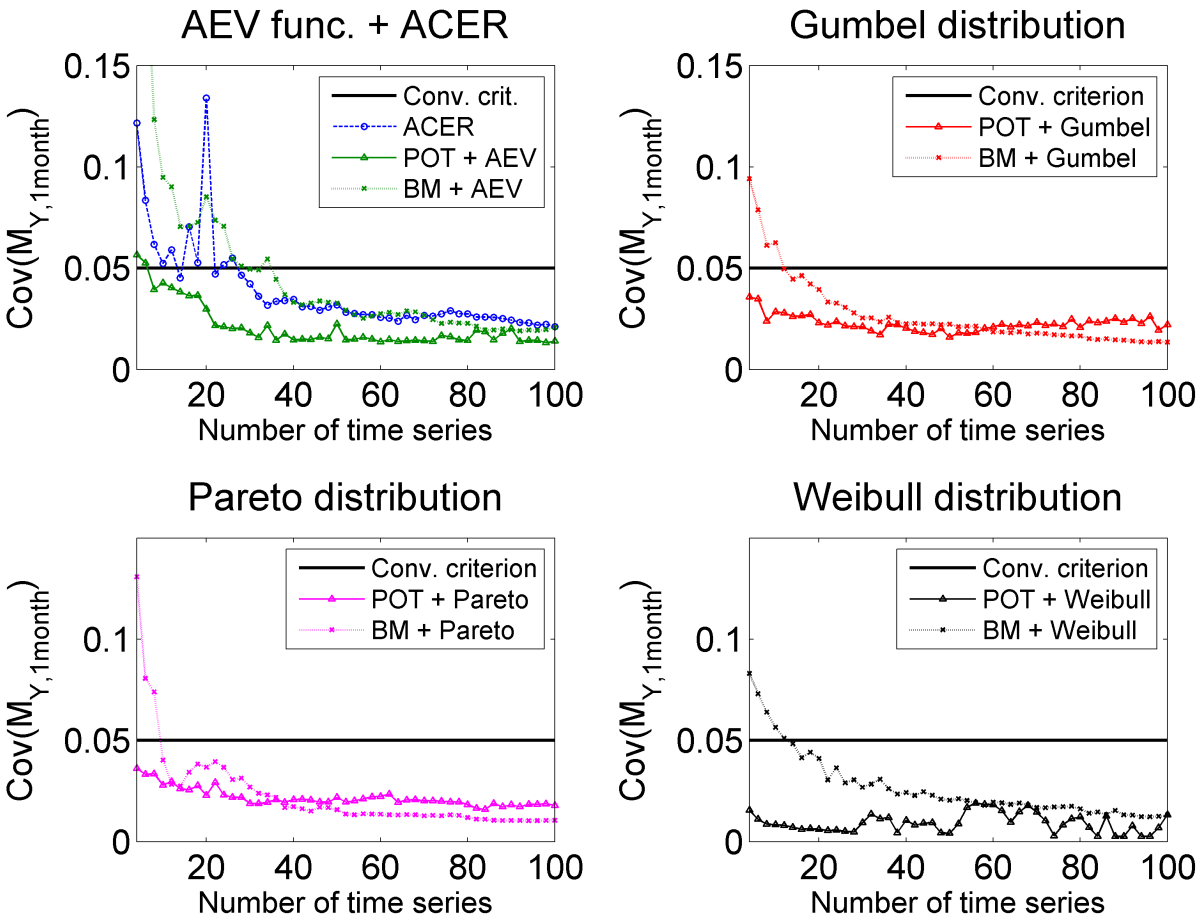

Figure 4.11: Rates of convergence of 1-month extreme blade root flapwise moment estimate for different extrapolation methods.

Table 4.3: Estimates of extreme blade root flapwise moment with 1-month recurrence period.

Method	Average estimate (mean of 30 runs)	95% CI	Time series necessary for convergence
ACER	1.026	(0.934, 1.117)	28
POT, AEV func.	1.038	(0.961, 1.115)	8
POT, Gumbel dist.	1.109	(1.039, 1.179)	4
POT, Pareto dist.	0.935	(0.864, 1.006)	4
POT, Weibull dist.	1.011	(0.980, 1.041)	4
Block Max., AEV func.	0.018	(0.931, 1.106)	36
Block Max., Gumbel dist.	1.081	(0.983, 1.178)	12
Block Max., Pareto dist.	0.934	(0.855, 1.013)	10
Block Max., Weibull dist.	1.044	(0.950, 1.139)	14

Convergence using the criteria from IEC61400-1 standard The data tables and graphs with convergence history shown above are based on the convergence criterion defined earlier, requiring that the width of the 95% confidence interval of the response at the target extrapolation level is not more than 20% of the estimate. The IEC standard 61400-1 defines a different convergence criterion, where the width of the 90% confidence interval at the 84% quantile of the response should be no more than 15% of the mean response with the same probability of occurrence. Figure 4.12 plots the convergence histories for extrapolating blade deflection based on the IEC convergence criterion. It is observed that the IEC criterion is satisfied in all calculation cases, for any amount of data used, including the calculations with the minimum amount of 4 data series.

4.1.3 Discussion on short-term load statistics

The results given in the previous section showed how different extrapolation methods can be used to predict wind turbine responses with a 1-month recurrence period. In general a well-confined estimate is obtained after using 3-4 hours of data, meaning that a reasonable prediction of the extreme response levels is possible by using time series covering approximately 0.5% or more of the target recurrence period.

Possibly the most clearly seen difference between the results shown in the previous section is the effect of distribution choice on the mean values of response estimates. The same trend is

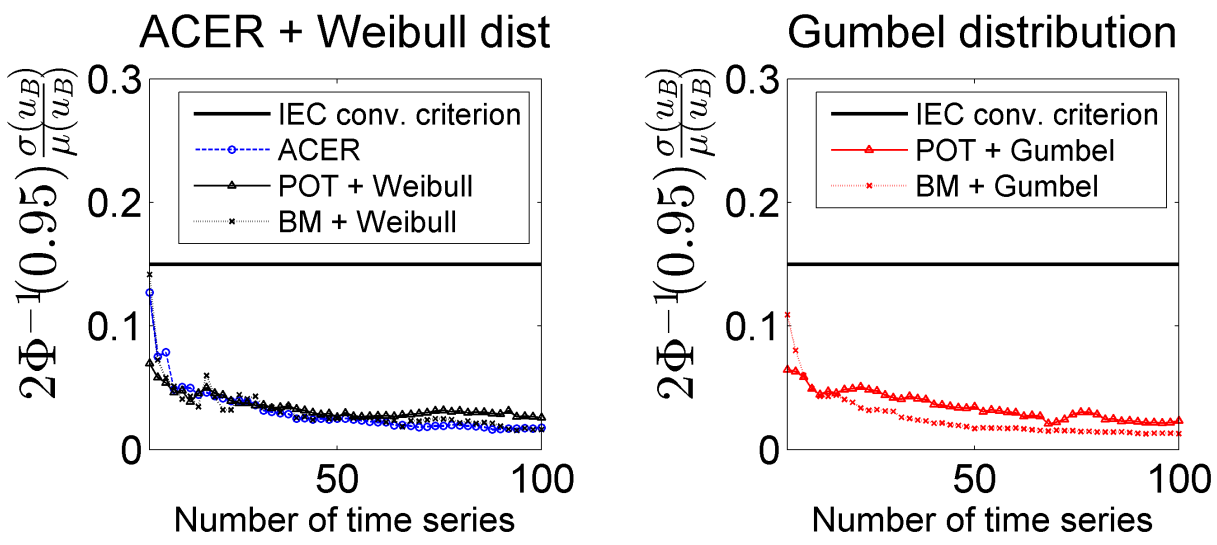

Figure 4.12: Rates of convergence of 1-month extreme blade deflection estimate; convergence criterion given by IEC61400-1.

visible both for blade deflection and flapwise moment: using the Weibull distribution and the ACER function results in better accuracy compared to using Gumbel or Pareto distribution, where fits using Gumbel distribution tend to overestimate, and the Pareto distribution tends to underestimate the response. Extrapolated values obtained using the Pareto distribution are the least accurate. The main cause for the low accuracy of the Pareto-distribution method can probably be found in the approach used to fit distribution parameters. In the present study, the Weibull, ACER, and Gumbel function parameters are all estimated by minimizing a least squares error, while the Pareto distribution parameters are based on the De Haan estimators. The De Haan estimators use simple algebraic functions and there is no error estimation. The poor performance of the extrapolation method which uses the Pareto distribution and De Haan estimators shows that using error minimization as a criterion for finding optimal distribution parameters should be preferred over closed-form approximate solutions without error control.

Extrapolation using the Gumbel distribution is the least accurate among the methods using least-squares fits for determining distribution parameters. The Gumbel distribution is represented by just two distribution parameters, which limits its flexibility. The existence of an additional parameter in the Weibull distribution function, and two additional parameters in the ACER function makes it possible to describe a larger family of statistical distributions and thus allows to model a wider range of responses with sufficient accuracy. However, using

even more parameters will not necessarily lead to further improvement of the results, because the presence of too many parameters can easily lead to overfitting. In fact, the impression is that the higher variance showed by the ACER method is mostly due to the use of four parameters, which occasionally leads to overfittng and less stable optimisation compared to other methods. However, it is not implied that the ACER method is less useful than other methods. On the contrary, it is one of the procedures that seem best suited for extrapolating wind turbine response, because the slower rate of convergence is compensated with higher accuracy and very good ability for extrapolating to very low probabilities of exceedance.

Three different methods for extracting independent response peaks were tested. The results indicate that as long as correlation between events is properly accounted for, all of these peak identification methods can be successfully used as part of an extrapolation procedure. As Figure 4.6 clearly shows, at high response levels the exceedance probabilities predicted by the ACER, peaks over threshold, and block maxima methods are almost identical. The extrapolated response levels are not strongly influenced by the peak extraction method either, with the exception of when using Pareto distribution.

The convergence when using the Block Maxima method is slower compared to other methods, which is a consequence of using less data points. The difference is especially visible for blade root flapwise moment (Figure 4.11). This has importance for standard design calculations, where not more than 15 time series per wind speed are usually used. For such cases it can be recommended to use methods which take more points into account, as for example peaks-over-threshold, ACER, or to reduce the size of the blocks, provided that the peaks extracted will still be uncorrelated.

During the process of analyzing data it was observed that tail threshold level has a certain degree of influence on fits performed using the Gumbel and Pareto distributions. Extrapolations using the Weibull distribution or ACER function are not that strongly influenced and can accommodate wider range of threshold levels. Again, the apparent reason for this behaviour is the greater flexibility of using distribution functions with more than two parameters, which allows capturing some of the sub-asymptotic behaviour thus reducing the dependence on the tail threshold value. Nevertheless, the automatic procedure for tail threshold identification defined in this chapter has worked well and it is not expected that threshold levels have any significant contribution to the uncertainties in the observed results, also for the fits with Gumbel and Pareto distributions.

Judging from the observed calculation results and from experiences gained during the analysis process, some of the factors that influence load extrapolation results have been

ranked according to their importance for successful implementation of a statistical load extrapolation procedure. Starting with the most critical factors, the list is the following:

1) Choice of statistical distribution and fitting method for approximating the tail response

2) Providing sufficient amount of data points

3) Data censoring: determining sample confidence intervals and excluding outliers

4) Choice of tail threshold level

5) Choice of method for decorrelation of data

Based on the criteria above, the recommendation is that either the Weibull distribution or ACER function are used as extrapolation functions, while any of the discussed peak extrapolation methods can be used, as long as enough data are available to ensure convergence of the results. If the asymptotic extreme value function $q\exp(a(\eta-b)^c)$ is chosen as extrapolation method, it is probably best to use the ACER method to extract average conditional exceedance rates from the data, as the ACER method and the AEV function are developed in connection with each other which makes the implementation easier and possibly more robust. Using the Gumbel distribution, while possibly leading to slight overprediction of the results, can still be a viable approach due to the great simplicity and high robustness. The use of Pareto distribution fit by De Haan estimators [13] is however not recommended, as it consistently provides non-conservative estimates. There are sources in literature (Naess, 2002 [54]) showing that other parameter estimation methods for the Pareto distribution might perform better than the De Haan estimators.

The convergence criteria that have been specified for this study are based on the width of the confidence interval evaluated at the target response level. This is a stricter approach compared to the recommendations in the IEC61400-1 standard [37], which state a convergence criterion based on the width of the confidence intervals at the 84% quantile of the response using a 10-minute reference period. A 84% quantile means a load level with probability of exceedance of approximately 1/6 over a 10-minute period, and it is expected that a number of data points exceeding that level will exist in a time history with 150 minutes duration (the minimum required by IEC) or longer. It is therefore not surprising that in the presented calculations the IEC convergence criterion was satisfied in all cases, even when using only 4 time series. However, despite that this convergence criterion is satisfied, the uncertainty at higher quantile levels such as the target extrapolation level can be quite high

as seen on Figures 4.8 and 4.9. The conclusion is that a stricter convergence criterion may be needed which should assess the model variance at higher quantile levels, preferably at the target extrapolation level.

4.2 Long-term statistical distribution of loads

4.2.1 Analysis methods

Modelling environmental conditions

Wind turbine rotors are subject to a 3-dimensional wind field randomly varying in space and time. For the purpose of wind turbine response simulations this is normally modelled as a random stochastic field superimposed on a corresponding set of mean values. The mean values of the field are usually described as function of a number of parameters such as the mean wind speed at hub height U_{hub}, horizontal and vertical wind shear exponents α_h and α_v, flow inclination angles φ_h and φ_v, etc. The random component of the wind field is given in terms of the standard deviation of the wind velocity σ_U, which is both spatially and time-correlated. In most practical cases the mean wind field is described as stationary within a reference period (typically 10 minutes).

In the present thesis three parameters: mean wind velocity at hub height, U_{hub}, turbulence intensity $I = \sigma_{U_{hub}}/U_{hub}$ and vertical wind shear exponent α_v are modelled with their statistical distribution. The three parameters mentioned are considered as the most significant environmental factors influencing the blade deflection, whose modelling is the objective of the present analysis. The IEC61400-1 [37] gives reference statistical distributions for wind speed and turbulence intensity. These standard values are used throughout the analysis in order to make it as general as possible. The standard however does not provide a reference distribution for the wind shear exponent. Therefore, the statistical distribution of wind shear is modelled based on site measurements from the Hovsøre turbine test site located in western Denmark.

All quantities under consideration are strongly correlated since they are describing properties of the same stochastic process. Correlation is taken into account by modelling the parameters as conditionally dependent on each other. The transformation between an independent set of random samples as generated by a computer random number generator, and a conditionally dependent sample is done using the Rosenblatt transformation [70]. For the purpose of using a Rosenblatt transformation the parameters are ranked according to

their assumed level of dependence. The mean wind speed at hub height is considered as the governing parameter and is modelled as statistically independent from other parameters. Turbulence intensity is modelled as conditionally dependent on the hub wind speed, and the wind shear exponent is considered dependent both on wind speed and turbulence intensity.

Statistical model for mean wind speed at hub height Wind conditions in IEC61400-1 are divided in classes according to the mean wind speed and turbulence. Here, a wind climate class 2B is used, which means an annual average wind speed of $U_{ave} = 8.5m/s$, and reference characteristic turbulence intensity at $15m/s$ wind speed of $I_{ref} = 0.14$. The statistical distribution of 10-minute average wind speeds is assumed to be Weibull distribution with scale parameter $A = 9.5912$ and shape parameter $k = 2$.

Model of turbulence intensity According to IEC61400-1, standard deviation of wind velocity is modelled as lognormally distributed and conditionally dependent on the 10-minute average wind speed, with

$$\begin{aligned} &E\langle\sigma_U|U_{hub}\rangle = I_{ref}(0.75U_{hub} + c); \qquad c = 3.8m/s \\ &Var\langle\sigma_U|U_{hub}\rangle = \left(I_{ref} \cdot 1.4m/s\right)^2 \end{aligned} \tag{4.38}$$

Wind shear model An exponential wind shear law is assumed, where the change of wind speed with height z is determined by the wind shear exponent α:

$$U(z) = U_{hub}\left(\frac{z}{z_{hub}}\right)^{\alpha} \tag{4.39}$$

The same law is assumed by IEC61400-1, however the standard does not provide reference for the statistical distribution of the wind shear exponent α. Therefore, the distribution of α is modelled here based on weather mast measurements data from the Hovsøre wind turbine test site. A total of about 5 months of data are included in the study, covering the period between 21-04-2012 and 18-09-2012. Wind shear exponent based on ten-minute reference period is determined by solving equation 4.39 for the ten-minute average wind speeds measured at heights of 29, 59 and 89m. In addition to the wind shear exponent, the average wind speed and turbulence intensity at hub height are taken because the modelling assumption is that wind shear will be conditionally dependent on these quantities. Only data coming from wind direction sectors free of turbine wake effects are chosen, meaning that the dataset is restricted to approximately the westerly half of the possible directions, 180° to 360° azimuth angle.

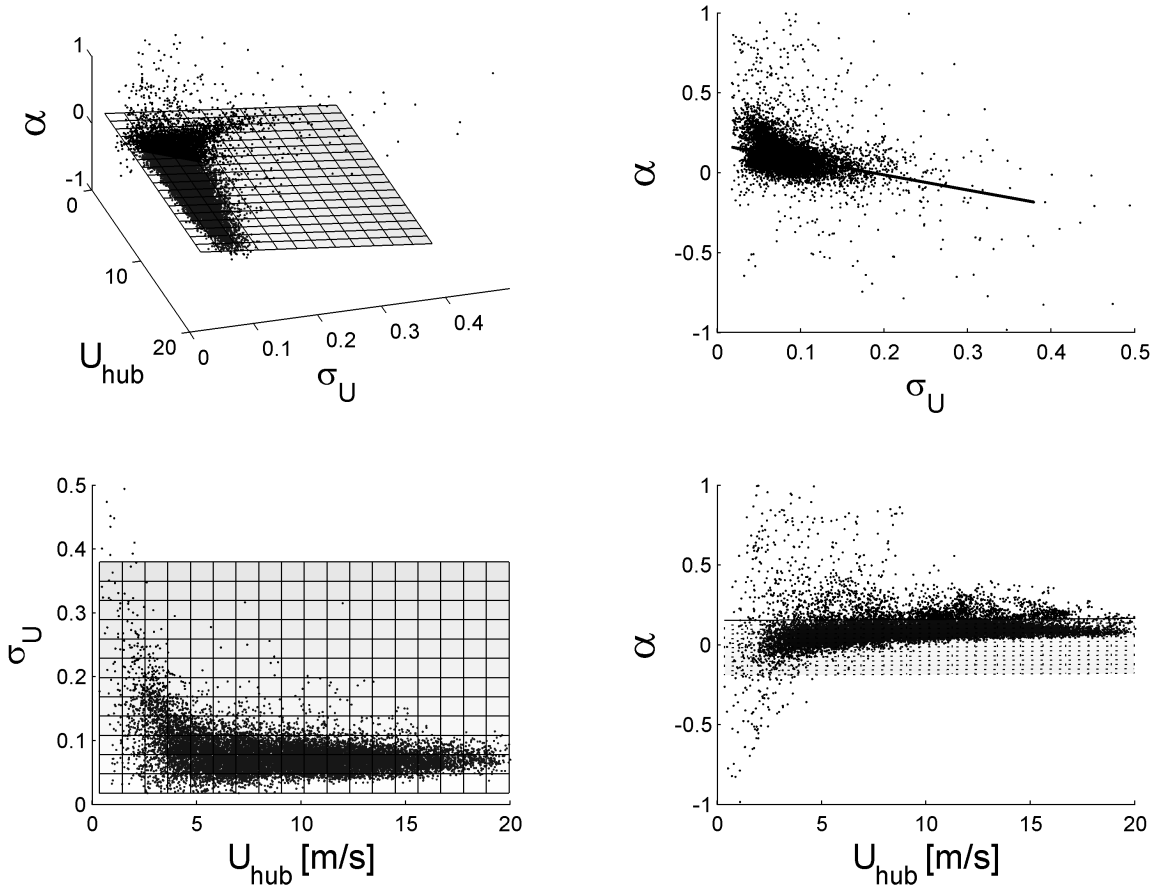

Figure 4.13: Wind shear data with reference planar surface $\bar{\alpha}$.

The statistical distribution of the wind shear exponent is modelled using an expression conditioned on the wind speed and turbulence:

$$\alpha = \bar{\alpha}(U, \sigma_U) + \varepsilon_\alpha \tag{4.40}$$

where $\bar{\alpha}$ is a reference mean value, and ε_α is the deviation from the reference. The reference value $\bar{\alpha}$ is considered to be a linear, deterministic function of the wind speed and turbulence, and it is determined by a least-squares fit to the data, see Figure 4.13. The deviation ε_α is randomly varying, and as visible from Figure 4.13, the distribution of the scatter is dependent on the wind velocity. In order to model such behaviour, the data are split into bins according to the 10-minute average wind velocity, and a separate statistical distribution for ε_α is fit to each of the data bins. The distribution of ε_α for most wind speeds is heavily-tailed, with a slight skewness, but very high kurtosis, see for example Figure 4.14, where the histogram of ε_α for $U = 2.5 \pm 0.5 m/s$ is shown. None of the more widely known

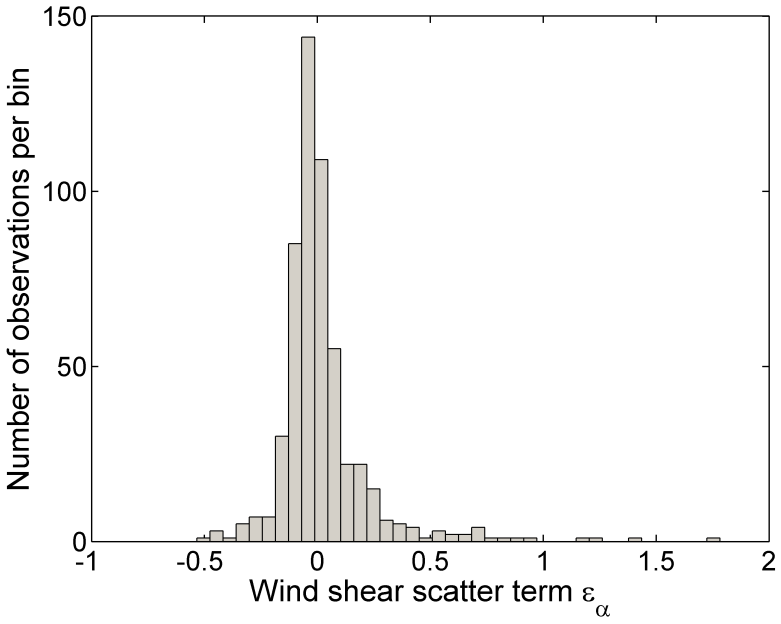

Figure 4.14: Histogram of wind shear deviation term ε_α for $U = 2.5 \pm 0.5m/s$.

continuous parametric statistical distributions (e.g., Gumbel, Weibull, Normal) are able to represent such a behaviour with sufficient accuracy. It is therefore chosen to approximate the distribution of ε_α using the Cornish-Fisher asymptotic expansion (Fisher, 1950 [24], Abramowitz & Stegun, [2]), which approximates a given distribution based on its first several statistical moments. The Cornish-Fisher expansion is preferred over other known expansions such as the Edgeworth expansion (see Wallace, 1958 [89], Abramowitz & Stegun, 1972 [2]) because it gives the inverse cumulative distribution function, $x = F^{-1}(F_x)$, which is more convenient for random sample generation and application to reliability methods, where the inverse distribution is required. Using the inverse cumulative distribution also allows for performing a fitting procedure using residual function based on the variate (i.e., ε_α in the present case). The variate (or its logarithm) usually has linearly spaced values which makes obtaining good fits in the tails of the distribution easier than with procedures where the logarithmically-behaving cumulative distribution $F(x)$ is used. The residual function used for this task is

$$r = \sum_{j=1}^{N} \left[\hat{\varepsilon}_{\alpha,j} - F_{\mathcal{A}}^{-1}(\hat{F}_j, \boldsymbol{\gamma}) \right]^2 \tag{4.41}$$

where $F_{\mathcal{A}}^{-1}(F, \boldsymbol{\gamma})$ is the inverse asymptotic probability distribution given by the Cornish-Fisher expansion, $\boldsymbol{\gamma} = [\gamma_1, \gamma_2, \ldots, \gamma_k]$ are the first k statistical moments acting as parameters

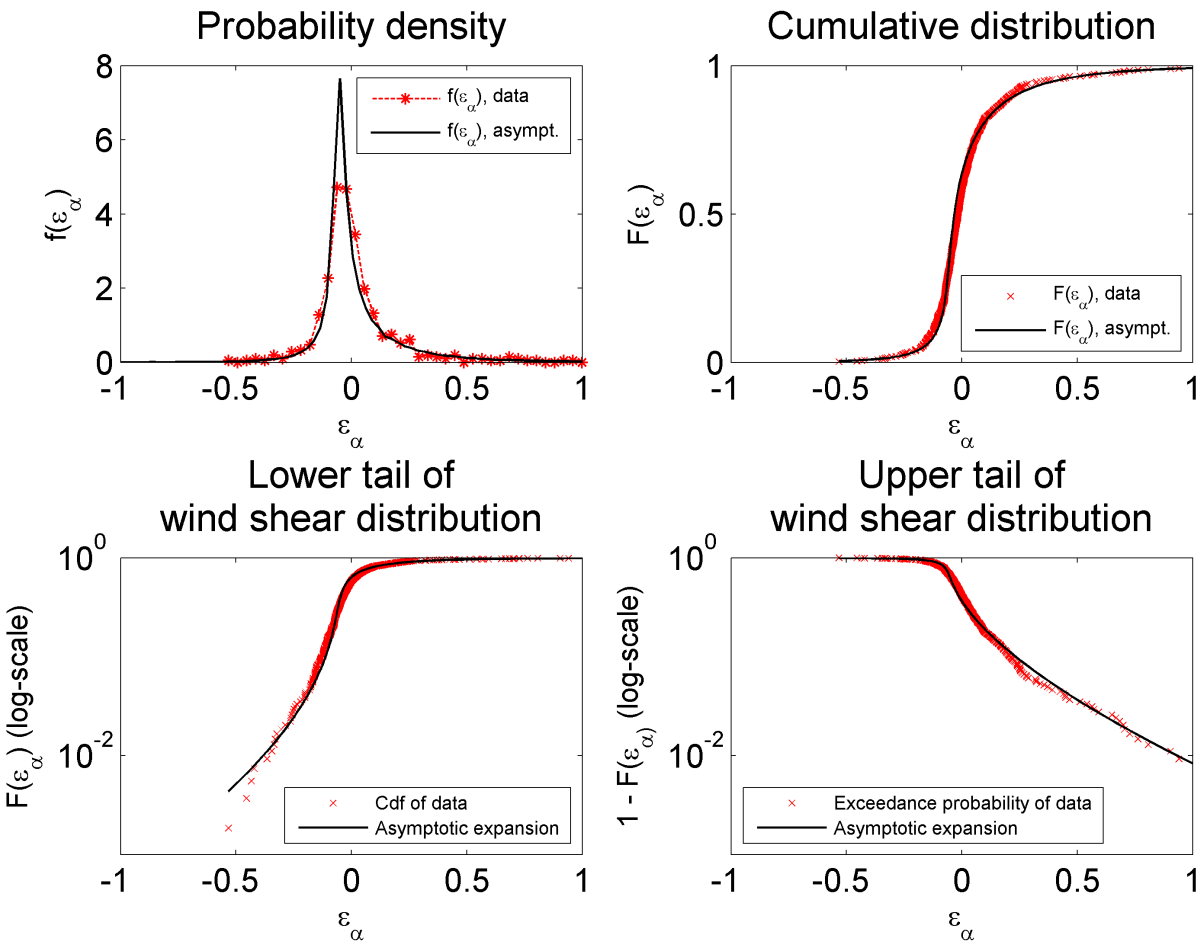

Figure 4.15: Wind shear distribution fit to data using asymptotic expansion.

of the expansion, $\hat{\varepsilon}_{\alpha,j}$, $j = 1 \ldots N$ is the set of observations, and $\hat{F}_j = P(X \leq \hat{x}_j)$ is the empirically observed probability of the variate not exceeding the value of $\hat{\varepsilon}_{\alpha,j}$. In the present analysis the asymptotic expansion is based on the first four statistical moments of data (mean, standard deviation, skewness, kurtosis). The wind speed range is divided into bins of 1m/s width, and a separate fit is carried out in each bin. Due to lack of data, the parameters of the asymptotic expansion for wind speeds higher than 22m/s could not be determined, therefore for $U_{hub} > 22m/s$ the model assumes the same distribution parameters as for $U_{hub} = 22m/s$. Figure 4.15 shows an example of the asymptotic expansion of ε_α fit to data for the wind speed bin $U = 2.5 \pm 0.5m/s$. Finally, Figure 4.16 shows a comparison between wind field data from observations and a wind field generated from random samples conditioned by equations 4.38, 4.39, and 4.40.

