Debora Cevasco

Failure diagnosis for offshore wind turbines with low availability of run-to-failure data

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Engineering

Glasgow, September 2022

Supervisors: Prof. Athanasios Kolios

University of Strathclyde,
Department of Naval Architecture Ocean and Marine Engineering

Renewable Energy Marine Structures Centre for Doctoral Training,
in collaboration with

Declaration of authenticity and author's rights

Abstract

Despite the efforts to achieve a through-life reliable design and the attempts to control the failures of wind turbines, some system failures are inevitable. The inherent requirement for cost, material, and weight optimisation, together with the extreme operating conditions, can lead to unexpected failures. This is true for land-based turbines and has an even greater impact on offshore wind systems, where the harsh environment and the high cost of the assets and logistics increase the importance of a proactive approach to the system's maintenance.

The smart management of an asset starts with the identification of the health status of its systems, to take cost effective decision on how and when maintain it. The first level of the detection of an anomality in the system comprises the recognition only of the failed status of the asset (level I). Following, the location of the failure should be identified (level II), followed by the detection of its degree of severity (level III) and consequences (level IV). Depending on the availability of continuous monitoring data, historical databases, and advanced numerical models, different frameworks can be established for the failure diagnostics and prognostics. This thesis investigates on the use data-driven, model-based, and digital twin solutions to support the diagnosis of failure events of offshore wind turbine systems characterised by a low availability of run-to-failure data. This topic is of major concern for either the current installations - for which the collection of data is restrained either to only few assets or to more cost-effective temporary monitoring campaign – and the new offshore wind technologies (e.g., floating wind, large-MW structures), for which no or only a limited amount of operating data has been gathered.

The mechanical failure of the components of the offshore wind speed conversion system can have a significant impact to the operational expenditure and can be associated to a significant loss of production of the offshore wind farm. The detection of their incipiency has been extensively investigated by machine and deep learning techniques on big sets of condition monitoring and operational data. By contrast, this research explores the implementation of transfer learning to detect anomalies in an offshore wind gearbox with low availability of representative failure data. To move towards the quantification of the consequences of such a failure (level IV), a case of study is used to explore then most suitable the model-reduction techniques to be applied to a full aero-servo-elastic model of the offshore wind turbine. Such a numerical model is the basis for the development of digital twin technology; it is aimed at capturing the only the essential dynamics while targeting the degree(s) of freedom indicating the presence of the failure mode.

The presence of a damage in the offshore wind foundation is not commonly recorded, yet structural failures can either lead to catastrophic consequence or considerably increase the cost of maintenance for the planning of expensive subsea inspections. In particular, the fatigue-driven offshore wind jacket foundation designs are sensitive to extreme site conditions, and their expected lifetime can decrease considerably if exposed for a long time to undetected phenomena such as scour and corrosion. This research demonstrates the feasibility of a vibration-based diagnosis (level II) of several damage scenarios for a jacket substructure of an offshore wind turbine. Considering than only a percentage of the assets in the farm are likely to be instrumented with a high-frequency structural health monitoring system, the feasibility of the detection (level I) of a structural failure mode via low-resolution operational data is additionally explored. These virtual monitoring frameworks are supported by the deployment of the digital twin technologies for their setup and their future field application.

Keywords:

Diagnostics, Failure detection, Offshore wind turbines, Digital twin, Run-to-failure data, Structural health monitoring.

Acknowledgements

This EngD research was carried out in collaboration with the Cranfield University and the University of Oxford. It was part of the Renewable Energy Marine Structures Centre for Doctorial Training (REMS) and received funding from the UK Engineering and Physical Sciences Research Council (EPSRC), under grant agreement EP/L016303/1, for the University of Strathclyde.

Many people have contributed to the success of this project with their ideas, critical questions, and support. In particular, I would like to thank:

My supervisor **Prof Athanasios Kolios** from the University of Strathclyde and **Ursula Smolka** from Ramboll. Thank you both for the scientific and emotional support, and valuable help, especially in the last years of these research. Athanasios, thank you for your endless energy, and your passion for the scientific research. You helped me to be persistent in achieving this goal, despite all the challenges. We made it! Thank you, Ursula, for your remarkable and inspiring creativity; you guided me in defining my professional interests.

My **REMS colleagues** (across all the cohorts), from all three universities, and that now are scattered all around the world. We are a bunch of great people! I am grateful that I had the chance to share this (roller coaster) experience with you, always being ready to help each other and have a laugh together.

My **colleagues at Ramboll**. Special thanks to the core (chronologically speaking) of the asset management team. Jannis Tautz-Weinert, thank you for being such a valuable mentor: your good attitude, your curiosity towards the scientific research, and your extraordinary understanding, allowed me to achieve a lot in the last years of my research. Thank to you, Simon Scott Siedler, my personal IT Yoda: your passion in software development and knowledge sharing your is contagious. Thank you, Carolin Wendelborn, for joining forces when running the simulations, and for the warm welcome to Hamburg, and to the company.

My friends in Cranfield, Glasgow, and Hamburg. A special thanks to you, Simon, for your patience, your encouragement, and your kind advice.

My family, ed i miei genitori, Aurora e Maurizio. Grazie mille per avermi sempre supportato e sopportato in questi anni, nonostante la distanza. *La vostra scienziata dalle crêuze di Sant'Ilario.*

Table of contents

List of figures

List of tables

Nomenclature

General acronyms

AHSE	Aero-Hydro-Servo-Elastic
API	Application Programming Interface
BEM	Blade Element Momentum
BT	Balanced Truncation
CMS	Condition Monitoring System
CNN	Convolutional Neural Network
CPU	Central Processing Unit
CV	Cross Validation
DD	Direct Drive
DDE	Direct Drive configuration with EESG generator
DDP	Direct Drivetrain configuration with a PMSG
DFIG	Doubly-Fed Induction Generator
DI	Indirect Drivetrain
DLC	Design Load Case
DNN	Deep Neural Network
DOF	Degree Of Freedom
EESG	External Excitation Synchronous Generator
EOC	Environmental and Operational Conditions
ETM	Extreme Turbulence Model
FA	Fore-Aft
FCM	Fuzzy C-Means
FEM	Finite Element Model
FEMU	Finite Element Model Updating
FFT	Fast Fourier Transform
FLS	Fatigue Limit State
FMECA	Failure Mode Effect and Criticality Analysis
IoT	Internet of Things
IRS	Improved Reduced System
LDA	Linear Discriminant Analysis
LR	Linear regression
LTI	Linear Time Invariant
MDE	Modal Decomposition and Expansion
ML	Machine Learning
mr	Minor Repair
Mr	Major Repair
MR	Major Replacements
MT	Modal Truncation
MTL	Multi-Task Learning
NFA	Natural Frequency Analysis
NN	Neural Network
OCSVM	One-Class Support Vector Machine

O&M	Operation and Maintenance
OMA	Operational Modal Analysis
OP	Operating Point
OPEX	Operational Expenditure
OWEZ	Offshore Windpark Egmond aan Zee
PC	Principal Component
PCA	Principal Component Analysis
PHM	Prognosis Health Management
PMSG	Permanent Magnet Synchronous Generator
RAM	Reliability, Maintainability, and Availability
RDS-PP	Reference Designation System for Power Plants
RELu	Rectified Linear unit
RF	Random Forest
RNA	Rotor Nacelle Assembly
ROM	Reduced-order model
SCADA	Supervisory Control and Data Acquisition
SCIG	Squirrel Cage Induction Generator
SHM	Structural Health Monitoring
SHMS	Structural Health Monitoring System
SISO	Single Input Single Output
SS	Side-side
SVM	Support Vector Machine
TDT	True Digital Twin
TOW	Tower
TP	Transition Piece
WF	Wind Farm
WK64	Wikinger reference wind turbine (at location 64)
WRIG	Wound Rotor Induction Generator
WRSG	Wound Rotor Synchronous Generator
WT, WTG	Wind Turbine, Wind Turbine Generator

Metrics and performance indicators

A_E	Energetic Availability
A_O	Operational Availability
A_T	Time-Based Availability
CF	Capacity Factor
CI	Confidence Interval
CM	Confusion Matrix
DEL	Damage Equivalent Load
DT	Downtime
KPI	Key Performance Indicator
MAC	Modal Assurance Criterion
MAPE	Mean Absolute Percentage Error
MART	Mean Active Repair Time
MDT	Mean Downtime

MSE	Mean Squared Error
MTTR	Mean Time To Repair
PSD	Power Spectral Density
RMS	Root Mean Square
ROC	Receiver Operating Characteristic
RPM	Rotation Per Minute
TI	Turbulence Intensity
WS	Wind Shear exponent

Projects and research programs

CREW	Continuous Reliability Enhancements for Wind
DOWEC	Dutch Offshore Wind Energy Converter
EVW	Erhöhung der Verfügbarkeit von Windenergieanlagen
GADS	Generating Availability Data System
GRC	Gearbox Reliability Collaborative
HOME	Holistic Operation and Maintenance for Energy from offshore wind farms
LWK	Land Wirtschafts-Kammer
OREDA	Offshore Reliability Data
REMS	Renewable Energy Marine Structures centre for doctoral training
ROMEO	Reliable OM decision tools and strategies for LCOE reduction
SPARTA	System Performance, Availability and Reliability Trend Analysis
WMEP	Wissenschaftliches Mess- und Evaluierungsprogramm

Institutes and companies

CWEA	California Water Environment Association
DIN EN	Deutsches Institut für Normung, Europäische Norm
DNV	Det Norske Veritas
EPRI	Electric Power Research Institute
EPSRC	Engineering and Physical Sciences Research Council
GL	Germanischer Lloyd
IEA	International Energy Agency
IEC	International Electrotechnical Commission
IEE	Fraunhofer-Institut für Energiewirtschaft und Energiesystemtechnik
ISO	International Organization for Standardization
IWES	Fraunhofer-Institut für Windenergiesysteme
NEDO	New Energy and industrial technology Development Organization
NREL	National Renewable Energy Laboratory
ORE	Offshore Renewable Energy Catapult research centre
VTT	Valtion Teknillinen Tutkimuskeskus

Chapter 1

Importance of health management in the offshore wind industry

The IoT-enabled smart asset management plays a fundamental role in the offshore wind industry. The digital twin technology is the one of the main recurring concepts of the industry 4.0. In their model- and data-drive form, having the access to a digital copy of the physical asset helps to unveil more information from the data collected, and to support a holistic health monitoring framework. This chapter provides a brief introduction to this research's background such as the concepts of monitoring, diagnostics, and the digital twin technology. Following, it outlines the current gaps in the knowledge, which are associated to some of the most pressing question from the offshore wind industry. The aim of this research is then defined, together with its main objectives. Its contribution to the scientific community – by the means of open access publications in scientific journals and conference proceedings – is documented and linked to this thesis chapters. Finally, the structure of the remaining of the manuscript is introduced.

1.1 Background

The fast growth of the offshore wind sector not only comes with design challenges, but it also needs to cope with the planning and risk mitigation for offshore Operations and Maintenance (O&M) activities. The two main contributors to operational expenditures (OPEX) are the cost for repairing and/or replacing the assets, and the cost for logistics. It has been forecasted that, with the employment of bigger power class turbines, installed further offshore, the share of the OPEX could reach the 39% of the lifetime costs of an offshore wind farm [1]. This estimate is associated to the implementation of the traditional strategies to maintain the offshore wind assets, based on a ***corrective*** (*run-to-failure*) and a merely ***predetermined*** (*time-based*) maintenance.

