Operational Data Mining for Offshore Wind Farm Maintenance

Fraser John Anderson

School of Engineering

University of Edinburgh

A thesis submitted for the degree of

Doctor of Philosophy

2024

...And we never say anything unless it is worth taking a long time to say.

Treebeard

Abstract

Utilising operational data from offshore wind farms is one lever which can be used to reduce the cost of energy in the industry. However, the value of operational data in reducing operations and maintenance costs is not yet fully leveraged. This is especially true for maintenance (O&M) records. Due to corporate sensitivity of the data; poor data collection and processing practices and; a lack of scrutiny into "data fusion" of maintenance data with other data streams, research into the effectiveness of maintenance interventions has significant room for improvement. This thesis therefore addresses the research question: "How can operational maintenance data be better leveraged to support decision making and therefore reduce O&M costs in the industry?".

This thesis presents a series of analyses which address the research question. First, there is a review of the offshore wind-data ecosystem to identify opportunities for improvement from data processes. Then, the available dataset is used to calculate and scrutinise relevant key performance indicators describing maintenance intervention. From the calculated key performance indicators, the thesis goes on to address two questions posed by the operator of the wind farm: "How effective are night shifts in increasing power production and availability?" and "What is the effect that annual services have on proceeding corrective works and in turn the reliability of wind turbines?".

Frequentist statistics are initially used to analyse a database of work procedures. The data is broken down into different maintenance types and

presented in terms of number of interventions per year and mean down-time. Two case studies are then presented which utilise Bayesian methodologies. Bayesian methodologies were selected as they present several advantages which map well to the problem of analysing maintenance data. A Bayesian hierarchical model was used to address the question “How effective are night shifts in increasing power production and availability?". The methodology allowed for conditional probability distributions to be derived for weekly lost production and technical availability from a small sample size, which could in turn be used for decision making. A Bayesian reliability analysis was used to address the question "What is the effect that annual services have on proceeding corrective works and in turn the reliability of wind turbines?". The methodology extends traditional reliability models by incorporating time-dependent variables, which could be used to quantify the effect of annual services on wind turbine time-to-failure. Employing a Bayesian regime in the reliability model provided in-built uncertainty quantification.

A high-quality database from a currently operational offshore wind farm in the UK is used for the above analyses. Results from the work procedure analysis show that tidally-restricted turbines reduce median availability by 0.89% and that failure rates range from below 1 to over 10 failures per turbine per year using different failure definitions. Results from the night shift analysis show a potential 0.64% increase in availability from a night shift strategy involving one crew transfer vessel employed at night. Results from the annual services analysis show a higher failure intensity for the first 6 days after an annual service takes place, after which the failure intensity is decreased up until 137 days after servicing. The case studies show where better processing of operational maintenance data can lead to insight and therefore aid decision making. The Bayesian hierarchical modelling approach facilitates the updating of prior beliefs as new, although limited, data becomes available. The Bayesian reliability analysis increased the complexity of failure modelling, which is useful in quantifying the impact of a maintenance procedure with a time-dependent effect.

Acknowledgements

I would like to acknowledge EPSRC for funding the work presented within this thesis through the Wind and Marine Energy Systems Centre for Doctoral Training under the grant number EP/L016680/1. A big thanks to the folk involved in running and organising the CDT program. Thanks especially to Drew Smith and Prof. Alasdair McDonald for looking after myself and my cohort during the first year. I was also grateful for the company of my cohort, who did not do a very good job of looking after my liver.

A huge thank you to my supervisors Dr. David García Cava, Dr. Rafael Dawid and Dr. David McMillan. Their experience was invaluable throughout the last 3 years. As first supervisor David García Cava showed seemingly endless patience for an often wayward PhD student. As the industrial supervisor, Rafael Dawid was essential to the PhD in providing the data and insight from the industrial partner. David McMillan also provided great support and guidance throughout the PhD.

I would also like to thank the offshore wind farm and data management software providers for providing me with the data I required for this thesis. I cannot name them for confidentiality reasons, but I am very grateful.

Lastly, thanks to my family, who are both mad and fantastic.

Declaration of Originality

I declare that I have composed this thesis myself and, except where otherwise noted, the work contained within it is entirely my own. Parts of this thesis have already been published, as noted in the text, but no part has been submitted previously for any other degree or professional qualification.

Fraser John Anderson

Contents

List of Figures

List of Tables

List of Symbols

Chapter 3

P	Power
$\bar{P}$	Average Power
P_{rated}	Rated Power
A_t	Time-Based Availability
A_O	Time-Based Operational Availability
A_E	Energetic/Production-Based Availability
E_a	Actual Energy Produced
E_p	Potential Energy Available
$A_{E,O}$	Production-Based Operational Availability
$A_{E,tech}$	Production-Based Technical Availability
λ	Failure rate
I	Number of Intervals for which data is collected
K	Number of assemblies
$n_{i,k}$	Number of failures for sub-assembly k and interval i
T_i	Total time period in hours
N_i	Total number of turbines
t	time
$t_{activerepair}$	Active repair time
$t^i_{pick-up}$	Time of technician pick up during a particular shift
$t^i_{drop-off}$	Time of technician drop-off during a particular shift

Chapter 5

i	Subscript denoting a data points
j	Subscript denoting a group in a hierarchical model
k	Subscript denoting a node in a Directed Acyclic Graph
$\boldsymbol{\theta_i}$	Generic set of variables in a Bayesian model
$\boldsymbol{y_i}$	Generic set of dependent variable data-points in a Bayesian model
$\boldsymbol{x_i}$	Generic set of independent variable data-points in a Bayesian model
$p(\boldsymbol{\theta_i})$	Generic prior distribution for variable of interest θ
$p(\boldsymbol{\theta_i}, \boldsymbol{y_i})$	joint probability density function for $\boldsymbol{y_i}$ and $\boldsymbol{\theta_i}$
$p(\boldsymbol{y_i}\|\boldsymbol{\theta_i})$	Generic likelihood function for variable of interest $\boldsymbol{\theta_i}$, based on data $\boldsymbol{y_i}$
$p(\boldsymbol{\theta_i}\|\boldsymbol{y_i})$	Generic posterior distribution for variable of interest θ, based on data y
X_k	Node in a Directed Acyclic Graph
$pa(X_k)$	Parent of node X_k in a Directed Acyclic Graph
K	Number of nodes in a Directed Acyclic Graph
α	Intercept term in a regression analysis
β	Effect (slope) term in a regression analysis
ϵ_i	Error term in a regression analysis
n_j	Number of data points in a hierarchical group
μ	Hyper-parameter for the mean of a variable of interest
σ_y	Within-group standard deviation y_i
σ_α	Standard deviation among group-level estimates
$\bar{y}$	Mean observed value for all data points (pooled mean)
μ_α	Hyperparameter representing the mean of group-level intercepts
μ_β	Hyperparameter representing the mean of group-level effects
$\boldsymbol{\Sigma}$	Covariance matrix
ρ	Correlation coefficient
κ_α	Rate parameter for hyperprior of σ_α
κ_β	Rate parameter for hyperprior of σ_β
$\boldsymbol{C}$	Correlation matrix
$\boldsymbol{L}$	Lower-triangular matrix representing the Cholesky decomposition of $\boldsymbol{C}$
η	Hyperparameter for the Lewandowski-Kurowicka-Joe (LKJ) distribution

l	Subscript denoting an event in a time-to-event dataset
m	Subscript denoting a B-spline knot vector entry
p	Subscript denoting covariate
q	Subscript denoting turbine frailty
$h_0(t)$	Baseline hazard rate
$h_l(t)$	Hazard function at time t
ϕ	Linear predictor describing the effect of the chosen covariates
a	Baseline failure rate
γ	Shape parameter in a Weibull model
$\boldsymbol{\beta}$	Vector of covariate effects
$\boldsymbol{X_l}$	Vector of covariates
$\boldsymbol{k}$	Vector knot locations
δ	Degree of B-spline
ζ_{p0}	Constant-effect covariate value
ζ_{pm}	Covariate value for a B-spline knot
$\boldsymbol{Z_{ij}}$	Vector of covariates for the i^{th} event for the q^{th} turbine
$\boldsymbol{b_q}$	Vector of turbine-specific effect parameters
$\mathcal{M}_R$	Set of models

Chapter 6

$seas$	Subscript denoting partial pooling by season
fr	Subscript denoting partial pooling by failure rate
y^{LP}	Lost production per week
y^{av}	Techical availability per week
ϵ	Model error
x_{NS}	Covariate denoting the presence of night shifts

Abbreviations

BBN	Bayesian Belied Network
BN	Bayesian Network
BHM	Bayesian Hierarchical Model
CBM	Condition Based Maintenance
CfD	Contract for Difference
CMS	Condition Monitoring System
CF	Capacity Factor
CoE	Cost of Energy
CTV	Crew Transfer Vessel
CV	Cross Validation
DAG	Directed Acycling Graph
ECMWF	European Centre for Medium-range Weather Forecasts
ELDP	Expected Log Point-Wise Predictive Density
GFS	Global Forecast System
H&S	Health and Safety
HMC	Hamiltonian Monte Carol
IEC	International Electrotechnical Commission
JUV	Jack-Up Vessel
KPI	Key Performance Indicator
LKJ	Lewandowski-Kurowicka-Joe
LCoE	Levelised Cost of Energy
LOO-CV	Leave-One-OUt Cross Validation
MCMC	Markov Chain Monte Carlo
MDT	Mean Down Time
ML	Machine Learning
MTBF	Mean Time Between Failures
MTTR	Mean Time to Repair
OM	Opportunistic Maintenance
O&M	Operations and Maintenance
OWDE	Offshore Wind Data Ecosystem
OWF	Offshore Wind Farm
NWP	Numerical Weather Prediction

OWF	Offshore Wind Farm
NWP	Numerical Weather Prediction
PDF	Probability Density Function
PM	Preventative Maintenance
R&M	Reliability and Maintenance
RBM	Reliability Centred Maintenance
RCM	Reliability Centred Maintenance
RIDDOR	Reporting of Injuries, Diseases and Dangerous Occurrences Regulations
SCADA	Supervisory Control and Data Acquisition
SOV	Service Operational Vessel
TRIR	Total Recordable Injury Rate
WT	Wind Turbine
WF	Wind Farm

Research Output

To date, the work completed for this thesis has lead to the following publications:

Peer Reviewed Journals:

(a) Anderson, F, McMillan, D, Dawid, R, Garcia Cava, D. A Bayesian hierarchical assessment of night shift working for OWFs. *Wind Energy.* 2023; 26(4): 402- 421. doi:10.1002/we.2806

(b) Anderson F, Dawid R, McMillan D., García Cava D. A Bayesian Reliability Analysis Exploring the Effect of Scheduled Maintenance on Wind Turbine Time-To-Failure. *Wind Energy.* 2023; 1-21. doi:10.1002/we.2846

Peer Reviewed Conference Papers:

(a) Anderson F, Dawid R, García Cava D, McMillan D. Operational Metrics for an Offshore Wind Farm Their Relation to Turbine Access Restrictions and Position in the Array. In: *Journal of Physics: Conference Series*; 2021

(b) Anderson F, Dawid R, McMillan D., García Cava D. On the Sensitivity of Wind Turbine Failure Rate Estimates to Failure Definitions In: *Journal of Physics: Conference Series*; 2023

Conference Presentations:

(a) Oral Presentation at the 20th EERA DeepWind Conference, Trondheim, Norway, 18–20 January 2023

(b) Oral Presentation at the 2022 Future Wind and Marine Conference, Edinburgh, Scotland, 17th February 2022

(c) Poster Presentation at the 2021 Future Wind and Marine Conference, Digital Conference, 18^{th} February 2022

(d) Poster Presentation at the 18^{th} EERA DeepWind Conference, Digital Conference, 13–15 January 2021

Industrial collaboration:

(a) An offshore wind farm operator provided data for analysis. Along with the organisation that provides their control room software, they provided expert knowledge and perspective to the analysis. Due to data sensitivity issues, the data provider wishes to remain anonymous.

1 Introduction

1.1 Background

1.1.1 Offshore Wind Power

The latest report by the Intergovernmental Panel on Climate Change (IPCC) confirms that *"Human activities, principally through emissions of greenhouse gases, have unequivocally caused global warming"* leading to *"widespread adverse impacts and related losses and damages to nature and people"* [1]. This evidence has led to the search for renewable alternatives to the generation of electricity from fossil fuels. The International Energy Agency (IEA)'s world energy outlook in 2022 estimated that renewable energy capacity is set to overtake fossil fuels by 2030 [2]. In fact, if we are to meet the objective of the Paris agreement of halving global greenhouse gas emissions by 2030, the IEA estimates that renewable energy capacity will have to triple. The majority of this growing contribution to renewable energy will come from wind and solar photovoltaics [2, 3].

In 2022, the installed wind power reached 909 GW worldwide [4]. The global wind energy council estimates that this number will increase by 680GW of new capacity in the next 5 years. Of this current 909GW capacity, offshore wind accounts for 64.3GW ($\approx$7%). By 2027, the share of offshore wind power is expected to rise to $\approx$12%, constituting 130GW new capacity [3]. Figure 1.1 presents IEA's estimates for how the share of renewable electricity generation will evolve in the coming years.

There are several features of offshore wind farms (OWFs) that drive this shift in the market. First, offshore sites often see higher average wind speeds and a more consistent wind resource, leading to generally higher capacity factors [6] . They

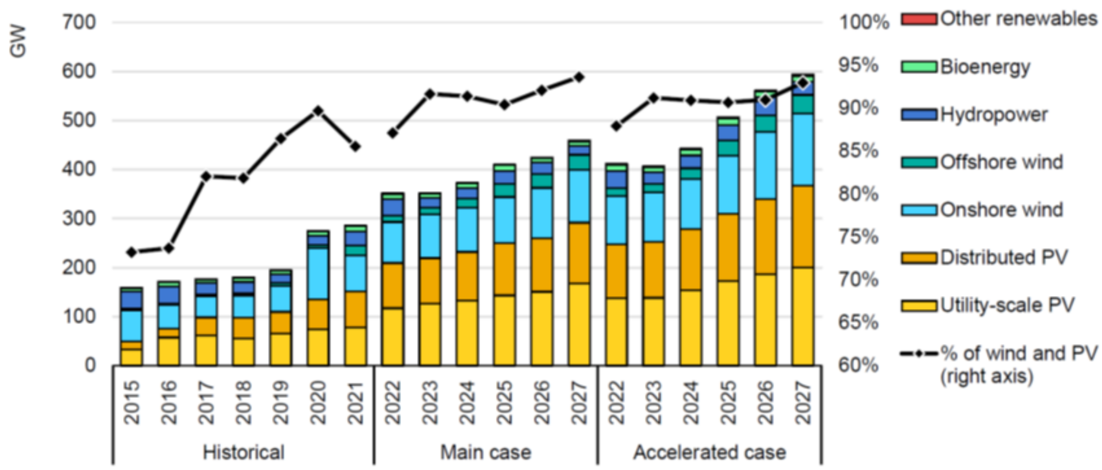


Figure 1.1: IEA renewable annual net capacity additions by technology, main and accelerated cases, 2015-2027. Taken from [5]

better support economies of scale via the enhanced possibility of developing larger turbines. The availability of large swathes of ocean which are suitable for offshore wind developments loosens the constraints imposed by scarcity of appropriate on-land installation sites [7]. They also increase the potential for public acceptability due to less visual and noise impacts [8].

In exploiting these potential benefits, however, OWFs are characterised by both increased capital costs and increased operations & maintenance (O&M) costs. The National Renewable Energy Laboratory (NREL)'s 2021 cost of energy review [9], for instance, estimates capital costs are 90% higher for offshore turbines compared with onshore (due to added costs for turbine substructures, electrical infrastructure and assemble/installation). They estimate that O&M costs are 142% higher for offshore turbines. The latter cost driver provides the context for this thesis. Since O&M costs are significant for offshore projects, cost reduction in this most lengthy stage of the project presents an opportunity for cost reduction [10].

1.1.2 Data Processes in Operation and Maintenance

From NREL's 2021 cost of energy figures, O&M currently makes up one third of a wind farm's cost of energy [9]. The potential for significant savings for offshore wind through a reduction of O&M costs has made O&M a prominent area of research in the offshore wind research community [11]. Leveraging this research to reduce operating costs is one way to ensure financial viability for OWFs under pressure from competitive Contract for Difference (CfD) auctions to place progressively low bids [12]. Operators can exploit this financial lever by continuously optimising their O&M strategy to address the unique challenges inherent in maintaining offshore wind turbines. This consists of reducing turbine unavailability in a cost-effective manner under the constraints of weather restrictions on safe access and expensive vessel charters.

OWFs generate vast amounts of operational data which can aid this optimisation process. Data sources come from meteorological sensors, turbine condition monitoring sensors, the Supervisory Control and Data Acquisition (SCADA) system, and maintenance records. These data can provide valuable insights into the performance, maintenance needs, and overall efficiency of the wind farm. However, on-going research focusing on analysis of all of these datasets implies it is as yet not leveraged to its full potential [13, 14, 15, 16].

The analysis and interpretation of these data streams have led to the development of O&M decision making tools of varying complexity, relating to both long-term strategy [17, 18] and day-to-day decision-making [19, 20, 21, 22, 23]. On top of this, the advent of increasingly sophisticated tools for on-site data collection, management, and analysis is becoming increasingly influential [13, 14, 15, 16]. The majority of these, however, are in the area of conditional monitoring - so much so that it now predominates in the wider O&M research space [24, 14]. Studies focussing directly on real-world operational maintenance data are comparatively scarce. There is a good reason for this: such data is generally hard to come by and can be unreliable [25]. Reder et al. [26] provide an insight of how this affects reliability analyses. In particular they highlight: the general lack of availability and poor quality of such data; the non-uniformity of its treatment, and the difficulty of augmenting reliability analyses with other data-streams. IEA's Wind task 33 also explored this theme, summarising that *"there is broad industry recognition of the relevance of reliability data collection and analyses for optimising both profit margins and LCoE* (Levelised

Cost of Energy) *. However, the lack of standards associated with reliability data is adversely impacting industry progress in addressing reliability issues."* [27].

A prominent problem for offshore wind farm data analysts is that offshore wind farm data may come in different formats and from various sources, making it difficult to integrate and analyse effectively [28, 29, 27]. There may also be limitations in data collection and storage systems at OWFs, which can hinder the collection of comprehensive and high-quality data [30, 28]. Factors such as limited sensor coverage [15], data transmission issues [15], and data storage capacity constraints [30] can contribute to gaps in the data collected, limiting its usability.

Such problems mean that OWF operators do not always have access to sufficient data to accurately describe their assets. This can lead to a variety of issues that can negatively affect operational efficiency. In the literature review of this thesis presented in chapter chapter 2, four key areas for development are highlighted. Namely, the potential for improving input data quality [27, 31, 15], data fusion [29, 15, 32], uncertainty quantification [33, 34, 35], and interpretability for decision makers [36, 11]. It follows that one way to improve operational & strategic decision making is to improve these data processes. In particular, Bayesian methods offer several unique features that map well to the areas for improvement listed above [37]: they are suitable for small and incomplete datasets [38], they have built-in uncertainty quantification [39], and they have the ability to combine different sources of knowledge [40]. Of particular interest to this thesis are the data that describe maintenance interventions and the data analysis tools which can be used to gain insight from them. Due to (i) corporate sensitivity of the data [41] (ii) generally poor data collection and processing practices [25] and (iii) a lack of scrutiny into "data fusion" of maintenance data with other data streams, data processing of maintenance interventions has significant room for improvement [14].

1.2 Motivation

There are two key motivations for this thesis. The first is a broader research context which encapsulates the growing role of offshore wind in the global energy mix; the significance of O&M costs in reducing the price of energy for offshore wind and the role that data processes play in that reduction. Based on the literature review conducted in section chapter 2, data processes for reliability and maintenance data are particularly under-researched [42, 33]. This has prompted a recent trend in attempting to address the uncertainty characterising reliability and maintenance data [43, 44, 45, 46]. On this topic, Shafiee [42] states that "*very few research has been conducted on the maintenance logistics planning under data uncertainty.*" Li et al. [25] elaborate on the issue, concluding that: *"The maintenance decisions determined under certainty are not adequate in modern MW/GW scale offshore wind farms, so a new series of solutions need to be developed to cope with uncertainty"*. Overall, the literature paints a picture that improvements are required in addressing this uncertainty when it comes to reliability and maintenance data. This research context might be bound in the following problem statement:

"*The value of operational data (especially maintenance records) in reducing O&M costs is not yet fully leveraged in the offshore wind industry.*"

The other key motivation is the involvement of the industrial partner, who has provided a large and easily accessible data set of relatively high quality for analysis. Advanced data management systems employed by said partner have gone a long way to eliminating the inhomogenity of the operational data available, removing a considerable obstacle to analysis of O&M activities of an offshore wind farm. Some studies have remarked on this obstacle. In particular, [47], in which data management techniques and workflow processes are explored for operational wind farm data. This PhD will turn its focus from the data management systems themselves towards data analysis possible from their results. In line with the trend in the literature [46, 43], it will focus on improving data processes via uncertainty quantification. More particularly it aims to provide a quantitative insight into the relationship between turbine performance and failure rates and the combination of factors that influence them. As depicted in figure 1.2, a number of factors will play a role and are typically described by the wind data ecosystem. Since data describing maintenance intervention has seen little scrutiny in the literature, and since there is a high quality operational maintenance dataset available for analysis, this thesis turns its focus to how various aspect maintenance intervention effects wind turbine performance.

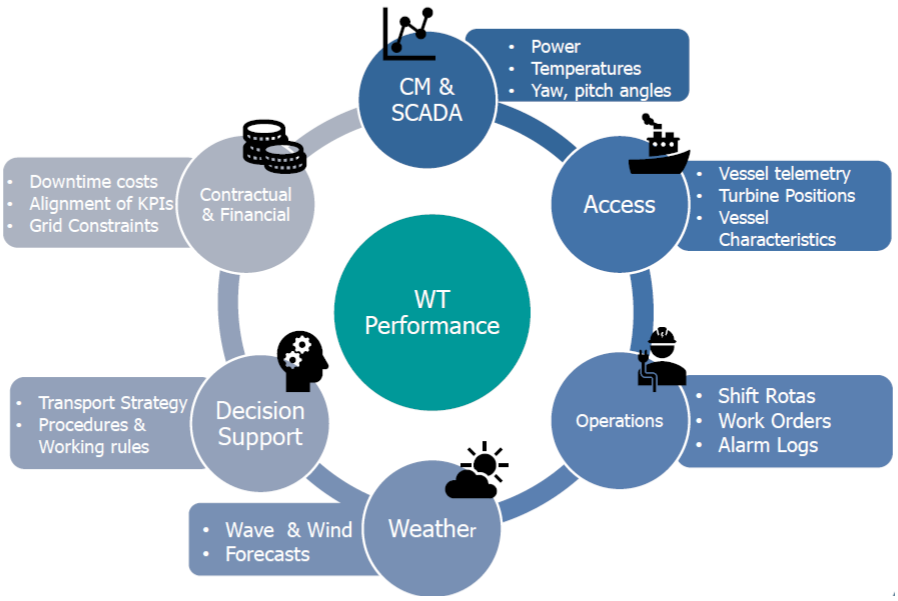


Figure 1.2: Summary of the the myriad aspects described within the 'Wind-Data Ecosystem' which effect wind turbine performance.

1.3 Aims & Objectives

Based on the problem statement that "*The value of operational data (especially maintenance records) in reducing O&M costs is not yet fully leveraged in the offshore wind industry*", the objective of this thesis is to answer the following research question:

"*How can operational reliability and maintenance data be better leveraged to support decision making and therefore reduce O&M costs in the industry?*"

This is a broad, open ended question, the answer to which encompasses many aspects. The aspects which are investigated within that open-ended question have been decided by a combination of literature review; input from the industrial partner and additional expert judgement towards the start of the project. The research question will be addressed by fulfilling the following objectives:

1. *Perform a review of the offshore wind data ecosystem to identify areas for improvement for data processes. Chapter chapter 2 presents a review of the wind data-ecosystem.*

2. *Use the available dataset to calculate relevant KPIs describing wind turbine maintenance intervention. Use those KPIs to explore the maintenance requirements of offshore wind turbines. The rationale behind this objective is presented in section section 2.6.1. The data mining methodology developed to calculate the KPIs is presented in section section 3.4. Section Frequentist Intervention Statistics presents a case study where the methodology is applied.*

3. *Investigate the utility of Bayesian data models in their application to operational maintenance data mining, and in leveraging value from operational maintenance data in general. The rationale for Bayesian reliability analysis as a solution to the areas for improvement identified in chapter chapter 2 is presented in section section 2.5. Chapter chapter 5 presents the methodologies developed in this thesis to leverage value from the data. Sections section 6.2 and section 6.3 present two case studies where the developed methodology is employed.*

4. *Evaluate the effectiveness of night shifts in increasing power production and availability based on up-to-date, real-world operational data. The rationale for this objective is qpresented in section section 2.6.2. Section section 6.2 presents a case study which addresses the objective.*

5. *Quantify the effect that annual services have on proceeding corrective works and in turn the reliability of wind turbines based on up-to-date, real-world operational data. The rationale for this objective is presented in section section 2.6.3. Section section 6.3 presents a case study which addresses the objective.*

The rationale for these objectives is presented in section section 2.6.

1.4 Contribution to Knowledge

The main contributions of this thesis to the research space are as follows:

1. It applies Bayesian models to operational maintenance data. In doing so, prior estimates of the impact of operational decisions can be quantified given limited data points. Uncertainty quantification is in-built, and multiple data-streams are used in the analysis. The models are white-box, and therefore interpretable. Two Bayesian methodologies are developed and utilised. The novelty of these can be summarised as follows:

    (a) A Bayesian hierarchical modelling approach is applied to a question of operational strategy - namely the question of the effectiveness of night shifts. Section section 6.2.4 discusses how the developed methodology allowed for uncertainty quantification with limited data points, and how the flexibility of the model allowed me to consider different operational scenarios.

    (b) A Bayesian reliability analysis is developed, the results of which are presented in section section 6.3. The methodology extends traditional reliability models into the Bayesian regime and incorporates time-dependent variables. Section section 6.3.4 discusses how the developed methodology is used to investigate the impact of scheduled maintenance on wind turbine time-to-failure in a new way.

2. It investigates the impact of maintenance intervention based on up-to-date, real-world evidence derived from a high-quality operational dataset. In doing so, several contributions to knowledge follow:

(a) Intervention metrics are presented for an offshore wind farm. Given the scarcity of publicly available maintenance data for OWFs, this in itself is valuable to the research community.

(b) The effect of tidal access restrictions on wind turbine maintainability are quantified. To my best knowledge, this is the first study to do so.

(c) It quantifies uncertainties associated with the pre-processing stage of a reliability analysis. An analysis of this type is valuable to the research community as it exposes the uncertainty surrounding published failure rates of wind turbines.

(d) It presents real world evidence for the effect of night shifts on wind farm performance. While previous studies have explored this subject in the past, none of the previous efforts have been based directly on real-world evidence.

(e) It presents real-world evidence for the effect of annual services on wind turbine reliability. While previous studies have explored this subject in the past, none have done so by reliability modelling of up-to-date, real-world data.

1.5 Thesis Outline

The rest of this thesis is structured as follows.

Chapter 2 presents a literature review. The first part of the literature review is dedicated to addressing objective 1. It presents a review of the offshore wind data ecosystem by first defines what a data ecosystem is (section section 2.2); then going on to summarise the data sources that contribute towards decision making at OWFs (section 2.3) and the decision making process tools used in decision making (section 2.4). key areas for improvement via better data analytics are summarised (section section 2.4.5), and these are used as a reference point throughout the thesis. The rest of the literature review is dedicated to presenting a rationale for the thesis objectives. Reasons for the use of Bayesian data modelling as a solution to the areas for improvement summarised in section section 2.4.5 are presented in section section 2.5. Last of all is a more focused review surrounding the focus of analysis within the thesis section 2.6.

Chapter 3 presents a methodology for extracting valuable information from the available dataset. It contains a definitions of Key Performance Indicators (KPIs) to be used in the following analyses (section section 3.2). Following these definitions is a data audit (section section 3.3) and a decsription of the data mining methodology used (section section 3.4).

Chapter 4 presents intervention metrics based on the methodologies developed in chapter chapter 3. The results presented here address objective 2. Section chapter 4 presents a case study of maintenance intervention data analysis using frequentist statistics. An anaysis of work procedures carried out on the site gives insight into the maintenance work carried, and this is used to derive assembly-level failure rates. The uncertainty in both turbine-level failure rates and assembly-level failure rates is explored by employing different data selection criteria.

Chapter 5 presents two Bayesian methodologies to be used in the subsequent analyses. Section section 5.2.1 presents the general principles of Bayesian analysis and introduces Bayesian networks and Bayesian hierarchical analysis. Section section 5.3 goes on to develop a Bayesian reliability analysis methodology for wind turbines.

Chapter 6 presents two case studies which emply the Bayesian methodologies developed in chapter chapter 5. Section section 5.2.1 uses the Bayesian hierarchical modelling technique developed in section section 5.2.1 to assess the effectiveness of night shifts in increasing power production and technical availability. Section section 5.3 goes use the methodology developed in section section 5.3 for

Bayesian reliability analysis of wind turbines. By using time-dependent covariates, the technique is employed to explore the effect of scheduled maintenacne on wind turbine time-to-failure.

2 Literature Review

2.1 Chapter Overview

This chapter presents a literature review. The first part of the chapter (sections section 2.2, section 2.3 and section 2.4) is dedicated to fulfilling the first objective listed in chapter chapter 1, namely:

"*Perform a review of the offshore wind data ecosystem to identify areas for improvement for data processes.*"

This chapter therefore provides a summary of the *Offshore Wind Data Ecosystem* (OWDE) for O&M. It does so via a number of steps:

1. **Section section 2.2** outlines what a data ecosystem is.
2. **Section section 2.3** describes the parts that make up the *Data Ecosystem* at OWFs that are pertinent to O&M.
3. **Section section 2.4** summarises decision making at OWFs, covering the role of input data to decision making tools in the operational, strategic and tactical echelons.

The second half of the chapter narrows the scope of the thesis from the initial literature review. It does so by two parts:

1. Section section 2.5 provides arguments for Bayesian data modelling as a solution to the four key themes identified from the review of the OWDE.

2. Section section 2.6 provides a more focused rationale for the thesis objectives presented in chapter chapter 1.

2.2 The Offshore Wind-Data Ecosystem

In its original use, an "*ecosystem*" is defined as "*a biological system composed of all the organisms found in a particular physical environment, interacting with it and each other*" [48]. In its extended use, it is a "*complex system resembling this*" [48]. The concept is therefore useful in describing any complex system characterised by the interaction of multiple participants, and has been adopted primarily in an industrial or business context. A *Data Ecosystem*'s purpose is to describe the process by which massive amounts of data are transformed via advanced analytics into useful information [49]. A Data Ecosystem borrows ideas from the business, digital and software metaphors. Its purpose is to describe the process by which massive amounts of data are transformed via advanced analytics into useful information [49].

Jansen et al. [50] state that three features define such ecosystems: network, platform and co-evolution. A *network* is a loose collection of stakeholders driven either by providing value to, or extracting value from, the given ecosystem. A *platform* is a means by which the various stakeholders generate benefits for themselves or other stakeholders. The name suggests a common software solution among actors, however its use is intended to note the broader family of services or tools by which value is extracted. *Co-evolution* refers to a process of increasing collective knowledge via collaboration and interaction between stakeholders.

An offshore wind farm can be thought of as a complex industrial ecosystem. It is characterised by the interaction of multiple stakeholders to produce low-carbon electricity from wind turbines [51]. In doing so, it is an example of an entity which gives rise to a data ecosystem. A prominent common objective of the stakeholders is to reduce the Cost of Energy (CoE) [52]. The data ecosystem's purpose is to support this goal. In the O&M phase of the project, cost reduction might be acheived via better decision making [42]. Better decision making does so via the production, analysis and visualisation of various data-streams which are important to key decision makers [49]. Specifically, the focus here is on the operational stage of a wind farm. In this context, the three terms *network*, *platform* and *co-evolution* which define an ecosystem can be summarised by the following points. A summary of the OWDE is shown in figure 2.1.

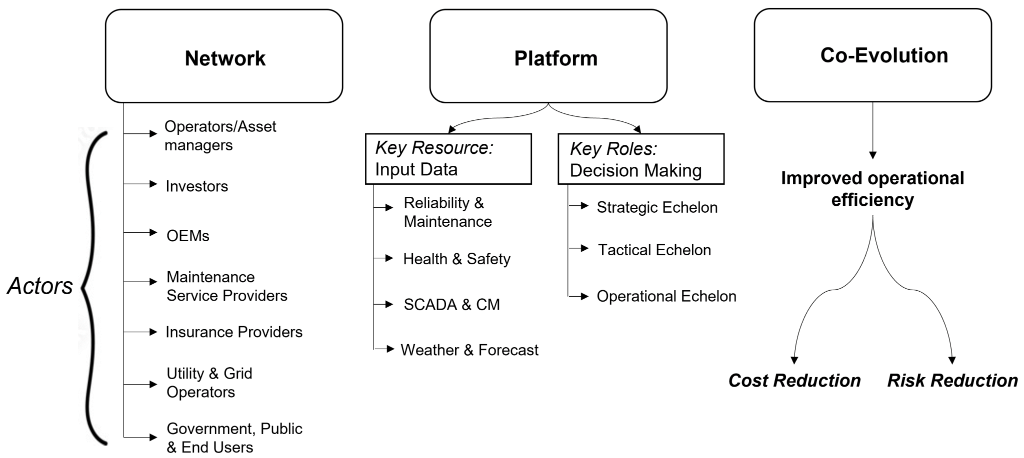


Figure 2.1: Summary of the key components of the offshore wind data ecosystem.

1. Network. The network is made up of (i) the primary stakeholders at an OWF and (ii) the wider research community. Between them, Gonzalez et al [53] and Tavner [54] define 6 stakeholders which are considered the primary actors, namely:

   (a) Operators and Asset Managers.

   (b) Investors.

   (c) Original Equipment Manufacturers (OEMs).

   (d) Maintenance Service Providers.

   (e) Insurance Providers.

   (f) Utility and Grid Operators.

   (g) Government, Public and End Users.

Refer to Gonzalez et al. [53] and Tavner [54] for a description of each actor's priorities. Note that these are broad categories and can be broken down further into lower-level actors.

2. Platform. The platform of the OWDE can be thought of as the collection of data mining tools that transform operational data into insight [50]. This review is essentially a review of the platform. The data mining tools and the data that drives them will be summarised in the following sections.

3. Co-Evolution. Co-evolution signifies the collective improvement of decision making throughout the operational phase of a wind farm within the wind industry. Indirect evidence for co-evolution can be obtained by considering several points.

   (a) Rapid expansion of the offshore wind market. Figure 2.2 shows the global installations of offshore wind from 2006 to 2021, as presented by the Global Wind Energy Council (GWEC). From 90MW of installed offshore wind capacity in 2006, new offshore wind installations have increased to over 20,000MW in 2021. The trend of growth is not expected to stop there. Collectively, European countries aim to increase offshore wind capacity by 89GW [55]. Asian countries (most prominently China, Taiwan and Vietnam) aim for up to 100GW by 2030 [56]. The US plan to increase their capacity from a single 12MW demonstration project to 28.1GW [55].

   (b) Rapidly decreasing costs of energy. This gearshift in the scale of deployment in the industry is made possible by recent progress in lowering the cost of OWFs. Figure 2.3 shows how prices have evolved in the UK. The black line represents figures from the Department for Business, Energy and Industrial Strategy (BEIS) shows the trend of LCoE for offshore wind in 2021 prices (estimated using the governments deflator [57] and figures

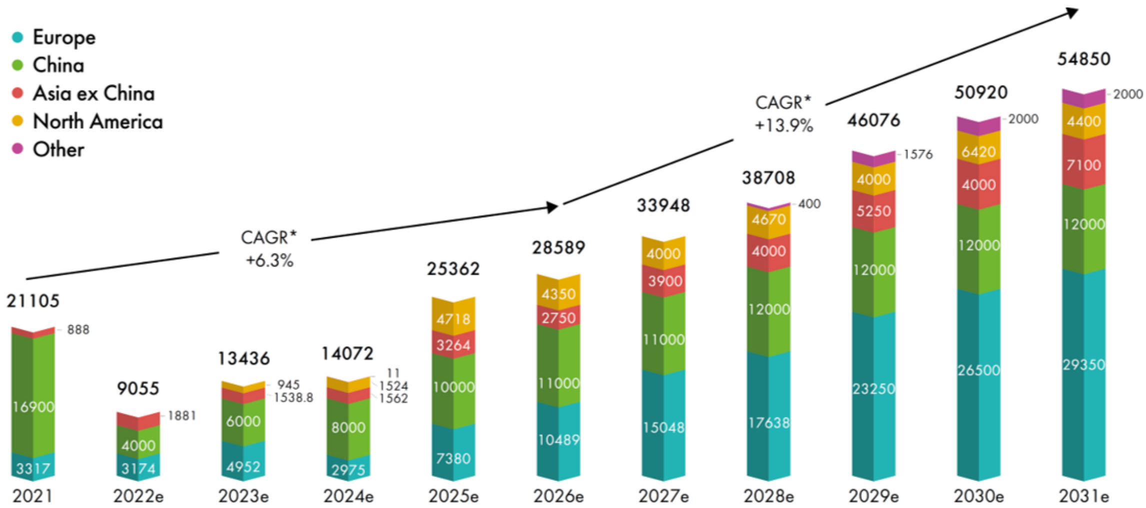


Figure 2.2: Global offshore wind market by region, taken from the Global Wind Energy Council's 2022 Global Offshore Wind report [4]. CAGR stands for Compound Annual Market Growth

from BEIS reports [12]). The orange line represents CfD strike prices published by the UK government, again in 2021 prices (calculated via the same method). Both show downward trends, providing evidence for decreasing costs of energy.

(c) Increasing research interest. Figure 2.4 shows the yearly research paper output for the Scopus search "offshore wind operations and maintenance". Taken cumulatively, there is a steady upwards trend in the size of the research space. Furthermore, this in itself is only indicative of a subset of the research space. Studies into wind turbine reliability [41, 58, 59], condition monitoring [15] and forecasting [13, 60] also see increasing scrutiny.

2.2.1 Resources, Actors, Roles and Relationships

As defined by Oliviera & Lóscio [61], there are five primary components that make up a Data Ecosystem:

1. An *Actor* is "*an autonomous entity such as an enterprise, institution or individual, which plays one or more specific roles in a Data Ecosystem*". Point 1

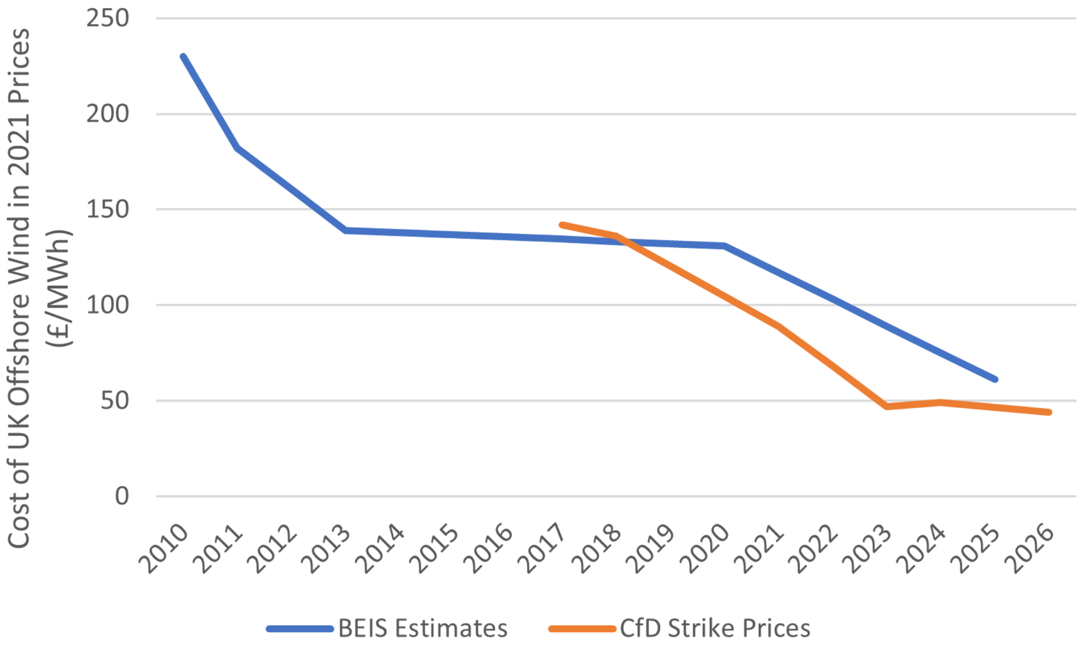


Figure 2.3: Cost reduction of OWFs in the UK.

of the list in the above subsection provides a suitable summary of the primary actors for a OWDE.

2. A *Role* is a function performed by an actor. These are not necessarily performed by one actor alone, and again may contain sub-roles within them. In this review, the key roles performed by decision makers are summarised in section 2.4.

3. A *Resource* is a valuable product utilised in the roles undertaken by actors. In generic data ecosystems, resources are datasets, software and infrastructure.

4. A *Relationship* is an interaction between actors within the Ecosystem. A prominent example would be arranging a maintenance contract between Maintenance Service Providers and Wind Farm Operators.

5. Additionally there are two essential *properties* of a data ecosystem: a networked character and self organisation.

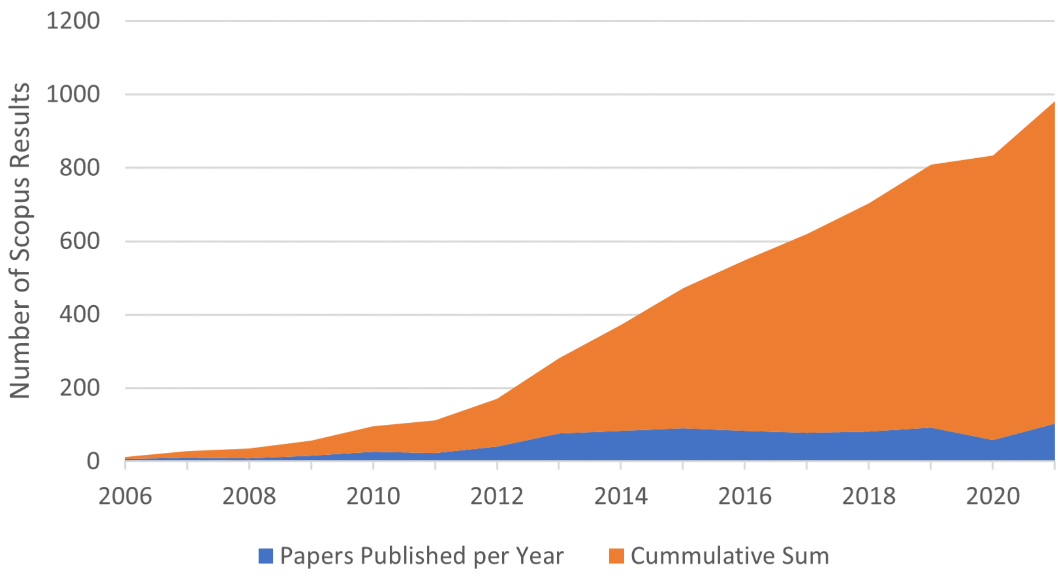


Figure 2.4: Results of the Scopus search "offshore wind operations and maintenance" in the abstract, title or keywords. Shows number of papers per year.

2.3 Resources - Input Data

The purpose of this section of the review is to assess the characteristics of the raw data, with a view to how they influence data analysis down the line. Pertinent datastreams for O&M researchers can be broadly categorised into 4 categories: *Reliability & Maintainablity*, *Weather and Accessibility*, *Peformance & Condition Monitoring* and *Health & Safety*. The quality of data in these categories can be assessed according to the 5V's of big data as laid out by Demchenko et al [62]. Namely:

1. *Volume* is related to the amount of data generated. This can refer to the amount of data in terms of memory requirements, number of tables or files, or length of time & scale of the data collection process. McKinnon et al. [63] provide an example of where the volume of input data is demonstrated to impact its effectiveness.

2. *Velocity* relates to the speed at which data is generated and processed. State-of-the-art tools such as condition monitoring systems require data to be processed

at or near-to real time. Gonzalez et al. [64] provide an example of where data velocity has the potential to impact decision making.

3. *Variety* is directly related to data complexity. It refers to the format of the data: whether it is structured (e.g. in a relational database with pre-defined data types), unstructured (consisting of different file formats, different storage platforms and incompatible data-types) or semi-structured (unstructured data provided a basic structure via some metadata). Leahy et al. [25] demonstrate how variety affects wind turbine reliability and condition monitoring.

4. *Value* relates to the added benefit afforded to the role being undertaken by an actor. Actors must balance the costs of collecting, storing and processing the data with the added-value it provides. Yang et al. [65] provide an example of an attempt to quantify the value of condition monitoring data to decision makers.

5. *Veracity* relates to both data consistency and data trustworthiness. Data consistency relates to statistical reliability. It describes the dataset in terms of format transactions, duplicated and missing data. Trustworthiness relates to the infrastructure which carries out data origin, collection, and processing methods. Tawn et al. [66] provide an example where the impact of data veracity is investigated on SCADA data.

The outcome of the input data review is summarised in tables 2.1 and 2.2.

2.3.1 Reliability & Maintenance

Reliability and Maintenance (R&M) data refers primarily to work orders and maintenance logs. Maintenance works at OWFs are recorded in different ways by different organisations (OEMs, operators, maintenance service providers). Traditionally, unstructured hand-written maintenance reports would record maintenance carried out. Examples of older research for onshore wind that were built on manual work orders are provided by WMEP (Wissenschaftliches Mess- und Evaluierungsprogramm) [76], WindStats [77], and Schleswig-Holstein Chamber of Agriculture (LWK) [77]. With

Table 2.1: Summary of data issues surrounding reliability and maintenance data (R&M) and H&S data.

Data-Stream	Issue	5Vs	Description	References
R&M	Data Quality	Veracity	Often incomplete and significant manual effort is required to draw valuable conclusions	[25, 67]
R&M	No Standard Taxonomy	Variety	No standard taxonomy for wind turbine failures	[25, 67, 68]
R&M	No Standard Failure Definition	Variety	No standard definition for wind turbine failures	[25, 69]
R&M	General lack of availability	Volume	Strict confidentiality practices make publicly-available failure data sparse	[25, 41]
H&S	Lagging indicators	Veracity	lagging indicators unlikely to be suitable for low frequency but high consequence events	[70, 71]
H&S	Shift in operating conditions	Veracity	Closer scrutiny required on how the industry shift towards tougher operating conditions impacts worker safety	[70]
H&S	No wind-specific legislation	Variety	Updating of the legislation relating to H&S in offshore wind required	[70]
H&S	Incompleteness	Volume	Not all wind farms accounted for in published statistics	[70, 72]

increased digitalisation in the industry, however, there is a shft towards digital work orders. This generally consists of a system where technicians can input maintenance tasks and materials consumed [25]. Such datasets can provide valuable insights into WT reliability and/or maintenance requirements for an OWF.

WT reliability analyses are a particularly integral part of decision making at OWFs [33]. They play a central role in reducing costs both at the design and operations phase of a wind farm. They indicate OEMs where resources should be directed to improve reliability [78], and indicate to wind farm operators where to direct their resources and how to optimise their operational strategy [79]. The value of reliability analyses have also been heightened by a prevailing lack of transparency in the industry [42]. Strict confidentiality practices are put in place by OEMs which they consider to protect their interests [25], but hampers every other stakeholder in making effective decisions to cut costs.

Table 2.2: Summary of data issues surrounding SCADA & CM and Weather & Forecast data.

Data-Stream	**Issue**	**5Vs**	**Description**	**References**
SCADA & CM	Standardised SCADA column names	Variety	There is no standard set of column names	[25]
SCADA & CM	Missing data	Volume	Uncertainties in data resolution, availability and completeness all can have significant effects on KPI calculation	[35, 66]
SCADA & CM	Standardised SCADA alarm codes	Variety	There is no standard set of SCADA alarm codes	[25]
SCADA & CM	SCADA alarm showers	Velocity	SCADA alarm showers prevalent around failure events make them difficult to analyse	[73]
SCADA & CM	Too many SCADA alarms	Volume	SCADA alarm codes are generated at too high a velocity to be useful to operators	[25, 74, 73]
Weather & Forecast	Spatial Resolution of wave climate	Veracity	Better understanding of the wave climate at a given site may be provided by better spatial resolution	[75]
Weather & Forecast	Post processing of forecasts	Veracity	Post processing of NWP forecasts for access	[13]

Despite this obstacle, there are several reliability analyses published in the literature. These are comprehensively reviewed by Pfaffel et al. [69] and Cevasco et al [41]. Since these review studies provide a wealth of information on the subject, a detailed review is not included here. However, two points are worth highlighting in gauging the value of this data for OWFs:

1. There are only two studies listed by Pfaffel et al. which provide failure rates for offshore turbines [80, 81]. Reliability analyses for offshore wind turbines are therefore especially valuable.

2. Despite the efforts of those review studies to extract trends from the publicly available data, the published failure rate figures are difficult to compare because the research space is so inconsistent [25]. This inconsistency stems not only

from different wind turbine concepts, environmental conditions, wind turbine ages and installation dates, but also from inconsistent data treatment [25].

Leahy et al. [25], Reder et al. [68] and Hahn et al. [27] all highlight the numerous issues associated with treatment of wind turbine reliability data. These can be summarised as follows:

1. General lack of failure data. See point 1 of the above list.

2. The lack of a standard failure definition in the wind industry [25]. What a researcher means when they say a turbine 'fails' varies from one study to the next [82]. Often, researchers are restricted by the format and detail of their available dataset. Failure definitions range from definitions which include automatic restarts [83, 84]; those which require human intervention to be considered [85]; and those which include conditions on downtime [86] or material usage [82]. This incongruity leads to a research space in which it is difficult to compare published failure rate data.

3. The lack of a standard taxonomy in the wind industry [68, 27]. There is also a large variation in how published failure rates are categorised by assembly. The Reliawind Taxonomy [86] is popular among academics: it is both specific enough to provide insight and simple to apply. The Reference Designation System for Power Plants (RDS-PP) is more popular in industry, especially in recent years [87]. The North American Electric Reliability Corporation provides the Generating Availability Data System (NERC-GADS) [88] which also sees use in industry.

4. Disparate data sources [25, 27]. Points 2 and 3 both stem from and influence a lack of consistency in exactly what reliability analysis looks like, to the point where relevant stakeholders (particularly operators) do not know the value of their own data [25].

5. General data quality issues [25, 27]. Failure data also tends to be incomplete and significant manual effort is required to draw valuable conclusions.

All of these points produce a measure of uncertainty in wind turbine failure statistics. All of them taken together produce a research space in there is significant uncertainty associated with failure rate figures, and this uncertainty is rarely addressed.