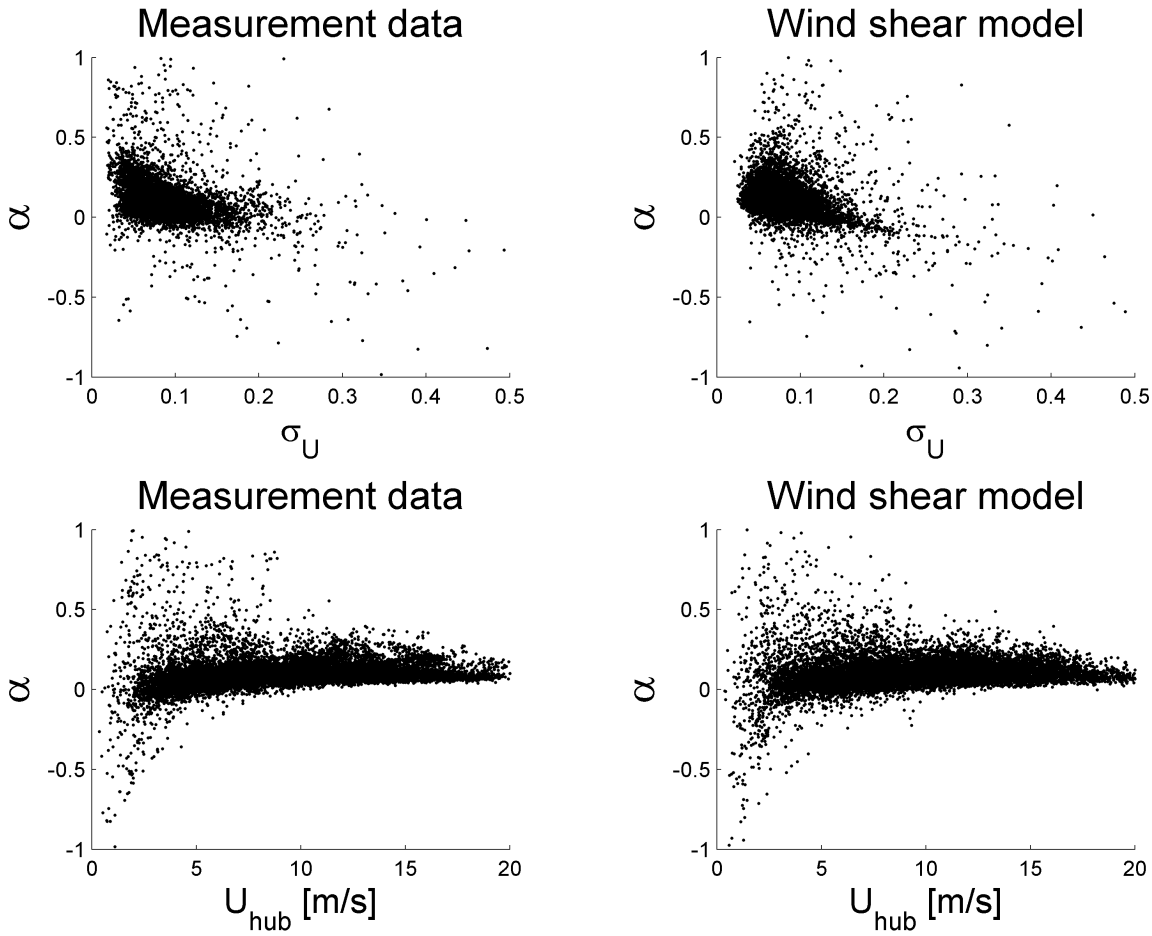

Figure 4.16: Comparison of a randomly generated wind field to measurement data.

Modelling variations in blade stiffness

The main quantity of interest in the present analysis is the blade tip deflection. Obtaining deflections does not require detailed stress and strength analysis of the blade interior, and if only tip deflection is sought, the exact displacement field in the rest of the blade has no great importance either. Therefore a very simple, single-parameter model for the variation in blade structural performance is used, where scatter in blade properties is taken into account by multiplying the blade mean stiffness to a random factor. According to Euler-Bernouli elastic beam theory, the tip deflection of a cantilever beam with length L loaded by a distributed force q is found by

$$w(x) = \frac{qx^2(6L^2 - 4Lx + x^2)}{24EI} \tag{4.42}$$

where L is the beam length, E is the stiffness modulus, and I is the area moment of inertia of the beam cross section. If one wants to change the resulting tip deflection by a compliance factor p, it will be sufficient to multiply the product EI by $1/p$. Therefore, modelling the variations of blade tip deflection due to scatter in stiffness comprises of representing the stiffness factor $1/p$ as a random variable. Here it is assumed that $1/p$ is Lognormally distributed, with a mean value of 1 and coefficient of variation of 0.03. This choice of variance represents a typical value for state-of-the-art, industrially produced large wind turbine blades. Introducing the stiffness adjustment factor in the actual structural model is done by multiplying the factor EI with $1/p$ in the structural properties input to the dynamic simulation program, where the structure of the blade is represented by beam elements.

Long-term statistical distribution of wind turbine response

The problem of obtaining short-term statistical distribution of loads was discussed earlier in this chapter. Based on the findings from the study on short-term distribution, the approach chosen is to use the peaks-over-threshold (POT) method for identifying extreme load events, and to fit a 3-parameter Weibull distribution to the extremes. Since the ultimate goal of the analysis is to evaluate the possibility of blade-tower collision, only the deflection at the moment of blade-tower passage has to be considered. This is done by extracting the points in time history where blade-tower passage events occur. Further analysis including peak extraction is performed on the new time series, formed by taking the distances between blade tip and tower surface observed during consecutive blade-tower passages events. The extrapolation procedure will result in what IEC61400-1 [37] defines as the local distribution of the response, denoted by F_{local}. The short-term distribution is obtained by transforming

the local distribution to match the desired short-term observation period, T, by

$$F_{short-term}(s\big|\mathbf{X};T) = F_{local}(s\big|\mathbf{X};T)^{n(\mathbf{X})} \tag{4.43}$$

where $\mathbf{X}$ is the vector of input variables that define the statistical distribution of the extremes, T is the observation time period, s is the magnitude of the extreme response, and $n(\mathbf{X})$ is the number of independent local extreme values for one time period T.

Long-term statistical distribution is obtained by integrating the short-term load distribution over the domain of input variables, $\mathbf{X} = [X_1, X_2, \dots X_m]$:

$$\begin{aligned} F_{long-term}(s\big|T) &= \int_{\mathbf{X}} F_{local}(s\big|\mathbf{X},T)^{n(\mathbf{X})}\mathbf{f}(\mathbf{X})\mathrm{d}\mathbf{X} \\ &= \int_{X_1}\dots\int_{X_m} F_{local}(s\big|\mathbf{X},T)^{n(\mathbf{X})} f(X_1)\dots f(X_m)\mathrm{d}X_1\dots\mathrm{d}X_m \end{aligned} \tag{4.44}$$

Input variables such as wind speed, turbulence etc., are often represented as discrete values, and in such cases equation 4.44 will read

$$F_{long-term}(s\big|T) = \sum_{j_1=1}^{k_1}\dots\sum_{j_m=1}^{k_m}\left(F_{short-term}(s\big|\mathbf{X}_J,T)\prod_{i=1}^{m}p_i\right) \tag{4.45}$$

where

m is the number of dimensions in $\mathbf{X}$;

$\mathbf{X}_J = [X_{1,j_1}, X_{2,j_2}, \dots, X_{m,j_m}]$ is the vector of input variables, where the i^{th} component of the vector takes discrete values with index j_i;

$p_i = f_{short-term}(s\big|\mathbf{X}_J,T)\Delta X_i$, $i = 1,\dots m$.

In principle, the integration approach is useful when the number of dimensions considered is limited to just a few, as for example in IEC61400-1, where the wind speed is the only component of $\mathbf{X}$ and equation 4.45 is used in its 1-dimensional form. For high number of dimensions numerical integration becomes impractical as the number of the required local distribution fits grows in the manner of a geometric progression. An alternative approach which is not sensitive to dimensionality is the Monte Carlo simulation, where a set of random samples are drawn directly from the long-term distribution of $\mathbf{X}$. Using the same mathematical approaches as for the short-term distribution, independent peaks are extracted from the response history obtained by the MC sample, and an extrapolation function is fit to the extrema. This extrapolation function will be an estimate of the long-term distribution $F_{long-term}(s\big|T)$.

4.2.2 Long-term distribution of blade deflection

The long-term distribution of the extremes of blade deflection during blade-tower passage events is obtained based on statistical extrapolation. Peaks-over-threshold method as described in Section 4.1 is used to extract the peak events, and a 3-parameter Weibull distribution is fit to the peaks. A total of four simulation data sets are generated with the purpose of solving the following three problems:

1) A 1-dimensional problem where the only variable is the wind speed, and the turbulence is taken as its characteristic, 90% quantile value. This is the standard approach described in IEC61400-1.

2) A 2-dimensional problem, where the wind speed and turbulence are random variables. The turbulence is conditionally dependent on the wind speed

3) A 4-dimensional problem, with random variables being the wind speed, turbulence, wind shear exponent, and blade stiffness multiplication factor. The turbulence is conditionally dependent on the wind speed, while the wind shear is conditionally dependent on both wind speed and turbulence.

The long-term distribution for problem 1) is obtained using 1-D integration of short-term distributions over simulated wind speeds. For problem 2) both the integration procedure given by equation 4.45 and the Monte Carlo approach are used, meaning that two data sets have been generated as the two methods require different simulations - the 2-D integration method is done over wind speeds and turbulence intensities with discrete values, while the Monte Carlo simulation uses a completely random set of input parameters. Using a 4-dimensional integration to solve problem 3) would require a very large amount of data, even if using sparse grids, therefore only the Monte Carlo approach is used for obtaining a solution. The amount and type of simulations used for generating the four discussed sets of data are given in table 4.4. A set of Monte Carlo simulations with random wind speed and turbulence covering a 5-year period is available as a reference, and the first three analyses presented here are designed in a way that it is possible to compare them with the reference 5-year data.

Table 4.4: Description of dynamic simulation data sets.

Set Nr.	Number of variables	Variable description	Simulations per bin	Number of bins	Total simulations
1	1	U_{hub}	100	22	2200
2	2	U_{hub}, σ_U	100	$12 \cdot 9 = 108$	10800
3	2	U_{hub}, σ_U	-	none (MC sim.)	10000
4	4	U_{hub}, σ_U, α, p_{stiff}	-	none (MC sim.)	10000

Long-term distribution from simulations with one and two input stochastic variables

Figure 4.17 a) shows a comparison between the long-term distributions obtained by extrapolation of 1-D and 2-D data sets. The 5 years of Monte Carlo simulations span over a short period compared to the typical target annual failure probabilities for which the structures are designed, but nevertheless they provide a valuable reference covering a significant part of the long-term distribution of the response. For the 1-D data set the integration of the short-term probability distributions is done only over the wind speed range, but for all simulations the characteristic 90% quantile of the turbulence I_{90} is used, which should result in loads similar to the ones obtained with randomly varying turbulence. Indeed, as visible from Figure 4.17 a), the long-term distributions from 1-D integration and 2-D integration do not differ from each other substantially and in general they both agree with the 5-year simulations. Despite that the 2-D integration takes the variation in turbulence into account and uses more data, it seems that the 1-D integration has a slight advantage over the 2-D integration. This can be attributed to the higher uncertainty associated with using much more (108 versus 22) separate extrapolation procedures to determine the short-terms statistical distributions for each wind and turbulence bin. The best agreement between 5-year simulations and extrapolated response however is achieved by the extrapolation based on the 10000-series Monte Carlo sample, Figure 4.17 b). In this case, the long-term distribution is described by a single extrapolation function, which itself is fit to a large amount of data, which reduces the uncertainties compared to the other methods. The difference between the peaks from 5-year data and the 10000 MC simulations for high probabilities of exceedance which can be seen on Figure 4.17 b) exists because all independent peaks from the 10000 time series are used, while for the 5-year data set only the global extrema from each time

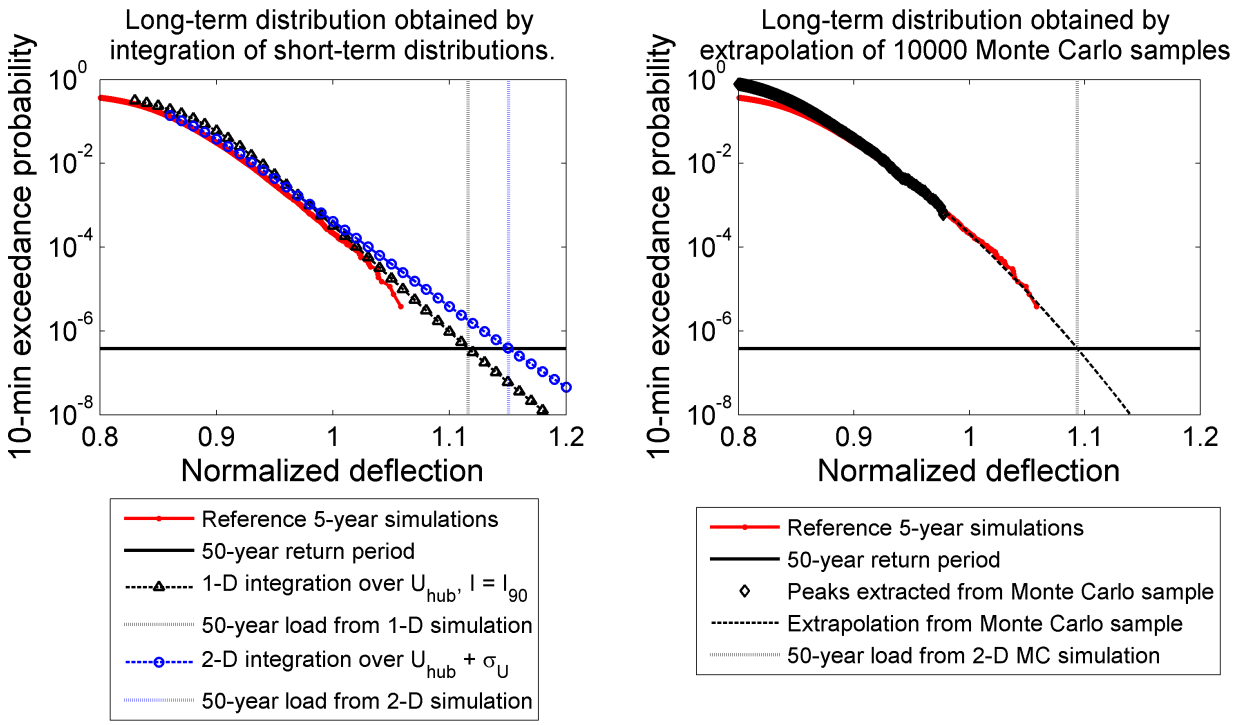

Figure 4.17: Extrapolated long-term distributions compared to 5-year reference Monte Carlo analysis.

series is used. For lower probabilities of exceedance however this difference vanishes. The method of directly simulating a sample from the long-term distribution of input parameters seems to come with the least uncertainty as only a single extrapolation is carried out based on a large amount of data (see the sampling history of the MC simulation on Figure 4.18). The method however has the disadvantage of being less flexible compared to the bin-wise integration procedure, which, once the short-term distributions have been obtained, can be used to model different types of long-term distributions.

Long-term distribution from simulations with four input stochastic variables

The long-term distribution of the turbine response obtained by Monte Carlo integration over wind speed, turbulence, wind shear and blade compliance is shown on Figure 4.19. The figure makes a comparison between the response based on 2-dimensional and 4-dimensional random input. Including a random distribution for the wind shear and the blade compliance has a visible effect on the predicted extreme response, with about 9% increase in the expected response level with 50-year return period. The two responses are not directly comparable,

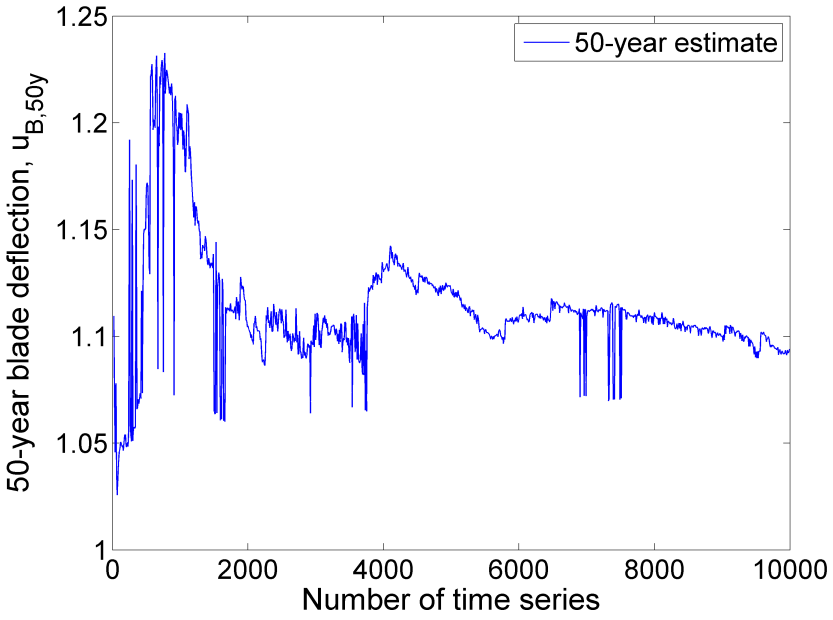

Figure 4.18: Sampling history showing the 50-year extreme deflection estimate as function of number of Monte Carlo samples.

because the 2-dimensional calculation uses a constant wind shear exponent of $\alpha = 0.2$ as recommended by the IEC 61400-1 standard, while the wind shear in the 4-dimensional calculation is determined by the random wind shear model defined in Section 4.2.1, where the mean shear exponent, as visible from Figure 4.16, is between 0 and 0.1. Nevertheless, the observed difference in exceedance probabilities shows that it is important to model the wind shear with its actual distribution or a characteristic value rather than the generic value fixed by the standard. This is especially valid for responses which are sensitive to wind shear as for example blade loads and blade deflection.

Blade-tower distance vs blade deflection

All calculations of blade deflection shown in this study until now are made with the purpose of obtaining a long-term statistical distribution of the blade response that can be used for evaluating the probability of blade-tower collision. Therefore only the deflection during blade-tower passage events is taken into account. The importance of this decision is shown on Figure 4.20, where the maximum observed blade deflection regardless of rotor asimuth, and the blade deflection at blade-tower passage are compared. For equal probabilities of exceedance the general blade deflection shows about 6% higher values than the

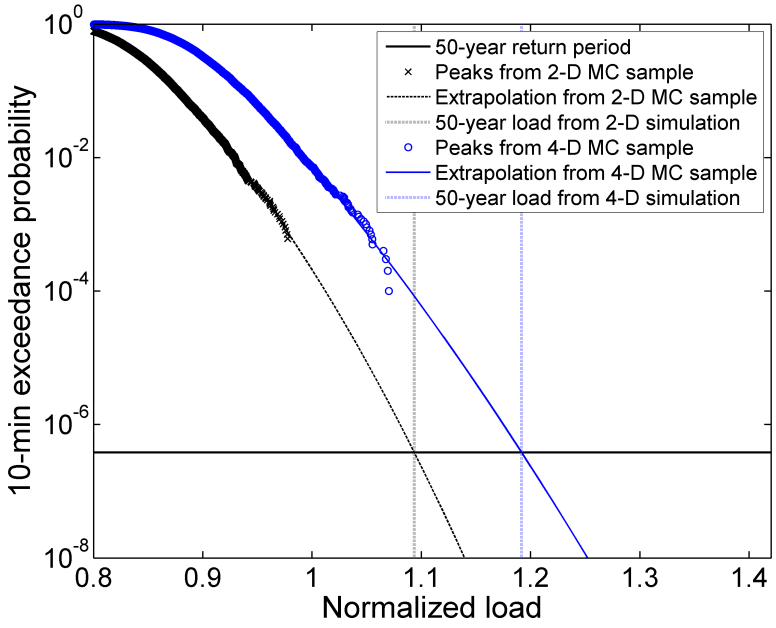

Figure 4.19: Comparison of extrapolated long-term distributions under two different sets of stochastic input variables.

deflection during blade-tower passing. Basing the design limits for blade-tower clearance on the absolute maximum blade deflection regardless of rotor asimuth is therefore a conservative approach.

4.2.3 Relative importance of environmental inputs

Importance factors as the α^2 values used in reliability analysis (Chapter 2) can be used as a measure of the relative importance of environmental and structural conditions to the extreme response. Here the relative importances of the 4 stochastic variables in the Monte Carlo sample shown on Figure 4.19 are estimated as function of a response threshold level, $u_{B,limit}$. The values of the importance factors are calculated using the approximate method given in equations 2.16 -2.18 from Chapter 2. For that purpose, a simple limit state equation is defined

$$g(\mathbf{X}, u_{B,limit}) = u_{B,limit} - u_B(\mathbf{X}) \tag{4.46}$$

For each threshold level $u_{B,limit} = u$ an estimation of the importance factors is done based on all points for which $g(\mathbf{X}, u_{B,limit}) \leq 0$.

The choice of threshold levels is done based on Figure 4.21 where it is seen that the

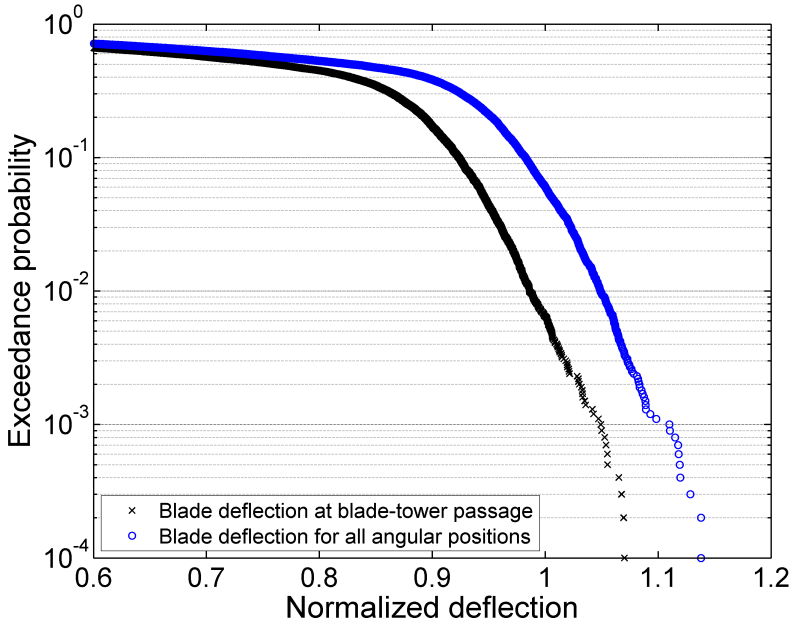

Figure 4.20: Difference between maximum overall blade deflection and maximum blade deflection observed during blade-tower passage.

magnitude of the response has a strong dependence on the wind speed, with a peak for wind speeds of around 10m/s. This is caused by the turbine control strategy, the peak in blade deflection corresponding to the wind speeds where the thrust force on the turbine is the highest, which reflects in high blade loads. Clearly the ten-minute wind speed is the most significant factor influencing the long-term statistical distribution of blade deflection, both because blade loading is dependent on wind speed as shown on Figure 4.21, but also because turbulence and wind shear are modelled as conditionally dependent on the wind speed. In order to evaluate importance factors for input parameters other than the wind speed, the threshold level has to be chosen sufficiently high so that most of the peaks above the threshold level occur in the same wind speed region of 8-12m/s. Figure 4.22 shows a graph of the importance factors $\boldsymbol{\alpha}$ as function of $u_{B,limit} \geq 0.8$.

Using such an arbitrarily defined limit state surface is a crude way to estimate importance factors, but it nevertheless reveals useful information. Given that the wind speed is in the range associated with highest blade loading, turbulence is probably the most significant contributor to the occurrence of extreme events. At lower threshold levels blade compliance and turbulence seem to have similar influence, however when only peaks above $0.85-0.9$ are considered, blade compliance gets less important and turbulence is clearly the dominating

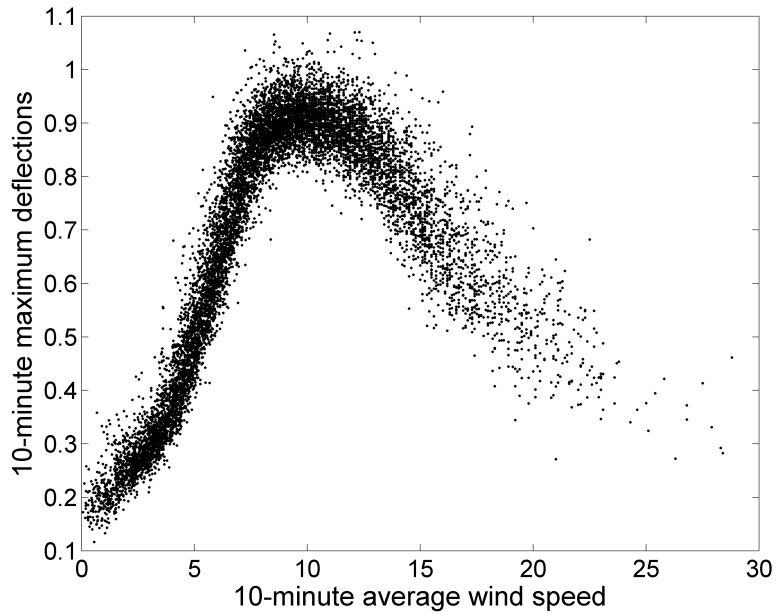

Figure 4.21: Influence of average wind speed on extreme wind turbine response.

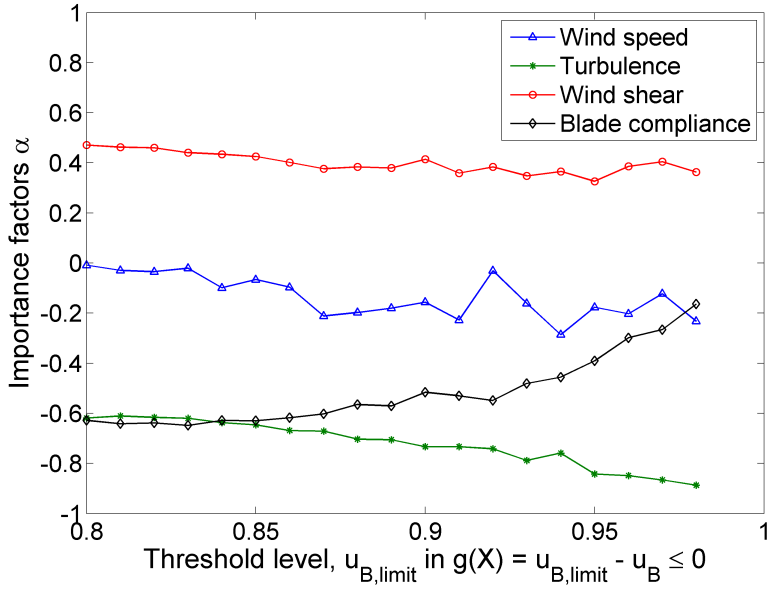

Figure 4.22: Importance factors from MC simulation estimated using arbitrary failure surface definition.

factor with importance of $|\alpha| > 0.8$. Wind shear also plays a role, and the opposite sign of its importance factor means that negative wind shear is more critical. This result confirms the observation made in connection with Figure 4.19, that it is necessary to take the wind shear statistics into account when analyzing extreme blade loads and deflections.

4.3 Probabilistic modelling of fatigue stress cycles

4.3.1 Introduction

The probabilistic aspects of fatigue load cycles have usually received less attention than the material properties, and the existing models are not always fully suitable for use with composite materials due to various reasons. For example, some publications such as Madsen, Krenk and Lind, 1986 [48], Moriarty, 2004 [52] define a single random parameter describing the statistical distribution of the total sum of stress cycles. Others as Friis-Hansen, 1994 [25], Kaminski, 2002 [41]) show similar models where the random parameter is the total accumulated damage at failure. These representations require assuming that the damage accumulated is only dependent on the stress range, and not on the mean values of the stress cycles. Using such an approach is simple and effective for some materials, however fatigue resistance of composite materials is often strongly dependent on the mean stress values, see Nijssen, 2004 [57], Toft, 2011 [86], which makes it necessary to use more advanced models for describing the distribution of stress cycles. Most often a 2-dimensional fatigue cycle spectrum for composites is modelled in the form of a "rainflow" matrix describing the accumulated number of stress cycles as function of stress range amplitudes and mean values.

4.3.2 Stress range binning

Defining a fatigue load cycle model requires dividing the stress cycles into bins. If the stress cycles are binned only according to their amplitudes a one-dimensional fatigue spectrum is obtained. If binning is done both according to amplitude and mean value, the result will be a fatigue cycle matrix (rainflow matrix). In principle the quantity of interest is the number of stress cycles in each bin accumulated over a given period of time. However, cycle counts have discrete values which limits the range of statistical analysis methods that can be used. On the other hand, the sum of all stress cycle values in each bin is a continuous variable which allows using all standard statistical techniques available for continuous data. If the value of the sum in the i^{th} bin, S_i, is known, the approximate number of stress cycles can

be estimated by dividing S_i to the stress range value r_i corresponding to the middle of the i^{th} bin. The binned stress cycle sums can be represented as a vector-valued multivariate function of dimension k, $\mathbf{S} = [S_1(r_1), \ldots, S_k(r_k)]$.

If the cycles need to be binned only according to amplitude, the dimension of the vector $\mathbf{S}$ will equal the number of amplitude bins, k_a. In case both the range and the mean value of the cycles is needed, the total dimension of the vector will be the product of the number of amplitude and mean value bins, $k_a \cdot k_m$.

4.3.3 Central limit theorem

Let us consider a function $Y = f(\mathbf{X})$ where $\mathbf{X}$ is a vector of random input variables. If a set of N independent samples are drawn from the statistical distribution of $\mathbf{X}$, the resulting function values $\mathbf{Y} = [Y(\mathbf{X_1}), Y(\mathbf{X_2}), \ldots Y(\mathbf{X_N})]$ will be independent and identically distributed (i.i.d). By means of the Central Limit theorem [40] the sample mean will for sufficiently large N approach a normal distribution with mean μ_y and standard deviation $\sigma_y/\sqrt{N}$:

$$\sqrt{N}\left(\frac{1}{N}\sum_{i=1}^{N} y(x_i) - \mu_y\right) \xrightarrow{d} \mathcal{N}(0, \sigma^2) \tag{4.47}$$

The expression in equation 4.47 can be extended to the multidimensional case, where for a k-dimensional function (Y_k) with $\boldsymbol{\mu} = E\langle \mathbf{Y} \rangle$ with i.i.d outcomes $\mathbf{Y_1}, \mathbf{Y_2}, \ldots, \mathbf{Y_N}$ where

$$\mathbf{Y_i} = \begin{bmatrix} Y_i(1) \\ \vdots \\ Y_i(k) \end{bmatrix} \tag{4.48}$$

The distribution of the vector of expected values will then converge to the multivariate normal distribution:

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N} [\mathbf{Y_i} - \boldsymbol{\mu}] \rightarrow \mathcal{N}_k(0, \boldsymbol{\Sigma}) \tag{4.49}$$

where $\mathcal{N}_k$ denotes the k-dimensional multivariate normal distribution with mean 0 and covariance matrix $\boldsymbol{\Sigma}$.