A corrective maintenance strategy is based on the full utilisation of an asset until this leads to its failure – defined as the 'inability of a system or component to perform its required functions within specified performance requirements' [2] – cf. Figure 1.1. A predetermined maintenance strategy relies on the time-based planning of the maintenance actions and the inspections. This strategy has the advantage of ensuring a reliable and predictable delivery of electricity. The drawbacks, however, include the possibility of over-maintaining the assets, by replacing the components well before the end of their nominal

life, and the scheduling of the inspections at the wrong moment in time – i.e., just shortly before a failure mechanism becomes visible and/or quantifiable by in-situ assessments.

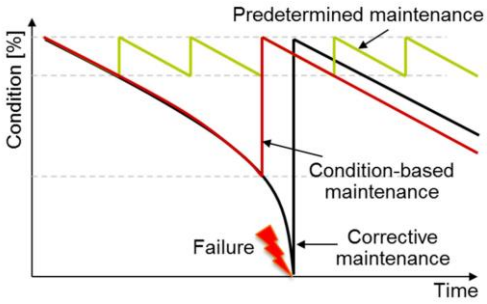


Figure 1.1: Influence of the maintenance strategy on asset condition [3].

Optimised maintenances strategies, and thus OPEX reductions, can be achieved, to varying extent, by deploying different strategies at the several stages of the offshore wind projects – cf. Figure 1.2. At the design phase, the cost of the maintenance can be controlled either by taking engineering decision on the reliability of the assets, and thus setting-up a ***reliability-centred*** maintenance strategy. To take no chances on the occurrence of a failure event, the source of the potential issues (i.e., failure cause) can be remove at the design conceptualisation phase (***design-out*** maintenance). During the operation, the better scheduling of the activities and the performance of multiple tasks per visit (***opportunity*** maintenance), can reduce the number of visits.

In the run to digitalisation, the use of the Internet of Things (IoT) enables the monitoring of the assets and their health, allowing it to switch to a ***condition-based*** maintenance. The maintenance of the turbines and their systems is scheduled on an as-needed basis. As illustrated in Figure 1.1, this strategy maximises the potential asset usage life in service, while serving to provide advanced warning of a developing failure mechanism.

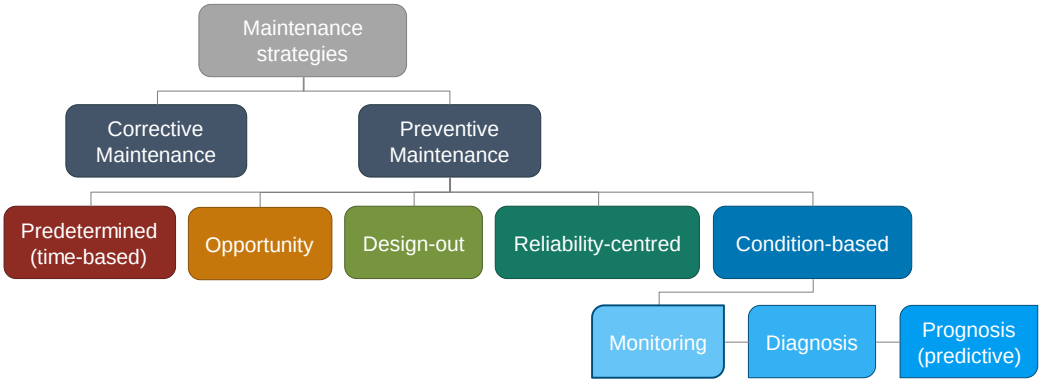


Figure 1.2: Classification of maintenance strategies, according to EN 13306 [4], and by extending it as suggested by Asmai et al. [5].

1.1.1 Health management via condition monitoring

A condition-based maintenance, as the name suggests, is achieved by knowing the actual conditions of the assets. This is enabled by applying a so-called ***Prognosis Health Management*** (PHM) framework [6]. The PHM consists of monitoring, diagnosis, and prognosis of the health of the systems – cf. Figure 1.2 – to support the decision-making process and asset management. The links between the several blocks are visualised in

Figure 1.3. The ***monitoring*** refers to the data collection process, which aim is to determine if the system is in a normal operating state. Normal behaviour models and anomaly detection techniques can help for this purpose, as long as the signals measured can give an indication of the presence of an anomaly. The ***diagnosis***, at its lowest level, refers to the identification of the presence of a failure event. Higher level of detection includes the recognition of the damage location, its corresponding degree, and the assessment of its consequences. Once the damage scenario is identified, the ***prognosis*** employs models to assess the performance degradation and further predicts the remaining useful life. Finally, the ***health management*** integrates outputs from monitoring, diagnosis, and prognosis for the decision making on the optimal maintenance and logistic strategies.

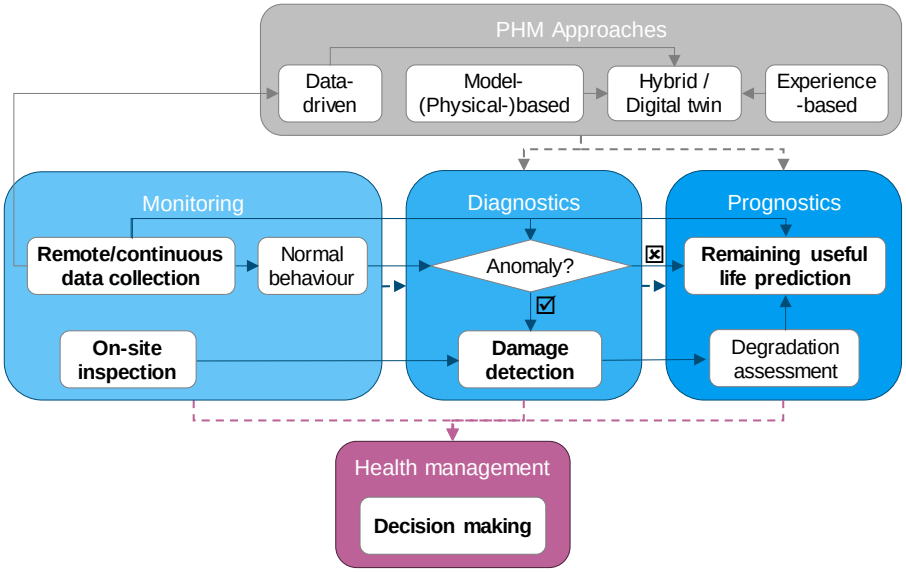


Figure 1.3: Relationship between the building blocks of PHM and approaches to support the diagnosis and prognosis on the systems health; adapted from [6], by extending it with the concepts from [5].

Several approaches have been applied for the purpose of the diagnosis and the prognosis of complex engineering systems, as described by Asmai et al. in [5]. These methods and their relationships are visualised in Figure 1.3 (in the "PHM Approaches" block). Although these approaches can be applied to the health management of systems from any field of the engineering, their description is following contextualized to the offshore wind application. It should be further noticed that the conventional classification is here broadened to include the concept of the digital twin technology.

Data-driven approaches

This approach use the routinely collected data to forecast the performances, the health and/or the operating status of the observed systems [7]. The data collected under nominal and degraded condition are associated with statistical or artificial intelligence techniques to generate an appropriate prognosis model [8]. This model is considered a "black box", because it learns patterns which exist in the data and cannot be tracked back to any physical relationship.

Statistical approaches can be divided into *parametric* and *non-parametric* approaches, depending on whether the information on the distribution of the data is accessible or not.

These methods are generally easy to apply, and they generally require less computational power than artificial intelligence methods. However, they do not easily accommodate the varying operating regimes and the high fluctuation of the environmental conditions. Additionally, they cannot capture the failure mechanism and they need for a large amount of failure data for training the detection [5].

Machine learning (ML) algorithms organize into *supervised* and *unsupervised* approaches, based upon if the acquired data contain indication of the health status. The ML models can be very flexible and adapt to the varying operating and environmental conditions. However, the goodness of their prediction is highly dependent on the quality of the data and their processing. The validation of these models, which affect the confidence in their prediction, is not easy to perform.

Model-based approach

A model, or physical, -based approach is applicable when an accurate *physical mathematical and/or numerical model* of the monitored system can be constructure to mimic its behaviour and response. The main advantage of this approach lies in its ability to provide a higher understanding of the physics of the failure [5].

However, this can be achieved, in the first place, only if the mechanistic knowledge of the system and its failure modes can be accessed and modelled. Furthermore, this approach is likely to be associated with high computational times. To capture the generally complex and holistic dynamics of big engineering systems, it is often required to couple and/or semi-couple different mathematical model [9]. Numerical models with varying *levels of fidelity* have been developed to represent the wind turbine dynamics depending on the purpose of the analysis – the reader is referred to ***Paper 7*** and ***Paper 11*** for more details. When these are deployed for the diagnosis and prognosis of a specific failure modes, it is however important that they can accurately capture the dynamics which are representative for the failure mechanism [10].

Experience-based approaches

Also known as knowledge-based approach, it is solely based on expert judgement based on *hand-code* and/or *rule-based methods* [7]. This approach does not generally require a continuous flow of data, but it rather relies on historical databases [5]. Methods such as the fuzzy logic, the Weibull distribution, the Bayesian approach, can be implemented to support the diagnosis. Despite its transparency, this approach is applicable to a simple process or system. Furthermore, the prediction of these models is generally not as accurate as the one from either the data-driven or the physical-based approach.

Hybrid approaches

A hybrid approach combines one or more of the aforementioned approaches, to offer more reliable and potentially more accurate prognostic results [7]. A framework for the application of an hybrid approach for the prognostic health management of offshore renewable assets is suggested in [11], based on the methodology developed in [8]. As it can be visualised in the workflow in Figure 1.4, this framework integrates the models of the physics of failure, the experience-based information from standards and/or historical databases, and the data-driven approach.

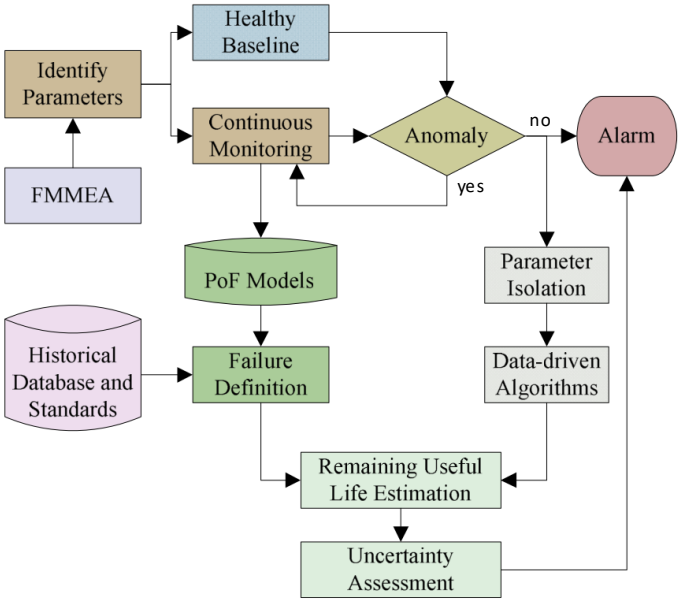


Figure 1.4: Hybrid PHM approach, adapted from Zhang et al. in [8].