As noted by Dao et al. [43], wind turbine manufacturers are likely to have databases of historical failure data and so the problem of limited and poor quality data may not be as severe for them. However, there is still exists a problem in that historical data must be used to predict the reliability of future wind turbines. As summarised by Jenkins et al. [36], the industry is evolving so quickly (in terms of turbine power rating) that current failure rate data is unlikely to capture future failure rates accurately. This means that the industry is faced with a miss-match between the pace of low-velocity reliability data and the fast development of new turbines. Under these circumstances, R&M data has too few samples to guarantee statistical significance. This is even remarked upon in the latest SPARTA portfolio review, where in comparing direct drive to geared turbines, they state: *"As the technology is a recent advancement, more data is required to create strong insights about the differences between direct drive turbines and turbines with gearboxes." This means that there are numerous sources of uncertainty surrounding maintenance and reliability data [46]; stemming from (i) inconsistent data treatment (ii) low data velocity and veracity and (iii) a quickly evolving industry.*

2.3.2 Health & Safety

Health & Safety (H&S) data consists of incident reports describing staff injury, near misses and work place hazards. Within the UK all employers are obliged to report H&S incidents according to the RIDDOR (Reporting of Injuries, Diseases and Dangerous Occurrences Regulations) legislation [89].

The literature on the subject is sparse. The authority on the subject is G+ [72] - a global health and safety organisation with four main aims: incident data reporting, good practice guidance, safe by design workshops and learning from incidents. Outwith the G+ reports, sources of offshore wind H&S data are not widely available. For a detailed review of the H&S for the offshore wind industry, see Rowell et al. [70]. They highlight several issues, namely:

1. The use of leading safety indicators over lagging indicators to assess safety at OWFs. To date, the H&S indicator used most of in the industry is the total recordable injury rate (TRIR) [72]. TRIR is an example of a *lagging indicator* - a measure of incidents that have already occurred [90]. Several authorities on H&S indicators have highlighted the inadequacy of lagging indicators to asses H&S standards [71, 91, 92]. These studies instead suggest the use of *leading indicators*, which measure occurrences that aim to predict future H&S performance [90].

2. Closer scrutiny on how the industry shift towards tougher operating conditions impacts worker safety. Rowell et al. [70] predict that the risk profile for the industry is going to increase. They put this down to factors such as increasing distances to shore; increasing number of technician transfers; the emergence of floating wind farms. Note that metrics describing maintenance interventions are presented in chapter chapter 4, which relate to the risk profile for technician transfers.

3. Updating of the legislation relating to H&S in offshore wind. H&S data veracity would likely be improved by offshore-wind-specific legislation. Like offshore wind R&M data, H&S data suffers from a lack of specificity for the wind industry.

Pfaffel et al., in their survey of KPIs within the industry, found that only 5 out of 27 respondents make use of HSE KPIs in their organization [93]. Additionally, there is a relative dearth in H&S related research compared to, for instance, studies seeking financial optimisation in the O&M phase. This would imply that H&S data is not as valuable to stakeholders as other data-streams. However, given the poor

H&S performance compared to the oil and gas industry [70], analysis of H&S data presents an opportunity for improvement in the industry. Additionally, given the risk profile of offshore operations is due to increase in the coming years [70], analysis of H&S data becomes increasingly important.

2.3.3 Performance & Condition Monitoring Data

Performance & condition monitoring data relates to 3 sub-categories:

1. SCADA Data. SCADA systems provide high level information to operators and OEMs at a turbine, farm or fleet level [47]. The implementation of SCADA systems in large-scale utility turbines is ubiquitous in the wind industry. Since offshore wind turbines are unmanned structures for the most part operating in far-off locations, remote measurement is a useful asset for monitoring their performance [47]. They encompass readings of a range of environmental & control variables, electrical characteristics and part temperatures [94]. Table 2.3 gives a summary of typical variables recorded - the precise list of measured parameters will vary from one SCADA system to another [94]. As summarised by Tautz-Weinert & Watson [94], vibrations, oil pressure level and filter statuses could be also be recorded by a WT SCADA system, but more often recorded separately in a dedicated condition monitoring system. There no standard set of monitoring equipment [15] or measurement nomenclature in the industry [25].

2. SCADA alarm codes. Additionally to the typical set of variables listed in table 2.3, SCADA systems provide alarm codes related to WT operational status. As described by Leahy et al. [31], there are generally three different types of alarm code:

   (a) *Information messages* are generated to communicate those changes in a WTs operational mode not associated with faults.

   (b) *Warning alarms* are generated when some measured variable comes close to its pre-defined operational limit.

(c) *Fault codes* are generated when those limits are exceeded by a given variable.

Gonzalez et al. [95] present a more detailed breakdown of alarm log types, including: grid conditions, environmental conditions, WT operational state, manual stops or restrictions, maintenance activities and component malfunction.

3. Condition Monitoring Signals. Condition Monitoring Systems (CMS) are similarly prevalent in OWFs, and serve a similar purpose to SCADA systems - that of tracking turbine health and performance [15]. However, their purpose goes beyond the high level performance monitoring of SCADA systems to provide a more detailed insight into turbine health at the sub-system level. The goal is to diagnose fault locations and severity accurately, and even to provide prognosis on time to failure [94]. Within the field of academic research for offshore wind O&M, condition monitoring is perhaps the primary contribution [33]. The term encompasses a wide range of input data and a similarly wide range of data analysis techniques. Several studies have attempted to capture this range via reviews: notably Stetco et al. [96], Garcia Márquez et al. [97], Tchakoua et al. [98], Rindali et al. [99] and most comprehensively (as of 2023) Badihi et al. [15]. Broadly, condition monitoring is the real-time measurement of a component's condition [100]. It can be further broken down into three levels of functionality [25]: fault detection, fault diagnosis and fault prognosis/prediction. The challenge lies in selecting which sensors to install and which data analysis techniques to use, as justifying the cost effectiveness of any condition monitoring is very difficult. For a more detailed overview of each of the CM techniques, see Badihi et al. [15].

SCADA data is often produced at a 1Hz sampling rate. However, in an effort to reduce the transmitted data bandwidth from the wind farm, it is usually averaged into 10 minute intervals [29]. Qui et al. [73] found that alarms tend to be generated in volumes too high to allow effective management from operators. This is especially pronounced during periods of faulty operation, in which alarm codes "alarm showers" are generated which are extremely difficult to untangle [73]. The sampling rate of condition monitoring tools (especially for vibration diagnostics) tends to be much higher, but ultimately depends on the condition monitoring system. Regarding dedicated CMS, vibration-based CM is the most prominent area of research in the field [15]. This type of monitoring is widely relied upon for detecting faults in components which produce a vibration signal. Most commercial CM systems rely on vibration signals of major components such as the blades, gearbox and drive-train bearings [101]. Most of these systems rely on accelerometers, which register a wide range of high frequencies (1-30kHz). At lower frequency ranges, there are velocity sensors; at lower frequency ranges still there are displacement sensors. At higher frequencies, there are spectral emitted energy sensors. In the high frequency range covered by accelerometers, the most common signal analysis technique is the Fast Fourier Transform (FFT). The aim of FFT in this scenario is to analyse the frequency spectrum of the vibration signal for variations in its harmonic components. By identifying faults in the component and relating them to unusual behaviour in the harmonic components, the signal can be used for fault detection in the future [102]. Other methods for detecting anomalous behaviour focus on the time-domain of vibration signals (namely envelope analysis [103] and the Hilbert Transform [104]) and time-frequency domain analysis.

SCADA and CMSs produce structured data which is easily incorporated into data analytics tools. SCADA systems record a set of standard parameters nearly everywhere they are equipped. However, some systems contain additional variables - Leahy et al. [25] recommend introducing a standard set of SCADA column names. They also recommend introducing standardised SCADA alarm codes, which can be difficult to process due to their variety [25]. The number and type of additional condition monitoring signals varies from one farm to the next. Performance and condition monitoring signals are therefore characterised by some industry-wide variety.

Just like any other component of the wind farm SCADA and CM signals are subject to occasional failure or communication loss. This leads to corrupted or missing parts of the dataset, which may in turn lead to biases or incorrect results. Results of subsequent data analytics would therefore be more robust if methods were put in place to handle missing data [66]. SCADA alarm codes have more substantial problems with veracity. As noted by Kaidis et al. [74] and later by Leahy et al. [25], interpreting SCADA alarm codes often requires significant manual effort from researchers if they are to act as a basis for future analysis.

SCADA systems are a cheap means of providing easily interpretable data which is indicative of turbine health and performance. This data is useful to wind farm operators and OEMs in providing high level information at a fleet or farm level [31]. It is, however, found largely lacking by the fault detection requirements of modern OWFs. Reasons for this are outlined by Yang et al. [105]: SCADA systems do not collect all signals necessary for in-depth condition monitoring analysis, and the data that is collected is often at too low a sampling frequency for use in fault diagnosis. Typical fault indications that arise from SCADA data changes are considered to have impractically poor prognostic lead times.

CMSs show substantial potential to predict wind turbine failures. The question for WF stakeholders is whether the savings from a particular CMS outweigh the cost of implementation. This is not at all an easy question to answer. However, with the trend in the industry towards bigger, higher capacity turbines [11], it becomes easier. As noted by Fischer & Coronado [106], the cost-benefit of a given CMS begins to lean further towards benefit as potential opportunity cost from turbine downtime increases. CM systems are discussed in more detail in Badihi et al.'s extensive review [15] on the subject.

SCADA alarm codes have potential utility for operators. Leahy et al. [25] summarise the efforts of the research community to use alarm codes both for condition monitoring purposes [94] and reliability analyses [64]. At present, however, issues with data volume, velocity and veracity effectively effectively nullify a large part of their potential.

Table 2.3: Typical parameters recorded by SCADA system [94]

Environmental	**Electrical Charactersitics**	**Part Temperatures**	**Control Variables**
Wind Direction	Active Power	Gearbox Bearing	Pitch Angle
Ambient Temperature	Reactive Power	Gearbox Lubricant Oil	Yaw Angle
Nacelle Temperature	Power Factor	Generator Bearing	Rotor Shaft Speed
Wind Speed	Generator Voltages	Generator Windings	Generator Speed
	Generator Phase Current	Main Bearing	Fan Speed/Status
	Voltage Frequencies	Rotor Shaft	Cooling Pump Status
		Generator Shaft	Number of Yaw Movements
		Generator Slip Ring	Set Pitch Angle/Deviation
		Inverter Phase	Number of Starts/Stops
		Converter Cooling Water	Operational Status Code
		Transformer Phase	
		Hub Controller	
		Top Controller	
		Converter Controller	
		Grid Busbar	

2.3.4 Weather & Forecast

Weather data provides information about the wind resource and wave conditions. As noted by Seyr and Muskulus [33], whether data is relatively easy to come by for researchers, as there are several meteorological measurement and re-analysis campaigns that provide data to the public. The critical application of weather data in O&M is in accessibility assessment. Significant wave height (H_{sig}) is often considered the limiting factor due to operational limits imposed on access vessels (usually 1.5m). The significant wave height is the mean of the third highest of the waves, measured from trough to peak, over a given sampling period [107]. However, other variables are of interest to schedulers. Namely, wind speed limits are imposed on certain maintenance works [108]; and more accurate weather conditions can be obtained by including wave period and wave direction in forecasting tools [13].

Decisions are typically made using publicly available forecast data from dedicated campaigns. In Europe, the European Centre for Medium-Range Weather Forecasts (ECMWF) [109] serves as the go-to for medium-range predictions. This model provides forecasts up to 10 days in advance with a horizontal resolution of 9km. Other prominent global forecasts include the GFS (Global Forecast System, USA) [110], and RHMC (Hydrometeorological Center of Russia). [111]

On top of forecast data, operators rely on various sources of real-time data both for short term maintenance planning and long term resource analysis. Wave conditions are typically measured by wave bouys. Draycott et al. [75] note that they provide good quality measurements of wave height, wave period and often directional measurements, but often suffer from poor spatial coverage. Wave Radar offer the potential for massive spatial coverage improvements. However, they are relatively new technology still being developed. See Draycott et al. [75] for novel wave measuring devices that could progress decision making in the offshore wind industry. Many Europe-based researchers depend on the FINO campaign [112], which provides weather measurements for a location in the North Sea.

For wind speed measurements, turbines themselves are equipped with anemometers and wind vanes at the top of their nacelles for supervisory control purposes. Mainly, these are used to determine if the wind speed is sufficient to start turbine operation [113]. Measurements are therefore reflected in 10-minute SCADA data on a turbine-by-turbine basis. Wind farms often also have on-site meteorological masts (met-masts) for wind resource characterisation pre-operation. Depending on the size of the farm, there could be one met-mast or multiple at different locations. A typical met mast will have a number of anemometers installed at different heights.

Current practice in the short-term relies primarily on campaigns like the ECMWF for medium-range wave height forecasts to be used in day-ahead decision making. The requirements of short term decision makers are therefore low-volume. Long-term decision making typically requires time-series data of wind speed and wave height.

This is also a relatively low volume data demand. However, post processing of this raw forecast data can provide improved forecast at a time horizon greater than 6 hours; and processing of wave bouy data can provide more accurate predictions for within-day decision making.

The ECMWF has a 1 hour forecast step and updates every 12 hours. Typically, wave bouys process time-series of wave height (and other parameters) on-board, before summary statistics of a selected period are transmitted via radio telemetry. The selected period is normally around 10-20 minutes [75]. Anemomemeters produce signals at a similar frequency.

Forecast data, wave and wind are usually presented as structured data types in a time-series format. They are therefore characterised by low variety. Where datasets may differ is in their resolution or in the number of variables measured.

The ECMWF is characterised by inevitable forecasting inaccuracies. However, it is consistently favoured to other global forecast campaigns for its accuracy. Draycott et al. [75] note that wave buoys provide accurate measurements of significant wave height, albeit at poor spatial coverage. Measurements from turbine anemometers, however, often leads to inaccurate wind speed measurements, due to the interaction of the rotor with the incident wind [113].

Forecast data is valuable to short term decision makers. As described by Gilbert et al. [13], additional value can be extracted from forecasts such as the ECMWF by post processing of their data. In doing so, analysers can engineer additional features [13] or increase spatial resolution [114, 115]. Accurate on-site measurement campaigns are valuable mostly at longer time-horizons. Time-series of the wind and wave climate, for instance, are key inputs for strategic decision making tools.

2.4 Roles - Decision Making

Decision making in the context of OWFs can refer to both long term strategising, medium-term alterations to the strategy and short term scheduling of maintenance tasks. Shafiee [116] classifies these time horizons into three echelons: strategic, tactical, operational respectively. Long-to-medium term decision making makes use of various decision support tools, as is comprehensively discussed by Seyr and Muskulus [14]. This is a multifaceted process, where various aspects must be considered if the tools are to be accurate and useful. The probability of failures, weather conditions and economical factors inform the choice of maintenance regime at the site. Short term decision making oversees the implementation of this maintenance regime on a daily basis and is a similarly complex problem, involving optimisation based on limited information and continuous review based on new information. This section is concerned with the current practice in decision making within the wind industry, and is discussed in more detail within the following subsections. It will focus more on how the available data sources support each stage of the decision making process - more information about the decision making process itself may be found from Shafiee [116].

2.4.1 Operational Decision Making

Shafiee designates three important aspects to operational decision making [116]: scheduling of maintenance tasks, routing of maintenance vessels and measuring the maintenance performance. The actors that influence the decisions and the data sources that are available to aid their decision making is presented in figure 2.5. Most significantly, the primary decision maker is the marine coordinator. They must make decisions on vessel dispatch, the routing of vessels and the prioritisation of maintenance tasks - all while implicitly assessing the value of the various data-streams available to them, safety measures and the priorities of the various stakeholders.

Most valuable to the marine coordinator in terms of input data is the weather forecast, point measurements of weather, condition monitoring signals and a list of work orders. Stock-Williams & Swamy [20] provide a succinct description of how the process plays out at an OWF. First, a list of service orders is produced by the asset management team containing corrective tasks, inspections, annual services and balance of plant activities. A scheduler uses this list of service orders to create a transfer plan on the morning of the day it is to be enacted, which assigns maintenance tasks to teams of technicians and teams of technicians to Crew Transfer Vessels (CTVs). Component replacements require the use of Jack-Up Vessels (JUVs), and so can be considered separately. Weather permitting, the transfer plan is carried out. This is where uncertainties stemming from incorrect fault diagnosis, human error, repair times and the weather forecast [13] take effect; the transfer plan nearly never goes exactly as expected. Those maintenance tasks which were completed are now recorded as having been completed, and those that are not are recorded as such by the asset management team.

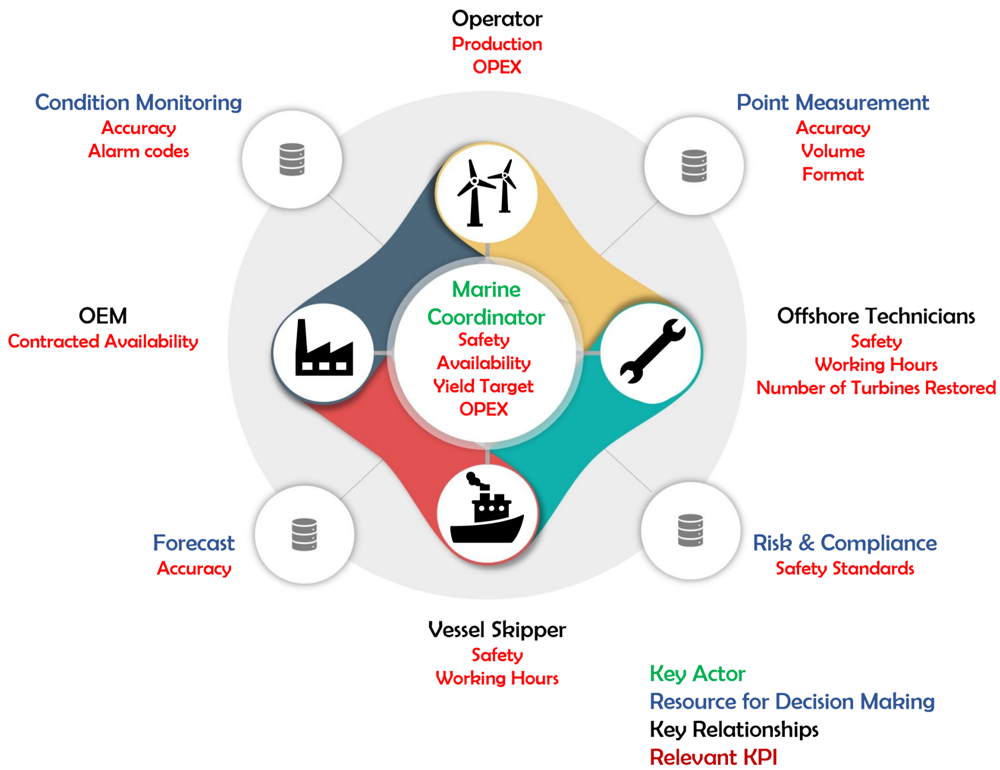


Figure 2.5: Summary of the offshore wind-data ecosystem, as it applies to short-term decision-making.

Improving the scheduling of maintenance tasks and the routing of maintenance vessels means addressing the uncertainties stemming from incorrect fault diagnosis, human error, repair times, the weather forecast [13]. Decision support for this area focuses on solving an optimisation problem. Stock-Williams & Swamy do so via meta-heuristic simulations [20]. Before this study, researchers focused on solving a set of equations. Dai et al [117] defined the problem as a Routing and Scheduling Problem of a Maintenance Fleet for Offshore Wind (RSPMFOWF). In this they set the objective of minimising cost of service vessels and production losses. They show that the proposed equations are difficult to solve, and so are only applied to small problems. Stålhane et al. [118] compare 2 formulations, finding that the one based on an efficient heuristic labeling algorithm outperforms the traditional solver. Dawid et al. [119] presented a number of novel additions to the previous methodologies in the field. First, they demonstrated the value of incorporating uncertainties into their optimisation framework. They concluded that "*a significant 14% increase in maintained turbines can be achieved by employing the proposed methodology highlights the importance of including uncertainties when modelling offshore vessel routing*". Second, they created visual aids to illustrate the policies produced by the tool. Third, they built in a "risk aversion factor" to facilitate tailoring of policies to different scenarios. Tan et al. [120] considered the impact of power loss, personnel costs, transportation costs, resource costs and other maintenance cost factors. They specified that their model "*belongs to the category of nonlinear optimization problems with constraints.*". Raknes et al. [121] highlighted three novel attributes of their mathematical optimisation model: the inclusion of several wind farms at different locations; the coverage of several shifts in the planning horizon; and modelling different types of vessels.

Dawid et al. [119] highlight three primary categories of uncertainty for maintenance task scheduling tools. Li et al. [46] argue that incorporating uncertainties into their maintenance strategy optimisation provides "*more feasible and reliable suggestions*" for the decision-maker. They suggest as part of their future works section that updating the uncertain input parameters to their model would be beneficial [46]. Based on Dawid et al. [119]'s primary uncertainty categories, the following input parameters have scope for improvement:

1. Uncertainty on task duration. Seyr et al. [45] demonstrate that uncertainty in the repair times of components has a significant effect on the modelled availability and lost production due to downtime for OWFs. They conclude that "*The use of an accurate representation of repair times is therefore strongly suggested and further research in distributions of repair times is encouraged. For multiple turbines in a wind farm, the effects on the results of using a distribution are expected to be even larger, as the repair times influence the scheduling of maintenance actions in a nonlinear way.*" Based on this conclusion, deriving conditional probability distributions for repair time estimates may be of benefit.

2. Uncertainty on the expected weather conditions at the wind farm. This can be addressed via probabilistic forecasting of weather windows & access indicators. The benefit of probabilistic access forecasting over the deterministic methods currently practiced at OWFs has been demonstrated by Taylor & Jeon [34], Browell et al. [122] and Gilbert et al. [13].

3. Possibility of failure misdiagnosis. Probability of failure misdiagnosis can be improved via more accurate condition monitoring tools. Badihi et al.'s review [15] suggests several avenues for improvement, namely: combining disparate condition monitoring signals for improved signal-based CM; hybrid approaches incorporating both signal-based and physics-based models; improved prognosis tools and farm-level modelling.

Improved communication of outputs would also advance operational decision making. Bessa et al. [123] highlight the importance of effectively communicating the uncertainty of weather forecasts as a means of promoting enhanced decision making; managing user expectations; promoting user confidence and reflecting the state of the science. Gilbert et al. [13] adhere to this conclusion and produce easily interpreted visualisation of their access forecast. For Bessa et al. [123], usefulness of information for short-term schedulers comes form a cost-loss model under uncertainty. Browell et al. [124] and Catterson et al. [125] exemplify this value in the field of access forecasting. For short-term scheduling tools, Stock-Williams and Swamy [20] present the first methodology to quantify the benefit of their scheduling tool.

The remaining aspect of operational decision making is measuring the maintenance performance. This is the opposite side of the operational coin: you can either improve decision making by improving planning in the first place (scheduling/vessel routing), or by retrospective analysis of operational data. Since operational data is generally hard to come by for wind energy researchers, the latter option has seen comparatively little attention in the literature. Pfaffel et al. [126] and Cevasco et al. [41] provide comprehensive reviews that include performance data for OWFs. However, a close scrutiny of how maintenance performance correlates to WF performance via retrospective analysis of operational data is virtually non-existent in the literature. Gonzalez et al. [53] provide a list of KPIs which would be suitable for the purpose.

2.4.2 Strategic Decision Making

Shafiee designates four important aspects to strategic decision making [116]: wind farm design for reliability, location and capacity of maintenance accommodations, selection of wind farm maintenance strategy and outsourcing of repair services. For strategic decision problems, either mathematical optimization models or O&M simulation models are used as a basis for decision making. Both depend fundamentally on input parameters, some of which are very difficult to come by for researchers. The most immediate improvement that can be made to the strategic decision making echelon is therefore in better quantification of O&M model input parameters. Most pertinently, reliability and cost data. Carroll et al. [82] provide the most comprehensive and up-to-date reliability analysis of offshore wind turbines. This covers the first 5 years of operations for turbines from one manufacturer. 5 years is only a small proportion of the lifetime of a wind farm. Additionally, failure rates can vary between different manufacturers. [127]. On top of this, costs (for vessels, technicians and costs of spare parts) are also sparse in the literature [33].

WF design for reliability refers to the design of the farm. Specifically, this is concerned with turbine layout in a grid pattern, and location of the turbines with respect to the incident wind/wave field. Cazzaro et al. [128] split this into three scales: (i) the macro-scale, concerned with wind farm location given a large swathe of land (ii) the meso-scale, concerning the shape of the wind farm and (iii) the micro-scale, concerned with placement of turbines within that shape. On the subject, Shafiee states: "*The literature on optimization of the design for OWFs is significant. When looking at the published works, there can be found several factors influencing the wind farm design (e.g. overall energy yield and the initial investment required). However, only a few number of models have taken the RAM* (Reliability, Availability, and Maintainability) *issues into consideration*". This is still largely the case. However, van den Brink et al. [129] provide at least one example where O&M costs are considered on the macro-scale. They concluded that the LCoE impact of variations in the O&M costs was considerable.

One way the resources described in section section 2.3 can aid integration of RAM data into WF design for reliability by quantifying the effect of weather conditions on WT failure rates and maintainability[130]. Several studies have explored the effect of meteorological conditions on wind turbine failure rate. Reder et al. [131] stress the potential benefit of altering the constant failure rate used to estimate wind turbine reliability by environmental covariates. In doing so, operators might alter their maintenance strategy to react to, for instance, seasonal periodicity in wind speed. The relationship between wind speed and wind turbine reliability was established relatively early on in the industry's development. Tavner et al. [132, 133] contributed the majority of this early research, identifying a significant cross-correlation between failure rates and wind speed. Faulstich et al.[134] corroborated these early findings. Early researchers also identified correlations between temperature, humidity and proximity to coast and failure rate, implying that the relationship was between reliability and weather, as opposed to solely wind speed [133]. This work was followed up by Wilson & McMillan [135], who quantified the impact of weather effects on failure rates by employing a Bayesian technique - namely they used Markov chains and Monte Carlo simulations. They went a step further by quantifying the effect that wind speed-dependent failure rates have on operational expenditure. Later, Reder et al. [136] devised a methodology for analysing environmental conditions ahead of turbine failures in more detail. They also explored an alternative approach using a naive Bayesian network to predict WT failures [131]. Seasonality is explored by a number of studies. Slimacek & Lindqvist [137] and Reder et al. [136] both observed higher failure rates in winter months, where wind conditions are harsher. Recently, Pelka et al. [127] incorporated weather variables into a reliability analysis of wind turbine power converters. On the micro-scale, quantification of reliability with respect to position in the array is something that has never been done before, to my best knowledge. This can be taken further by also exploring the effect of a given site's bathymetry on turbine maintainability [138].

On the location and capacity of maintenance accommodations, again there exists a potential for improvement from refinement of model input parameters for simulation/optimisation models. An interesting aspect here is that the wind industry is developing rapidly. Future sites most probably will use WTs that are far bigger, and have a higher power rating than those currently operational [36, 3]. On top of this, more sites are likely to be located further from shore and in deeper waters [70, 3]. Seyr and Muskulus [14] note that simulation models tend not to be very well suited to this task, as optimisation involves running the model several times under different input assumptions. Optimisation models, such as the ones proposed by Besnard et al. [139] are more suitable.

For wind farm operators, a significant strategical aspect to decision making is the selection of the maintenance strategy. O&M activities fall into one of two categories (i) preventive maintenance (PM) and (ii) corrective maintenance (CM). The former describes maintenance which is undertaken to lower the risk of a failure in the future. The latter describes maintenance which performed upon the failure of a component. To this day, corrective maintenance accounts for most maintenance actions at OWFs

[33]. Scheu et al. [140] focus solely on corrective maintenance in their model. However, most models include scheduled maintenance in some form, but prioritise the importance of corrective works [141, 19, 142, 143].

Studies which have scrutinised the impact of scheduled maintenance in the literature tend to optimise the time between preventative maintenance actions. Carlos et al. [144] derived an optimal 113 days between scheduled maintenance visits using a Monte Carlo simulation method, using a Spanish failure database. Zhong et al. [145] devised a decision making tool based on a nondominated sorting genetic algorithm which allows the user to balance the objectives of turbine reliability and maintenance cost. Besnard et al. [146] used a stochastic optimisation model to show that preforming service maintenance tasks opportunistically during periods of low wind could save 32% of the transportation and production losses when compared to more conventional approaches. Byon et al [147] used a partially observed Markov decision process to derive an optimal preventative maintenance policy. Pattison et al. [148] include annual service campaigns in their novel architecture and system for the provision of Reliability-Centred Maintenance (RCM) such that cost-effective PM management could be achieved. Incorporating the effects of service history (preventative maintenance history) into decision making tools therefore sees some scrutiny. This might be addressed better by the input data reviewed in section 2.3 by incorporating preventative maintenance into reliability models fit to the data.

Within the categories of preventative and corrective, one might also include the sub-categories of condition-based maintenance (CBM) [149], opportunistic maintenance (OM) [150], and risk-based maintenance (RBM) [151]. In the first of the sub-categories, maintenance is performed based on condition indicators (e.g. signals from the CM system or inspection reports),[149]. Opportunistic maintenance describes the situation where multiple components are repaired at once in an attempt to fully utilise a given maintenance intervention [150]. Risk-based maintenance strategies use techniques such as the criticality of the failure modes is determined through a failure mode and effects analysis (FMEA) [10] to formally optimise their maintenance strategy based on the relative risk of component failures. Despite this prevalence of CM studies in the literature, attempts to quantify the impact of the developed tools prove to be difficult. The earliest was by McMillan & Ault [152], who used discrete-time Markov chains to evaluate effectiveness of a CM system on various metrics such as failure rate, O&M costs and availability. The work of May et al. [153] followed up, using a Hidden Markov Model to model reduced failure types, false alarms, detection rates and 6 month failure warnings form CM systems.

In reality, state-of-the-art WF maintenance strategies often consist of some combination of the two categories and three sub-categories. Some research in the field attempts to capture this. Lu et al [154], for instance, propose a combined opportunistic CBM optimization approach in which economic dependence exists among components. Tain et al. [155] use a similar approach, and both present financial benefits to their strategies compared to corrective maintenance. Other studies have explored the combination of CBM, periodic maintenance and risk-based maintenance.

In their review of condition-based maintenance strategies, Kang et al. [156] highlighted 3 aspects which are important to such modelling: the influence of degradation, the multi-component problem, and maintenance constraints. All three might provide opportunities for better data/data analysis tools to prompt better decision making:

1. Degradation modelleing. Leites et al. [157] stress that wind turbine prognosis (and RUL estimation) is a field very much still in its infancy. The potential benefits are therefore largely unexplored. They highlight that "*according to [158] prognostic techniques are among the most important areas to be addressed and improved in the field of WT O&M.* They go on to also highlight several points for improvement for wind turbine RUL: improved availability of run-to-failure data; exploration of accelerated ageing environments; development of real-time prognostic algorithms; uncertainty representation and management techniques and prognostic performance evaluation.

2. Multi-component degradation models via modelling of component dependencies. See Rinaldi et al. [159] for a description of the various dependencies. Niu et al.

[160] give an example of how RUL prediction of a system can be improved by considering stochastic dependence among components.

3. Maintenance constraints. Improved detail of work order data could shed light on the impact of maintenance constraints (such as limited technicians, specific vessels, maintenance tools, and changeable marine climate) on maintenance performance.

The final aspect of Strategic decision making is outsourcing of repair services. Wind farms typically have a 5-year warranty in which the OEM handles maintenance [161]. After this, they may either extend their original contract, take maintenance in-house or employ another 3rd party maintenance service provider, [161]. Shafiee stated that this "*has tremendous potential for research*" as wind farm operators are often interested in contracting out maintenance activity. As shown by Hawker and McMillan [161], contractual arrangements can have a significant impact on the performance of a wind farm. They found that distance to sites from the maintenance depot is a key consideration, as well as the form of warranty taken (i.e. time- or energy-based availability). Marugń et al. [162] present a techno-economic model for avoiding conflicts of interest between owners and maintenance suppliers of OWFs. They showed the value of the model in addressing the conflict of interest between the two parties via adjusting penalisation, incentives, resources, and adequate control of availability in the contract. Zhang and Zhang [163] attempt to maximise the interests of a common service provider and multiple operators. They argue that the representing the different interests of all parties was important in reducing the impact of maintenance activities on power productions of OWFs. A comparison of maintenance and performance KPIs of wind farms which are in/out of warranty would be of benefit here. SPARTA explored this issue in their 2017/18 portfolio review, concluding that wind farms with No OEM have a higher average production based availability but wind farms with Full OEM have a higher capacity factor [164]. More examples of this kind of analysis from industry benchmarkers or researchers working with operators would be of benefit [42].

2.4.3 Tactical Decision Making

Shafiee designates four important aspects to tactical decision making [116]: spare parts inventory management, maintenance support organisation and purchasing or leasing decisions. An important general consideration for tactical decisions is that outputs from strategic decision making tools can be regularly updated with new operational data as it becomes available. These decisions are dynamic in that they require regular synthesis of old information with new data. In this regard, Bayesian methods, which can quickly update model parameters, are particularly well suited [165]. See section section 2.5 for more details on Bayesian updating.

The problem of spare parts management is a pertinent one for the wind industry. It is made more complicated by burgeoning strain on the wind turbine supply chain [166]. Tusar et al. [167] provide a concise review on spare parts inventory management. They frame the problem as balancing the cost for holding excessive spare parts and the cost of downtime because of spare part shortages. Many of their recommendations also translate to the strategic echelon: more accurate statistical failure modelling and learning from historical failure data bases. However, they also recommend regular stock inspections, awareness of lead times and the implementation of multi-echelon inventory system. They break down the problem in into three stages: demand forecasting, inventory management, and data processing [167].

Maintenance support organisation means working out resource requirements in terms of vessels and technicians. OWFs rely primarily on CTVs for daily operations. Removal of main components requires hiring very expensive JUVs. Depending on the circumstances of the farm in terms of distance to shore, water depth etc., they might also opt for the use of helicopters and service operational vessels (SOVs) in carrying out their maintenance. On top of this, wind farms are a complex collection of mechanical and electrical components that require technicians with different skill sets to maintain. The issue has therefore seen a reasonable level of scrutiny in the literature. Besnard et al. [168], for instance, optimise the number of vessels, helicopters and length of maintenance coverage (12h vs. 24h) in their maintenance support model.

Dalgic et al. [169] also compare various configurations of vessels on both the day shift and the night shift. They conclude that "*10 out of the most costly 17 configurations do not have CTV for Night Shift' and that 'the number of CTVs during Day Shift and Night Shift is distributed evenly (or close to even) in best configurations such as 4-4 and 5-4, because the resources are utilised in an optimum manner with minimum redundancy*". Where improved data processing could improve maintenance resource allocation is by developing better methodologies to integrate operational data into maintenance support organisation. A good first step would be to see how relevant KPIs (e.g. number of vessels per turbine, average number of technicians per vessel) relates to maintenance performance.

Purchasing or leasing decisions can be seen as an extension of the above point. The challenge is in determining the optimum combination of vessel/helicopter purchases and leases to cut OPEX costs. A significant part of the problem is addressing jack-up hiring strategy, which are expensive both to purchase and to hire in the spot market [170]. Dalgic et al. [171], Stålhane et al. [172] provide decision making tools related to hiring decisions for jack-up vessels. An interesting prospect is resource sharing, as explored by the Crown Estate [173].

2.4.4 Summary: Importance of the 5Vs

This section summarises the above points relating to the relative impact and influence of the 5Vs within the offshore wind data ecosystem.

1. Volume. The volume of O&M data required by the use cases outlined in section section 2.4 varies. The typical ability of a given OWF data ecosystem to meet that requirement also varies, and typically varies in different was depending on the dataset in question. The volume requirements for reliability and maintenance data are such that the researcher can calculate a statistically significant value for the mean number of turbine/system/subsystem failures, and also be able to characterise the statistical uncertainty in that mean. Having enough

samples of historical failures of a turbine similar enough to the one in question (w.r.t power rating, turbine concept etc.) and similar enough to the site in question (w.r.t environmental conditions) to be able to define a probability distribution would likely meet both requirements. Given (i) the speed at which the industry is evolving and (ii) the general lack of availability of reliability data in the industry, the typical data ecosystem falls far short of this requirement. Health and safety data is also an area where the data volume is currently insufficient. This is because stakeholders tend to underestimate the importance of collecting relevant KPIs [93, 70]. In contrast, some authors have commented on the potential for SCADA alarm codes if they were not characterised by such impractically high volumes.

2. Velocity. The question of velocity is perhaps most prevalent for SCADA and condition monitoring monitoring data. While SCADA at 10-minute data is sufficient for a broad overview of turbine health, it is often considered as insufficient for the purposes of condition monitoring [15]. Some researchers have suggested that high frequency SCADA data could offer an opportunity for cost cutting in condition monitoring [174]. However, tailored condition monitoring systems based on higher frequency measurements seems to offer the current best solution to condition monitoring [15]. SCADA alarm showers around failure events presents another velocity-related issue.

3. Veracity. Broadly speaking there are 3 areas which are impacted by data veracity. The first is the data quality of reliability and maintenance data. Often this is incomplete and significant manual effort is needed to draw valuable conclusions [25]. The second is health and safety data, for which Rowell et al. [70] has highlighted several veracity-related issues. The third is the spacial coverage of wave buoys/wave radar & the post processing of forecasted weather data.

4. Variety. The impact of variety stems predominantly from a prevailing lack of standards in many areas of data collection and processing within the wind industry. This is especially relevant for reliability and maintenance data, which lacks both a standard taxonomy and and standard on data collection and processing. However, issues in variety also affect SCADA data (by not having a standard set of SCADA alarm codes) and health and safety data (by not having any legislation relating directly to offshore wind).

5. Value. Each of the data streams listed above is very valuable to OWF operators and contribute to good performance. As things stand, reliability and maintenance data is highly valuable because high quality and freely accessible data-sources are very difficult to come by. While many stakeholders do not value health and safety data as highly, Rowell et al. [70] argues that the wind industry is lagging behind the oil and gas industry in this respect. They also argue that health and safety data will become increasingly important as the industry transitions to further offshore sites in deeper waters, some of which are likely to be floating. The value of condition monitoring data is indisputable given the proportion of scrutiny it receives in the OWDE research space. Given the shift towards higher power ratings, the value of condition monitoring data is likely to continue to increase. Weather data collected on-site is arguably the area with least potential for further value to be extracted via research. However, post-processing of forecast data has a significant potential for improvement and therefore value to OWF stakeholders.

2.4.5 Summary: Improving Data-Assisted Decision Making

From the above review, this thesis defines four broad areas to improve upon if more value is to be extracted from the offshore wind-data ecosystem: *Data Quality. Data Fusion*, *Uncertainty Quantification* and *Applicability to Decision Makers.*

1. Data Quality. Improving data collection and processes presents the most immediate opportunity for improvement in data-assisted decision making. All analytics tools are only as good as the input data used to drive them. The following points can be taken from the literature:

   (a) Reliability and maintenance data. For recommendations of improvements in reliability data, refer to the points made by IEA task 33 [27]. These points are still widely relevant to the industry. Namely [27]:

      i. Make sure you get access to all relevant data;
      ii. Identify your use-case and be aware of the resulting data needs;
      iii. Map all wind turbine components to one taxonomy / designation system Map;
      iv. Align operating states to IEC 61400-26;
      v. Train your staff understanding, what data collection is helpful for;
      vi. Support data quality by making use of computerized means;
      vii. Share reliability data to achieve a broad statistical basis;
      viii. Develop comprehensive wind-specific standard based on existing guidelines/standards;
      ix. Develop component- / material-specific definition of faults, location, and severity

   (b) SCADA. Leahy et al. [31] only highlight one key issue pertaining to SCADA data. Namely, they call for standardised column names. However, Tawn et al [66] showed that wind power data from SCADA systems have typical median levels of missing data of 2.70% for the power variable and 1.57% for wind speed, and that methods for missing data improve the quality of subsequent data analysis. Pfaffel [35] also note that uncertainties in data resolution, availability and completeness all can have significant effects on KPI calculation. These data-centric uncertainties should therefore be considered.

(c) SCADA alarm codes. SCADA alarm codes are negatively affected by issues in data volume, velocity and variety. In order for operators to make better use of alarm codes for fault detection/reliability analyses, they overall volume needs to be reduced. The velocity (or so-called "alarm shower") at the approach of a failure needs to be addressed for there to be any discernible context. Lastly, Leahy et al. [31] recommend standardised alarm code signals, which would alleviate problems in variety.

(d) A prominent challenge for the quality of data from CM systems is in selecting the optimal collection of sensors [15]. This applies to (i) which method of CM is used; (ii) how many sensors are installed and (iii) integrating different CM systems with the turbine control and SCADA system. Increasing the veracity of CM signals can come from increasing the variety of signals within an integrated CM system. However, the question of how much variety is a difficult one to answer. Operators especially want to avoid information overload, where there are problems with volume & velocity (too much data too fast).

(e) Weather and Forecast. For the purposes of O&M, real-time weather parameters are accurate. Better understanding of the wave climate at a given site may be provided by better spatial resolution, as potential provided by several emerging wave-measuring technologies [75]. Post processing of NWP forecasts shows promise to improve short-term decision making. A notable study in this area is from Gilbert et al. [13].

(f) Health & Safety. Rowell et al.'s [70] review of H&S data highlights a number of important points related to H&S data. In particular they recommend (i) the use of leading safety indicators over lagging indicators to assess safety at OWFs; (ii) closer scrutiny on how the industry shift towards tougher operating conditions impacts worker safety; (iii) continued collaboration with G+; (iv) updating of the legislation relating to H&S in offshore wind.

2. Data Fusion. There is a significant, and mostly untapped potential improvement of various data-analytics via combining data-streams which are usually analysed separately. An important point in this regard is that improved data management systems are required for this purpose. The variety of the data ecosystem as a whole means that various sources making it up are often made up of different formats. Integrating them therefore often takes significant manual effort from practitioners [29]. Opportunities in this area can be summarised as follows:

   (a) Combining condition monitoring signals. Multi-parameter modelling is 1 of 12 recommendations made by Bahidi et al. [15] for improving WT condition monitoring. Since maintenance scheduling depends on interpretation of CM signals by experts, decision making is made easier when corroboratory evidence of a failure comes from multiple sources and the attendant uncertainty is reduced. Bahidi et al. go on to say: "*According to the reviewed literature, relying on more than one monitoring sensor (e.g., multiple sensors in different locations and of different types) or multiparameter monitoring often improves the chances of successful detection and diagnosis of incipient damages or faults at an early stage.* Fischer & Coronado [32] echo this point of view in their earlier review of CM user experience. They state that "*data from multiple sensors and possibly different monitoring techniques are combined to allow inferences that would not be possible based on a single sensor or technique.*"

   Data fusion of CM signals could be especially powerful if various cheap and reliable signals were analysed together. Taking, for example, vibration signals together with temperature signals may still be cheaper than several other CM methods. The most obvious 'easy win' would be to use SCADA signals to augment other CM signals, as they are available anyway. Turnbull et al. [175] provide an example of vibration signal analysis augmented with 10-min SCADA data.

(b) Incorporating leading safety indicators into access forecasting tools. A step on from the vessel motion tool presented by Gilbert et al. [13] would be to incorporate leading safety indicators into the decision making process. Given the recommendations of Rowell et al. [70] that (i) the demand for more emphasis on H&S measures and (ii) the transition to leading indicators, incorporating H&S considerations into decision making tools could be of benefit. This could also translate to the tactical and strategic time horizons.

(c) Combining failure prognosis with access forecasting. Having more precise information on both when a turbine will fail *and* when a suitable whether window is available to perform a preventative maintenance task; and combining these data-streams would make for easier decision making for the scheduler. Simple progress could be made just by overlaying the outputs of access forecasts and component prognosis tools. This would formalise (and therefore potentially reduce human error) the process already implicitly undertaken in the brain of the scheduler.

(d) Incorporating weather data and service history into reliability models. As outlined in section 4.1.5, the effects of many covariates have previously been explored in the literature. Incorporating environmental, operational and design parameters to make more comprehensive reliability models would provide more insight into reliability drivers [127]. In this regard, reliability analysis techniques which have been implemented in fields other than wind turbine reliability (e.g. time-dependent covariates [176, 177], hierarchical models [178], frailty effects [179]) could provide improved analyses.

3. Uncertainty Quantification. Uncertainty quantification has seen relatively little scrutiny in all of the above aspects of the above review [33, 15]. However, good operational decision making involves understanding operational risk. Operational risk is, in turn, uncertainty quantification. The following considerations fall under this category:

(a) Modelling of uncertainty of input variables (failure rates, repair times and cost parameters) of operational decision making tools. In Seyr & Muskulus' review of decision support models for operations and maintenance for OWFs [33], they repeatedly state the need of improved handling of uncertainties. To quote them: "*The existing models can capture most influencing factors; however, handling of uncertainty in the inputs is still an area that can be improved upon. Different wind farm layouts and locations can be included in the existing models, albeit being complex in implementation. The existing models lack an accessible way to handle uncertainties in the input and provide information about the variability in the output. Currently, mean values are reported and sensitivity analyses must be done by manually varying.*"

(b) Probabilistic access forecasting. Taylor & Jeon [34] and Browell et al. [122] demonstrate the value of probabilistic time series modelling over their deterministic alternatives. Gintautas & Sørensen [180] demonstrate that more sophisticated methodologies in accounting for weather forecast uncertainties lead to more available operational hours for maintenance. Gilbert et al. [13] demonstrate the value of probabilistic forecasts using statistical post-processing of Numerical Weather Prediction (NWP) for OWF access is by Gilbert et al. [13].

(c) Addressing uncertainties in KPI calculation. Pfaffel et al. [35] explore how uncertainties in data handling affect KPI calculations. They highlight that "*changes in the data resolution, data availability, as well as missing inputs, can cause considerable uncertainties*". Similar research into maintenance KPIs, in particular failure rate calculations and cost parameters, would be beneficial.

4. Interpretability to decision makers. The final key point lies in interpretability. The output of data analytic tools is more valuable when presented in an easily understandable format to decision makers [123]. It also must be relevant to

their particular set of requirements. The following considerations come under this category

(a) Translating improvements in accuracy to improvements in cost. Data analytics tools are more likely to be taken up in the industry if their usefulness can be demonstrated. In the field of access forecasting, Browell et al. [122] and Catterson et al. [181] have done so via the use of cost-loss models. In the field of reliability analysis, Carroll et al. [78] have used an operational cost model to explore the consequence that different reliability statistics have for cost estimates of different WT concepts.

(b) Applicability of reliability statistics to new wind farms. Reliability data for wind turbines is limited in the public domain. On top of this, the offshore wind industry is growing rapidly. This rapid growth has manifested in rapidly increasing turbine sizes and a transition towards direct-drive turbines [36]. This means that current reliability statistics are outdated for modern wind farms. Given the issues with velocity of reliability data in the industry, this presents a problem. Jenkins et al. [36] outline a method for tackling the problem via expert elicitation. Such methods for translating outdated data to new turbines are useful to all stakeholders. The industry is in particular need of tools which can build on partially outdated failure data, while incorporating new insight as new data becomes available. In this respect, Bayesian models show promise as a potential solution [182].

(c) Interpretability of Machine learning and AI. Machine learning and AI have a huge potential to advance almost all aspects of the wind-data ecosystem [11]. However any value that is to be derived from them depends on how easily the results can be explained by the analyst and subsequently interpreted by the decision maker [15].

2.5 Bayesian Modelling in Relation to Operational Data Mining

The rapid development of the offshore wind industry has coincided with the advent of increasingly sophisticated statistical approaches - in particular machine learning (ML) methodologies [183]. In modelling the O&M phase of an OWF, several researchers in recent years have called for better handling of uncertainty in data analytics tools. To quote some prominent examples:

1. Li et al. [46]: *"The maintenance decisions determined under certainty are not adequate in modern MW/GW scale offshore wind farms, so a new series of solutions need to be developed to cope with uncertainty."*

2. Dao et al. [43]: *"Although input data uncertainty is a critical factor that wind energy system operators need to deal with, 19 existing simulation models vastly ignore the uncertainties of data in reliability and O&M cost evaluation."*

3. Seyr et al. [33]: *"The existing models can capture most influencing factors; however, handling of uncertainty in the inputs is still an area that can be improved upon."*

In addressing these points, a researcher would ideally consider both epistemic and aleatory uncertainty [141]. In endeavouring to do so, Bayesian estimation and inference presents a number of unique advantages. Some of these advantages have been explored in previous studies within the field of wind energy research [184], however the utility of their application is by no means fully realised in the wind industry [165].

2.5.1 Aleatory Uncertainty Quantification

The first advantage relates to the modelling of aleatory uncertainty, describing the natural uncertainty inherent in a given system. While this is largely unavoidable, there is value in quantifying it and in turn understanding the risk it poses to operational decisions. In their extensive review of decision support models [14], Seyr and Mulkulus stressed that this statistical uncertainty in input factors needed to be explored more thoroughly to advance the field. For instance, with respect to failure rates, production losses, forecast accuracy, and repair times. They went on to do so, first investigating the effect of uncertainty in repair times [185] on lost production, then combining this with uncertainty in forecasts used for decision making [186, 187]. The significance of uncertainty in input variables was a conclusion corroborated by Scheu et al [188], who investigated the influence of uncertainty of component failure rates on WT availability. Dao et al. [56] also explored this theme, by using fuzzy logic as a method to handle data uncertainty in operational simulation models.

The Bayes method provides a means of easily quantifying uncertainty in input variables [39]. To my best knowledge, Bayesian methods have not been used to undertake a thorough investigation of WT failure rate, and the associated aleatory uncertainty. Rather, they have been used in a predictive capacity to estimate the impact of various features on failure rate. This is done by Reder & Melero [131] as part of their investigation into the dependence of failure rates on meteorological conditions. Likewise, Wilson & McMillan used Bayes' rule to derive wind-speed dependent failure rates [135]. Another important factor to consider in this regard is that aleatory uncertainty is increased for reliability estimates based on data from different sites - or even the same site operating under different strategies [46]. The flexibility afforded by Bayesian models could be useful for accounting for this uncertainty. Hierarchical models, in which distinct groups (e.g. turbines from different wind farms) share information between themselves but are still given individual probabilistic specifications, could be particularly useful [189, 190]. Regarding decision making hierarchical modelling has the potential to improve decision making by sharing statistical strength among distinct groups [191]. Section section 5.2.1 elaborates on this point.

2.5.2 Model Flexibility in Relation to Epistemic Uncertainty

The second group of advantages relate to epistemic uncertainty - that which arises from modelling assumptions and gaps in domain knowledge. There are three key points to consider in using Bayesian models to address epistemic uncertainty in offshore wind O&M.

1. **Suitability for Small & Incomplete Datasets**. Data relating to reliability and maintainability of wind turbines is generally difficult to come by for researchers [68]. When it is available, data quality is often an issue [25]. This point has been emphasised and elaborated upon by prominent review studies in the field, namely: Seyr & Muskulus [14], Reder et al. [26], IEA Wind task 33 [27], Pfaffel et al. [126] and Cevasco et al. [41]. The issues that can be drawn from these studies are summarised in section section 2.3.1. This thesis argue that these difficulties point directly to a Bayesian outlook as a solution. In contrast to almost all ML methodologies, Bayesian models have an inherent suitability for small datasets [38]. Robust inferences are possible with datasets for which frequentest approaches are inadequate. Despite this, there have been limited applications of Bayesian models in WT reliability analysis. Huang et al. [192] provide a case study for a generic reliability analysis using fuzzy lifetime data. Examples of the use of Bayesian networks for reliability analyses are provided by Wilson & Huzurbazar [193], Boudali & Dugan [194] and Nannapaneni & Mahadevan [195]. These papers use prior distributions and hierarchical modelling strategies to estimate component and system reliability based on some observed conditional probability. Nannapaneni & Mahadevan [195] note the advantages of such approaches in modelling both epistemic and aleatory uncertainty.