4.3.4 Load cycle counts for a stationary process

If the function $Y = f(\mathbf{X})$ represents the outcome of a stationary stochastic process, the distribution of Y will not change over time, meaning that any sample-based estimates of μ_y and σ_y obtained over a simulation period T_0 will also be representative of the process

over any other time interval T. For the case when the function value $\mathbf{Y_i}$ in equation 4.49 represents the sum of stress amplitudes $\mathbf{S_i}$ discussed above, the values of $\boldsymbol{\mu_S}$ and $\boldsymbol{\Sigma_{SS}}$ will scale with the sample size due to the summation over different number of terms. If a Monte Carlo sample of size T_0 is simulated, the obtained sample statistical moments can be used to estimate the statistical moments of a Monte Carlo sample with an arbitrary size T:

$$\boldsymbol{\mu_S}^{(T)} \approx \tfrac{T}{T_0}\boldsymbol{\mu_{\hat{s}}}^{(T_0)} \qquad ; \qquad \boldsymbol{\Sigma_{SS}}^{(T)} \approx \sqrt{\tfrac{T}{T_0}}\boldsymbol{\Sigma_{\hat{s}\hat{s}}}^{(T_0)} \tag{4.50}$$

It follows that if a Monte Carlo sample is generated which is large enough to properly represent response statistics and for the central limit assumption to be valid, the sample moments can be used to estimate what will be the distribution of a Monte Carlo sample with different size. To illustrate that with a real-world example, if the outcome of a 10-minute dynamic loading time history is taken as a single sample point, then a set of, say, 1000 such time histories can be used to obtain an estimate of the distributions of stress cycle sums S_i. Then equation 4.50 can be used to scale the loads and estimate what is the expected statistical distribution of a larger data set, for example covering 1 year (52560 simulations, or about 50 times longer period). This statistical distribution can then be used within a reliability model to estimate the annual reliability index of the structure.

4.3.5 Load cycle counts for a non-stationary process

Structural load time histories are not always stationary processes. Wind turbine response statistics for example vary significantly with the wind conditions which the turbine is subjected to. In order for the model outlined above to be valid for a given non-stationary process, the process needs to show a degree of mixing, meaning that the dependency within the process decreases with time separation, and process statistics will ultimately approach constant values if observed over a long period of time. For such processes a solution which preserves the validity of the central-limit theorem is drawing a random sample directly from the long-term distribution of the input variables. However, due to the non-stationarity, the sample size necessary for reaching convergence towards the normal distribution will be larger compared to the case of a stationary process.

4.3.6 Extrapolation of model statistics

As all predictive models the fatigue load model presently under consideration needs to be able to predict events that have not been observed in the sampled data. In the present case

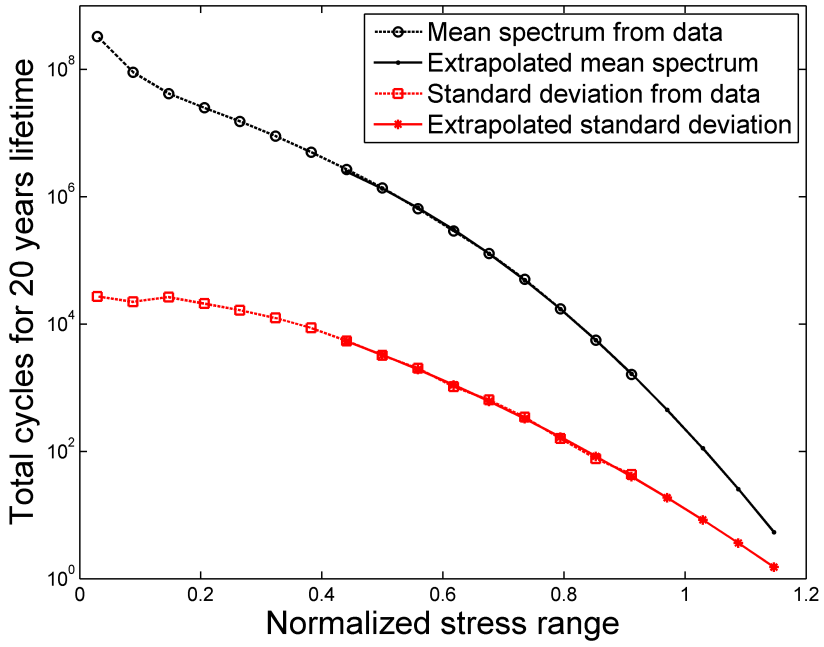

Figure 4.23: Extrapolation of fatigue spectrum distribution parameters.

such unobserved events are stress cycles with amplitudes larger than what has been recorded, or with amplitude range where only a few events have been recorded and the information is not enough to obtain reliable statistics. A solution for such cases is extrapolating the statistics from range bins where sufficient data exists, to range bins where not enough data are available. The results from such an operation performed on the data used later in section 4.3.8 are shown on Figure 4.23, where the mean and standard deviation of the number of cycles in a given bin are extrapolated using quadratic functions. The quadratic functions extrapolate the mean and standard deviation of the range bin sums $\mathbf{S} = [S_1, S_2, \ldots S_k]^T$, while the off-diagonal entries of the covariance matrix $\boldsymbol{\Sigma_{SS}}$ that correspond to the extrapolated ranges are all set to zero.

4.3.7 Procedure for probabilistic model of fatigue load cycles

Summarizing the information given above, the following step-by-step procedure for obtaining a probabilistic fatigue load cycle model from response time history is outlined:

1. Using simulation or measured response data, obtain a set of m time history samples with equal length T_0, where T_0 is sufficiently long as to be representative for the

statistical distribution of the modelled stochastic process

2. Apply an algorithm for counting load hysteresis loops, for example the rainflow counting method

3. Divide each of the collected cycle time histories into bins according to cycle range and mean values

4. Sum the total range values in each bin, for each time history:

$$\mathbf{S} = \begin{bmatrix} \sum r_{1,1}^{(1)} & \cdots & \sum r_{1,1}^{(m)} \\ \vdots & & \vdots \\ \sum r_{1,k_2}^{(1)} & \cdots & \sum r_{1,k_2}^{(m)} \\ \sum r_{2,1}^{(1)} & \cdots & \sum r_{2,1}^{(m)} \\ \vdots & & \vdots \\ \sum r_{k_1,k_2}^{(1)} & \cdots & \sum r_{k_1,k_2}^{(m)} \end{bmatrix} \tag{4.51}$$

where k_1 and k_2 are the number of bins for cycle range and mean value, respectively, m is the number of time histories evaluated, and $r_{i,j}^{(k)}$ are the ranges of the individual stress cycles obtained by the cycle counting algorithm.

5. Find the vector of sample mean values $\boldsymbol{\mu}_{\hat{\boldsymbol{S}}}$ and the covariance matrix $\boldsymbol{\Sigma}_{\hat{\boldsymbol{S}}\hat{\boldsymbol{S}}}$ evaluated over the sample size m

6. If necessary, extrapolate the values of $\boldsymbol{\mu}_{\hat{\boldsymbol{S}}}$ and $\boldsymbol{\Sigma}_{\hat{\boldsymbol{S}}\hat{\boldsymbol{S}}}$ to larger stress amplitudes or mean values.

7. Obtain estimates for the statistical distribution of the fatigue loads accumulated over the design time period T by using the expression from equation 4.50:

$$\boldsymbol{\mu}_{\boldsymbol{S}}{}^{(T)} \approx \tfrac{T}{T_0}\boldsymbol{\mu}_{\hat{\boldsymbol{S}}}{}^{(T_0)} \qquad ; \qquad \boldsymbol{\Sigma}_{\boldsymbol{S}\boldsymbol{S}}{}^{(T)} \approx \sqrt{\tfrac{T}{T_0}}\boldsymbol{\Sigma}_{\hat{\boldsymbol{S}}\hat{\boldsymbol{S}}}{}^{(T_0)} \tag{4.52}$$

4.3.8 Probabilistic cycle counts of wind turbine blade response

The procedure for obtaining a probabilistic fatigue cycle count model is illustrated here by estimating the number of stress cycles caused by the flap-wise (along-wind) bending of a wind turbine blade for a 2.3MW commercial wind turbine under simulated normal operating conditions. Fatigue load time histories are obtained using the aeroelastic simulation code BHawC.

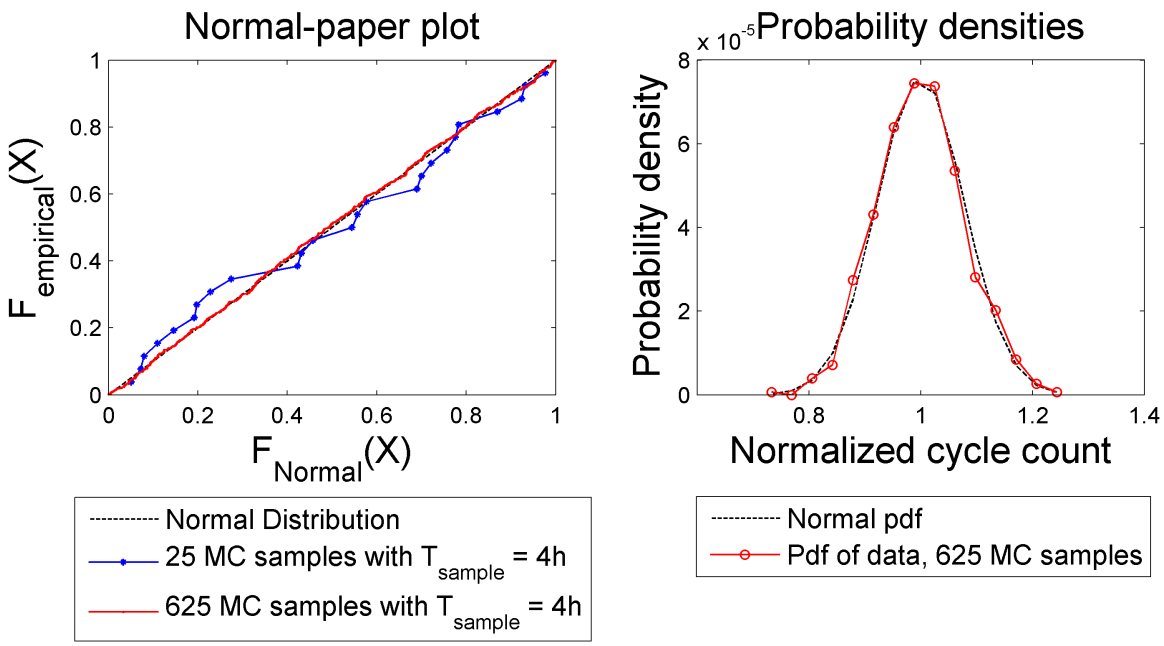

Figure 4.24: Statistical distribution of stress range sums.

Cycle count distribution under stationary conditions

The first problem considered is the short-term load distribution obtained from time histories generated under fixed input parameters: average hub height wind speed $U_{hub} = 11m/s$ and turbulence intensity $I = 0.15$. The constant input parameters mean that the simulated process will be a stationary one. A total of 15000 time series with 10 minute duration are simulated, amounting to about 104 days of time history (this is the same data set used in Section 4.1 for assessment of load extrapolation methods). Using the procedure outlined in section 4.3.7, two probabilistic load models are fitted. The first one uses a relatively short time history of 100 hours, which is divided into 25 samples of 4h duration. The second model is fit to the entire 2500h of available data divided into 625 samples of 4h duration each, and is used as reference for evaluating the quality of the fit obtained with the model using limited amount of data. Figure 4.24 shows an example of the statistical distribution of the stress range sums for an arbitrarily chosen bin. It is clearly visible how in accordance with the central limit theorem for large sample sizes the distribution is converging towards the normal distribution.

Another comparison between the limited-data and full-data models is shown on Figure 4.25, where the fatigue range spectra are shown. On the left side of the figure are plotted

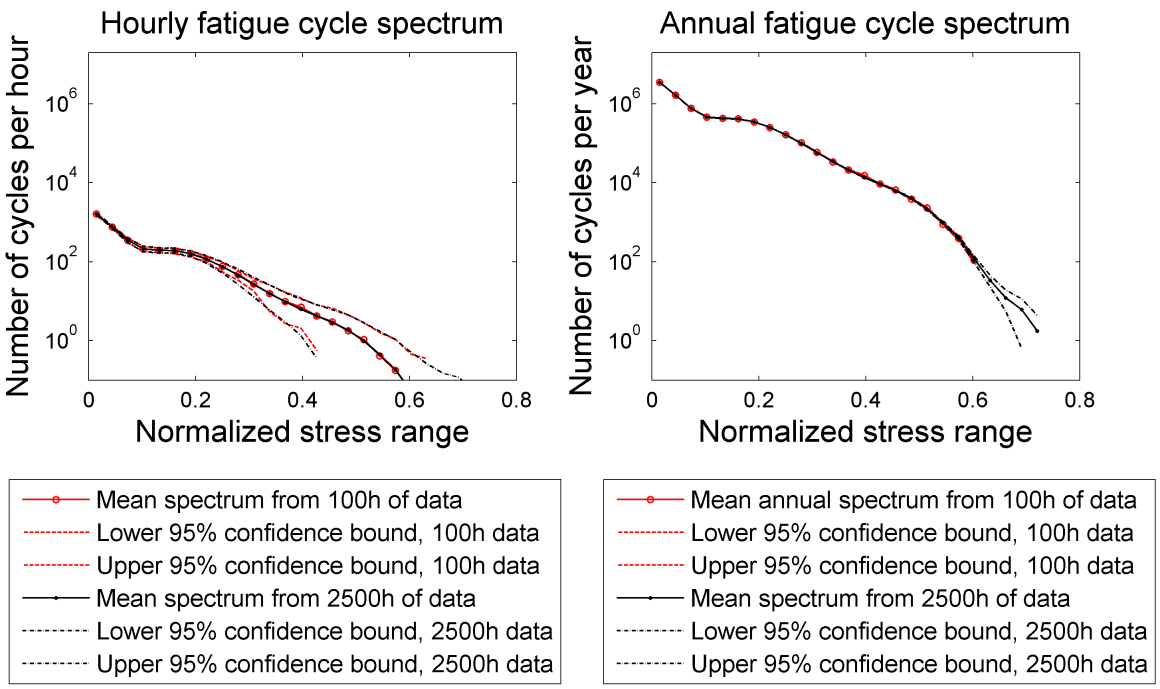

Figure 4.25: Statistical distribution of fatigue spectra under stationary conditions.

the unmodified sample statistics, while the right hand side shows the estimated annual fatigue cycle accumulation, derived using equation 4.52. The uncertainty in the fatigue loads becomes much smaller for longer time periods as seen from Figure 4.25. This can be expected because the mean number of cycles scales proportionally to the length of the time period, while the standard deviation grows with the square root of the same. Therefore a longer time interval allows better convergence of sample statistics towards the actual mean value of the process.

Figure 4.26 shows a 3-D plot of the rainflow matrix, where the annual number of accumulated fatigue cycles is given as function of both mean and range values. Similarly to the 1-D fatigue spectrum, the uncertainty in the annual fatigue loads in the matrix is relatively small, and is limited to the bins where a small number of cycles are expected to occur. Generating random outcomes from the fatigue models shown above requires also that correlations between random variables are taken into account. The correlation matrices are also obtained from data and shown on Figure 4.27, both for the 1-D fatigue spectrum and the 2-D fatigue matrix. The rainflow matrix contains a large number of empty bins which causes some of the correlation coefficients to be undefined. Therefore the correlation matrix for the 2-D fatigue model shown on Figure 4.27 has rank significantly lower than the full rank corre-

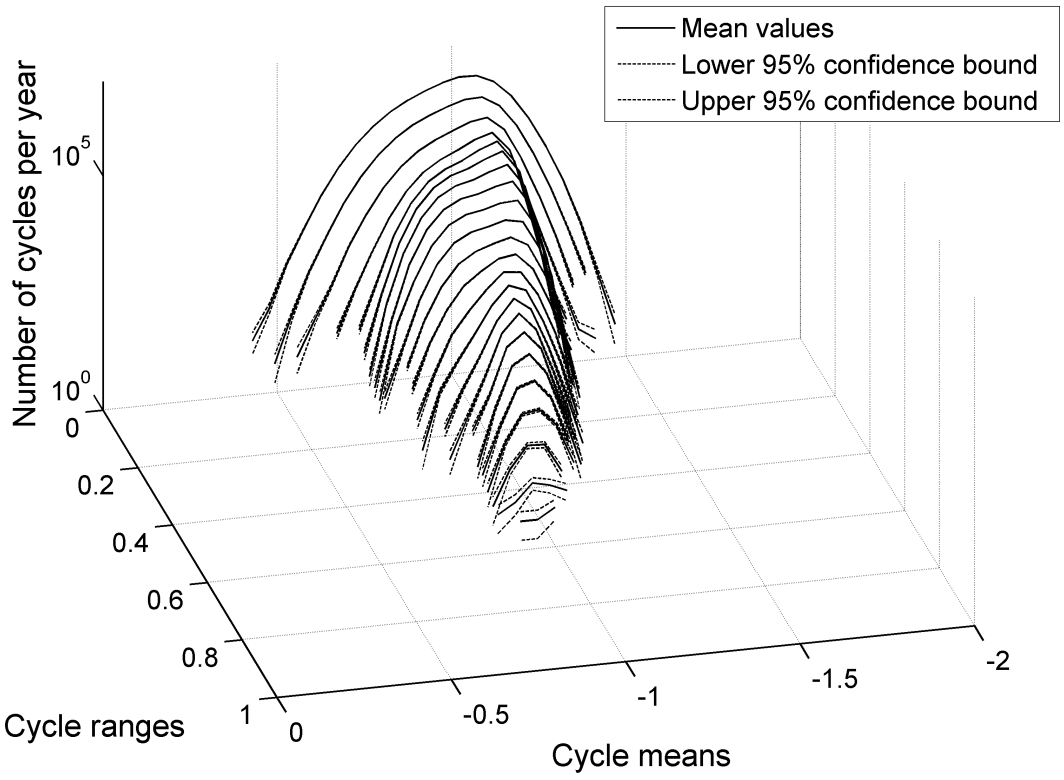

Figure 4.26: Stochastic annual load cycle accumulation matrix under stationary conditions.

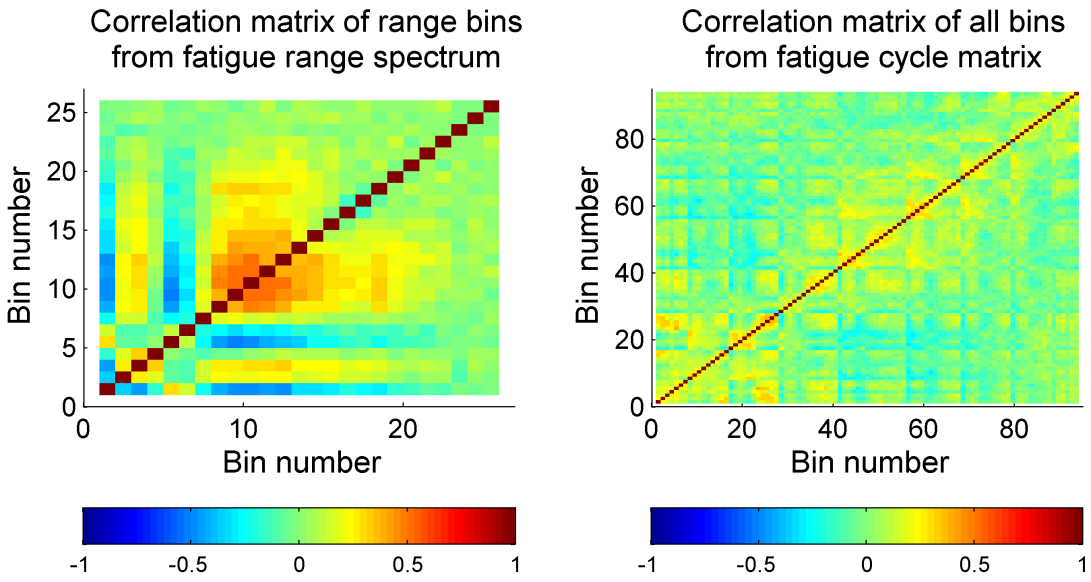

Figure 4.27: Correlation between stress ranges under stationary conditions.

sponding to the total number of bins, which would equal 624. While the correlation matrix corresponding to a 1-D spectrum contains some terms showing significant though not very strong correlation, in the 2-D model the correlations get very scattered and less significant, and their practical influence on the model outcomes is limited.

Long-term fatigue load

Long-term distribution of fatigue loads can be obtained either by direct sampling of the long-term distribution, or by integration of a number of short-term distributions obtained under varying input conditions. Here direct sampling of the long-term distribution is used, which is the less computationally-intensive method when the long-term distribution is a function of two or more parameters. The same turbine model and structure as for the short-term distribution described in section 4.3.8 are used, however now the environmental input parameters instead of being fixed will vary according to their long-term statistical distribution. The influence of two random variables is taken into account: average wind speed at hub height, and turbulence intensity. The distribution of both parameters corresponds to an environmental condition class 2B as given by the international standard IEC61400-1 [37]. A total of 8400 independent simulations with 10-minute duration corresponding to 1400h of time history are simulated using the aeroelastic code BHawC. Sample statistics are then obtained by splitting the data into 84 samples each containing 100 independent simulations (a total of 16.7h of data per sample). The resulting probabilistic long-term rainflow matrix is shown on Figure 4.28, and the correlation matrices for both a 1-D and 2-D model are shown on Figure 4.29. The long-term fatigue cycle matrix shows similar load uncertainties compared to the ones obtained under stationary conditions. The main difference is that due to the non-stationarity of the process a larger amount of time series is required in order to obtain a good model. The effects of non-stationarity are also visible in the correlation matrix plots on Figure 4.29 where stress ranges from the long-term distribution show much stronger correlation structure compared to the stationary process considered in section 4.3.8.

Figure 4.30 shows the coefficient of variation (c.o.v) of the fatigue damage predicted by the probabilistic cycle count model using the two spectra shown on Figure 4.25, the S-N curve from equation 3.20 in Chapter 3, and summing the damage contributions of individual cycles using Palmgren-Miners's rule (equation 3.31). Both in the case of short-term and long-term fatigue loads the total accumulated damage is normalized so that the damage for $m = 10$ has an average of 1. The uncertainty in the short-term damage is significant, while for a lifetime of 20 years the uncertainty is small, with c.o.v between 0.01 and 0.1% and

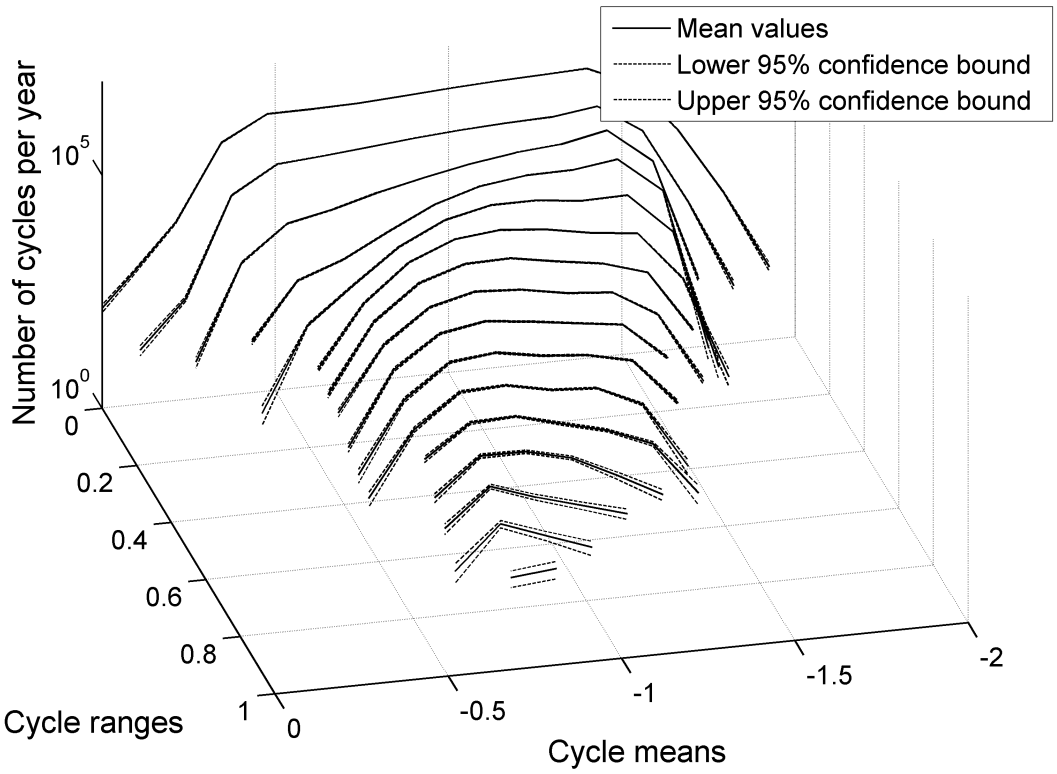

Figure 4.28: Stochastic annual load cycle accumulation matrix for simulated long-term fatigue load distribution.

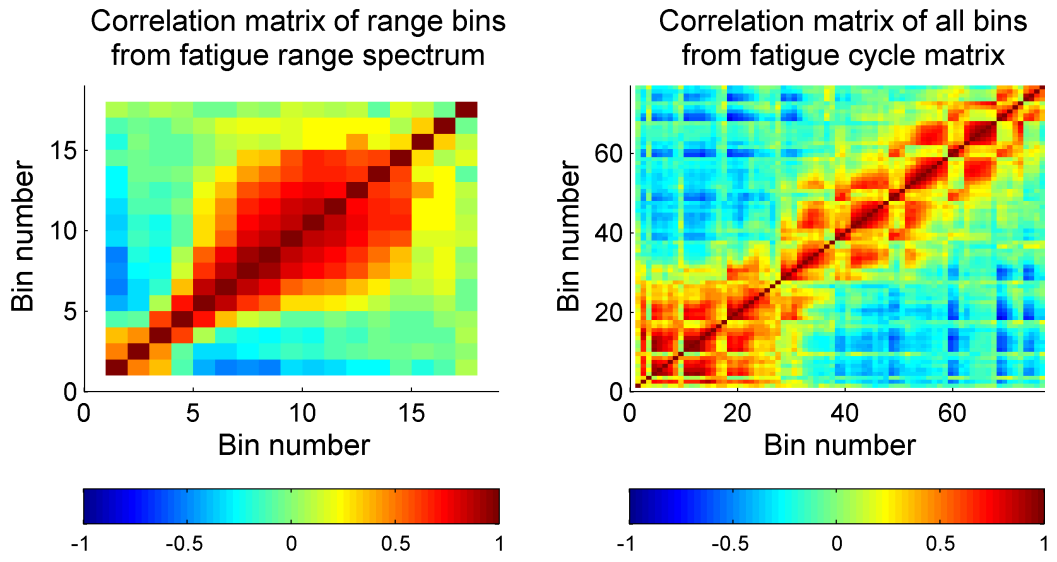

Figure 4.29: Correlation between stress ranges for simulated long-term fatigue load distribution.

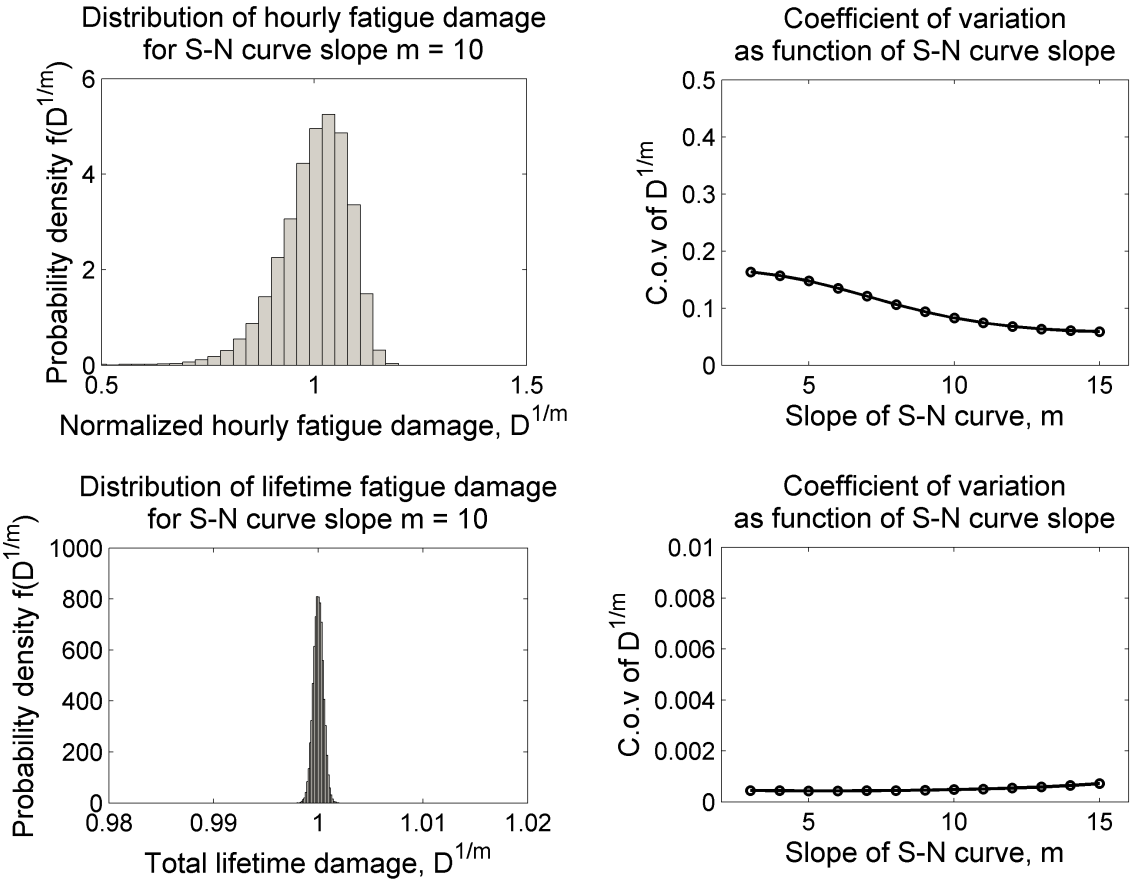

Figure 4.30: Coefficient of variation of total lifetime fatigue damage as function of S-N curve slope.

slightly growing for larger values of m. For values of m in the range between 10 and 15 which is typical for glass-fiber reinforced composites, the c.ov. is between 0.05 and 0.1%. The larger uncertainty for higher fatigue slopes can be explained with the growing influence of high-amplitude cycles for higher values of m. The load model gives higher uncertainty for load cycles with high amplitude, hence the increased uncertainty in the accumulated lifetime damage when the high-load cycles have dominating contribution to the overall damage. Figure 4.30 shows that considering the uncertainty in the number of stress cycles is only relevant when analyzing short time periods, while for calculations spanning years or decades the effects of the uncertainty vanish.

4.4 Discussion on statistical distribution of loads

The present chapter showed that the statistical distribution of extreme loads under normal operating conditions can be reasonably well approximated by statistical load extrapolation. The tail of the turbine response seems to be well-behaved and linear for low exceedance probabilities, which allows for efficient and accurate extrapolation. It has to be noted though that load extrapolation only covers a limited range of conditions. Discrete events such as emergency stops, faults, grid drops or storms need to be covered in additional analysis. Another source of uncertainty in the extrapolation is the presence of controller, which is not guaranteed to have a completely linear behaviour and to react to extreme load events in the same way as to regular loads. Example of such a response feature could be load or vibration sensors employed in an active load curtailment functionality that is triggered when extreme loads are detected.