Digital twin technology

The term of "digital twin" is currently applied in a multitude of forms and variety of ways – for instance refer to the discussion of Wright and Davidson [12], and reviews of Wagg et al. [13] and Worden et al. [14]. The concept used within the framework of this thesis is the so-called Ramboll's "True Digital Twin" (TDT) developed by Tygesen et al., in [15], [16], for the oil and gas industry. This technology is extended to the offshore wind industry by Augustyn [17], to be applied for the condition monitoring framework proposed by the ROMEO research project [18]–[20]. The TDT distinguishes itself from all the other digital twin concepts being based on the idea of creating a coupling between the virtual prototype – or a reduced-order model (ROM) of it – and its physical twin, by the means of calibration methods based on data from field measurements. A digital twin based on a ROM has a the advantage of the computational efficiency, while providing a high-fidelity model of the full system [21].

1.1.2 State-of-the-art of monitoring and diagnosis concepts for offshore wind turbine's systems

A key enabler of smart asset management in the offshore wind industry is the ability know the status of the wind turbine assets. The monitoring of the health, or the eventual detection of the incipiency of some failure mechanisms in the systems of the asset, is the first step towards the development of prognostics models.

In this section, the scientific literature on the monitoring and diagnosis of the main systems of offshore wind turbines is briefly reviewed, to outline the current developments and challenges. In this process, the gaps in the knowledge are identified – following highlighted in ***bold italic*** –, and consequently used to formulate the purpose and goals of this research.

Monitoring systems

The monitoring of the condition of an asset can be obtained from in-situ, non-invasive tests, and operating measurement [22]. Especially in the case of structural anomalies, the current practice for the detection strongly relies on practical on-site assessments [23]. These inspections can be associated with significant costs and risks due to the offshore environment – in the event of structural failures below sea-water level and/or at height (for the components of the rotor-nacelle assembly) [24]. For this reason, research activities have focused their effort on the investigation of remote and continuous monitoring concepts for assessing the health status of wind turbine (WT) systems [11].

To support the continuous condition monitoring of offshore WTs, a great variety of sensing devices and several analysis methods can be implemented [23]. These have been extensively reviewed in [22], and classified in [3] as follow:

- *local* or *direct sensing*, aimed at direct measurement and strongly dependent on the monitored structure;
- *global* or *indirect sensing* (also called *vibration-based*), for the indirect detection of deviations in the response of the system, and applicable to any type of structure.

Direct sensing for offshore WTs can include, for instance, the use of strain gauges for the local monitoring of fatigue damage, the installation of sonar sensors for the monitoring of the scour phenomena, but also the non-destructive test evaluations to be performed on-site [25]. Although the post-processing of the measured signals for the detection of anomalies is relatively straightforward, the cost for the installation and maintenance of this ad-hoc sensor, and/or the performing a visit to the WT, is generally high for offshore applications [22]. Furthermore, it is not easy to scale these monitoring strategies to the whole WT structure, and to the other assets in the farm [3].

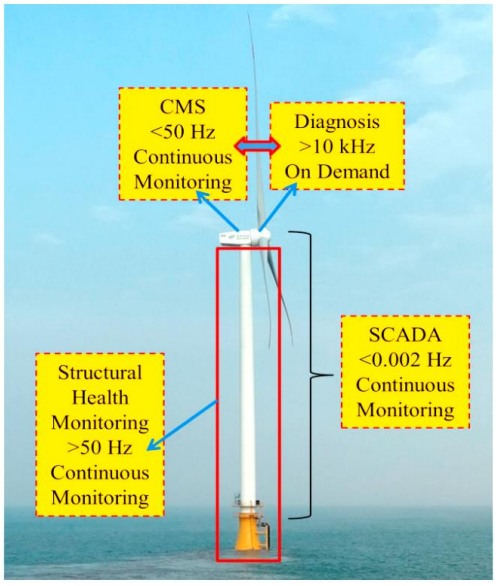


Figure 1.5: Condition monitoring systems for offshore wind turbines [23]. These systems are, by listing them system from low to high frequency signals: SCADA, SHMS, CMS, and additional diagnostic data.

Indirect sensing, which make use of vibrational data for the monitoring of the system global dynamics, is relatively cheaper and more flexible than direct sensing. Accelerometers are generally used for this purpose, in alternative to strain gauges, being more practical for installation and removal, and unaffected by the temperature [26]. A set of accelerometers is distributed on a structure, and when this is excited by environmental loads, the vibrations

are measured over time and are used to extract modal properties - e.g., mode shape, natural frequencies and damping on a global scale.

A set of direct and indirect sensing devices constitutes a monitoring system. The monitoring system of the structural components of a WT, such as the support structure or rotor blades, is often referred to as *Structural Health Monitoring system* (SHMS). In contrast, the set of sensors use for the WT mechanical and electrical components, i.e., the ones of the drivetrain, is termed *Condition Monitoring system* (CMS). The SHMS and CMS often record high-frequency vibrational measurements, and they are installed on a limited amount of assets, depending on the project specific necessity, classifications requirements, and possible country regulations [3]. The operating and environmental conditions are recorded in a *Supervisory Control and Data Acquisition* (SCADA) system, which often have a 10-minute resolution. The SCADA data collection can increase to 1 Hz frequency if required for some specific tasks – e.g., WT control strategy. The reader is referred to Figure 1.5 for a visualisation of the location of these systems.

Challenges of detection methods based on indirect monitoring

When relying on monitoring system recording vibrational data from SHMS and CMS, it should be noticed that the detection of anomalies is challenged by the fact that:

- the recording of changes in the dynamic behaviour is potentially associated with several factors and/or coexisting modes of failure of the structures – e.g., [22],
- the application of system identification techniques – for the extraction of the physical properties of the structure (i.e. modal frequencies and damping, mode shapes, and stiffness) – on operational systems in nonstationary conditions is not always applicable [21], and it has some intrinsic uncertainties [27], [28]

To cope with this latter issue, *non-parametric* methods can be deployed to detect the damage from the vibration measurements. These methods include time series analysis and statistical methods for the detection of the damage from features that are not attributed to physical changes of the structure.

However, due to the complex dynamics of offshore WT structures, the *parametric* vibration-based methods are still the prime focus of the research. Consequently, extensive effort has been made to improve the system identification methods. In the context of offshore wind applications, an operational (output-only) modal analysis (OMA) approach is often used to derive the system's modal properties. To capture a sufficient number of modes of vibration, a number sensors should be installed at several levels and/or locations [29]. An optimal sensor placement study can be conducted to achieve this in a cost-effective way; for instance, Richmond et al. [30] demonstrated that, for an offshore support platform, this can be done without sacrificing quality of modes. Data normalisation and clustering are generally applied on the raw measurement and/or the OMA extracted modes to reduce the effect of operational conditions [31], [32]. Novel approaches and techniques are continuously proposed to improve the accuracy of the identification results. In [33], [34] researchers investigated a modified stochastic subspace system identification. In [35] other authors suggested the use of data from the blades to improve the observability of aerodynamically damped global modes.

Current practise and knowledge gaps

This section briefly reviews the current practise for the monitoring of the mechanical components of the drivetrain and the structural components of the offshore wind turbine

substructures. The scope is to identify the gaps in the knowledge and, consequently, to formulate this thesis' research aim and objectives.

Health monitoring of drivetrain components

The components of the drivetrain, such as those of the power generation and conversion, and the electronics of the control systems, have been of most concern due to their generally low reliability [36]. This has motivated the investigation of their failure modes – e.g. [37] –, of virtual prototypes for their condition monitoring – e.g. [38] –, and the development of novel approaches to detect their abnormal behaviour – e.g. [39].

Due to the high costs, and thus generally low data availability, of the dedicated CMS, the detection of failure events in the drivetrain systems of installed assets has been extensively investigated by using of the SCADA system data [40]. In contrast to the CMS, which often record temperature and vibrational measurements, the SCADA system captures information from various systems dynamics. Additionally, these data are generally collected for every asset in the farm and with a resolution varying from 0,002 Hz to 1 Hz.

Thanks to the availability of these big datasets, data-driven detection approaches have been extensively applied to the signals from the operational turbines, as revised and discussed in [41]. The majority of the studies based their analysis on less than 30 turbine years of SCADA data [40]. To the knowledge of the author, no study applied a data-driven approach for the damage detection task on very small datasets – i.e., for less than 3 turbine years. This issue affects new installations – generally associated to new wind turbine technologies, which are increasingly more complex and bigger –, for which only a very limited amount of data might be available.

As opposed to data-driven approaches, high fidelity physical models, which can represent the full dynamics with respect to the failure mode, can be created to capture the physics of failure of the system and calculate their fatigue consumption. Several researchers analysed the physics of failure of the drivetrain mechanical systems by simulating the complex kinematics and suggesting alternative ways of capturing their malfunctioning, such as in [38], [42]. These studies were done for simplified simulation of turbine loads. To couple the complex aero-hydro-servo-elastic (AHSE) dynamics of the offshore wind turbine's code to these specialised virtual prototypes, their model-order should be reduced to find a good compromise between the accuracy and the computational speed.

To move towards the development of a digital twin of an offshore wind farm, which is able to capture the mechanism of the failure, and the presence of a damage in the drivetrain systems, model-order reduction techniques of the turbine dynamics need to be investigated. Most of the research work on this matter has been aimed at reducing the complex, multi-body model of the specific components, e.g., [21]. However, to manage to deal with several wind turbine models and their interaction across the location (via their aerodynamics), the reduction approaches should act on the AHSE code of dynamics.

Monitoring of offshore wind structures

With regard to structural failures, the damage to the turbines' blades (erosion and delamination) are relatively common in the offshore environment [43]. The vibration-based monitoring of the wind turbine blades is generally supported by testing in a laboratory environment. Examples of such a setup are in [44] for a parametric, and in [45] for a non-parametric detection approach. If experiments are not possible and/or only a limited number of accelerometers is available, the authors of [46] proposed a non-parametric

method based on the analysis of the timeseries of the blades' edgewise frequencies via the statistical control of their residual errors from a Gaussian-process prediction.

On the other hand, although structural failures of towers and foundations are relatively unlikely, due to design conservatism, their presence could result in dramatic consequences if undetected [47]. It might also not always be possible to design the support structure in an inspection-free manner; thus, it is sometimes necessary to perform potentially expensive inspection activities during its lifetime. Therefore, the ability to detect and reliably monitor structural anomalies and failures can have a huge impact on the maintenance costs and on the decision-making of lifetime extension strategies [48]. For the investigation of the monitoring of offshore wind tower and foundation structures, it is neither easy to scale the assets and the representative loads for laboratory testing, nor access run-to-failure data from field applications, since they are rare in these safety-critical systems. Some speculations have been made on the feasibility of the detection of structural anomalies based on the trends and the scatter identified from vibration-based measurements from the field [49], [50]. Most of the proposed methods identify the presence of anomalies by tracking the natural frequencies, the mode shapes, and their derivatives, such as the displacement modal curvature [26]. In [51], [52], the authors adopted a model-based approach to investigate the feasibility of the detection of structural anomalies of different nature. In [51], they analyse the deviations in the modal parameters (frequencies and mode shapes) caused by the three damage types implemented in a finite element model of a gravity-based wind turbine tower. However, they conclude that the change induced by these failure modes in the global properties is likely be hidden from the variation of wave loadings. In [52], they propose the use of a vibration-based artificial neural network for the estimation of location and severity of simulated structural damage in onshore turbine towers. The main findings outlined that a detection algorithm trained on frequencies only performed better for the assessment of the severity.