2. **Bayesian Updating**. This is an advantage which has gained some momentum in wind energy research. By implementing a Bayesian model, prior beliefs are constantly updated as new information becomes available. Bayesian models therefore provide a natural means by which to incorporate new measurements into decision making tools [196]. Most relevant here is Bayesian updating in O&M planning. There is an immediate application in updating parameters associated with condition monitoring of turbines. Herp et al. [197] used Bayesian

classifiers in power curve modelling, providing a case study for updating inspection plans in risk-based decision making. Likewise, Song et al. [198] used Bayesian updating in health state monitoring for wind turbines. Plumley et al. [199] and Asgarpour & Sørensen [200] provide more particular diagnostic and prognostic tools involving Bayesian updating of features due to sensor measurments and maintenance actions respectively. Bayesian updating can also be particularly useful in incorporating data from inspections, as has been done by [201] and [151].

For the most part these studies deal with health monitoring for particular components. This thesis proposes that the same concept of Bayesian updating can also be useful when applied to higher level reliability analyses, and longer term tactical decisions. For instance, Bayesian updating of failure rates at the turbine level could justify changes in strategy or resource allocation. This thesis, for example, presents a model to estimate the advantage provided by night shifts in lowering opportunity cost in chapter chapter 6. This is one example of decision making on mid-term time-scale. The strategy might be altered depending on whether the data differs significantly from prior beliefs. This can assessed quickly on (e.g.) a week-to-week or night-to-night basis, as new data becomes available.

3. **Combining Different Sources of Knowledge**. Operating and maintaining an offshore wind farm involves decision making based on a myriad of different factors. In the short term, marine co-ordinators rely on various sources of information to make decisions. This might involve forecasts of significant wave height and wind speed; outputs of diagnostic and/or prognostic tools; availability of spare parts, vessels and technicians (with different experience levels and qualifications); and the co-operation of different organisations, each with different priorities. In the mid-to-long-term, operators can amend their strategy to improve operational efficiency. This can be a similarly multi-faceted process, with different sources of information to consider. Their foresight can be aided

by advanced analytics of various historical data streams, however synergies between different sources of knowledge are as yet not fully realised. Operators might also be interested in hypothetical what-if scenarios involving new strategies. In this case, historical datasets would need to be augmented by expert knowledge or outputs from decision making tools.

Bayesian models are well suited to dealing with this multifaceted nature [40]. They are able to incorporate knowledge of different accuracies and from different sources in a mathematically coherent way [202]. Li et al. [203] provide an example of their dataset being augmented by previously published data and expert opinion in their reliability analysis of floating offshore WTs. Dinwoodie et al. [141] used Bayesian Belief Networks (BBNs) in combination with an operational model to simulate dynamic decision making. Lazakis & Kougioumtzoglou [204] also combine BBNs with other decision making tools - in their case a hazard identification and failure mode, effects and criticality analysis. There is, however, some novelty in combining data-streams via Bayesian models, as well as the decision making tools associated with them [165]. This novelty presents an opportunity for operational efficiency improvement, especially as more advanced data management systems begin to have more of an influence in the industry.

These advantages suggest a set tools which are flexible enough for researchers to extract value from operational data from offshore wind farms, while addressing some of the epistemic challenges associated with doing so.

2.5.3 A Bayesian Perspective on Decision Making

Making improvements on operational decision making tools for the future would benefit from a more thorough treatment of both aleatory and epistemic uncertainties [14]. Previously, Zitrou et al. [205, 206, 207] explicitly set out to tackle both of these categories within the field of offshore wind. In the first of the studies cited [205], epistemic uncertainty is handled by expert knowledge, but statistical uncertainty is also accounted for in the model. The second [206] focuses specifically on epistemic uncertainty via an expected value of the perfect information approach. The third again highlights how epistemic and aleatory uncertainty is modelled using a general marked point process model [205]. While these models are probabilistic (in that they use Monte Carlo simulations to derive probability distributions of variables), they are not Bayesian, as they do not depend on Bayes rule. Likewise, cost modelling tools which simulate the O&M phase via Monte Carlo simulations are probabilistic, but do not necessarily make use of the advantages of Bayesian modelling.

These tools are designed for use in the pre-operational stage of a wind farm. This thesis proposes that there is value to be had in an alternative approach - in retrospective analysis of operational data. This is where the work in this thesis presents novelty and improvements to data mining processes. In outlook, this is more closely aligned with the work of Dinwoodie et al. [141] alluded to above. Decisions at offshore wind farm are dynamic [42]. Changes in strategy require synthesis of historical data and some form of hypothesised solution [208]. Bayesian modelling provides a means to do this with small datasets, allowing a framework of new data integration and integration of different sources of knowledge. By design, it is flexible and allows for uncertainty quantification [38].

Figure 2.6 shows how the advantages of Bayesian models map well to the themes for improvement for the wind-data ecosystem identified in chapter 2.

2.6 Rationale of Thesis Objectives

The primary focus of this thesis is in extracting value from the *Operations* dataset. This objective was arrived at by 3 key considerations:

1. Analysis of operational maintenance data in the literature is comparatively sparse.

2. Due to the efforts of the industrial partner, the dataset *Operations* is of a higher quality than is typically available to researchers.

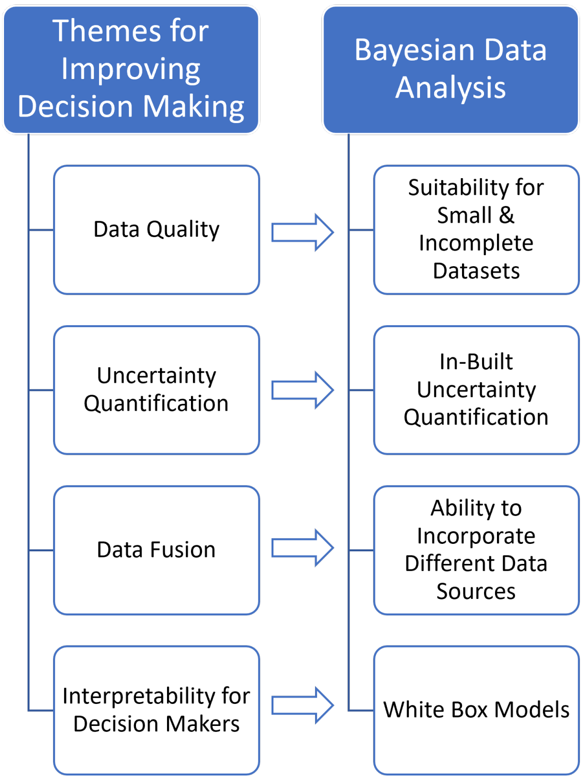


Figure 2.6: Summary of the themes for improving decision making identified in chapter 2 (left hand side), and how they map well to the advantages provided by Bayesian models (right-hand side).

3. There is an established relationship with the wind farm operator, which presented an opportunity for real-world strategy to inform research and research to inform real-world strategy.

With the scope narrowed to analysis of *Operations*, the next step was to scrutinise more closely which area or areas of the current research space could be improved, bearing in mind the priorities of the industrial partner. Referring to section section 2.4 of Chapter chapter 2, a retrospective analysis of operational data has the potential to improve either or all of the strategic, tactical and operational echelons. More precise aims are described in the following subsections. Each subsection describes a work package, and is split into 3 parts: **Motivation**, **Literature** and **Novelty**. **Motivation** provides a description of what influenced the analysis; **Literature** gives an overview of the research space for each point; **Novelty** gives an overview of the novelty provided by this thesis on each point.

2.6.1 Maintenance KPI Monitoring

This subsection elaborates on objective 2 outlined in chapter 1, namely:

"*Use the available dataset to calculate relevant KPIs describing wind turbine maintenance intervention. Use those KPIs to explore the maintenance requirements of offshore wind turbines.*"

Chapter 4 addresses this point firstly by defining and calculating relevant KPIs. KPIs that could be categorised as *intervention* descriptors are explored in more detail.

There are 3 key motivators for calculating maintenance KPIs, namely:

1. The sparcity of publicly available reliability and maintainability data from the wind industry. See Section section 2.4 and Table table 2.1.

2. There are large uncertainties surrounding the data processing of reliability and maintainability data. The effect of this uncertainty on failure rate calculations has not been explored in great detail in the literature.

3. A subset of turbines in the OWF which the dataset describes have tidal access restrictions, meaning that they are located in shallow water depths and are sometimes inaccessible by CTV. This presents an opportunity to assess how reduced accessibility is related to turbine performance.

Feng et al. [209] were the first to publish data relating to the reliability of offshore wind turbines. They present lists of turbine failure modes and operational issues experienced by round 1 offshore wind farms in the UK, namely: Barrow, North Hoyle, Scroby Sands and Kentish Flats. [210] built on this early work in their review of UK experiences and offshore operational challenges in the offshore wind industry. Their study extended the analysis of [209] to round 2 offshore wind farms, and presented annual failure rate and downtime figures for turbine subsystems. To this day, Carroll et al. [82] provide the most comprehensive reliability analysis of offshore wind turbines. They present an estimation of failure rates, repair times and repair costs of unscheduled maintenance activities among a population of approximately 350 turbines rated at 2-4MW. Estimates are summarised both by subcomponent and failure severity, using the categories *minor failure*, *major failure* and *major replacement*. Besides these academic publications, there are two industrial initiatives that monitor reliability statistics for offshore turbines, both of which commenced in 2013. The first is the UK's SPARTA (System Performance, Availability and Reliability Trend Analysis) initiative, which provides anonymised operator benchmarking of performance and reliability KPIs [211, 212, 213, 214]. However, any reliability figures published in their annual portfolio review are quite high level. In one of their reports, they present a downward trend of 'Forced outages' per turbine, reaching around 2 per month by 2020. An older report shows 'number of repairs' per turbine per month, however it is unclear how this is defined. More recent reports show number of replacements for major components [215]. WInD-Pool (Wind-Energy-Information-Data-Pool) [216], the successor to WMEP (Wissenschaftliches Mess- und Evaluierungsprogramm) [76] is the German equivalent. See Pfaffel et al. [126] and Cevasco et al. [41] for in-depth reviews about RAM and performance data of wind turbines.

Novelty comes from the following considerations:

1. This thesis presenting up-to-date, real-world reliability and maintenance data, which in itself is a valuable and novel contribution.

2. To my best knowledge, no studies have investigated the sensitivity of failure rate figures to different failure definitions/data pre-processing methods. Given the lack of standards for failure collection and definition and the inconsistency of the research space, this is valuable.

3. To my best knowledge, there are no studies in the literature which investigate tidal access restrictions.

2.6.2 Night Shifts

This subsection elaborates on objective 3 outlined in chapter 1, namely:

"*Explore the effectiveness of night shifts in increasing power production and availability based on up-to-date, real-world operational data.*"

Chapter chapter 6 addresses this point by employing a Bayesian hierarchical model (developed in chapter chapter 5) to the question of night shift strategy. The motivation, literature and novelty behind this chapter are outlined below.

Night shifts are an aspect of maintenance intervention which might have a significant impact on accessibility and in turn on performance. The question of whether to employ them exemplifies well the trade-off constantly undertaken at offshore wind farms to increase profit while maintaining H&S standards. It presents an opportunity to reduce turbine downtime and improve operational efficiency, however it comes with the potential for increased risk for technicians.

There are two primary benefits of adopting a night shift for corrective maintenance. The first is to avoid lost production due to downtime from turbine faults. This benefit is presumably most prevalent in months characterised by high wind speeds, where access conditions are difficult and the opportunity cost associated with unplanned downtime is high [217]. There may also, however, be an indirect cost associated with spreading resources and avoiding redundancy, as noted by Dalgic et al. [169] - a benefit which depends largely on the planning and foresight of the operator.

The question of night shifts arose in this thesis because the wind farm employed night shifts over the period of data collection. There is therefore an opportunity to explore the theme via a retro-active analysis of operational data. 24/7 working is explored as it was utilised at the site: by reporposing CTVs for corrective maintenance activities for a night shift. This strategy was used throughout the months of November, December, January and February in an attempt to improve accessibility in times where turbine failures are potentially most frequent and most costly.

The above point establishes an opportunity to address the question of night shifts by data mining. In order to take advantage of the features described in section section 2.5, this thesis proposes Bayesian modelling as a solution. Regarding which of the family of Bayesian methods to propose for this analysis, refer to Li and Shi's review of applications of Bayesian methods in wind energy conversion systems. [218]. In their "*discussion and prospects*" section, they recommend a hierarchical Bayesian approach as an advance on the methods typically employed in the field of wind energy. One of their recommendations is that the method by applied to cost-effectiveness analysis. Their recommendation provides a motivation for the proposed methodology. There are a few studies which address 24/7 working in the literature. Night shifts have been analysed using the StrathOW-OM cost-modelling tool [169], Besnard et al.'s maintenance support organisation model [168] and via a business case by Poulsen et al. [219]. All studies presented advantages to employing 24/7 working. Dalgic et al.[169] simulated different configurations of CTV utilisation, presenting their results in a £/MWh format. They concluded that "*10 out of the most costly 17 configurations do not have CTV for Night Shift*". Besnard et al. [168] concluded that "*the availability increases by almost 1% for each logistic solution by using 24/7 work shifts instead of 12/7 work shifts*", and report a financial benefit amounting to 5-15% of O&M costs. Poulsen et al [219] concluded that by implementing 24/7 working savings of approximately 1,800,000€per year can be realized via a series of expert interviews and dedicated focus groups.

This thesis addresses the same question, but provides novelty by employing a data-modelling approach in the form of a Bayesian hierarchical model. This means that, in contrast to previous studies examining night shifts for OWFs, the methodology employed in this thesis provides real-world evidence of night shift effectiveness. To my best knowledge, it is also the first time that Bayesian hierarchical modelling has been applied to cost-effectiveness analysis of a maintenance strategy based on retro-active analysis of operational data.

2.6.3 Exploring the Effect of Scheduled Maintenance on Unscheduled Maintenance

This subsection elaborates on objective 4 outlined in chapter 1, namely:

"*Explore the effect that annual services have on proceeding corrective works and in turn the reliability of wind turbines based on up-to-date, real-world operational data.*"

Chapter 6 addresses this point by employing a Bayesian reliability analysis. The motivation, literature and novelty behind this chapter are outlined below. Scheduled (preventative) maintenance of wind turbines is routine [220]. Its purpose is to decrease the possibility of turbine failures in the future. This is typically done by completing an annual service campaign, in which a list of tasks are completed at each turbine. Given their importance to wind turbine reliability [99], the industrial partner expressed an interest in quantifying the effect of annual services on proceeding corrective works. Part of their input in this respect was that annual services seemed to increase the chance of failures in the short term. This is a feature of preventative maintenance which has not been covered in the previous literature. The industrial partner therefore showed interest in the effect of annual service campaigns on turbine reliability both in the short-term and long-term.

Wind turbine reliability is commonly captured in the wind industry via some form of reliability modelling. Reliability modelling of wind turbines is an established method [221]. It might prove useful in fulfilling the second objective in this thesis if some covariate effect describing annual services' effect on WT reliability can be incorporated. The literature surrounding reliability modelling of WTs is therefore presented below, and where this thesis can provide novelty on the point.

2.6.3.1 Poisson & Weibull Processes

A central part to many operational decision making tools is the representation of WT failure, more formally known as failure modelling [33]. It is important to capture the frequency and probability of turbine failures accurately, as strategy for responding to them will dictate maintenance costs [82]. Many studies have in common their reliance on historic failure data to do so. However, there is plenty of license for different interpretations of that data. In simple cases, failures are modelled deterministically [33]. Turbines are assumed to fail after a constant period of time, taken as the Mean Time Between Failures (MTBF). While this is broadly indicative of failure behaviour, a recent study by Scheu et al. [188] has shown the importance of considering the uncertainty of failure rate around a mean value. It is therefore advisable to adopt a probabilistic method, where failures are assumed to occur at random intervals as described by some PDF, which can be fit to the data. Most popular of these is the Poisson process, of which there are two variants: the homogeneous and non-homogeneous Poisson process (HPP & NHPP respectively). The HPP assumes that time between failures is Poisson distributed according to a constant mean failure rate throughout time. The NHPP is an extension of this regime, where the mean failure rate is assumed to be time-dependent. The usefulness of this assumption is demonstrated by Slimacek and Lindqvist [137], who use it to model seasonal effects, and numerous studies which use a power law process to model wear-in, wear-out and serial defect effects [222].

Another commonly used method of modelling wind turbines is modelling time-to-failure via a Weibull distribution. Like the NHPP, Weibull time-to-failure modelling allows for changing failure intensity through time. In the case where covariates are not included in the model (or where the effect of covariates is constant), the Weibull formulation leads to a failure intensity that is monotonically increasing or decreasing through time.

Such methods might be categorised under the broader heading of reliability analysis in the field of engineering, and survival analysis in the fields of biology and medicine. There is a depth of theory in survival analysis in the data science research community which has seen broad application in these fields [223, 224, 225]. However, extensions of the NHPP which permeate survival analysis literature are yet to be fully utilised in the reliability analysis of wind turbines. To the authors' best knowledge, there are two studies which have made use of such extensions thus far. The first is by Slimacek and Lindqvist [137], who extended a NHPP by including a frailty model to capture heterogeneity unexplained by model covariates. They also employed the model covariates to explore the effect of 4 factors on turbine reliability, namely: harshness of local environment, turbine concept, date of installation and seasonality. The second is by Ozturk et al. [226], who also explored the impact model covariates affecting turbine reliability. This list was slightly more comprehensive, encompassing climatic regions, elevation location, distance to coast, mean annual wind speed, turbine age, turbine type, number of previous failures and scheduled maintenance history. However, for the most part they used simpler non-parametric methods, which differ from the semi-parametric methods which are closely related to NHPPs.

2.6.3.2 Annual Servicing As a Model Covariate

As summarised as part of section 2.4.2, annual services have previously been included in operational/simulation tools. Those tools tend to optimise the time between preventative maintenance actions [144, 145, 146]. None of these studies have used work orders to build reliability models which include preventative maintenance as a covariate. Annual services might be included in reliability models as *time-dependent* covariates. In this way, the time dependence of their effect on wind turbine failure-rate might be captured.

Time-dependent covariates are (to my best knowledge) a novel consideration for wind turbine reliability analyses. However, they are a statistical tool which is sometimes employed by survival analysis in fields other than WT reliability analysis [176, 177]. The methodology presented in chapter chapter 5 employs these methods in the field of wT reliability analysis for (to my knowledge) the first time.

The novelty in the methodology comes from extending the typical Poisson Process-type approach by two means. First, a transition of traditional reliability analysis methods into the Bayesian regime is proposed. This has the advantages characteristic of all Bayesian methodologies, as outlined in section section 2.5. Second, a data-modelling approach which has been employed by numerous recurrent time-to-event data analyses is proposed, which is capable of capturing the effect of non-corrective works on the likelihood of proceeding failures. This was inspired by anecdotal evidence of increased turbine failures after annual services from the operator of a large offshore wind farm. In Chapter chapter 6 this relationship is established and quantified, which is a novel consideration for wind farm O&M modellers. Time since annual service is therefore considered as an impacting covariate which has not been explored by previous studies. The effects of seasonality, year of operation and position in the array are also considered as covariates in the model.

2.7 Chapter Conclusion

This chapter had 2 objectives. The first was to fulfill objective 1 defined in chapter chapter 1, namely:

> "*Perform a review of the offshore wind data ecosystem to identify areas for improvement for data processes.*"

The review culminated in four areas for improvement: data quality, data fusion, uncertainty quantification and applicability to decision makers.

Figure 2.7 summarises the scope and points of novelty for the analyses covered in this section. Each of these analyses from left to right in figure 2.7 are presented in Chapters chapter 4, chapter 6 respectively.

The objectives were arrived at by a combination of literature review, expert input from the industrial partner. An early decision was made to focus on reliability and maintenance data. From this decision, three analyses were proposed. The first proposes calculating maintenance KPIs to measure maintenance intervention and effectiveness at the site. The second explores the effectiveness of night shifts in increasing power production and availability. The third explores the effect of annual services on proceeding corrective works.

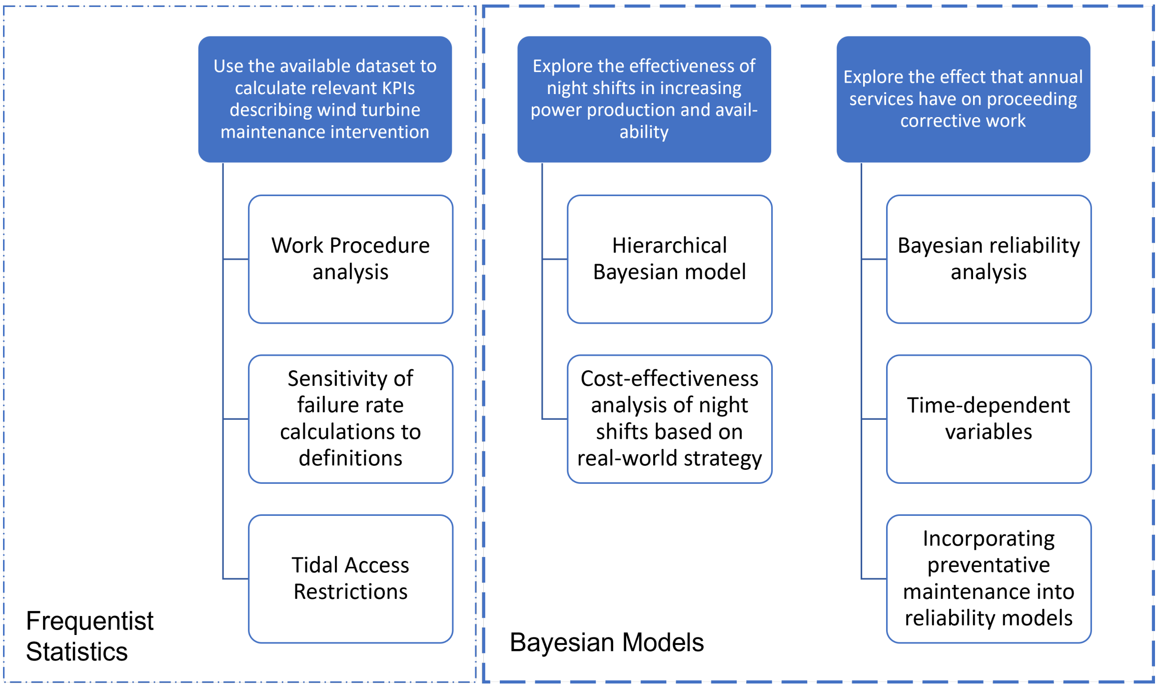


Figure 2.7: Objectives (black boxes with white writing) and methods/novel contributions for each objective (white boxes with black writing).

The chapter also made an argument for Bayesian models as a solution to the four points for improvement highlighted in Chapter chapter 2. Based on these arguments, a Bayesian framework for data analysis is developed and presented in section chapter 5. The advantages that this thesis are suitable to addressing those issues are: a suitability for small and incomplete datasets; in-built uncertainty quantification; the ability to incorporate different sources of data and their white-box nature. This argument led to Bayesian methodologies being adopted for the latter two analyses listed above. Bayesian hierarchical models were suggested by Li and Shi [184] as a tool for cost-effectiveness analysis. A Bayesian hierarchical model was therefore adopted for the night shifts analysis - this is designed and presented in section section 5.2.1.. For the question of annual services, some methodology that could robustly incorporate time-dependence into the analysis is preferable. For this reason, a Bayesian reliability analysis with time-dependent covariates is opted for. This is designed and presented in section section 5.3. This borrows from statistical models employed by survival analyses, but that (to my best knowledge) have not been employed by wind turbine reliability analyses.

3 Data Mining Methodology

3.1 Chapter Overview

This chapter presents a data mining methodology which is employed in the case study presented in chapter chapter 4. The data mining methodology achieves data fusion by combining the *Operations*, *SCADA*, *Weather* and *Turbine Properties* data tables described in section section 3.3.

The first part, contained within section section 3.2, defines KPIs to be used in the following analyses. Section section 3.3 gives an overview of the available dataset. Section section 3.4 describes the data mining methodology used to derive those KPIs from the available dataset. The data mining methodology outlined in section section 3.4 is developed to address the objective described in section section 2.6.1.

3.2 Key Performance Indicator Definitions

As described by Gonzalez et al. KPIs are tools for measuring the progress of a business towards its goals. They go on to provide a review of the KPIs used in the wind industry, which might in turn be used for the work in this thesis. They categorise KPIs as either *Performance*, *Reliability*, *Maintenance*, *Finance* and *Safety*. Since the database described in section section 3.3 does not contain any information regarding safety or finance, this thesis focuses on the former three categories. These are elaborated upon below.

3.2.1 Performance Indicators

With reference to an offshore wind farm, performance typically equates to efficiency [53]. Most pertinent to the operators is the question: *is the wind farm producing as much energy as it could?* [53] Variants on time- and energy-based availability are common, as well as capacity factor. [126, 227].

The simplest of these to define and calculate is the capacity factor (CF), which is the ratio of the turbine's actual average power output ($\bar{P}$) to its rated power (P_{Rated}):

$$CF = \frac{\bar{P}}{P_{Rated}}. \tag{3.1}$$

Gonzalez et al. [53] consider the CF to have value during feasibility and project development stages. However, since the denominator is a constant and does not adjust for on-site wind conditions, they conclude that it is insufficient for evaluating operational efficiency.

One of the most commonly employed metric is Time-Based Availability [69], denoted A_t. This gives a measure of how much time a turbine is available to produce power compared to the total time in operation [227]:

$$A_t = \frac{\textit{Time available}}{\textit{Total time in consideration}} \tag{3.2}$$

Availability is used for many purposes, representing the different interests of the various stakeholders involved in the operation of an OWF [227]. Namely: energy estimates, revenue projections, turbine design performance evaluation, warranties, and performance bonuses or penalties [227]. This quite naturally leads to different definitions of *time available* and *total time in consideration.* IEC 61400-25-1 [228] define two categories: *Operational* and *Technical Availability*, such that Operational Availability (A_O):

- Considers *Time available* to contain:
  - Hours of full and partial generation (including low wind);

- and *Total time in operation* may contain hours of:
  - Technial standby;
  - Requested Shutdown;
  - Out of environmental specification;
  - Out of electrical specification;
  - Scheduled maintenance;
  - Planned Corrective Action:
  - Forced outage;
  - Suspended;
  - Force Majeure.

And Technical Availability (A_{tech}):

- Considers *Time available* to contain:
  - Hours of full and partial generation (including low wind);
  - Technical Standby;
  - Out of environmental specification;
  - Requested Shutdown;
  - Out of electrical specification;
- *Total time in operation* contains hours of:
  - Planned Corrective Action:
  - Forced outage,
- and time not included in the calculation contains:
  - Scheduled maintenance;
  - Suspended;

- Force majeure;
- Information Unavailable;

IEC 61400-25-1 [228] alternatively refers to A_O as the WT user's view and A_{tech} as the WT manufacturer's view, which is perhaps more indicative of the utility of the two measures. A_{tech} is the most meaningful in considering the reliability of turbine in terms of the frequency and consequence of unexpected failures [41]. However, A_O might be a more practical measure for the operator.

Alternatively, energetic or production-based availability (A_E) can be used, defined as the ratio between the actual energy produced (E_a) and the potential energy available (E_p) [41] :

$$A_E = \frac{E_a}{E_p} \tag{3.3}$$

As highlighted by Gonzalez et al. [53], this measure has the advantage of capturing the percentage of energy captured from that available, which they consider to best represent most accurately the efficiency of a WT. IEC standard 61400-26-2 [229] provides similar sub-categories of A_E to those defined for A_t, which is designate $A_{E,O}$ (production-based operational availability) and $A_{E,tech}$ (production-based technical availability) respectively.

The utility of A_E comes with a caveat, however. Namely, there is no standard procedure for defining and calculating the actual available energy for the given period [53]. IEC standard 61400-26-2 provides various means of doing so, which are categorised into two groups: those methods employing wind speed measurements together with power curves, and those employing some direct estimation of power by comparison to other turbines in the farm. Section section 3.4.2 describes how this thesis calculates opportunity cost for failed turbines.

There are various initiatives which have published performance statistics for offshore wind farms. Pfaffel et al. [126] provide a concise review of these studies. In their conclusions, they state that: *"Results on the performance of WT can be considered to be reliable. There is no reason to doubt the comparison of capacity factors which* *shows a high location dependency. Offshore WT reach the highest capacity factors closely followed by onshore WT in the US. When it comes to availability, the situation is a bit more ambiguous. Publications lack detailed definitions of the availability calculation which weakens the significance of the comparison. Still, it should be easy to choose the right source for specific applications and basic statements can also be made. For example, the time of very low availability in offshore wind energy seems to be overcome".*

3.2.2 Reliability & Maintenance Indicators

Wind turbine reliability is concerned with two questions: (i) *how often does a turbine fail?* And (ii) *what are the impacts of those failures in terms of downtime and resources?* [53] Reliability indicators are those which address these questions. Maintenance indicators are concerned with quantifying an answer to question (ii).

3.2.2.1 Number of Interventions

Number of interventions is described by Conzalez et al. [53]. An *intervention* is classified here as at least one displacement of a maintenance crew from a vessel onto a turbine within a period on turbine downtime. Note that multiple work orders carried out within that period of downtime still count as one intervention. This is presented in a number of visits per year format. Less visits should be needed in the case of higher reliability and/or in the case of a well optimised maintenance strategy.

3.2.2.2 Failure Definition

What researchers mean when they say that a turbine *fails* varies, as there is no standardised definition of a failure within the wind industry [82]. The definition used largely depends on the quality and quantity of data available to the analyst, which varies widely from one study to the next [78]. Here, use the following definition is used:

"*A turbine downtime event accompanied by an unscheduled visit to that turbine*"

This definition is itself open to interpretation, however, depending on the data selection criteria used. This problem is discussed and investigated more in section section 3.4.3.

3.2.2.3 Failure Rate

WT failure rate is commonly reported in a failure per turbine per year format, according to following equation, as introduced by [82]:

$$\lambda = \frac{\sum_{i=1}^{I} \sum_{k=1}^{K} \frac{n_{i,k}}{N_i}}{\sum_{i=1}^{I} \frac{T_i}{8760}} \tag{3.4}$$

where I is the number of intervals for which data are collected, K is the number of assemblies, $n_{i,k}$ is the number of failures for sub-assembly k and interval i, N_i is the total number of turbines and T_i is the total time period in hours. The numeber 8760 is the number of hours in a year.

3.2.2.4 Failure Categorisation

WT failure modelling often considers four kinds of failure:

1. Manual Restarts (*man*). Manual restarts refer to instances where the turbine has tripped, and requires a technician to manually restart it on-site. However, no assembly-repairs are required to return it to an operational state.

2. Minor failures (*mr*). Simple assembly failures which do not require any replacements of major components and can typically be carried out in one day.

3. Major failures (*Mr*). Failures which are more labour and cost intensive to resolve, but which still do not require any major components to be replaced.

4. Major replacements (*MR*). Labour and cost intensive replacements of major components such as a gearbox or generator.

Depending on the analysis, manual restarts (*man*) may be included as a separate category, or included as minor failures. Carroll et al. [82] categorise failures by cost of consumable materials, such that: *mr* repairs are those corresponding to those with a repair material cost of less than €1,000; *Mr* failures have repair material cost between €1,000 and €10,000 and *MR* failures over €10,000. This categorisation yields 6.2 minor repairs, 1.1 major repairs and 0.3 major replacements per year, with 0.7 failures per turbine per year have no cost data so could not be categorized.

Alternatively, Scheu et al. [10] define various criticality categories via a number of other measures within their failure modes & criticality analysis of offshore wind turbines. These are summarised in table 3.1. Such measures are useful in the case that the database does not contain cost information - as is the case here.

Table 3.1: Failure category criteria as defined by [10]

	mr	***Mr***	***MR***
Downtime	$\leq$ 7 days	7 days $<$ downtime $\leq$ 15 days	$>$ 15 days
Spare Part Cost	0–7.5 k€	0–7.5 k€	$>$ 100 k€
Type of Intervention	1 CTV, SOV or helicopter mobilization and use for up to 1 day, 3 or less technicians	1 CTV, SOV or helicopter mobilization and use for up to 7 days, 6 or less technicians	1 CTV, SOV or helicopter mobilization and use more than 7 days, 1 jack-up vessel mobilization and use for min 1 day.

Failure severity is defined by the following logic:

1. Each maintenance intervention is categorised by the work procedure attached to it.

2. For each work procedure (e.g. removal and replacement of converter modules), the average number of technicians, number of visits and downtime is calculated.

3. That work procedure is assumed to be a *mr* repair unless any of the conditions defined in table 3.1 are met by the average number of technicians, number of visits or downtime for that work procedure. If they are met, that work procedure is categorised as either *Mr* or *MR*.

3.2.2.5 (Mean) Time Between Failures

Time between failures is the elapsed time between the resolution of one failure and the onset of the next. In WT reliability analyses, the time between failures is often assumed to follow an exponential distribution:

$$f(t) = \lambda \mathrm{e}^{-\lambda t}, \tag{3.5}$$

where $f(t)$ represents the density function of time between failure, t time, and λ the failure rate. The Mean Time Between Failures (MTBF) is the average time elapsed between one failure's resolution and another's onset. It is described by [54]:

$$MTBF = \int_0^\infty t f(t) dt. \tag{3.6}$$

In the case that time between failures is exponentially distributed, MTBF becomes the corollary of failure rate:

$$MTBF = \frac{1}{\lambda}. \tag{3.7}$$

However, this assumption does not always hold true, and in failure modelling of wind turbines other distributional assumptions may be more accurate. The effect of using different distributional assumptions for modelling time-between-failures is investigated in section section 6.3.

3.2.2.6 Mean Down Time

Mean Down Time (MDT) is the average time to return a wind turbine to its functional state. It can refer to both corrective (in which case it is similar to Mean Time to Repair (MTTR)) and non-corrective works. It is calculated here by averaging the downtime from downtime events associated with certain types of maintenance task.

While Pfaffel et al. [126] identify 7 studies that provide information on MDT, they are all for onshore turbines. Here MDT per failure severity category is calculated as discussed in section section 3.2.2.4, as well as distributions of downtimes for the respective categories.

3.2.2.7 Active Repair Time

The *repair time* is similar to the mean time to repair defined by Gonzalez et al. [53]. It refers to the total duration taken to complete the entire repair process, including both active repair activities and any associated idle time or waiting time.

In contrast, *active repair time* specifically refers to the period when the actual repair work is being performed on a turbine assembly. Active repair time is defined here as the sum of differences in pick-up/drop-off times for technicians to/from a given turbine during a given downtime event:

$$t_{activerepair} = \sum_{i}^{I}(t^{i}_{pick-up} - t^{i}_{drop-off}) \tag{3.8}$$

where I is the total number of work shifts throughout the period of downtime in question, $t^{i}_{drop-off}$ is the time a technician team is dropped-off to the turbine in the given work shift and $t^{i}_{pick-up}$ is the time a technician team is picked-up from the turbine in the given work shift.

3.3 Data Audit

The analysis of presented in this thesis is based on operational data provided by a large offshore wind farm. This is the role of the industrial partner in the project. Due to data confidentiality, the industrial partner cannot be named. The WF consists of a fleet of modern, geared wind turbines with a multi-MW power rating. The database available for analysis consists of four and a half years of SCADA data, maintenance activity logs and weather data, spanning the time period from July 2018 to January 2023. It is assumed that this is beyond the 'early failures' period on the bathtub curve commonly considered to represent WT failure intensity through time [230]. Of the factors influencing the planning and cost of maintenance of offshore wind farms as listed in [14], the metadata presented can be used to represent: power production, inspections, repairs, failures of turbines, repair time, wave height, wind speed, weather windows, travel time, environmental conditions (dependent on time and season), types of maintenance and distance from shore.

Here, 4 categories of data are designated which will be used throughout the thesis: *Operations, SCADA, Weather, Turbine Properties* Data. The metadata of each of these categories is presented in the following subsections.

3.3.1 Operational Data

Operational data applies to data that provides insight into maintenance actions. It details technician movements on vessels to-and-from the farm, and what kind of maintenance tasks they performed when they reached the turbine. The data mining process described in this thesis is primarily concerned with four primary operational data tables: *Work Procedures*. *Tasks/Task Types*, *Operations Planned Movements*, *Vessel Stops* and *Operations Shift Tasks*. Table 3.2 summarises the information derived from each of these tables, as well as their inherent flaws.

It should be noted that *Work Procedures* was only incorporated into the analyses presented in the following chapters at a late stage of the research project. Chapters chapter 5 and chapter 6 therefore only present utility of those analyses at the turbine level, and they could be improved by incorporating work procedures.

There are two noteworthy pre-processing steps for maintenance data. Firstly, the 79 maintenance task types listed in the table *Task Types* were further categorised into 9 categories:

1. Annual. Annual describes periodic annual servicing works. It is common practice in the industry to have an annual campaign where a set of scheduled maintenance actions are carried out on turbines. The exact nature of these actions vary from one service provider to the next and depend largely on maintenance contract arrangements [161]. This generally includes tasks such as lubrication of mechanical parts (e.g. gear oil, hydraulic oil, greasing), measurement of part temperatures, a torque tensioning of bolts and basic inspection of parts within the nacelle.

2. Balance of Plant (BoP). Balance of plant contains the task types labelled *Corrective - BoP* and *Defects & Tasks*.

Table 3.2: Summary of the operational data-tables available for analysis.

Data Table	**Information Derived**	**Disadvantages**
A. Work Procedures	• Text description of work carried out	• Sometimes difficult to map to assembly taxonomy
B. Tasks/Task Types	• Task descriptions • Task categories	• Most corrective task descriptions only contain an alarm code • Not always indicative of actual assembly failure
C. Operations Planned Movements	• Manual acknowledgement/card swipe times for technician transfer of control on/off turbine • Estimated times for technician transfers from work plans • Repair times	• Incomplete: some pick-ups have to drop-off and vise-versa • Some transfers are 'planned', but not acknowledged.
D. Vessel Stops	• Stop types (pick-up/drop off) • Stop order • task ID label	• Timestamps often inaccurate.
E. Operations Shift Tasks	• Links other data tables • Info. on shift times	• No detail on work carried out during shift

3. Corrective. Corrective maintenance actions entail a repair. It implies the turbine (or at least part of the turbine) is not in a functional state.

4. Inspection. Likewise, it is common to regularly inspect certain key components - for example a visual inspection of blades for cracks or erosion. The majority of inspection works come under the heading of 'statutory inspections', as summarised in section section 4.4,

5. Retrofitting. Retrofitting activities refer to activities where obsolete technical components with are replaced new ones to prevent the chance of malfunction. A prominent example (as evidenced by the analysis in section section 4.4) is converter software updates.

Second, planned movements were separated according to whether they had been acknowledged by technicians, and the resulting list of time-stamps were used as the basis of a condensed 'Operations' data table (see section section 3.4).

3.3.2 SCADA Data

SCADA data was available in two forms: 10-minute SCADA data (typical of most SCADA datasets) and SCADA sampled at an increased resolution of approximately 2Hz. Ostensibly the high resolution SCADA data is an advantageous option for many analyses because it provides more detail. However, it requires significantly more processing power. 5 months of high resolution data, for instance, takes up approximately 3.7Gb of memory - compared to approximately 179Mb for the low-resolution data. The choice of format therefore depends on the use case. If condition monitoring is the focus of the research, higher resolution data would present an advantage [15]. For instance Gonzalez et al. [174] use 0.25Hz SCADA data for condition monitoring, and note benefits over 10-minute data. If, however, the focus of the research is on analysis of operational data, it may be preferable to incorporate the low-computational-cost/low-resolution option presented by the 10-minute data. For reasons outlined in section section 2.6 the focus of the study was on analysing operational data. Where *SCADA* data is referred to it therefore refers to 10-minute SCADA data. Table 3.3 lists the column names of the 10-minute SCADA data tables.

Table 3.3: Variables included in the 10-minute SCADA data.

SCADA Variable
Turbine
Mean wind speed
Mean wind speed bin
Mean active power
Mean nacelle direction
Mean wind direction
Mean generator rpm
Mean pitch angle
Mean tower vibration
Rotor rpm
Time ready to operate
Time in operation

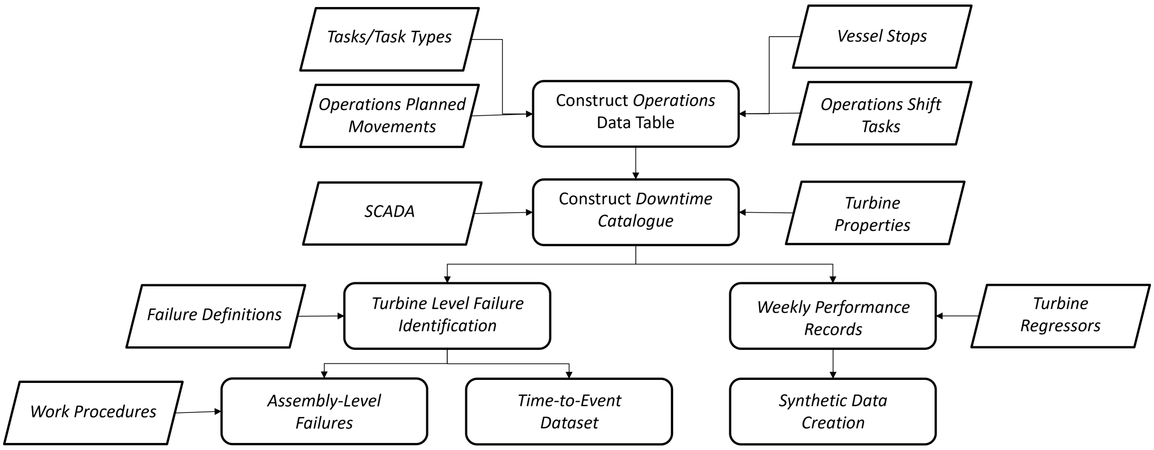


Figure 3.1: Summary of the data mining process employed in this thesis.

3.3.3 Weather Data

Weather data consists of time-series of significant wave height collected by two wave buoys at the site and wind speed/direction measurements collected from an on-site met-mast. Both time-series have a resolution of 1 hour. It is assumed that the mean of the 2 significant wave height signals was representative of the site.

3.3.4 Turbine Properties

Turbine properties refers to data-tables which contained turbine-specific information. Importantly for the analysis carried out in this thesis, it contained each turbine's location and water depth recordings. It also categorised turbines that were "*tidally-restricted*" or "*severely tidally-restricted*". Tidally restricted turbines are characterised by shallow water depths, meaning that they could only be accessed at high tide.

3.4 Data Mining Methodology for KPI Calculations

This section provides a methodology to derive the KPIs defined in section section 3.2 from the data-tables summarised in section section 3.3 (see tables 3.2 and 3.3 for the metadata). The overall procedure is summarised in figure 3.1.

3.4.1 Downtime Catalogue Construction

The data tables outlined in section section 3.3 provide the basis of the analyses presented in this thesis. To reiterate:

1. *Operations* data describes the maintenance works undertaken at the sites. The key data tables, along with the information provided by them, are summarised in table 3.2.

2. SCADA is typical 10-minute SCADA data with the variables described in table 3.3.

3. *Weather* data provides hourly wind-speed and wave height data.

4. *Turbine Properties* provides turbine-specific location (i.e. location and tidal access restrictions).

The above data tables were combined into a more compact data-table named *Downtime Catalogue*. *Downtime Catalogue* contained a series of maintenance actions at each turbine spanning July 2018 to January 2022. It was constructed via the following steps:

1. *Operations Planned Movements* was cross-referenced with *Tasks/Task Types* and *Vessel Stops* to create a time-series of maintenance interventions. From this the following could be identified:

   (a) What kind of work was undertaken at the turbine and when;

   (b) The approximate repair time for the maintenance action;

   (c) How many technicians worked on the maintenance action.

   Once *Work Procedures* was incorporated into the analysis, more detailed information could be derived regarding the intervention.

2. Periods of turbine downtime were identified from the SCADA data. These periods of downtime were cross referenced with the result from point 1. This meant that downtime events could be mapped to specific maintenance actions.

3. Information from *Turbine Properties* was merged-in to the result of point 2.

From *Downtime Catalogue*, most of the KPIs defined in section section 3.2 could be calculated. However, additional steps were required to estimate lost production and failure rates. The process to calculate them are outlined in the following subsections.

3.4.2 Lost Production Calculation

Lost production is calculated via a linear regression to the turbine's nearest neighbours. The process is as follows:

1. A reference table was created to hold the regression parameters of each turbine's mean active power to the active power of its 4 nearest neighbours; and to the farm average. The regression parameters were calculated from periods of uninterrupted up-time for the two turbines in question. Note that this method is similar to the "*average production of representative comparison turbines*" as defined by [229]. However, it differs by two means:

   - In this thesis, regression parameters are used over the actual power production of similar turbines;
   - only one turbine is used to calculate the lost production, not a group of turbines.

   This has the advantage that it only needs one turbine operating at full capacity to make an estimate. However, the method could be more prone to outlier power values.

2. A turbine downtime is identified by the SCADA flag "time ready to operate". If the time ready to operate is less than 600, we know the turbine is down for at least some of the time that entry.

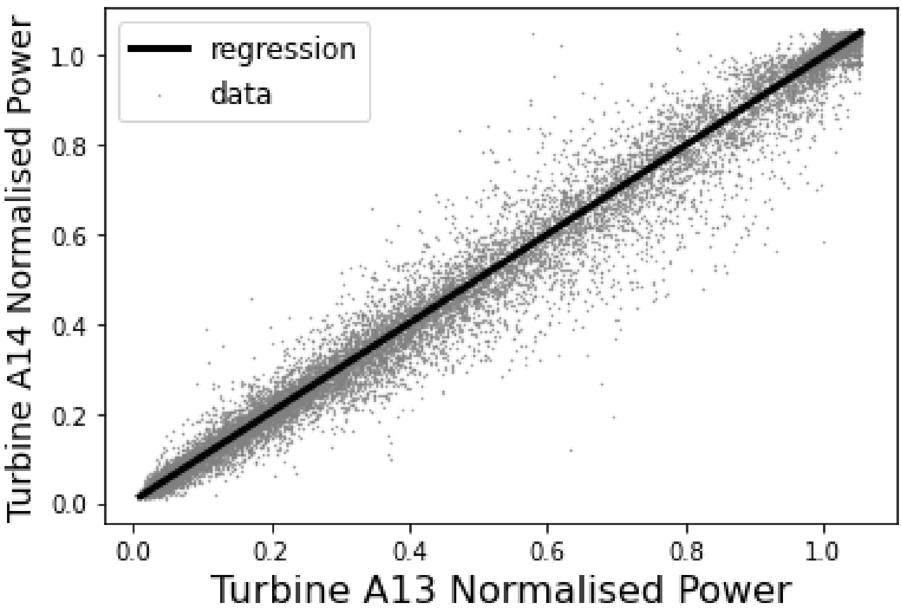


Figure 3.2: Example of 10-minute mean power relationship between neighbouring turbines.

3. For each timestamp in which the turbine is down, SCADA data is used alongside *regressors* to estimate the power according to the regression to the nearest turbine producing power.

An example regression between two neighbouring turbines is shown in figure 3.2. Note that, not every turbine-pair-regression will look the same. The very straight regression line shown in figure 3.2 is perhaps expected for turbines operating in the same row/in close proximity. However, local variations in climate will likely produce different regression relationships.

3.4.3 Identification of Turbine-Level Failures

In the creation of the *Downtime catalogue*, three things are accomplished:

1. Turbine downtime events are extracted from the SCADA data.

2. Those downtime events which are accompanied by an intervention from the maintenance team are extracted from SCADA and *Operations Planned Movements*.

3. Corrective maintenance interventions are extracted from *Task Types* together with SCADA and *Operations Planned Movements.*

These three points are the basis upon which this thesis can define a failure, namely:

"*a turbine downtime event accompanied by an unscheduled visit to that turbine*"

However, exactly how *downtime event*, *intervention* and *corrective maintenance action* are interpreted presents an uncertainty in the failure definition. The sensitivity of this interpretation is explored in chapter chapter 4.

3.4.4 Identification of Assembly-Level Failures

Assembly-level failures are identified by the following steps:

1. *Downtime Catalogue* is subject to the various data selection criteria outlined in section section 3.4.5.

2. If the subsystem or assembly is immediately evident from the text description in *Work Procedures*, that task is labelled accordingly. Example: "Main bearing removal and replacement"

3. If the subsystem or assembly is not evident from *Work Procedures*, the text description in *Tasks* is used. Tasks have two types of descriptor:

   (a) Tasks labelled with *service department* are closer to work procedure descriptors. These descriptors are sought out first.

   (b) The remaining task descriptors contain a single alarm code, corresponding to what the turbine manufacturer thinks is wrong with the turbine. These descriptors have a higher level uncertainty because they depend on accuracy of diagnosis at the beginning of the data mining process and accuracy of interpretation at the end.

4. Once tasks are labelled with a subsystem and assembly, they are mapped to the enhanced taxonomy defined by Reder et al. [68]. This is shown in figure 3.3 Note that the selection of taxonomy is also a sort of "metaparameter" which has a significant impact on results - this is discussed more in chapter chapter 4. This taxonomy was selected as a modernised version of the Reliawind taxonomy which has been popular in previous academic studies [82]. In their own words, Reder et al. [26] state the advantage of such an approach as ensuring "*the comparability of this study to older studies*".

Subsystem	Assembly	Subsystem	Assembly	Subsystem	Assembly
Power Module		**Control & Communications**		**Auxiliary System**	
	Frequency Converter		Sensors		Cooling system
	Generator		Controller		Electrical Protection and Safety
	Switch Gear		Communication System		Human Safety
	Soft Starter		Emergency Control & Communication Series		Hydraulic Group
	MV/LV Transformer		Data Aquisition System		WTG Meteorological Station
	Power Feeder Cables	**Nacelle**			Lightning Protection
	Power Cabinet		Yaw System		Firefighting System
	Power Module Other		Nacelle Cover		Cabinets
	Power Protection Unit		Nacelle Bed plate		Service Crane
Rotor & Blades		**Drive train**			Lift
	Pitch System		Gearbox		Grounding
	Other Blade Brake		Main Bearing		Beacon/Lights
	Rotor		Bearings		Power Supply Auxiliary Systems
	Blades		Mechanical Brake		Electrical Auxiliary Cabling
	Hub		High Speed Shaft	**Structure**	
	Blade Bearing		Silent Blocks		Tower
			Low Speed (Main) Shaft		Foundations

Figure 3.3: Summary of the taxonomy defined by Reder et al. [68]. Taken from [68].