Not surprisingly, turbulence was identified as one of the dominating factors that may cause extreme response events. However, it was seen that wind shear can also have a significant influence, specifically for the blade-tower distance. The guideline IEC 61400-1 [37] gives specific modelling suggestions for turbulence, however variable wind shear is only required in a limited number of scenarios without turbulence, and for general simulations a fixed exponent $\alpha = 0.2$ is specified. This value is favorable for blade-tower distance and can lead to non-conservative estimates for blade-tower collision probabilities. Therefore a modification of the wind shear model in future versions of the IEC standard could lead to improved design criteria.

In general, the accuracy of load simulations is only guaranteed provided that the external conditions are exactly as modelled. This is especially valid for the extreme loads which were

shown to be strongly influenced by inflow conditions. Although probably to a less extent, this discussion does also apply to fatigue loads. Due to the time-averaging effects the long-term distribution of number of stress cycles has very low uncertainty, however, any bias in the calculation of load magnitudes or in the effect of environmental conditions will also reflect in the fatigue life estimate. The presence of uncertainties in the models of loads and external conditions is treated in the following chapter of this thesis.

Chapter 5

Treatment of Uncertainties

5.1 Types of uncertainty

A general way of classifying uncertainties is dividing them into aleatory and epistemic ones. Epistemic uncertainty is associated with not having sufficient knowledge of a certain phenomenon, and can in principle be reduced by improving our state of knowledge, for example by using more precise measurements or models. An uncertainty which we have no means of controlling and reducing by mathematics or measurements, as for example the ultimate strength of a bundle of glass fibers, is characterized as aleatory. Der Kiureghian & Ditlevsen, 2007 [14] note that aleatory uncertainty does not necessarily exist as such, because theoretically any apparently random phenomenon can be explained by a sufficiently sophisticated model, and what will be left is only epistemic uncertainty. Nevertheless it is for obvious reasons practically convenient to characterize some of the uncertainties as aleatory. Further division can be made according to the source of the uncertainty, distinguishing between physical (aleatory) uncertainty caused by naturally occurring random variations which can be measured, uncertainty coming from modelling errors, statistical uncertainties due to limited amount of data used for model parameter tuning or for quantifying inputs, and measurement uncertainties. Other types of uncertainties such as the ones arising from human activities (gross errors, decision making, and other) can also be part of a risk analysis consideration, however they will not be part of the present thesis.

The presence of uncertainty is accounted for in a reliability analysis by including additional random variables. This increases the variability in the model and usually results in decreased reliability estimate. If there is any epistemic uncertainty source included in the model, the resulting reliability index will not represent the actual physical reliability of the

structure, but it can be considered instead as a lower bound of the reliability. The difference between this "lower bound" and the actual reliability is a penalty to the design resulting from lack of sufficient knowledge and confidence. Reducing the epistemic uncertainty by e.g., improving models and measurements can thus increase the estimated reliability of the structure. Such an increase of the reliability estimate will not mean higher reliability of the structure itself, but it means that the "lower bound" estimate has come closer to the actual reliability, due to the improved state of information.

Consider a simple R-S limit state equation, where there are uncertainties in the estimations of both the demand S and the capacity R:

$$g(\mathbf{X}) = R(\mathbf{X}, X_{capacity}) - S(\mathbf{X}, X_{demand}) \tag{5.1}$$

Both the demand and the capacity are uncertain and functions of the stochastic variables $\mathbf{X}$, and the additional variables X_{demand} and $X_{capacity}$ have to account for uncertainties that are not accounted for in the calculation of the demand and capacity itself. These are the uncertainties arising from insufficient knowledge and model inaccuracy - i.e., a combination of model, statistical, and measurement uncertainties.

Tarp-Johansen et al., 2002 [79] have used the following definition of X_{demand} for ultimate loads:

$$X_{demand} = X_{exp}X_{st}X_{aero}X_{dyn}X_{str}X_{sim}X_{ext} \tag{5.2}$$

where X_{exp} accounts for the uncertainties in the exposure (terrain roughness, flow inclination, obstacles), X_{st} is a statistical uncertainty connected with the evaluation of the wind climate, X_{aero} is the model uncertainty due to errors in estimation of the blade aerodynamics in terms of lift and drag coefficients, X_{dyn} takes into account possible imprecise simulation of the dynamic response, X_{str} stands for model uncertainties in stress calculations using FEM or other models, X_{sim} accounts for statistical uncertainty due to the limited number of simulations used to assess the wind turbine response, and finally X_{ext} is the uncertainty in the extrapolation procedure such as distribution fits and peak extraction methods. In a study from 2003 [78] the same author suggests a similar set of uncertainties for fatigue life modelling, including terms related to both the load and the resistance of a structure:

$$X = \frac{X_{exp}^m X_{aero}^m X_{dyn}^m X_{str}^m X_{RFC}^m}{X_{SN}} \tag{5.3}$$

The additional terms in equation 5.3, X_{RFC} and X_{SN}, correspond to uncertainties in the stress cycle counting, and in the parameterization of the S-N curve, respectively. The exponent m in the numerator is the fatigue slope m. It appears in equation 5.3 because the

Table 5.1: Uncertainties in estimating extreme and fatigue loads.

Name	Description	Distribution	C.O.V		Mean
			Ultimate	Fatigue	
X_{exp}	Exposure	LN	0.1	0.1	$\mu = 1$
X_{st}	Wind statistics	LN	0.05	-	$\mu = 1$
X_{aero}	Airfoil coefficients	GB	0.1	0.1	$\mu = 1$
X_{dyn}	Dynamics	LN	0.05	0.05	$\mu = 1$
X_{sim}	Simulation	N	0.05	-	$\mu = 1$
X_{str}	Stress analysis	LN	0.03	0.05	$\mu = 1$
X_{ext}	Extrapolation	LN	0.05	0.05	$\mu = 1$
X_{RFC}	Rainflow count	LN	-	0.02	$\mu = 1$
X_{SN}	Material (S-N curve)	LN	-	0.05	$\mu = 1$

accumulated fatigue damage is also an exponential function of the stress amplitudes. Table 5.1 lists the distribution parameters of the uncertainties for extreme and fatigue limit states as given by [79].

Based on available measurement data and on the models defined in the previous chapters of this thesis, it is possible to quantify some of the uncertainties from Table 5.1. The following sections of this chapter describe the process of calibrating some of the uncertainties to a particular application (combination of wind turbine and site conditions), based on data from observed actual turbine behaviour. The uncertainty definitions used here are based mainly on the part of the model where they are included, and are not necessarily divided into physical, statistical, modelling or measurement uncertainties, but are usually a combination of several of these types.

5.2 Statistical uncertainty in wind climate assessment

Wind climate on a given turbine site is usually assessed by carrying out a measurement campaign over multiple years. Due to year-to-year variations in the wind climate, wind conditions measured over successive years can differ. In general, the longer the measurement period is, the smaller the uncertainty in the wind climate assessment will be. Theoretically, the variance of a sample mean $\bar{x}$ is given by $\sigma_{\bar{x}}^2 = \sigma_x^2/N$, where N is the sample size. The uncertainty in estimating the mean wind climate will show the same behaviour and will be

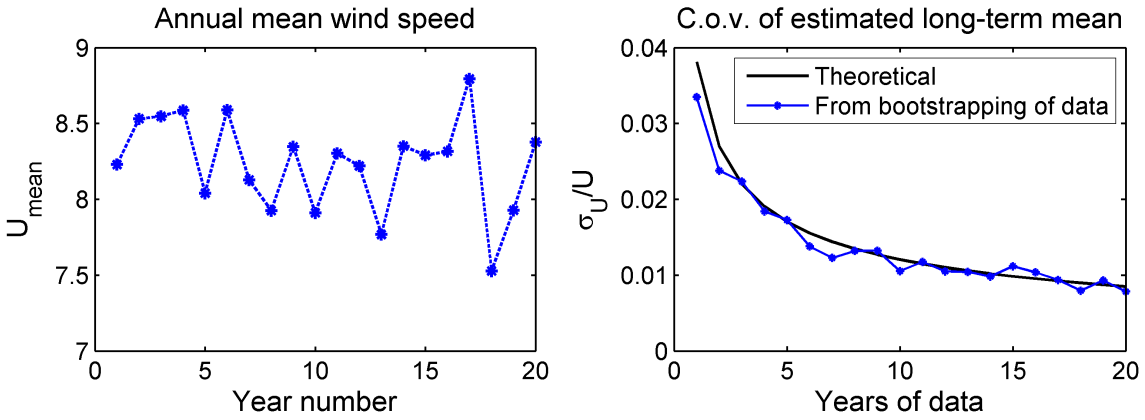

Figure 5.1: Uncertainty in the estimation of average wind speed as function of the number of years of data.

decreasing proportionally to the number of measurement years. This is illustrated on Figure 5.1, created using 10-minute average wind speed data taken over 20 years at the Danish island of Sprogø.

Once the uncertainty in quantifying a given wind climate parameter has been assessed, the next step is to translate that uncertainty into uncertainty in load levels. This is done by evaluating the load distribution as function of the given environmental parameter. Figure 5.2 shows an example where the annual extreme blade flapwise load is evaluated as function of the average annual wind speed following the Rayleigh distribution. The extreme load is in general a nonlinear function of the average wind speed and other environmental parameters, but in the near vicinity of a given point on the curve, it can be sufficiently well represented by a linear approximation. Using a linear approximation also allows for easy extension of the problem to a multi-dimensional one, and the uncertainty in the extreme load can be represented as a function of any number of parameters using the expression

$$\sigma_F = \nabla F(\mathbf{P})^T \mathbf{\Sigma_{PP}} \nabla F(\mathbf{P}) \tag{5.4}$$

where $\mathbf{P}$ is the vector of parameters, $F(\mathbf{P})$ is the extreme load with a given recurrence period (for example annual extreme load or the 98% quantile of the annual extreme), and $\mathbf{\Sigma_{PP}}$ is the covariance matrix of the parameters. Applying this expression to the mean wind speed data shown on Figure 5.1 results in expected coefficient of variation of the annual extreme load level of 0.2% or less, even after only 1 year of measurements. The low coefficient of variation is due to the small effect that changes in average wind speed have on the extreme

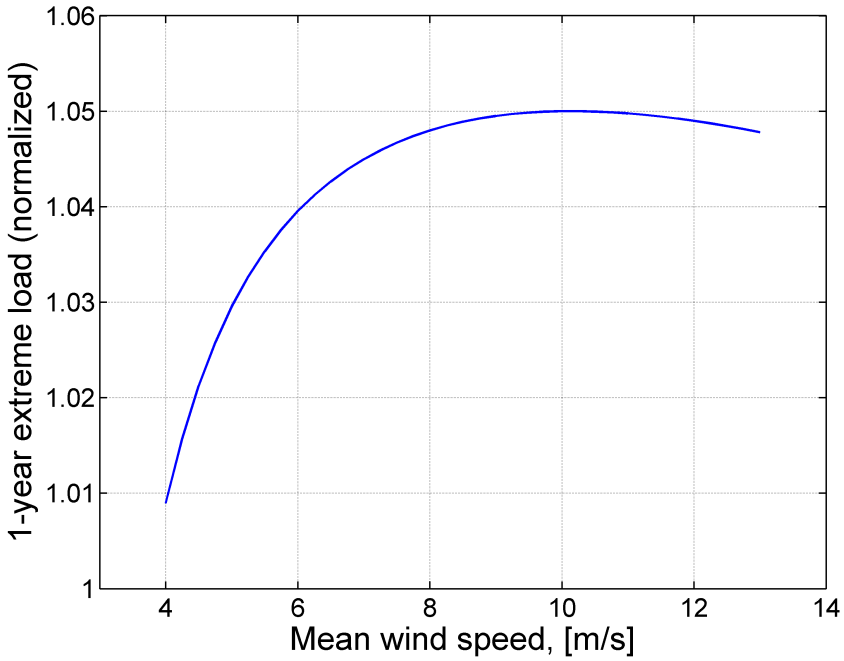

Figure 5.2: Extreme blade flapwise load with 1-year period of recurrence as function of the mean of Rayleigh-distributed wind speed.

loads. The highest loads on the wind turbine blades in many cases occur at about rated or slightly below rated wind speed, and therefore the probability of observing high loads is connected with the amount of time the 10-minute average wind speed at the site is close to the wind speed at which maximum loads occur. This can be seen on Figure 5.2, where the expected extreme load has a peak at about $10m/s$ mean wind speed. A slightly larger uncertainty could in principle be expected for fatigue loads, because a difference in average wind speed will not only result in changing load magnitudes, but also in a varying number of stress cycles due to the different number of operational hours under certain conditions.

As turbulence and wind shear have a high influence on the variance in the extreme loads, it is expected that the statistical uncertainty associated with these might be larger. The effect of statistical uncertainty in turbulence and wind shear estimations is however not quantified in the present work due to lack of available long-term statistical data for turbulence and wind shear. Instead, the value of statistical uncertainty quoted by Tarp-Johansen et. al, 2002 [79] will be used, which is represented by a log-normally distributed variable X_{st} with mean of 1 and c.o.v of 0.05.

5.3 Uncertainty in ultimate limit state

5.3.1 Model uncertainty in ultimate resistance of the structure

The ultimate limit state in the reliability models discussed in the present thesis is either extreme blade deflection, or failure of composite lamina. For blade deflection the "resistance" of the structure is the available blade-tower distance for an undeformed blade. The uncertainty in this value is mainly due to possible deviations in the geometry of the turbine, and it is accounted for by including the variable X_{geom}, which is considered as normally distributed with mean of 1 and standard deviation of 0.03. In the case when material failure is used as limit state, the scatter in composite strength is accounted for by probabilistic modelling of lamina strength. There are uncertainties associated with the geometry of the structures (e.g., lamina thicknesses), which could cause errors in stress estimation. In the present thesis this is accounted for by considering plate thicknesses as random variables. There are also uncertainties associated with using a failure criterion to identify failure, and with the definition of a failure event. This could be modelled by introducing another uncertainty variable, X_{FC}. However, in the present thesis the loading in all ultimate strength analysis is compressive, and the Hashin criterion expression for defining compressive strength is just the lamina compressive strength itself, therefore the variance of the uncertainty variable X_{FC} is considered to be part of the variance of the material compressive strength, X_c.

5.3.2 Statistical uncertainty in extreme load distribution

Statistical uncertainties in the estimation of extreme loads exist due to the limited number of time series used to assess the load distribution and due to the subsequently applied extrapolation. Section 4.1 of this thesis already discussed this uncertainty in the context of convergence analysis of the extrapolation methods, where convergence was considered sufficient when the coefficient of variation of a given extreme load estimate was less than 0.05. Based on that discussion, it is considered that an appropriate value for the c.o.v. of X_{ext} is 0.05.

5.3.3 Statistical uncertainty due to time- and space-averaging

Ten-minute time period is the reference duration for the majority of wind turbine-related statistics. The standard procedures for determining wind climate, load distribution, power curve, and subsequent design parameters are all based on statistics from ten-minute reference

periods. This choice of period duration is based on the common form of weather records, where the ten-minute measurement periods have been used for many years. Such a period is practical for long-term statistics where measurements are taken over several years or decades, however ten minutes is a period long enough for significant changes in conditions due to turbulence to occur. Change in wind conditions will subsequently result in changes in turbine behaviour, and it can be expected that the period-to-period variation in loads and load extremes will be large. The uncertainty is further increased by the use of just two values - mean and standard deviation of wind speed in a single point - which does not take into account the spatial variation of the wind field across the rotor. This may cast doubts on whether the use of ten-minute reference period and a single reference location is the approach best suited for representing turbine load statistics, especially in the case of variable-speed, pitch-regulated turbines where several control strategies are employed depending on the wind speed. The parameters controlled are usually power, torque, blade pitch angle, and rotor angular velocity. Table 5.2 shows an overview of the three operational regimes typically employed on a variable-speed turbine:

1) Constant-pitch regime for low wind speeds, where the pitch angle is held at its optimal value, and the rotor angular velocity is controlled to yield the highest possible power. This regime is active for all wind speeds below V_{Ω_r} which marks the point at which the rated rotor angular velocity, Ω_r, is achieved.

2) Torque-control regime, which is transitional between the constant-pitch and constant-power regime. For wind speeds between V_{Ω_r} and V_r the nominal rotational speed is reached, however the nominal power is not reached yet. Optimal power output is ensured by controlling generator excitation and torque, while keeping the angular velocity constant.

3) Pitch-control (constant-power) regime is activated at wind speeds above the rated speed, V_r. At wind speed above V_r both the rotational velocity and power are at their rated values, and turbine control is achieved through changing the pitch angle of the blades. There is an excess of available energy in the wind, and the blades are pitched out in order to keep power constant and avoid over-speeding.

When different control strategies are used it is natural to expect that loads on the blades will differ too. This can be seen on Figure 5.3, where the distribution of load extremes is compared for the three different operating regimes. Load distributions are generated by simulating 100 time series with ten-minute duration, using mean wind speed of $11m/s$.

Table 5.2: Operating regimes of a variable-speed, pitch-controlled wind turbine.

Regime	Wind speed	Power	Pitch angle	Rotational speed
Constant pitch	$V_{cut-in} - V_{\Omega_r}$	$0 - P(\Omega_r)$	$0°$	$0 - \Omega_r$
Torque control	$V_{\Omega_r} - V_r$	$P(\Omega_r) - P_r$	$0°$	Ω_r
Pitch control	$V_r - V_{cut-out}$	P_r	$0° - 90°$	Ω_r

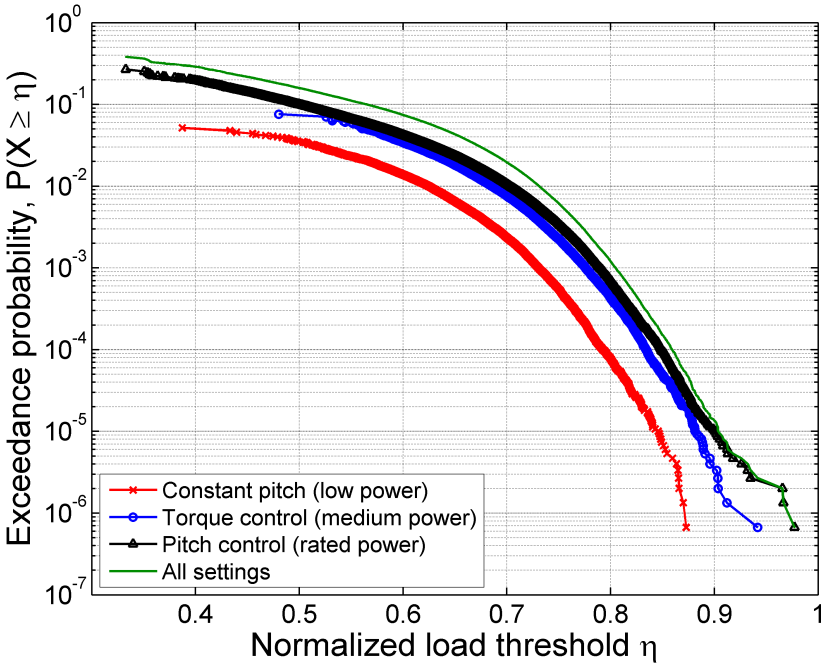

Figure 5.3: Extreme blade flapwise loads under different control strategies.

Although the average wind speed is the same in all time series, due to the turbulence the actual instantaneous wind speed varies in wide ranges and all three operating regimes become activated. Figure 5.4 shows an example time series where all three regimes are present in a ten-minute period.

Having seen that turbulence can cause significant changes in the response, two questions need to be answered: a) Does the time- and space-averaging of wind speed and turbulence have any effect on the long-term distribution of extreme loads, and b) Would using a shorter reference period reduce the scatter.

The effect of using 10-minute-averaged wind data at the hub location as reference is investigated by comparing measured and simulated time series of wind turbine response.

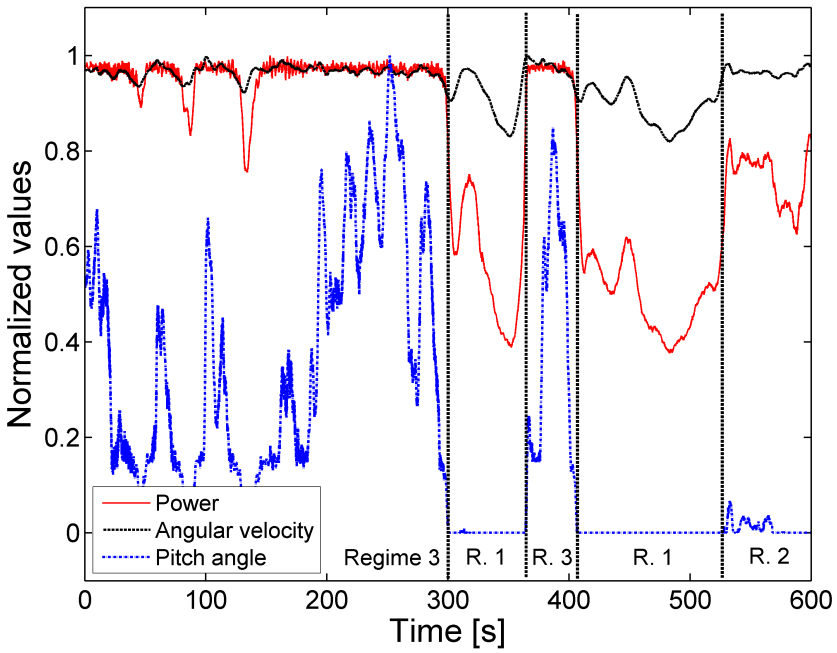

Figure 5.4: Example time series with different operational regimes.

This is followed by a brief study on what the effect of shortening the reference period would be. The latter study is based solely on simulations.

Comparison between measured and simulated extreme loads

The extreme values have the highest importance for reliability analysis purposes. The comparison between measured and modelled loads is therefore made based on extreme values of blade loads measured in 10-minute intervals. A data set containing 1932 measurements on the maximum blade root flapwise moment of a $2.3MW$ turbine is collected over a 22-day period. In parallel with the load data a number of environmental inflow conditions - average hub height wind speed, turbulence intensity, wind direction and wind shear are collected and used as input to aeroelastic simulations providing a comparison between simulated and measured data. As seen on Figure 5.5 the statistical distributions of measured and simulated extreme loads match well; the mean of the ratio between measured and simulated extremes is $E\langle M_{max,Measured}/M_{max,Model}\rangle = 0.988$, and the value of standard deviations is practically identical, the ratio between them being 0.999. There is also a good correlation between the individual extremes, which is visible from the one-to-one comparison shown on Figure 5.6.

Naturally, there is some scatter visible on the correlation diagram shown on Figure 5.6.

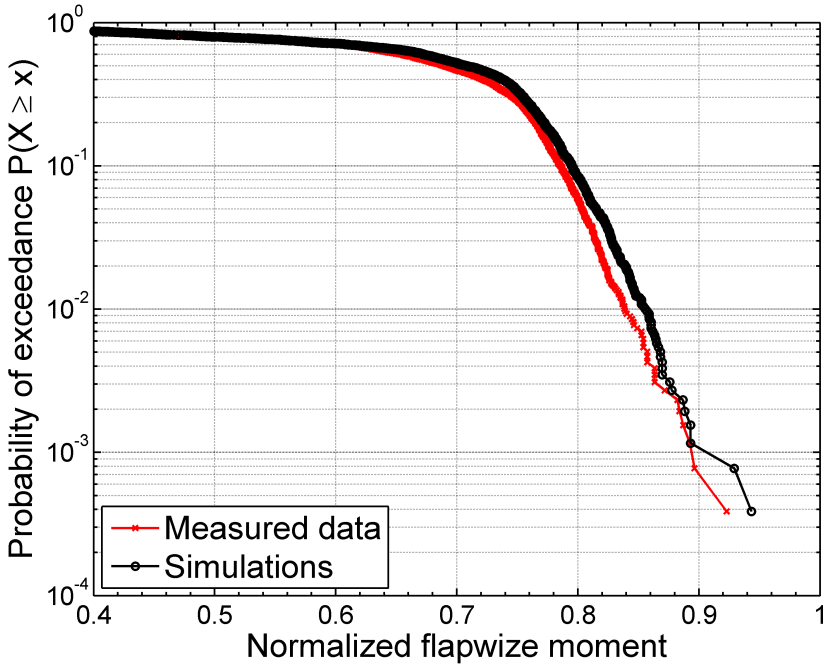

Figure 5.5: Statistical distribution of measured and simulated extreme blade loads.

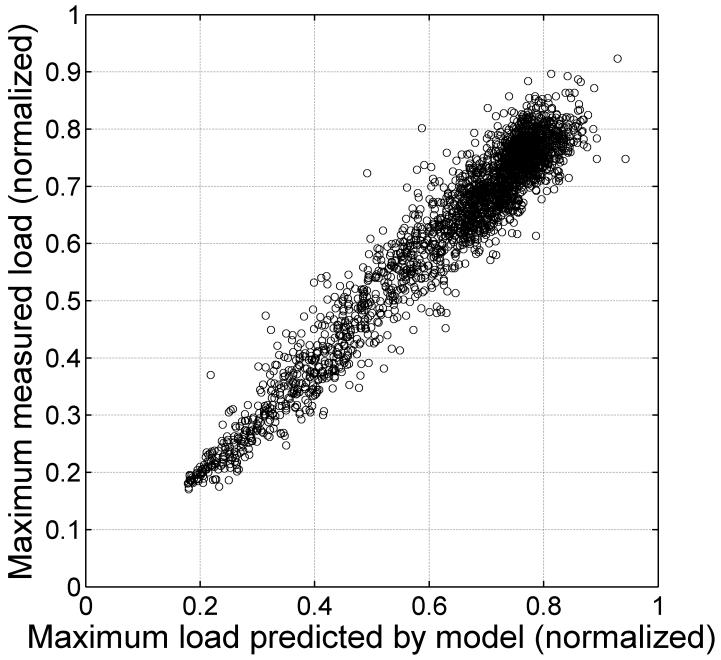

Figure 5.6: Correlation diagram for modelled and measured 10-minute extreme loads.

This scatter can to a large extent be attributed to the effects of averaging the wind speed time history, which makes it impossible to recreate the exact measurement conditions in the simulations. Other factors that can contribute to the difference between measured and simulated loads could be uncertainties in the load measurements due to e.g., strain gauge calibration, or uncertainties in the load model due to errors in the blade aerodynamic coefficients, in the Blade Element Momentum method (Hansen, 2008 [31]) which is used to model rotor aerodynamics, uncertainties in the Mann turbulence model (Mann, 1994 [49]) used to simulate a random turbulence field, uncertainties in modelling the dynamic response of the turbine, and other. However, if any of the mentioned errors had a significant presence in the model, its influence would be also seen as a bias between measurements and model predictions. The good agreement between the statistical distributions of measured and modelled loads under all observed wind conditions shows that errors due to airfoil coefficients, turbulence spectrum modelling, dynamic response modelling, and similar, are negligible.

Just as with any other statistical population, a larger sample size results in lower standard deviation of the sample mean, i.e., over long periods the average ratio between simulated and measured extreme loads will converge to a single value. Based on the available data, the expected long-term mean ratio is very close to 1. As a result, the variable X_{load} representing the ratio between measured and modelled extreme loads, $X_{load} = \frac{M_{max,Measured}}{M_{max,Model}}$, can be considered deterministic with value of 1.

Considering equation 5.2, the variable X_{load} will represent a combination of several uncertainty factors listed in the equation. These are X_{aero}, X_{dyn}, and X_{sim}. As these variables are being replaced by a deterministic value, the joint statistical distribution of the updated model uncertainties $X_L = X_{exp}X_{st}X_{load}X_{ext}$ will have a smaller variance compared to the joint distribution of the uncertainty variables suggested by Tarp-Johansen et al., 2002 [79], which is $X_0 = X_{exp}X_{st}X_{aero}X_{dyn}X_{str}X_{sim}X_{ext}$. Figure 5.7 shows the probability density function of X_L compared to the probability density of X_0.

The set of environmental parameters used in the comparison of model to measurements includes horizontal and vertical wind shear assessed by combining measurements from a met mast and from a LIDAR instrument. As a result the exposure uncertainty is largely reduced and it is not considered to be part of the combined variable X_{load} discussed above. However, in wind farms the wind conditions might vary over the park area. Since for most wind turbine installation sites the climate is evaluated based on only one or a few masts, the wind conditions estimated from mast measurements might not be completely representative for the conditions experienced by some turbines placed relatively far away from the mast. This

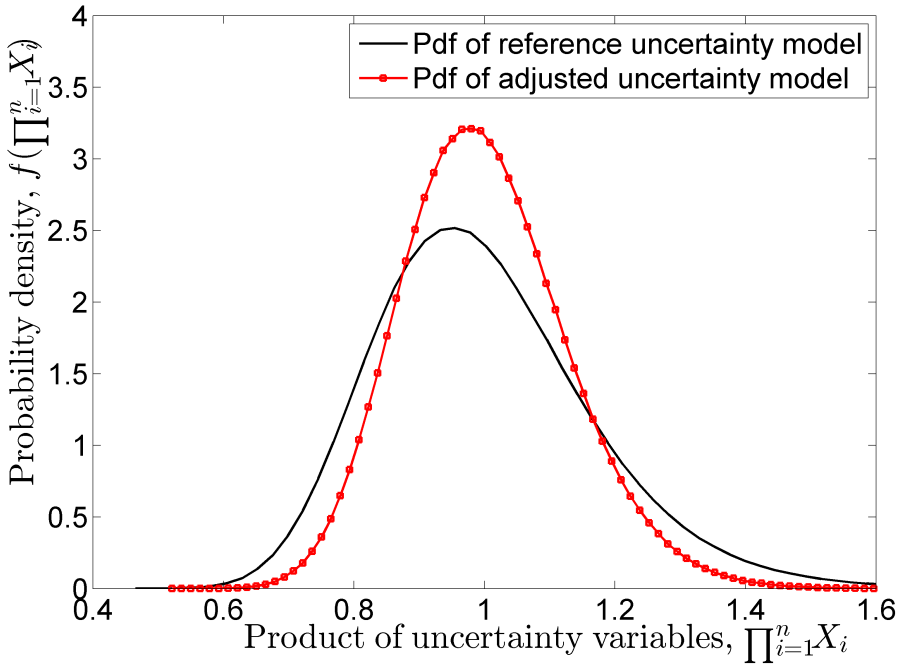

Figure 5.7: Comparison of updated model uncertainty to reference values prescribed in literature.

means that exposure uncertainty still has to be taken into account. X_{exp} is considered to be Lognormally distributed with mean of 1 and c.o.v of 0.1, the same values as assumed in [79].