For what concerns the offshore wind jacket substructures, the natural frequencies of the global modes have been shown to have a low sensitivity to structural damage (e.g. member loss, scour, corrosion); in contrast, the mode shapes have greater sensitivity to damage [53]. The feasibility of the vibration-based diagnostics methods for offshore wind jacket support structures should be further investigated, to prove the viability of the monitoring systems to detect and distinct the failure events of different nature.

However, it should be noted that the assessment of a mode shape requires the equipment of high-resolution sensors on the support structure and tower of the wind turbine. The installation of global monitoring systems is generally recommended but done only on a fraction of the assets in the farm. Besides, the issues related to the application of traditional OMA to wind turbine dynamics still impose, to date, difficulties in utilizing mode shapes to capture a damaged status in practical applications - refer to [50], [54].

Low-resolution statistics from the SCADA system data have been already employed for the prediction of the fatigue loads acting on the WT tower and substructure [55], [56]. Therefore, these data-driven models should be able to monitor the presence of a structural damage and an extreme environmental condition, especially in the case these events affect the fatigue life of the structures. The feasibility and the potential applicability of such approach based on the use of SCADA data statistics for the detection of structural anomalies has not been explored yet.

1.2 Research aim and objectives

The aim of this thesis to investigate diagnostics approaches for offshore wind turbine systems, based on data-driven, model-based and hybrid methods, with an eye towards the digital twin technology. The contribution to the knowledge is the investigation of the detectability of abnormal behaviour and failure events in systems whose run-to-failure data are limited and/or cannot be collected.

This thesis delivers on the following objectives:

1. Review and classify, in a comprehensive study, the reliability, maintainability, and availability (RAM) data from onshore and offshore wind turbines, to identify the systems holding the highest criticalities and trends based on deployment parameters and environmental conditions of offshore wind turbines.
2. Investigate model-reduction techniques for a numerical model of wind turbine dynamics by targeting a drivetrain failure mode, to support the future development of digital twin technology for its remote monitoring.
3. Experiment a data-driven approach for the detection of anomalies in a component of the drivetrain system with low availability of run-to-failure data.
4. Demonstrate the feasibility of vibration-based detection of several damage scenario in the substructure of an offshore wind turbine, with the support of the digital twin technology, and develop a framework for the future field application.
5. Explore the feasibility of the detection of a structural failure via low-resolution data and develop a framework for the future field application with the support of the digital twin technology.

1.3 Scientific publications and collaborations

This thesis is composed of a portfolio of works, that have been published in peer-reviewed scientific journals and conference proceedings. The main publications, whose material is used for compiling this thesis, are following listed in chronological order.

Paper 1 **D. Cevasco**, S. Koukoura, A. J. Kolios, (2021), "Reliability, availability, maintainability data review for the identification of trends in offshore wind energy applications", *Renewable Sustainable Energy Reviews*;136. https://doi.org/10.1016/j.rser.2020.110414

Paper 2 Z. Lin, **D. Cevasco**, M. Collu (2020), "A methodology to develop reduced-order models to support the operation and maintenance of offshore wind turbines", *Applied Energy 2020.* https://doi.org/10.1016/j.apenergy.2019.114228

Paper 3 **D. Cevasco**, J. Tautz-Weinert, U. Smolka, A.J. Kolios, (2020), "Feasibility of machine learning algorithms for classifying damaged offshore jacket structures using SCADA data", presented at *EERA DeepWind2020* (Trondheim, Norway), published in *IOP Conference Series: Journal of Physics.* https://doi.org/10.1088/1742-6596/1669/1/012021

Paper 4 **D. Cevasco**, J. Tautz-Weinert, A.J. Kolios, U. Smolka, (2020), "Applicability of machine learning approaches for structural damage detection of offshore wind jacket structures based on low resolution data", presented at *The*

Science of Making Torque from Wind: TORQUE 2020 (Delft, Netherlands), published in *IOP Conference Series: Journal of Physics.* https://doi.org/10.1088/1742-6596/1618/2/022063

Paper 5 **D. Cevasco**, J. Tautz-Weinert, M. Richmond, A. Sobey, A.J. Kolios, (2022), "A damage detection and location scheme for offshore wind turbine jacket structures based on global modal properties", *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering.* https://doi.org/10.1115/1.4053659

Paper 6 I. M. Black, **D. Cevasco**, A. Kolios, (2022), "Deep neural network hard parameter multi-task learning for condition monitoring of an offshore wind turbine", presented at *The Science of Making Torque from Wind: TORQUE 2022* (Delft, Netherlands), published in *IOP Conference Series: Journal of Physics.* https://doi.org/10.1088/1742-6596/2265/3/032091

Paper 7 M. Richmond, **D. Cevasco**, A. Kolios, (2022), "A review of modelling methods for arrays of offshore wind turbines", *Submitted to Renewable Sustainable Energy Reviews.*

The co-authors played any of a variety of roles in the preparation of these manuscripts. To reduce the ambiguity, and to achieve greater clarity in contributions of all authors, the contributions matrices are reported in Section A.3 of the Appendices, by using nowadays standard for the taxonomy of the contribution roles (so called "CrediT").

Other collaborative publications were achieved by contributing to the investigation on a holistic model of dynamics for the HOME-Offshore project [9]. These publications are following listed and only cited in the course of the thesis.

Paper 8 **D. Cevasco**, M. Collu, Z. Lin, (2018), "O&M cost-based FMECA: identification and ranking of the most critical components for 2-4MW geared offshore wind turbines", presented at *Global Wind Summit: WindEurope 2018* (Hamburg, Germany), published in *IOP Conference Series: Journal of Physics.* https://doi.org/10.1088/1742-6596/1102/1/012039

Paper 9 Z. Lin, **D. Cevasco**, M. Collu, (2018), "Progress on the development of a holistic coupled model of dynamics for offshore wind farms, phase I: aero-hydro-servo-elastic model, with drive train model, for a single wind turbine", presented at the *37th International Conference on Ocean, Offshore and Arctic Engineering* (Madrid, Spain), published in the *Proceedings of the ASME OMAE2018* - 77886

Paper 10 Z. Lin, A. Stetco, J. Carmona-Sanchez, **D. Cevasco**, M. Collu, G. Nenadic, O. Marjanovic, M. Barnes, (2019), "Progress on the development of a holistic coupled model of dynamics for offshore wind farms, phase II: study on a data-driven based reduced-order model for a single wind turbine", presented at the *38th International Conference on Ocean, Offshore and Arctic Engineering* (Glasgow, Scotland), published in the *Proceedings of the ASME OMAE2019* - 95542

Paper 11 J. Carmona-Sanchez, Z. Lin, M. Collu, M. Barnes, O. Marjanovic, **D. Cevasco**, (2019), "An analysis of the impact of an advanced aero-hydro-servo-elastic model of dynamics on the generator-converter dynamics, for an offshore fixed 5MW PMSG wind turbine", presented at the *15th IET International*

Conference on AC and DC Power Transmission (Coventry, United Kingdom), published in the *Proceedings of the ACDC2019.*
https://doi.org/10.1049/cp.2019.0080

A list of additional publications, which were prepared in the context of the EngD programme, and related to other academic collaborations, is provided in Section A.1 of the Appendices, together with a list of the oral and poster conferences presentations.

1.4 Thesis outline

The remaining manuscript is structured as follows:

Chapter 2 assesses and discusses the criticality of current and future offshore wind turbine systems, by integrating the findings from an examination of the literature on RAM data with wind farm availability simulations. Challenges and limitations associated with the collection and use of offshore RAM statistics are finally provided. The results of this investigation have been already published in Paper 1.

Chapter 3 provides case studies for tackling the challenges related to the monitoring of offshore wind drivetrain components. The findings of Paper 8 are used to support the problem statement. Based on the work of Paper 7 and Paper 2, model-order reduction techniques are investigated towards the development of a digital twin technology to monitor failures of the gearbox. This chapter continues with the investigation of data-drive techniques for detection of gearbox anomalies based on the low availability of failure data, as presented in Paper 6.

Chapter 4 introduces to the models - a virtual prototype, and a digital twin model – of the jacket substructures of an offshore wind turbine, applied for the investigation on the detection of the structural failure events of the following chapters. It additionally assesses the impact of the various damage scenario on the structure dynamics and its remaining useful life, by running natural frequency analysis and fatigue limit state simulations.

Chapter 5 present a case study to demonstrate the feasibility of detection of structural anomalies of different nature in an offshore wind turbine jacket substructure, by the support of the digital twin technology for the simulation of the vibration-based measurements. This chapter is based on the work of Paper 5.

Chapter 6 provides a case study for the detection of structural anomalies in offshore jacket substructures based on low-resolution data from the operating wind turbine. This detection approach is supported by the used of the wind turbine models. The virtual prototype of the turbine is used to prove the feasibility of the monitoring, which has been documented in Paper 3. Then, as presented and discussed in Paper 4, the digital twin technology is deployed to outline the challenges and possibility of future field applications.

Chapter 7 discusses the main findings of each chapter, by additionally extending on their challenges. It additionally provides recommendations, for each of the case study, on the readiness of the technologies to the field applications.

Chapter 8 concludes the thesis by presenting a summary of the main findings, giving advice on future works, and posing the contribution to the knowledge of each of the studies conducted in this thesis.

Chapter 2

Review of RAM databases and identification of failure trends

The fast growth of the offshore wind sector not only comes with design challenges, but it also needs to cope with the planning and risk mitigation for offshore maintenance activities. For this reason, research activities have focused their effort on investigating what, and how, failure events manifest in a wind turbine system. In this chapter, the initiatives, and the findings of the historical collection of reliability, availability, and maintainability (RAM) data are presented. At first, a systematic review of the figures for offshore and onshore wind projects is offered. Trends based on the deployment parameters are then extracted, to speculate on the impact of design characteristics and environmental conditions on the systems' RAM. Finally, an O&M simulation tool is deployed to estimate the operational availability of a next-generation offshore wind farm. The material of this chapter has been peer reviewed and published in a peer-reviewed journal [1].

2.1 Background

Since the early-stage wind farms, a considerable effort was made in collecting indicators for their reliability, maintainability and availability (RAM) statistics and putting them into databases [57]. National and international initiatives were mainly directed at creating repositories for onshore wind turbines [58]. Only recently have some initiatives focused on the collection of RAM data from modern [59], [60] and/or offshore systems [61], [62]. Although large and heterogeneous, the populations of some of the most well-known campaigns (e.g. [63], [64]) generally include statistics of outdated configurations and small rated wind turbines, compared to modern installations and offshore trends. Nonetheless, the collection of historical data has shown to be useful for benchmarking critical

[1] **D. Cevasco**, S. Koukoura, A. J. Kolios (2021), "Reliability, availability, maintainability data review for the identification of trends in offshore wind energy applications", *Renewable Sustainable Energy Reviews*;136. https://doi.org/10.1016/j.rser.2020.110414

components to support monitoring concept development and a systematic service life performance analysis.