3.4.5 Failure Definitions

This study defines a failure as *a turbine downtime event accompanied by an unscheduled visit to that turbine*. Often reliability analyses of wind turbines will use some variation of this definition. Carroll et al. [80], who provide the more extensive reliability analysis of offshore wind turbines in the literature, use the definition of '*a visit to a turbine, outside of scheduled operation, in which material is consumed*'. In the absence of any material usage data, this thesis uses turbine downtime as a qualifier for failure.

The variation in turbine-level failures initially focuses on the three sources of uncertainty highlighted in section section 3.4.3. To elaborate:

1. What is meant by a *downtime event*? This has two aspects to it.

    (a) Some studies (e.g. Wilkinson et al. [86]) impose a lower limit on downtime for that event to be considered a failure. This seems sensible for low values - if for instance a turbine was down for less than an hour it would indicate the turbine had not yet lost its capability to function and was visited when a maintenance team already at the farm for a minor repair. We will refer to this as a *downtime limit* henceforth.

    (b) Other studies argue that several sequential downtime events can be brought on by one failure, and that there should be some limit on the time elapsed between those events for them to be grouped under one failure. We will refer to this as a *grouping limit* henceforth.

2. What is meant by an *intervention*? There are three aspects

    (a) For some maintenance tasks, there is a *drop-off* of a technician and no corresponding *pick-up*, and vice-versa. The question is to only include tasks which have both (effectively imposing a lower limit on repair time) or to include tasks which have one or the other.

(b) As an extension of the above point, and analogously to a downtime limit, we might impose a lower limit on active repair time. We will refer to this henceforth as a *repair limit*. A low repair time implies a manual restart of a turbine, which reliability analyses often aim to filter out.

3. What is meant by a *corrective maintenance action*? This is obvious at first glance: include all maintenance actions labelled as corrective in the database in the database. However, it is conceivable that different studies will expand or contract the range of their included tasks subject to a number of considerations. These are listed below.

   (a) **Opportunistic tasks**. Downtimes attributed to non-corrective works, such as annual services, sometimes contain corrective works.

   (b) **Multiple Components**. Sometimes, downtimes that are considered *failure* could be attributed to multiple components. The failure could either by attributed to one component or multiple components.

   (c) **Retrofitting**. Some retrofitting works are labelled as such in the dataset, but others are labelled as corrective.

   (d) **Annual Services**. As above, some tasks are labelled as corrective but appear to be part of the annual service.

   (e) **Balance of Plant (BoP) tasks**. BoP jobs are either labelled as 'defects' or as corrective - bop. Similar to the above, a certain amount of BoP work may be covered by contractual arrangements, and may or may not be included in failure estimates.

   (f) **Unlabelled tasks**. Some tasks are labelled as corrective but contain no work-procedure descriptor. Others might have a vague work procedure associated with them and not a task descriptor, so similarly cannot be classified by assembly.

(g) **Fault Finding missions**. Some corrective tasks have the work procedure *Fault Finding* - it is unclear whether a repair was conducted directly from the fault finding mission. However, some fault finding tasks have an alarm code attached to them, and might be further categorised.

Here a baseline definition is defined by the conditions:

1. No downtime limit;
2. No grouping limit;
3. No opportunistic maintenance included (each downtime event is one failure);
4. All tasks that are recorded as corrective are included. This includes some tasks which might otherwise be recorded as retrofitting and annual service.

3.5 StrathOW-OM

StrathOW-OM is an O&M cost model designed for strategic planning purposes [231]. It is used to calculate key performance indicators which are used in the subsquent analyses of chapter 6. The model was developed at the University of Strathclyde by Dinwoodie [231]. It was subsequently validated against three other cost modelling tools [232]. Since then, it has been further developed and frequently utilised by other Strathclyde researchers [233, 78, 30].

StrathOW-O&M models a series of work shifts simulated in the time domain. Three simulated time-series feed into the central simulation: one describing significant wave height and wind speed; one the ideal power production for the farm; and another the probability of a subsystem failure in each time-step. These are all derived from user inputs describing the met-ocean climate, power curve and turbine reliability estimates respectively. The weather conditions are generated using a correlated, Multivariate Auto-Regressive approach (MAR). A Non-Homogeneous Poisson Process (NHPP) is used to model reliability through time. For each time step, the conditional reliability of a subsystem is compared to a randomly generated number to determine if that subsystem has failed. When a failure occurs, repairs are carried out dependant on availability of required resources (in terms of vessels, staff and materials) and climate restraints on vessel operational usage (significant wave height and wind speed limits). Again, the resource and operational limits are user-defined.

Once the shift is simulated, the model records the condition of the wind farm in terms of turbines available and resources utilised. The process is repeated for the specified lifetime of the farm, and the lifetime power production and availability are calculated and stored. This is repeated until there is convergence of availability estimates on cross-simulation values. Cross-simulation calculations are passed to model outputs for post-processing. Outputs consist of a list of Key Performance Indicators (availability, power production and number of failures), cost estimates (revenue, lost production costs, vessel and staff costs, costs of spare parts) and vessel specific information (CTV utilisation, number of JUV charters).

3.6 Chapter Summary

This chapter presents a data mining methodology for transforming the available dataset into useful information. The chapter starts with a list of performance and reliability/maintainability KPIs before turning attention to the data handling. The majority of these KPIs are well defined in the literature [234]. However, from the available literature [41, 126], it is not well defined and applications of different failure definitions create an inconsistent research space [25].

The data mining methodology consisted of procedure to calculate the KPIs from the available dataset. The first step in this procedure consists of a data audit. From the data audit, it is evident that parts of the *Operations* dataset are more detailed to that which is commonly available to researchers. Given the quality of the dataset and the comparative predominance of analysis into other data-streams in the literature [33], *Operations* is to be the focus of analysis in this thesis.

Elements of *Operations* are used as a base out of which a *Downtime Catalogue* is constructed. *Downtime Catalogue* cross-references *Operations* with SCADA data and turbine properties, such that every WT downtime event can be attributed to a set of maintenance actions. A methodology for lost production is developed where a linear regression to one of the turbine's nearest neighbours is used to estimate lost production.

Finally, different failure definitions are discussed. The definition of a failure in this thesis hinges on addressing 3 sources of uncertainty in the data-processing stage. Namely: *What is meant by a downtime event?*, *What is meant by an intervention?* and *What is meant by corrective maintenance action?*. These failure definitions facilitate the sensitivity analysis conducted in the next chapter, which explores the most immediate results which can be derived from the methodology developed in this chapter.

4 Work Order Data Mining

4.1 Chapter Overview

This chapter addresses objective 4 defined in section section 1.3, namely:

"*Use the available dataset to calculate relevant KPIs describing wind turbine maintenance intervention. Use those KPIs to explore the maintenance requirements of offshore wind turbines;*"

according to the rationale presented in section section 2.6.1 and the methodology developed in chapter chapter 3. The results presented in this chapter are an updated version of the results published in [235, 236].

The chapter starts using the frequentist statistics derived from the process developed in chapter chapter 3 are used to scrutinise the effect of tidal-access restrictions on wind turbine reliability in section section 4.2. The rest of the chapter is dedicated to reliability analysis and exploring the uncertainty surrounding WT reliability analysis due to data processing. Section section 4.3 explores the uncertainty in turbine-level failure rates due to different failure definitions. The analysis makes its way towards an assembly level reliability analysis by first presenting an analysis of the work procedure data (section section 4.4) and then mapping the work procedures to a failure taxonomy (section section 4.5). In mapping the work procedures to assemblies, different data selection procedures are explored.

4.2 Tidal Access Restrictions

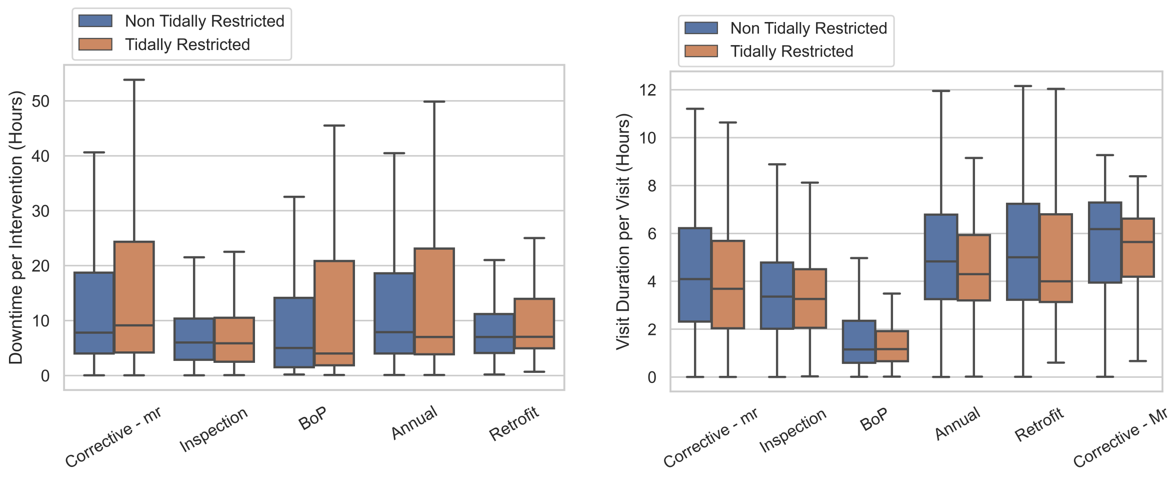
The most immediate objective that the data mining methodology of chapter chapter 3 can address is exploring the effect of tidal access restrictions.

Figure 4.1 (a) shows that tidal access restrictions lead to shorter visit durations when compared to their non-tidally-restricted counterparts for all maintenance categories. Figure 4.1 (b) shows that this leads to generally longer downtimes per intervention. This leads to 12% difference in the mean downtime form failures for tidally-restricted turbines. The disparity increases as the severity of the fault increases. This is evident from scrutinising the fault severity categories described in section section 3.2.2.4. For *mr* corrective works, there is an 8% difference between the mean downtimes of the two groups; For *Mr* corrective works the difference rises to 13%. For *MR* corrective works, the downtime of tidally restricted turbines is 12% less than non-tidally restricted. However, the number of samples is so small for major replacements that this result is not robust. Downtime per intervention is also generally higher for BoP jobs and for scheduled jobs save inspections. However, the mean values of each distribution of downtimes is markedly different from the median, implying that there are are a lot of extreme values skewing the distributions. The restrictions also entail more visits per repair, in an environment operators want to reduce the number of technician transfers and the consequent safety risk. Again, this effect is amplified with the severity of the fault.

Figure 4.2 shows how this reduced maintainability translates to a reduced technical availability. This analysis shows that tidal access restrictions have the potential to cause lower availabilities. This is true for the median value of technical availability, but also increases the chance of outliers at lower availabilities (as shown by the leg of the box plot in figure 4.2).

4.3 Turbine-Level Failure Rate Calculation

The rest of this chapter is dedicated to WT reliability analysis. WT reliability analysis, refers to the process of first calculating the failure rates based on the failure



(a) (b)

Figure 4.1: Comparison of tidally restricted and non-tidally restricted (a) visit duration and (b) downtime per intervention (as defined in section section 3.2.2.1) for different maintenance categories (as defined in section section 3.3.1). The *Mr* category is only shown in figure (b) as it would change the scale of figure (a) and obscure the other categories.

definition and data selection criteria summarised in sections section 3.4.4 and section 3.4.5. By doing so, failure rates are presented at the turbine level then the assembly level. This section presents the uncertainty in turbine-level failure rates, according to the data selection criteria defined in section section 3.4.5.

The baseline turbine-level failure rate estimate was calculated using the *Downtime Catalogue* and the baseline definition used in section section 3.4.5.

It comes in at 8.94 failures per turbine per year. Figure 4.3 explores the sensitivity of that figure to the various points outlined in section section 3.4.5. Figure 4.3 (a) shows the failure rate estimate falling sharply from the baseline with increasing downtime limit up until around 10 hours, after which the decline starts to slow. The limit of the x-axis corresponds to the most extreme downtime limit in the literature of 72 hours [69]. At the more reasonable limit of 1 hour, the failure rate estimate reaches 8.31 failures per turbine per year, a 4% reduction on the baseline. However, it is evident from figure 4.3 that dropping these failures does not incur a significant increase in availability. Figure 4.3 (c) shows the dropping failure rate estimate and

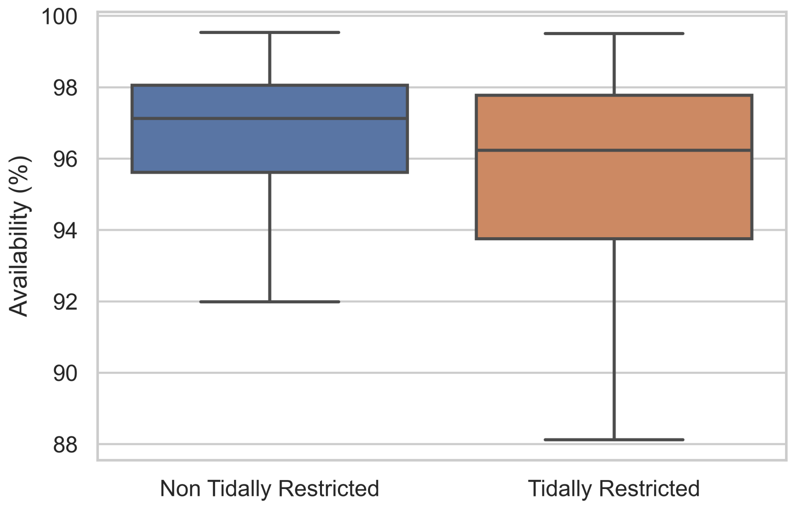


Figure 4.2: Comparison of time-based availability for tidally restricted and non-tidally restricted turbines.

rising technical availability up to a downtime limit of 5 hours. Only slight variations in technical availability are evident in this range.

Figure 4.3 (b) shows a similarly sharp fall-off from the baseline case with increasing grouping limit up until the 24 hour mark, after which the curve flattens. At 24 hours, the failure estimate reaches 6.54 failures per turbine per year. Note that the availability estimate remains constant independent of the grouping limit.

Repair time limit is explored in figure 1 (d). This shows an approximately linear relationship between failure rate estimates and repair time limits. At 1 hour, the baseline drops to 7.98; at two hours to 7.13. The effect of increasing repair time limit on technical availability is more significant, implying more care should taken in imposing even small repair time limits.

Figures 4.3 (a), (b), and (c) address the question: *what is meant by an intervention?*. Including zero repair time jobs (dashed lines) increases the baseline failure rate estimates to 10.97. This disparity decreases with increasing downtime limit.

Figure 4.4 explores the question: *what is meant by 'corrective maintenance'*. Essentially the baseline failure definition can be expanded or contracted based on additional data selection criteria. Of the data selection criteria explored, the failure rate ranges from 7.07 to 12.15. Inclusion of opportunistic jobs increases the failure rate to

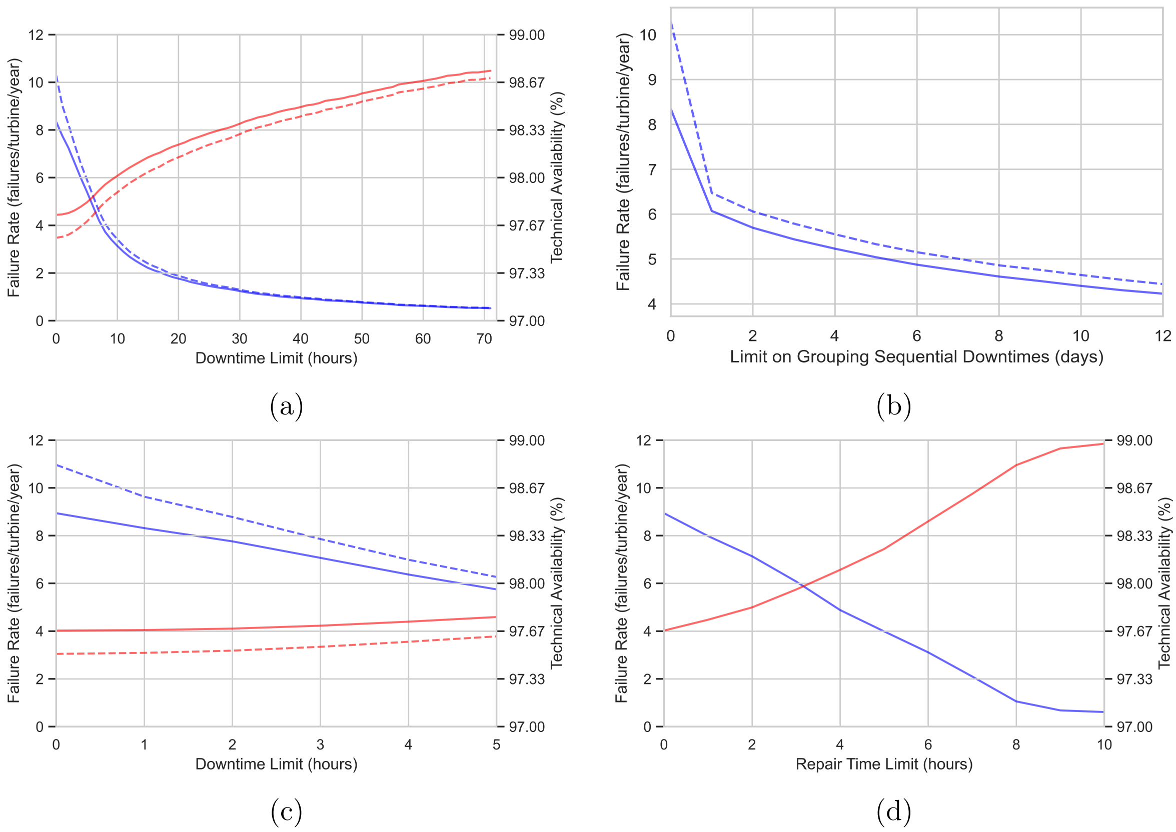


Figure 4.3: Sensitivity of baseline failure rate (blue lines) to (a) downtime limit, (b) grouping limit and (d) repair limit. (c) Also shows the downtime limit from 0 to 5 hours on the x-axis. Red lines show the corresponding technical availability. Solid lines include only downtime events with *both* manually acknowledged drop-off and pick-up; dashed lines include downtime events with *either* a manually acknowledged drop-off or pick-up.

10.17 failures per turbine per year. This has a particularly significant effect on technical availability, reducing the baseline estimate by 1.55%. Note that this availability estimate contains all downtimes where a corrective action was carried out, even if the majority of the downtime was due to (e.g.) an annual service. BoP jobs increase the baseline to 9.78 failures. However, the baseline estimate can also be decreased via plausible data selection measures. Filtering 'no-assembly' and 'fault finding' missions have a similarly significant effect to the inclusion of opportunistic jobs, reducing the baseline to 7.07 and 7.25 respectively. Filtering out jobs which could be alternatively labelled 'retrofit' or 'annual service' reduces the baseline to 7.54. The technical availability reduction in all of these scenarios is similar, at around 2%.

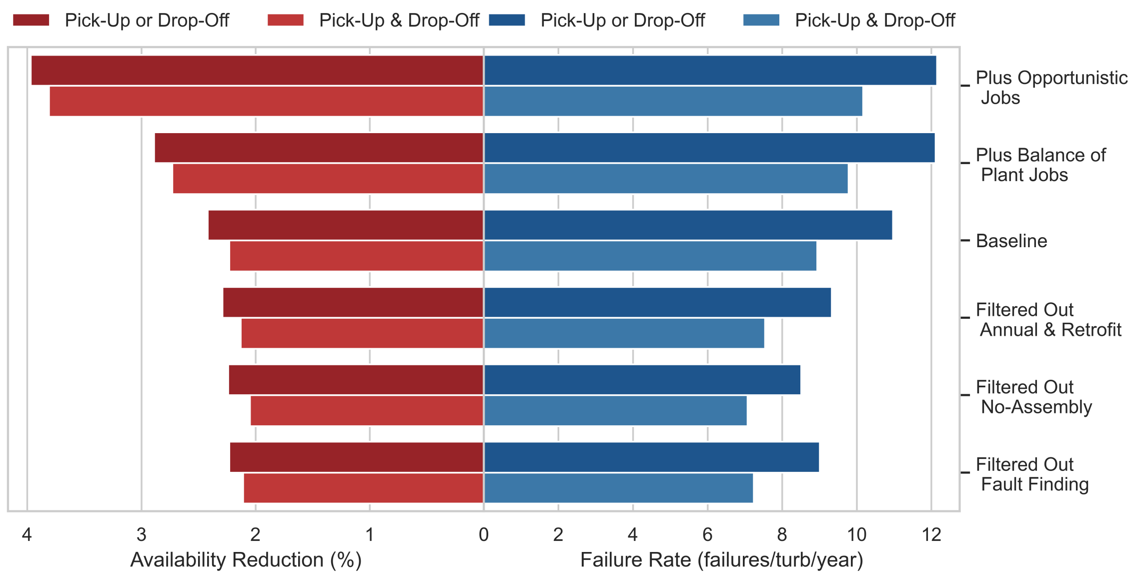


Figure 4.4: Bar chart exploring the sensitivity of the failure rate estimate (blue) and technical availability drop (red) to various data selection criteria defined in section section 3.4.5.

4.4 Work Procedure Analysis

This section expands on the previous section's analysis on failure rates. It shifts perspective from trying to define a 'failure' to scrutinising the work procedures in more detail. It therefore presents an analysis of the various work procedures that make up *Downtime Catalogue*. In doing so, the analysis can move towards a reliability analysis at the assembly-level.

The same process for defining turbine level failures as defined in section section 3.4.3. The analysis stops short at the work procedure name - i.e. there is no attempt at this stage to map work procedures to a specific failure taxonomy. The baseline failure definition is altered by the following conditions:

1. A downtime limit of 48 hours is placed on consecutive downtime events *only if the downtime events were due to the same work procedure*. Note that this is slightly different to the downtime limit applied in the above analysis looking at sensitivity of failure rates.

2. A downtime limit of 1 hour.

3. A repair time limit of 0.

4.4.1 Maintenance Types

Figure 4.5 shows the contribution of each maintenance type to total downtime for the period covered. It shows that the majority of downtime is attributable to corrective works (61%); and works without a work procedure account for 15% of total downtime. That leaves 24% for scheduled works. Of the corrective works, major replacements collectively make up the majority of downtime (51%), followed by collective major repairs (24%), minor repairs (14%) and BoP jobs (10%). Of scheduled works, downtime from inspections is the main contributor (48%), followed closely by inspection works (47%) and not so closely by retrofits (5%).

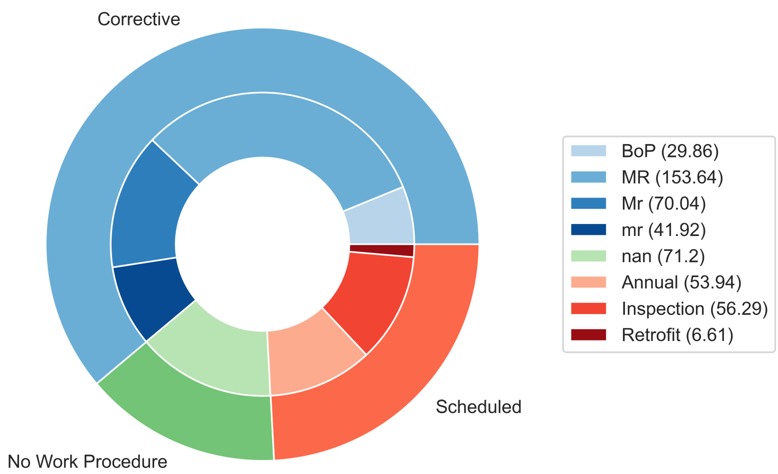


Figure 4.5: Contribution of different maintenance types to turbine downtime. The key for the inner ring is shown on the right-hand side. The outer ring is classified by whether it is scheduled or corrective. The number in the brackets show the average downtime per turbine per year due to each maintenance type in units of hours.

Figure 4.6 shows the fraction of interventions attributable to each maintenance category. In contrast to figure 4.5, just under half of interventions are due to scheduled works (44%), 8% do not have a work procedure and 48% from corrective works. This adds up to a total of 15.58 interventions per turbine per year on average. The majority of corrective interventions are for minor repairs (56%), followed by major repairs (26%), major replacements (13%), and BoP jobs (5%)

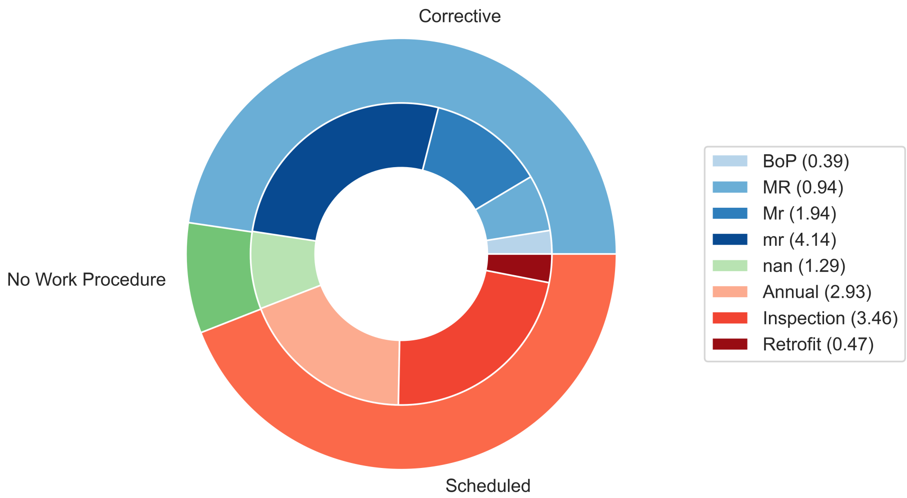


Figure 4.6: Contribution of different maintenance types to maintenance interventions. The key for the inner ring is shown on the right-hand side. The outer ring is classified by whether it is scheduled or corrective. The number in the brackets show the average number of interventions per turbine per year due to each maintenance type.

4.4.2 Work Procedure Statistics

Figure 4.7 presents statistics for each work procedure. Average intervention rates are shown on the left-hand side of the graph. On the right hand side of the graph, mean downtimes per intervention are shown. Each bar is split into two segments. The inner, solidly-coloured segment shows jobs. where only that work procedure was carried out during the downtime event. The outer, hatched segment shows additional interventions/downtime that were performed where more than one work procedure was recorded during the downtime event. Statistics for work procedures with extreme values for average intervention rate and downtime are presented in figures 4.8 and 4.9 respectively. Figure 4.9's work packages are all major replacements of components.

Corrective works present the greatest variety in work procedures. Most frequent are the 'fault finding' and tasks, followed by 'electrical panels'. Downtimes where only fault finding missions were recorded make up 12% of corrective works and account for

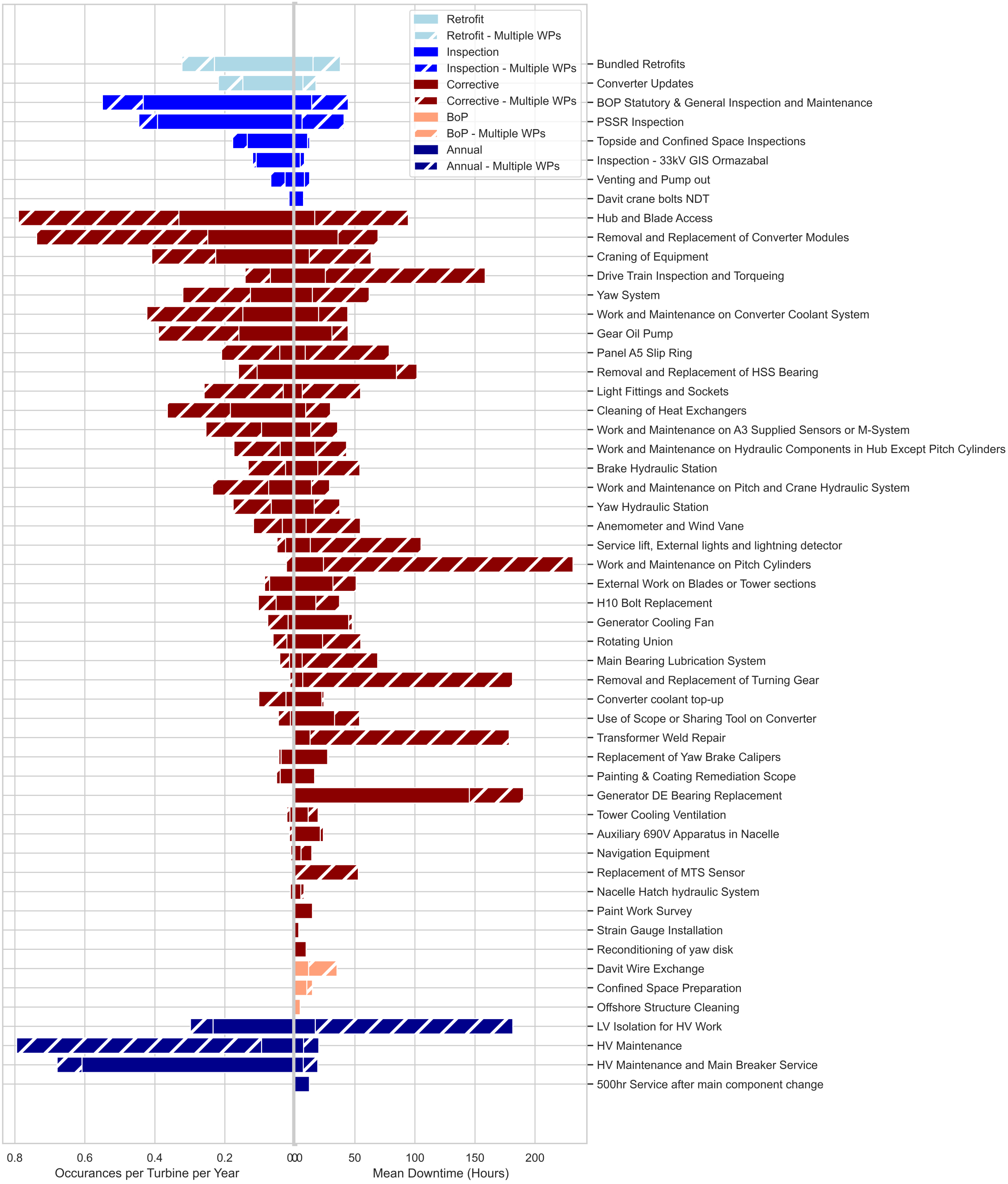


Figure 4.7: Mean intervention rate (number per turbine per year) and downtime (hours) for the most frequent work procedures in the dataset. Each bar is color coded according to maintenance type. Work procedures within each maintenance type are ordered (from top to bottom) by their total contribution to downtime. Dashed segments of the graph represent jobs where more than one work procedure is recorded.

6% of the downtime from corrective works. Jobs containing a fault finding mission make up 35% of corrective works and account for 50% of the downtime from corrective works. On their own, 'electrical panels' jobs account for 10% of corrective works and only 3% of downtime from corrective works. Jobs which contain an electrical panels work procedure account for 30% of corrective works and 49% of downtime from minor repairs. Other significant minor repairs in terms of contribution to total downtime are: hub and blade access; removal and replacement of delta modules; and craning of equipment

Of the scheduled works, statutory inspections (at around 2.2 interventions per turbine per year) and annual services (at around 2.3 interventions per turbine per year) are the most significant contributors to downtime. In contrast to corrective tasks, information on the sub-tasks for these non-corrective tasks is limited. The current analysis is therefore limited to these general work procedure descriptors for scheduled tasks.

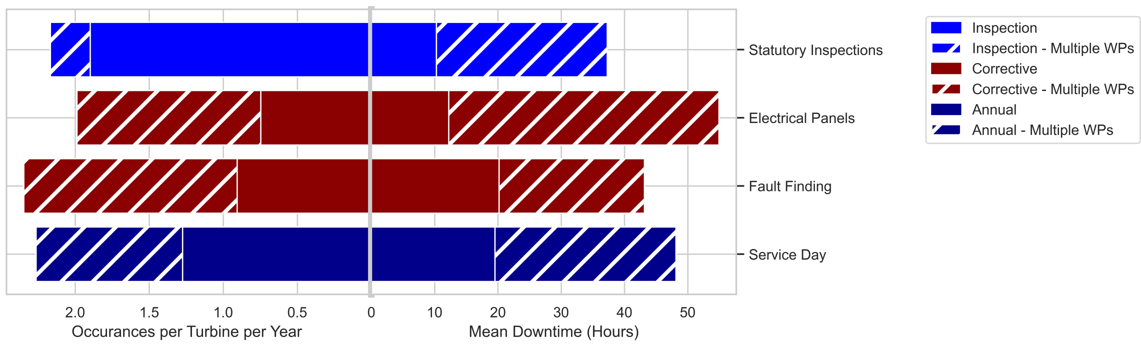


Figure 4.8: Mean intervention rate (number per turbine per year) and downtime (hours) for frequently occurring work procedures in the dataset. Each bar is color coded according to maintenance type. Work procedures within each maintenance type are ordered (from top to bottom) by their total contribution to downtime. Dashed segments of the graph represents opportunistic jobs.

Figure 4.9 presents metrics maintenance tasks which require a JUV. Their planning and execution is therefore entirely differently than the other tasks. Note that these only make up a portion of the 'major replacements' category. Together, JUV failures constitute less than 0.04 failures per turbine per year (less than 0.2% of all interventions). Notably, the mean downtime where multiple work procedures are present is at least twice as large as the stand alone repair time for all categories of major replacement. The generator and gearbox, in particular, have around 7 times longer downtime with multiple work procedures than stand-alone. It might be assumed that the operator takes advantage of the periods forced downtime to conduct (often many) other repairs on the turbine. The most significant contributor to downtime is the generator, followed by the gearbox, then transformer.

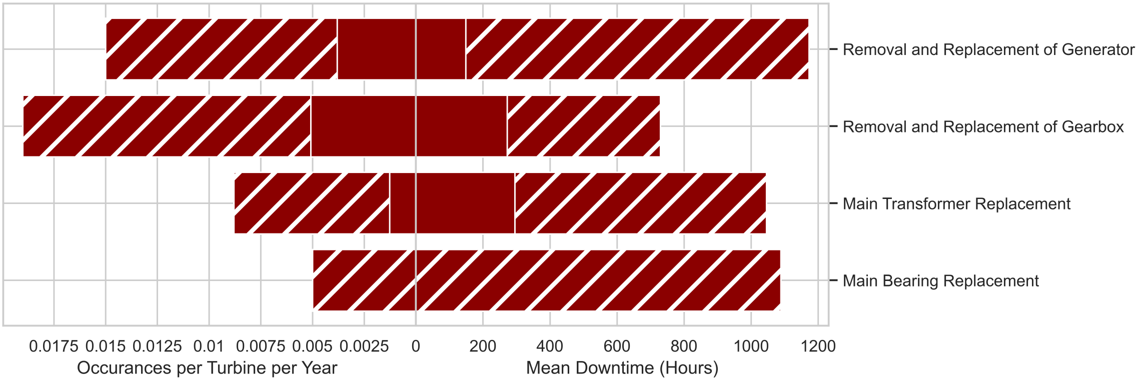


Figure 4.9: Mean intervention rate (number per turbine per year) and downtime (hours) for work procedures with long downtimes in the dataset. Each bar is color coded according to maintenance type. Work procedures within each maintenance type are ordered (from top to bottom) by their total contribution to downtime. Dashed segments of the graph represents opportunistic jobs.

4.5 Assembly-Level Failures

This section charts the transition from work procedures to assembly-level failure rate estimates. The definition of assembly-level failures follows the methodology laid out in section section 3.4.4. The baseline for assembly level failures is altered to reflect the results of section section 4.3. Namely, the following is added to the baseline definition:

1. Jobs which contain task descriptions in line with retrofits and annual services are labelled as such.

2. 'No assembly' jobs are filtered out.

Note, the baseline is re-defined by the above parameters so that the additional failures from different scenarios can be made obvious. It is not redefined to capture the 'best' failure definition.

4.5.1 Unclear Work Procedures

The majority of corrective work procedures presented in figures 4.7, 4.8, 4.9 can be categorised straight-forwardly by assembly. However, some work procedures could be categorised into several of the assembly categories summarised in figure 3.3. These unclear work procedures are as follows:

- Craning of Equipment;
- Drive Train Inspection and Torqueing;
- Fault Finding;
- Hub and Blade Access;
- Electrical Panels

The results of the methodology to categorise these work procedures into assembly-level failures is shown in figure 4.10. Craning of equipment and Panel A1 do not contribute a significant amount to any assembly failure rate - most of those work procedures will be categorised as 'nan' in the reliability analysis. Drive train inspection and torqueing mainly contributes to the gearbox and main bearing categories. Fault Finding is contributes significantly to multiple components: most prominently the converter, gearbox, yaw system, cooling system, and hydraulic group. Hub and blade access contributes mostly to blades, hub and hydraulic group, as one might expect. However, it also contributes a significant amount to the gearbox and other components in the drive train. This might be used as another argument for better data collection practices - it seems there is a contradiction here and the true reason for the downtime is ambiguous. Electrical Panels is spread out across multiple components.

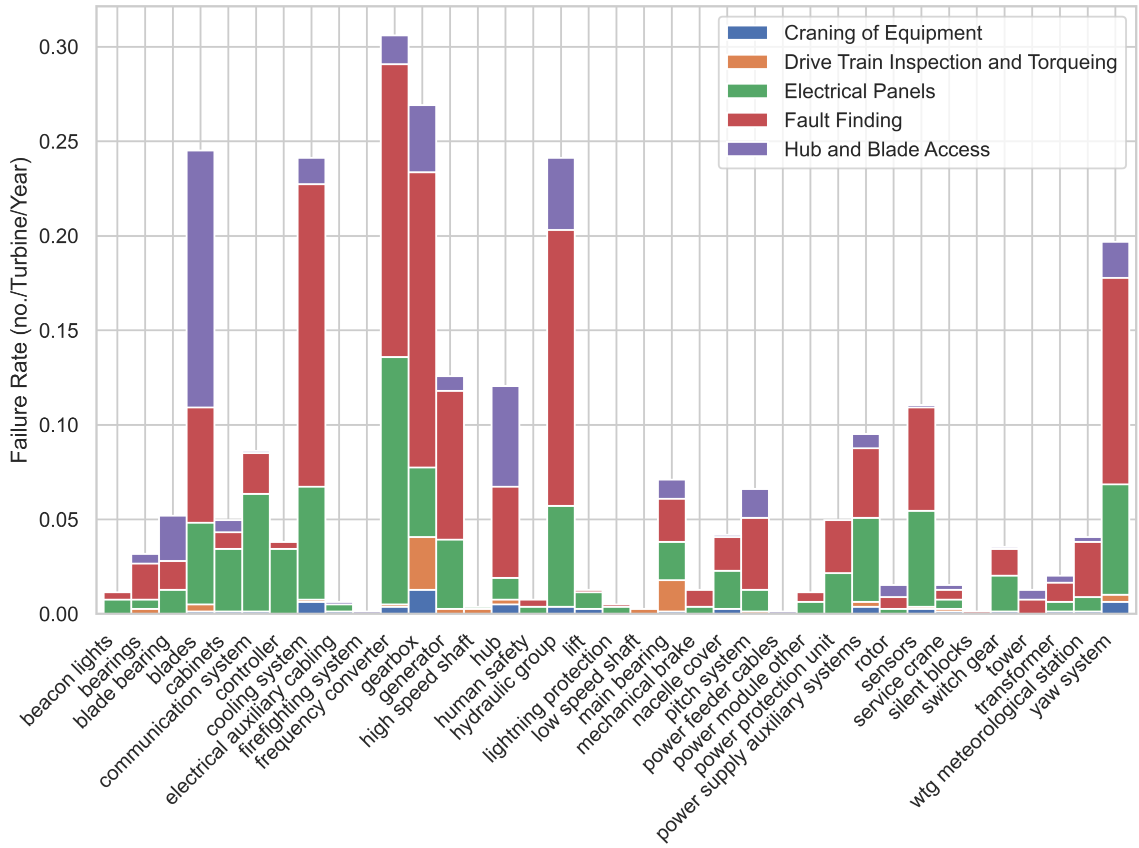


Figure 4.10: Breakdown of unclear work procedures by assembly, as estimated using the methodology defined in section section 3.4.4.

4.5.2 Assembly Failure Rate Estimates

Figure 4.12 shows the baseline estimates of assembly-level failures. Most failures arise from the various wind turbine cooling systems and converter, followed by the hydraulic group, gearbox, blades and yaw system. Note that, assemblies are defined by function. This means that assemblies like (e.g.) the cooling system, hydraulic group and drive train bearings might otherwise be categorised under different assemblies.

There is a significant proportion of repairs where multiple components could be identified in almost all categorise. This presents another source of uncertainty at the assembly-level. Compared to the figures presented by Carroll et al. [82], the proportion of MR and Mr repairs is high for several categories.

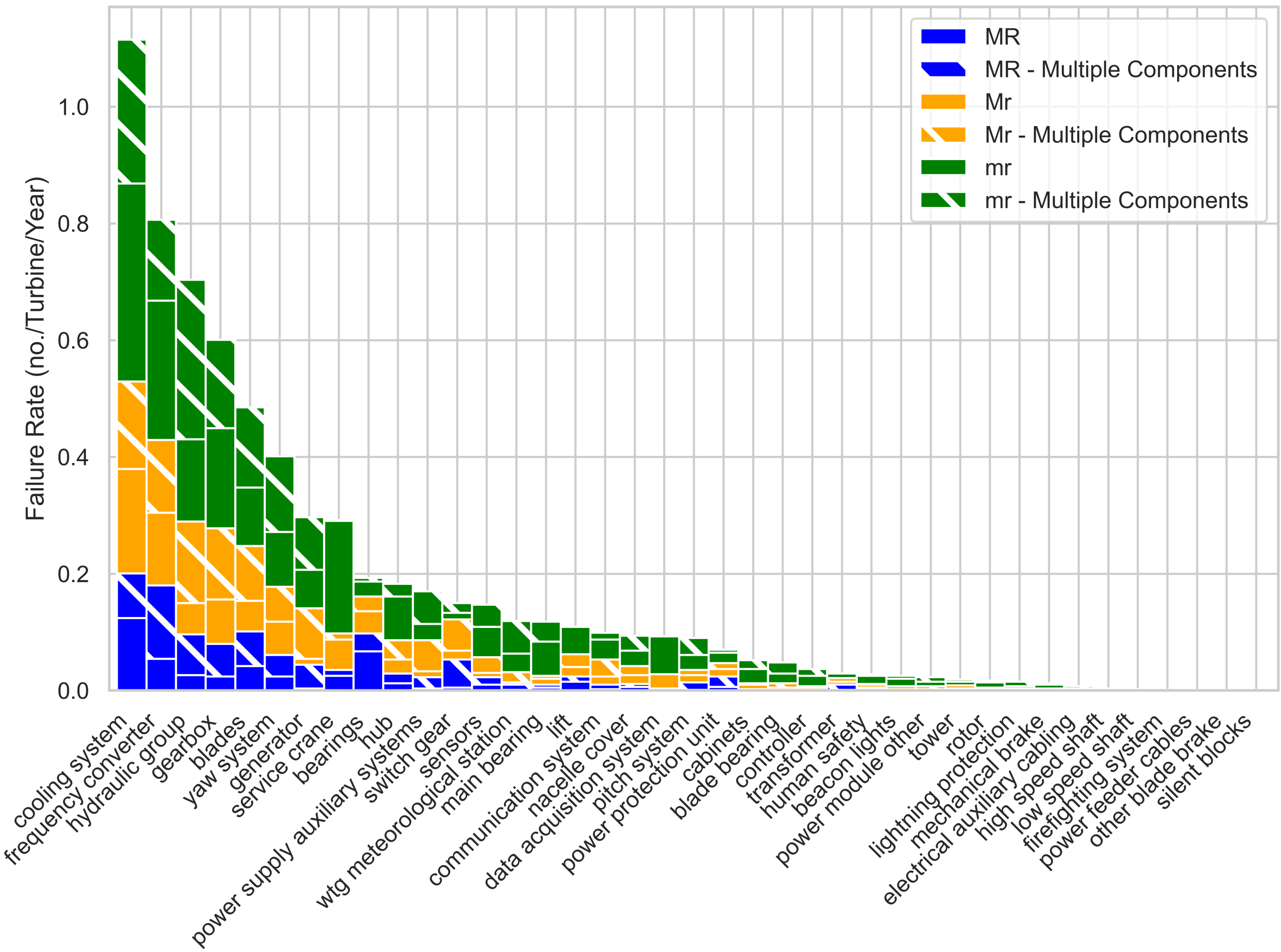


Figure 4.11: Assembly-level failure rates for when the repaired component is clearly identifiable (solid bars) and when multiple components could be identified during the one downtime (hathced bars).

4.5.3 Uncertainty in Assembly Failure Rate Estimates

Opportunistic jobs are shown by the non-blue bars. For most categories of failure, they do not have a very significant effect on the failure rate of the given assembly. The cooling system. frequency converter, service crane and lift present exceptions. BoP jobs, as shown by the purple bars, are srestricted to the service crane and lift. Retrofits are shown by orange bars - they have the most evident effect on the converters and cooling system, assemblies which are retrofitted regularly. The switch gear sees a lot of failures categorised under annual services - likely due to regular high voltage maintenance carried out as part of those campagins.

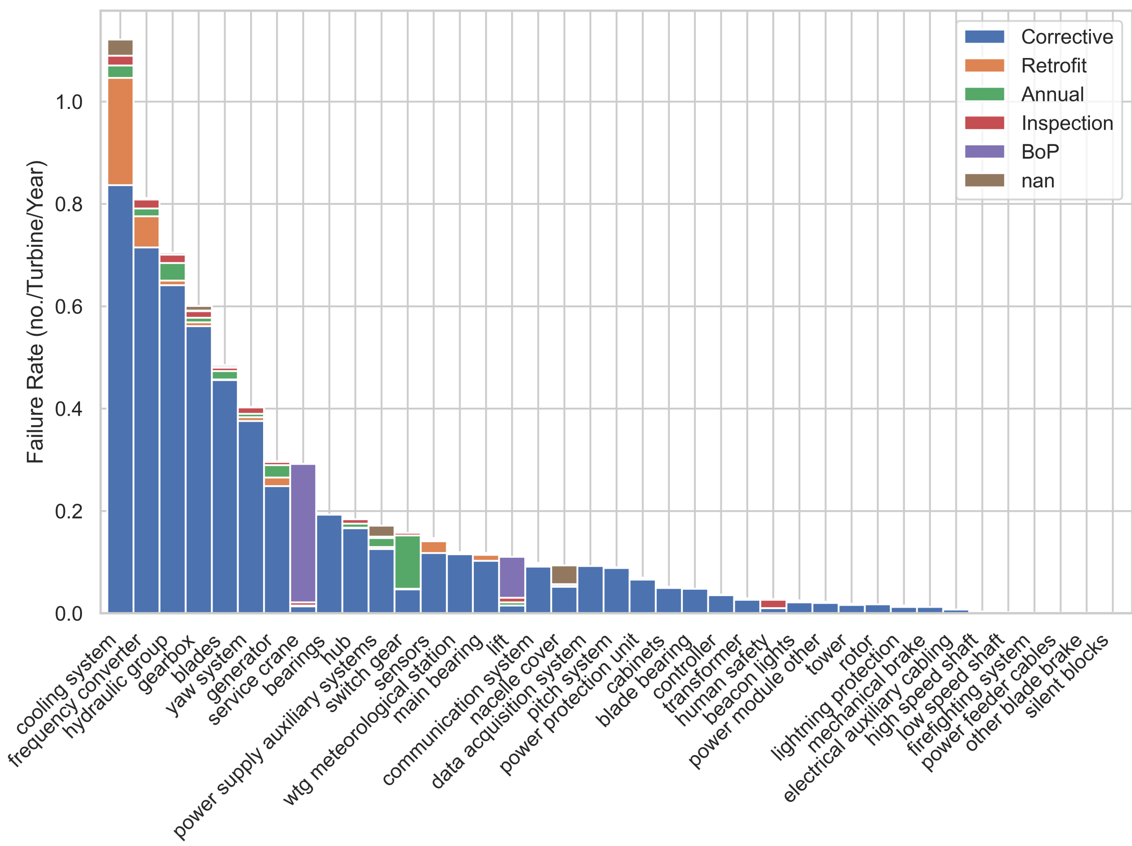


Figure 4.12: Assembly-level failure rates under different data selection criteria, shown in units of failures per turbine per year.

4.6 Discussion

This chapter presents frequentist statistics that could be calculated as a result of the data mining methodology presented in chapter chapter 3. The utility of the method is first demonstrated in section section 4.2, where the effect of tidal access restrictions were scrutinised. All of these statistics point to a reduced maintainability for tidally restricted turbines. This is important for wind farms which may be deployed in regions characterised by shallow water depths. For instance, currently operational offshore wind farms in Vietnam, Japan, Finland and Sweden, have an average water depth of just 2m, 6m, 8.5m and 8.5m respectively [237]. The implication for potential future sites situated in shallow water depths is that energy production may be less efficient than would be expected otherwise. For this particular site, it amounts to a 0.89% reduction in median technical availability. As shown in figure 4.2. It would be interesting to scrutinise the progression of these turbines' reliability over time, i.e. whether the comparative lack of opportunity for preventative works has a greater impact towards the end of their lifetime.

Sections section 4.3, section 4.4, section 4.5 are unusual as they effectively present the messy data pre-processing stage of a wind turbine reliability analysis. Aan analysis of this sort is valuable to the research community as it exposes the uncertainty surrounding reliability analysis of wind turbines. From this one dataset, results in section section 4.3 show derived failure rates ranging from below 1 failure per turbine per year to over 10 failures per turbine per year using failure definitions which have previously been used in the literature. Main sources of uncertainty for the failure rate estimate come from (1) data-preprocessing and (2) any limits placed on potential failures, evidence for which is presented in section section 4.3. Figure 4.4 explores the range of uncertainty in point (1); failure rate estimates range from 7.07 to 12.15 depending on which tasks are included/excluded. Figure 4.3 shows the range of uncertainty in point (2); the *downtime limit* reduces the baseline failure rate estimate by approximately 0.5 failures per turbine per year with every hour added to the limit, up until around 10 hours. The *active repair time limit* reduces the baseline failure rate approximately 1 failure per turbine per year with every hour added to the limit, up until around 8 hours. The most significant reduction in the *grouping limit* is in the first 24 hours, where the baseline estimate is reduced to 6.54 failures per turbine per year.

A lack of any standard failure definition presents a significant epistemic uncertainty to any wind turbine reliability analysis as this chapter has shown that different failure definitions can lead to very different failure rate estimates. Not only failure definitions, but different methodologies for data processing can lead to different failure rate estimates. The uncertainty revealed via the above results indicate a standard failure definition would be insufficient, and that there is a need for a recommended practice on data collection.

As reliability analyses are rare and valuable in the wind industry, one has been undertaken on the available dataset. The process of deriving assembly-level failures from turbine-level failures starts starts with an analysis of the work procedures in section section 4.4, then to the various means of assigning downtimes to a specific taxonomy outlined in section section 4.5. There are additional uncertainties which affect the results and have not been addressed here and are worth addressing. Firstly, there is uncertainty around data interpretation. Another (hypothetical) group of researchers may have produced different results for assembly-level failures, as many of them depend on alarm codes to ascribe them to a specific assembly. Given the dataset does not record the complete history of alarm codes, but simply the most likely culprit for the fault, this is something of a risky tactic. Secondly, there is no standard taxonomy in the wind industry. The taxonomy therefore becomes a sort of hyper-parameter which effects how the data is interpreted. A similar study could be undertaken to map the work procedures to other taxonomies used in the literature, most notably the RDS-PP taxonomy [41].