Based on the observations and comparisons given above, the uncertainty model for ultimate limit state is updated. The distributions of the updated uncertainty variables are summarized in Table 5.3.

Influence of using a 30-second reference period

Figure 5.8 shows a comparison of the extreme loads obtained using the standard 10-minute reference periods to the load distributions obtained using 30-second reference periods. Both distributions are based on the same data set of 2200 simulations, where 100 simulations with ten-minute duration are carried out for each wind speed from cut-in to cut-out. Only the global extreme from each simulation is considered with the 10-minute reference period, while for the 30-second reference period a total of 20 maxima per simulation are taken. For the sake of comparison the quantile values for the data with 30-second resolution shown on Figure 5.8 are chosen as equivalent to 50% and 98% quantiles in a 10-minute period. The shorter reference time period does not introduce any significant change in the load

Table 5.3: Modelling uncertainties in estimating ultimate loads. Coefficients of variation: "TJ" as defined by Tarp-Johansen et al. [79]; "Adjusted" represents updated values.

Name	Description	Distribution	C.O.V		Mean
			TJ	Adjusted	
X_{exp}	Exposure	LN	0.1	0.1	$\mu = 1$
X_{st}	Wind statistics	LN	0.05	0.05	$\mu = 1$
X_{aero}	Airfoil coefficients	GB	0.1	-	$\mu = 1$
X_{dyn}	Dynamics	LN	0.05	-	$\mu = 1$
X_{sim}	Simulation	N	0.05	-	$\mu = 1$
X_{str}	Stress analysis	LN	0.03	0.03	$\mu = 1$
X_{ext}	Extrapolation	LN	0.05	0.05	$\mu = 1$
X_{load}	Combined load	Deterministic	-	0	$\mu = 1$

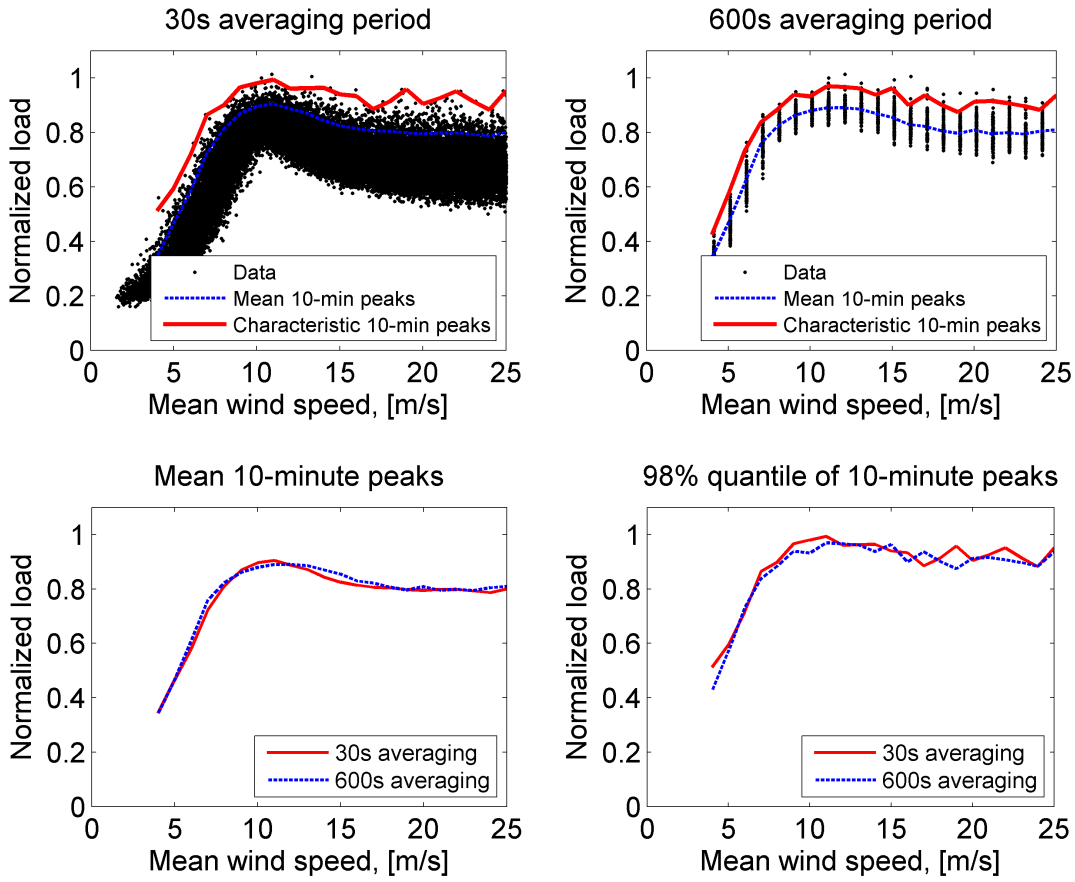

Figure 5.8: Extreme blade flapwise loads under different control strategies.

distribution - both the ten-minute-equivalent mean and characteristic (98% quantile) values remain practically unchanged. It is therefore not expected that shortening the reference time period to a value smaller than ten minutes will result in any change in the load distribution. In fact, as seen from Figure 5.8 the use of 30-second reference period instead of 10 minutes results in less correlation between the mean wind speed and the extreme load. This shows that the turbulence-related uncertainty comes not only from the long time-averaging period, but mostly from using a single reference point for describing the wind field. Considering this in the context of the calculation of the uncertainty variable X_{load} discussed in the previous paragraphs, if the discrepancy between measurements and model predictions has to be reduced further, it is necessary to employ detailed 2-dimensional representation of the actual flow field which the turbine experiences.

5.4 Uncertainties in fatigue limit state modelling

Some of the uncertainties associated with fatigue life of composites have already been discussed in the present thesis, where probabilistic models for material fatigue properties and number of stress cycles were demonstrated in Chapters 3 and 4 respectively. Material properties and cycle counts are however not the only source of uncertainty in fatigue life estimations. For most structures load amplitudes are also random, and there are other contributing factors as for example modelling, measurement, statistical, and numerical errors.

Fatigue life estimations are usually based on cycle-by-cycle calculations of incremental damage, which are then summed over the time history by applying a damage evolution model as for example Palmgren-Miner's rule, [58], [51]. There are many indications (Schön, 2002 [72], Nijssen, 2004 [57]) that Palmgren-Miner's rule is non-conservative when estimating fatigue life of composites subjected to variable amplitude spectrum loading. This has to be taken into account in a reliability analysis by defining a probabilistic model for the uncertainty in the accumulated damage at failure, as shown by Toft & Sørensen, 2011 [86], Toft et al., 2012 [84].

Works by Ronold et al., 1999, [69] and Toft et al., 2012 [84] are probably among the most complete studies on uncertainties in fatigue life modelling of composite structures, both taking into account several uncertainty effects, and demonstrating a procedure for tuning partial safety factors. There are however some details which are considered important by the author of the present thesis, and which are not included in the publications mentioned - the former study does not take into account the effect of the mean values of stress cycles

and the bias in Palmgren-Miner's sum, while the latter does not distinguish between the uncertainty in the S-N curves and in the damage accumulation rule, and does not consider the uncertainty in the number of stress cycles.

As the overview of literature references from the previous paragraphs and from Chapters 3 and 4 suggests, many aspects of probabilistic fatigue have been studied. Most studies however deal with only parts of the problem, and, more important, a distinction and detailed comparison between the significance of different sources of uncertainty has not been made in full. The relative significance of different uncertainty factors is a valuable knowledge for a designer, as it can indicate where efforts for reducing the uncertainty should be directed. In addition to the uncertainty in material parameters and in stress cycle counts, the following shows a demonstration of quantifying uncertainties caused by load amplitude estimation errors and errors in fatigue damage evolution modelling. This is followed by an overall comparison of the relative importance of the uncertainties for the fatigue life of a composite laminate.

5.4.1 Types of uncertainties in fatigue life modelling

Four types of uncertainty are considered, divided according to the stages of the analysis where their influence appears:

- Uncertainty in the number of load cycles
- Uncertainty in load amplitudes
- Uncertainty in fatigue properties of the material
- Uncertainty in the damage evolution model

The uncertainty in the number of load cycles takes into account the possible variation in number of cycles with certain amplitude and mean value which the structure will experience over its lifetime. It is an uncertainty caused by physical factors, and can be considered as aleatory. Load amplitudes and fatigue properties of the material (S-N curve parametrization) can be uncertain both due to physically occurring random variations, and due to modelling, measurement and statistical errors, thus carrying a combination of epistemic and aleatory uncertainty. The damage evolution model as its name suggests is mainly subject to modelling uncertainties. As a result the uncertainties associated with it can be considered epistemic, provided that the natural variation of fatigue properties in the material is fully accounted for by the model for uncertainties in the material properties. The uncertainties in the number

of load cycles over the lifetime of the structure, and in the material properties are discussed in the previous chapters of this thesis. In the following, the quantification process for the remaining uncertainty types mentioned above is illustrated with the help of models and test data.

5.4.2 Uncertainty in load amplitudes

Load cycles in fatigue are collected in bins, and the number of cycles in a given bin is a random number. Therefore the uncertainty in the amplitude of a given load cycle will be reflected into the rainflow matrix as an uncertainty to the location in the matrix where this particular cycle will be accumulated. This matter was already addressed in Section 4.3. The remaining uncertainty will be associated with how precisely the value of the load amplitude is determined, i.e., an epistemic uncertainty which can in principle be reduced by improving the quality of models and measurements. For wind turbine blades possible factors contributing to such an uncertainty in load predictions are errors in the estimation of the inflowing wind field (wind speed, turbulence, shear), airfoil coefficients, dynamic modelling, and others.

As discussed previously in Sections 4.3 and 5.2 of this thesis, the uncertainty of the mean value of a given sample tends asymptotically to zero as the number of samples tends to infinity. Fatigue loads are accumulated over the entire lifetime of the structure, and therefore the same dependence is valid. As a result it can be expected that the variance in the loads will not result in any significant additional uncertainty. What can still cause errors in the estimation of the fatigue damage is a possible bias in the load estimation due to steady-state errors.

The fatigue damage accumulated in a composite structural component over a ten-minute period is strongly correlated with the highest loads in this ten-minute period. This can be demonstrated by analyzing the 10000 simulated time-series which are also used for determining the fatigue load spectra shown in section 4.3. For each time series a relative fatigue damage estimate is determined using rainflow counting and a fatigue life law described by the constant-life diagram shown later on Figure 5.9. Then, the correlation coefficient between the logarithm of the damage variable and time series statistics is determined as a measure of dependence. Table 5.4 lists the correlation coefficients determined in this way. The ten-minute maxima of measured loads show highest correlation with fatigue damage. The reason for this behaviour is the relatively high slopes of the $S-N$ curve characteristic for composite materials, typically $m \approx 10-20$. A high value of the slope m means that

Table 5.4: Correlation between time series statistics and accumulated fatigue damage per series.

Type of statistics	Correlation with damage
Mean	0.875
Standard deviation	0.851
Maximum (absolute values)	0.955

relatively few large load ranges have the most significant contribution to the accumulated fatigue damage. Therefore the maximum loads per 10-minute period will actually give a realistic measure of the uncertainty in the fatigue damage.

In Section 5.3.3 it was observed that the ratio between the measured maximum moment and the maximum moment predicted by the load model under the same conditions, $X_{load} = M_{max,measurement}/M_{max,model}$, has a mean value of approximately 1. Therefore it can be concluded that there is no significant uncertainty in long-term fatigue loads (amplitudes and cumulative numbers of cycles) due to the time- and space- averaging of the wind field.

X_{load} explains only part of the total uncertainty in load modelling, related mainly to the averaging of the turbulence field, and airfoil coefficients and dynamic modelling. Some other errors as for example in the terrain-related variations of the wind field, or stress calculations might also have a bias. Therefore, the overall uncertainty with relation to load modelling is given by

$$X_L = X_{exp} X_{st} X_{load} X_{str} X_{geom} \tag{5.5}$$

where X_{exp} accounts for uncertainties in the actual wind field due to terrain variation and obstacles, X_{st} is the statistical uncertainty connected with the evaluation of the wind climate, X_{str} takes into account possible inaccuracies in determining the stress distribution in the structure, and X_{geom} describes possible variations in structural geometry such as plate thickness, which will cause variations in the stresses. Table 5.5 lists the distribution properties for all parameters. X_{geom} has the normal distribution with mean of 1 and variance 0.03, which is the value obtained by analyzing thickness data for test specimens of the type "MD2" from the OptiDat database [56]. Including X_{geom} in a reliability analysis is though only necessary if the limit state equation does not consider itself variations in material thickness. The distributions of X_{exp}, X_{st} and X_{str} are the same as the ones used for describing uncertainty in ultimate loads. The extrapolation uncertainty X_{ext} is not included in the fatigue calculations, because it reflects the uncertainty in the value of cycles with very long

Table 5.5: Model uncertainty parameters for fatigue load model.

Parameter	Distribution	Mean	C.o.v.
X_{load}	Deterministic	1	-
X_{exp}	Lognormal	1	0.1
X_{st}	Lognormal	1	0.05
X_{str}	Lognormal	1	0.03
X_{geom}	Normal	1	0.03

recurrence periods measured in years, while the fatigue life is influenced by a wider range of load levels most of which occur more frequently.

5.4.3 Uncertainty in material fatigue properties

Uncertainty in material fatigue properties was already discussed in chapter 3 of this thesis, where probabilistic models for estimating the fatigue life of laminates and sandwich cores were developed. The variance in these models corresponds to the uncertainty associated with variations in the material fatigue properties.

5.4.4 Uncertainty in damage evolution

As stated in section 3.3.5, the damage contribution from load cycles is calculated using Palmgren-Miner's rule:

$$\sum_i \frac{n_i}{N(S_i)} = 1 \tag{5.6}$$

where n_i is the actual number of cycles in the i^{th} range bin S_i as predicted by the fatigue load model, and $N(S_i)$ is the number of cycles to failure under constant-amplitude load S_i as predicted by the fatigue life model.

A number of literature sources (see Nijssen, 2006 [57] for an extensive list of references on the topic) state that Palmgren-Miner's rule is non-conservative when estimating fatigue life of composites subjected to variable amplitude spectrum loading. Toft, 2011 [86], using the same OptiDat database which the present study refers to, reports average ratio between actually observed and expected lifetime of ≈ 0.3, and provides an estimate for the variance of this ratio. This result however cannot be directly used here, because the model from [86]

represents the overall variance as a single value which takes into account uncertainties in the S-N relationship as well as variations in the damage evolution. The present study uses probabilistic models that already take into account the uncertainty in the S-N fatigue law under constant-amplitude loading, therefore any additional variance added to the fatigue life model should only account for the uncertainty related to the damage accumulation law. What is needed is assessing the uncertainty associated with using a linear damage accumulation law for predicting fatigue life of composite specimens subjected to variable-amplitude loading, for the case when probabilistic constant-amplitude fatigue life models are used. The "MD2" laminate from the OptiDat [56] database is chosen as a basis for this study, because a large number of constant-amplitude as well as variable-amplitude test data are available for this layup.

The OptiDat public database [56] provides an extensive amount of static and fatigue test data for glass-fiber reinforced laminates and can be used as the source of reference observations for inferring the model parameters. The laminate named "MD2" has a layup which resembles laminates often used in wind turbine blades - a mixture of UD (0°) and biaxial (±45°) fabrics and a total of 14 layers: $[\pm 45°0° \pm 45°0° \pm 45°0° \pm 45°0° \pm 45°]$. A large number of constant-amplitude data with different R-ratios are available for this laminate, providing possibility for defining a constant-life diagram(CLD). The CLD is a model predicting the damage increments (the number of cycles to failure) as function of cycle amplitude and mean values, by interpolating between S-N curves obtained from constant-amplitude tests with several different load ratios. A constant-life diagram for the MD2 laminate is constructed using test data for 7 load ratios $R = [-2.5, -1, -0.4, 0.1, 0.5, 2, 10]$. A probabilistic S-N curve is obtained for each of the 7 data sets using the Bayesian inference method mentioned above. Table 5.6 gives the parameters of the 7 S-N curves. The constant-life diagram obtained in this way is shown on Figure 5.9.

Variable-amplitude tests on the MD2 laminate are carried out using the WISPER and WISPERX standardized load sequences (ten Have, 1992, [81], [80]). Using the rainflow count of the spectrum sequences, the constant-life diagram from Figure 5.9, and applying damage summation by Palmgren-Miner's rule, the number of spectrum repetitions (passes) which will lead to specimen failure are determined. Model results are then compared to actual test outcomes by the ratio $\Delta N_{passes} = N_{measured}(\sigma)/N_{model}(\sigma)$, where N is the number of passes, and σ is the maximum stress range reached in the test sequence. The results for 15 tests with WISPER and 13 with the WISPERX sequences are shown on Figure 5.10. The average of ΔN_{passes} is $\mu_{\Delta_N} = 0.33$, a result which agrees well with Nijssen, 2006 [57] and Toft, 2011

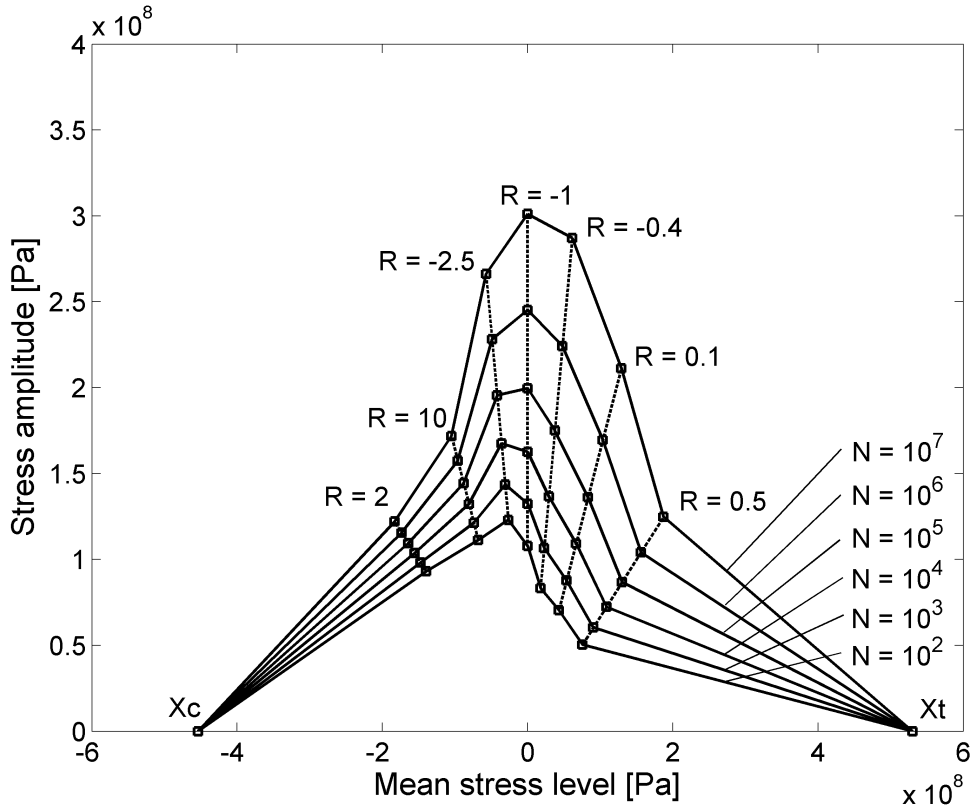

Figure 5.9: Constant life diagram describing fatigue life for different load ratios.

Table 5.6: Parameters of probabilistic S-N curves used to construct a constant-life diagram for the MD2 laminate.

R-ratio	Su		q		m	
	Mean	C.o.v.	Mean	C.o.v.	Mean	C.o.v.
-2.5	454MPa	0.053	1.141	0.6	14.91	0.1
-1	454MPa	0.053	1.000	0	11.21	0.033
-0.4	530MPa	0.047	1.269	0.37	9.30	0.067
0.1	530MPa	0.047	1.374	0.133	10.48	0.035
0.5	530MPa	0.047	1.355	0.315	12.67	0.090
2	454MPa	0.053	1.199	0.959	42.3	0.588
10	454MPa	0.053	1.000	0	26.5	0.037

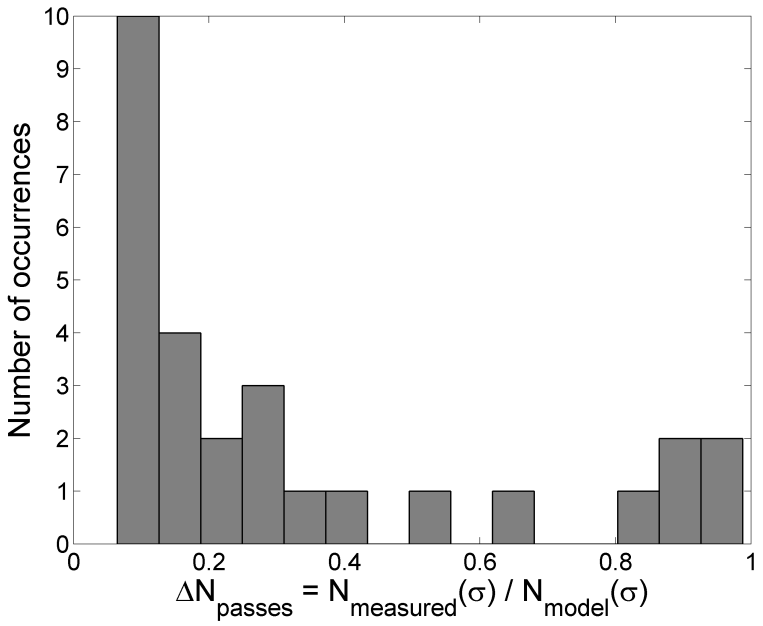

Figure 5.10: Histogram of observed damage to failure.

[86]. In order to match the mean predictions of the fatigue model with the mean of the test outcomes, the model is modified by introducing a proportionality coefficient, equal to μ_{Δ_N}. As shown on Figure 5.11 a), this is equivalent to shifting the S-N curve of the fatigue model towards lower fatigue lifetime. With the introduction of the proportionality coefficient a good agreement between modelled and measured mean values is obtained. What remains to be considered is the variance of Δ_N.

The constant-amplitude S-N curves that constitute the CLD shown on Figure 5.9 are all probabilistic. As a result model predictions for the number of spectrum repetitions until failure under variable-amplitude loading will also vary randomly. Figure 5.11 b) shows a comparison between the variance of variable-amplitude test data, and the variance of the fatigue model modified by multiplying the number of passes of failure with the proportionality coefficient μ_{Δ_N}. Visually the model variance matches well with the observed variance from data. This can be also seen on Figure 5.12, where the distribution of the deviation between the mean S-N curves and actual outcomes, $\varepsilon_N = \log N(\sigma) - \mu_{\log N}(\sigma)$, is plotted for both the test data and the modified model. The standard deviation of ε_N is $\sigma_{\varepsilon_N,data} = 0.80$ for the measured data, and $\sigma_{\varepsilon_N,model} = 0.89$. This shows that the variance caused by the randomness in the constant-life diagram results in model variance at least as large as what is seen in measurement data. It is therefore not necessary to include additional sources of uncertainty

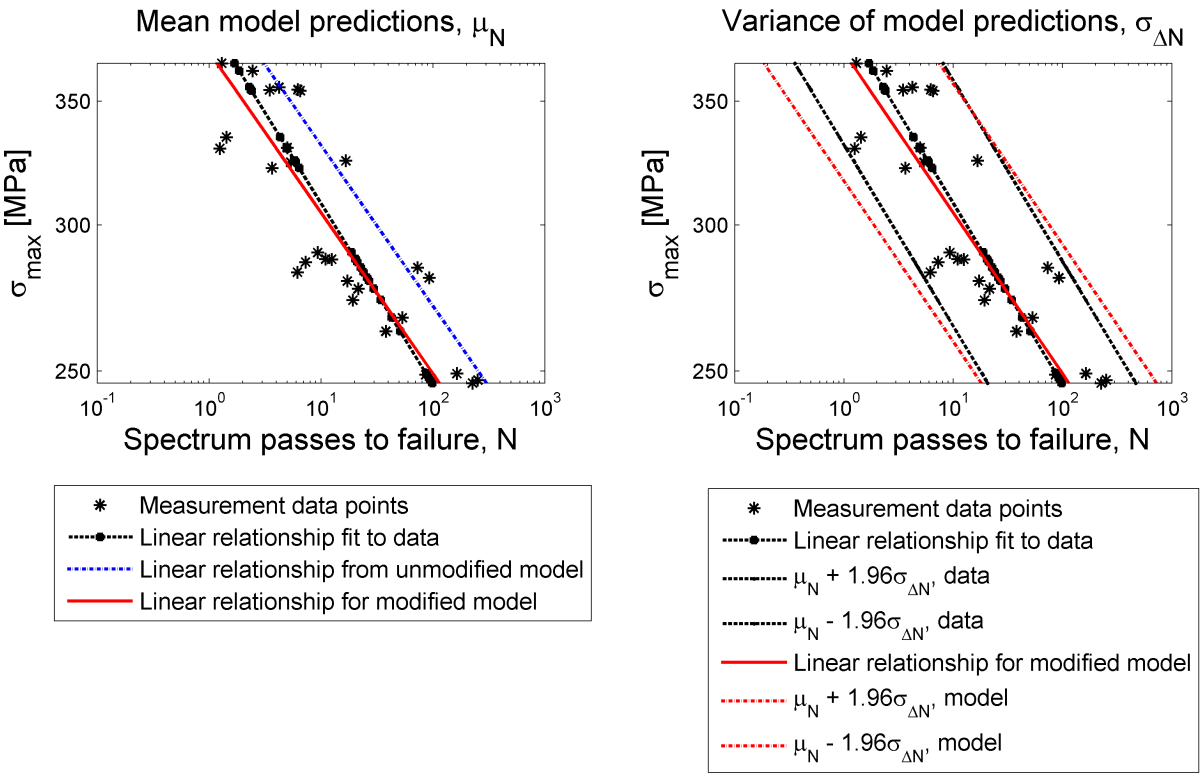

Figure 5.11: Adjusting the fatigue life prediction model by shifting the S-N curve.

in the model, and the uncertainty in the damage accumulation law can simply be accounted for by using a deterministic proportionality coefficient, $X_{PMR} = 0.33$.

The main conclusions from studying fatigue life of laminates subjected to variable-amplitude loading are that Palmgren-Miner's rule applied to composite materials under variable-amplitude loading produces non-conservative estimates, and that using a probabilistic constant-life diagram together with a proportionality factor results in a model whose predictions have variance very similar to the one observed in tests. While one can expect these findings to be also valid at least to some extent in other cases with similar loading and material conditions, their overall validity is not guaranteed. The lifetime predictions under variable amplitude loading are very sensitive to the parameters used in the constant-life diagram, and the type of spectrum that is used in the tests can also have a significant influence on the outcomes.

While realizing the shortcomings of the presently described uncertainty model, it is not possible to substantially improve the findings only based on the limited data available even if other models for damage evolution or for fatigue life under constant-amplitude loading are

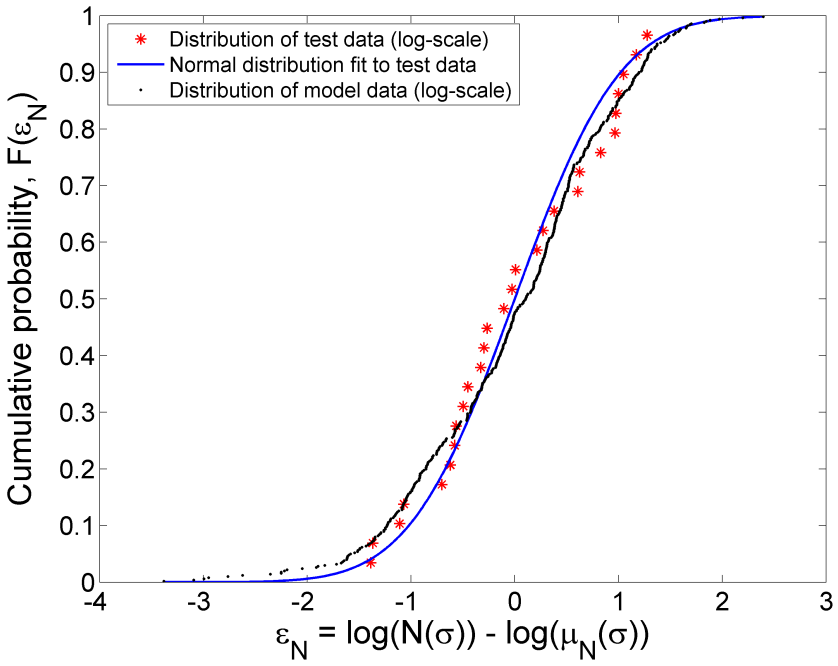

Figure 5.12: Cumulative distribution of accumulated damage at failure: comparison between data and adjusted model.

applied. It is not possible to avoid the natural scatter in fatigue properties, and most other models for damage evolution (see Nijssen, 2006 [57]) can result in similarly large errors and uncertainties, while being more complex than Palmgren-Miner's rule. Therefore the combination of probabilistic constant-life diagram, Palmgren-Miner's rule and a proportionality coefficient applied to the fatigue damage is chosen as the procedure that will be used in the present thesis for determining the fatigue life of laminates subjected to variable-amplitude loading. It is expected that the model will perform reasonably well in the example study considered in Chapter 6, because the sandwich structure which is the subject of analysis has layup and material almost identical to the test laminate discussed here, and the loading spectrum is similarly based on wind turbine blade loads.

5.4.5 Influence of model uncertainties on total accumulated lifetime damage

The following illustrates the influence of the presently discussed uncertainties on the variability in the total accumulated damage over a structure's lifetime. For the purpose a number of simulations are carried out where model parameters and variables associated with one of the uncertainty factors are varied randomly according to their statistical distributions, while all other quantities are kept constant. The constant life diagram for the MD2 laminate is used for obtaining the uncertainty related to material properties, while varying the fatigue model parameters listed in Table 5.6, and 4 elastic constants - the elastic moduli in axial and transverse direction E_1 and E_2, the Poisson ratio ν_{xy}, and the shear modulus G_{12}. The uncertainty related to load cycles is determined by taking random realizations of the model shown on Figure 4.28. The uncertainty in load amplitudes is evaluated in a similar fashion by taking random samples of the parameter X_L. As the uncertainty in Palmgren-Miner's rule is modelled as deterministic coefficient, it is not considered in this comparison. 10000 simulation samples are generated for each of the three cases mentioned, and the total lifetime damage is calculated for every sample. The overall stress amplitudes are tuned in a way that the total damage using the mean values of all model variables and parameters equals 1. The resulting empirical probability densities are shown on Figure 5.13.