For the purpose of this literature review, only fully operational data are collected, rather than test data of single components – i.e. gearbox reliability collaboration [65] and blade reliability studies [66], [67]. Starting from a chronological overview of the main initiatives for onshore systems, the main findings are then outlined for the more recent offshore data collections. Despite including insights on offshore risks, the industry-driven databases are generally not available to the public for confidentiality reasons. Consequently, publications from other independent authors are reviewed, to obtain an appreciation of the RAM experience of various offshore wind farms installed in European waters.

2.1.1 Chronological overview to data collection initiatives

Onshore data collection and statistics

One of the first reliability databases for onshore turbines was compiled by Lynette [68], who analysed the trend in availability and costs for the maintenance of various types of small-scale units installed in California until the end of the 1980s. Similarly, the Electric Power Research Institute (EPRI) collected, during 1986 and 1987, failure data for a portion of its Californian population [57]. These statistics were reported by the DOWEC (Dutch Offshore Wind Energy Converter) research program in [69], which was one of the pioneers in the documentation of wind turbine reliability figures. Due to the outdated technologies of the Californian population, these data were not integrated in their comparative study. They instead cross-analysed the yearly statistics, recorded around the end of the 1990s and first years of the 2000s, of the largest European data collection campaigns – the German and Danish Windstats newsletter, and the German LWK (Land Wirtschafts-Kammer) and WMEP (Wissenschaftliches Mess- und Evaluierungsprogramm) databases. By plotting the results per size classes (within the single initiatives) and across the databases, the DOWEC team simply observed a significant scatter in the trend, justifying through the unknown age and the outdated type of some installations the causes of their statistics' discrepancy. In contrast, some years later, Ribrant et al. [70], [71] outlined some similarities in the share of the components to the turbines' failure rate and downtime, when comparing the Swedish and Finnish databases – Vindstat and Felanalys, and VTT, respectively – with the current WMEP results.

In the same period, a substantial contribution to the collection and understanding of the turbine reliability statistics was made by Durham University (UK) and Fraunhofer IWES (Germany). For the first time, Tavner et al. [72], [73] analysed and compared the time-trend of the reliability results from EPRI, Windstats, LWK and WMEP, plotting them against the historical data from other industrial turbines. Tavner et al., compared the components' failures of the LWK population, grouping them by layout and size [63], [72], [74]. Furthermore, Tavner et al. investigated the effect of weather and location on the turbines' failures, drawing some preliminary conclusions on their correlation. In [75], [76] they observed a periodicity in the failure frequency for some of the Danish Windstat components, while in [77] they investigated the possible dependency of failures on the wind farm location for a population of Enercon E30-33 turbines. In parallel, and consistently with what is shown by [31]–[34], Faulstich et al. analysed the effect of the turbines' configuration and location (onshore, coastal, and offshore) on the reliability figures of the WMEP database

[78], [79]. In [80], they additionally explored the possible link between WMEP failures and wind speed, developing further the first observations of Hahn [81].

From the experience of the above-mentioned databases, and because of the increasing number of wind farm installations in Europe, more structured RAM data collections were launched. Aimed at moving towards a design-for-reliability approach, and targeting improved condition monitoring techniques [64], the European ReliaWind project ran for three years from 2008. Starting from reviewing previous European projects (EUROWIN and EUSEFIA [58]) and national-level initiatives, this project collected and analysed heterogeneous data from the turbines operation and maintenance (O&M) activities, based on the joint effort of industry (e.g. Siemens Gamesa), technical experts (Garrad Hassan, now DNV GL) and academia (Durham University, among others). Contemporarily, in Germany, the Fraunhofer Institute continued the WMEP database activity in the *"Increasing the availability of wind turbines"* (EVW) project [82], [83], which ended in 2015. Despite these projects representing two of the most recent and complete databases from the European experience, due to confidentiality issues, only their final reports and relative values of the total statistics are accessible.

From the first decade of the 21st century, other data collection initiatives from the rest of the world have contributed to documenting and tackling the reliability of onshore installed systems. Academic and industrial researchers in India, China and Japan published the first reliability and availability statistics reports for specific wind farms [84]–[86] and turbine manufacturers [59]. The CREW Database and Analysis Program [87], in the USA, which is an ongoing activity coordinated by Sandia Laboratories, has become more extensive and structured.

Offshore data collection and statistics

Little failure data exist in the public domain for offshore wind systems. Performance from UK's offshore round 1 wind farms, with evidence of wind farms' availability indicators (see Appendix B) and capacity factors (*CF*), were first reported by Feng et al. [88]. In this work, as in [61], maintenance records and operational issues of four selected wind farms were analysed. Similarly, the reliability figures for the Egmond aan Zee offshore wind farm (OWEZ) were derived by Crabtree et al. [89] by accessing the operational report from Noordzee Wind [90], for the first three years from installation. In [89], they additionally updated the results from the early experience of round 1 wind farms, which were affected by technological-immaturity failure events. Besides, they collected the performance indicators of round 2 wind farms, showing a growth in the average *CF* for the more modern offshore wind turbines, in line with results presented by the SPARTA [91] and Offshore-WMEP [92] projects.

With regard to reliability and maintainability data, one of the most complete contributions is the dataset published by Carroll et al. [62], for a population of 350 offshore wind turbines. Despite the results presented being from a single manufacturer, the detailed definition of the failure, and the further results on the repair time, material costs, and required technicians per subassembly, are provided.

In terms of availability, onshore wind turbines are shown to reach values in a range of 95-97% for modern systems [58]. For offshore projects, however, the location and associated challenges (i.e., accessibility and exposure to extreme weather conditions) can considerably lower availability. As observed by [91], [92], older farms – comprising turbines with relatively low nominal capacity, and relatively close to the coast – exhibit an

availability in the range of the onshore average one. Newer farms – bigger and generally located further from the coast – are characterised by an increase in maintenance efforts [1]. Despite the higher *CF*s [89], the technical availability (A_E), – as defined in Section B.1 of the Appendices – of offshore wind farms can fluctuate across the years, depending on the distance from shore and hence the ease to perform the required maintenance operations [60]. Additionally, the fluctuation of A_E among several surveys can be related to the varying maintainability of components for the different wind turbines' concepts and designs [93], [94].

Industry-led databases

Following the compilation of the first databases (e.g. LWK, WMEP), and recognising the limits of these earlier data collection initiatives, several authors [87], [95], [96] suggested possible improvements to the techniques and processes used for gathering and analysing data. Among these, Hameed et al. [97] proposed an optimal RAM database to be applied to offshore wind turbines, after having identified the shortcomings of the historical databases from the onshore wind and the offshore oil and gas (Offshore Reliability Data, OREDA) industries. The IEA Wind Task 33 started compiling recommended practices in [98], observing how the lack of standards associated with reliability data adversely impacts industry. In line with these conceptual examples are two recently launched RAM databases for onshore and offshore systems:

- The SPARTA (System Performance, Availability and Reliability Trend Analysis) initiative [99], started in 2013 by The Crown Estate (UK) under the supervision of the Offshore Renewable Energy (ORE) Catapult research centre. SPARTA is gathering KPIs (at wind farm level) and reliability figures (at subsystem level) from the participating operators, outputting a monthly benchmark.
- The German equivalent, WInD-Pool (Wind-Energy-Information-Data-Pool), with Fraunhofer IEE as trustee [100]. It can be seen as the successor to WMEP [78], where additional (but never published) information on the cost of the maintenance services was collected. It continues and merges the EVW (Erhöhung der Verfügbarkeit von Windenergieanlagen) [82], [83] and Offshore-WMEP [101]–[103] research projects, gathering historic and recently collected data for both onshore and offshore wind turbines.

2.1.2 Recent review effort

The onshore and the few offshore available data were already analysed, and cross compared by several authors. In 2011, Sheng and Wang [104] compiled the first extensive survey of the various databases available until that year. Three more recent studies have significantly contributed to gathering and comparing the data available until the year 2018. Pfaffel et al. [58] presented a comprehensive collection of the to-date available RAM statistics (a total of 23, of which 20 are onshore), updating the historical data comparison initiatives with the results from offshore wind farms, and including datasets from outside the European continent. The failure frequencies and downtime are presented in an all-in-one comparison according to standardised key performance indicators (KPIs), in their normalised and non-normalised form. Artigao et al. [105] cross-compared some of these reliability statistics (13 initiatives, of which two are offshore) with the purpose of identifying the critical components of wind energy converters across all technologies, sizes and locations, in order

to suggest condition monitoring strategies, techniques and technologies. Similarly, Dao et al., in [106], used the averaged statistics of [58] to visualise the trends in reliability and maintainability figures of offshore wind turbines, as opposed to the onshore ones, and assess their impact on operational cost, to assist operators to identify the optimal degree of reliability improvement to minimise the levelized cost of energy.

2.2 Problem statement and aim

Previous studies identified some trends in the averaged data, depending on the survey location (on- and offshore, and governing country), population size and mean power rating (e.g. [106]). However, two main issues related to the use of cumulative statistics still subsist:

- As highlighted by Leahy et al. [107], and previously by Sheng and O'Connor [108], the currently accessible RAM data lack a harmonised practice for their collection, processing and publication. The absence of standardisation in the type of data, and methods for their collection, leads to different levels of data quality. Furthermore, it is challenging to compare data among studies if project-specific and/or undocumented taxonomies are employed for the systems and subsystems characterisation.
- As noticed in [57], [109], technologies of different maturity, in different operating years, are expected to fail differently. The deployment location and the varying environmental conditions can play an important role in the lifespan reliability of the wind turbine systems and subsystems [75], [110], [111]. When comparing the statistics in terms of averages among the heterogeneous population [58], [105], [106], this level of detail is not considered.

To tackle the first issue, Leimeister at al. [109] recognised that fuzzy set and/or evidence theories can help deal with the uncertainties of vague data. Nonetheless, these methods cannot cope with the same level of detail and information as for a RAM database. In 2011, the Continuous Reliability Enhancements for Wind (CREW) Database and Analysis Program, supported by the US government, introduced a consistent approach for the collection of high-resolution supervisory control and data acquisition (SCADA) data with the aim of characterising the reliability and performance of the country's fleet. Motivated by the standardisation intent, the members of IEA Wind Task 33 created, in 2013, the *"Reliability Data Standardization of Data Collection for Wind Turbine Reliability and Operation & Maintenance Analyses: Initiatives Concerning Reliability Data"*. Similarly, industrially-led repositories are currently collecting data with a higher level of detail, adopting structured and harmonised procedures to accommodate the different types of data from wind farms [99], [100], while providing sufficient information for a consistent comparison of the structure per typology, age and location. With regard to the second point, some research effort was put into finding a correlation between the turbines' failure rates and the associated environmental conditions [75], [77]. In this regard, Barabadi et al. [112] suggested a methodology for RAM data collection of engineering structures in Arctic conditions, showing its applicability to the offshore and marine industries.