Secondly, the dataset upon which this study is based is from a single offshore wind farm. As a consequence: (a) failure rates are from one manufacturer and turbine model, and therefore represent a small subsection of turbines available on the market;

(b) the dataset only covers a small portion of the farm's lifetime, and may misrepresent the average failure rates of assemblies over the turbine's lifetime.

Thirdly, the scenario 'Filtered out annual service and retrofit' is largely a data quality issue. According to the operator, most tasks which are recorded as corrective, but whose work procedure imply a retrofit or annual service, are corrective in nature. This is why they were included in the 'baseline' calculation. Often when a turbine fails, and the solution is to retrofit it. With increasing take-up of better data standards on the industry, these kind of interpretation issues will be obsolete. As it stands, at least for this data set, this is another instance which requires manual interpretation, and could be taken either way depending on the mindset of the analyst.

4.7 Chapter Conclusion

This chapter presented a solution to the objective:

> "*Use the available dataset to calculate relevant KPIs describing wind turbine maintenance intervention. Use those KPIs to explore the maintenance requirements of offshore wind turbines.*"

In addressing this goal, this chapter relied upon the data mining methodology developed in chapter chapter 3. The utility of the data mining methodology was first demonstrated by comparing tidally-restricted to non-tidally restricted turbines. Tidally restricted turbines are characterised by generally shorter visits, resulting in differences in maintainability. This disparity was evident by a higher MTTR, and higher number of visits for tidally restricted turbines compared to non-tidally restricted.

This chapter then presents a reliability analysis of wind turbines at the turbine level and the assembly level. Its aim is to present failure data which is rarely presented and therefore valuable within the research community. However, it also aims to lay bare the considerable uncertainty associated with data-preprocessing of reliability data. It does so by exploring the sensitivity of failure rate estimates to different failure definitions and data selection criteria. The baseline failure definition used is 'A turbine downtime event accompanied by an unscheduled visit to that turbine'. Different interpretations of *downtime event* were explored by imposing a lower limit on the downtime of an event for it to be considered a failure; a similar lower limit on repair time; and a limit on the amount of time allowed to elapse between sequential downtime events at the same turbine to be grouped into one failure. The baseline failure rate estimate of 8.94 failures per turbine per year showed a considerable sensitivity to all of these factors. Including opportunistic jobs (where corrective tasks are included in scheduled works) in the estimate increases the baseline to 10.17; including corrective balance of plant jobs to 9.78. Filtering the corrective maintenance actions to exclude those which could also be interpreted as retrofitting or annual service activities reduces the baseline to 7.54. Filtering out jobs which couldn't be fit into the assembly taxonomy reduced the baseline to 7.25 and filtering 'fault finding' missions reduced the baseline to 7.07.

The chapter then presented an analysis at the assembly level, and likewise explored the uncertainty in assembly level estimates. Assembly failure rates showed high values for the frequency converter and cooling system, which both increased by including retrofitting jobs. The analysis also showed that many of the failures could be attributed to multiple assemblies, due either to multiple work procedures present in the downtime or multliple task descriptions/alarm codes.

5 A Bayesian Framework for Data Analysis

5.1 Chapter Overview

This chapter presents the methodologies developed within a Bayesian framework to address objective number 3 defined in section 1.3:

"Explore the utility of Bayesian data models in their application to operational maintenance data mining, and in leveraging value from operational maintenance data in general."

A significant motivator for the use of Bayesian methods is in their ability to quentify uncertainty [38], which has been a theme for O&M analysts in recent years [45, 43, 44]. The rationale behind Bayesian models as a solution to the areas for improvement identified in section section 2.4.5 are outlined in section section 2.5. To reiterate, Bayesian methods map well to the areas for improvement identified in chapter chapter 2. Their suitability for small and incomplete datasets [38] maps to characteristic of generally poor data quality of reliability and maintenance data [25]. They provide inherent uncertainty quantification [39] and can incorporate multiple sources of data [40].

Two Bayesian methods are developed. The first, presented in section section 5.2.1 relates to short-to-medium term Bayesian updating of prior estimates of relavent KPIs based on limited data-points. A Bayesian hierarchical modelling approach is developed for this purpose. The second, presented in section section 5.3 relates to a longer time-horizon reliability model which incorporates time-dependent variables.

The utility of the developed methodologies is demonstrated by the case studies of chapter chapter 6.

5.2 Bayesian Hierarchical Modelling Framework

5.2.1 The Fundamentals of Bayesian Inference

Bayesian inference presents as a powerful framework for making decisions and drawing conclusions in the face of uncertainty [38]. Unlike frequentist statistics, where probabilities are associated with the long-run behavior of random processes, Bayesian statistics treats probability as a measure of belief or uncertainty. This paradigm shift allows practitioners to incorporate both prior information and observed data in a systematic manner to arrive at posterior probabilities, which represent the updated beliefs after accounting for new evidence [38]. See Gelman [38] for a detailed overview of Bayesian data analysis. The following subsection relies heavily on their work.

Bayesian Data Analysis consists of three main steps:

1. Constructing a full probability model. Knowledge of the problem at hand is incorporated into a joint probability distribution for all quantities (both observable and non-observable) with which the model is concerned. A model providing a joint probability density function for some observed variable y and some parameter θ can be written as a product of two density functions as follows:

$$p(\theta, y) = p(\theta)p(y|\theta), \tag{5.1}$$

where $p(\theta)$ and $p(y|\theta)$ are referred to as the *prior* distribution and *likelihood* function respectively.

2. Conditioning on observed data. This second stage is the key principle of Bayesian inference, and is neatly encapsulated within Bayes' theorem:

$$p(\theta|y) = \frac{p(\theta)p(y|\theta)}{p(y)}, \tag{5.2}$$

where the conditional probability distribution of the unobserved quantities of interest, $p(\theta|y)$, is estimated. This is commonly referred to as the *posterior* distribution. $p(y)$ is the non-zero probability of observation y, which is given by a sum over all possible values of θ:

$$p(y) = \sum_{\theta} p(\theta)p(y|\theta) \quad \text{or} \quad p(y) = \int_{\theta} p(\theta)p(y|\theta)d\theta, \tag{5.3}$$

i.e. depending on whether θ represents a discrete or continuous variable respectively. Since this acts as a normalising constant within the model, Bayes rule can also be seen in its alternative form:

$$p(\theta|y) \propto p(\theta)p(y|\theta). \tag{5.4}$$

3. Evaluating the fit of the model and interpretation the consequences of the posterior distribution. Asking questions such as: *do inferences from the model make sense in light of the substantive knowledge?*, *how sensitive are results to modelling assumptions?* and *what aspects of reality are* (inevitably) *not captured by the model?* are important for statistical analyses in general. They are especially important for anything under the broad heading of machine learning, and specifically for Bayesian analysis for the reason that not everything can be encapsulated in a set of probability distributions.

5.2.2 Bayesian Networks

A Bayesian Network (BN) is a model of causal influence [165]. It models causal influence via a qualitative part known as a Directed Acyclical Graph (DAG), and a quantitative (or probabilistic [238]) part based on a collection of conditional probability functions [239]. The DAG is a series of links and nodes which encode the relationships between a set of random variables. Nodes represent variables within the model, while links represent conditional dependencies of a *descendant* to a *parent* [240]. Figure fig. 5.1 is an example of this, with X_1 being a parent to both X_2 and X_3; X_3 being a parent to both X_4 and X_5. Crucially, a lack of a link between nodes implies conditional independence between variables.

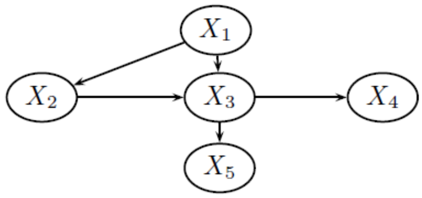


Figure 5.1: Example of a Directed Acyclic Graph.

The quantitative part of the model is therefore described by the set of conditional probability functions describing each node's relationship with their parents:

$$p(X_k|pa(X_k))_{k=1}^{n}, \tag{5.5}$$

where X_k represents some variable, $pa(X_k)$ represents the parents of that variable and n represents the number of nodes. In the scenario that the links represent discrete variables, these relationships are encoded by conditional probability tables (CPTs) [241] rather than conditional probability distributions (CPDs) [242]. The joint probability function can be given by the chain rule of probability:

$$p(X_k, \ldots, X_K) = \prod_{k=1}^{K} p(X_k|pa(X_k)) \tag{5.6}$$

The term *Bayesian network* can in of itself be misleading, as it does not necessarily guarantee a commitment to Bayesian statistics. In many instances, their construction relies on frequentest statistics to create CPDs or CDPs, and it is only in in their role as a basis for probabilistic inference that they become Bayesian. They are, however, extremely useful for representing Hierarchical models, which are widely applicable to a wide range of problems [243, 244, 245].

5.2.3 Bayesian Hierarchical Modelling

Hierarchical modelling is a generalisation of the typical Bayesian Network (BN) [184]. It differs from BNs in that they directly characterise the relationships manifest in structured data types. This is represented by figure 5.2, where a simple BN consisting of variables A, B and C takes on 3 different structural forms in an attempt to capture dependence between the subcategories BI and BII. The relationship between 'sub-nodes' BI and BII is referred to as a 'part-of' relationship by Gyftodimos and Flach [238]. The left-most diagram characterises this particular part-of relationship via a 'nested node'. The middle diagram represents a "*tree structure*" where the variables of the model are descendants of a top-level composite node t. The right-most diagram shows how the structure can be flattened back into a regular BN.

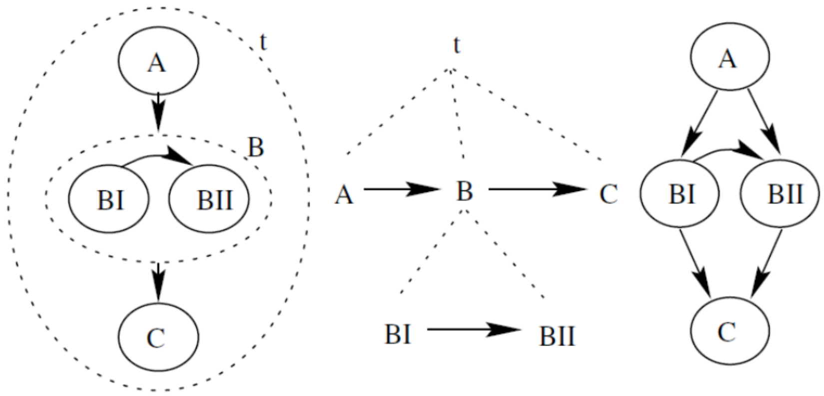


Figure 5.2: Various visualisations of related variables within a BN. Taken from [238].

This step from the non-hierarchical BN to the hierarchical model introduces some key differences which allow us to approach many multi-parameter problems in a much more natural and intuitive way [38]. In particular, the concept of "*partial pooling*" becomes useful in many applications [246, 247], as an alternative to pooled or unpooled models. This relates to phenomena which may be assumed to have a "*global*" effect, the influence of which varies under different conditions or between different subgroups [38]. To give a more explicit definition [38]:

1. A *pooled* model assumes that all data points, regardless of their group or category, are generated from a single common distribution, with no consideration for group-specific variations.

2. An *unpooled* model treats each group or category of data as if it has its own separate distribution or set of parameters, without sharing information across groups.

3. A *partially pooled* model simultaneously models both individual-level and group-level effects, allowing for the sharing of information across groups while accommodating group-specific variations.

To take an example from the wind industry: all wind turbines share a similar set of failure mechanisms, however the frequency or consequence of those failures are assumed to vary with wind speed/turbulence [217], maintenance strategy [248], WT type [78] and access conditions [14]. In the absence of hierarchical modelling, there are two options to account for these variations:

1. Estimate a global average from the broad pool of all wind turbines.

2. Consider failure rates and downtimes separately for (e.g.) each category of WT type.

Taken more generally, approach (1) is an example of a pooled model, and approach (2) an unpooled model [249]. Both will lead to inaccuracies in prediction. The first is insufficient for large datasets because it discards information relating to the inherent relationships in the data, and the second is in danger of over-fitting, meaning that it fits existing data very well but lacks any predictive power. A Bayesian hierarchical model finds a compromise such that information can be shared between groups whilst also maintaining the benefit of the global averages [250]. The 3 approaches are summarised in figure 5.3, which represents some modelling of parameters θ_i based on observations y_i. The options 5.3 (a) and (b) can be achieved by complete pooling and no pooling respectively. In Bayesian modelling, option 5.3 (c) is achieved not by specifying probabilistic distributions for θ_i in themselves but via so called hyperprior distributions for the parameters μ, σ. In this way the conditional probability distributions of separate groups (θ_i)'s are viewed as a sample from a common population distribution, and share information via their common hyperpriors [38]. This results in shrinkage of group means away from their individual sample towards the mean of the collective, an effect which is particularly useful for groups with small sample sizes [38]. In specifying our priors, the researcher therefore has influence over the mean and variance of posterior estimates for the hyperpriors μ, σ; influence over the mean and variance of parameters θ_i, and influence over the shrinkage effect of those θ_i parameters from the global mean.

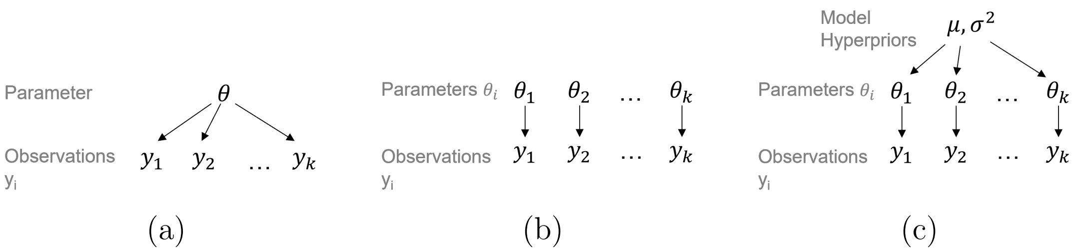


Figure 5.3: An example of (a) pooled (b) unpooled and (c) partial pooled models for some parameters θ_i, observations y_i and hyperpriors μ and σ^2.

5.2.4 Hyperpriors

In hierarchical models, prior information is introduced via so-called hyperpriors which regularise global averages and variances [38]. Broadly speaking, there are a two categories of prior to consider:

1. Non-informative or weakly informative priors [251]. Non-informative or weakly informative priors are used when there is no or little existing data or firm expert judgement available which relates to the problem in question. In this case it might seem convenient to assign equal probabilities to all possibilities (a uniform prior [252]), or to use (e.g.) a very diffuse normal distribution [251]. Non-informative or weakly informative priors are discouraged when using with small datasets [253]. To let the data dominate the posterior may lead to poor inferences, in the same way that frequentest approaches with small sample lack statistical significance. However, if there are enough samples such that allowing the data to dominate the posterior is statistically sound, they may be used [251].

2. Informative priors are at the other end of the scale - they convey precise information about a variable [254]. This is achieved either by expert elicitation or by some empirical Bayesian method [254]. The latter category encompasses

methods by which to estimate priors from the dataset itself, before undertaking a formal Bayesian analysis. This is something of a compromise between frequentist and Bayesian approaches which negates some of the benefits of the proposed methodology, so it is deemed unsuitable for this study. Of the former category, a metadata analysis of other studies may be hindered by common confidentiality practices of the industry, and surveyed expert opinions are outwith the scope of the project. In this case, it was judged that the most robust treatment of prior knowledge can be provided by a O&M cost modelling tool, of which there are many examples in the literature [14].

Deriving hyperpriors for the model can be acheived via a two-stage Bayesian updating approach as described by Yu et al.[254]. The process is summarised in figure 5.4. First, historical data for the parameters of interest are derived from the O&M cost model. The results of that modelling are used to perform Bayesian inference with weakly-informative priors for all parameters of the data-model. The means and variances of the posterior distributions of that initial Bayesian inference stage act as informative priors incorporated in the later inference procedure, which is where priors are condition on observed data.

5.2.5 Sampling Method

Since researchers cannot typically rely on exact posterior inference via an analytical solution of Bayes's rule (equation 5.2), they turn to numerical integration techniques, of which there are two categories: deterministic and stochastic [38]. Within the family of stochastic techniques, a general class of algorithms named Markov Chain Monte Carlo (MCMC) has been particularly important in making Bayesian inference practical for generic hierarchical models [38]. MCMC allows the Bayesian researcher to draw a series of correlated samples that will converge in distribution to the specified target distribution of the model [255].

Among this family of algorithms, Hamiltonian Monte Carlo (HMC) is often considered effective [256]. HMC works by treating probabilistic systems as if they are

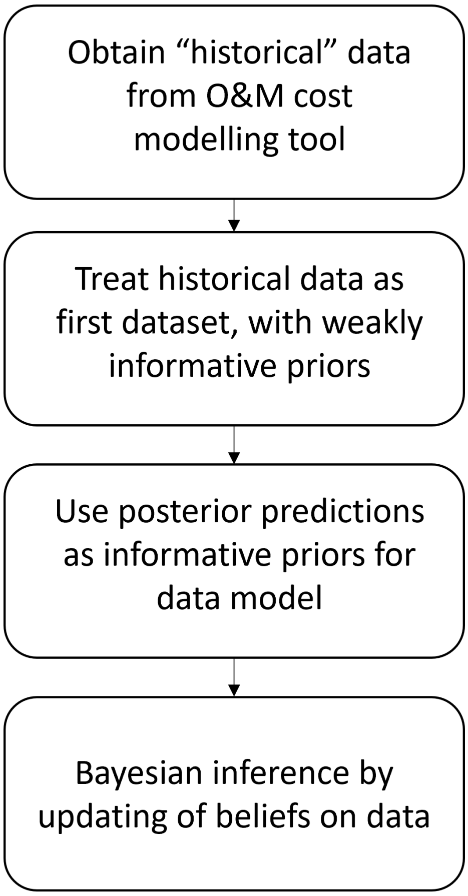


Figure 5.4: Summary of two-stage Bayesian uodating approach defined by Yu et al.[254].

physical systems, shifting the problem from that of sampling from a target distribution to the simulation of Hamiltonian dynamics [257]. This allows for for much more efficient sampling, as it avoids for the computationally inefficient random walk behaviour characteristic of algorithms such as Metropolis [258] or Gibbs [259] sampling to be done away with. A particular advantage of using PYMC3 software is the application of the No-U-Turn Sampler (NUTS), a variation on the typical HMC algorithm [256] which eliminates the need to hand-tune HMC via a parameter which defines the number of steps in the simulation. This provides a certain accessibility to practitioners, as tuning of the model usually requires some prior experience or expertise, and can be time-consuming.

5.2.6 Defining a Bayesian Hierarchical Model

Referring the the points outlined in section section 5.2.1, the first step of a Bayesian analysis consists of constructing a full probability model. In doing so, we can draw from the works of Gelman and Hill [260] and [254]. The model is summarised in figure 5.6. Layer A represents informative input from an O&M model. This is not discussed in the following subsections but is discussed in section section 6.2.2.1, where the case studies are presented. Layer B represents the hyperpriors for the covariance matrix, as described in section section 5.2.6.2. Layer C presents the inputs to the regression model, also described in section section 5.2.6.2. Layer D represents the regression framework itself, as decribed in section section 5.2.6.1. Finally, layer E represents the posterior specification.

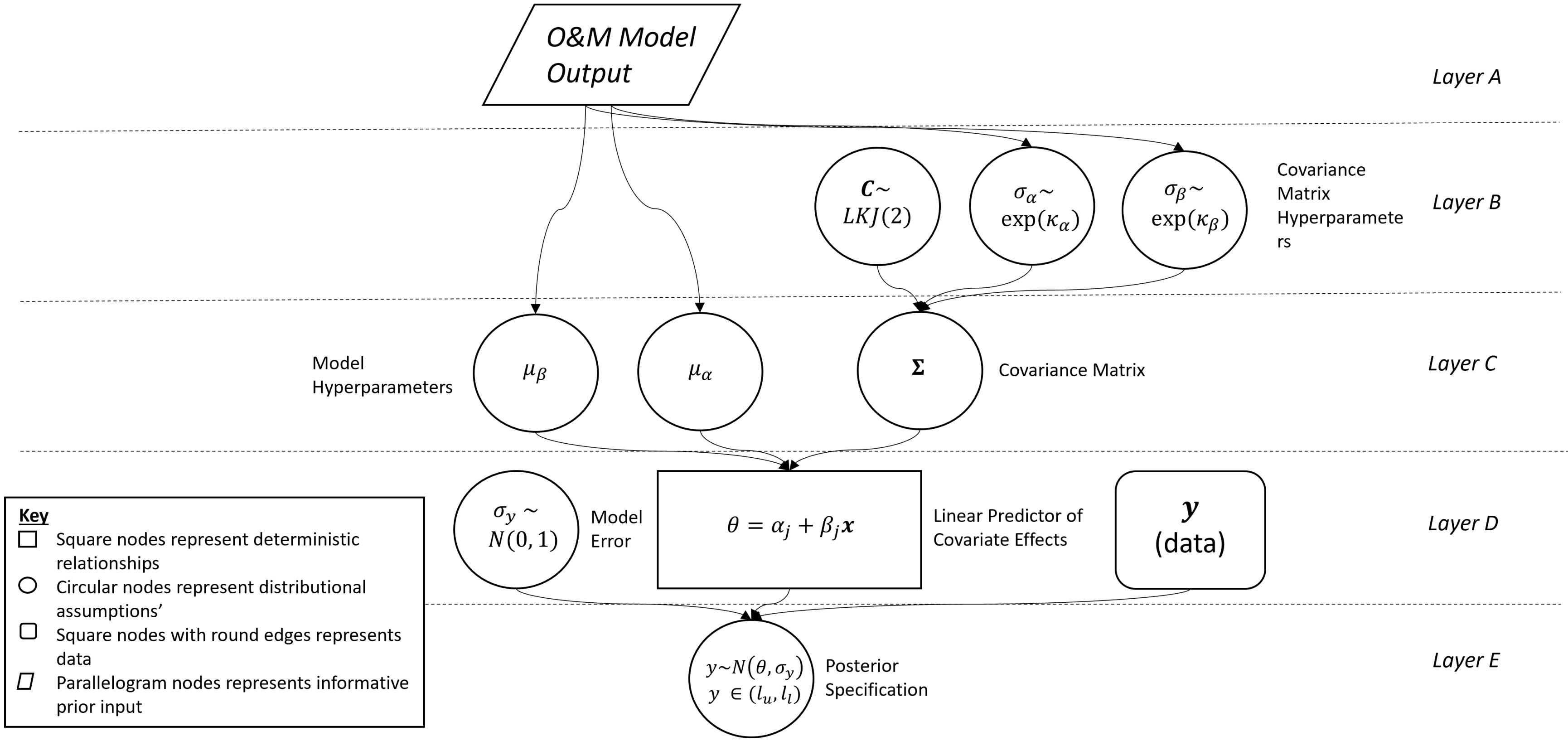


Figure 5.5: Diagram showing the relationships between model parameters and hyper-parameters in the Bayesian hierarchical model as described in sections section 5.2.6.1 and section 5.2.6.2.

5.2.6.1 Regression Framework

In the following, consider a dataset consisting of $\boldsymbol{y}$ data points describing the dependent variable of interest θ with a set of corresponding points $\boldsymbol{x}$ describing the independent variable. The size of the effect of the independent variable on y is given by a coefficient β and there is an intercept α which represents the value of $\boldsymbol{y}$ in the absence of $\boldsymbol{x}$. Also consider that $\boldsymbol{y}$ can be grouped in a hierarchical nature into j groups. There are two conventional approaches to modelling $\boldsymbol{y}$. The first is an unpooled model:

$$\theta = \alpha + \beta\boldsymbol{x} + \epsilon_i, \tag{5.7}$$

where ϵ is the model error. The second is a pooled model:

$$\theta = \alpha_j + \beta\boldsymbol{x} + \epsilon_i. \tag{5.8}$$

The simplest hierarchical model we could create to model θ ignores the effect of β and imposes partial pooling on α, such that a partially-pooled α (denoted $\hat{\alpha}$) is approximately a weighted average between groups:

$$\hat{\alpha} \approx \frac{(n_j/\sigma_y^2)\overline{y}_j + (1/\sigma_\alpha^2)\overline{y}}{(n_j/\sigma_y^2) + (1/\sigma_\alpha^2)}, \tag{5.9}$$

where n_j is the number of data points in group j, $\overline{y}_j$ is the mean observed value of each group (the unpooled estimate), $\overline{y}$ is the mean observed value for all data points (the pooled estimate), σ_y^2 is the within-group variance in values and σ_α^2 is the variance among group-level estimates. This represents both the relative information available about individual groups and the average of all groups. This means that:

1. Estimates for groups with smaller sample sizes will shrink towards the overall average;

2. Estimates for groups with larger sample sizes will be closer to the unpooled estimates and will influence the the group-wide average.

This is done in the hierarchical Bayesian framework by defining a hyperparameter on the α_j's, for example:

$$\begin{aligned} &\theta = \alpha_j; \\ &\alpha_j \sim \mathcal{N}(\mu_\alpha, \sigma_\alpha), \end{aligned} \tag{5.10}$$

where their mean μ_α and standard deviation σ_α are estimated from the data (via Bayes rule). The common distribution $\mathcal{N}(\mu_\alpha, \mu_\beta)$ has the effect of pulling the estimates of α_j towards the common mean in such a way that statistical strength is shared between groups. Note that the distributional assumption of a normal distribution ($\mathcal{N}$) can change based on the nature of the model. Also note that the normal distribution is defined in section section 5.2.6.2.

Equation 5.10 can now be expanded to also include a varying effect parameter, which behaves in the same way as the common hyper-parameter on α_j, e.g.;

$$\begin{aligned} &\theta = \alpha_j + \beta_j \boldsymbol{x_i}; \\ &\alpha_j \sim \mathcal{N}(\mu_\alpha, \sigma_\alpha); \\ &\beta_j \sim \mathcal{N}(\mu_\beta, \sigma_\beta), \end{aligned} \tag{5.11}$$

where again their mean μ_β and standard deviation σ_β are estimated from the data. We can see that now both the intercept and the effect vary by group. However, the model can be further expanded to include covariation between intercepts and effects. This can be done by specifying a multivariate Normal distribution over α_j and β_j, such that:

$$\begin{pmatrix} \alpha_j \\ \beta_j \end{pmatrix} = \mathcal{N}\left(\begin{pmatrix} \mu_\alpha \\ \mu_\beta \end{pmatrix}, \boldsymbol{\Sigma} \right), \tag{5.12}$$

where:

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_\alpha^2 & \rho\sigma_\alpha\sigma_\beta \\ \rho\sigma_\beta\sigma_\alpha & \sigma_\beta^2 \end{pmatrix} \tag{5.13}$$

and where ρ is a between-group correlation parameter.

5.2.6.2 Probability Model

In the Bayesian model, we have to make assumptions as to the probability distributions of the hyper-priors α, β, and the covariance matrix $\mathbf{\Sigma}$. In this thesis, μ_α, μ_β are defined by the two-stage Bayesian updating procedure described by [254]. They are represented by a normal distribution:

$$\begin{aligned} \alpha &= \sqrt{\frac{1}{2\pi\sigma_\alpha^2}} \exp\left(-\frac{1}{2\sigma_\alpha^2}(x_i - \mu_\alpha)^2\right) \\ \beta &= \sqrt{\frac{1}{2\pi\sigma_\beta^2}} \exp\left(-\frac{1}{2\sigma_\beta^2}(x_i - \mu_\beta)^2\right) \end{aligned} \tag{5.14}$$

In the first stage of the two-stage Bayesian updating procedure, the parameters μ_α, μ_β, σ_α and σ_β are selected to define a defuse distribution. In the second stage μ_α, μ_β, σ_α and σ_β are defined by running the values of the cost model as if they were historical data to obtain informative values.

Covariation between the groups is modelled by a multivariate Normal distribution:

$$\begin{pmatrix} \alpha_j \\ \beta_j \end{pmatrix} \sim \sqrt{\frac{1}{2\pi\mathbf{\Sigma}}} \exp\left(-\frac{1}{2}(\boldsymbol{x_i} - \boldsymbol{\mu})^T \mathbf{\Sigma}^{-1} (\boldsymbol{x_i} - \boldsymbol{\mu})\right), \tag{5.15}$$

where:

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_\alpha \\ \mu_\beta \end{pmatrix} \tag{5.16}$$

In order to induce a prior on $\mathbf{\Sigma}$, we need to define a prior on the correlation matrix of α_j and β_j and a prior for the standard deviations on σ_α and σ_β [261]. Again, these parameters are defined using the two-stage updating procedure as described above. In this thesis, σ_α and σ_β are defined by exponential distributions:

$$\begin{aligned} \sigma_\alpha &= \kappa_\alpha \exp(-\kappa_\alpha \boldsymbol{x_i}) \\ \sigma_\beta &= \kappa_\beta \exp(-\kappa_\beta \boldsymbol{x_i}) \end{aligned} \tag{5.17}$$

Lewandowski-Kurowicka-Joe (LKJ) distribution [262] is used to set a prior on the correlation matrix of $\mathbf{\Sigma}$. $\mathbf{\Sigma}$ can be written as [263, p. 442]:

$$\mathbf{\Sigma} = \begin{pmatrix} \sigma_\alpha & 0 \\ 0 & \sigma_\beta \end{pmatrix} \boldsymbol{C} \begin{pmatrix} \sigma_\alpha & 0 \\ 0 & \sigma_\beta \end{pmatrix}, \tag{5.18}$$

Where C is the correlation matrix. We can then induce a prior for $\mathbf{\Sigma}$ by setting a prior for the correlation matrix $\boldsymbol{C}$ and the standard deviations. The LKJ distribution is governed by a hyperparameter η such that:

$$\begin{aligned} &\boldsymbol{C} \sim LKJ(\eta); \\ &p(\boldsymbol{C}|\eta) \propto |\boldsymbol{C}|^{\eta-1} \end{aligned} \tag{5.19}$$

This means that $\eta = 1$ leads to a uniform distribution on correlation matrices and the magnitude of correlations between components decreases as $\eta \to \infty$.

Finally, a sampling distribution (y_{samp}) must be specified for the data for every group j. This is defined such that:

$$\begin{aligned} &y_{samp} \sim \mathcal{N}(\theta_j, \sigma_y); \\ &y_{samp} \in (l_u, l_l); \\ &\sigma \sim exp(\kappa_y), \end{aligned} \tag{5.20}$$

where, l_u, l_l represent upper and lower limits imposed by restrictions on the data (e.g. availability cannot be lower than 0 or greater than 1).

5.3 Reliability Analysis of Wind Turbines

While the above methodology is able to account for variability across multiple levels or groups in the dataset, the functionality of modelling is altered when considering time-to-event datasets. While the goal of the hierarchical models developed above is to describe the relationships and variability at each level of a hierarchy while also accounting for uncertainty [38], here the goal is to understand the time it takes for an event to occur and how various factors influence this time.

5.3.1 Hazard Scale Models

Modelling time between failures facilitates the use of hazard scale models [224]. Within a hazard scale formulation, the hazard function for a given event l at time t from our designated point of origin can be represented via a regression model:

$$h_l(t) = h_0(t) \exp\left(\phi_l(t)\right), \tag{5.21}$$

where $h_0(t)$ is the baseline hazard rate and $\phi_l(t)$ is some linear predictor describing the effect of the chosen covariates on event l at time t. The baseline hazard is therefore the estimated hazard rate in the absence of covariate effects. Typically wind turbine reliability modelling relies on exponential models:

$$h_l(t) = \lambda_l(t); \quad \lambda_l(t) = a \exp\left(\phi_l(t)\right). \tag{5.22}$$

Here, when covariate effects are assumed to be time-invariant a constant hazard rate is defined by a. Note the similarity to the formulation used by Slimacek and Lindqvist [137] to define their Homogeneous Poisson Process (HPP). In some instances [159, 264] researchers opt for Weibull distributed failures. In this case:

$$h_l(t) = \gamma t^{\gamma-1} \exp\left(\phi_l(t)\right) \tag{5.23}$$

where $\gamma > 0$ is a shape parameter.

The above model formulations are akin to the popular Cox model. The model presented here, however, differs by three means - via time-dependent covariate effects, frailty (random) effects and Bayesian estimation, which are elaborated on in the following subsections.

5.3.2 Covariate Effects

Covariate effects are wrapped up in the linear predictor
$phi_l(t)$:

$$\phi_l(t) = \boldsymbol{\beta}^T(t) \boldsymbol{X}_l(t), \tag{5.24}$$

where $\boldsymbol{X}_l(t) = [1, x_{l,1}(t), ..., x_{l,P}(t)]$ is a vector of covariates with $x_{l,p}(t)$ being the observed value of the p^{th} covariate for the l^{th} event at time t. $\boldsymbol{\beta}(t) = [\beta_0, \beta_1(t), ..., \beta_P(t)]$ is a vector of coefficient values in which β_0 is an intercept term and $\beta_p(t)$ is the coefficient for the p^{th} variable. Individual covariates contribute an *effect* via the term $\exp(\beta_p(t))$ from the exponential term in equation 5.23, otherwise known as the hazard ratio. This represents a relative increase or decrease in the baseline hazard due to the inclusion of a specific parameter. A unit increase in the covariate corresponds to a unit increase in the hazard, proportional to the hazard ratio.

Typically within WT reliability analyses, $\beta_p(t)$ is time-invariant, i.e.:

$$\beta_p(t) = \zeta_{p0}, \tag{5.25}$$

such that ζ_{p0} is a constant effect. This is commonly referred to as the proportional hazards assumption, and is key in formulating the Cox-model. Note that this assumption is implicitly held by Slimacek and Lindqvist [137] and any previous studies which employ covariate effects [159]. Some covariates may violate this assumption, however, by having a time-dependence. Such variables will be referred to as having a time-varying effect, since their hazard ratio is a function of time. In this thesis the time-dependence is modelled via a B-spline, or basis spline, so called because it combines a number of weighted basis functions to represent the given relationship. Thus, the covariate effect is modelled as:

$$\beta_p(t) = \zeta_{p0} + \sum_{m=1}^{M} \zeta_{pm} B_m(t; \boldsymbol{k}, \delta) \tag{5.26}$$

where ζ_{p0} is a constant, $B_m(t; \boldsymbol{k}, \delta)$ is the m^{th} basis term of a degree δ B-spline function subject to a vector of knot locations $\boldsymbol{k}$, and ζ_{pm} the corresponding m^{th} B-spline coefficient. Such a formulation utilising B-splines has been shown to be effective for modelling time varying effects, as shown by (e.g.) Perperoglou [265], Andrinopoulou et al. [266] and Gao et al [267]. Their effectiveness, however, depends on two hyperparameters; namely δ and the knot vector $\boldsymbol{k}$. Regarding the choice of δ, cubic basis functions (i.e. $\delta = 3$) are most popular in the literature [268]. Given the property of continuous first and second derivatives, they provide smooth interpolation when compared to linear and quadratic basis functions [269]. They also avoid some of the issues associated with higher order polynomials, for example Runge's problem.

Knot locations present a more complex problem. Improper selection of knot locations can lead to poor predictions via either overfitting (in the case where there are too many splines) or underfitting (in the case where too few knot locations) [270]. The literature presents two main ways of handling this issue. The first employs some form of penalty against over-fitting, as is common to many machine learning methodologies. This involves the selection of another smoothing parameter or penalty weight, which is optimised to achieve the best 'smoothness' between adjacent splines [265, 271]. The alternative approach is opted for, whereby several configurations of knot locations are considered and compared via a model selection criteria [272], such that a model is selected parsimoniously. More information on model comparison is given in section 3.3.3.

5.3.3 Frailty Models

Wind turbine reliability is difficult to properly quantify. There are myriad different covariates that have an influence. On top of this, the complete set of influential environmental, operational or design parameters that might be considered to improve the accuracy of any reliability model is rarely available to researchers - or indeed even operators. Often there is an additional heterogeneity between systems which is unexplained by model covariates. In survival or reliability analysis this hierarchical nature is commonly accounted for by including *frailties* in the model [273]. In effect this clusters observations into groups sharing a common characteristic; here, it takes the form a turbine-by-turbine frailty. This means that any i^{th} event belonging to any j^{th} given turbine are correlated, which can be modelled via a random effects term in the linear predictor:

$$\phi_{lq}(t) = \boldsymbol{\beta^T X_{lq}}(t) + \boldsymbol{b_q^T Z_{lq}}, \tag{5.27}$$

where $\boldsymbol{X_{lq}}$ has been given a subscript to denote events belonging to the turbine q, $\boldsymbol{Z_{lq}}$ is a vector of covariates for the i^{th} event for the q^{th} turbine and $\boldsymbol{b_q}$ is the associated vector of turbine-specific random effect parameters with covariance matrix $Cov(\boldsymbol{b_q}) = \Sigma_b$ and expected value $E\{\boldsymbol{b_q}\} = 0$.

Including frailty effects not only has the potential to improve the accuracy of the model, but addresses how to alter the traditional Cox-like parameterisation to include recurrent events for individual turbines. As explained by Amorim & Cai [274], there are two options for datasets of this type. The first is to employ a Markov process, where future events depend only on the immediate past. The other is via shared random effects or frailties, as is the base here. This is a natural choice for wind turbines: failure rates may differ significantly even among fleets of the same model and it is often difficult to attribute any discrepancies to observable covariates.

5.3.4 Priors

As with the hierarchical model specified in section section 5.2.6.2, a full-probability model must be constructed by specifying priors for model parameters. In contrast to the Bayesian hierarchical modelling framework, it is assumed that there are enough failures at the turbine level to allow for weakly informative priors. Additionally, in the absence of structured expert opinion it is difficult to define informative priors for many parameters. The time-varying effect of maintenance works on failure rate, for instance, is a novel covariate for which no data points exist. Neither is the effect accounted for by state-of-the-art decision tools. The covariance matrices which characterise the frailties of individual turbines would vary from farm-to-farm.

Having determined that priors will be weakly-informative, this thesis makes the following assumptions:

1. **Intercept value** β_0. Models contain an intercept term in the linear predictor which partly characterises the baseline hazard. It takes the form of a diffuse normal distribution such that $\beta_0 \sim N(\mu, \sigma)$, where $\mu = 0$ and $\sigma = 5$. This specification implies little known effect by centering the prior mean coefficient

values on 0 and allowing them enough scope to deviate from that value via a relatively large standard deviation.

2. **Time-Invariant Regression Coefficients** ζ_{p0}. As above, a normal distribution is assumed such that $\theta_{p0} \sim N(\mu, \sigma)$, where $\mu = 0$ and $\sigma = 5$. This specification implies little known effect by centering the prior mean coefficient values on 0 and allowing them enough scope to deviate from that value via a relatively large standard deviation

3. **Time-Varying B-Spline Coefficients** θ_{pl}. The B-spline coefficients are decribed via a random walk where $\zeta_{p,1} \sim N(0, 1)$ and $\zeta_{p,m} \sim N(\zeta_{p,m-1}, \tau_p)$ for $m = 2, ..., M$, where M is the total number of basis terms. Using this specification also requires a prior for τ_p, which is defined as an exponential such that $\tau_p \sim \exp(\lambda_\tau)$, where $\lambda_{\tau]} = 5$. τ_p acts as a smoothing parameter for the B-Spline which can be used as a form of regularisation. Thus, the exponential form of prior is used to encourage low values and prevent over-fitting.

4. **Wiebull Scale Parameter** γ. Where a Weibull model is employed a Weibull model, the scale parameter is assumed to follow a half normal distribution, such that $\gamma \sim N(\mu, \sigma)$, where $\mu = 0, \sigma = 1$ and $\gamma > 0$.

5. **Covariance Matrix**. The covariance matrix is made up of a series of decompositions. First the covariance matrix $\boldsymbol{\Sigma}_b$ is decomposed into a correlation matrix $\boldsymbol{\Omega}$ and a vector of variances $\boldsymbol{D}$, such that:

$$\boldsymbol{\Sigma}_b = \boldsymbol{D}^{1/2}\boldsymbol{\Omega}\boldsymbol{D}^{1/2} \tag{5.28}$$

The variances are in turn decomposed into the product of a simplex vector $\boldsymbol{\pi}$ and the trace of the matrix, such that:

$$Var(x_i) = \pi_i Tr(\boldsymbol{\Sigma}_b) \tag{5.29}$$

where $Var(x_i)$ is the variance of some variable x_i and π_i is the i^{th} element of the vector $\boldsymbol{\pi}$. The trace is the product of the order of the matrix and the square of a scale parameter ξ. Again, the LKJ distribution [262] is used as a prior for $\boldsymbol{\Omega}$. The regularisation parameter is set to 1, meaning that the prior distribution is jointly uniform over all possible correlation matrices. A symmetric Dirichlet prior is used for π, which has a single concentration parameter $\omega > 0$. Again the default $\omega = 1$ is used, meaning jointly uniform probability for all simplexes. A Gamma distribution is specified for ξ, such that:

$$\xi = \frac{\beta^{\alpha} x^{\alpha-1} \exp(-\beta x)}{\Gamma(\alpha)}, \tag{5.30}$$

where $\alpha = 1$ is a shape parameter $\beta = 1$ is a rate parameter and $\Gamma(\alpha)$ is a Gamma function.

Figure 5.6 summarises the relationships between model variables for the Weibull model. Layer A is the outermost layer representing the random walk hyperparameter for time-dependent variables, as specified in point 3 of section section 5.3.4. Layer B represents covariate hyperparameters for time-dependent variables, also specified in point 3 of section section 5.3.4. Layer C represents the range of variables that make up the covariate effects and covariance matrix, as described in sections section 5.3.4 and section 5.3.4 respectively. Label D represents those two components. Layer E represents the regression framework described in section section 5.3.3. Layer F represents the hazard rate specified in section section 5.3.1.

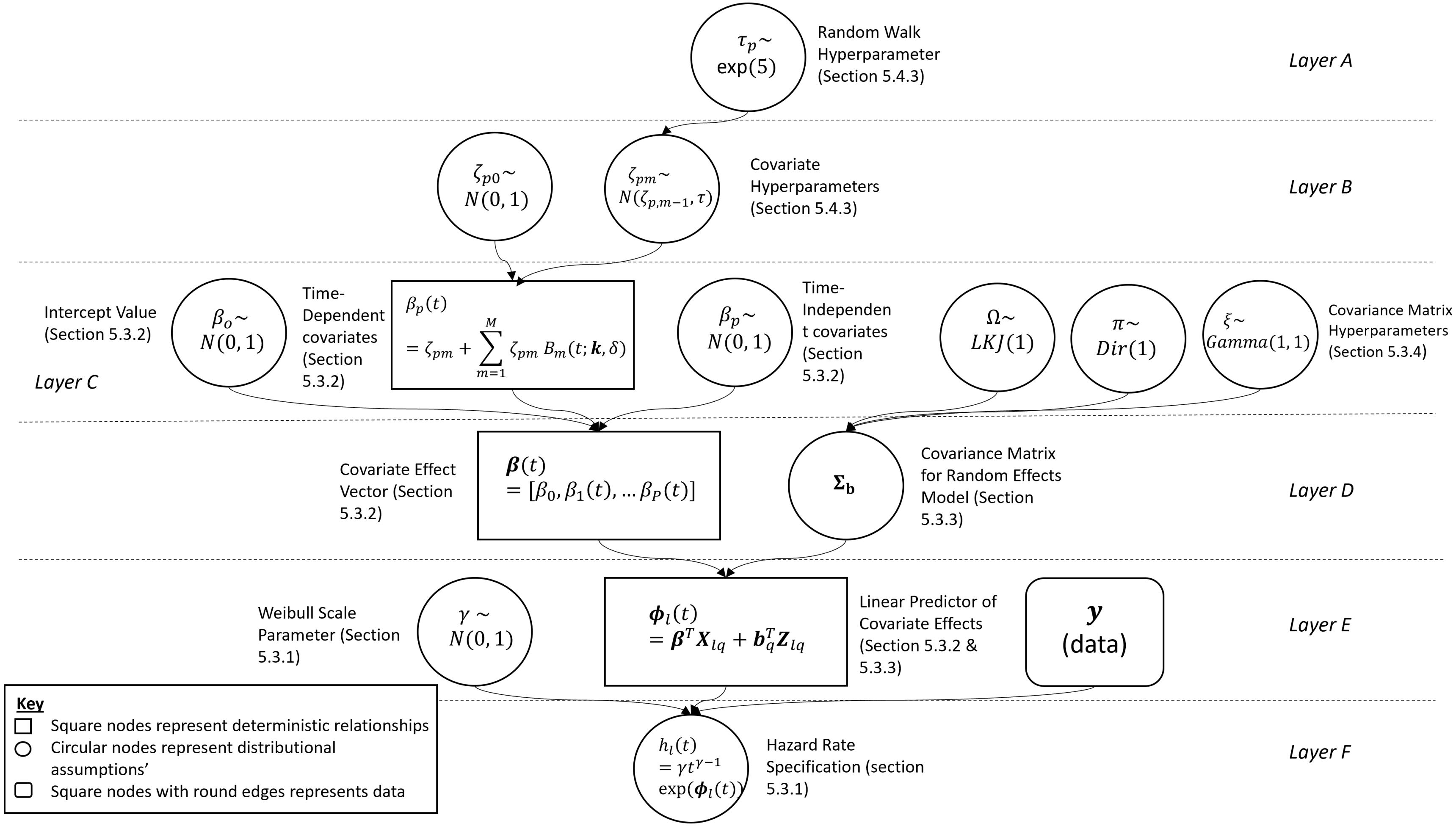


Figure 5.6: Diagram showing the relationships between model parameters & hyper-parameters as described in section section 5.3, along with their associated prior specifications.

5.3.5 Model Comparison

This analysis uses leave-one-out cross-validation (LOO-CV) to undertake model comparison [275]. LOO-CV was developed specifically for estimating out-of-sample predictive accuracy from the posterior sample of fitted Bayesian models. Consider a dataset $\boldsymbol{y} = (y_1, y_2, ..., y_n)$, which has been generated by some theoretical true generating mechanism for $\boldsymbol{y}$, $p_t(\boldsymbol{y})$. Also consider a new dataset $\tilde{\boldsymbol{y}} = (\tilde{y}_1, \tilde{y}_2, ..., \tilde{y}_n)$. $\tilde{\boldsymbol{y}}_{\boldsymbol{i}}$ which is independent from $\boldsymbol{y}$, but is also assumed to be generated by $p_t(\boldsymbol{y})$. In the Bayesian setting we arrive at a posterior predictive distribution for the new dataset $\tilde{\boldsymbol{y}}$ fitted on the orginal dataset $\boldsymbol{y}$:

$$p(\tilde{\boldsymbol{y}}|\boldsymbol{y}) = \int p(\tilde{\boldsymbol{y}}|\theta)p(\theta|\boldsymbol{y})d\theta \tag{5.31}$$

Evaluating a given set of models $\mathcal{M}_r \in \{\mathcal{M}_1, \mathcal{M}_2, ..., \mathcal{M}_R\}$ involves estimating the expected log point-wise predictive density (ELPD), which acts as a measure of predictive accuracy for the dataset $\tilde{\boldsymbol{y}}$. The ELDP is given by [276]:

$$elpd(\mathcal{M}_k|\boldsymbol{y}) = \sum_{i=1}^{n} \int p_t(\tilde{y}_i) \log p_k(\tilde{y}_i|\boldsymbol{y}) d\tilde{y}_i, \tag{5.32}$$

where $p_k(\tilde{y}_i|\boldsymbol{y_i})$ is the posterior predictive density for the model M_k. In practice. however, what one refers to as the *true* data generating process is unknown. Some means of approximation for equation 5.32 is therefore necessary. In many machine learning methodologies, out-of-sample test data is used for a similar purpose. In the Bayesian framework, cross-validation (CV) [277] has become popular. CV involves splitting the data into K parts which are respectively used as out-of-sample validation sets for the model fit with the remaining data. In LOO-CV the analysts estimates the predictive accuracy of the n data points taken one at a time, such that $K = n$. This is given by:

$$elpd_{LOO}(\mathcal{M}_k|\boldsymbol{y}) = \sum_{i=1}^{n} \log p_{\mathcal{M}_k}(_i|y_{i-1}), \tag{5.33}$$

where

$$\log p_{\mathcal{M}_k}(y_i|y_{i-1}) = \log \int p_{\mathcal{M}_k}(y_i|\theta) p_{\mathcal{M}_k}(\theta|y_{i-1}) d\theta \tag{5.34}$$

is the leave-one-out predictive log density for the i^{th} data point of the model $\mathcal{M}_k$. In practice we are only concerned with the comparison of two models, so it is sufficient to estimate the difference in their expected predictive power. For two models $\mathcal{M}_1$ and $\mathcal{M}_2$ (e.g. representing models with different knot vectors $\boldsymbol{k}$), the difference in their expected predictive accuracy is estimated as:

$$\begin{aligned} \widehat{elpd}_{LOO}(\mathcal{M}_1, \mathcal{M}_2|\boldsymbol{y}) &= \widehat{elpd}_{LOO}(\mathcal{M}_1|\boldsymbol{y}) - \widehat{elpd}_{LOO}(\mathcal{M}_2|y) \\ &= \sum_{i=1}^{n} (\log p_{\mathcal{M}_1}(y_i|y_{i-1}) - \log p_{\mathcal{M}_2}(y_i|y_{i-1})) \\ &= \sum_{i=1}^{n} \widehat{elpd}_{LOO,i}(M_1, M_2|\boldsymbol{y}). \end{aligned} \tag{5.35}$$

Now that there is a model comparison criterion in place, I employ a two stage step-wise procedure for model building as developed by Hofner et al. [278]. This consists of the following steps:

1. **Starting Model definition.** The starting model consists simply of the baseline hazard rate with turbine frailties. There are two choices for baseline hazard rate; namely exponential and Weibull. This is the initial model selection problem. Note that frailties must be included from the outset to account for recurrent events among the same turbine.

2. **Initial Choice Set.** Define a set of variables to be respectively added to the model. These are defined in section section 6.3.1.1. The primary problem is in selecting the number of knots for time-varying effects. This is done by increasing the number of knots to the point where the increase in model accuracy is insignificant. If any respective variable improves the model's predictive accuracy it is included.