5.5 Discussion on uncertainty modelling

The present chapter demonstrated how random effects and uncertainties influencing fatigue and ultimate limit states can be taken into account. The uncertainty in the load estimations

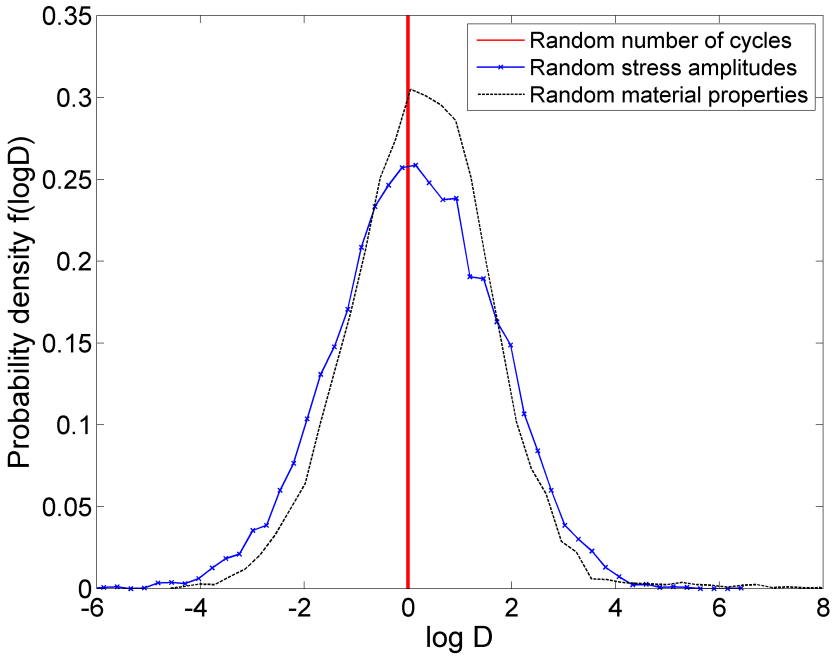

Figure 5.13: Comparison of influence of uncertainties on total accumulated lifetime damage.

defined in the present analysis is smaller than the reference values by Tarp-Johansen et al., 2002 [79]. A possible explanation is that the present analysis uses an advanced load simulation environment, which might allow for more precise modelling of turbine response, compared to what have been available to the authors of [79] ten years ago.

One of the clear conclusions that can be drawn from the study on uncertainties is that some aleatory uncertainties such as the number of fatigue stress cycles which were discussed in Chapter 4, and the uncertainty in load amplitudes due to the averaging of the wind field, are dependent on the design period, and practically vanish for longer time periods. This does not happen to uncertainties such as material fatigue properties, which are considered as constant over the entire lifetime of the structure. On the contrary, disregarding the material degradation over time can actually increase the uncertainty by adding a modelling error to it.

A direct comparison of the variance in total lifetime fatigue damage showed that uncertainties in material properties and in load amplitudes have almost equal influence. The scatter in material properties is a naturally occurring phenomenon and the variance of fatigue life associated with it can only be reduced by improving the materials. As discussed above, the aleatory part of the uncertainty in the loads becomes minimal over very long time

periods, and as a result the remaining uncertainty is mostly epistemic in nature. Improving the models can thus help reduce the model uncertainty. Changing the models will of course not improve the actual reliability of the physical structure, however it will reduce the penalty on the estimated reliability caused by not having sufficient knowledge. As a result, reducing model uncertainty can ultimately lead to more optimal design.

The uncertainty in Palmgren-Miner's rule is probably the variable whose properties are most difficult to explain, because it is the only influence factor which was found to have a significant bias, i.e., its mean value does not approach 1. At first glance, there are at least two factors that could be the possible causes of this behaviour. The first of them is the fact that Palmgren-Miner's rule is invariant to the sequence in which load cycles occur, and material degradation is not taken into account. An argument against this is that a given large cycle could initiate some damage in the form of micro-cracks, the result of which is that subsequent cycles will result in faster damage progression compared to what is expected from fully intact material. As S-N curves are typically calibrated for cycle amplitudes well below the yield strength of the material, the constant-amplitude tests would be unable to reveal the effect of large stress cycles. Another possible explanation of the bias is the use of rainflow (or any other cycle counting) algorithm in order to convert the random, variable-amplitude time series into a damage-equivalent set of cycles with defined amplitudes. Instead of being damage-equivalent, the resulting fatigue spectrum might be introducing a bias by inaccurately considering as cycles some features in the time series that do not have actual influence on the lifetime. Discussing these problems in more detail and finding a physical model which accurately explains this behaviour is a demanding task which could not take place as part of the present thesis as it would consume time resources needed for carrying out other important parts of the study. Therefore the solution chosen was to introduce a correction coefficient that will compensate the bias in the fatigue damage model.

Chapter 6

Reliability Analysis

The present chapter demonstrates two examples of estimating the reliability of a wind turbine blade. The first example considers the entire blade and estimates the probability of blade-tower collision, while the second example estimates the reliability against fatigue and ultimate material failure of a sandwich panel which is part of the aerodynamic shell and is located close to the blade root. Both examples use the same load distributions for a 2.3MW wind turbine defined in Chapter 4 of this thesis. With these two examples, three failure modes which are typically part of the standard blade design assessment are covered - fatigue failure, blade-tower collision, and buckling-driven ultimate material failure. The majority of commercially produced wind turbines are designed according to the guidelines given in standards such as the IEC61400-1 [37] or its equivalents. Whenever possible, the analysis is based on the conditions given in the IEC standard, which gives the opportunity for comparison between the reliability model and the deterministic design equations using partial safety factors defined in the standard.

6.1 Reliability against blade-tower collision

6.1.1 Limit state formulation

The limit state function $g(\mathbf{X})$ has the typical demand-capacity type definition discussed in Chapter 2. When considering blade-tower collision as a failure mode, the demand will be the blade deflection, while the capacity will be the available blade-tower clearance for undeformed blade. The equation will have the form

$$g(\mathbf{X}) = X_{capacity}R - X_{demand}S \tag{6.1}$$

where S is the demand (deflection), R is the capacity (available clearance), and X_{demand} and $X_{capacity}$ are random factors accounting for the uncertainties associated with modelling the demand and capacity, respectively.

6.1.2 Uncertainty modelling

Model uncertainty variables X_{demand} and $X_{capacity}$ have to account for uncertainties that are not already included in the calculation of the demand and capacity itself. The modelled capacity, or in other words the available blade-tower clearance, can be influenced by deviations in the geometry of the structure - for example the precision of the alignment between blade root and hub flange surfaces can have an influence on the blade coning angle and hence on the available blade-tower clearance. Therefore $X_{capacity} = X_{geom}$, where X_{geom} is a random variable accounting for deviations in turbine geometry. Based on engineering judgement X_{geom} is considered as normally distributed with mean of 1 and a C.O.V (coefficient of variation) of 0.03.

The uncertainty in load evaluation X_{demand} is taken as defined in Chapter 5. Two cases are considered - the values referred to as "low" uncertainty model by Tarp-Johansen et. al, 2002 [79], and the adjusted values defined in the present thesis:

$$\begin{aligned} \text{Reference:} \quad X_{demand} &= X_{exp}X_{st}X_{aero}X_{dyn}X_{str}X_{sim}X_{ext} \\ \text{Adjusted:} \quad X_{demand} &= X_{exp}X_{st}X_{load}X_{str}X_{ext} \end{aligned} \tag{6.2}$$

Table 6.1 lists the distribution parameters for X_{demand} and $X_{capacity}$.

6.1.3 Design equation

The design equation that corresponds to the limit state equation 6.1 and complies with the requirements in the IEC61400-1 standard is

$$R_c \geq \gamma_n \gamma_m \gamma_f S_c \tag{6.3}$$

where

R_c is the characteristic capacity;

S_c is the characteristic demand;

γ_n is a partial safety factor accounting for consequences of failure;

Table 6.1: Modelling uncertainties in estimating extreme design loads. Coefficients of variation: "TJ" as defined by Tarp-Johansen et al., 2002 [79]; "Adjusted" represents updated values.

Type	Name	Description	Distribution	C.O.V		Mean
				TJ	Adjusted	
Capacity	X_{geom}	Geometry	N	-	0.03	$\mu = 1$
Demand	X_{exp}	Exposure	LN	0.1	0.1	$\mu = 1$
	X_{st}	Statistical	LN	0.05	0.05	$\mu = 1$
	X_{aero}	Airfoil coefficients	GB	0.1	-	$\mu = 1$
	X_{dyn}	Dynamics	LN	0.05	-	$\mu = 1$
	X_{sim}	Simulation	N	0.05	-	$\mu = 1$
	X_{str}	Stress analysis	LN	0.03	0.03	$\mu = 1$
	X_{ext}	Extrapolation	LN	0.05	0.05	$\mu = 1$
	X_{load}	Combined load	Deterministic	-	-	$\mu = 1$

γ_m is a partial safety factor accounting for uncertainty in materials (blade-to-blade variation)

γ_f is a partial safety factor accounting for uncertainty in loads

According to IEC61400-1, the values of the partial safety factors for calculations involving statistical extrapolation are $\gamma_f = 1.25$, $\gamma_m = 1.1$, and $\gamma_n = 1$, where the value of γ_n is the one for non-fail-safe structural components. Typically, the characteristic values are given as the 5% quantile for the capacity, and the 98% for the loads. In the case of blade-tower clearance however the design equation uses $R_c = \mu_R$, and the partial safety factor for materials is applied to the demand side of equation 6.3.

6.1.4 Results from reliability analysis of blade-tower collision

Reliability against blade-tower collision is calculated based on the extrapolation function fit to the Monte Carlo sample with four random variables discussed previously in Chapter 4. Two limit state equations are used, based on the two uncertainty models given in equation

Table 6.2: Results from reliability analysis based on target design values determined using IEC61400-1 safety factors.

Uncertainty type	Reference	Modified
Reliability Estimates		
Annual probability of failure $P_f \cdot year^{-1}$	0.1008	$1.348 \cdot 10^{-6}$
Reliability index	1.2772	4.6928
Importance factors		
X_{geom}	0.13	0.22
X_{exp}	-0.43	-0.68
X_{st}	-0.21	-0.34
X_{aero}	-0.79	-
X_{dyn}	-0.21	-
X_{str}	-0.13	-0.21
X_{ext}	-0.21	-0.31
u_B	-0.17	-0.48

6.2:

$$g(\mathbf{X}) = X_{geom} d_{limit} - X_{exp} X_{st} X_{aero} X_{dyn} X_{str} X_{sim} X_{ext} u_B \quad \text{(non-modified uncertainty)}$$

$$g(\mathbf{X}) = X_{geom} d_{limit} - X_{exp} X_{str} X_{ext} X_{load} u_B \quad \text{(modified uncertainty)} \tag{6.4}$$

where d_{limit} is the available clearance, and u_B is the blade deflection at blade-tower passage events estimated by the extrapolation function. A number of reliability calculations using different values of d_{limit} are made using the FORM method. Estimated probabilities of failure are given on Figure 6.1 as functions of the design blade-tower distance at rest. The figure reveals that the uncertainty model has very significant effect on the reliability estimates, as the estimated reliability for the case with modified uncertainties is several orders of magnitude higher than the reliability with non-modified uncertainties.

A second reliability calculation is carried out for a problem where the design clearance, $d_{design} = 2.185$ is determined by solving equation 6.3 using the partial safety factors recommended by IEC61400-1, $d_{design} = \gamma_n \gamma_m \gamma_f D_{char} = 1.0 \cdot 1.1 \cdot 1.25 D_{char} = 1.375 D_{char}$. The design clearance is then included in the limit state functions in equation 6.4. A reliability analysis with this settings results in an estimation about what is the reliability of a structure

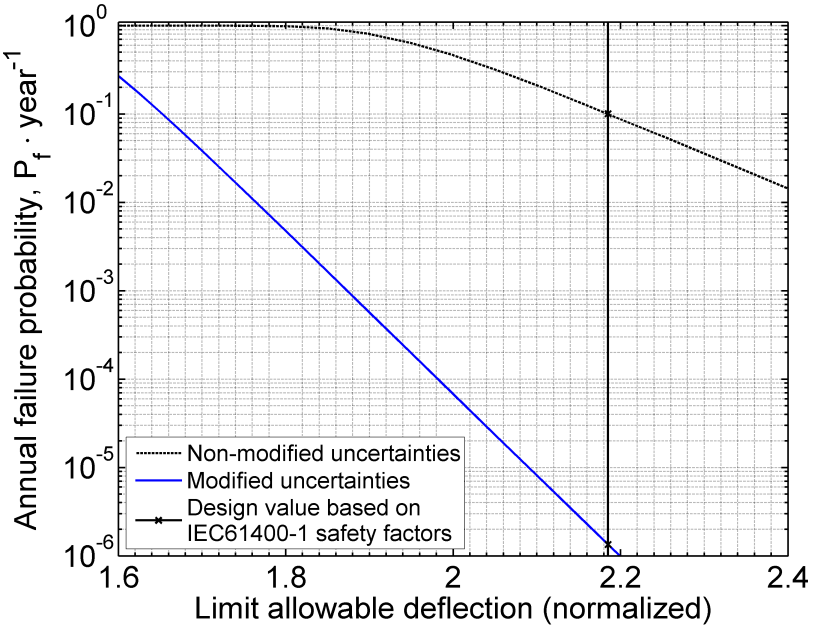

Figure 6.1: Blade-tower collision probabilities based on FORM analysis.

designed according to the IEC standard. Table 6.2 lists the reliability estimates from these calculations, together with the importance factors of the stochastic variables.

6.1.5 Discussion on reliability against blade-tower collision

Figure 6.1 and table 6.2 illustrate what is probably the most noticeable result from the reliability analysis, which is that the choice of model uncertainties completely dominates the outcome of the analysis. The considerable efforts that have been put in obtaining precise estimates of the long-term distribution of blade deflection have little effect on the final result from the analysis, as most of the uncertainty parameters considered have higher influence on the model than the actual load distribution. As visible from table 6.2, for analysis using the non-modified set of model uncertainties the importance factor of the blade deflection u_B is among the lowest. Using the modified uncertainty model results in u_B having the second highest importance factor, however the uncertainty variable X_{exp} still has a higher importance. Toft [82] has reported similar findings that model uncertainties have very large influence on the reliability.

6.2 Reliability of a sandwich panel subjected to compression and transverse pressure

6.2.1 Panel description

In the present example the reliability analysis of a sandwich panel is demonstrated. The panel under consideration represents the geometry and material composition typical for the aerodynamic shells on a wind turbine blade in the area close to the blade root (Figure 6.2). In this particular location the outer shell of the blade is subjected to high loads, because the central spar usually does not span along the whole blade length. Therefore in the first several meters from the root of the blade the entire load is carried by the outer shell, which is specifically reinforced for that purpose. The specifics of the structure and the high bending moments in this area make it one of the most likely failure locations, and as a result it is a good choice for carrying out reliability analysis.

The part of the shell considered in the present analysis is placed on the downwind side of the blade. It has rectangular shape, with constant curvature of $R = 1.2m$ along its short side, and a constant twist of $2deg/m$ over its length. The dimensions of the panel are $a = 2m$ and $b = 0.5m$ (Figure 6.3). The panel is loaded by in-plane compressive force Q_A, and a transverse pressure Q_T. More details about the calculation of Q_A and Q_T are given below in section 6.2.2.

The sandwich panel has glass-epoxy face laminates, and balsa core. The two face laminates have identical, symmetric layups with 8 layers, $[\pm 45^\circ\ 0^\circ_4\ \pm 45^\circ]$, and average thickness of $4.3mm$. The average core thickness is $20mm$. Layer thicknesses are considered as normally distributed random variables, with distributions given in table 6.3.

6.2.2 Load description

Based on the location of the sandwich panel close to the root of the blade it is assumed that the in-plane forces acting in the sandwich will be a function of the flapwise bending moment. As a simple approximation the in-plane force is assumed to be linearly dependent on the bending moment, and uniformly distributed along the width of the sandwich plate. Taking the long side of the sandwich plate as the x-coordinate, and the short one as y, the magnitude of the force per unit length N_x is found by

$$N_x[N/m] = (y_2 - y_1)\frac{M_y}{I_{yy}} \tag{6.5}$$

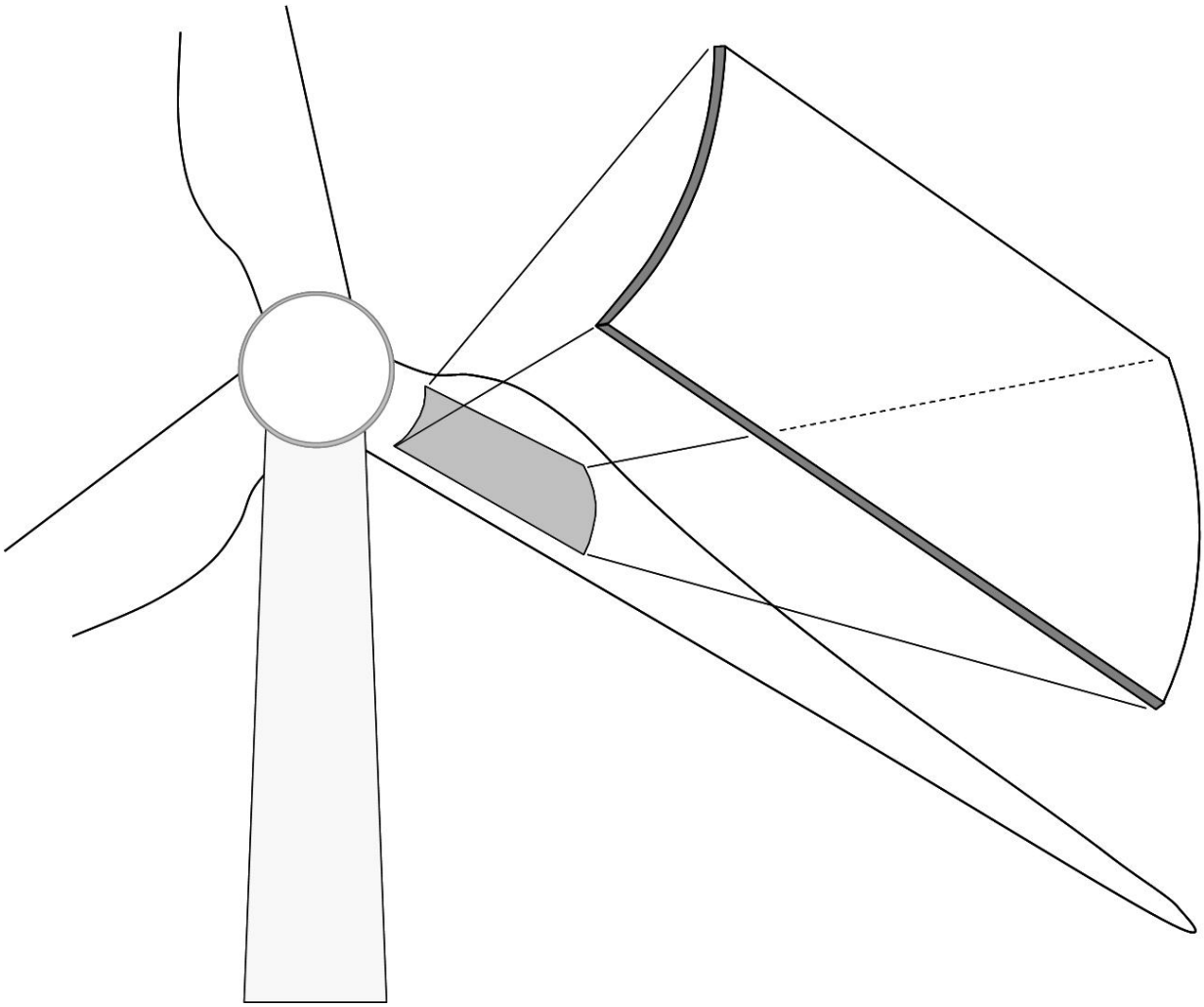

Figure 6.2: Location of panel.

Table 6.3: Sandwich plate layer orientations and thicknesses.

Location	Material type	Orientation	Mean thickness	C.o.v. of thickness
Upper face	Biax	$\pm 45°$	0.75mm	0.03
Upper face	4xUD	0°_4	2.8mm	0.03
Upper face	Biax	$\pm 45°$	0.75mm	0.03
Core	Balsa wood	$0°$	20mm	0.03
Lower face	Biax	$\pm 45°$	0.75mm	0.03
Lower face	4xUD	0°_4	2.8mm	0.03
Lower face	Biax	$\pm 45°$	0.75mm	0.03

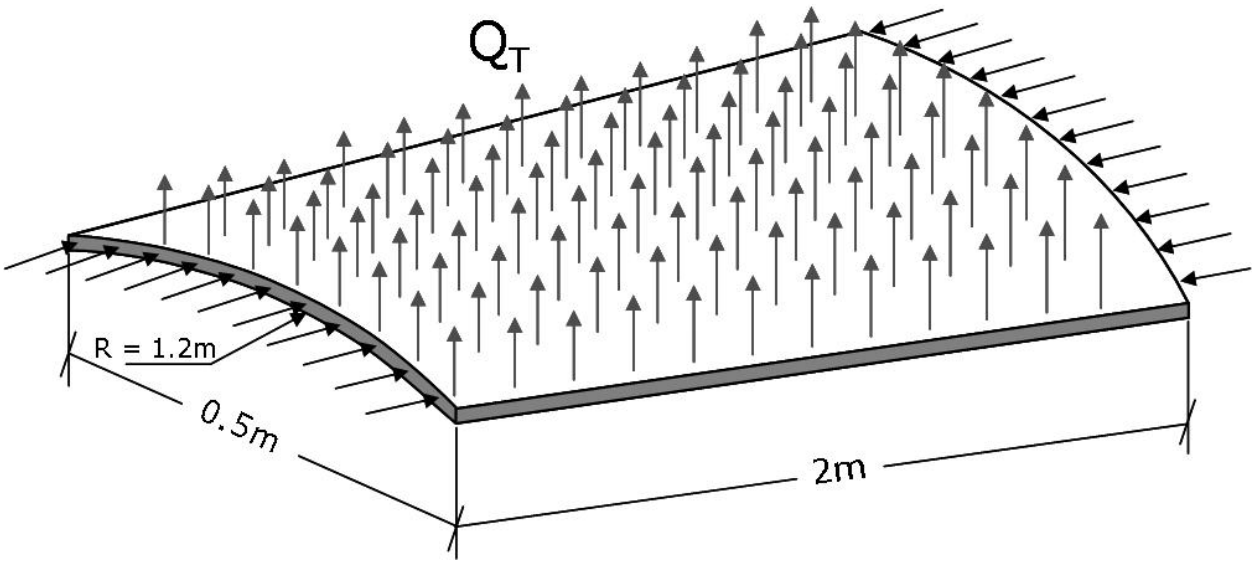

Figure 6.3: Plate geometry and loading.

where M_y is the bending moment, I_{yy} is the area moment of inertia for bending about the y-axis, and y is the y-coordinate of the point of the cross section where the stresses are evaluated. The value of y is taken as $y = R = 1.2$ and I_{yy} depends on the plate thickness, with a mean value $\mu(I_{yy}) = 0.1553m^4$.

Load magnitudes are based on the same 10000 simulations used for calculating the reliability against blade-tower collision in section 6.1. For the ultimate material failure limit state the statistical distribution of load extremes is obtained using the methods discussed in Chapter 4. The Peaks-over-threshold method is used to identify peaks, and the extrapolating distribution is the Weibull distribution. Fatigue cycle counts are determined by applying rainflow-counting on the 10,000 series, and subsequently extrapolating to a lifetime of 20 years. The resulting fatigue spectrum is shown in Figure 6.4.

The majority of the load cycles in the spectrum shown on Figure 6.4 are in compression. In order to simplify the calculations for the present example, the rainflow matrix is converted to a 1-dimensional spectrum assuming a constant load ratio of $R = \sigma_{min}/\sigma_{max} = 10$ which corresponds to purely compressive load cycles and is the ratio which is most representative for the actual load conditions. This will allow using a single S-N curve instead of a constant-life diagram.

The flapwise bending moment gives a real-world load case example for the in-plane loads. However, there are no similar data available for the transverse pressure. The approximate upper limit of the magnitude of the aerodynamic pressure can be estimated using the formula

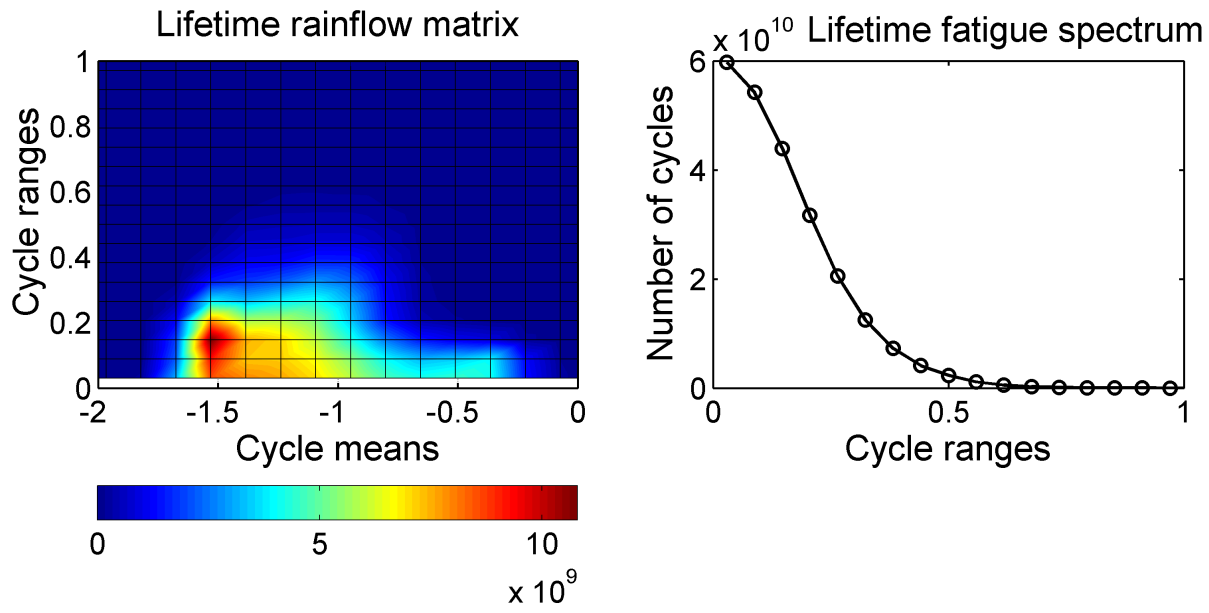

Figure 6.4: Probabilistic fatigue load model.

for the lift force:

$$L = \frac{1}{2}\rho V^2 A C_L \tag{6.6}$$

where ρ is air density, A is the airfoil area, C_L is the lift coefficient, and V is inflow speed. The maximum lift coefficient of an airfoil is usually in the range of $1.2 - 1.5$, while the maximum inflow speeds are up to about $V = 100m/s$, which is rather conservative as such flow speeds are more characteristic for the tip of the blade rather than the root. Using these input values, and an air density of $1.15kg/m^3$ will result in a maximum estimated pressure of up to $\approx 10000N/m^2$, which is significantly less than the in-plane loads caused by the bending moment. It is therefore expected that the transverse pressure will not have a significant impact on the reliability of the panel. A fixed value of $Q_T = 40000N/m^2$ is therefore assumed. This pressure will not have a direct influence on the reliability of the model, however its presence is useful when analyzing the buckling behaviour of the panel, as the pressure will provide a slight out-of-plane displacement that will initiate the buckling process.

6.2.3 Fatigue limit state

In the case of fatigue failure the capacity is usually taken as $D_{failure} = 1$, and the demand is the total accumulated damage over the lifetime of the structure, $D(\mathbf{X})$:

$$g(\mathbf{X}) = D_{failure} - D(\mathbf{X}) = 1 - D(\mathbf{X}) \tag{6.7}$$

The damage D is calculated using Palmgren-Miner's sum (equation 3.31). It is an exponential function of the stress - the stress variable S_i is raised to the power of the fatigue slope m. This exponentiation causes $D(\mathbf{X})$ to vary in wide ranges. The behaviour of the limit state equation is stabilized by raising $D(\mathbf{X})$ to the power of $1/m$. Another introduced modification is the replacement of the expression $1 - D(\mathbf{X})$ with $1/D(\mathbf{X}) - 1$. With this adjustment the limit state equation gets the form which is most suitable for using the Model Correction Factor method.

$$g(\mathbf{X}) = D(\mathbf{X})^{-\frac{1}{m}} - 1 \tag{6.8}$$

The limit state function has to also include the effects of model uncertainties. Adding the previously defined (see Chapter 5) uncertainty terms $X_L = X_{exp}X_{st}X_{load}X_{str}$ and X_{PMR} to equation 6.8 results in the following

$$g(\mathbf{X}) = \left[\frac{1}{X_{PMR}} D(\mathbf{X}, X_L)\right]^{-\frac{1}{m}} - 1 \tag{6.9}$$

The load uncertainty term is included in the model by multiplying the load variables Q_A and Q_T with X_L.