Having established the need for more representative RAM databases, this chapter aims to perform a comprehensive review of existing published data related to the reliability, the availability, and maintainability of onshore and offshore wind turbines, with a view to critically discussing commonalities and distinguishing correlation aspects between modern and more early-stage assets.

2.3 Methodology

To consistently compare the RAM statistics and unveil potential trends, three methods are employed in this study. First, a cataloguing activity is used to gather the information available from the literature by using a standardised terminology for: the statistics, the type of data, and the wind turbine typology. This process facilitates access to the RAM figures and allows the evaluation of the completeness and quality of the data collected by each initiative. Next, the adaptation of RAM repositories' data to a unique taxonomy based on the most widely adopted reference designation permits a fair comparison of the data across initiatives. Finally, calculations are carried out to uniformly compare the onshore and offshore reliability and maintainability data in terms of operational (time-based) availability for a hypothetical offshore wind farm scenario.

2.3.1 Statistics cataloguing

Categories and convections

To compile a comprehensive catalogue of the most important (and accessible) information for onshore and offshore RAM statistics, it is first necessary to identify what characteristics are worth being collected. The database and the population size are usually reported to provide an indication of the statistical significance of the data collected. However, consistency issues when discussing and comparing the results can arise from the lack of sufficient details on the population of wind turbines [69].

Table 2.1: Wind turbine configuration types, adapted from the classification of [113], and integrated with the details from [114], [115].

Concept	Type	Sub-Type		Speed	Control	Gearbox	Generator	Grid connection
Danish	**A**	**A0**		Fixed (dual)	Stall [1]	Multi-stage [3]	SCIG	Capacitor
Advanced Danish	**A**	**A1** **A2**		Fixed	Stall [2] Pitch	Multi-stage [3]	SCIG	Capacitor
Variable-speed	**B**			Limited Variable [4]	Pitch	Multi-stage	WRIG	Capacitor
Variable-speed	**C**			Variable	Pitch	Multi-stage	DFIG	Partial-scale power converter [5]
Variable-speed	**D**	**DI**	**DImW** **DImS** **DImP**	Variable	Pitch	Multi-stage	WRSG SCIG PMSG	Full-scale power converter
Variable-speed	**D**	**DI**	**DI1P**	Variable	Pitch	Single stage	PMSG	Full-scale power converter
Variable-speed	**D**	**DD**	**DDP**	Variable	Pitch	None	PMSG	Full-scale power converter [6]
Variable-speed	**D**	**DD**	**DDE**	Variable	Pitch	None	EESG	Full-scale power converter [6]

[1] passive stall regulation
[2] active stall regulation
[3] generally four-stages and up to two-stages gearbox
[4] with a variable resistance in the rotor windings
[5] converter feed back to the generator
[6] double feed back to the generator

For this reason, the classification applied here differentiates the turbines by:

- ***Power rating.*** As several studies have already shown (e.g. [69], [116], [117]), the number of failures (per turbine and/or per component) can be dependent on the

dimension of the turbines, making it necessary to present the results by power class grouping.

- ***Age of the installation.*** It is common knowledge that the failure rate is a time-dependent variable [76]. Therefore, the age of the system, whenever known, is reported to account for the possible influence on the population statistics (e.g., infant mortality and end-of-life wear out failure events).
- ***Technical concepts and drivetrain configuration.*** A deep understanding of the results also comes from the knowledge of the type of structure and configuration analysed. It is proposed here to identify the drivetrain layouts according to the concept classes, as presented in Table 2.1. These configurations are similar to those of [113]; however, the sub-types acronyms are arranged to meet the concepts used by the WMEP [115] and LWK projects [74]. The generators' description and acronyms are retrievable from [118] and [119], respectively. Schematics of the configurations, can be found in [114].

The terminology introduced in Table 2.1 is used to state if, and which, ***RAM indicators*** are provided in each reference. The level of detail reported is specified according to the conventions introduced in Table 2.2. Even though a constant (averaged) value of the failure rate over time is generally reported – based on the generally assumed homogeneous Poisson process for the components' useful life, information on the time variance of the failure rate and the other RAM indicators is reported where possible. It must be noted that, although the best match between the provided definitions and single initiatives terminology is pursued, the classification is completed based on the engineering judgement of the authors.

Table 2.2: Legend of symbols adopted in the synoptic tables.

Symbol	Description
$\%_*$	Available only as a percentage/share of the $*$, per component (without total/absolute values)
N	Number of stops
$*$ (t)	Information on the time distribution of the RAM variable ($*$) is given
$*_i$ (t)	Information on the time distribution of the RAM variable ($*$) is given, per component
✓/X	Information given but not possible and/or easy to access
●	Findings for failure frequency (top three)
○	Findings for failure downtime (top three)
Ave	Averaged information and/or results
Det	Detailed results
Info	Additional/other information
WT/WF	Wind turbines/farms in the population
On	Onshore installations
Off	Offshore installation

To implicitly suggest the definition of failure used, the classification of the typology of the ***O&M data sources*** is specified if possible. As reported by Kaidis et al. [120] the use of different methods for the data collection is associated with multiple RAM information and data quality. While in the IEA Wind Task 33 [98], the four main groups of equipment, operating, failure and maintenance/inspection data are distinguished, a higher level of detail is here given by following the grouping of [120]. As for [107], the data sources are classified in Table 2.3. Furthermore, the difference between the collection scheme

typologies – either ***raw*** or ***results data*** approach, as proposed by [58] – is additionally made, whenever possible.

To assist the understanding of incomplete and public-restricted analysis, an additional column is added for summarising the ***main findings***. The list of references is then included by specifying the type of information given in the case of multiple citations. Other basic information, such as ***country***, ***study period*** and ***database size***, is also reported for completeness.

Table 2.3: Reliability and maintenance data sources and their characteristics, by adapting and extending the [54], [65].

Type of data	Information derived	Notes (*disadvantages*)
(i) Maintenance logs (incident reports)	• Accurate information on failures • Information for downtimes • Cost of repair/ replacement	• Can be in hard copy (*difficult to read and/or incomplete*)
(ii) Operation and alarm logs	• Number of stops • Stops duration	• Unknown alarm code • Numerous stops for same failure • (*No information on environment. conditions*) • (*No description of maintenance activity*)
(iii) a. 10-minutes SCADA b. SCADA alarms	• Failure data (frequency and downtime) • Information for further analysis (e.g., root cause analysis) • Environmental conditions	• Large amount of data (*consuming processing*) • Not all alarms are associated with failures • (*No description of maintenance activity*)
(iv)Service provider bills	• Maintenance cost • Indications of failures	• (*No detailed information*)
(v) Component purchase bills (work orders)	• Cost information for components 'repair/replacement	• (*No information on failure data*)

Synoptic tables

The characteristics of the 24 databases found in the literature were accessed either through the initiatives' original publications (if possible), or the review works mentioned in Section 2.1. Due to the low level of detail in some of these works, their cataloguing remained partially completed, not providing a sufficient description for the classification of some quantities. It should be noted that further analysis is required to integrate the results from the European EUROWIN and EUSEFIA projects [58], the Japanese NEDO initiative [58], and the Fraunhofer's EVW and recently launched WInD-Pool database (see Section 2.1). Therefore, only the initiatives that could be fully accessed are reported in the synoptic Table 2.4.

Table 2.4: Synoptic table cataloguing the details on the literature databases, and the main findings of these initiatives.

Database and Organisation	Country	Period	WTs	Data source and type	Size	WT characteristics			RAM / Performance data			Notes (cf. Table 2.2)	Ref.
						Rating [MW]	Age [year]	Type	λ	DT	KPI		
VVT *Finnish Wind Power Association*	Finland	From 1991 Ongoing	On	Performance and failure annual/monthly reports. Reports available from 1999 at [121]	Ave n° WTs (2000-2004): 71.2 N° WFs (in 2006): ~92	✓/X (info in [121])			λ	MDT	CF	2000-2004 statistics • *hydraulics, blades/ pitch, gears* ○ *gears, blades/pitch, hydraulics*	[70], [71] (accessing 2000/2004 stats from [121])
					N° WTs (1996-2008) < 1 MW: 35 ≥ 1MW: 37	X	1÷20 (early fail. excluded)	X	$(\%)_{\lambda}$ $\lambda(t)$	$(\%)_{DT}$ MDT(t)	X	In the original study [121] (in language) $*_i$ (t) are also reported	[122], [123] *from* [124]
Vindstats *ELFORSK*	Sweden	From 1997 Ongoing	On	Performance and failure annual reports. Online available at [125] (from 2002)	Ave n° WTs (until 2005): 723	✓/X (info in [125])	✓/X (info in [125])		X	$\%_{DT}$	X	2000-2004 statistics (the two initiatives have some overlap) • *electrical sys., sensors, blades/ pitch* ○ *gears, control sys, electrical sys.*	*(2000/4)* [70], [71] *(2009)* [122], [123]
					(in 2009)	≥ 0.15 ≤ 3.0				X	A_0		
Felanalys *Vattenfall Power Consultant (prev. Swedpower AB)*		From 1989 to 2005		Type (i)	Ave n° WTs (in 2005): 786	< 0.5 0.5 ÷ 1 > 1.0	1÷19	X	$\%_{\lambda}$ $\lambda(t)$	X	$A_0(t)$		
Windstats Newletter *Haymarket Media Group*	Denmark	From 1987 Ongoing	On	Failure monthly reports	Ave n° WTs (1999-2001): ~2,000	X	X	X	$(\%)_{\lambda}$ $\lambda_i(t)$	X	X	*other, control, yaw sys.*	[69]
					N° WTs (1994-2004): 2,345-851	≥ 0.1 ≤ 2.5	X	✓/X (all concepts)	λ(t)	X	X	*other, yaw sys., hydraulic sys.*	[73]
												Influence of wind speed on WT λ	*Det* [76] *Info* [126]
	Germany			Failure quarterly reports	Ave n° WTs (1999-2001): ~2,750	X	X	X	$(\%)_{\lambda}$ $\lambda_i(t)$	X	X	*other, electric, control*	[69]
					N° WTs (1994-2004): 1,291-4,285	≥ 0.1 ≤ 2.5	X	✓/X (all concepts)	λ(t)	X	X	*electrical sys., other, yaw sys.*	*Ave/Det* [73] *Info* [126]
					1999's WTs n°: ~7,000 2008's WTs n°:~20,000	X	X	X	$N_i(t)$ $(\%)_N$	$DT_i(t)$		Based on data from Bill Canter (2010), Editor at WindStats.	[122]
	–			(Finland, Sweden, Denmark and Germany ave. results)	N° WTs (in 2012) Denmark: ~5,000 Germany: ~24,000 Sweden: ~1,200	X	X	X	X	DT	X	Averaged 2003-2007 and 2008-2012. *gearbox, elect. sys., generator, rotor*	[123]

Table 2.4: Synoptic table cataloguing the details on the literature databases, and the main findings of these initiatives.