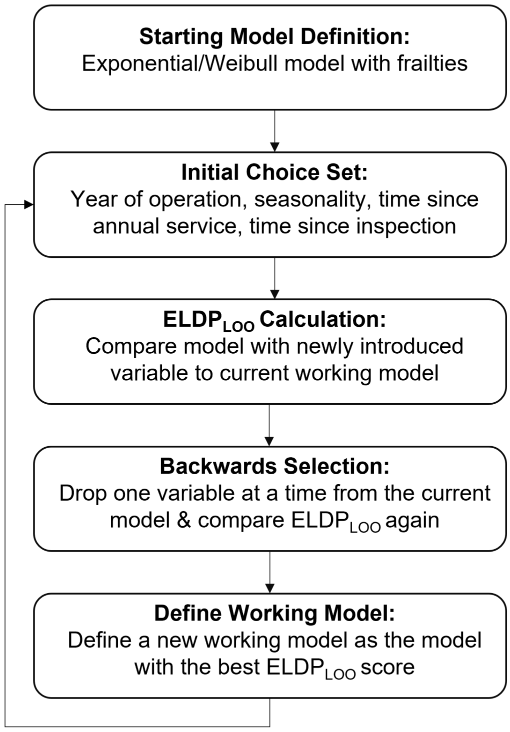


Figure 5.7: Flow chart summarising the model selection criteria

3. **Backwards Selection.** Perform a backward deletion step on the current model, i.e., estimate all hazard regression models obtained from the current model by dropping one covariate at a time. If an improvement of the model comparison criterion can be achieved, make the reduced model with optimal model comparison criterion the working model.

The process is summarised by figure 5.7.

5.4 Chapter Conclusion

This chapter presents two Bayesian methodologies which have been developed to extract value from the operational dataset. Bayesian methodologies were selected for their ability to handle small and incomplete datasets, their inherent uncertainty quantification capabilities, their white-box nature and their ability to incorporate different sources of knowledge.

Within the Bayesian regime, two modelling techniques were developed. The unique features of these techniques are summarised in table 5.1. The first, a Bayesian hierarchical modelling technique, utilises a two-stage Bayesian updating approach to derive informative priors and then update those priors based on relevant data. The second extends traditional reliability models into the Bayesian regime (for robust uncertainty quantification) using time-dependent variables.

Table 5.1: Summary of the key features of the Bayesian methodologies presented in this chapter

	Bayesian Hierarchical Model	**Bayesian Reliability Analysis**
Useful application	Modelling the relationships and variability of hierarchical data-types	Understaning the time it takes for an event to occur and how various factors influence this time
Data Type	Hierarchical (Multi-level) data	Time-to-event data
Key Components	Informative hyperparameters, hierarchical variables	Weakly-informative priors, Baseline hazard function, time-dependent variables, constant variables, frailty effects
Intended Use Case	Bayesian updating to assess maintenance strategy	Accurate representation of turbine time-to-failure
Intended time-frame	Short-to-mid term	Mid-to-long term

6 Bayesian Case Studies

6.1 Chapter Overview

This chapter presents a series of case studies based on the methodologies presented in chapter chapter 3. A version of the results that are presented here are published in [279] and [280]. Namely:

1. Section section 6.2 presents an application of the Bayesian hierarchical modelling methodology developed in section section 5.2.1 to the question of night shifts. This section therefore addresses the objective: "*Explore the effectiveness of night shifts in increasing power production and availability based on up-to-date, real-world operational data,*" the rationale for which is laid out in section section 2.6.2.
2. Section section 6.3 presents an application of the Bayesian reliability modelling methodology developed in section section 5.3 to the question of annual services and their effect on WT reliability. This section therefore addresses the objective: "*Explore the effect that annual services have on proceeding corrective works and in turn the reliability of wind turbines based on up-to-date, real-world operational data,*" the rationale for which is laid out in section section 2.6.3.

6.2 Bayesian Hierarchical Assessment of Night Shifts

This section presents a Bayesian hierarchical model to assess the effectiveness of night shifts in increasing power production and technical availability. The rationale for the analysis is presented in section section 2.6.2. To re-iterate, the question of night shifts was of interest to the industrial partner of this project as they were in operation at the WF from November to February inclusive. While previous studies [219, 150, 169] have investigated night shifts before (and found various advantages of using the strategy), no cost benefit analysis is provided on real-world, up-to-date data. Providing that cost benefit analysis is the aim of the following analysis.

6.2.1 Synthetic Data Creation

Two final steps of data manipulation on from *Downtime Catalogue* defined in section section 3.4.1 was required for the night shift analysis. This consisted of combining *Downtime Catalogue* and *Turbine Regressors*, and splitting the data into either monthly or weekly time periods to create *Monthly Performance Records* and *Weekly Performance Records* respectively. While it is typical for operators to track monthly Key Performance Indicators (KPIs) for offshore wind farms, *Weekly Performance Records* was created so that there would be more samples to inform the following Bayesian analysis. *Weekly Performance records* is therefore used as the input to the Bayesian model.

The impact of night shifts on lost production due to repairs was analysed using two pieces of logic which act on *Weekly Performance Records.*

The first addressed the period of November through February (inclusive), where night shifts were in operation at the site. This involved estimating the reduction of downtime achieved by the work which was undertaken at night, which involved reusing CTVs for a shift starting at 10pm. This time of year is characterised by a higher wind energy content, so the opportunity cost from any given failure is expected to be greater. Also, the industry tends to regard the winter months as a time of high failure rates. Research work has corroborated this notion [135, 136]. Algorithm 1 provides the psuedocode used for estimating a hypothetical lost production for each event, which was used to update *Weekly Performance Records.* Effectively the work undertaken at night is displaced to the next available shift, then assume a knock-on effect of subsequent works. There is a clause in the algorithm which estimates the effect of having different numbers of technicians in the night team, which is evaluated at 12, 9 and 6. These numbers were selected to represent fractions of the maximum number of technicians that could fit on one CTV. For technician teams of 6 or 9, *Night Shift Lookup Table* was used to estimate which maintenance actions could still go ahead with less technicians available, and would therefore lead to reduced lost production.

The second piece of logic addressed the period between March and October (inclusive), where night shifts were not in operation. This switches the question from *how much 'lost production' was avoided by employing night shifts?* to *how much 'lost production' could be avoided by employing night shifts over the entire year?* Ilook at all corrective events which had overnight downtime. It is assumed that one vessel would be repurposed to address corrective works at night, and that there is suitable resources in terms of technicians to do so. The algorithm employed (algorithm 2) incorporates the *Weather* data table, seeks out weather windows during the night where the significant wave height does not exceed 1.3m, and 're-orders' the timeline of repair works such that turbine downtime is reduced. 1.3m was used as it is the threshold used by the site. There is an obvious shortcoming with this approach: namely that it is unable to capture all of the operational factors which are inherent in deciding which turbines to visit during a given shift. This will differ given availability of spare parts, severity of fault and/or availability of appropriately qualified technicians. Since it is assumed that one vessel is in use during any given night shift, faults are selected simply by their proximity to the O&M service base. Again, we evaluate the strategy with a team of 12 technicians, 9 technicians and with a team of 6.

6.2.2 Bayesian Hierarchical Model Specification

Here, two models are built to explore KPIs of interest, namely:

1. Opportunity cost, presented in units of £/MW installed capacity. Calculated by estimating lost production due to corrective maintenance per week via the model, applying a constant price of energy (see section section 6.2.2.4) to each

Algorithm 1 Pseudocode for estimating avoided lost production during winter

```
for each row in Downtime Catalogue do
    Cut Operations down to between start & end of downtime event
    Cut Operations again to jobs that occurred during night shift
    Re-sample resulting Dataframe by night shift
    Group by turbine and shift
    if no. techs available = 12 then
        Calculate number of night-tasks completed & total hours completed
    else no. techs available < 12
        Use Night Shift Lookup Table to estimate how many tasks would have been completed
        Calculate number of night-tasks completed & total hours completed
    end if
end for
if no. night shifts > 0 then
    Cut Weather down to above end of downtime event
    Label subsequent day-shift hours as work hours or non-work hours
    Count out the requisite number of extra day work hours required to complete the task in question
    Store hypothetical task end time
else return 0
    Employ turbine regressors to calculate lost production and hypothetical lost production
```

posterior density point and dividing by the capacity of the farm. Lost production data points are denoted $\boldsymbol{y}_i^{LP}$

2. Time-based technical availability. Defined as the ratio of downtime due solely to corrective maintenance time as defined in section section 3.2.1. Availability data points are denoted $\boldsymbol{y}_i^{av}$.

Since both models have the same formulation with different priors. The superscript index is dropped and define generic variables for both. The model specification follows the Bayesian hierarchical methodology developed in section section 5.2.6. The covariate of interest is whether night shifts are in operation or not. This is labelled x_{NS} and takes a value of 0 or 1. The hierarchical groups are the months of the year such that $\alpha_j = \alpha_{seas}$ represents seasonality as a hierarchical construct. The model can therefore be summarised as:

$$\boldsymbol{y} = \alpha_{seas} + \beta_{seas}\boldsymbol{x_{NS}} + \epsilon_i; \tag{6.1}$$

Algorithm 2 Pseudocode for estimating the potential for avoided lost production for the months March through October

for each row in *Downtime Catalogue* **do**
  Cut *Operations* down to between start & end of downtime event
  Create daily lookup table with shift starting time and amount of time spent undertaking works for each downtime event
  Cross-reference look up table with *Weather* to determine hourly wave-height readings throughout duration of downtime
  Update lookup table to include fictitious night shifts
  Determine whether fictitious night shifts are weather-appropriate for undertaking maintenance works
  Lookup entries of fictitious night shift already recorded to assess whether there are enough technicians to perform work
  for each row in *Lookup Table* **do**
    if there exists available night shifts **then**
      if there are day shifts where weather conditions are fine but no work has been carried out **then**
        Assume insufficient resources and return 0 change in lost production
      else
        Reshuffle lookup table so that most immediate shifts are utilised in maintenance work
        Employ *turbine regressors* to calculate lost production and hypothetical lost production

$$\begin{pmatrix}\alpha_j\\ \beta_j\end{pmatrix} = \mathcal{N}\left(\begin{pmatrix}\mu_\alpha\\ \mu_\beta\end{pmatrix}, \mathbf{\Sigma}\right), \tag{6.2}$$

As part of the sensitivity analysis, the data is also categorised via failure rate. The influence of failure rate is explored by altering the hierarchical model to include it as another feature. In order to do so, turbines are discretised into 'high', 'medium' and 'low' failure rates. This is based on the sample quantiles, such that there are an approximately equal number of turbines in each bin. The priors shown in figure 6.1 are altered to accommodate the change in the model. This followed the same process as before, but with fewer turbines in the simulation. Where failure rates are introduced, the hierarchical structure is extended to accommodate two levels. In this case, the model shifts from modelling a farm-wide lost production/availability towards modelling sub-groups of turbines. This introduces another level in the hierarchy which must be represented in the data model. Equation 6.1 must be altered to reflect another level of hierarchy:

$$\boldsymbol{y} = \alpha_{seas,fr} + \beta_{seas,fr} x_{NS}, \tag{6.3}$$

where the index fr denotes failure rate quantiles.

6.2.2.1 Hyperprior Elicitation

Hierarchical models can overcome some of the issues associated with small data sample sizes by employing informative hyperpriors. In this case, the hyperpriors are the α and β parameters.

6.2.2.2 Hyperprior Elicitation

As summarised in section section 5.2.4, this thesis relies on an O&M cost model for hyperprior elicitation. This analysis relies on the use of the StrathOW-OM tool [17] to derive hyperpriors. The model is described in section section 3.5. Prior estimates for α and β, σ_α, σ_β and λ are obtained via a two-stage Bayesian updating approach as described in section section 5.2.4, where the output of the StrathOW-OM is considered as historical data. First, weakly informative *pre-priors* are specified, in the form of normal distributions for α and β, σ_α and σ_β with large variances in comparison to the means, and exponential distribution for σ_y with a weakly informative rate parameter. The StrathOW-OM was run with the input assumptions described in table 6.1. Outputs were adjusted so that the lost production and availability per week could be retrieved. More informative priors could then be estimated by retrieving posterior inferences for model parameters from the Bayesian hierarchical model, using the StrathOW-OM model output as the likelihood. The refinement of the parameters α and β for the lost production model, from the weakly informative pre-priors to the priors, and eventually to the posterior estimates is shown in figures 6.1 (a) & (b).

6.2.2.3 Sampling

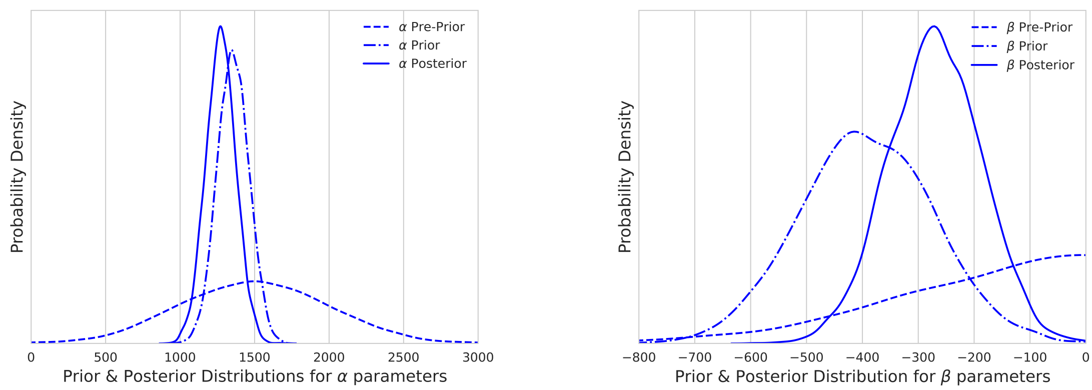
As described in section **??**, the NUTS sampler is readily applicable to the model. However, there are a few parameters which need to be specified before samples can be drawn from the posterior. 1,000 tuning steps are used, which effectively act as

Table 6.1: Input assumptions for the StrathOW-OM tool used for prior elicitation.

Variable	**Description & Source**
Failure Rates, Repair Times	The assumption of zero prior knowledge of failure rates at the current site before the implementation of the Bayesian model is made. Failure rates and repair times therefore come from Carroll et al. [82].
Vessels	It is assumed there are 6 CTVs operating at the site via a long-term charter, which is closest to the average of 6.33 vessel used per day, each having a 12-person capacity.
Weather	Weather readings from the on-site met and wave measurement equipment are fed into the tool, which in turn simulates weather conditions with similar characteristics.
Power Curve	The power curve was estimated using the SCADA readings from the site, by binning data by wind speed and averaging power output within each bin.
Governing Weather Criteria	CTV: wave height - 1.3m, SOV: wave height - 1.3m, Heavy-Lift vessel: wind speed - 10m/s.

an initial training period for the sampler to converge to the target distribution by optimising the step size parameter of the HMC. the number 1,000 was selected by experimentation - for computational efficiency it is best to minimise this while maintaining a robust model. Once this initial tuning has taken place, these samples are discarded as there is no guarantee that they have asymptotically come from the target distribution. The program goes on to sample 4,000 draws from the target posterior distribution using 4 chains (pymc3 uses 1 CPU core per chain - 4 chains reflects a 4 core computer). Again 4,000 draws is something of a compromise between computational efficiency and having enough posterior samples to draw from to reflect the target distribution. The number of chains specifies the number of Markov chains to run, and is useful for checking that the model has converged to a stationary target distribution.

Figure 6.2 explores the convergence of the NUTS algorithm given the above parameters. As described by Betancourt [281], when the marginal energy and energy transition distributions of any given Hamiltonian Markov transition are well matched,



(a) (b)

Figure 6.1: Kernel density estimate plots showing pre-prior, prior and posterior distributions for the hyperparameters (a) α and (b) β, representing the weekly lost production average value and the average savings value afforded by night shifts. X-axis figures are irrelevant - it is only important to see that the two distributions are well matched.

the Markov chains are likely to be performing robustly and will present small auto-correlations.

6.2.2.4 Cost Modelling Assumptions

There are additional expenses and savings to be considered in the case of night shifts, for which assumptions must be made. These relate to cost modelling, and are detailed below:

1. The number of vessels needed and the cost of their charters. It is assumed that the effect of night shifts is to 'displace' a vessel from the day shift to the night, effectively allowing the opportunity to avoid a vessel charter by re-using a day-time vessel. Further, It is assumed a daily vessel charter rate of £1,750, as is assumed by Dinwoodie et al [17].

2. The value of electricity. This will vary from farm to farm depending on the support mechanism employed in each case. For simplicity a constant value of 120 £/MWh is considered as the base case, and consider varying values of constant prices between 40-150 £/MWh within the sensitivity analyses.

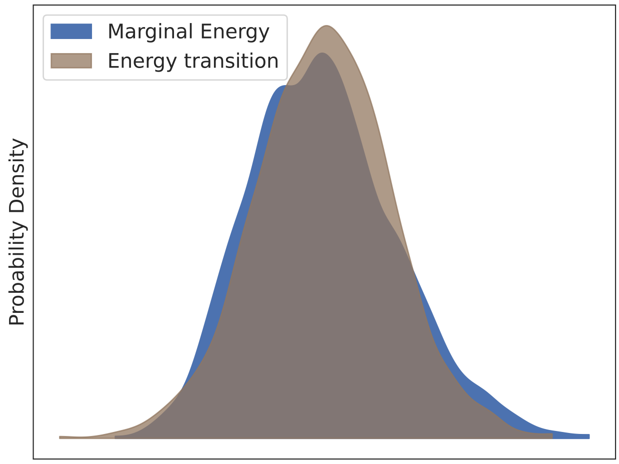


Figure 6.2: Energy plot showing the marginal energy and energy transition distributions of the Markov chains simulating model variables. Similar density functions imply a well sampled parameter space. The x-axis is irrelevant - what is important is how much the marginal energy and energy transition overlap [281].

3. Technician and support staff wages. There are three scenarios to select from: one that supports 6 technicians per night shift, one that supports 9 and one that supports 12, as well as onshore support staff and vessel skippers. For each scenario, It is assumed that there are two onshore staff and two vessel operators. It is assumed that all staff receive £20 an hour (based on the average technician salaray [282]), and assume a baseline case that staff receive time and a half for their work. This amounts to a weekly added cost of £13,400 in the case of 12 technicians, £10,920 in the case of 6 technicians and £8,400 in the case of 6 technicians. It is assumed that there are night workers on call for the entirety of the period that night shifts are in operation.

4. Months per year working 24/7. In the sensitivity analysis below only profitable months are included in yearly calculations. For example, if night shifts are not found to be profitable in the Summer months, they are excluded.

These assumptions are summarised in table 6.2 Where cost parameters are considered in the model, they are added to the posterior predictive distributions as a deterministic value.

Assumption	Baseline	High Value	Low Value
Extra Vessel Cost	£1,750/day	£1,750/Day	£0/Day
Electricity Price	120 £/MWh	150 £/Mwh	40 £/Mwh
Staff Wages (12 techs)	£13,400/week	£13,400/week	£0/week
Staff Wages (9 techs)	£10,920/week	£13,400/week	£0/week
Staff Wages (6 techs)	£8,400/week	£13,400/week	£0/week
Months per Year	N/A	N/A	N/A

Table 6.2: Cost modelling assumptions used to quantify a cost benefit to night shifts.

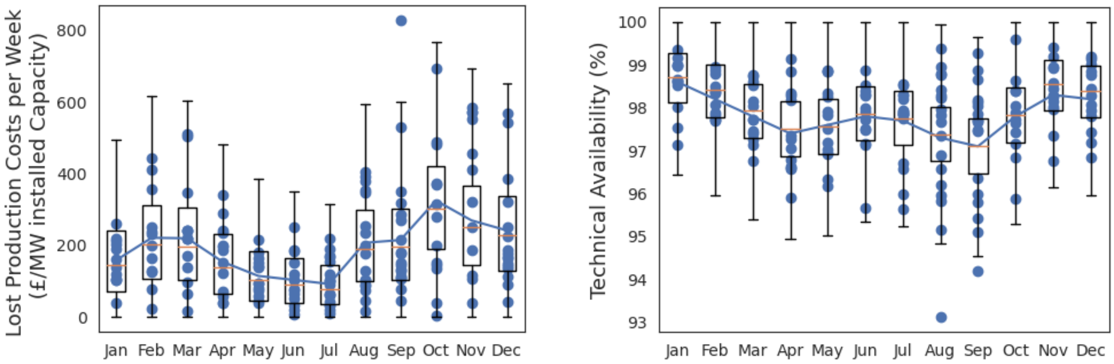
6.2.3 Night Shift Analysis - Results

6.2.3.1 Posterior Predictive Checking

Posterior predictive checking is the Bayesian's way of evaluating the fit of the model and interpreting the consequences of the posterior distribution.

Figures 6.3 (a) and (b) provide an immediate sanity check that the model is behaving as it should. Here, the posterior predicted values for (a) lost production and (b) availability in each month are plotted according to the operational strategy actually employed at the site (without any inclusion of 'expert knowledge') . They confirm that data generated form the posterior parameters of the model reflects that observed from the data. There is more lost production in the winter months, where wind energy content is higher, access conditions least favourable and it is thought that wind turbine failures are more likely to occur [135]. Winter months (where night shifts are in operation) are characterised by the highest time-based availability. The number of repairs carried out per month roughly reflects the trend observed by the SPARTA initiative [81] - the operator takes advantage of the favourable access conditions to carry out most minor repairs in the months with higher accessibility. Interestingly, a significant uptick in costly failures is observed in August, where wind conditions are still comparably light, and continues throughout the Autumn months. This might be a peculiarity of this site.

Figures 6.4 (a) and (b) explore the effect of Bayesian updating on the probability distribution (PDFs) of weekly lost production in the months of July and October respectively. The posterior predictive distribution in both cases is significantly different



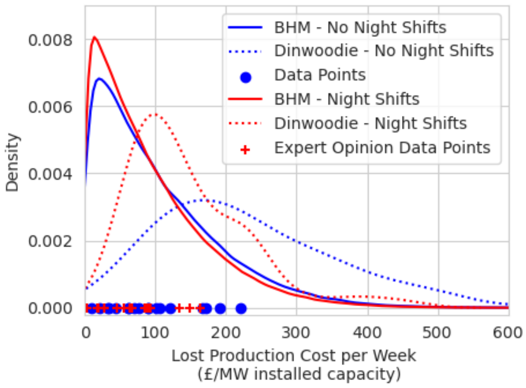
(a) (b)

Figure 6.3: Posterior density functions describing (a) lost production per week . Mean likelihood values calculated from the data are shown by the blue line, and individual data points by the blue dots.

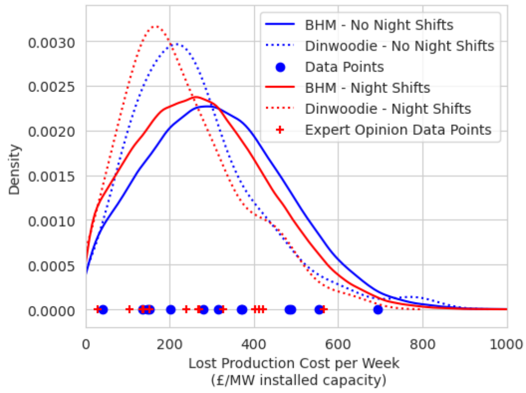
to the prior beliefs which were used to derive estimates of priors for variables. This illustrates the usefulness of the Bayesian Hierarchical approach in re-assessing strategy based on (statistically speaking) limited data-points. Given an operator considering extending night shifts from winter to other times of the year, a state-of-the-art operational cost model would estimate quite significant savings of opportunity cost in July. Likewise, they would have a misleading perception of October, where opportunity cost has been higher than expected. On the other hand, the data alone would be an insufficient means upon which to base future operational strategy, as the small number of samples in each month also provide an incomplete picture. The Bayesian hierarchical approach is a compromise between the two, such that samples in different months can be partially pooled to share statistical strength.

The difference in the probability distributions shown in figures 6.4 implies that the data points provide some detail where the Dinwoodie model is lacking that detail. For instance, the data points are characterised by certain features that are not captured in the StrathOW-OM model. Namely:

1. Varying vessel and technician resource. The StrathOW-OM model assumes that there is a constant number of CTVs in operation throughout the lifetime of the farm. However, the number of CTVs vary at the site to accommodate extra



(a)



(b)

Figure 6.4: Conditional probability density describing lost production for the months of (a) July and (b) October. Probability density functions derived via the StrathOW-OM model are shown via dotted lines, those derived via the hierarchical model by solid lines. Night shift scenarios are shown in red, non night shifts in blue.

maintenance work in the summer and less maintenance work in the winter. This leads to there being from 8-10 CTVs in summer at times and from 4-5 CTVs in winter. This goes some way to explain the difference in lost production probability distribution in July.

2. Seasonal failure rates. The StrathOW-OM model assumes constant failure rates for components. Additionally, the failure rates used in the initial modelling stage are taken from the work of Carroll et al. [82]. Failure rates in the data are likely to be different than those figures. When those failures occur is also important. This particular site, for instance, sees high autumnal failure rates, which is what leads to the different probability distributions in figure 6.4.

3. Opportunistic maintenance/multiple work procedures. The StrathOW-OM does not account for opportunistic maintenance. There is potential in the data for corrective works to be included as part of annual service campaigns, and therefore not get picked up as part of the corrective maintenance work. Annual services are likely to be undertaken in the summer, which may also lead to different predictions in those months than the StrathOW-OM model.

6.2.3.2 Night Shift Cost Savings

Results of the baseline scenario are shown in figure 6.5 and column 1 of table 6.3. Figure 6.5 shows the mean weekly lost production value throughout the year for 3 of the different scenarios considered: where there are no night shifts in operation, and where there are night shifts in operation with either 6 or 12 technicians. The PDFs with and without night shifts show similar shapes. The 'Night Shift' scenario PDF shifts the most likely values for each month to lower values of lost production. Profits from employing night shifts are predicted to be greater in the winter than summer months, as expected. However, profits are also substantial during the months of August, September and October, where the wind energy content tends to be lower than winter. The benefit of night shifts on corrective works therefore does not solely vary with wind speed, but also resource organisation throughout the year and potentially on the non-corrective works being undertaken at the site. Applying the baseline scenario assumptions lead us to infer a financial yearly saving of £1,625 per MW installed capacity.

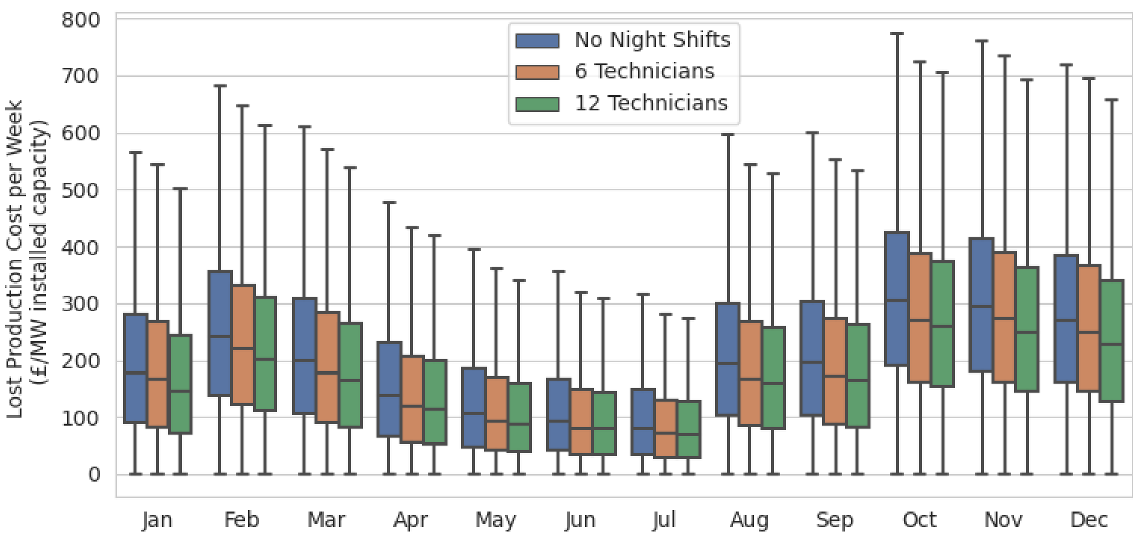


Figure 6.5: Seasonal variation of predicted lost production values in the scenarios where night shift are in operation with 12 technicians (blue), in operation with 6 technicians (orange) and out of operation (green).

Additional simulations were undertaken so that the impact of cost modelling assumptions on the viability of the night shift regime could be assessed. The first of these was the price of electricity, which at a constant rate can be assumed to have a linear effect on the estimated costs during any given month. This is explored in columns 4-6 of table 6.3. Months characterised by higher lost production values are more sensitive to the price of electricity variable. The months of June and July, for instance, are mainly influenced by savings which can be achieved via vessel leasing costs. The months of September and October vary much more. As do the winter months of January and February characterised by high energy content. The yearly costs, as described in the final row of table 6.3, vary significantly. At lower costs of energy the additional risk of undertaking work at night may not be worthwhile. However, there are still significant savings to be had at lower costs of energy with fewer technicians on night shift.

Table 6.3: Table exploring sensitivity of the baseline scenario to modelling assumptions.

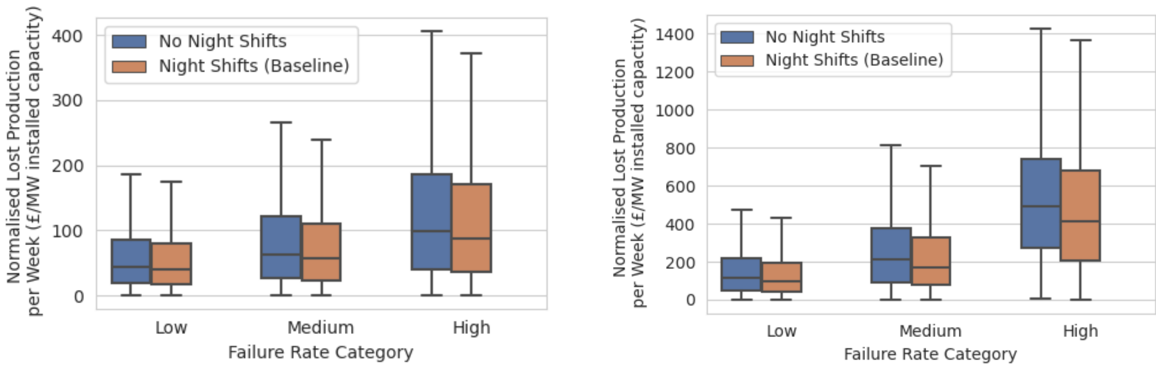
	Weekly mean Δ Night Shift - no Night Shift (£/MW installed capacity)								
	Failure Rate			**Electricity Price**			**Support Staff Wages**		**Vessel Savings**
	Baseline	**Low**	**High**	**£40 /MWh**	**£80 /MWh**	**£150 /MWh**	**(x2)**	**(x1)**	**No Savings**
Jan	27	12	42	11	17	33	15	47	8
Feb	33	20	57	13	21	41	16	54	14
Mar	29	24	37	14	19	37	16	50	10
Apr	22	17	30	14	18	28	12	42	6
May	19	14	21	12	15	20	10	36	4
Jun	18	14	18	13	15	20	9	35	3
Jul	17	11	14	12	14	19	7	33	1
Aug	28	14	32	16	22	37	15	49	9
Sept	30	21	32	15	21	34	15	51	11
Oct	39	20	69	17	27	49	20	60	20
Nov	37	16	84	14	24	49	18	57	17
Dec	37	17	50	13	23	45	17	57	17
Yearly Total	1,625	930	2,162	745	1,202	2,000	806	2,520	677

Results are similarly sensitive to support staff and vessel leasing costs. In the scenario where no vessel savings were achieved, there is a significant reduction in the yearly savings. At lower prices of energy, employing a night shift where no savings in vessel leasing costs could be assumed could become completely unfeasible for low income months such as June and July. Similar conclusions can be drawn in the more costly case where staff are paid double time for night shifts (the column (x2)). The effects of higher staff wages are mitigated, however, by optimising for the most profitable number of technicians per night shift. Where technicians are paid the same as they would be on day shifts (x1), savings are most significant. Providing no incentive for night shift work would be bad practice for businesses considering the disruption to life it causes staff.

The final sensitivity parameter was failure rate, which was explored by extending the model to include another level in the hierarchy. The 'Low' and 'High' categories are explored in columns 3 and 4 of table 6.3. There is some variability in how this effects results month-to-month. In the high wind speed months of January and February, the disparity is significant between high and low failure rate turbines. Likewise in the months October, November and December. For the yearly simulated data, the difference in the means was substantial. Savings in the summer months are more consistent. Turbines in the 'low' failure rate category see just over half the savings as the high. Figures 6.6 (a) and (b) explore how failure rate categories effect uncertainty. Introducing night shifts has a more significant effect on higher failures. This is down not only to a reduction in the mean expected value, but also to a reduction in the likelihood of high values of opportunity cost. These plots show how including additional parameters such as failure rate can lead to a more thorough uncertainty quantification. The uncertainty for the 'High' failure category is far significantly greater than that observed in figure 6.5.

6.2.3.3 Availability Savings

Results of the baseline scenario for availability are shown in figure 6.7 and columns 1 and 5 of table 6.4. Figure 6.5 shows the weekly availability posterior density throughout the year for 3 of the different scenarios considered: where there are no night shifts in operation, and where there are night shifts in operation with either 6 or 12



(a) (b)

Figure 6.6: Box-plots exploring the effect of night shift baseline scenario in (a) July and (b) October.

technicians. The 'Night Shift' scenario PDF shifts the most likely values for each month to higher availabilities. In contrast to the lost production model, the benefit of night shifts on availability is more consistent throughout the year. Repairs during the favourable access conditions of the summer are preferable to repairs in the winter. There is a drop going from June and July into August and September which is surprising. This may have to do with a closer focus on scheduled maintenance in the winter months, which puts off any corrective actions. Such a speculation might be supported by the results of the second Bayesian case study, which show that failures are general much more likely to occur immediately after annual service campaigns.

The first half of table 6.4 explores mean availability savings further. Differences in availability due to the number of technicians on night shifts is explored. On top of this, the percentage of weeks where the availability exceeds 98% is calculated. This reflects the structure of contracts fer turbines which are still under warranty, where OEMs must meet a contracted availability. Employing just one team of 6 technicians itself has an appreciable effect on availability, ranging between 0.36 per week in January to 0.57 in August, and a yearly mean difference of 0.48. The scenario likewise has a significant effect on yearly percentages of an arranged lower availability estimate, which are met 93% more of the time than the non-night shift scenario As expected, more technicians on night shift translates to better availabilities. Employ-

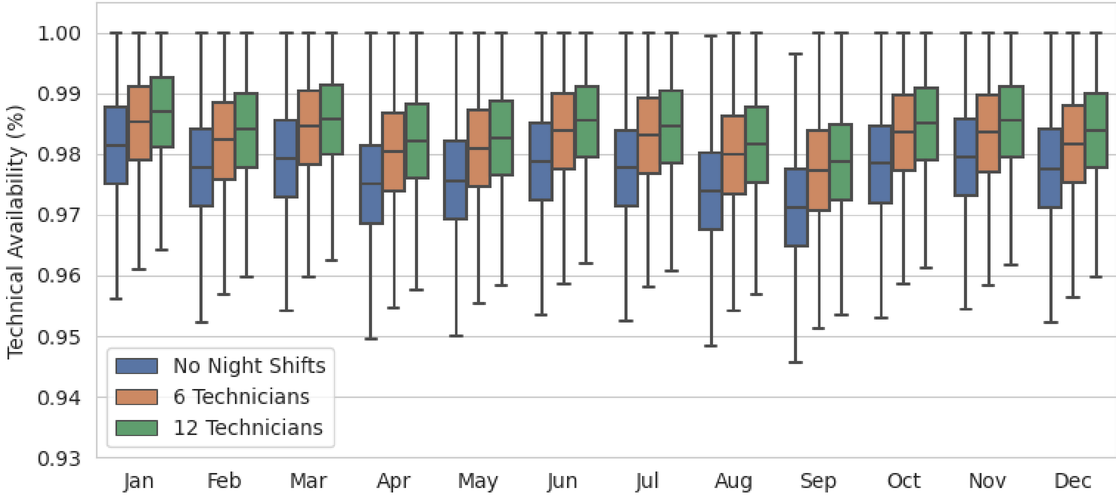


Figure 6.7: Seasonal variation of predicted availability in the scenarios where night shift are in operation with 12 technicians (green), in operation with 6 technicians (orange) and out of operation (blue).

ing 12 technicians means that the farm could expect to surpass a yearly time-based technical availability of 99 times out of a hundred. Such a guarantee could have a significant effect on the negotiation of maintenance contracts between operators and OEMs.

6.2.4 Night Shift Analysis - Discussion

6.2.4.1 Operational Risk

Offshore wind turbine maintenance is a balancing act. On the one hand, operators try to lower their cost of energy by reducing direct O&M costs and maintaining a high power output. On the other, they must take care to ensure the safety and well-being of their staff. The question of introducing night shifts exemplifies this risk-profit trade-off well, as the potential financial boost comes with it potential safety implications to consider for technicians.

Consideration of these safety implications for night shifts on OWFs is hindered by a lack of scrutiny by H&S experts, as it is currently an under researched area. However, there are a few factors characterising the current state of the offshore wind industry which may stimulate interest in H&S for 24/7 working. namely:

Table 6.4: Table exploring night shift scenarios on availability modelling outputs. Figures in bold represent months where the average weekly technical availability difference exceeded 0.5%.

	Mean Availability Difference (%)		
(All fr Quantiles)	**12 techs**	**9 techs**	**6 techs**
January	**0.53**	0.36	0.36
February	**0.59**	0.43	0.43
March	**0.61**	**0.56**	0.48
April	**0.70**	**0.64**	0.53
May	**0.68**	**0.63**	**0.51**
June	**0.64**	**0.59**	0.48
July	**0.67**	**0.61**	**0.52**
August	**0.74**	**0.69**	**0.57**
September	**0.74**	**0.71**	**0.61**
October	**0.64**	**0.57**	0.49
November	**0.56**	0.40	0.39
December	**0.60**	0.40	0.40
Yearly Average	**0.64**	**0.55**	0.48
(fr = High)			
January	0.44	0.28	0.27
February	**0.76**	**0.61**	**0.58**
March	**0.54**	0.47	0.42
April	**0.70**	**0.65**	**0.56**
May	**0.74**	**0.67**	**0.56**
June	**0.53**	0.48	0.34
July	**0.58**	0.53	0.44
August	**0.88**	**0.82**	**0.72**
September	**0.95**	**0.86**	**0.77**
October	0.49	0.44	0.42
November	**0.70**	**0.53**	0.47
December	**0.74**	**0.59**	**0.53**
Yearly Average	**0.68**	**0.56**	0.50

1. Where safety regulations have typically restricted operational practices to be conducted in the daylight (as suggested by Dalgic et al. [169]), there is evidence of at least one large operational OWF moving beyond this regulatory regime.

2. Where early wind farm installations are situated in the most convenient locations closest to shore, the industry will need to explore further offshore sites if it is to meet the ambitious targets it has set. This means that operators may have to look for new ways to improve accessibility.

3. The potential advent and use of so-called flotels, mothership concepts, Service Operation Vessels (SOVs) and walk-to-work schemes for O&M activities for far-offshore wind farms can facilitate 24/7 working with shorter distances to travel, potentially making it more appealing to operators.

4. It can be argued that the competitive nature of CfD-like auctions which are proliferating many European countries have driven the cost reduction in offshore wind. This culture might be assumed to continue into the near future and offshore wind farms will continue to adopt money-saving measures to get a leg-up on the competition. This argument becomes especially pertinent when offshore wind farms transition to post-subsidy financing models.

Publications in academia, or indeed industrial technical reports, about the H&S implications of night shifts for offshore wind are sparse. Garrido et al. [283] come the closest by examining the effect of shift schedules on sleep patterns of offshore wind farm workers in Germany via an online questionnaire. Surprisingly, they found no effect of night shift working on sleep disorders. To the authors' best knowledge, this is the only publication to explore the subject for offshore wind, however there are publications focusing on 24/7 working in the oil & gas industry and offshore working in general which may provide a more comprehensive frame of reference.

A notable review study [284] of offshore workers in the petroleum industry found that more sleep problems were consistently reported by night workers compared to day workers, with workers generally adapting to night shifts within one to two weeks' work.

Concerning accidents occurring during night shift, it reported inconsistent findings. Two studies recorded higher H&S incidents rates during night than day [285, 286], while another [287] reported no such disparity. Interestingly, effects of reaction times (a good indicator of job performance) due to night shifts were consistently negligible across four studies [288, 289, 290, 291]. Reaction times only suffered during the first day transitioning to night shift in a study by Waage et al. [289], and the first day transitioning back to day shift by Bjorvatn et al. [288]. A publication from the Health & Safety Executive in the UK [292] encourages a 14D/14N shift pattern for offshore oil platforms "*unless there are very strong reasons why it cannot be implemented in particular circumstances*", in order to allow for readjustment of circadian rhythm for workers. While in this case some of the stresses of the night shift pattern will be mitigated by the fact that technicians will be sleeping on land, these studies might prove informative for the offshore wind industry.

6.2.4.2 Further Discussion

The new findings of this chapter consist of a real-world case study of night shift working. The proposed methodology, Bayesian hierarchical modelling, is novel in the sense that it has not previously been applied to failure, work order and SCADA data to aid decision making at an OWF. As a result of the methodology, uncertainty quantification for various night shift scenarios have been presented, allowing for differences in the probability distributions of various operational contexts to be explored. This is what makes the proposed methodology well suited to the analysis of operational data. Given the input data is collated on a weekly basis, there are too few samples in each month to build statistically robust inferences. A combination of informative priors and sharing of statistical strength across categories directly addresses the issue. Further novelty is provided over previous studies in the field via three means. There is something to be said of each of these in terms of benefits afforded and limitations created.

First, a night shift strategy which is employed by a currently operational offshore wind farm is assessed. It therefore supports a data-driven model, which provides real world evidence, over a simulation model. To the authors best knowledge, this is the first time such real world evidence has been presented in the literature. A disadvantage of the current methodology is that the case study is restricted to the data describing this one wind farm - the results cannot be generalised. As a result, it is difficult to compare results with previous studies, although this would be desirable. A rudimentary attempt can be made by comparing the baseline results arrived at here to those of arrived at by Besnard et al. [168]. Subtracting scenario 8 from 5 & dividing difference by the installed capacity used, we arrive at a figure of £2,000/MW installed capacity to compare to our £1,625/MW installed capacity. . Their conclusion that each "the availability increases by almost 1% for each logistic solution by using 24/7 work shifts instead of 12/7 work shifts" is somewhat supported by this analysis, where the baseline scenario presents 0.64% difference. Due to confidentiality reasons, the results cannot be compared to the generic figure given by Poulsen et al. [219] of 1.8M€per year. If one considers take hypothetical wind farms of 200MW, 500MW and 1GW, respective savings of £325,000, £812,500 €, £1,625,000 result.

Second, a sensitivity analysis to explore the impact of modelling assumptions on results is undertaken. While the baseline scenario shows reasonable savings for the strategy explored, these are quickly negated by increasing support staff wages and in the case where savings in vessel charters cannot be assumed. These will grow more substantial with reduced prices of electricity or in the case of turbines with low failure rates. Each wind farm will have a unique set of circumstances to contend with, so these considerations are significant.

Third, evidence is provided of the benefit of Bayesian hierarchical modelling in retroactive analysis of offshore wind farm operational data. A review of Bayesian methods in the wind industry [218] has previously suggested that this could be an advantageous avenue to explore in the wider field of Bayesian analysis. Given there is *some* requisite data available to inform the model, the methodology provides a simple and computationally undemanding approach to draw inferences (and value) form operational data. For this specific question of night shifts, the approach was useful. However, more sophisticated model parameterisations will be needed to explore complex models with more variables. This dependence on prior information is another potential limitation of the approach. If that prior information is not accurate, it means that an inaccurate bias is effectively being introduced into the results which could make for dubious inferences. In the case where extreme values are derived from the cost model, for instance, the value to be drawn from the data would be negated.

Fourth, a significant drawback from the data model is that safety implications are not considered. As is the case with many OWF datasets [70], incident reporting or HSE data was not available to integrate into the current analysis. This offers a potential for future work. The current model could be expanded to explicitly capture the trade-off between lowering costs and increasing safety risk. This could be done by setting up a Bayesian hierarchical model with a similar structure to the one defined in this chapter, but with "number of incidents" or something similar as the relevant KPI.

6.3 Bayesian Reliability Analysis

This section presents a case study based on the methodology developed in section section 5.3, based on the rationale presented in section section 2.6. The methodology is used to explore the effect of scheduled maintenance works on wind turbine time-to-failure.

In contrast to the case study presented above, where the subject of the analysis was modelling the relationships and variability of hierarchical data-types, here the subject is to understand the time it takes for an event to occur and how various factors influence this time. The time varying effect of scheduled maintenance on time-to-failure is a novel consideration for reliability models of wind turbines.

6.3.1 Model Specification

This case study follows the methodology developed in section section 5.3.

6.3.1.1 Selected covariates

The key contribution of this chapter in in quantifying the effect of operational covariates on WT reliability. The time-varying effect of scheduled maintenance works on failure rates is a particularly novel consideration. The full list of covariates considered is as follows:

1. **Seasonality**. Simalarly to Slimacek and Lindqvist [137], Seasonality is employed as a model covariate. This is either coded by month of the year, (i.e. 'Jan', 'Feb'...) such that each month has a constant covariate effect with respect to a reference month (which is arbitrarily chosen to be April), or season of the year under the same assumption (reference month is arbitrarily Autumn).

2. **Year of Operation**. Year of operation is of interest because WT failure behaviour is often assumed to vary over time, most frequently by a power law process. Failure intensity is often assumed to follow the bathtub curve [293], which includes a wear-in period and wear-out period. One study by Stiesdal et al. also included a serial defect period for Siemens turbines [222]. It is assumed that the data covers the normal operations period in the wind farms life, however there are few results in the literature to conform that this is actually flat. Also of interest is the fact that the site deployed their advanced data management system in 2018, which in itself may have improved operational efficiency.

3. **Turbine Location**. Turbine location may be assumed to effect reliability primarily due to the effects of turbulence. This effect is included via a frailty term, either using the turbine row or individual turbine names as categorical variables.

4. **Time Since Annual Services**. It is common practice in the industry to have an annual campaign where a set of scheduled maintenance actions are carried out on turbines. The exact nature of these actions vary from one service provider to the next and depend largely on maintenance contract arrangements

[161]. This generally includes tasks such as lubrication of mechanical parts (e.g. gear oil, hydraulic oil, greasing), measurement of part temperatures, a torque tensioning of bolts and basic inspection of parts within the nacelle. Analysis of the effect of annual services on failure rates was a key motivation for introducing time-dependent effects, as the data provider for this study reported allegorical evidence of repeated failures after annual services from technicians.

5. **Time Since Inspection**. Likewise, it is common to regularly inspect certain key components - for example a visual inspection of blades for cracks or erosion. Again there is an assumed time-dependence for this variable; it is conceivable that a repair is more likely to take place soon after an inspection, but thereafter the risk of failure might be assumed to decrease.

6.3.1.2 Time-to-event Dataset

The *Downtime Catalogue* acts as the basis of this study. However, it requires an additional transformation step in order to convert the information into a data format that is compatible with the methodology presented in section section 5.3. This is referred to as the time-to-event data-set, a subset of which is shown in table 6.5. Most important are the columns **tgap** and **failure**, upon which the model fundamentally depend. **tgap** represents the time since the last maintenance action, i.e. the difference between **tstart** (the time stamp at which the turbine is restored to fully operational after the previous maintenance action) and **tstop** (time time stamp at which the turbine stops producing power due to some maintenance intervention). The rest of the columns are covariate values. Time is recorded in units of days.

6.3.2 Model Comparison Results

Before the effect of covariates can be explored, the model selection procedure outlined in section section 5.3.5 is used to explore the effect of variable inclusion/exclusion on predictive power. Table 6.6 summarises the results of the step-wise model comparison procedure. The table summarises the effect of each covariate as they are

Table 6.5: First five rows of the time-to-event dataset. Here **ID** is the event ID, **Turbine** is the turbine where the event took place, **tstart** is the time of when the last maintenance intervention took place, **tstop** is the time where the event occured, **AS** defines whether the previous intervention was an annual service, **Inspection** defines whether the previous intervention was an inspection, **failure** specifies that the event was a failure, **month** is the month the event took place, **year** is the year the event took place and **season** is the season the event took place.

ID	Turbine	tstart	tstop	tgap	AS	inspection	failure	month	year	season
1	A1	0	15	15	0	0	1	7	2018	summer
2	A1	29	49	20	0	1	1	8	2018	summer
3	A1	50	87	37	0	0	1	9	2018	autumn
4	A1	99	115	16	1	0	1	10	2018	autumn
5	A1	115	117	2	0	0	1	10	2018	autumn

introduces. Note that, during the backwards selection stage, some of these variables may change. Immediately from the starting model definition it can be seen that a Weibull model significantly outperforms the exponential model in representing the time-to-event dataset. From visual inspection of figure 6.8, it is evident that the Weibull model also matches more accurately the non-parametric Kaplan-Meier estimate [294]. This is in itself an interesting result. According to Seyr & Muskulus [14], WT failures are most commonly modelled exponentially via a Poisson Process. The results in figure 6.8 imply that a Weibull distributed time-to-failure would be more accurate at the turbine level, since the Wiebull model matches the Kaplan Meier estimate more accurately than the exponential. This means that that failures are more likely to occur immediately after a repair is conducted. Scheu et al. [188] conclude that the difference in modelled availability between exponential and Weibull distributed failures is as much as 10%. This distributional assumption can therefore have a significant impact on the outputs of cost models - it may benefit the research community to investigate this disparity further.

Seasonality has a lesser, though still significant effect on modelling performance, as does year of operation. Time since annual service also improves model performance - the degree to which it does so depends on the knot vector. There is a quite significant improvement when adding 5 internal knots (compared to 0), but the improvement quickly tails off as additional knots are added. The peak value is at 7, after which

Table 6.6: Summary of covariate effects on WT reliability. TVE stands for time-varying effect. No covariate signifies the previously best performing model.

Modelling Group	**Added Variable**	**Modelling Alternative**	**LOO-score (ELPD)**	**LOO-score (SE)**
1	Baseline	Weibull	0	0
		Exponential	-588.3	32.8
2	Seasonality	Seasons	0	0
		Monthly	-11.9	6.4
		No Covariate	-38.5	10.2
3	Year of Operation	Yearly (Constant effect)	0	0
		Yearly (year-by-year)	-2.0	0.9
		No Covaiates	-47.6	9.3
4	Time Since Annual Service	TVE (7 Internal knots)	0	0
		TVE (8 Internal knots)	-0.4	0.3
		TVE (9 internal knots)	-0.6	0.4
		TVE (10 Internal knots)	-1.1	0.5
		TVE (6 Internal knots)	-4.1	0.7
		TVE (5 Internal knots)	-4.6	1.1
		No Covariate	-28.1	5.9
5	Time Since Inspection	TVE (5 Internal knots)	0	0
		TVE (6 Internal knots)	-1.3	0.3
		TVE (7 internal knots)	-1.4	0.4
		TVE (4 Internal knots)	-3.4	0.8
		TVE (2 Internal knots)	-9.2	1.9
		No Covariate	-18.7	5.1

there is a very small depreciation in model performance. Given that all internal knot configurations containing greater than seven have a small $ELPD_{LOO}$ (equation 5.32) score compared to the model error (SE), they have similar predictive power and the least complex model should be chosen. Modelling both time since annual servicing and time since inspection provides the most accurate results. Time since inspection is most accurately modelled with 5 internal knots - perhaps this is fewer than time since annual service since it is characterised by a less dramatic spike in the first few weeks after performing the task. The Effects of increasing number of internal knot locations is shown in figure 6.9.