Partial safety factors for fatigue limit state

When failure is described in terms of a non-dimensional failure criterion, the criterion value will be function of both the material properties and the loads, and the limit state equation is usually not split into demand and capacity parts (or, the ultimate criterion value of 1 can be considered as a deterministic "capacity"). In such cases it is convenient to use a single safety factor value. In case of rule-based fatigue design partial safety factors γ_f and γ_n are usually applied as a single value $\gamma = \gamma_f \cdot \gamma_m$ multiplied to the stress levels S_i (see for example the IEC 61400-1 standard, [37]). Using the limit state equation 6.9, the corresponding design equation including the safety factor γ and a design parameter z is

$$g = D(\mathbf{X_c}, (\gamma z \mathbf{S_c}))^{-\frac{1}{m_c}} - 1 \tag{6.10}$$

where the subscript c denotes characteristic values. It should be noted that the load levels $\mathbf{S} = [S_1, S_2, \cdots]$ are also a part of the vector of variables $\mathbf{X}$, but here $\mathbf{S}$ is considered separately as these are the only values in $\mathbf{X}$ which are multiplied by the safety factor γ. Writing out in full the formula for accumulated damage, equation 6.10 becomes

$$g+1 = \left(\sum \frac{n_i}{N(\mathbf{X_c}, \gamma z S_{i,c})}\right)^{-\frac{1}{m_c}} = \left(\sum \frac{n_i \gamma z S_{i,c}^{m_c}}{(S_{u,c} q_c)^{m_c}}\right)^{-\frac{1}{m_c}} = \frac{1}{\gamma}\left(\sum \frac{z n_i S_{i,c}^{m_c}}{(S_{u,c} q_c)^{m_c}}\right)^{-\frac{1}{m_c}} \quad (6.11)$$

The variables S_u, m and q in equations 6.10 and 6.11 are the coefficients of the $S-N$ curve representing the fatigue life of the panel. As equation 6.11 shows, the safety factor γ is a common multiplier and can be taken out of the damage evaluation function. Therefore the value of γ can be estimated with the following:

$$g = \frac{1}{\gamma}\left(\frac{1}{X_{PMR}} D\big(\mathbf{X_c}, z\mathbf{S_c}\big)\right)^{-\frac{1}{m_c}} - 1 \quad (6.12)$$

All values in g outside the damage term are deterministic, and it is known that the design limit is achieved when $g = 0$, i.e., when $D^* = X_{PMR}$. Therefore, if $z = 1$, i.e., no tuning of the target reliability is required, the value of γ can be found by determining the characteristic value of the damage function $D(\mathbf{X_c}, \mathbf{S_c})^{1/m}$. The value of γ for a specific target reliability index is found by calibrating equation 6.9 to the desired reliability using the design parameter z, followed by solving the design equation 6.12 with the tuned value of z.

Based on a single design equation it is only possible to determine one safety factor at a time. If two or more distinct partial safety factors are present in the equation, there will be a possibly infinite number of combinations of partial safety factor values that satisfy the equality. Therefore, if it is required that the safety factors are split into two components, one of the partial safety factors need to be picked arbitrarily or determined by other means. Here it is chosen to split the overall safety factor γ into partial safety factors for material properties and load conditions, $\gamma = \gamma_m \cdot \gamma_f$. The value of γ_m is defined based on the difference between the characteristic $S-N$ curve and the $S-N$ curve at the design point, i.e., the ratio

$$\gamma_m = \frac{S(N_{ref}, \mathbf{X}_{char})}{S(N_{ref}, \mathbf{X}^*)} \quad (6.13)$$

where $S_{ref}(N_{ref}, \mathbf{X})$ is an $S-N$ fatigue law giving a relation between a reference number of cycles to failure, N_{ref}, and the stress level S_{ref} which is expected to cause failure after the reference number of cycles. $\mathbf{X}_{char}$ is the set of material properties defining the characteristic $S-N$ curve, and $\mathbf{X}^*$ is the design point from reliability analysis. The characteristic values

of material properties are taken as the values defining the characteristic S-N curve marking 95% survival probability, i.e., there is 5% possibility that the actual $S-N$ curve will result in lower fatigue life than the characteristic curve.

The load S_i is given by the product of the in-plane load Q_A and the model uncertainty X_L. Its characteristic value is defined using a characteristic model uncertainty $X_{L,char} = 1$, and the characteristic value of the in-plane load Q_A taken as its 98% quantile. Once the values of γ and γ_m have been found by equations 6.12 and 6.13 respectively, the remaining partial safety factor γ_f is found by $\gamma_f = \gamma/\gamma_m$. The value of γ_m accounts for the uncertainty in material properties, while γ_f combines the uncertainty in loads and geometry.

Limit state equation in terms of equivalent loads

The equivalent load $S_{eq}(\mathbf{X})$ defines the magnitude of an imaginary constant-amplitude loading which for a given reference number of cycles N_{ref} will result in exactly the same accumulated damage as what is obtained from applying the actual load spectrum:

$$S_{eq} = \left(\frac{\sum n_i S_i^m}{N_{ref}} \right)^{\frac{1}{m}} \tag{6.14}$$

Equivalent fatigue loads are often used as a design criterion rather than the accumulated damage. It can be therefore useful to express the limit state equation in terms of equivalent loads. An approach which leads to a limit state equation equivalent to equation 6.9 is replacing the damage expression $D^{-1/m}$ with the ratio of the actual equivalent load $S_{eq}(\mathbf{X}, D(\mathbf{X}))$ and the equivalent load which will lead to failure, $S_{eq}(\mathbf{X} \big| D(\mathbf{X}) = 1)$. Using this ratio, the limit state function takes the following form:

$$g(\mathbf{X}) = \frac{S_{eq}\Big\{\mathbf{X}, X_{PMR}, X_L \Big| \big((X_{PMR})^{-1} D(\mathbf{X}, X_L) = 1\big)\Big\}}{S_{eq}\Big\{\mathbf{X}, X_L, X_{PMR}, D(\mathbf{X}, X_L)\Big\}} - 1 \tag{6.15}$$

6.2.4 Ultimate strength limit state

The limit state equation for material strength-related failure of the panel is based on the ratio between the ultimate resistance of the panel and the load applied to it:

$$g(\mathbf{X}) = \frac{R(\mathbf{X})}{X_{exp} X_{st} X_{load} X_{str} X_{ext} S(\mathbf{X})} - 1 \tag{6.16}$$

The load is multiplied with model uncertainties $X_L = X_{exp} X_{st} X_{load} X_{str} X_{ext}$, the values of which are as defined in the column "Adjusted" in Table 5.3 in Chapter 5. Other uncertainties

such as material strength and stiffness, layer thicknesses, and distribution of load extremes are included in the vector of stochastic variables $\mathbf{X}$.

Partial safety factors for ultimate strength limit state

The load variable $S(\mathbf{X})$ is proportional to the axial load Q_A and to the model uncertainty X_L, while the resistance variable $R(\mathbf{X})$ is a function of several of the components in $\mathbf{X}$. The main contributing factors are the compressive strength and layer thickness, X_c and t_{UD}. The presence of at least three random variables that influence the limit state equation requires the use of a minimum of three partial safety factors, γ_f, γ_m, γ_g, corresponding to Q_A, X_c, and t_{UD}, respectively. Just as in the case of fatigue limit state discussed above, a single design equation makes it only possible to tune one partial safety factor at a time. The approach chosen for the present case is to apply a design parameter z to the ultimate compressive strength X_c. The value of γ_f is determined by $\gamma_f = (Q_A^* X_L^*)/(Q_{A,char} X_{L,char})$, while γ_g is taken equal to 1. The characteristic values of X_L and t_{UD} are taken as their mean values, $X_{L,char} = 1$ and $t_{UD,char} = E\langle t_{UD} \rangle$. Adopting such an approach results in the following limit state equation

$$g(\mathbf{X}) = \frac{R(z \cdot X_c, t_{UD}, \mathbf{X})}{X_{exp} X_{st} X_{load} X_{str} X_{ext} S(\mathbf{X})} - 1 \tag{6.17}$$

The design equation corresponding to the limit state equation 6.17 has the following form:

$$g = \frac{R(\frac{1}{\gamma_m} z \cdot X_{c,char}, t_{UD,char}, \mathbf{X}_{char})}{\gamma_f S(Q_{A,char}, \mathbf{X}_{char})} - 1 \tag{6.18}$$

For most of the calculations presented here, $z = 1$ and the limit state equation will represent the actual reliability of the panel. Tuned limit state equation with z different from 1 is used only for calibrating safety factors.

6.2.5 Annual vs. lifetime reliability

The structural reliability index is always based on a reference time period. The probability of failure obtained by a reliability analysis is defined as the likelihood of observing a failure within one reference time period. Fatigue damage is cumulative in nature, therefore it makes most sense to use the entire lifetime of the structure as a reference time period. For ultimate strength analysis however the reliability is usually given in terms of an annual reliability index, i.e., the reference period is one year.

Comparison of the reliability against failure in different limit states is only possible if the calculations are based on the same reference period. For the sake of comparison, a time frame of 20 years, equal to the design lifetime of the structure, is used in the present example for calculating the reliability against both fatigue and ultimate-strength failure. An exception are the calculations involving safety factor calibration for ultimate strength limit state, where a reference time period of one year is used. This ensures comparability of the partial safety factors derived from the reliability analysis to the safety factors recommended by standards and guidelines, where quoted safety factors for ultimate strength design are usually based on one year reference period. The reliability index based on a reference period equal to the entire lifetime of the structure will be referred to as "lifetime reliability", while the reliability with one-year reference period will be referred to as "annual reliability".

6.2.6 Stochastic variables

Probabilistic model parameters, uncertainties and input variables are all input as stochastic variables in the reliability analysis. Their statistical distributions and correlations were described in the previous sections of this study. Due to the layer-wise description used for sandwich faces, a considerable amount of stochastic variables is required. Table 6.4 summarizes all the variables necessary for defining the reliability model. Each of the sandwich components - upper face, lower face, and core, requires a slightly different set of variables. However, some of the variables referring to one of the sandwich components have to be used in the reliability analysis of other components, because they influence the reliability too. For example, the core thickness will have some influence on the reliability of the faces, because the stresses in the faces are dependent on core thickness.

6.2.7 Structural models

The primary model of the panel is a nonlinear finite element model built of 8-node layered shell elements in the commercial FE program ANSYS. A geometrically nonlinear analysis is carried out in a number of steps with gradually increasing loads. For each step the stresses are evaluated for each layer in each element of the model, and depending on whether a fatigue or ultimate strength limit state is considered, the results are used to evaluate failure criteria (ultimate strength limit state), or as input to a fatigue life model (fatigue limit state).

A second, simplified model is also implemented for the purpose of applying the Model Correction Factor method (Section 2.2.4. The simplified model of the panel is the analytical,

Table 6.4: Overview of stochastic variables included in reliability model.

Variable type	List	Variable count			
		Face 1	Face 2	Core	All
Lamina elastic constants	$[E_1, E_2, \nu_{12}, G_{12}]_i,\ i = 1, \ldots, N$	24	24	4	52
Lamina material strength	$[X_t, X_c, Y_t, Y_c, S_{12}]_i,\ i = 1, \ldots, N$	30	30	5	65
Thickness of UD layers	t_{UD}	1	1	0	2
Thickness of Biax layers	t_{Biax}	1	1	0	2
Thickness of core	t_{core}	0	0	1	1
Fatigue parameters	q, m	2	2	2	6
Fatigue model error	ϵ_m	1	1	1	3
In-plane load	Q_A	-	-	-	1
Transverse load	Q_T	-	-	-	1
Model uncertainties	$X_{exp}, X_{load}, X_{str}, X_{ext}, X_{PMR}$	-	-	-	5

linear elastic solution for a rectangular, orthotropic, simply supported sandwich plate, as described in Section 3.1.3. In general the correspondence between the advanced and the simplified model is good both in terms of stress fields and displacements (see Figure 6.5 where the stress fields are compared).

Fatigue failure

Stresses in the faces and core are evaluated by carrying out a nonlinear analysis with a number of steps with increasing loads. The load magnitudes in each step equal the loads corresponding to the middles of the load bins in the fatigue load spectrum shown on Figure 6.4 b). The highest stress values found in each load step are then used as input to the fatigue life model described in 3.3. Only the fatigue life of the faces is considered, because in the presently discussed load case the transverse shear stresses in the core are practically negligible in the linear part of the panel response. With the simplified model the fatigue life is estimated using an almost identical approach, by evaluating stresses in all layers. Due to the linear behaviour of the simplified model it is not necessary to perform multiple analysis steps - a single step is used to determine the stress field for a given load magnitude, and the stresses for other load levels are found proportionally. For the fatigue limit state the advantage of the "realistic" model in comparison with the simplified one is the better accuracy of boundary conditions and stress calculations, and the geometrically nonlinear

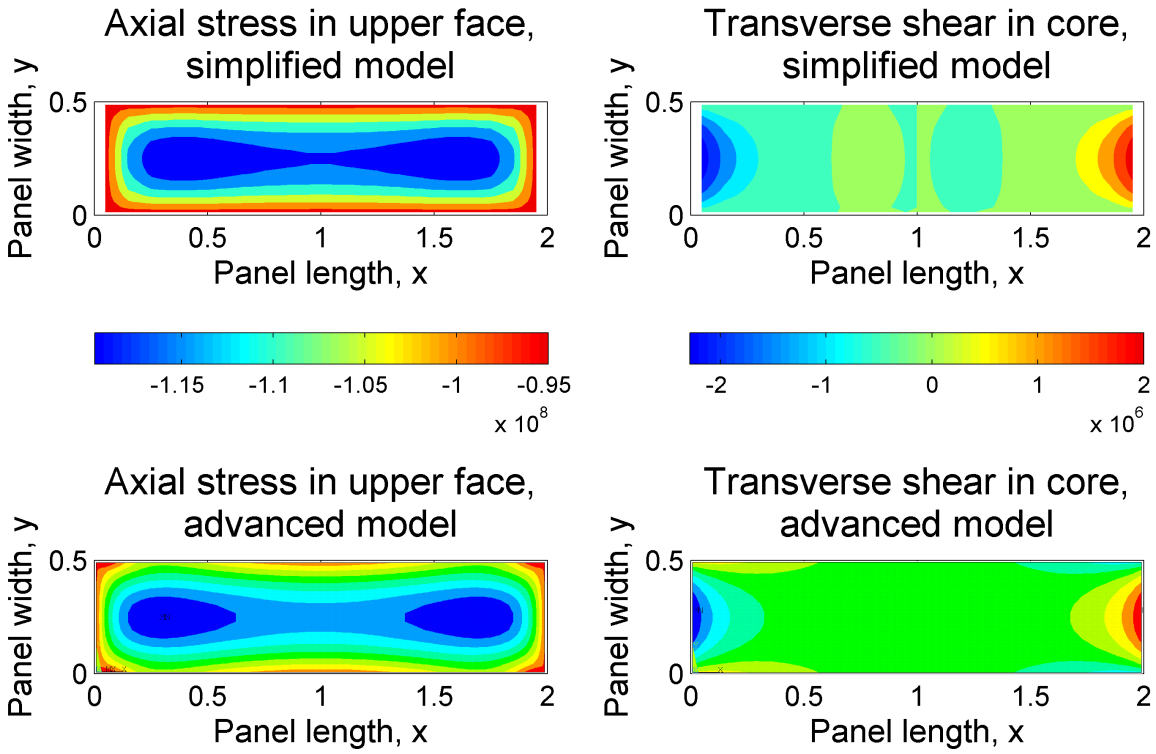

Figure 6.5: Axial stresses in the upper face, and transverse shear stresses in the core: comparison between stress field predictions from the simplified and advanced models.

analysis which is able to capture the nonlinearities in the panel response which occur when large out-of-plane displacements have taken place.

Failure due to exceeding the ultimate strength

The ultimate strength of the panel is determined using the method described in Section 3.1.5. Ultimate failure is reached when the first fiber failure occurs in any layer and location in the panel. The point of failure is obtained by applying a progressive failure analysis which takes into account matrix material degradation happening prior to fiber failures. In contrast to the fatigue limit state where the simplified and advanced model show good agreement, this is not entirely true when modelling ultimate material failure. The simplified model can either provide an estimate of the eigenvalue buckling limit, or an estimate of ply failure due to compressive load, but it is not able to capture the combination of the nonlinear interaction between buckling and material effects, see Figure 6.6. For moderate load magnitudes the response of the panel is linear and the two models correspond well to each other. The stress field is similar in both models, with close to uniform compressive stress distribution in the layers with zero-degree orientation (see Figure 6.7. The nonlinear response of the panel which is in the form of increased out-of-plane displacement under high loads causes increased bending of the plate which results in linearly varying stress field in the zero-degree layers. The difference reflects in the predictions of the layer location where first ply failure occurs. This is shown on Figure 6.8, where the location of first-fiber failure is shown for 1000 random samples. The simplified model predicts that failures will be more evenly distributed among layers, compared to the predictions from the advanced model which estimates that most failures will be located in the layers that are subject to highest bending stresses.

6.2.8 System reliability

The sandwich structure consists of three main components (the faces and the core), each associated with a separate failure mode, making the sandwich a three-component reliability system. If failure of any of the faces or the core is considered an overall structural failure, the sandwich panel can be considered to be a series system. The layered structure of the faces means that a laminated face will itself also be a reliability system. Depending on the failure mode, the system reliability behaviour of the face laminates is treated in a different way. The model for fatigue limit state is based on failure of the entire face, and it is therefore assumed that the outcome of the model represents the system failure probability for the entire laminate under consideration. Ultimate material failure however is defined as failure

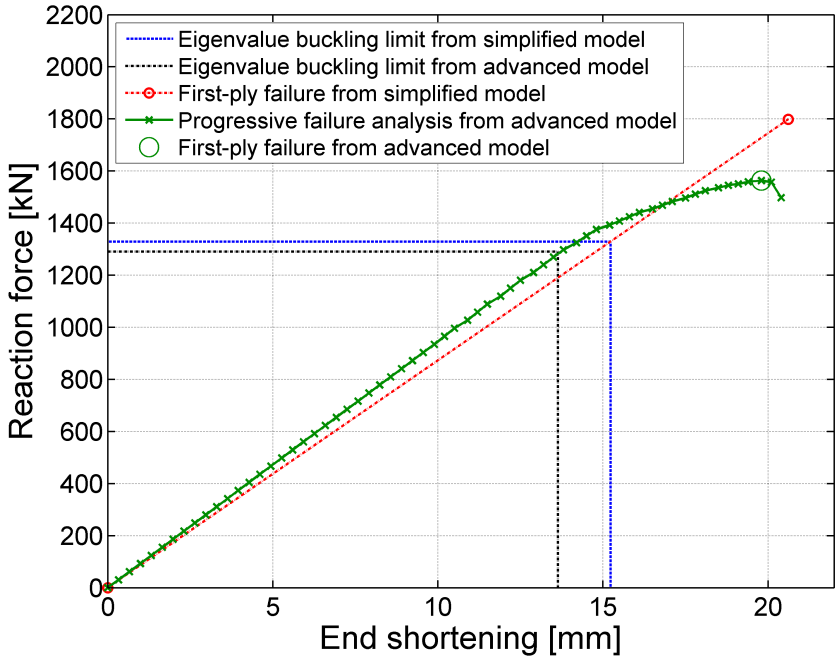

Figure 6.6: Comparison of ultimate failure predictions from the advanced and simplified plate models.

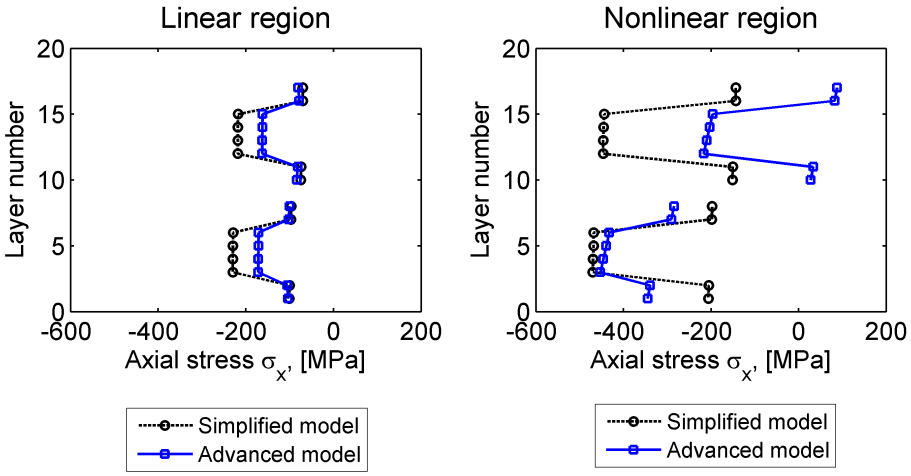

Figure 6.7: Through-the thickness stress distribution in the sandwich panel: a) Linear model behaviour b) Effects of nonlinearity under high loads.

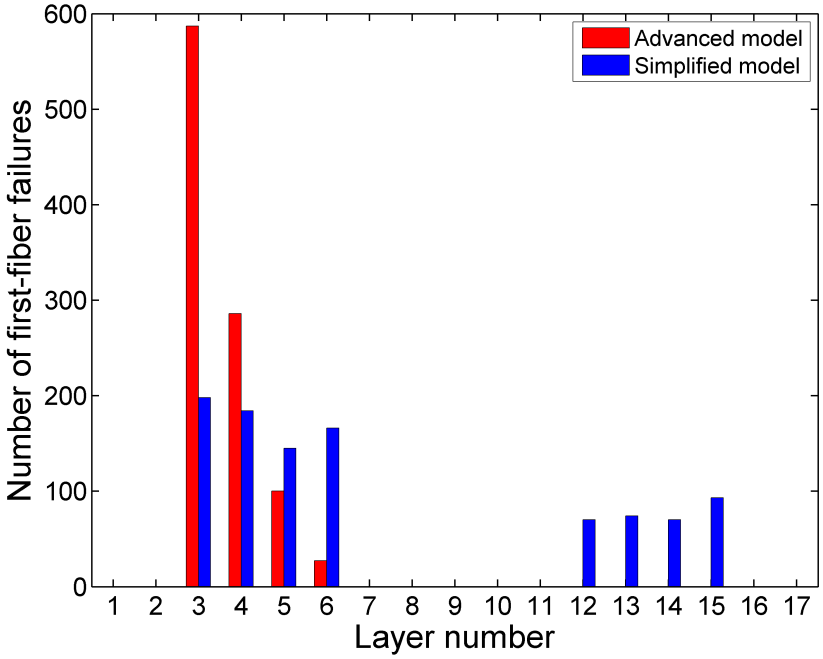

Figure 6.8: Distribution of failure locations per 1000 random plate model samples loaded until ultimate material failure.

of any of the layers within the face, and as a result the face laminate has to be treated as a series system when this limit state is considered.

If reliability analysis is carried out on each of the system components, the probability of failure for a series system is then obtained by applying the multivariate normal distribution to the component reliability indexes (equation 2.24):

$$P_f = 1 - \Phi_k(\boldsymbol{\beta}, \mathbf{R}) \tag{6.19}$$

where $\boldsymbol{\beta}$ is the vector of component reliability indexes, $\mathbf{R}$ is the matrix containing the correlation coefficients between component reliabilities, and $\Phi_k()$ is the k-dimensional multivariate standard normal distribution. The correlation coefficient between components i and j can be estimated from the importance factors resulting from the component reliability analysis, $R_{ij} = \boldsymbol{\alpha}_i^T \boldsymbol{\alpha}_j$.

Another approach to obtaining the system reliability is to apply a method which directly evaluates the failure probability over the entire failure domain - i.e., simulation-based methods.

For the present analysis both approaches are utilized. The reliability against fatigue

Table 6.5: Component reliability results overview, fatigue limit state, 20 years lifetime

Component	Result	Iteration number			
		$I = 0$	$I = 1$	$I = 2$	$I = 3$
Upper face	R_{REAL}/S	2.3981	0.9979	0.9999	-
	R_{IDEAL}/S	2.4191	1.0086	1.0106	-
	$\nu(\mathbf{X})$	0.9913	0.9894	0.9894	-
	β_{FORM}	5.4772	5.4699	5.4698	-
	P_f	$2.16 \cdot 10^{-8}$	$2.25 \cdot 10^{-8}$	$2.25 \cdot 10^{-8}$	-
Core	P_f	$\rightarrow 0$			
Lower face	R_{REAL}/S	2.0824	0.9817	0.9978	0.9998
	R_{IDEAL}/S	2.3311	1.1196	1.1405	1.1431
	$\nu(\mathbf{X})$	0.8933	0.8768	0.8749	0.8747
	β_{FORM}	4.8744	4.7801	4.7685	4.7671
	P_f	$5.46 \cdot 10^{-7}$	$8.76 \cdot 10^{-7}$	$9.28 \cdot 10^{-7}$	$9.34 \cdot 10^{-7}$

failure is estimated using FORM analysis and the Model Correction Factor method. For ultimate material failure the same approach is applied to each of the layers in the two laminated faces, and in addition asymptotic sampling analysis is carried out.

6.2.9 Results

Reliability analysis

Fatigue limit state The reliability against fatigue failure of the two sandwich faces is determined using the Model Correction Factor method. For both components fast convergence is achieved, using not more than 4 function calls to the advanced model, showing that the Model Correction Factor method is capable of ensuring excellent computational efficiency. Under the loading conditions in the present analysis the probability of fatigue failure of the core tends to zero, and a detailed reliability analysis is not carried out. Table 6.5 shows an overview of the component reliability analysis for fatigue limit state. The upper face has higher reliability, which is a result of the stress distribution pattern - the transverse pressure causes bending stresses, which, added to the stress caused by the in-plane force, result in higher compressive stresses in the lower face compared to the upper face.

Table 6.6: Component reliability results overview, ultimate strength limit state, 20 years reference period

Layer number	$\nu(\mathbf{X})$	β_{FORM}	P_f
3	1.021	5.561	$1.34 \cdot 10^{-8}$
4	1.023	5.592	$1.12 \cdot 10^{-8}$
5	1.026	5.630	$9.02 \cdot 10^{-9}$
6	1.032	5.687	$6.47 \cdot 10^{-9}$
12	1.161	6.749	$7.44 \cdot 10^{-12}$
13	1.165	6.766	$6.61 \cdot 10^{-12}$
14	1.169	6.785	$5.82 \cdot 10^{-12}$
15	1.170	6.786	$5.78 \cdot 10^{-12}$

Ultimate strength limit state The stress pattern shown on Figure 6.7 reveals that the stress in layers with 0-degree orientation is higher than the stress in the Biax layers. As a result, a significant failure probability exists only for the 0-degree layers. Component reliability analyses are carried out for all layers with this orientation, using the Model Correction Factor method. The results are summarized in Table 6.6.

The design points obtained by the layer-wise reliability analysis are very similar to each other. In fact, the only difference in the design values for different layers is in the compressive material strength in that particular layer, while the design values for the load-related variables are practically the same. This makes it possible to run an efficient importance sampling analysis by using a "common" design point. For variables where the design point is common for all layers, the common design value is used as a sampling center point, while for the material strength of individual layers the sampling center is placed at the mean values, $U = 0$. In addition to the component reliability analyses and the importance sampling which are all to some extent dependent on the use of a simplified model, an Asymptotic sampling (AS) calculation is carried out. The AS analysis directly estimates system reliability using only the advanced model, eliminating any uncertainties associated with the use of the simplified model. Six support points are used, $f = [0.25, 0.3, 0.35, 0.4, 0.45, 0.5]$. The number of data samples for each support point is based on convergence considerations, where the criterion for convergence is if at least 20 failure events have been observed. Table 6.7 summarizes the results from the AS analysis, and Figure 6.9 shows graphically the support points and the asymptotic reliability index.

Table 6.7: Asymptotic sampling analysis results, ultimate strength limit state, 20 years reference period

f-value	Number of samples	Failure outcomes	$P_f(f)$	$\beta(f)$
0.25	300	32	$1.07 \cdot 10^{-1}$	1.244
0.3	500	35	$7.00 \cdot 10^{-2}$	1.476
0.35	1000	31	$3.10 \cdot 10^{-2}$	1.866
0.4	1500	30	$2.00 \cdot 10^{-2}$	2.054
0.45	2000	25	$1.00 \cdot 10^{-2}$	2.326
0.5	2500	19	$6.33 \cdot 10^{-3}$	2.493
1	-	-	$1.10 \cdot 10^{-7}$	5.181

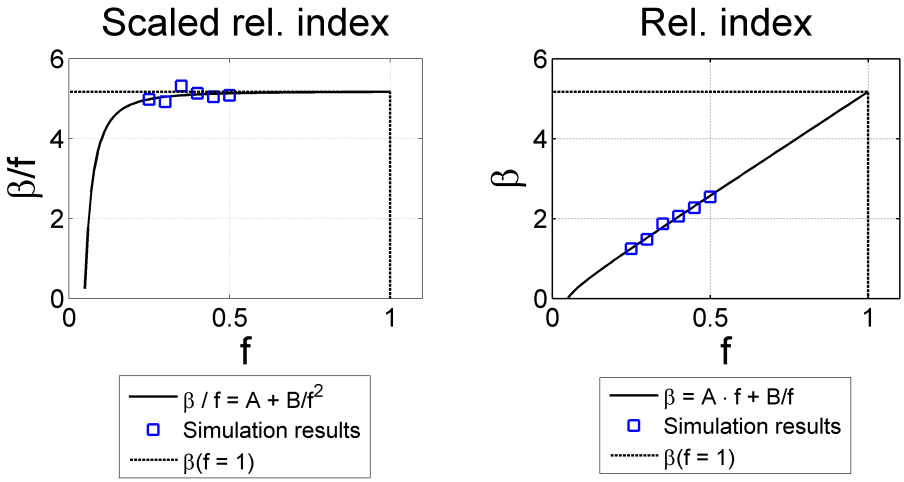

Figure 6.9: Results from Asymptotic Sampling analysis with 6 support points.

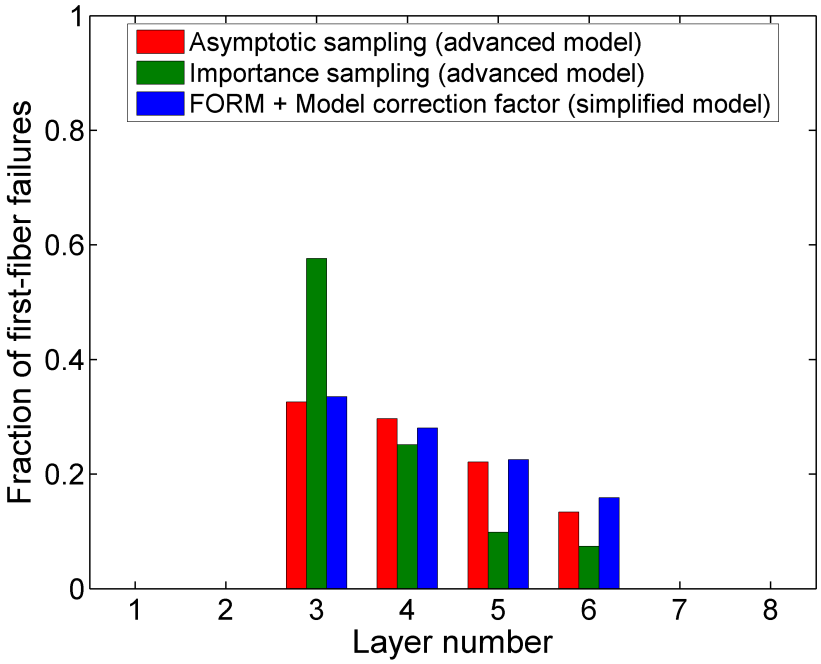

Figure 6.10: Estimated distribution of failure locations per 1000 failure events, as predicted by Asymptotic Sampling analysis, by component reliability analysis using the Model Correction Factor method, and by Importance Sampling.