Database and Organisation	Country	Period	WTs	Data source and type	Size	WT characteristics			RAM / Performance data				Ref.
						Rating [MW]	Age [year]	Type	λ	DT	KPI	Notes (cf. Table 2.2)	
WMEP *Fraunhofer IWES (prev. ISET)*	Germany	From 1989 to 2008	On	Type (i), (iv) for about 17 years (~64,000 reports)	Ave n° WTs (1998-2000): ~1,435	0 ÷ 1.5 0.56 ÷ 1.5	X	X	$\lambda_i(t)$	X	X	Results per size	[69]
					Ave n° WTs (2004-2005): 865	X	X	X	λ	MDT	X	• *electrical sys., control sys., hydr.* ○ *generators, gears, drivetrain*	[70], [71]
					Ave n° WTs (until 2006): ~1,500	< 0.5 0.5 ÷ 1 ≥ 1.0	1÷15	X	λ λ(t)	X	X	Results per age, and size	[127]
								all (by concept)	N(t) $N_i(t)$	X	X	- Results per age, size, and concept - Only stops for unplanned maint.	[128]
					Ave n° WTs (in 2008): over 1,500	< 0.5 0.5 ÷ 1 ≥ 1.0	X	all (by concept)	λ	MTTR	$A_T(t)$	Results per concept, and location	[79], [115], [117], [129], [130]
				Additional weather data from meteorological masts	N° of E32/33 WTs: 32 (~24% of E32/33 in WMEP survey)	0.3 0.33	1÷10	DImW	λ	X	X	- Periodicity of λ with wind speed - Influence of site	[75], [77]
LWK *Land Wirtschafts-Kammer*	Germany (Schleswig- Holstein)	From 1993 to 2006	On	Failure annual report	N° WTs (1999-2000): 510	0 ÷ 1.5 0.56 ÷ 1.5	X	X	λ	X	X	Results per size	[69]
					N° WTs (1994-2004): 158-653	0.225 ÷ 1.5 (in 3 groups)	Average age(t)	A0/A1/A2 B, C, DDE		MTTR	X	Time averaged results per size, and type	*Det* [126] *Ave* [63], [72], [74]
					In 13 years ~5,800 turbine years	X	X	X			X	Average result (per component) against the WMEP survey	[58], [61], [64]
MECAL *MECAL Experts*	Nether-lands	From ~2010 to 2014	On	Type (iii.a/b) *raw data*	N° WTs in WF A: 23 N° WTs in WF B: 36 N° WTs in WF C: 4	3.0 0.85 ÷ 1.75 2.0	1÷2-4 (early fail. excluded)	A2, B, C, DI	$(\%)_\lambda$	$(\%)_{MART}$	X	Differentiate in failure severity and MART or MDT	[120], [131]
RGU *Robert Gordon University*	United Kingdom	From ~1997 to 2006	On	Type (i)	N° WTs: 77 N° WFs: 26	0.6	X	X	X	X	X	Weibull distrib. for drivetrain components	[132]
Strathclyde (On) *University of Strathclyde*	United Kingdom	From ~2009 to 2014	On	Type (i), (v) from one manufacturer	Total n° WTs C: 1822 DImP/DI1P: 400	1.5 ÷ 2.5	1÷5	C DImP DI1P	λ(t)	RT	CF	- Broken into cost categories - Info. per type (only drivetrain sub-sys.)	[133]

Table 2.4: Synoptic table cataloguing the details on the literature databases, and the main findings of these initiatives.

Database and Organisation	Country	Period	WTs	Data source and type	Size	WT characteristics			RAM / Performance data				Ref.
						Rating [MW]	Age [year]	Type	λ	DT	KPI	Notes (cf. Table 2.2)	
CIRCE *Universidad de Zaragoza*	Spain	From ~2013 to 2016	On	Type (i) and (iii.a), *raw data* from 14 manufacturers	Ave n° WTs: <4,300 Ave n° WFs: 230 Ave n° WTs C < 1 MW: 2,130 C ≥ 1MW: 2,270 DD: 215	0.3 ÷ 1	X	C	% λ	% TTR	X	• *gearbox, blades, blade brake* ○ *gearbox, generator, blades*	[60]
						1 ÷ 3		C				• *gearbox, control, pitch sys.* ○ *gearbox, generator, blades*	
						0.6 ÷ 2		DD				• *controller, meteo. stat., yaw sys.* ○ *generator, blades, controller*	
				Type (iii.a/b), *raw data*	Ave.n° DI WTs: 383 Ave.n° DD WTs: 57	0.85 ÷2 2	X	C DD	X	X	X	Comparison of failures and alarm logs (per size/type)	*Det* [60] *Info* [134]
ReliaWind *European Union founded project*	Europe	From 2008 to 2010	On	Type (ii), (iii.a), (iv), (v)	Ave n° WTs: ~350	> 0.85	> 2	Variable-speed	$(\%)_{\lambda}$	$(\%)_{TTR}$	X	- Unpublished absolute values Insights on impact of single systems	[64], [135]
EPRI *Electric Power Research Institute*	United States (California)	From 1986 to 1987	On	Type (i)	Ave.n° WTs: 290	0.04 ÷ 0.6	X	✓/X (info in [69])	λ	MART	X	• *gearbox, blades, blade brake* ○ *gearbox, generator, blade*	[69]
CREW *Sandia National Laboratory*	United States	From 2007 Ongoing	On	Type (iii.a/b), *raw data*	Ave.n° WTs: ~900 N° WFs: 10 (until 2012)	X	X	X	$(\%)_{NA}$	MDT	A_0 CF	2012/13 results in unavailability (NA)	*Det* [136] *Info* [137]
				Type (i) and (iii. a/b), results *data*	(in 2016)	X	X		X	X	X	General information on the initiative	[87]
Muppandal *Noorul Islam University*	India	From 2000 to 2004	On	Type (ii), (iii)	N° WTs: 15	0.225	X	A0	λ(t) λ_i(t)	MTTR(t)	CF(t) A_0(t) A_T(t)	- *mech./ electr. system* downtime by maint. type - Cost analysis	[85]
CWEA *Chinese Wind Energy Association*	China	From 2010 to 2012	On	Performance and failure data from 47 manufacturers	2010's WTs n°: 111 2011's WTs n°: 560 2012's WTs n°: 640	✓/X (only top 9 manufact.)	~ 2-3 (WTs installed in 2008/2009)	C DImP DDP	λ	X	A_T(t)	- λ derived by [58] - A_T according to [58] per drivetrain type (in 2011/12)	[59]

2.3.2 Taxonomy adaption

As for [58], [105], it was necessary to select a uniform and convenient language to identify the equipment in a wind turbine to coherently compare the several statistics. As outlined by [98], different lists of terms can be used for categorising aspects of components, failures, maintenance tasks etc. as "taxonomies". Several equipment taxonomies have been developed in the past. Among these are the VTT components' breakdown presented by Stenberg in [124], the Sandia Laboratories taxonomy in [138], and the GADS (Generating Availability Data System) used for North America wind plants. Another two, more recent, sophisticated and comprehensive classifications are the ReliaWind project taxonomy (see e.g. [60], [64]) and the RDS-PP® (Reference Designation System for Power Plants) taxonomy, published by the VGB PowerTech e.V. in [139].

The first was created for an extensive failure data analysis and is internationally recognised, while the second one, also widely accepted, is currently employed for the offshore data collection schemes of SPARTA and WinD-Pool (see Section). However, the ReliaWind complete taxonomy is not publicly available, and there is no ongoing development to maintain it. In contrast, the RDS-PP® offers open access to the draft document [139], and a high level of detail for both system and subsystem identification and components' technical information. For these reasons, the RDS-PP® was adopted to unify the statistical reliability and maintainability data in this analysis. The information necessary for the adaptation of the taxonomies is accessed from the VGB PowerTech e.V. draft document [139], and summarised in Table 2.5.

Table 2.5: RDS-PP® taxonomy adopted for system and subsystem with numbered labels added for the presentation of the results.

RDS-PP® Acronyms	Labels
MDA	1. Rotor System
MDA10	1a. Rotor Blades
MDA20	1b. Rotor Hub Unit
MDA30	1c. Rotor Brake System
-	1d. Pitch System
MDK	2. Drivetrain System
MDK20	2a. Speed Conversion System
MDK30	2b. Brake System Drivetrain
MDL	3. Yaw System
MDX	4. Central Hydraulic System
MDY	5. Control System
MKA	6. Power Generation System
MS	7. Transmission
MSE	7a. Converter System
MST	7b. Generator Transformer System
MUD	8. Nacelle
MUR	9. Common Cooling System
CKJ10	10. Meteorological Measurement
UMD	11. Tower System
UMD10	11a. Tower
UMD80	11b. Foundation System
-	12. Others

The author mapped, to the best of their knowledge, the initiative-specific taxonomies to RDS-PP®. When a proper mapping was not possible, a higher share was given to the introduced "Other" category. On the other hand, the generic "electrical systems" category, usually adopted in the earlier data collections, is integrated here into the transmission group.

2.3.3 Availability calculation

Availability assessment tool

The variance in the failure rate of the wind turbines results in a variance in their estimated availability. This can affect decision making for the accurate planning of future offshore wind projects. The openO&M assessment tool, developed by the authors in [140], is employed to estimate and investigate the impact of the different failure rates provided in the literature on the potential availability of an offshore wind farm. The tool was built with the aim of supporting the development of wind farms' maintenance strategies. As for other O&M management tools [141], [142], they have a modular structure consisting of the following core modules: (1) reliability, (2) power, (3) weather forecasting, (4) maintenance and (5) cost. The flowchart of the processes and steps of the tool are shown in Figure 2.1.

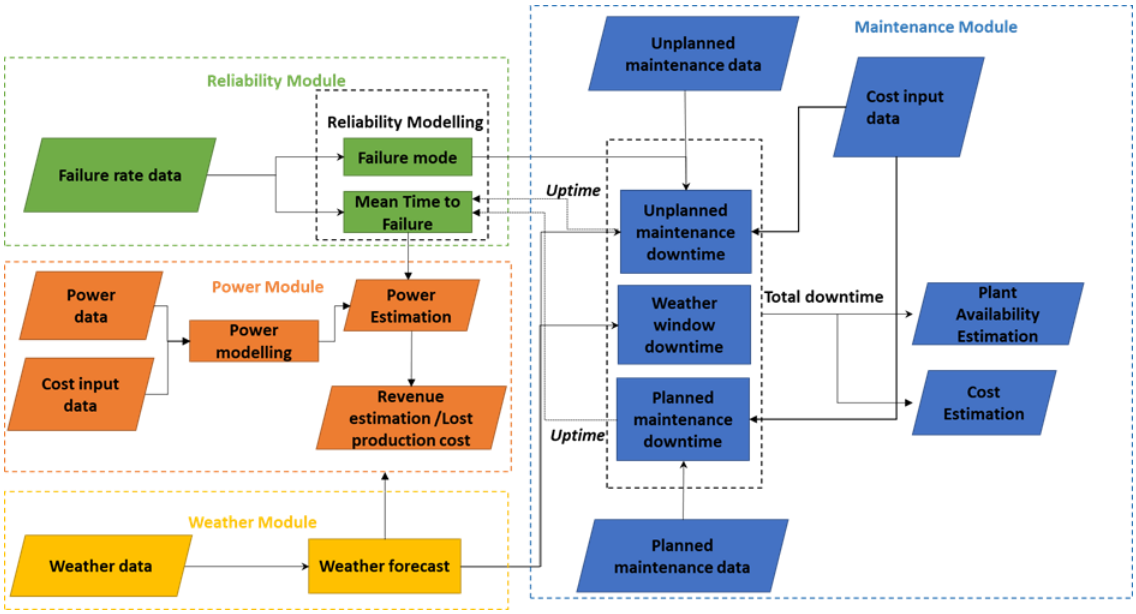


Figure 2.1: Workflow of openO&M tool.