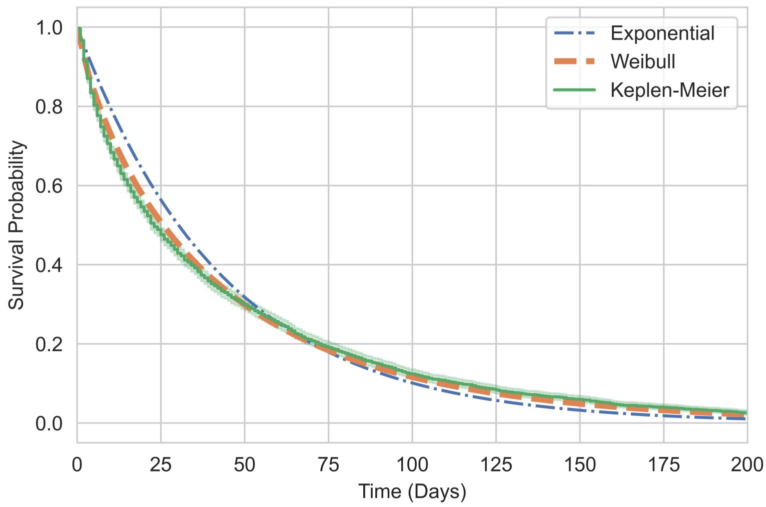


Figure 6.8: Survival Curves for both baseline hazard rate distributions. For reference the non-parametric Kaplan-Meier curve is also shown.

6.3.3 Covariate Values

The effect of seasonality is shown in figure 6.10. This most likely indicative of the strategy employed at the sight rather than any wind speed effects. The operator favours the low wind speed months of June and July to perform repairs. This corroborates the seasonality pattern of repair rate presented by the SPARTA initiative [81]. It does not align with the results presented by Slimacek and Lindqvist [137], who observe a similar uptick in repairs during summer, but increased failure rates throughout winter. The disparity may be indicative of improving operational efficiency with growing experience in the industry. The high hazard rate in Autumn is unexpected given the patterns previously presented in the literature. *However, Fischer et al. [295] present evidence for high failure rates for Scandanavian power converters in August and in Autumn. The results of section section 4.5.2 show that a significant proportion of failures in this population is due to the converter. This high Autumnal failure rate may therefore be attributable to a higher-than-average proportion of converter failures.*. It is unclear exactly what causes this - it has been suggested that there may be a 'fatigue effect' after the busy summer months, but this is difficult to verify.

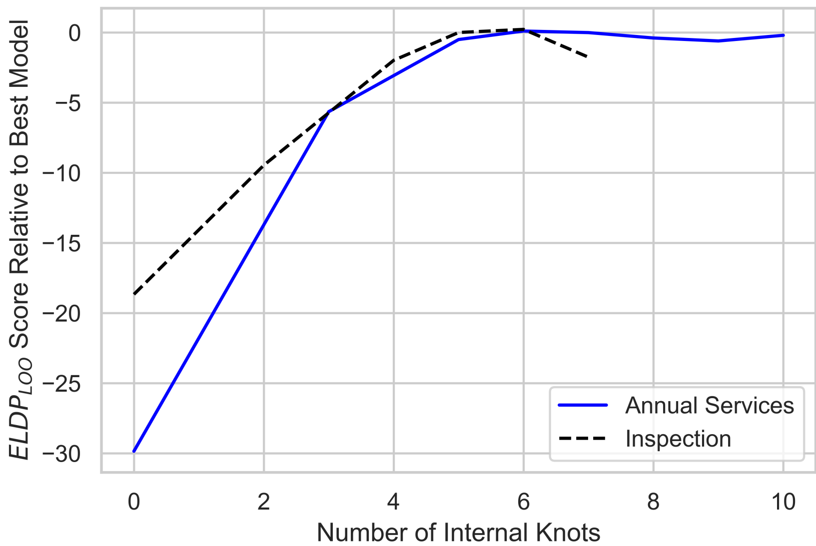


Figure 6.9: Cubic fitted curve estimations of $ELPD_{LOO}$ values for models of varying internal knot locations.

The effect of year of operation is shown in figure 6.11. The effect on the hazard rate is with reference to the first year for which data is available, 2018. There is a consistent and quite significant reduction of the hazard rate through time. Interestingly, this coincides with the introduction of an advanced data management system to the site in 2018. It may therefore present evidence of the value of data management systems regarding operational efficiency.

Frailty effects are explored in figure 6.12. Figure 6.12 (a) shows how the means of all monte-carlo simulations for each turbine are distributed - approximately normally around a median of 1. There are quite significant deviations from the mean. The most unreliable turbine is predicted to fail 1.65 times more that average, the most reliable 0.5 times the average. This is quite a significant heterogeneity in failure intensity that is never taken into account in costs models. Again, further research into how this effects cost modelling outputs may be useful for researchers in the field. Figure 6.12 (b) shows how all of the frailty estimates (i.e. from every monte-carlo simulation) are distributed, which is useful for checking the assumption that frailties are indeed gamma distributed. Evidently the random effect estimates can be

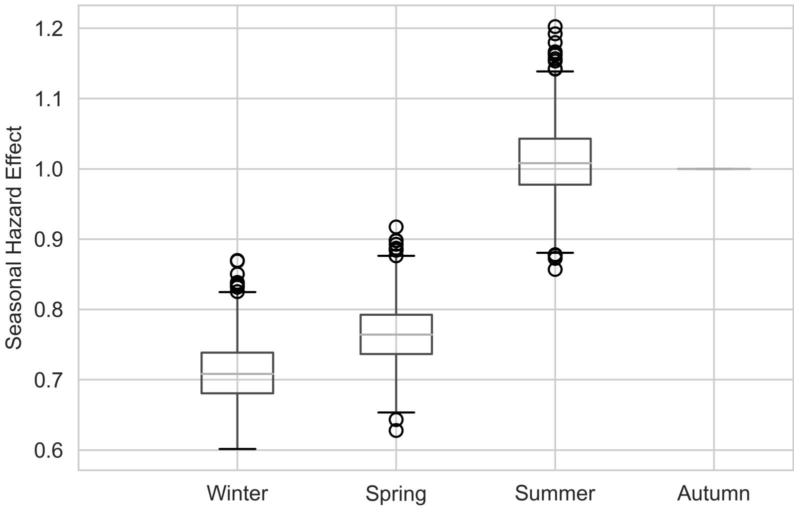


Figure 6.10: Effect of year of seasonality on WT hazard rate, relative to Autumn.

accurately approximated by a gamma distribution (5.30), as the fitted gamma curve fits the histogram very well.

Frailties are also useful for exploring the effect of position in the array. Figure 6.13 explores the effect that turbine row has on frailty values. It does so by fitting a cubic spline function to the mean frailty values of turbines in each row, as well as the 25^{th} and 75^{th} percentile values for each row. Values towards the left of the plot on the x-axis represent rows which are towards the front of the array with respect to the prominent wind direction. The first row is characterised by a relatively low failure intensity and relatively low uncertainty in the failure intensity. The next foremost rows are characterised by both an increasing hazard rate and increasing uncertainty in the hazard rate. This trend continues to a point in the middle of the array. After this, successive rows retain high uncertainty in hazard rates, with mean values decreasing towards the back of the array. The observed trend is expected - turbines downwind of others are subject to higher turbulence, which has been shown to be detrimental to WT reliability [296]. At the same time downwind turbines operate in lower speeds, a decreasing median trend is observed from the rows in the middle to the rows at the back. Interestingly these 'middle' rows are characterised by the highest heterogeneity in reliability. This has an interesting potential application in wind farm design for

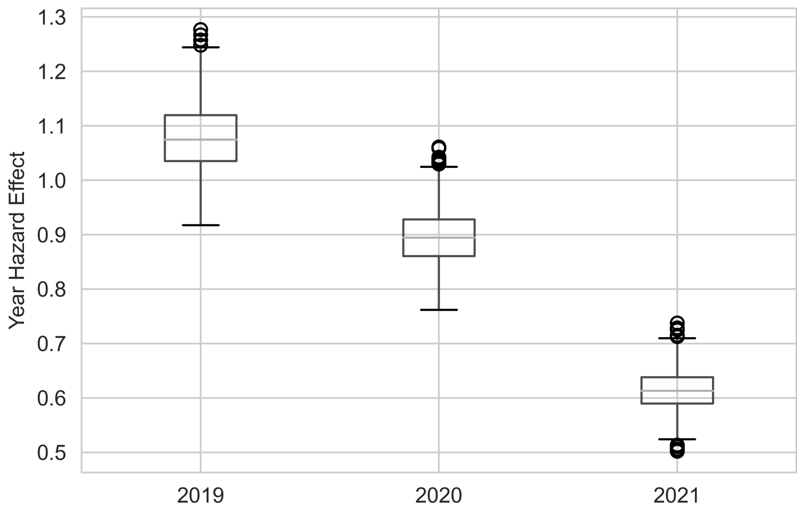
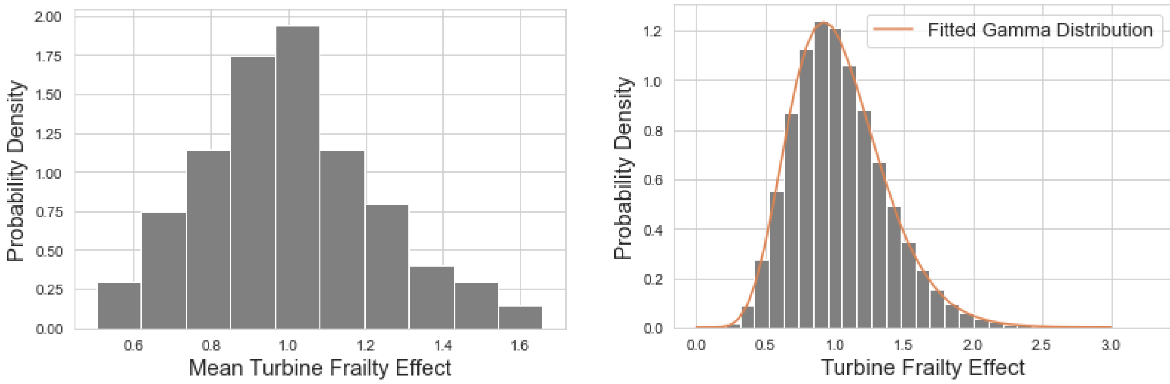


Figure 6.11: Effect of year of operation on WT hazard rate, relative to the year 2018. Since data-collection commenced in 2018, there has been a consistent downward trend.

reliability. If "middle" row failure intensity is higher, it would discourage designs of wind farms in which turbines are constantly operating in the wake of others. Further investigation into this phenomenon, and how it varies with varying turbine/farm size and inter-array spacing, would be an interesting line of further enquiry.

Figure 6.14 shows the time-dependent covariate *time since annual service*. It is characterised by a sharp initial upward trend peaking at 2 days. The maximum value for mean Hazard ratio estimate at this point is 1.57. Turbine failure intensity is higher for the first 6 days after the service, after which there is a reduced failure intensity until 156 days after. Beyond this point, failure intensity is estimated to increase, as might be expected for turbines far away from their annual service. Lower values along the time axis have lower uncertainty in their estimate as the majority of turbines will most likely require corrective maintenance again in the next few weeks than in the next few months. Hence, estimates further along the time-axis are more uncertain. The results back up the theory of the operational team - initially annual services can be thought to lead to more failures, after which they prevent failures as they are supposed to. However, the effect is short-lasting, and the higher failure intensity may not justify any change in strategy.



(a) (b)

Figure 6.12: Histograms of (a) the mean frailty effect for each turbine and (b) the turbine frailty effect over all monte-carlo simulations. The assumption of gamma-distributed random effects is a reasonable one.

Figure 6.15 shows the time-dependent covariate *time since inspection*. The shape is similar to *time since annual service*: there is an initial peak in failure intensity (where inspections report a problem leading to a corrective maintenance action), after which the hazard rate is below 1. Again the initial peak is at 2 days after inspection, with a slightly smaller hazard ratio of 1.30. 6 days after servicing, the hazard rate falls below 1. It remains below 1 throughout the entire year, rising gradually as time goes on.

6.3.4 Discussion

There are two ways to consider the utility of this study. The first has to do with the context and quality of the input data, the second to do with the methodology itself. Regarding the input data: the data-table *Downtime Catalogue* results from an operational database provided by a currently operational OWF which employs an advanced data management system. Given the scarcity of reliability data available for offshore wind turbines (see Reder et al. [26] for details), this is valuable to the research community. However, the dataset is limited in that it is not broken down into a component or subcomponent taxonomy. This means that no conclusions can be made on the time-to-failure of individual components, the effect of seasonality on individual components or the effect of year of operation on individual components.

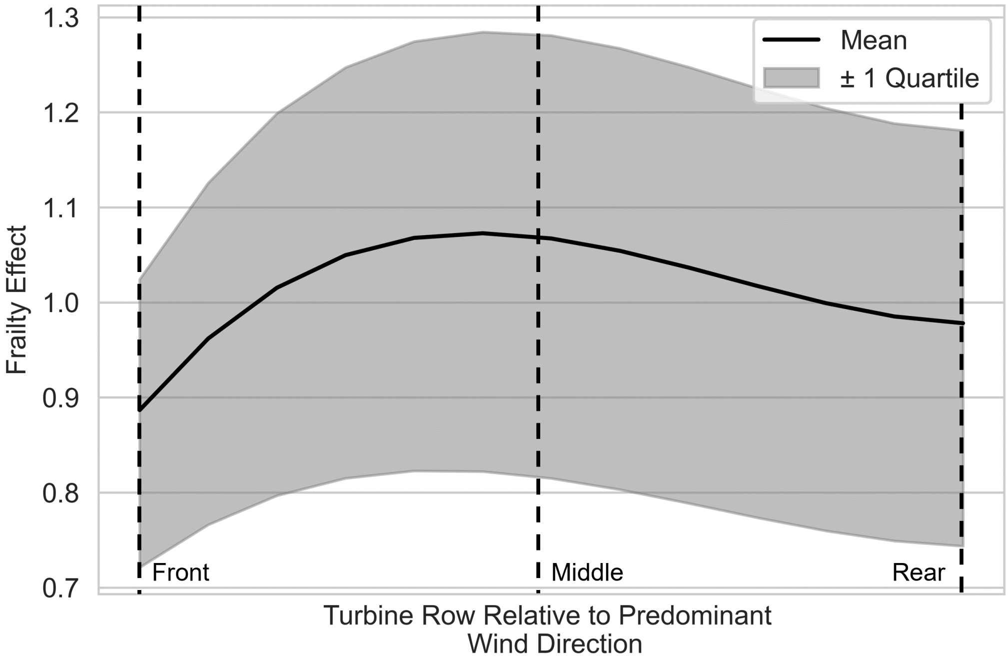


Figure 6.13: Fitted cubic function approximating frailty effects' relationship to turbine position in the array with respect to the prominent wind direction.

Since some modern cost modelling tools consider reliability at the component level, such an analysis would be useful to the research community. This is especially pertinent for the time varying effect of inspections. Figure 6.15 does not reveal anything particularly useful - as expected, the inspection reveals the need for corrective action and so there are likely to be corrective actions following. However, future work plotting figure 6.15 on a component by component basis with a view to seeing how inspections of a given component correlate to that component's failure rate would be more valuable. It may also be of value to consider the time varying effect of annual services on specific components, such that operators can consider changes in how scheduled maintenance is carried out. That being said, the approach used here can be easily adapted to a subcomponent taxonomy, or indeed can be applied to repairable systems other than wind turbines. Also, uncovering this relationship at the turbine level is also valuable, as it has previously not been formally quantified.

Regarding the methodology itself: the alterations proposed to the traditional Non-Homogeneous Poisson Process (NHPP) are novel in the context of WT reliability.

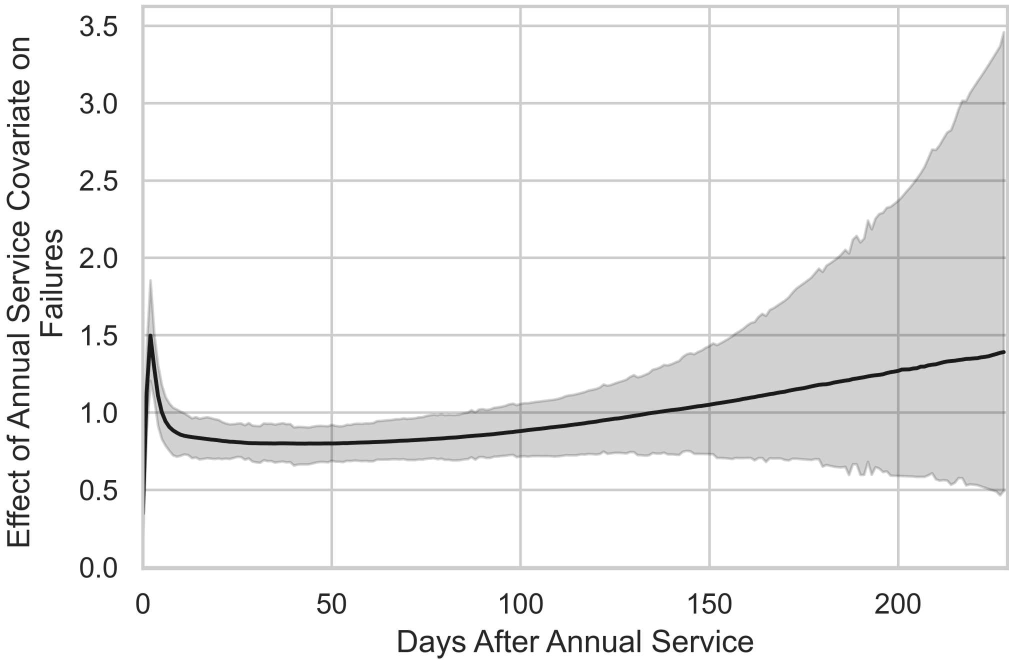


Figure 6.14: Time-dependent relationship between annual servicing and wind turbine failure rate. A hazard ratio < 1 siginifies an increased failure rate, < 1 decreased.

Given the growing recognition in the wind research community that more care must be taken uncertainty handling, extending the NHPP into the Bayesian regime is a natural extension. The advantages of implicit uncertainty handling by Bayesian models are well documented and often cited [37]. Interpreting uncertainty via Bayesian interval estimates, which are interpreted as the interval containing the true parameter with some probability, are generally seen as more intuitive than confidence intervals, which are interpreted as the range of values containing the true parameter a certain percentage of the time given a large sample approximation.

Inclusion of time dependent variables is the second extension. In this case it proved to be a useful one, as it allowed us to explore the effect of varying failure intensity proceeding scheduled maintenance works which was posited by the operational team. There is one important caveat in employing time-dependent variables, though. Namely, the selection of knot locations in B-Splines takes care, as different assumptions might lead to quite significantly different results. The parsimonious model selection methodology used was relatively computationally expensive, considering the

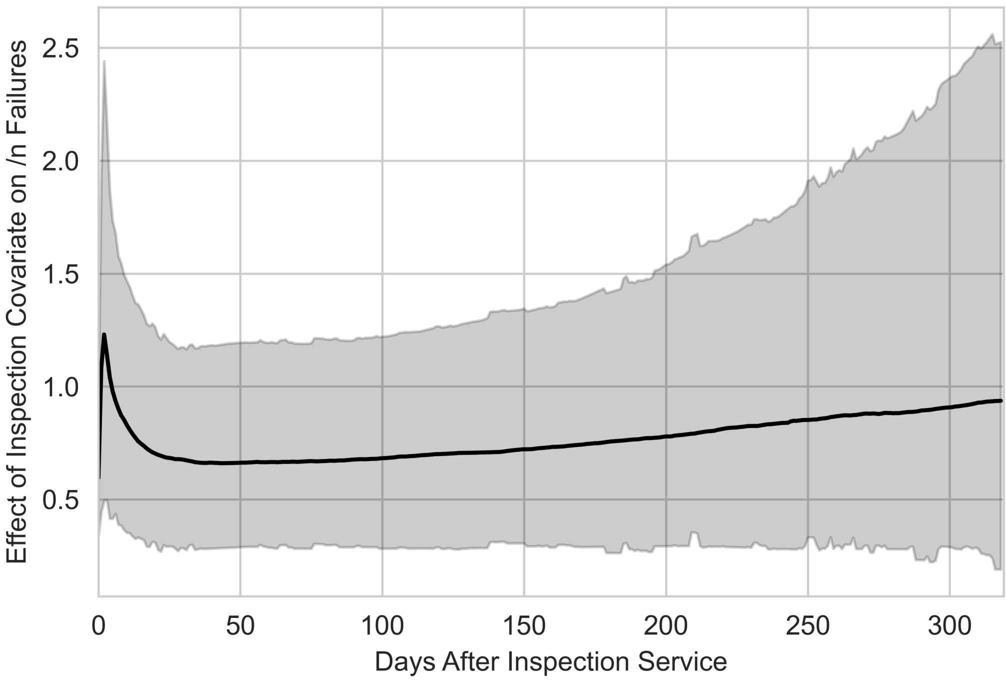


Figure 6.15: Time-dependent relationship between inspection and wind turbine failure rate. A hazard ratio < 1 signifies an increased failure rate, < 1 decreased.

usual run-time of survival models.

The final methodological point has to do with failure definition. As noted by Leahy et al. [25], there is no standard definition of a failure in the wind industry. The choice of failure definition often depends on the data available to the researcher. Here, downtime is used as a qualifier for a failure, when that downtime coincided with a maintenance action recorded as corrective in the database. In the absence of a standard failure definition in the industry, this choice of failure definition becomes a model hyper-parameter in itself which may significantly affect the results. Of particular importance is the timing of failure onset, which is not necessarily the onset of downtime recorded here (especially for minor repairs). However, for many operational datasets it would be difficult to decipher any failure on-set time beyond the onset of downtime.

6.4 Case Study Applicability

The above case studies presents an application of the Bayesian models developed in section chapter 5. The rationale for a Bayesian framework is presented in section section 2.5. The goal of applying those Bayesian methodologies was to explore 2 aspects of operations that were of import to the operators of the OWF, as presented in sections section 2.6.2 and section 2.6.3. Given that these analyses are dedicated to analysing the operations at one OWF, it begs the question of the results: *what is specific to this wind farm and what can be extrapolated to the wider O&M research community?*

Regarding the Bayesian hierarchical exploration of night shifts for offshore wind farms, there are several important considerations. First, there results presented in section section 6.2 are specific to this wind farm. As noted by Poulson et al, [297], the set of factors likely to vary the effectiveness of night shifts include:

1. Number of vessels needed for working daytime and night;
2. Night-shift add-on salary for technicians;
3. Night shift add-on to cover on-shore support for monitoring purposes;
4. Months per year working 24/7;
5. Capacity Factor;
6. Time for production stop per WTG as a result of faults/repair time;
7. Number of errors/faults/stoppages per wind turbine per month;
8. Price of electricity during and after subsidy period.

Since it is the combination of these factors that matters, each wind farm would have to consider night shifts on a case-by-case basis. However, table 6.3 at least gives an indication of the impact of factors 2, 3, 4, 7 and 8. For instance it may prove informative to wind farms who are considering implementing night shifts that a combination of low electricity price, high support staff wages, low failure rates and no savings from vessel hiring logistics would make night shifts unfeasible - at least for the summer months. The results might therefore be extrapolated to the O&M research space as a whole. Also, the distributions derived in section section 6.2 could be used in future analyses as prior information which could be updated based on observations from the new wind farm. In this sense, the methodology can also be extrapolated to the wider research community. The notion that prior estimates from O&M models can be dynamically updated as new data become available in a computationally effective way presents an opportunity to improve upon those existing tools [33].

There are also several key points of consideration for the Bayesian reliability analysis presented in section section 6.3. Again, since the data that goes into the model is taken from one wind farm, the results should be regarded with caution. Factors likely to alter the results include:

1. The set of tasks completed during annual servicing:
2. Turbine power rating;
3. Climatic conditions (humidity, wind speed, rain, temperature);
4. Turbine manufacturer;
5. Effectiveness of inspections;
6. Maintenance strategy.

Again, since it is the combination of these factors that matters, each wind farm would need to update these results with their own observations. However, the results exploring the time-dependence of scheduled maintenance works are novel. As such, they can be extrapolated to the wider community as an indicative relationship. Likewise, the methodology which incorporates time-dependent variables for wind turbine reliability analysis might be extrapolated to the wider community.

To summarise, both case studies have similar characteristics in terms of their applicability to the wider research community. The actual results presented in sections section 6.3 and section 6.2. The results should be viewed with caution by other operators, as there are several factors in each case study which might alter the conclusions. However, the results could extrapolate to the wider research community if they were to feed into future analyses via prior distributions. At least, the methodologies and their rationale could be extrapolated to the wider community. Namely, the rationale of dynamically updating prior estimates with new information. Using the Bayesian framework does so with robust uncertainty quantification included. In the case of the Bayesian hierarchical modelling framework, this has the potential to share statistical strength among different operational contexts. In the case of Bayesian reliability analysis, it means incorporation of time-dependent covariates.

6.5 Chapter Conclusion

This work presents an application of the Bayesian methodologies developed in chapter chapter 5. The first method, a Bayesian hierarchical modelling approach was used to assess the cost-benefit of night shifts. The second method, a Bayesian reliability modelling approach, was used to assess the effect of scheduled works on WT reliability.

Section section 6.2 presents an assessment of night shifts for offshore wind farms. A particular strategy is considered where one CTV was re-purposed to carry out corrective maintenance work during a night shift. A hierarchical model was used to model the two operational metrics: opportunity cost and technical time-based availability. The hierarchical nature of the model allowed me to explore the variation of these values throughout the year, and to introduce estimates of lost production and availability from hypothetical alterations in strategy. The Bayesian nature of the model allowed me to make inferences from a relatively small sample size via the use of informative priors derived from the StrathOM-OW tool. The more generic advantage that Bayesian models provide inherent uncertainty handling was also evident in the results. There is something to be said of this feature in itself, as statistical uncertainty in maintainability has been identified in recent studies as both significant and under-investigated for offshore wind turbines. These hypothetical scenarios were in the form of retroactive algorithms applied to the data to assess the benefit of night shifts undertaken at the site. Samples from the posterior of this model provided financial implications of the lost production. Assumptions were made as to the cost of extra staff wages and avoided vessel costs to retrieve estimates for a baseline case when night shifts were in operation and out of operation. Comparing the mean values of yearly simulated data it was found that £1,625 per MW installed capacity could be saved. There are a number of factors which influence this figure, which are explored via a sensitivity analysis. Altering the price of energy reduced the difference in means to £745/MW capacity at £40/MWh. It increased the difference in means to £2,000 at £150/MWh. Support staff wages and assumed savings in vessel charters also had a significant impact. The assumptions made for these variables were rudimentary, and could be explored further. Extending the hierarchical model to include failure rates also proved informative. Those turbines in the lowest failure rate category saw roughly half the savings from those in the high failure rate category, at £745/MW and £2,162/MW installed capacity respectively. Availability savings were more consistent. Maintenance service providers would see a significant improvement in the number of times they would exceed a time-based technical availability of 98%.

Section section 6.3 presents a Bayesian reliability analysis of wind turbines. A model is proposed which incorporates time-dependent variables in response to the operational team's reasoning that turbine failures are more frequent following annual services. This was indeed the observed behaviour, with a higher turbine failure intensity for 6 days after annual service works take place. There is a peak failure intensity reaching 1.57 times the average 2 days after servicing. A similar, though less significant effect was observed with the covariate representing time since inspection. The higher failure rate lasted for 6 days after the maintenance work was carried out, reaching a peak of 1.302 days after. Seasonality effects similar to those presented by the SPARTA initiative were also observed, and a consistent year-on-year reduction of failure intensity. Significant turbine frailties were observed. Plotting these by turbine row, a dependence of reliability on turbine position with respect to the prominent wind direction is observed.

7 Conclusion

7.1 General Conclusions

Based on the problem statement that "*the potential of operational maintenance data is not yet fully leveraged in decision making*", the objective of this thesis was to answer the following research question:

"*How can operational maintenance data be better leveraged to support decision making and therefore reduce O&M costs in the industry?*"

To answer this primary research question several research objectives were considered. This chapter summarises how each of these research questions was addressed in each of the chapters in this thesis.

Chapter chapter 2 presented a literature review. The purpose of the literature review is twofold. First, it presents a review of the wind data ecosystem to identify areas for improvement. Second, it provides a rationale for the objectives of the analyses presented in this thesis. Four themes of improvement were identified to improve data-assisted decision making from the first part of the review:

1. Improved data quality. Improved data quality in reliability data has frequently been suggested as an areas much in need of improvement in the industry [27, 25, 33, 41, 68]. Prominent issues include strict confidentiality practices on reliability data, the lack of standard failure definition or taxonomy, and the generally poor quality of failure data.

2. Uncertainty quantification. Uncertainty quantification is becoming something of a trend in the more recent literature surrounding O&M [46, 45, 43]. In Seyr

et al.'s review of O&M modelling, for instance, a key conclusion is that more effort needs to be put into quantifying uncertainty in O&M modelling inputs.

3. Data Fusion. As mentioned by [29], generally poor data management tools at offshore wind farms means that integrating data-streams often takes significant manual effort from practitioners [29]. Condition monitoring is the area of the literature where there has been the most suggestions for combining various data-streams [15, 32]. However, there is also scope outwith condition monitoring for data fusion, which has not seen much scrutiny so far in the literature.

4. Usability for decision makers. Decision making tools are more likely to be of use to the industry if they are (i) presented in an easily understandable format to decision makers [123] and (ii) relevant to their particular set of requirements.

In the second part of the review, Bayesian models are proposed as a solution to those areas for improvement. They map well to the areas for improvement due to several characteristics:

1. Suitability for small and incomplete datasets. In contrast to almost all ML methodologies, Bayesian models have an inherent suitability for small datasets [260]. This means that robust inferences are possible with datasets for which frequentest approaches are inadequate. In this case that means the generally low-velocity, low-volume reliability and maintenance data.

2. Their inherent uncertainty quantification capabilities. Bayesian models have in-built uncertainty quantification [39], making them a natural solution to the need for uncertainty quantification identified in the literature review.

3. Their ability to incorporate multiple data sources. Bayesian models are able to incorporate knowledge of different accuracies and from different sources in a mathematically coherent way [202], making them a natural solution for the need for data fusion in decision making for OWFs.

Chapter chapter 3 presents a data mining methodology used to calculate maintenance KPIs. Since reliability and maintenance data is so sparse in the literature, this was the most immediate opportunity to address the research question. Gonzalez et al. [53] present a list of KPIs which are suitable for the purpose. The performance and reliability maintenance KPIs used in this thesis are defined in section section 3.2. Since operational maintenance data is understudied in the literature, it was made the focus of this thesis. The data mining methodology therefore used the operational maintenance data as a base for data analysis. As presented in section section 3.3, the operational maintenance data pertinent to this thesis comes from the data-tables *Work Procedures*, *Tasks/Task Types*, *Operations Planned Movements*, *Vessel Stops*, *Operations Shift Tasks*. Taken together, these data tables provide information on when maintenance interventions were carried out, and what kind of work was undertaken when it was. Section section 3.4 describes the process of supplementing this maintenance data with SCADA and turbine property data to create *Downtime Catalogue*. From *Downtime Catalogue*, the failure rates of wind turbine assemblies were calculated, along with the associated downtimes, repair times, number of technicians, number of visits and visit duration. A methodology was also presented to calculate opportunity cost for those downtimes based on a regression of active power between neighbouring turbines. From the surrounding literature [126, 25] it was evident that multiple failure definitions could be defined from downtime catalogue. These were listed in this chapter, so that they could provide the basis of a sensitivity analysis in the subsequent chapter. By defining a data mining methodology which incorporated multiple data-streams, data fusion was exploited as an opportunity to improve data assisted decision making.

Chapter chapter 4 presents a case study of frequentist statistics based on the methodology presented in chapter chapter 3. First, the utility of the *Downtime Catalogue* is demonstrated by comparing tidally-restricted and non-tidally restricted turbines. As shown in section section 4.2, Tidally-restricted turbines are characterised by longer times-to-repair and shorter visit durations for all maintenance categories. This reduction in maintainability leads to a lower availability of tidally-restricted turbines.

These findings imply that any offshore wind site located in shallow water depths may have to adjust their maintenance strategy accordingly.

The rest of the chapter is dedicated to quantifying "*intervention*" for the offshore wind farms. First, the sensitivity of turbine-level failure rates to the different failure definitions defined in chapter chapter 3 is investigated. Results showed that calculated failure rate figures show a lot of variation (from 7.07 to 12.15 failures per turbine per year) depending on the definition used. This range in values shows that there is potential for significant uncertainty in failure rate figures. Work procedures are then categorised by work type (corrective vs. schedules) and mean downtimes and number of interventions per year are presented per work procedure. This initial work procedure analysis showed that the majority of corrective works are due to 'Fault Finding' missions. Fault finding missions, along with other unclear work procedures (hub and blade access, electrical panels, craning of equipment and drive train inspection and torqueing) are examined in more detail such that the work procedures could be attributed to assembly.

Then, the work procedure analysis was mapped to a failure taxonomy as defined by [68]. By considering the different definitions of failure and methodologies for deriving failure rates, the uncertainty in the failure rate figures is explored at the assembly-level. The results show some variation in the derived failure rates. The least reliable systems are the cooling system, frequency converter and hydraulic group. However, reliability of wind turbines may change radically between different manufacturers and with different climate conditions [127].

Chapter chapter 5 presents two Bayesian methodologies which tie in with the objective to explore the utility of Bayesian models for extracting value from the data-set described in section chapter 3. The first method is based on Bayesian hierarchical modelling. Informative priors are provided by an O&M cost model and the parameters of the model are conditioned on a limited dataset via Bayes' rule. The second is a Bayesian reliability analysis developed to capture time-to-failure behaviour for wind turbines by incorporating time-dependent variables. The developed Bayesian methodologies provide the primary advantage of robust uncertainty quantification. Beyond that, the methodologies are intended for different use cases. The first (Bayesian hierarchical modelling) can be utilised to update prior beliefs about strategy changes as new information becomes available. The variability between different operational scenarios can be assessed in this way. The second (Bayesian reliability modelling) is intended for time-to-event dataset, and is useful for reliability modelling.

In chapter chapter 6, the utility of the developed methodologies are demonstrated via two case studies. The first used Bayesian hierarchical analysis to explore the question of night shifts, the second used Bayesian reliability analysis to explore the effect of annual services on wind turbine reliability. In section section 6.2, baseline results show that night shifts could increase technical availability by 0.64% and potentially increase revenue by £1,625/MW installed capacity. These numbers were shown to vary substantially depending on the modelling assumptions, and whether or not to employ night shifts is a question that needs considered on a site-to-site basis. Nevertheless, there are several novel points that can be extrapolated from the case study to the wider research community. First, the methodology is useful in the low-velocity, low-volume context that characterises offshore wind maintenance analysis. The case study shows how prior estimates from O&M models can be dynamically updated as new data becomes available to aid decision making. Second (and as an extension to the first point), the distributions derived in section section 6.2 can be used as prior estimates for other wind farms. Third and finally, the results show how the benefit of night shifts contract or expand with factors such as number of technicians working night shifts, electricity price, support staff wages, failure rates and vessel hiring logistics.

In section section 6.2.4, the Bayesian reliability analysis with time-dependent variables showed that WT failure intensity reaches 1.57 times the baseline in the six days directly proceeding annual servicing. WT failure intensity is lower than the baseline from 6 days to 137 days after annual servicing. Also observed were a significant year-on-year reduction of failure intensity since the introduction of the site's data management system in 2018, a clear preference for modelling time-to-failure via a Weibull distribution and a dependence on location in the array with respect to the prominent wind direction. While these results would likely vary on a farm-by-farm (or indeed operator-by-operator) basis, they might extrapolated to the wider community as an indicative relationship. Likewise, the methodology which incorporates time-dependent variables for wind turbine reliability analysis is novel, and might be adopted by future wind farms with more advanced data management systems.

7.2 Limitations & Future Work

There are limitations to all of the analyses presented in this thesis. Consequently, they might be improved upon by future works. To elaborate:

- Generally, there is one immediate limitation which affects all analyses. Namely, the data is all from one wind farm. Several potentially important features are therefore left out of the analyses. The failure rates presented in chapter chapter 4 are derived from a single farm and a single manufacturer. However, failure rates are likely to change depending on manufacturer [298] and meteorological conditions [217]. The night shift analysis presented in section chapter 6 would likely have different results depending on accessibility, maintenance strategy and power rating. The results of section section 6.3 are also likely to change depending on wind turbine concept and maintenance provider. A useful extension to all of the analyses presented in this thesis would therefore be to include data from multiple sites and WT technologies. If this were the case the methodologies developed in chapter chapter 5 could be expanded to include more covariates and more complex Bayesian network models. This would enhance the benefits of the Bayesian methodologies proposed. If, for example, a new wind farm were to start operating based on a strategy of the group-level reliability statistics of multiple farms, it could then update these estimates with its own data within a Bayesian framework. This would imply increased transferability of old data to new farms. The implication for this study are largely outlined in section

section 6.4. Results could be taken as indicative and are useful as prior information. However, updated information from the wind farm in question would improve the accuracy of inferences.

- Chapter chapter 4 is largely an exploration of the uncertainty surrounding failure rate estimates. There are several future works which could provide value to the research community. Most immediately, value could also be extracted by replicating Carrol et al.'s reliability study for offshore turbines [80]. Since SCADA data is available, downtime and lost production could also be included.

  While the chapter explored the uncertainty in the calculated failure statistics, it did nothing to address those uncertainties. A methodology could be employed to assess uncertainty both inherent in the data itself and in the calculated metrics. While this thesis presents Bayesian models which account for uncertainty *once the metrics have already been calculated*, they do not address the underlying uncertainty in the failure rate calculation itself. Novel Bayesian techniques could be used to address uncertainty in the data collection process itself. For instance, Bayesian Measurement Error Models could be used together with expert elicitation [299]. Bayesian model averaging could be used combine results from multiple models, each representing a different aspect of data collection uncertainty [300]. This would provide more robust means for accounting for uncertainty, given the issues with data collection itself as recorded in the literature. The implication for this study is that the results of chapter chapter 6 are limited in their quantification of uncertainty *up to the point* where data collection is concerned. They could therefore be improved by more robust uncertainty quantification beyond this point.

- The Bayesian hierarchical model is limited in that the cost modelling is rudimentary. While the cost modelling assumptions are illustrative of potential of the methodology (and indeed, are in line with common O&M modelling assumptions [33]), there is a lot of room for modelling assumptions to be more thoroughly explored, especially in modelling staff wages and vessel leasing costs. Such figures can be difficult to come by in the public domain. Uncertainty in their values might still be elicited by expert judgement - a method used almost routinely in Bayesian modelling in other fields like the social sciences. Another step forward would be to more accurately capture price of energy fluctuations by some form of price modelling. This would more accurately represent the nature of CfDs, or indeed post subsidy financing. A more comprehensive Bayesian network model could also provide an estimate for LCoE. This would require the estimation of a number of additional cost parameters which were not available during the current analysis. The focus of the study was to extract value from the data that is available, and showcase the benefit of the proposed methodology in doing so. In opting to explore opportunity cost, the lost production as measured from the data was utilised. Likewise, measurements of technical availability can be used to update prior estimates directly. Via these means, one could explore the KPIs which are most pertinent to the Operator and OEM. An exploration of LCoE in the Bayesian framework, on the other hand, would require a more detailed dataset. The implication for this study is that the results are both limited and facilitated by the dataset. Future studies could therefore build on the approach used in this thesis by padding out these relatively uncomplicated Bayesian network models.

- There are three ways to gain further insight from the results presented results presented in section section 6.3. The first is by the use of a cost modelling tool, for instance the one developed by Dinwoodie et al. [17]. To begin with, a comparison of the different baseline distributional assumptions could be informative. This would build on the work of Scheu et al. [188], which investigated different baseline hazard rate assumption. Two factors would provide novelty: the fact that the results presented here are based on real world data from a currently operational wind farm, and an investigation into not only availability but the impact on Levelised Cost of Energy (LCoE) or some other financial metric. The results of such a study would further highlight the importance of modelling statistical uncertainty accurately.

There could also be value in a similar study where the impact of heterogeneity in turbine reliability is investigated. Given that mean values of turbine-by-turbine frailty are approximately normally distributed, this may not have such a significant impact as if there were gamma distributed random effects with a long tail. For instance, it would be interesting to see whether inclusion of random effects in cost models would significantly effect the financial outputs, especially for a large OWF.

The second avenue for further research is in optimising the timing of scheduled maintenance campaigns and the resources dedicated to them. Figure 6.14 has three stages to consider. First, there is the initial spike in failure intensity which occurs during the first few weeks after a scheduled maintenance work. This may not inspire any drastic changes to when services are performed, as typically annual service campaigns are carried out in Summer where lost production is low and access favourable. However, if operators could quantify the additional risk in terms of corrective maintenance, they could make a more informed decision about how many technicians and vessels they require.

The third avenue is in including more detailed data about exactly what work was carried out on the turbine. By doing so, the proposed methodology would be able to uncover dependencies between the selected covariates and specific failure modes. This data was not available for research presented in chapter chapter 6.

7.3 Recommendations to Operators

While the analyses presented in this thesis are limited by the fact that they represent one wind farm, there are several conclusions which might translate to other operators. These are summarised as follows.

1. The first step to more insight from operational maintenance data is better data quality [25]. Beyond this, the literature review presented in section chapter 2 identifies uncertainty quantification a key area for improvement for decision making tools [45, 43]. Data fusion also offers a route for better decision making tools, however its implementation is dependent on better data management tools and better data architectures. [15].

2. Operational data mining is potentially insightful, but data processing can be messy. This is demonstrated in chapter chapter 4, where failure rate estimates range from under 1 failures per turbine per year to over 10 failures per turbine per year. Point two of Hahn et al's Recommended practices for wind farm data collection [27] is especially pertinent to this point. Namely: *Identify your use-case and be aware of the resulting data needs.*

3. Bayesian methods provide a framework to update prior beliefs with new information in the form of operational data specific to the site in question. As shown in chapter chapter 5, Bayesian hierarchical models provide a framework to build conditional probability distributions from limited data points, given some prior knowledge of the problem. This facilitates the potential to scrutinise decisions relating to maintenance strategy dynamically (i.e as new data becomes available) while considering uncertainty.

4. Night shifts have the potential to increase availability. Whether this translates to significant savings in O&M cost depends on the electricity price, wages of technicians and potential for redundancy avoidance. Moreover, the H&S consequences of night shift working has not seen sufficient scrutiny in the literature to quantify additional operational risk. So, while night shifts have the potential to boost profitability, it very much depends on the individual circumstances of the wind farm in question. Since H&S is an under-scrutinised area for offshore wind and since the H&S statistics do present poor H&S performance for the

industry [70], increasing operational risk for technicians should not be taken lightly.

5. Since it is (to my best knowledge) the first result of its type, relationship established in section section 6.3 between annual servicing could be used to more accurately capture failure behaviour in decision making tools such as O&M cost models. While the relationship is very likely to change for any given wind farm, again the parameters of the model could be updated using a Bayesian regime. Outwith modelling, operators could be aware that annual servicing has the potential to lead to increased failure rates in the short term.

Bibliography

[1] Ian Havercroft. Brief-ccs in the ipcc sixth assessment (ar6) synthesis report. 2023.

[2] IEA. World energy outlook 2022, 2022.

[3] Global offshore wind report 2022. Technical report, Global Wind Energy Council, 2022.

[4] Global offshore wind report 2022. Technical report, Global Wind Energy Council, 2022.

[5] Renewables 2022 analysis and forecast to 2027. Technical report, International Energy Agency, 2022.

[6] JK Kaldellis and D Apostolou. Life cycle energy and carbon footprint of offshore wind energy. comparison with onshore counterpart. *Renewable Energy*, 108:72–84, 2017.

[7] Mehmet Bilgili, Abdulkadir Yasar, and Erdogan Simsek. Offshore wind power development in europe and its comparison with onshore counterpart. *Renewable and Sustainable Energy Reviews*, 15(2):905–915, 2011.

[8] JK Kaldellis and M Kapsali. Shifting towards offshore wind energy—recent activity and future development. *Energy policy*, 53:136–148, 2013.

[9] Tyler Stehly and Patrick Duffy. 2020 cost of wind energy review. Technical report, National Renewable Energy Lab.(NREL), Golden, CO (United States), 2021.

[10] Matti Niclas Scheu, Lorena Tremps, Ursula Smolka, Athanasios Kolios, and Feargal Brennan. A systematic failure mode effects and criticality analysis for offshore wind turbine systems towards integrated condition based maintenance strategies. *Ocean Engineering*, 176, 2019.

[11] Emanuel PP Soares-Ramos, Lais de Oliveira-Assis, Raúl Sarrias-Mena, and Luis M Fernández-Ramírez. Current status and future trends of offshore wind power in europe. *Energy*, 202:117787, 2020.

[12] Beis electricity generation costs (2020). https://www.gov.uk/government/publications/beis-electricity-generation-costs-2020. Accessed: 2023-06-27.

[13] Ciaran Gilbert, Jethro Browell, and David McMillan. Probabilistic access forecasting for improved offshore operations. *International Journal of Forecasting*, 2021.

[14] Helene Seyr and Michael Muskulus. Decision support models for operations and maintenance for offshore wind farms: A review. *Applied Sciences*, 9:278, 2019.

[15] Hamed Badihi, Youmin Zhang, Bin Jiang, Pragasen Pillay, and Subhash Rakheja. A comprehensive review on signal-based and model-based condition monitoring of wind turbines: Fault diagnosis and lifetime prognosis. *Proceedings of the IEEE*, 110(6):754–806, 2022.

[16] Tenghui Li, Xiaolei Liu, Zi Lin, and Rory Morrison. Ensemble offshore wind turbine power curve modelling – an integration of isolation forest, fast radial basis function neural network, and metaheuristic algorithm. *Energy*, 239:122340, 2022.

[17] Iain Dinwoodie, Ole Erik Endrerud, Matthias Hofmann, Rebecca Martin, and Iver Sperstad. Reference cases for verification of operation and maintenance simulation models for offshore wind farms. *Wind Engineering*, 39, 2015.

[18] Kolios Athanasios and Feargal Brennan. Romeo - deliverable report - d8.1: Development of a high-fidelity cost/revenue model for impact assessment. Technical report, ROMEO, 2018.

[19] Rafael Dawid, David McMillan, and Matthew Revie. Decision support tool for offshorewind farm vessel routing under uncertainty. *Energies*, 11:1–17, 2018.

[20] Clym Stock-Williams and Siddharth Krishna Swamy. Automated daily maintenance planning for offshore wind farms. *Renewable Energy*, 133:1393–1403, 2019.

[21] Chandra Ade Irawan, Djamila Ouelhadj, Dylan Jones, Magnus Stålhane, and Iver Bakken Sperstad. Optimisation of maintenance routing and scheduling for offshore wind farms. *Wind Engineering*, 39, 2015.

[22] Iraklis Lazakis and Shahroz Khan. An optimization framework for daily route planning and scheduling of maintenance vessel activities in offshore wind farms. *Ocean Engineering*, 225, 2021.

[23] Lijuan Dai, Magnus Stålhane, and Ingrid Utne. Routing and scheduling of maintenance fleet for offshore wind farms. *Wind Engineering*, 39, 2015.

[24] Idriss El-Thalji and Jayantha P. Liyanage. On the operation and maintenance practices of wind power asset: A status review and observations. *Journal of Quality in Maintenance Engineering*, 2012.

[25] Kevin Leahy, Colm Gallagher, Peter O'Donovan, and Dominic T. J. O'Sullivan. Issues with data quality for wind turbine condition monitoring and reliability analyses. *Ocean Engineering*, 209:107381, 2020.

[26] M. D. Reder, E. Gonzalez, and J. J. Melero. Wind turbine failures - tackling current problems in failure data analysis. *Journal of Physics: Conference Series*, 753, 2016.

[27] Berthold Hahn, Thomas Welte, Stefan Faulstich, Pramod Bangalore, Cyril Boussion, and Keith Harrison. Recommended practices for wind farm data collection and reliability assessment for om optimization. *Energy Procedia*, 2017.

[28] Alexios Koltsidopoulos Papatzimos. *Data-driven Operations & Maintenance for Offshore Wind Farms : Tools and Methodologies*. PhD Thesis, Industrial Doctorate of Offshore Renewable Energy, 2019.

[29] Alexios Koltsidopoulos Papatzimos and PR Thies. Towards automated and integrated data collectionstandardising workflow processes for the offshore wind industry. *Offshore Wind Energy*, 2017, 2017.

[30] Jade McMorland, Callum Flannigan, James Carroll, Maurizio Collu, David McMillan, William Leithead, and Andrea Coraddu. A review of operations and maintenance modelling with considerations for novel wind turbine concepts. *Renewable and Sustainable Energy Reviews*, 165:112581, 2022.

[31] Kevin Leahy, Colm Gallagher, Peter O'Donovan, and Dominic T.J O'Sullivan. Issues with data quality for wind turbine condition monitoring and reliability analyses. *Ocean Engineeringy*, 209:107381, 2020.

[32] Katharina Fischer and Diego Coronado. Condition monitoring of wind turbines: State of the art, user experience and recommendations. *VGB PowerTech*, 2015:51–56, 07 2015.

[33] Helene Seyr and Michael Muskulus. Decision support models for operations and maintenance for offshore wind farms: A review. *Applied Sciences*, 9:278, 2019.

[34] James W. Taylor and Jooyoung Jeon. Probabilistic forecasting of wave height for offshore wind turbine maintenance. *European Journal of Operational Research*, 267, 2018.

[35] Sebastian Pfaffel, Stefan Faulstich, and Kurt Rohrig. Considering uncertainties of key performance indicators in wind turbine operation. *Applied Sciences*, 10(3):898, 2020.

[36] Brian Jenkins, Ian Belton, James Carroll, and David McMillan. Estimating the major replacement rates in next-generation offshore wind turbines using structured expert elicitation. *Journal of Physics: Conference Series*, 2362:012020, 11 2022.

[37] Peter Congon. *Bayesian Statistical Modelling*. 2006.

[38] Andrew Gelman, John B. Carlin, Hal S. Stern, David B Dunson, Aki Vehtari, and Donald B Rubin. *Bayesian Data Analysis*. Chapman and Hall/CBC, 2013.

[39] Pierre-Alexandre Mattei. A Parsimonious Tour of Bayesian Model Uncertainty, 2019.

[40] Alyson G. Wilson and Kassandra M. Fronczyk. Bayesian reliability: Combining information. *Quality Engineering*, 29(1):119–129, 2017. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/08982112.2016.1211889.

[41] Debora Cevasco, Athanasios Kolios, and Sofia Koukoura. Reliability, availability, maintainability data review for the identification of trends in offshore wind energy application. *Renewable and Sustainable Energy Reviews*, 136:110414, 2021.

[42] Mahmood Shafiee. Maintenance logistics organization for offshore wind energy: Current progress and future perspectives. *Renewable Energy*, 77, 2015.

[43] Cuong D. Dao, Behzad Kazemtabrizi, and Christopher J. Crabtree. Offshore wind turbine reliability and operational simulation under uncertaintiess. *Wind Energy*, 23, 2020.