Figure 6.10 shows how the different reliability analysis methods predict the distribution of the layer locations where first-ply failures will occur. In comparison with Figure 6.8 it is seen that the layer-wise tuning of the Model Correction Factor in the component reliability analyses has resulted in better prediction of the ratio between component failure probabilities compared to the prediction with solely using the simplified model. However, the Model Correction Factor approach still differs in this aspect from the other two methods which evaluate solely the advanced model.

Importance factors

The input variables with highest importance factors are listed in Table 6.8. The scatter in material fatigue properties, represented by the model error ϵ_m, the variance of the model error σ_{ϵ_m}, and the fatigue parameter q_A, have very high importance factors for the fatigue limit state. In addition to that, model uncertainty parameters X_{exp}, X_{st} and X_{ext} also influence the reliability. The ultimate strength limit state is entirely dominated by the load uncertainties, with X_{exp}, X_{st}, X_{ext} and X_{str} having the highest influence.

Table 6.8: Overview of variables with largest importance factors.

Fatigue		Ultimate, FORM		Ultimate, A. Sampling	
Variable	α	Variable	α	Variable	α
ϵ_m	-0.59	X_{exp}	-0.75	X_{exp}	-0.67
X_{exp}	-0.52	X_{st}	-0.37	X_{st}	-0.57
q_A	-0.30	X_{ext}	-0.34	X_{ext}	-0.31
σ_{ϵ_m}	-0.28	X_{str}	-0.22	X_{str}	-0.21
X_{st}	-0.26	t_{UD}	0.18	t_{UD}	0.11
m_A	0.22	X_c	0.17	Q_A	-0.11
$E_{1,core}$	0.17	Q_A	-0.16	t_{core}	0.10
X_{str}	-0.16	X_t	0.14	-	-

System reliability

Based on importance factors from component reliability results, the correlations between component reliabilities are estimated with the expression $r_{ij} = \boldsymbol{\alpha_i}^T \boldsymbol{\alpha_j}$. For the fatigue limit state the correlation between the reliability of the upper and the lower face is $\rho = 0.29$. For ultimate strength limit state the layerwise reliabilities have much higher correlation, and the correlation coefficient between any of the 0-degree layer component reliabilities is $\rho \approx 0.97$. The correlation between ultimate strength and fatigue reliability is $\rho = 0.57$.

Using the multivariate normal distribution, the system probability of failure is determined by $P_f = 1 - \Phi_k(\boldsymbol{\beta}, \mathbf{R})$ (eq. 6.19). The result is given in Table 6.9. The system probability of failure in ultimate strength limit state predicted using the series-system representation and the Model Correction Factor method is similar to the failure probability predicted by the Asymptotic Sampling method. Based on the probability of occurrence of the different failure modes it is clear that the fatigue limit state is the most critical failure mode for the design of the structure discussed in the present example.

Safety factor calibration

Two target reliability levels are chosen for partial safety factor calibration. The first target level is $P_f = 5{\cdot}10^{-4}$, corresponding to $\beta \approx 3.30$. This reliability is similar to what is expected from the IEC 61400-1 standard [37] and what has been reported by previous studies which were discussed in Chapter 1. A higher reliability level, $\beta = 4.1$ is chosen as comparison.

Table 6.9: Component and system reliabilities, 20 years reference period

		Reliability index	Failure probability
		β	P_f
Fatigue	Upper face	5.4698	$2.25 \cdot 10^{-8}$
	Lower face	4.7671	$9.34 \cdot 10^{-7}$
	System	4.7623	$9.57 \cdot 10^{-7}$
Ultimate	Layer 3	5.561	$1.34 \cdot 10^{-8}$
	Layer 4	5.592	$1.12 \cdot 10^{-8}$
	Layer 5	5.630	$9.02 \cdot 10^{-9}$
	Layer 6	5.687	$6.47 \cdot 10^{-9}$
	Layer 12	6.749	$7.44 \cdot 10^{-12}$
	Layer 13	6.766	$6.61 \cdot 10^{-12}$
	Layer 14	6.785	$5.82 \cdot 10^{-12}$
	Layer 15	6.786	$5.78 \cdot 10^{-12}$
	System, FORM	5.3666	$4.01 \cdot 10^{-8}$
	Importance Sampling	5.3729	$3.87 \cdot 10^{-8}$
	Asymptotic Sampling	5.1808	$1.10 \cdot 10^{-7}$
Overall system		4.7540	$9.97 \cdot 10^{-7}$

Table 6.10: Safety factors for ultimate strength and fatigue limit states, calibrated on independent random variables $\mathbf{X}$. ULS factors are based on annual reliability index, FLS factors on lifetime reliability

β_{target}	$\beta = 4.10$			$\beta = 3.30$			IEC		
$P_{f,target}$	$P_f = 2.5 \cdot 10^{-5}$			$P_f = 5 \cdot 10^{-4}$					
Limit state	γ_m	γ_f	γ	γ_m	γ_f	γ	γ_m	γ_f	γ
Ultimate, 5% char. strength	1.01	1.58	1.59	1.0	1.44	1.44	1.3	1.25	1.63
Ultimate, 50% char. strength	1.07	1.58	1.69	1.06	1.44	1.53	-	-	-
Fatigue, 5% char. S-N curve	1.19	1.3	1.55	1.23	1.11	1.36	1.2	1.0	1.2
Fatigue, 50% char. S-N curve	1.35	1.3	1.76	1.41	1.11	1.56	1.7	1.0	1.7

For ultimate strength limit state the component reliabilities of different layers have very high correlation, $\rho \approx 0.98$ as shown by the component reliability analysis using the Model Correction Factor method. It is therefore considered that the changes in the reliability of one component will result in proportional changes in the reliabilities of other components. As a result, a single component limit state should be representative enough to be used for safety factor tuning. Results in Table 6.6 show that layer 3 accounts for about 35% of the overall failure probability. Based on this observation, the limit state equation for layer 3 is tuned so that it predicts failure probability which is 35% of the target failure probability. It is expected that this approach will result in an overall probability of ultimate failure approximately equal to the target probability of failure. Then, safety factors are calibrated using equations 6.12 and 6.18.

Usually the design standards recommend the use of characteristic material properties corresponding to at least 95% survival probability. However, it is often more convenient for designers to use characteristic values for material stiffnesses and strengths corresponding to 50% survival probability, because such data are more readily available and because the use of the mean stiffness makes it possible to apply the models to problems other than validation of structural capacity. The IEC 61400-1 standard allows using the mean material properties in fatigue limit state instead of a lower quantile value, provided that a safety factor of $\gamma \geq 1.5$ ($\gamma \geq 1.7$ for composites) is used. Based on that, two separate calibrations for safety factors are carried out, one using material properties (ultimate material strength and S-N curves) with 95% survival probability, and another with 50% survival probability. In Table 6.10 the results from these calibrations are compared to the corresponding safety factor values given by the IEC standard.

The ultimate strength limit state is dominated by the model uncertainties related to loads, which is reflected in the partial safety factors: the values of γ_m are much smaller than the values of γ_f. This repeats the tendency seen in the importance factors, where, as Table 6.8 showed, load-related model uncertainties have the highest four importance factors. A somewhat opposite tendency can be seen for fatigue, where the material-related partial safety factors are in most cases higher than the load-related ones. The cause of this behaviour is the larger scatter of material properties in fatigue compared to the ultimate strength of the material. This is supported by the importance factors (Table 6.8), where material properties have among the largest importance factors in fatigue.

The results from safety factor calibration show lower safety factors compared to the requirements in IEC61400-1 [37]. An exception is the case of fatigue limit state using a

Table 6.11: Safety factors for ultimate strength and fatigue limit states, calibrated on dependent variables $R(\mathbf{X})$ and $S(\mathbf{X})$.

β_{target} $P_{f,target}$	$\beta = 4.10$ $P_f = 2.5 \cdot 10^{-5}$	$\beta = 3.30$ $P_f = 5 \cdot 10^{-4}$
Limit state	γ	
Ultimate, 5% char. strength	1.57	1.42
Ultimate, 50% char. strength	1.68	1.52

characteristic S-N curve with 95% survival probability, where the overall safety factor is higher than the one recommended in IEC 61400-1. In this case, using an overall factor of 1.2 as recommended in the standard would result in reliability index of $\beta \approx 2.4$. This is an effect of the presence of the uncertainty variables related to load modelling, which result in a relatively high value of the load-related partial safety factor, $\gamma_f = 1.11$. The value recommended by IEC is $\gamma_f = 1.0$, which makes it unclear whether the IEC standard design guidelines take load modelling uncertainties in fatigue into account at all. However, it cannot be concluded that load modelling uncertainty is not accounted for - as discussed earlier, when two or more partial safety factors are applied in the same equation, an infinite number of correct solutions exists, and one of the partial safety factor values have to be determined by other means. Therefore, it might be the case that γ_m accounts for all combined uncertainties, while γ_f equals one for convenience.

The results given in Table 6.10 are based on applying safety factors directly on the basic random variables $\mathbf{X}$. As discussed in Chapter 2, this approach ensures that the safety factors are invariant to the formulation of the limit state function. The validity of this approach is illustrated by the results in Table 6.11, where another calibration of the overall safety factors for ultimate strength limit state $\gamma = \gamma_m \cdot \gamma_f$ is shown, this time by applying the safety factors directly on the dependent variables, the resistance $R(\mathbf{X})$ and the load $S(\mathbf{X})$. Using the dependent variables results in lower partial safety factors. However, the difference is not large, which shows that for this particular limit state the dependent variables have close to linear relation to the basic input variables.

6.3 Discussion on reliability analysis results

The present chapter demonstrated the process of estimating the reliability of a wind turbine blade against blade-tower collision, and the reliability of a sandwich panel against fatigue and ultimate strength failure. The outcome of the analyses has shown that two groups of factors have the highest impact on the reliability - a) uncertainties in load modelling, which can be seen as a major factor for all three limit states considered in the present chapter, and b) the scatter in material properties, which has a significant influence on the reliability against fatigue failure.

The interpretation of the results from the safety factor-related analyses depends on the choice of an acceptable reliability level. The "low" target reliability level used here, $\beta = 3.3$, is already higher than the range quoted in literature, $\beta \approx 2.7 - 3.1$. If it is assumed that a reliability of $\beta = 3.3$ is sufficiently high, the results showed that in most cases the calibrated safety factors are smaller than the ones used in the standards, or that if the safety factors recommended by the standard are used, the resulting reliability is higher than the target reliability. In general, it can be concluded that with the present level of modelling detail, including material, load and uncertainty models, the use of the standard safety factors results in sufficiently safe structures, and in some cases the actual reliability even exceeds the target reliability levels significantly.

Structural reliability estimates and partial safety factor values are highly dependent on the variables describing uncertainties in load modelling. Due to that, the ratio between partial safety factors for loads and materials found in the present analysis is different from the one used in the IEC 61400-1 standard: the values of the strength reduction factor γ_m are smaller than the ones in the standard, and the opposite is true for the load-related factor γ_f. Large model uncertainties are also the cause for the only result where the presented reliability models require higher safety factors than the standard, which is the case of fatigue limit state with an $S - N$ curve representing 95% survival probability. This ultimately brings the same conclusion which was already seen in Chapter 5 - that quantifying model uncertainties is an important task which could potentially lead to optimized design or reduced safety factors.

It should be noted that the models used in the present analysis only account for the natural variation in materials which are assumed to be free of defects and discontinuities on macroscopic level. As a result, the partial safety factors for material strength represent only the additional safety needed to account for the variability in the properties of a defect-free material with constant geometry and layup. The partial safety coefficients in IEC61400-1 are based on the same assumption. If these conditions are not satisfied, adding additional

safety is required. This can for example be done as suggested by the DNV-OS-J102 standard [16], where additional partial safety factors are introduced for material degradation due to ageing, ply drop-offs, casting process, and other.

The Model Correction Factor method was used in determining the reliability of a sandwich panel against fatigue and ultimate material failure. The method proved to be computationally very effective as just a few calls to the advanced model per failure mode were necessary. In the case of fatigue limit state where the linear part of the panel response has most importance the agreement between the advanced and the simplified model is very good. As the simplified model is not able to take nonlinear effects into account, there is some doubt that the Model Correction Factor method would correctly predict the reliability in ultimate strength limit state. However, as results show, carrying out component reliability analyses where the model correction factor $\nu(\mathbf{X})$ is independently calibrated for each of the layers in the laminate has resulted in a system reliability estimate which is relatively close to the estimate obtained by Asymptotic Sampling using only the advanced model. This was further verified by carrying out an Importance Sampling run centered at the design point from the MCF analysis. Even if it is assumed that the reliability index predicted by the AS method is closest to the actual reliability, the performance of the Model Correction Factor method is impressive as it predicts a similar order of magnitude of the failure probability with making only 24 calls (3 per layer) to the advanced model, while the Asymptotic Sampling has required the use of 8800 function calls. Nevertheless, using the Model Correction Factor method should always be used with caution, and it has to be made sure that the analysis has converged to a realistic design point and that the behaviour of the advanced model is properly approximated by the simplified model in the region of the design point.

Chapter 7

Conclusions and Recommendations for Further Work

The present thesis addressed the reliability-based design of composite structures and wind turbine blades in particular. The reader was introduced to the main aspects of reliability of composites, models for ultimate static and fatigue strength of composites were developed, the process of estimating the statistical distribution of load extremes was demonstrated, and example reliability analyses were shown.

Based on the findings in the thesis, the following main conclusions are drawn:

a) the problem of estimating the reliability of wind turbine blades has been addressed in detail, and suitable methodologies for carrying efficient and robust reliability analysis have been identified

b) model uncertainties have a very high influence on the reliability estimate, and an effort to reduce model uncertainties can be rewarded with improved structural reliability or lower safety factors

c) wind turbine design using present state-of-the-art models and standard safety factors results in sufficiently safe structures, and in some cases the actual reliability exceeds the assumed sufficient reliability levels significantly.

Three years is a long period giving the chance for addressing a problem extensively. Despite that, it is rarely that a given topic of that size and complexity can be studied in its completeness. The present project is no different, and there are a number of aspects that the author considers as deserving further attention.

The first suggestion for further work is an investigation of additional factors that might have deteriorating effect on the material properties and the strength of the structure. These are for example geometrical discontinuities such as ply drops, material degradation due to exposure to severe environmental conditions and ageing, and defects in the materials. The DNV-OS-J102 standard [16] addresses this issue to some extent by providing a number of additional partial safety factors, including factors accounting for ageing of material, for the presence of ply drop-offs, and for the material curing process. Defects in the structure of composite materials are however not dealt with. Analyzing the influence of defects in the material is probably often avoided due to fears of opening a "Pandora's box", because various defects are present on different scale levels, their location and size are random, and taking them into account requires significant changes in the models and assumptions. Nevertheless, flaws can have high influence on the safety of the structure, and therefore they are one of the important matters that need to be taken further. Inspiration can be taken for example from the work on modelling the spatial and size distribution of defects by Toft et. al, 2011 [83], while studies by Leong et. al, 2012 [44] and Hayman et. al, 2007 [36] discuss the effect of wrinkles which are among the defects with most severe consequences for the safety of composite structures.

Another phenomenon that requires more understanding is the mismatch between the predicted and observed fatigue life of composites subjected to variable amplitude loading. In the present thesis this has been taken empirically into account by adjusting the model to fit the observations, however the generality of such an approach can be doubted. Therefore finding physical model that correctly predicts the phenomenon will be a significant improvement.

A major part of the uncertainties observed in the analysis are a consequence of uncertainties in modelling of loads and external conditions. The uncertainty can be reduced by improving the knowledge on turbine controller behaviour, extreme operational events and frequency of faults, and assessment of the wind field across the entire rotor.

A final recommendation for further work is carrying out a risk analysis which includes the consequences of different modes of failure. This information can be used to identify the causes of failure which are potentially most critical from economical point of view, and provide directions for eventual efforts towards improvement of the reliability.

Bibliography

[1] 3AComposites, www.3acomposites.com. *Baltek select grade structural balsa data sheet*, 2011.

[2] Abramowitz, M., and Stegun, I. *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*. Dover Publications, 1972.

[3] Augusti, G., Baratta, A., and Casciati, F. *Probabilistic Methods in Structural Engineering*. Chapman and Hall, 1984.

[4] Bjerager, P. Probability integration by directional simulation. *Journal of Engineering Mechanics, ASCE 114*, 8 (1988), 1285–1302.

[5] Bourgund, U., and Bucher, C. G. Importance sampling procedure using design points. Tech. Rep. 8-86, Institute of Engineering Mechanics, University of Innsbruck, Austria, 1986.

[6] Box, G. E. P., and Tiao, G. C. *Bayesian Inference in Statistical Analysis*. Adison-Wesley, Reading, Mass., 1992.

[7] Breitung, K. Asymptotic approximations for multinormal integrals. *Journal of Engineering Mechanics, ASCE 110*, 3 (1984), 357–366.

[8] Bucher, C. Asymptotic sampling for high-dimensional reliability analysis. *Probabilistic Engineering Mechanics 24* (2009), 504–510.

[9] Bucher, C. G., and Bourgund, U. A fast and efficient response surface approach for structural reliability problems. *Journal of Structural Safety 7* (1990), 57–66.

[10] Cetin, K. O., and Der Kiureghian, A. Probabilistic models for the initiation of seismic soil liquefaction. *Journal of Structural Safety 24* (2002), 67–82.

[11] Daniels, H. E. The statistical theory of strength of bundles of threads. *Proceedings of the Royal Society of London, Series A 183*, 995 (1945), 405–435.

[12] Davison, A. C., and Smith, R. L. Models for exceedances over high thresholds (with discussion). *Journal of the Royal Statistical Society B 52* (1990), 393–442.

[13] Dekkers, A., Einmahl, J., and de Haan, L. A moment estimator for the index of an extreme-value distribution. *The annals of statistics 17*, 4 (1989), 1833–1855.

[14] Der Kiureghian, A., and Ditlevsen, O. Aleatory or epistemic? Does it matter? *Journal of Structural Safety 31*, 2 (2009), 105–112.

[15] Det Norske Veritas. DNV classification notes: Structural reliability analysis of marine structures. Tech. Rep. CN30.6, Det Norske Veritas, 1992.

[16] Det Norske Veritas. Offshore standard DNV-OS-J102: Design and manufacture of wind blades, offshore and onshore wind turbines, 2006.

[17] Dimitrov, N., Friis-Hansen, P., and Berggreen, C. Reliability analysis of a composite wind turbine blade section using the Model Correction Factor method: numerical study and validation. *Journal of Applied Composites 20*, 1 (2013), 17–39. DOI 10.1007/s10443-011-9246-3.

[18] Ditlevsen, O., and Arnbjerg-Nielsen, T. Model Correction Factor method in structural reliability. *Journal of Engineering Mechanics 120*, 1 (1994), 1–10.

[19] Ditlevsen, O., and Madsen, H. O. *Structural Reliability Methods*. Wiley, 1996.

[20] Drezner, Z. Computation of the multivariate normal integral. *ACM Transactions of Mathematical Software 18*, 4 (1992), 470–480.

[21] Efron, B. Bootstrap methods: another look at the jackknife. *Annals of Statistics 7*, 1 (1979), 1–26.

[22] European Committee For Standartization. European Standard EN 1990: Eurocode - Basis of structural design, 2002.

[23] Favaz, Z. Fatigue life model for fibre-reinforced materials under general loading conditions. *Journal of Composite Materials 28*, 15 (1994), 1432–1451.

[24] Fisher, R. A. *Contributions to Mathematical Statistics*. Wiley, 1950.

[25] Friis-Hansen, P. *Reliability Analysis of a Midship Section*. PhD thesis, Technical University of Denmark, Department of Naval Architecture and Offshore Engineering, 1994.

[26] Gardoni, P., and Der Kiureghian, A. Probabilistic capacity models and fragility estimates for reinforced concrete columns based on experimental observations. *Journal of Engineering Mechanics 128*, 10 (2002), 1024–1038.

[27] Gardoni, P., and Der Kiureghian, A. Probabilistic models and fragility estimates for bridge components and systems. Tech. Rep. 2002/13, Pacific Earthquake Engineering Center (PEER), University of California, Berkeley, 2002.

[28] Gollwitzer, S., and Rackwitz, R. An efficient solution to the multinormal integral. *Probabilistic Engineering Mechanics 3*, 2 (1988), 98–101.

[29] Gollwitzer, S., and Rackwitz, R. On the reliability of daniels systems. *Journal of Structural Safety 7* (1990), 229–243.

[30] Gumbel, E. J. *Statistics of Extremes*. Columbia University Press, 1958.

[31] Hansen, M. O. L. *Aerodynamics of Wind Turbines*, second ed. Earthscan, 2008.

[32] Hashin, Z. Failure criteria for unidirectional fiber composites. *Journal of Applied Mechanics, ASME 47* (1980), 329–334.

[33] Hashin, Z. Fatigue failure criteria for unidirectional fiber composites. *Journal of Applied Mechanics, ASME 48* (1981), 846–852.

[34] Hasofer, A. M., and Lind, N. C. Exact and invariant second-moment code format. *Journal of the Engineering Mechanics Division, ASCE 100*, EM1 (1974), 111–121.

[35] Hayman, B., Berggreen, C., Lundsgaard-Larsen, C., Delarche, A., and Toftegaard, H. L. Studies of the buckling of composite plates in compression. *Ships and Offshore Structures 6*, 1-2 (2011), 81–92.

[36] Hayman, B., Berggreen, C., and Pettersson, R. The effect of face wrinkle defects on the strength of frp sandwich structures. *Journal of Sandwich Structures and Materials 9* (2007), 377–404.

[37] IEC. International Standard IEC61400-1: Wind Turbines - Part 1: Design Guidelines, 2005.

[38] ISO. International standard ISO 2394: General principles on reliability for structures, 1998.

[39] Jones, R. *Mechanics of Composite Materials*. Taylor & Francis, 1998.

[40] Kallenberg, O. *Foundations of Modern Probability*. Springer-Verlag, 1997.

[41] Kaminski, M. On probabilistic fatigue models for composite materials. *International Journal on Fatigue 24* (2002), 477–495.

[42] Kang, W. H., and Song, J. Evaluation of multivariate normal integrals for general systems by sequential compounding. *Journal of Structural Safety 32* (2010), 35–41.

[43] Lekou, D. J., and Philippidis, T. P. Pre- and post-thin: A tool for the probabilistic design and analysis of composite rotor blade strength. *Journal of Wind Energy 12*, 7 (2009), 676–691.

[44] Leong, M., Overgaard, L. C. T., Thomsen, O. T., Lund, E., and Daniel, I. M. Investigation of failure mechanisms in gfrp sandwich structures with face sheet wrinkle defects used for wind turbine blades. *Journal of Composite Structures 94* (2012), 768–778.

[45] Lepage, P. G. A new algorithm for adaptive multidimensional integration. *Journal of Computational Physics 27* (1978), 192–203.

[46] Liu, Y., and Mahadevan, S. Probabilistic fatigue life prediction of multidirectional composite laminates. *Journal of Composite Structures 69* (2005), 11–19.

[47] Love, A. E. H. On the small free vibrations and deformations of elastic shells. *Philosophical transactions of the Royal Society (London) A17* (1888), 491–549.

[48] Madsen, H. O., Krenk, S., and Lind, N. C. *Methods of Structural Safety*. Prentice-Hall, 1985.

[49] Mann, J. The spatial structure of neutral atmospheric surface-layer turbulence. *Journal of Fluid Mechanics 273* (1994), 141–168.

[50] Melchers, R. E. Search-based importance sampling. *Journal of Structural Safety 9* (1990), 117–128.

[51] Miner, A. M. Cumulative damage in fatigue. *Journal of Applied Mechanics 12* (1945), A159–A164.

[52] Moriarty, P. J., Holley, W. E., and Butterfield, S. P. Extrapolation of extreme and fatigue loads using probabilistic methods. Tech. Rep. NREL/TP-500-34421, National Renewable Energy Laboratory, Colorado, USA, 2004.

[53] Naderi, M., and Maligno, A. R. Fatigue life prediction of carbon/epoxy laminates by stochastic numerical simulation. *Journal of Composite Structures 94*, 3 (2011).

[54] Naess, A., and Clausen, P. Combination of the peaks-over-threshold and bootstrapping methods for extreme value prediction. *Structural Safety 23* (2001), 315–330.

[55] Naess, A., and Gaidai, O. Estimation of extreme values from sampled time series. *Structural Safety 31* (2009), 325–334.

[56] Nijssen, R. *Optidat Reference Document*. WMC knolwedge centre, The Netherlands, http://www.wmc.eu, 2006.

[57] Nijssen, R. P. L. *Fatigue Life Prediction and Strength Degradation of Wind Turbine Rotor Blade Components*. PhD thesis, Delft University of Technology, Faculty of Aerospace Engineering, 2006.

[58] Palmgren, A. Die lebensdauer von kugellagern. *Zeitschrift der Vereines Deutches Ingenieure 68*, 14 (1924), 339–341.

[59] Pandey, M. D. An effective approximation to evaluate multinormal integrals. *Journal of Structural Safety 20* (1998), 51–67.

[60] Papanikos, P., Tserpes, K. I., and Pantelakis, S. Modelling of fatigue damage progression and life of cfrp laminates. *Fatigue & Fracture of Engineering Materials & Structures 26* (2003), 37–47.

[61] Paris, P. C., Gomez, M. P., and Anderson, W. E. A rational analytic theory of fatigue. *The Trend in Engineering 13* (1961), 9–14.

[62] Passipoularidis, V. A., and Phillipidis, T. P. Strength degradation due to fatigue in fiber dominated glass/epoxy composites: a statistical approach. *Journal of Composite Materials 43*, 9 (2009).

[63] Philippidis, T. P. Fatigue strength prediction under multiaxial stress. *Journal of Composite Materials 69* (1999), 11–19.

[64] Pickands, J. Statistical inference using extreme order statistics. *The Annals of Statistics 3*, 1 (1975), 119–131.

[65] Reddy, J. N. *Theory and Analysis of Elastic Plates and Shells*. CRC Press, Taylor and Francis, 2007.

[66] Reisner, E., and Stavsky, Y. Bending and stretching of certain types of heterogeneous aelotropic elastic plates. *Journal of Applied Mechanics 28* (1961), 402–408.

[67] Richards, F. S. G. A method of maximum likelihood estimation. *Journal of the Royal Statistical Society 23* (1961), 469–475.

[68] Ronold, K. O., and Larsen, G. C. Reliability-based design of wind-turbine rotor blades against failure in ultimate loading. *Journal of Engineering Structures 22* (2000), 565–574.

[69] Ronold, K. O., Wedel-Heinen, J., and Christensen, C. J. Reliability-based fatigue design of wind-turbine rotor blades. *Journal of Engineering Structures 21* (1999), 1101–1114.

[70] Rosenblatt, M. Remarks on a multivariate transformation. *Annals of Mathematical Statistics 23* (1952), 470–472.

[71] Rubinstein, R. Y. *Simulation and the Monte Carlo Method*. Wiley & Sons, 1981.

[72] Schön, J., and Nyman, T. Spectrum fatigue of composite bolted joints. *International Journal of Fatigue 24* (2002), 273–279.

[73] Shooman, M. L. *Probabilistic Reliability: an Engineering Approach*. McGraw-Hill, 1968.

[74] Silva, A. D., and Kyriakides, S. Compressive response and failure of balsa wood. *International Journal of Solids and Structures 44* (2007), 8685–8717.

[75] Soden, P. D., and McLeish, R. D. Variables affecting the strength of balsa wood. *Journal of Strain Analysis 11*, 4 (1976), 225–234.

[76] Song, J., and Der Kiureghian, A. Bounds on system reliability by linear programming. *Journal of Engineering Mechanics 129*, 6 (2003), 627–636.

[77] Sørensen, J. D., Frandsen, S., and Tarp-Johansen, N. J. Effective turbulence models and fatigue reliability in wind farms. *Probabilistic Engineering Mechanics 23* (2008), 531–538.

[78] Tarp-Johansen, N. Examples of fatigue lifetime and reliability evaluation of larger wind turbine components. Tech. Rep. Risø-R-1418, Risø national laboratory, Roskilde, Denmark, 2003.

[79] Tarp-Johansen, N., Madsen, H., and Frandsen, S. T. Partial safety factors for extreme load effects. Tech. Rep. Risø-R-1319, Risø national laboratory, Roskilde, Denmark, 2002.

[80] ten Have, A. A. WISPER and WISPERX, a summary paper describing their background, derivation and statistics. Tech. Rep. NLR 92410 U, National Aerospace Laboratory NLR, The Netherlands, 1992.

[81] ten Have, A. A. WISPER and WISPERX final definition of two standardized fatigue loading sequences for wind turbine blades. Tech. Rep. NLR 91476 U, National Aerospace Laboratory NLR, The Netherlands, 1992.

[82] Toft, H. S. *Probabilistic Design of Wind Turbines*. PhD thesis, Aalborg University, Department of Civil Engineering, 2010.

[83] Toft, H. S., Branner, K., Berring, P., and Sørensen, J. D. Defect distribution and reliability assessment of wind turbine blades. *Journal of Engineering Structures 33*, 1 (2011), 171–180.

[84] Toft, H. S., Branner, K., Jr., L. M., and Sørensen, J. D. Uncertainty modelling and code calibration for composite materials. *Journal of Composite Materials, doi 10.1177/0021998312451296* (2012).

[85] Toft, H. S., Naess, A., Saha, N., and Sørensen, J. D. Response load extrapolation for wind turbines during operation based on average conditional exceedance rates. *Wind Energy 14*, 6 (2011), 749–766.

[86] Toft, H. S., and Sørensen, J. D. Reliability-based design of wind turbine blades. *Journal of Structural Safety 33* (2011), 333–342.

[87] Tsai, S., and Wu, E. A general theory for strength of anisotropic materials. *Journal of Composite Materials 5* (1971), 55–80.

[88] Veldkamp, H. F. *Chances in Wind Energy: A Probabilistic Approach to Wind Turbine Fatigue Design*. PhD thesis, Delft University, Wind Energy Research Institute, 2006.

[89] Wallace, D. L. Asymptotic approximations to distributions. *Annals of Mathematical Statistics 29* (1958), 635–654.

[90] Weibull, W. A statistical distribution of wide applicability. *Journal of Applied Mechanics 18* (1951), 293–297.

[91] Wöhler, A. Über die festigkeitsversuche mit eisen und stahl. *Zeitschrift für Bauwesen 20* (1870), 73–106.

[92] Zenkert, D. *An Introduction to Sandwich Constructions*. Engineering Materials Advisory Services Ltd., 1995.

[93] Zenkert, D. *Foundations of Fibre Composites: Notes for the Course: Composite Lightweight Structures*. DTU, 2006.