[44] Matti Niclas Scheu, Athanasios Kolios, Tim Fischer, and Feargal Brennan. Influence of statistical uncertainty of component reliability estimations on offshore wind farm availability. *Reliability Engineering and System Safety*, 168, 2017.

[45] Helene Seyr and Michael Muskulus. Value of information of repair times for offshore wind farm maintenance planning. *Journal of Physics: Conference Series*, 753, 2016.

[46] Mingxin Li, Xiaoli Jiang, James Carroll, and Rudy R. Negenborn. A multi-objective maintenance strategy optimization framework for offshore wind farms considering uncertainty. *Applied Energy*, 321:119284, 2022.

[47] Alexios Koltsidopoulos Papatzimos. *Data-driven Operations Maintenance for Offshore Wind Farms : Tools and Methodologies*. PhD thesis, Industrial Doctorate of Offshore Renewable Energy, 2019.

[48] Masaharu Tsujimoto, Yuya Kajikawa, Junichi Tomita, and Yoichi Matsumoto. A review of the ecosystem concept — towards coherent ecosystem design. *Technological Forecasting and Social Change*, 136:49–58, 2018.

[49] Jay Lee, Hossein Davari Ardakani, Shanhu Yang, and Behrad Bagheri. Industrial big data analytics and cyber-physical systems for future maintenance & service innovation. *Procedia cirp*, 38:3–7, 2015.

[50] Slinger Jansen, Michael A Cusumano, and Sjaak Brinkkemper. *Software ecosystems: analyzing and managing business networks in the software industry*. Edward Elgar Publishing, 2013.

[51] Offshore Renewable Energy Catapult. Guide to an offshore wind farm: Updated and extended, 2019.

[52] Barry Johnston, Aoife Foley, John Doran, and Timothy Littler. Levelised cost of energy, a challenge for offshore wind. *Renewable Energy*, 160:876–885, 2020.

[53] E. Gonzalez, EM. Nanos, H. Seyr, L. Valldecabres, Nurseda Y. Yürüşen, Ursula Smolka, Michael Muskulus, and Julio J. Melero. Key performance indicators for wind farm operation and maintenance. *Energy Procedia*, 137:559–570, 2017.

[54] Peter Tavner. *Offshore wind turbines: reliability, availability and maintenance*, volume 13. IET, 2012.

[55] Global offshore wind report 2020. Technical report, Global Wind Energy Council, 2020.

[56] Cuong D. Dao, Behzad Kazemtabrizi, and Christopher J. Crabtree. Offshore wind turbine reliability and operational simulation under uncertaintiess. *Wind Energy*, 23, 2020.

[57] Gdp deflators at market prices, and money gdp. `https://www.gov.uk/government/collections/gdp-deflators-at-market-prices-and-money-gdp`. Accessed: 2023-06-27.

[58] Cuong Dao, Behzad Kazemtabrizi, and Christopher Crabtree. Wind turbine reliability data review and impacts on levelised cost of energy. *Wind Energy*, 22, 2019.

[59] René MM Slot, John D Sørensen, Bruno Sudret, Lasse Svenningsen, and Morten L Thøgersen. Surrogate model uncertainty in wind turbine reliability assessment. *Renewable Energy*, 151:1150–1162, 2020.

[60] Soklin Tuy, Han Soo Lee, and Karodine Chreng. Integrated assessment of offshore wind power potential using weather research and forecast (wrf) downscaling with sentinel-1 satellite imagery, optimal sites, annual energy production and equivalent co2 reduction. *Renewable and Sustainable Energy Reviews*, 163:112501, 2022.

[61] Marcelo Iury S. Oliveira and Bernadette Farias Lóscio. What is a data ecosystem? In *Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age*, dg.o '18, New York, NY, USA, 2018. Association for Computing Machinery.

[62] Yuri Demchenko, Paola Grosso, Cees de Laat, and Peter Membrey. Addressing big data issues in scientific data infrastructure. In *2013 International Conference on Collaboration Technologies and Systems (CTS)*, pages 48–55, 2013.

[63] Conor McKinnon, Alan Turnbull, Sofia Koukoura, James Carroll, and Alasdair McDonald. Effect of time history on normal behaviour modelling using scada data to predict wind turbine failures. *Energies*, 13(18), 2020.

[64] Elena Gonzalez, Maik Reder, and Julio J Melero. Scada alarms processing for wind turbine component failure detection. In *Journal of Physics: Conference Series*, volume 753, page 072019. IOP Publishing, 2016.

[65] Wenxian Yang, Peter J. Tavner, Y. Feng Christopher J. Crabtree, and Y. Qiu. Wind turbine condition monitoring: technical and commercial challenges. *Wind Energy*, 17, 2014.

[66] Rosemary Tawn, Jethro Browell, and Iain Dinwoodie. Missing data in wind farm time series: Properties and effect on forecasts. *Electric Power Systems Research*, 189:106640, 2020.

[67] Berthold Hahn, Thomas Welte, Stefan Faulstich, Pramod Bangalore, Cyril Boussion, and Keith Harrison. Recommended practices for wind farm data collection and reliability assessment for O&M optimization. *Energy Procedia*.

[68] M. D. Reder, E. Gonzalez, and J. J. Melero. Wind Turbine Failures - Tackling current Problems in Failure Data Analysis. *Journal of Physics: Conference Series*, 753, 2016.

[69] Sebastian Pfaffel, Stefan Faulstich, and Kurt Rohrig. Performance and reliability of wind turbines: A review. *Energies*, 10:1904, 2017.

[70] David Rowell, David McMillan, and James Carroll. A review of health and safety in the offshore wind industry. *Manuscript under review*, 2023.

[71] Andrew Hopkins et al. Thinking about process safety indicators. 2009.

[72] Global offshore wind health and safety organisation. `http://https://www.gplusoffshorewind.com`. Accessed: 2023-03-09.

[73] Yingning Qiu, Yanhui Feng, Peter Tavner, Paul Richardson, Gabor Erdos, and Bindi Chen. Wind turbine scada alarm analysis for improving reliability. *Wind Energy*, 15(8):951–966, 2012.

[74] Christos Kaidis, Bahri Uzunoglu, and Filippos Amoiralis. Wind turbine reliability estimation for different assemblies and failure severity categories. *IET Renewable Power Generation*, 9(8):892–899, 2015.

[75] S. Draycott, B. Sellar, T. Davey, D. R. Noble, V. Venugopal, and D. M. Ingram. Capture and simulation of the ocean environment for offshore renewable energy. *Renewable and Sustainable Energy Reviews*, 104:15–29, 2019.

[76] Stefan Faulstich, Berthold Hahn, and Peter J Tavner. Wind turbine downtime and its importance for offshore deployment. *Wind energy*, 14(3):327–337, 2011.

[77] F. Spinato, P. J. Tavner, G. J.W. Van Bussel, and E. Koutoulakos. Reliability of wind turbine subassemblies. *Wind Energy*, 3:387–401, 2009.

[78] James Carroll, Alasdair McDonald, and David Mcmillan. Reliability comparison of wind turbines with DFIG and PMG drive trains. *Energy Conversion, IEEE Transactions on*, 30:663–670, June 2015.

[79] Debora Cevasco, Maurizio Collu, and Z Lin. O&m cost-based fmeca: Identification and ranking of the most critical components for 2-4 mw geared offshore wind turbines. In *Journal of Physics: Conference Series*, volume 1102, page 012039. IOP Publishing, 2018.

[80] James Carroll, Alasdair McDonald, and David McMillan. Failure rate, repair time and unscheduled O&M cost analysis of offshore wind turbines. *Wind Energy*, 19:1107–1119, 2019.

[81] Sparta Portfolio Review 2016/17, 2017.

[82] James Carroll, Alasdair McDonald, and David McMillan. Failure rate, repair time and unscheduled O&M cost analysis of offshore wind turbines. *Wind Energy*, 19:1107–1119, 2019.

[83] Chun Su, Yang Yang, Xiaolin Wang, and Zhaoyong Hu. Failures analysis of wind turbines: Case study of a chinese wind farm. In *2016 Prognostics and System Health Management Conference (PHM-Chengdu)*, pages 1–6, 2016.

[84] Johan Ribrant and Lina Bertling. Survey of failures in wind power systems with focus on swedish wind power plants during 1997–2005. *IEEE Transactions on Energy Conversion*, 22:167–173, 2007.

[85] Christos Kaidis, Bahri Uzunoglu, and Filippos Amoiralis. Wind turbine reliability estimation for different assemblies and failure severity categories. *IET Renewable Power Generation*, 9(8):892–899, 2015.

[86] Michael Wilkinson, Keir Harman, Ben Hendriks, Fabio Spinato, and Thomas van Delft. Measuring wind turbine reliability, results of the reliawind project. *Proceedings of European Wind Energy Conference EWEA2011, Brussels, Belgium: European Wind Energy Association*, 2011.

[87] Danish Standards Foundation. *A Guide to RDS - Reference Designation Systems. TAG Numbers for Systems in Accordance with the ISO/IEC 81346 Standard Series*, 2020.

[88] NERC. *Wind Turbine Generation: Data Reporting Instructions*, 2018.

[89] Health and Safety Executive. Riddor—reporting of injuries, diseases and dangerous occurrences regulations 2013. 2013.

[90] Fred A Manuele. Leading & lagging indicators. *Professional Safety*, 54(12):28–33, 2009.

[91] Matthew Hallowell, Mike Quashne, Rico Salas, Brad MacLean, and Ellen Quinn. The statistical invalidity of trir as a measure of safety performance. *Professional Safety*, 66(04):28–34, 2021.

[92] Andrew Hopkins. Lessons from esso's gas plant explosion at longford. *Warwick Pearse, Clare Gallagher and Liz Bluff*, 1, 2001.

[93] Sebastian Pfaffel, Stefan Faulstich, and Shawn Sheng. Recommended key performance indicators for operational management of wind turbines. In *Journal of Physics: Conference Series*, volume 1356, page 012040. IOP Publishing, 2019.

[94] Jannis Tautz-Weinert and Simon J. Watson. Using scada data for wind turbine condition monitoring - a review. *IET Renewable Power Generation*, 11, 2017.

[95] Elena Gonzalez, Maik Reder, and Julio Melero. SCADA alarms processing for wind turbine component failure detection. *Journal of Physics Conference Series*, 753, September 2016.

[96] Adrian Stetco, Fateme Dinmohammadi, Xingyu Zhao, Valentin Robu, David Flynn, and Mike Barnes. Machine learning methods for wind turbine condition monitoring: A review. *Renewable Energy*, 133, 2019.

[97] Fausto Pedro García Márquez, Andrew Mark Tobias, Jesús María Pinar Pérez, and Mayorkinos Papaelias. Condition monitoring of wind turbines: Techniques and methods. *Renewable Energy*, 46:169–178, 2012.

[98] Pierre Tchakoua, René Wamkeue, Mohand Ouhrouche, Fouad Slaoui-Hasnaoui, Tommy Andy Tameghe, and Gabriel Ekemb. Wind turbine condition monitoring: State-of-the-art review, new trends, and future challenges. *Energies*, 7(4):2595–2630, 2014.

[99] Giovanni Rinaldi, Philipp R. Thies, and Lars Johanning. Current status and future trends in the operation and maintenance of offshore wind turbines: A review. *Energies*, 14(9), 2021.

[100] Lina Alhmoud and Bingsen Wang. A review of the state-of-the-art in wind-energy reliability analysis. *Renewable and Sustainable Energy Reviews*, 81:1643–1651, 2018.

[101] Wei Qiao and Dingguo Lu. A survey on wind turbine condition monitoring and fault diagnosis—part ii: Signals and signal processing methods. *IEEE Transactions on Industrial Electronics*, 62(10):6546–6557, 2015.

[102] DJ Lekou, F Mouzakis, A Anastasopoulos, and D Kourousis. Emerging techniques for health monitoring of wind turbine gearboxes and bearings. In *Proc. EWEC*, pages 16–19, 2009.

[103] Nathan Weller. Acceleration enveloping–higher sensitivity, earlier detection. *Orbit*, pages 10–19, 2004.

[104] Dingguo Lu and Wei Qiao. Frequency demodulation-aided condition monitoring for drivetrain gearboxes. In *2013 IEEE Transportation Electrification Conference and Expo (ITEC)*, pages 1–6. IEEE, 2013.

[105] Wenxian Yang, Peter J. Tavner, Y. Feng Christopher J. Crabtree, and Y. Qiu. Wind turbine condition monitoring: technical and commercial challenges. *Wind Energy*, 17, 2014.

[106] Katharina Fischer and Diego Coronado. Condition monitoring of wind turbines: State of the art, user experience and recommendations. *VGB PowerTech*, 2015:51–56, 07 2015.

[107] Edgar L. Andreas and Sinny Wang. Predicting significant wave height off the northeast coast of the United States. *Ocean Engineering*, 34(8):1328–1335, 2007.

[108] David McMillan and Graham W. Ault. Quantification of Condition Monitoring Benefit for Offshore Wind Turbines. *Wind Engineering*, 31(4):267–285, 2007.

[109] European centre for medium-range weather forecasts (ecmwf). public datasets. `https://www.ecmwf.int/`. Accessed: 2022-11-11.

[110] Global forecast system (gfs).

[111] European directory of marine organisations (edmo).

[112] Bundesamt für seeschifffahrt und hydrographie (bsh). fino datenbank. `https://www.bsh.de/DE/THEMEN/Beobachtungssysteme/Messnetz-MARNET/FINO/fino_node.html`. Accessed: 2022-11-11.

[113] Lucy Y. Pao and Kathryn E. Johnson. A tutorial on the dynamics and control of wind turbines and wind farms. In *2009 American Control Conference*, pages 2076–2089, 2009.

[114] Jaroslav Frnda, Marek Durica, Jan Nedoma, Stanislav Zabka, Radek Martinek, and Michal Kostelansky. A weather forecast model accuracy analysis and ecmwf enhancement proposal by neural network. *Sensors*, 19(23), 2019.

[115] Jaroslav Frnda, Marek Durica, Jan Rozhon, Maria Vojtekova, Jan Nedoma, and Radek Martinek. Ecmwf short-term prediction accuracy improvement by deep learning. *Scientific Reports*, 12(1):1–11, 2022.

[116] Mahmood Shafiee. Maintenance logistics organization for offshore wind energy: Current progress and future perspectives. *Renewable Energy*, 77, 2015.

[117] Lijuan Dai, Magnus Stålhane, and Ingrid B Utne. Routing and scheduling of maintenance fleet for offshore wind farms. *Wind Engineering*, 39(1):15–30, 2015.

[118] Magnus Stålhane, Lars Magnus Hvattum, and Vidar Skaar. Optimization of routing and scheduling of vessels to perform maintenance at offshore wind farms. *Energy Procedia*, 80:92–99, 2015. 12th Deep Sea Offshore Wind RD Conference, EERA DeepWind'2015.

[119] Rafael Dawid, David McMillan, and Matthew Revie. Decision support tool for offshorewind farm vessel routing under uncertainty. *Energies*, 11:1–17, 2018.

[120] HZ Tan, Wei Lv, LW Jin, ZC Liu, and JS Feng. Modeling and solution of offshore wind farm maintenance scheduling. *DEStech Trans. Environ. Energy Earth Sci*, 2017.

[121] Nora Tangen Raknes, Katrine Ødeskaug, Magnus Stålhane, and Lars Magnus Hvattum. Scheduling of maintenance tasks and routing of a joint vessel fleet for multiple offshore wind farms. *Journal of Marine Science and Engineering*, 5(1):11, 2017.

[122] J. Browell, I. Dinwoodie, and D. McMillan. Forecasting for day-ahead offshore maintenance scheduling under uncertainty. In *Risk, Reliability and Safety: Innovating Theory and Practice - Proceedings of the 26th European Safety and Reliability Conference, ESREL 2016*, 2017.

[123] Ricardo J. Bessa, Corinna Möhrlen, Vanessa Fundel, Malte Siefert, Jethro Browell, Sebastian Haglund El Gaidi, Bri-Mathias Hodge, Umit Cali, and George Kariniotakis. Towards improved understanding of the applicability of uncertainty forecasts in the electric power industry. *Energies*, 10(9), 2017.

[124] J. Browell, I. Dinwoodie, and D. McMillan. Forecasting for day-ahead offshore maintenance scheduling under uncertainty. In *Risk, Reliability and Safety: Innovating Theory and Practice - Proceedings of the 26th European Safety and Reliability Conference, ESREL 2016*, 2017.

[125] VM Catterson, D McMillan, I Dinwoodie, M Revie, J Dowell, J Quigley, and K Wilson. An economic impact metric for evaluating wave height forecasters for offshore wind maintenance access. *Wind Energy*, 19(2):199–212, 2016.

[126] Sebastian Pfaffel, Stefan Faulstich, and Kurt Rohrig. Performance and reliability of wind turbines: A review. *Energies*, 10:1904, 2017.

[127] Karoline Pelka and Katharina Fischer. Field-data-based reliability analysis of power converters in wind turbines: Assessing the effect of explanatory variables. *Wind Energy*, 26(3):310–324, 2023.

[128] Davide Cazzaro, Alessio Trivella, Francesco Corman, and David Pisinger. Multiscale optimization of the design of offshore wind farms. *Applied Energy*, 314:118830, 2022.

[129] I Maassen van den Brink, N Olijve, and W Pustjens. Determination of the cost levels of wind farms (and their grid connections) in new offshore wind energy search areas. In *Technical Report*. Blix Consultancy & Pondera, 2020.

[130] Maik Reder, Nurseda Y. Yürüşen, and Julio J. Melero. Data-driven learning framework for associating weather conditions and wind turbine failures. *Reliability Engineering and System Safety*, 169, 2018.

[131] M. Reder and J. J. Melero. A bayesian approach for predicting wind turbine failures based on meteorological conditions. *Journal of Physics: Conference Series*, 1037, 2018.

[132] Peter Tavner, Clare Edwards, Andy Brinkman, and Fabio Spinato. Influence of wind speed on wind turbine reliability. *Wind Engineering*, 30(1):55–72, 2006.

[133] P.J. Tavner, Rosa Gindele, S. Faulstich, G Whittle, and David Greenwood. Study of effects of weather location on wind turbine failure rates. *European Wind Energy Conference*, 16, 01 2010.

[134] S Faulstich, P Lyding, and PJ Tavner. Effects of wind speed on wind turbine availability. In *Proceedings of the 2011 European Wind Energy Conference*, 2011.

[135] G Wilson and D. McMillan. Assessing wind farm reliability using weather dependent failure rates. *Journal of Physics: Conference Series*, 524, 2014.

[136] Maik Reder, Nurseda Y. Yürüşen, and Julio J. Melero. Data-driven learning framework for associating weather conditions and wind turbine failures. *Reliability Engineering and System Safety*, 169, 2018.

[137] Vaclav Slimacek and Bo Lindqvist. Reliability of wind turbines modeled by a poisson process with covariates, unobserved heterogeneity and seasonality. *Wind Energy*, 19:n/a–n/a, 02 2016.

[138] Chandra Ade Irawan, Djamila Ouelhadj, Iver Bakken Sperstad, and Dylan Jones. A combined tactical and operational framework for maintenance scheduling and routing in offshore wind farms. *Journal of the Operational Research Society*, pages 1–20, 2022.

[139] François Besnard, Katharina Fischer, and Lina Bertling Tjernberg. A model for the optimization of the maintenance support organization for offshore wind farms. *IEEE Transactions on Sustainable Energy*, 4:443–450, 2013.

[140] Matti Scheu, Denis Matha, Matthias Hofmann, and Michael Muskulus. Maintenance strategies for large offshore wind farms. *Energy Procedia*, 24:281–288, 2012.

[141] Iain Dinwoodie, David McMillan, Matthew Revie, Iraklis Lazakis, and Yalcin Dalgic. Development of a combined operational and strategic decision support model for offshore wind. *Energy Procedia*, 35:157–166, 2013. DeepWind'2013 – Selected papers from 10th Deep Sea Offshore Wind RD Conference, Trondheim, Norway, 24 – 25 January 2013.

[142] Ole-Erik V. Endrerud, Jayantha P. Liyanage, and Nenad Keseric. Marine logistics decision support for operation and maintenance of offshore wind parks with a multi method simulation model. In *Proceedings of the Winter Simulation Conference 2014*, pages 1712–1722, 2014.

[143] Matthias Hofmann and Iver Bakken Sperstad. Nowicob – a tool for reducing the maintenance costs of offshore wind farms. *Energy Procedia*, 35:177–186, 2013. DeepWind'2013 – Selected papers from 10th Deep Sea Offshore Wind RD Conference, Trondheim, Norway, 24 – 25 January 2013.

[144] Onshore wind farms maintenance optimization using a stochastic model. *Mathematical and Computer Modelling*, 57(7):1884–1890, 2013.

[145] Shuya Zhong, Athanasios A. Pantelous, Michael Beer, and Jian Zhou. Constrained non-linear multi-objective optimisation of preventive maintenance scheduling for offshore wind farms. *Mechanical Systems and Signal Processing*, 104:347–369, 2018.

[146] François Besnard, Michael Patriksson, Ann-Brith Strömberg, Adam Wojciechowski, Katharina Fischer, and Lina Bertling. A stochastic model for opportunistic maintenance planning of offshore wind farms. In *2011 IEEE Trondheim PowerTech*, pages 1–8, 2011.

[147] Eunshin Byon, Lewis Ntaimo, and Yu Ding. Optimal maintenance strategies for wind turbine systems under stochastic weather conditions. *IEEE Transactions on Reliability*, 59(2):393–404, 2010.

[148] D. Pattison, M. Segovia Garcia, W. Xie, F. Quail, M. Revie, R. I. Whitfield, and I. Irvine. Intelligent integrated maintenance for wind power generation. *Wind Energy*, 19(3):547–562, 2016.

[149] John Dalsgaard Sørensen. Framework for risk-based planning of operation and maintenance for offshore wind turbines. *Wind Energy*, 12(5):493–506, 2009.

[150] François Besnard, Michael Patriksson, Ann-Brith Strömberg, Adam Wojciechowski, Katharina Fischer, and Lina Bertling. A stochastic model for opportunistic maintenance planning of offshore wind farms. In *2011 IEEE Trondheim PowerTech*, pages 1–8, 2011.

[151] Mihai Florian and John Dalsgaard Sørensen. Risk-based planning of operation and maintenance for offshore wind farms. *Energy Procedia*, 137:261–272, 2017. 14th Deep Sea Offshore Wind RD Conference, EERA DeepWind'2017.

[152] David McMillan and Graham W. Ault. Quantification of condition monitoring benefit for offshore wind turbines. *Wind Engineering*, 31(4):267–285, 2007.

[153] Allan May, David McMillan, and Sebastian Thöns. Economic analysis of condition monitoring systems for offshore wind turbine sub-systems. *IET Renewable Power Generation*, 9(8):900–907, 2015.

[154] Yang Lu, Liping Sun, Xinyue Zhang, Feng Feng, Jichuan Kang, and Guoqiang Fu. Condition based maintenance optimization for offshore wind turbine considering opportunities based on neural network approach. *Applied Ocean Research*, 74:69–79, 2018.

[155] Zhigang Tian, Tongdan Jin, Bairong Wu, and Fangfang Ding. Condition based maintenance optimization for wind power generation systems under continuous monitoring. *Renewable Energy*, 36(5):1502–1509, 2011.

[156] Jichuan Kang, Jose Sobral, and C Guedes Soares. Review of condition-based maintenance strategies for offshore wind energy. *Journal of Marine Science and Application*, 18:1–16, 2019.

[157] Gustavo de Novaes Pires Leite, Alex Maurício Araújo, and Pedro André Carvalho Rosas. Prognostic techniques applied to maintenance of wind turbines: a concise and specific review. *Renewable and Sustainable Energy Reviews*, 81:1917–1925, 2018.

[158] Abhinav Saxena, Indranil Roychoudhury, Jose Celaya, Sankalita Saha, Bhaskar Saha, and Kai Goebel. Requirements specification for prognostics performance-an overview. *Aiaa infotech@ aerospace 2010*, page 3398, 2010.

[159] G. Rinaldi, P.R. Thies, R. Walker, and L. Johanning. A decision support model to optimise the operation and maintenance strategies of an offshore renewable energy farm. *Ocean Engineering*, 145:250–262, 2017.

[160] Huifang Niu, Jianchao Zeng, Hui Shi, Xiaohong Zhang, and Jianyu Liang. Degradation modeling and remaining useful life prediction for a multi-component system with stochastic dependence. *Computers & Industrial Engineering*, 175:108889, 2023.

[161] Graeme Hawker and David Mcmillan. The impact of maintenance contract arrangements on the yield of offshore wind power plants. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 229, 07 2015.

[162] Alberto Pliego Marugán, Fausto Pedro García Márquez, and Jesús María Pinar Pérez. A techno-economic model for avoiding conflicts of interest between owners of offshore wind farms and maintenance suppliers. *Renewable and Sustainable Energy Reviews*, 168:112753, 2022.

[163] Bingying Zhang and Zijun Zhang. A robust model for scheduling power productions of multiple offshore wind farms using one-to-many maintenance services. *IET Renewable Power Generation*, 15(13):2997–3013, 2021.

[164] Sparta Portfolio Review 2017/18, 2018.

[165] Tosin Adedipe, Mahmood Shafiee, and Enrico Zio. Bayesian Network Modelling for the Wind Energy Industry: An Overview. *Reliability Engineering and System Safety*, 202, 2020.

[166] Riccardo Mogre, Srinivas Talluri, and Federico D'Amico. A decision framework to mitigate supply chain risks: An application in the offshore-wind industry. *IEEE Transactions on Engineering Management*, 63(3):316–325, 2016.

[167] Md Imran Hasan Tusar and Bhaba R Sarker. Spare parts control strategies for offshore wind farms: a critical review and comparative study. *Wind Engineering*, 46(5):1629–1656, 2022.

[168] François Besnard, Katharina Fischer, and Lina Bertling Tjernberg. A model for the optimization of the maintenance support organization for offshore wind farms. *IEEE Transactions on Sustainable Energy*, 4:443–450, 2013.

[169] Y. Dalgic, I. Lazakis, I. Dinwoodie, D. McMillan, M. Revie, and J. Majumder. The influence of multiple working shifts for offshore wind farm om activities - strathow-om tool. In *RINA, Royal Institution of Naval Architects - Design and Operation of Wind Farm Support Vessels 2015, Papers*, 2019.

[170] Michiel A.J. uit het Broek, Jasper Veldman, Stefano Fazi, and Roy Greijdanus. Evaluating resource sharing for offshore wind farm maintenance: The case of jack-up vessels. *Renewable and Sustainable Energy Reviews*, 109:619–632, 2019.

[171] Yalcin Dalgic, Iraklis Lazakis, Osman Turan, and Sol Judah. Investigation of optimum jack-up vessel chartering strategy for offshore wind farm o&m activities. *Ocean Engineering*, 95:106–115, 2015.

[172] Magnus Stålhane, Marielle Christiansen, Odin Kirkeby, and Andreas Jebsen Mikkelsen. Optimizing jack-up vessel strategies for maintaining offshore wind farms. *Energy Procedia*, 137:291–298, 2017. 14th Deep Sea Offshore Wind RD Conference, EERA DeepWind'2017.

[173] Crown Estate. Jack-up vessel optimisation: improving offshore wind performance through better use of jack-up vessels in the operations and maintenance phase, 2014.

[174] Elena Gonzalez, Bruce Stephen, David Infield, and Julio J Melero. Using high-frequency scada data for wind turbine performance monitoring: A sensitivity study. *Renewable energy*, 131:841–853, 2019.

[175] Alan Turnbull, James Carroll, and Alasdair McDonald. Combining scada and vibration data into a single anomaly detection model to predict wind turbine component failure. *Wind Energy*, 24(3):197–211, 2021.

[176] Margarita Moreno-Betancur, John B Carlin, Samuel L Brilleman, Stephanie K Tanamas, Anna Peeters, and Rory Wolfe. Survival analysis with time-dependent covariates subject to missing data or measurement error: Multiple imputation for joint modeling (mijm). *Biostatistics*, 19(4):479–496, 2018.

[177] David M Zucker and Alan F Karr. Nonparametric survival analysis with time-dependent covariate effects: a penalized partial likelihood approach. *The Annals of Statistics*, 18(1):329–353, 1990.

[178] Valen E Johnson, Ann Moosman, and Paul Cotter. A hierarchical model for estimating the early reliability of complex systems. *IEEE Transactions on Reliability*, 54(2):224–231, 2005.

[179] V. Slimacek and B. H. Lindqvist. Reliability of wind turbines modeled by a Poisson process with covariates, unobserved heterogeneity and seasonality. *Wind Energy*, 19(11):1991–2002, 2016. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/we.1964.

[180] Tomas Gintautas and John Dalsgaard Sørensen. Improved methodology of weather window prediction for offshore operations based on probabilities of operation failure. *Journal of Marine Science and Engineering*, 5(2), 2017.

[181] I Dinwoodie, VM Catterson, and D McMillan. Wave height forecasting to improve off-shore access and maintenance scheduling. In *2013 IEEE Power & Energy Society General Meeting*, pages 1–5. IEEE, 2013.

[182] David Wilkie and Carmine Galasso. A Bayesian model for wind farm capacity factors. *Energy Conversion and Management*, 252:114950, 2022.

[183] Adrian Stetco, Fateme Dinmohammadi, Xingyu Zhao, Valentin Robu, David Flynn, and Mike Barnes. Machine learning methods for wind turbine condition monitoring: A review. *Renewable Energy*, 133, 2019.

[184] Gong Li and Jing Shi. Applications of Bayesian methods in wind energy conversion systems. *Renewable Energy*, 43, 2012.

[185] Helene Seyr and Michael Muskulus. Value of information of repair times for offshore wind farm maintenance planning. *Journal of Physics: Conference Series*, 753, 2016.

[186] Helene Seyr and Michael Muskulus. Interaction of repair time distributions with a weather model. In *29th International Congress on Condition Monitoring and Diagnostics Engineering Management (COMADEM 2016)*, 2017.

[187] Helene Seyr and Michael Muskulus. How does accuracy of weather forecasts influence the maintenance cost in offshore wind farms? In *Proceedings of the International Offshore and Polar Engineering Conference*, 2017.

[188] Matti Niclas Scheu, Athanasios Kolios, Tim Fischer, and Feargal Brennan. Influence of statistical uncertainty of component reliability estimations on offshore wind farm availability. *Reliability Engineering and System Safety*, 168, 2017.

[189] Elias Gyftodimos and Peter A. Flach. Hierarchical Bayesian Networks: An Approach to Classification and Learning for Structured Data. In *Methods and Applications of Artificial Intelligence*, 2013.

[190] Maryam Ashrafi and Hamid Davoudpour. A hierarchical Bayesian network to compare maintenance strategies based on cost and reliability: A case of onshore wind turbines. *International Journal of Industrial Engineering : Theory Applications and Practice*, 26:361–375, 2019.

[191] Alessio Ishizaka and Ashraf Labib. Analytic hierarchy process and expert choice: Benefits and limitations. *OR insight*, 22(4):201–220, 2009.

[192] Hong Zhong Huang, Ming J. Zuo, and Zhan Quan Sun. Bayesian reliability analysis for fuzzy lifetime data. *Fuzzy Sets and Systems*, 157, 2006.

[193] Alyson G. Wilson and Aparna V. Huzurbazar. Bayesian networks for multilevel system reliability. *Reliability Engineering System Safety*, 92(10):1413–1420, 2007.

[194] H. Boudali and J.B. Dugan. A discrete-time bayesian network reliability modeling and analysis framework. *Reliability Engineering System Safety*, 87(3):337–349, 2005.

[195] Saideep Nannapaneni and Sankaran Mahadevan. Reliability analysis under epistemic uncertainty. *Reliability Engineering System Safety*, 155:9–20, 2016.

[196] David J. Hand and Keming Yu. Idiot's Bayes - Not so stupid after all? *International Statistical Review*, 69, 2001.

[197] Wind turbine performance analysis based on multivariate higher order moments and bayesian classifiers. *Control Engineering Practice*, 49:204–211, 2016.

[198] Zhe Song, Zijun Zhang, Yu Jiang, and Jin Zhu. Wind turbine health state monitoring based on a bayesian data-driven approach. *Renewable Energy*, 125:172–181, 2018.

[199] Charles E. "Plumley, Graeme K. Wilson, Andrew D. Kenyon, Francis Quail, and Athena" Zitrou. Diagnostics and prognostics utilising dynamic bayesian networks applied to a wind turbine gearbox. *9th International Conference on Condition Monitoring and Machinery Failure Prevention Technologies 2012, CM 2012 and MFPT 2012*, 1, 2012.

[200] Masoud "Asgarpour and John Dalsgaard" Sørensen. Bayesian based prognostic model for predictive maintenance of offshore wind farms. *International Journal of Prognostics and Health Management*, 9, 2018.

[201] Jannie S. Nielsen and John D. Sørensen. Bayesian estimation of remaining useful life for wind turbine blades. *Energies*, 10(5), 2017.

[202] Laura Uusitalo. Advantages and challenges of bayesian networks in environmental modelling. *Ecological Modelling*, 203, 2007.

[203] He Li, C. Guedes Soares, and Hong-Zhong Huang. Reliability analysis of a floating offshore wind turbine using bayesian networks. *Ocean Engineering*, 217:107827, 2020.

[204] Iraklis Lazakis and Maria A Kougioumtzoglou. Assessing offshore wind turbine reliability and availability. *Proceedings of the Institution of Mechanical Engineers, Part M: Journal of Engineering for the Maritime Environment*, 233(1):267–282, 2019.

[205] A. Zitrou, T. Bedford, and L. Walls. Quantification and modelling of epistemic uncertainties for availability risk of future offshore wind farms using expert judgment. In *Safety and Reliability: Methodology and Applications - Proceedings of the European Safety and Reliability Conference, ESREL 2014*.

[206] A. Zitrou, T. Bedford, and A. Daneshkhah. Robustness of maintenance decisions: Uncertainty modelling and value of information. *Reliability Engineering System Safety*, 120:60–71, 2013.

[207] Athena Zitrou, Tim Bedford, and Lesley Walls. Modeling epistemic uncertainty in offshore wind farm production capacity to reduce risk. *Risk Analysis*, 0, 2021.

[208] Mihai Florian and John Dalsgaard Sørensen. Risk-based planning of operation and maintenance for offshore wind farms. *Energy Procedia*, 137:261–272, 2017.

[209] Y Feng, P. J Tavner, and H Long. Early experiences with uk round 1 offshore wind farms. *Proceedings of Institution of Civil Engineers: Energy*, 163, 2010.

[210] Christopher J Crabtree, Donatella Zappalá, and Simon I Hogg. Wind energy: Uk experiences and offshore operational challenges. *Proceedings of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy*, 229(7):727–746, 2015.

[211] Sparta portfolio review 2016/17, 2017.

[212] Sparta portfolio review 2017/18, 2018.

[213] Sparta portfolio review 2018/19, 2019.

[214] Sparta portfolio review 2020/21, 2021.

[215] Sparta portfolio review 2020/21, 2021.

[216] Stefan Faulstich, Sebastian Pfaffel, Berthold Hahn, and Johanna Hirsch. Performance and reliability benchmarking using the cross-company initiative windpool. In *Proceedings of the RAVE Offshore Wind R&D Conference, Bremerhaven, Germany*, volume 14, 2015.

[217] G Wilson and D. McMillan. Assessing wind farm reliability using weather dependent failure rates. *Journal of Physics: Conference Series*, 524, 2014.

[218] Gong Li and Jing Shi. Applications of bayesian methods in wind energy conversion systems. *Renewable Energy*, 43, 2012.

[219] Thomas Poulsen, Charlotte Bay Hasager, and Christian Munk Jensen. The role of logistics in practical levelized cost of energy reduction implementation and government sponsored cost reduction studies: Day and night in offshore wind operations and maintenance logistics. *Energies*, 10, 2017.

[220] Shuya Zhong, Athanasios A. Pantelous, Michael Beer, and Jian Zhou. Constrained non-linear multi-objective optimisation of preventive maintenance scheduling for offshore wind farms. *Mechanical Systems and Signal Processing*, 104:347–369, 2018.

[221] Peter Tavner and Peter J. Tavner. *Offshore wind turbines reliability, availability and maintenance / Peter Tavner.* Renewable energy series ; 13. Institution of Engineering and Technology, London, 2012. Publication Title: Offshore wind turbines reliability, availability and maintenance.

[222] Henrik Stiesdal and Peter Hauge Madsen. Design for reliability. In *Proc. Copenhagen Offshore Wind International Conference, Copenhagen, Denmark*. Citeseer, 2005.

[223] Viani Biatat Djeundje and Jonathan Crook. Dynamic survival models with varying coefficients for credit risks. *European Journal of Operational Research*, 275(1):319–333, 2019. Publisher: Elsevier.

[224] Michael J Crowther and Paul C Lambert. A general framework for parametric survival analysis. *Statistics in medicine*, 33(30):5280–5297, 2014.

[225] Frank Emmert-Streib and Matthias Dehmer. Introduction to survival analysis in practice. *Machine Learning and Knowledge Extraction*, 1(3):1013–1038, 2019.

[226] Samet Ozturk, Vasilis Fthenakis, and Stefan Faulstich. Assessing the factors impacting on the reliability of wind turbines via survival analysis—a case study. *Energies*, 11(11), 2018.

[227] DNV GL. *Definitions of Availability Terms for the Wind Industry*, 2017.

[228] International Electrotechnical Commission (IEC). *IEC 61400-26-1: Time Based Availability for Wind Turbines*, 2010.

[229] Pd iec/ts 61400-26-2:2014: Wind turbines. production-based availability for wind turbines, 2015.

[230] F. Spinato, P. J. Tavner, G. J.W. Van Bussel, and E. Koutoulakos. Reliability of wind turbine subassemblies. *Wind Energy*, 3:387–401, 2009.

[231] Iain Dinwoodie, David McMillan, Matthew Revie, Iraklis Lazakis, and Yalcin Dalgic. Development of a Combined Operational and Strategic Decision Support Model for Offshore Wind. *Energy Procedia*, 35:157–166, 2013.

[232] Iain Dinwoodie, Ole Erik Endrerud, Matthias Hofmann, Rebecca Martin, and Iver Sperstad. Reference cases for verification of operation and maintenance simulation models for offshore wind farms. *Wind Engineering*, 39, 2015.

[233] Iain Dinwoodie, David McMillan, Matthew Revie, Iraklis Lazakis, and Yalcin Dalgic. Development of a combined operational and strategic decision support model for offshore wind. *Energy Procedia*, 35:157–166, 2013.

[234] E. Gonzalez, EM Nanos, H. Seyr, L. Valldecabres, Nurseda Y. Yürüşen, Ursula Smolka, Michael Muskulus, and Julio J. Melero. Key Performance Indicators for Wind Farm Operation and Maintenance. *Energy Procedia*, 137:559–570, 2017.

[235] Fraser Anderson, Rafa Dawid, David García Cava, and David McMillan. Operational metrics for an offshore wind farm their relation to turbine access restrictions and position in the array. In *Journal of Physics: Conference Series*, 2021.

[236] Fraser Anderson, Rafa Dawid, David McMillan, and David García Cava. On the sensitivity of failure rate estimates to different failure definitions. In *Journal of Physics: Conference Series*, 2023.

[237] H. Díaz and C.Guedes Soares. Review of the current status, technology and future trends of offshore wind farms. *Ocean Engineering*, 209:107381, 2020.

[238] Elias Gyftodimos and Peter A. Flach. Hierarchical bayesian networks: An approach to classification and learning for structured data. In *Methods and Applications of Artificial Intelligence*, 2013.

[239] Tim Bedford, John Quigley, Lesley Walls, Babakelli Alkali, Alivesa Daneshkhah, and Gavin Hardman, editors. *Advances in Mathematical Modelling for Reliability*. IOS Press, Incorporated, 2008.

[240] Bruce G. Marcot and Trent D. Penman. Advances in bayesian network modelling: Integration of modelling technologies. *Environmental Modelling Software*, 111:386–393, 2019.

[241] Haiqin Wang and Marek J. Druzdzel. User interface tools for navigation in conditional probability tables and elicitation of probabilities in bayesian networks. *CoRR*, abs/1301.4430, 2013.

[242] Judea Pearl. Bayesian networks. 2011.

[243] Jeffrey N Rouder and Jun Lu. An introduction to bayesian hierarchical models with an application in the theory of signal detection. *Psychonomic bulletin & review*, 12(4):573–604, 2005.

[244] Anna E McGlothlin and Kert Viele. Bayesian hierarchical models. *Jama*, 320(22):2365–2366, 2018.

[245] Gianluca Baio and Marta Blangiardo. Bayesian hierarchical model for the prediction of football results. *Journal of Applied Statistics*, 37(2):253–264, 2010.

[246] Yu Feng, Ke Gao, and Suzanne Lacasse. Bayesian partial pooling to reduce uncertainty in overcoring rock stress estimation. *Journal of Rock Mechanics and Geotechnical Engineering*, 2023.

[247] Tae-Jeong Kim, Hyun-Han Kwon, and Carlos Lima. A bayesian partial pooling approach to mean field bias correction of weather radar rainfall estimates: Application to osungsan weather radar in south korea. *Journal of Hydrology*, 565:14–26, 2018.

[248] Mhammed Sahnoun, David Baudry, Navonil Mustafee, Anne Louis, Philip Andi Smart, Phil Godsiff, and Belahcene Mazari. Modelling and simulation of operation and maintenance strategy for offshore wind farms based on multi-agent system. *Journal of Intelligent Manufacturing*, 30, 2019.

[249] John Salvatier, Thomas V. Wiecki, and Christopher. Fonnesbeck. Probabilistic programming in Python using PyMC3. *PeerJ Computer Science*, 4, 2016.

[250] Antonio Pacifico. Robust open bayesian analysis: Overfitting, model uncertainty, and endogeneity issues in multiple regression models. *Econometric Reviews*, 40(2):148–176, 2021.

[251] Nathan P Lemoine. Moving beyond noninformative priors: why and how to choose weakly informative priors in bayesian analyses. *Oikos*, 128(7):912–928, 2019.

[252] David Draper. Bayesian hierarchical modeling, 2000.

[253] Sindhu R. Johnson, George A. Tomlinson, Gillian A. Hawker, John T. Granton, and Brian M. Feldman. The prior can often only be understood in the context of the likelihood. *Journal of Clinical Epidemiology*, 63, 2010.

[254] Rongjie Yu and Mohamed Abdel-Aty. Investigating different approaches to develop informative priors in hierarchical bayesian safety performance functions. *Accident Analysis and Prevention*, 56, 2013.

[255] Probabilistic inference using markov chain monte carlo methods. Technical report, University of Toronto, 1993.

[256] Matthew D. Hoffman and Andrew Gelman. The no-u-turn sampler: Adaptively setting path lengths in hamiltonian monte carlo. *Journal of Machine Learning Research*, 15, 2014.

[257] Michael Betancourt. The convergence of markov chain monte carlo methods: From the metropolis method to hamiltonian monte carlo. *Annalen der Physik*, 531, 2019.

[258] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, and Augusta H. Teller. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21, 1953.

[259] Stuart Geman and Donald Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence,*, 6, 1984.

[260] Andrew Gelman and Jennifer Hill. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Analytical Methods for Social Research. Cambridge University Press, 2006.

[261] William T Vetterling, William T Vetterling, William H Press, William H Press, Saul A Teukolsky, Brian P Flannery, and Brian P Flannery. *Numerical recipes: example book C*. Cambridge University Press, 1992.

[262] Generating random correlation matrices based on vines and extended onion method. *Journal of Multivariate Analysis*, 100(9):1989–2001, 2009.

[263] Richard McElreath. *Statistical rethinking: A Bayesian course with examples in R and Stan*. Chapman and Hall/CRC, 2018.

[264] Hadi Abdollahzadeh, Karim Atashgar, and Morteza Abbasi. Multi-objective opportunistic maintenance optimization of a wind farm considering limited number of maintenance groups. *Ocean Engineering*, 88:247–261, 2016.

[265] Aris Perperoglou. Cox models with dynamic ridge penalties on time-varying effects of the covariates. *Statistics in Medicine*, 33(1):170–180, 2014.

[266] Eleni-Rosalina Andrinopoulou, Paul H. C. Eilers, Johanna J. M. Takkenberg, and Dimitris Rizopoulos. Improved dynamic predictions from joint models of longitudinal and survival data with time-varying effects using p-splines. *Biometrics*, 74(2):685–693, 2018.

[267] Dexiang Gao, Gary K. Grunwald, John S. Rumsfeld, Lynn Schooley, Todd MacKenzie, and A. Laurie W. Shroyer. Time-varying risk factors for long-term mortality after coronary artery bypass graft surgery. *The Annals of Thoracic Surgery*, 81(3):793–799, 2006.

[268] Peter J Green and Bernard W Silverman. *Nonparametric regression and generalized linear models: a roughness penalty approach*. Crc Press, 1993.

[269] Akhil Prasad, Atul Manmohan, Prabhakar Karthikeyan Shanmugam, and Kothari D.P. Application of cubic spline interpolation technique in power systems: A review. In Young Kinh-Nhue Truong and Muhammad Sarfraz, editors, *Topics in Splines and Applications*, chapter 7. IntechOpen, Rijeka, 2018.

[270] Viani Biatat Djeundje and Jonathan Crook. Dynamic survival models with varying coefficients for credit risks. *European Journal of Operational Research*, 275(1):319–333, 2019.

[271] Eilers P.H.C. and Marx B.D. Flexible smoothing with b-splines and penalties. 11(2):89 – 102, 1996.

[272] Friedman J.H. and Silverman B.W. Flexible parsimonious smoothing and additive modeling. *Technometrics*, 31(1):3 – 21, 1989.

[273] Andreas Wienke. *Frailty models in survival analysis*. Chapman and Hall/CRC, 2010.

[274] Leila DAF Amorim and Jianwen Cai. Modelling recurrent events: a tutorial for analysis in epidemiology. *International Journal of Epidemiology*, 44(1):324–333, 12 2014.

[275] Aki Vehtari, Andrew Gelman, and Jonah Gabry. Practical bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, 27(5):1413–1432, aug 2016.

[276] Tuomas Sivula, Måns Magnusson, Asael Alonzo Matamoros, and Aki Vehtari. Uncertainty in bayesian leave-one-out cross-validation based model comparison, 2020.

[277] Seymour Geisser. The predictive sample reuse method with applications. *Journal of the American Statistical Association*, 70(350):320–328, 1975.

[278] Benjamin Hofner, Thomas Kneib, Wolfgang Hartl, and Helmut Küchenhoff. Building cox-type structured hazard regression models with time-varying effects. *Statistical Modelling*, 11(1):3–24, 2011.

[279] Fraser Anderson, David McMillan, Rafael Dawid, and David Garcia Cava. A bayesian hierarchical assessment of night shift working for offshore wind farms. *Wind Energy*, 2023.

[280] Fraser Anderson, Rafa Dawid, David McMillan, and David García Cava. A bayesian reliability analysis exploring the effect of scheduled maintenance on wind turbine time-to-failure. *Manuscript under review*, 2023.

[281] Michael Betancourt. A conceptual introduction to hamiltonian monte carlo. *arXiv: Methodology*, 2017.

[282] Offshore+wind+turbine+technician salary in united kingdom - average salary.

[283] Marcial Velasco Garrido, Janika Mette, Stefanie MacHe, Volker Harth, and Alexandra Marita Preisser. Sleep quality of offshore wind farm workers in the german exclusive economic zone: A cross-sectional study. *BMJ Open*, 8, 2018.

[284] Ingrid Nesdal Fossum, Bjørn Bjorvatn, Siri Waage, and Ståle Pallesen. Effects of shift and night work in the offshore petroleum industry: A systematic review. *Industrial Health*, 51, 2013.

[285] K. R Parkes. Shiftwork, job type, and the work environment as joint predictors of health-related outcomes. *Journal of Occupational Health Psychology*, 4, 1999.

[286] Valdo F. Rodrigues, Frida Marina Fischer, and Brito Mozar J. Shiftwork at a modern offshore oil rig. *Journal of Occupational Health Psychology*, 4, 1999.

[287] Ø Lauridsen and T Tonnesen. Injuries related to the aspects of shift working: a comparison of different offshore shift arrangements. *Journal of Occupational Accidents*, 12(1-3):167–176, 1990.

[288] Bjørn Bjorvatn, Kristine Stangenes, Nicolas Øyane, Knut Forberg, Arne Lowden, Fred Holsten, and Torbjørn Åkerstedt. Subjective and objective measures of adaptation and readaptation to night work on an oil rig in the north sea. *Sleep*, 29(6):821–829, 2006.

[289] Siri Waage, Anette Harris, Ståle Pallesen, Ingvild B Saksvik, Bente E Moen, and Bjørn Bjorvatn. Subjective and objective sleepiness among oil rig workers during three different shift schedules. *Sleep medicine*, 13(1):64–72, 2012.

[290] A Harris, S Waage, H Ursin, AM Hansen, B Bjorvatn, and HR Eriksen. Cortisol, reaction time test and health among offshore shift workers. *Psychoneuroendocrinology*, 35, 2010.

[291] Bjørn Bjorvatn, Kristine Stangenes, Nicolas Øyane, Knut Forberg, Arne Lowden, Fred Holsten, and Torbjørn Åkerstedt. Randomized placebo-controlled field study of the effects of bright light and melatonin in adaptation to night work. *Scandinavian journal of work, environment & health*, pages 204–214, 2007.

[292] Offshore working time in relation to performance, health and safety. Technical report, Health and Safety Executive, 2010.

[293] F. Spinato, P.J. Tavner, Gerard van Bussel, and E. Koutoulakos. Reliability of wind turbine subassemblies iet proceedings. *IET Renewable Power Generation*, 3, 12 2009.

[294] E. L. Kaplan and Paul Meier. Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, 53(282):457–481, 1958.

[295] Katharina Fischer, Karoline Pelka, Arne Bartschat, Bernd Tegtmeier, Diego Coronado, Christian Broer, and Jan Wenske. Reliability of power converters in wind turbines: Exploratory analysis of failure and operating data from a worldwide turbine fleet. *IEEE Transactions on Power Electronics*, 34(7):6332–6344, 2019.

[296] Peter Tavner and Peter J. Tavner. *Offshore wind turbines reliability, availability and maintenance / Peter Tavner.* Renewable energy series ; 13. Institution of Engineering and Technology, London, 2012.

[297] Thomas Poulsen, Charlotte Bay Hasager, and Christian Munk Jensen. The role of logistics in practical levelized cost of energy reduction implementation and government sponsored cost reduction studies: Day and night in offshore wind operations and maintenance logistics. *Energies*, 10, 2017.

[298] V. Slimacek and B. H. Lindqvist. Reliability of wind turbines modeled by a poisson process with covariates, unobserved heterogeneity and seasonality. *Wind Energy*, 19(11):1991–2002, 2016.

[299] Stefanie Muff, Andrea Riebler, Leonhard Held, Håvard Rue, and Philippe Saner. Bayesian Analysis of Measurement Error Models Using Integrated Nested Laplace Approximations. *Journal of the Royal Statistical Society Series C: Applied Statistics*, 64(2):231–252, 07 2014.

[300] Larry Wasserman. Bayesian model selection and model averaging. *Journal of mathematical psychology*, 44(1):92–107, 2000.