The inputs are weather data, cost data, reliability data, turbine specific data (power curve), wind farm layout data (distances from shore), and repair information, such as number and type of vessels, and the crew required for the restoring of the system. The wind farm lifecycle and maintenance activities are simulated, and a number of KPIs are produced, such as the time-based availability, power produced, and operating costs.

Wind farm availability estimation

For the calculation of the lifetime availability of a wind farm, both planned and unplanned maintenance activities are considered. On the one hand, scheduled maintenance happens at

yearly intervals and is performed for each subsystem of the turbines in the farm following grouping and prioritisation. The downtimes are calculated based on the maintenance activity duration which is assumed to be fixed. On the other hand, unplanned maintenance downtime is sensitive to the availability of spare parts, vessels, and personnel for the repair of damaged subsystems.

With regard to the distribution of unforeseen failures in time, this information is modelled from the reliability module based on the reliability data from the literature. The input failure rates are grouped into minor repair (*mr*), major repair (*Mr*) and major replacement (*MR*), according to the material costs indicated by Carroll et al. in [62]. When this information is not provided, the downtime statistics are used for the classification of the failure rates, similarly to what is suggested by [47] (cp. Table 2.6). When a failure occurs, the turbine status varies depending on the failure type. For *mr*, the turbine is assumed to continue operating even after the failure detection, and the shutdown is only assumed during the repair time. For *Mr* and *MR*, the turbine is stopped after the detection of malfunctioning, going back to service only after the system is restored. The time to failure associated with each failure mode, for a particular subsystem *i*, is assumed to be distributed by an exponential probability density function *f(t)* (Equation 2.1) with parameter $\lambda_{i,mode}$ being the failure rate for subsystem *i* under a particular failure mode (i.e., *mr*, *Mr*, or *MR*).

$$f(t)= \lambda_{i,mode}\, e^{-\lambda_{i,mode}t} \qquad \textbf{Equation 2.1}$$

The cumulative distribution function is the probability of failure (*PoF*) of the subsystem according to the exponential reliability theory and is given in Equation 2.2. The *PoF* of the whole wind turbine is the *PoF* of all subsystems considering all failure mode classifications, as explained further in [140].

$$PoF=1-e^{-\lambda_{i,mode}t} \qquad \textbf{Equation 2.2}$$

Further inputs for the availability estimation are the farm layout and the forecast of the environmental conditions during its lifetime. For the purpose of this analysis, the wind farm reference layout is based on Bak et al. [143]. The weather data simulated throughout the lifetime of the farm for its operation (wind speed for power production) and accessibility (wave height for the mobilisation of the vessels) are based on the FINO3 database. The stochasticity of the weather module is obtained by the implementation of a Markov model trained on the historical wind speeds and wave heights. Finally, information on the times and logistics for the performance of the unplanned maintenance – including repair times and resources needed – are based on [62]. The further assumptions in the maintenance module and additional information can be retrieved from [140].

Table 2.6: Criteria for the classification of the reliability databases in minor and major repair and major replacements.

Classification	mr	Mr	MR
Material cost	≤ 1,000 €	1,000 € < cost ≤ 10,000 €	> 10,000 €
Downtime (onshore)	≤ 3 days	3 days < downtime ≤ 7 days	> 7 days
Downtime (offshore)	≤ 7 days	7 days < downtime ≤ 15 days	> 15 days

2.4 Results

A more in-depth view of the data repositories enables a cross-comparison of the statistics and critical discussion. Initially, the quality and consistency of the averaged reliability and maintainability figures are evaluated in an all-in-one comparison. A detailed discussion of the effect of the deployment parameters on the reliability and performance of onshore wind turbines is then suggested, leading to either further supporting the trends already identified in historical repositories or updating them based on the experience of the more recent surveys.

2.4.1 Trends in the averaged reliability and maintainability statistics

In Figure 2.2, the data from all the complete and accessible initiatives, collected in Table 2.4, are presented as dimensional quantities, in failure frequency against time lost to restore the system after failure. Due to poor documentation, confidentiality reasons, or the lack of a standardised approach, a significant spread across these averaged results is immediately observable. These statistics generally collate the data over broad populations, for varying characteristics of the units. Based on the detailed analysis of Table 2.4 and the homogeneous taxonomy adaption activity, it is possible to discuss and draw the main conclusions from the plot.

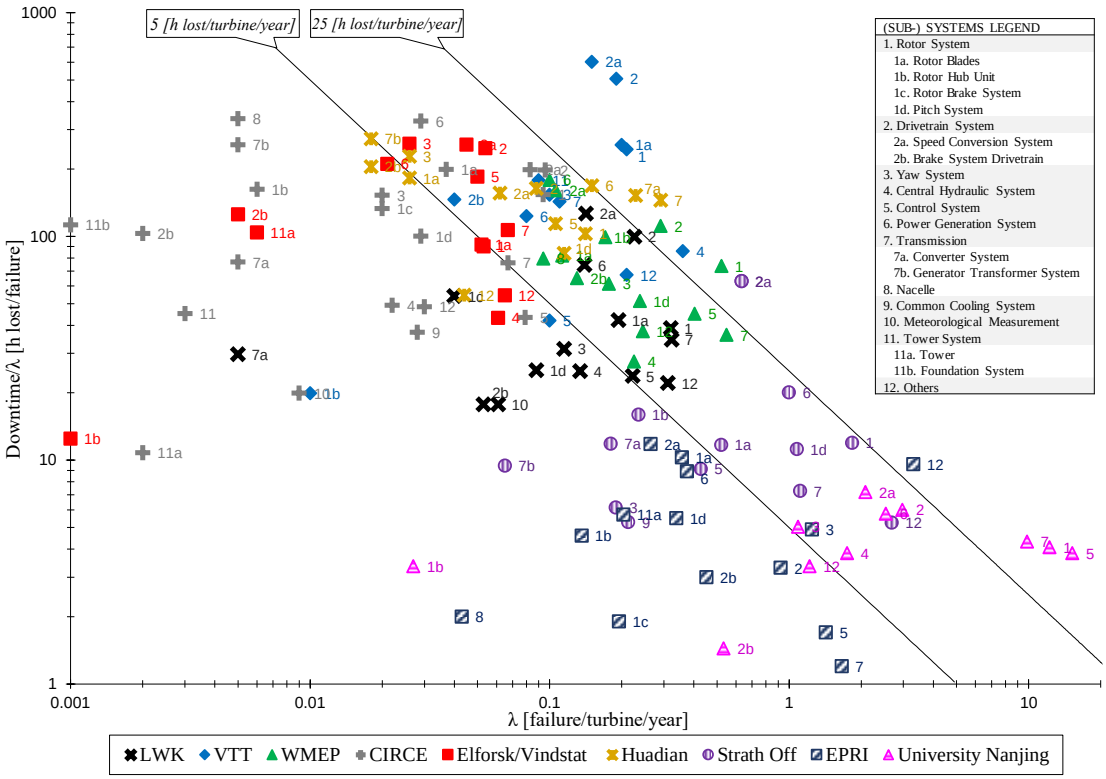


Figure 2.2: All-in-one comparison of the most complete reliability statistics, as rate of failure vs. hours lost per failure per components (bi-logarithmic scale). Some data are retrieved from [58], while others are integrated according to Section 2.3.1.

It is noticeable that the results from the WMEP and Huadian databases are largely in the medium range of lost hours per failure of the components (from 5 to 10 h/turbine/year), whilst the data from the other sources are distributed over a wider range. For instance, the majority of Spanish (CIRCE) results are distributed below five hours per turbine and year limit, while the VTT population shows values largely outside the average range, as observed by [113]. A possible contributor to the inconsistencies between these initiatives can be the definition of "failure". This can indeed vary from being just a required visit to the turbine, considering only when a maintenance activity is required (e.g. [62], [128]), to when an event has a downtime over a certain threshold (e.g. [120]), or, conversely, to account for alarm logs and remote resets (e.g. [60], [84]). When counting remote resets, the failure rates recorded are higher.

Likewise, the mean time lost to restore the turbine operation is lower, because of the presence of these small downtime events. This is the case with the data collected by the Southeast University Nanjing [84], whose outlier behaviour can be traced back to either the use of SCADA alarms or the very short period of the survey [106]. The unrealistically high failure rate and small downtime of the EPRI statistics can be associated with the infant mortality of its early-stage technologies. With regard to the recent CIRCE statistics, while the time lost per failure is generally comparable to that of the older data collected (Vindstat and Felanalys, VTT, LWK and WMEP), the lower frequencies of the malfunctions can be related to the higher maturity of the technology and to the fact that only components' (internally caused) failures were considered, excluding from the analysis all the other outage events.

Regarding the Strathclyde offshore (Strath-Off) statistics, their skewed behaviour is associated with the use of the mean active repair time (MART) as an indication of the downtime. Nonetheless, the gap between these MARTs and the mean downtime (MDT) of the other statistics – except for Huadian and EPRI collections, already suggests the possible high impact of logistics and technical delays on the downtime for maintenance actions. Refer to Section B.1 of the Appendices for the definition of the MART and the MDT statistics.

Looking then at the single initiatives' systems' share, the drivetrain failures seem to be, in general, the highest contributor to the hours lost per turbine per year, due to the presence of a gearbox. While this is true for the European initiatives, the Chinese statistics of Huadian and Nanjing contrapose a higher criticality for the control and electrical transmission system. This is in line with what is reported in the CWEA study, where converters yield the highest failure rate among all the subsystems. As explained by the authors [59], this could be related to the harsh environment (very low temperature reached) in the area where these wind farms are installed. However, Artigao et al. [105] suggest that this is a common trend among the Chinese statistics.

2.4.2 Trends based on design parameters

At a turbine and a farm level

The authors from the DOWEC project [69] were the first to identify the need to separate the data also by power classes when plotting and comparing the reliability statistics. By using the data available at the time from the WMEP statistics, they observed that the turbines rated between 0.56-1.5 MW fail significantly more often than the smaller turbines; however, the population is 95% represented by lower rating units. Based on the German "250 MW Wind" program, more detailed and complete results for the WMEP project were collected

and published. From the data of the first 15 years of the initiative, Hahn et al. [127] and Echavarria et al. [128] observed a time invariant increase of the turbine failure frequency with power rating.

In the LWK survey, the distribution of failure intensity among 12 different turbine models was sorted by turbine size [113]. The same authors who, in [73], already intuitively appreciated a lower reliability in the newer German turbines comparing them to their smaller predecessors, reaffirmed in [63] the general trend of failure rate to increase proportionally with the turbines' rating. This was shown to be particularly true for the type A1 turbines [113], while the direct-drive technologies (DDE type) seemed not to follow the rule, maintaining an almost constant overall failure rate of about 2.5 failures per turbine and year for larger units. Nothing can be stated for other concepts, such as type B, due to the lack of data for large units. Looking at the averaged failure rates for type C, the more recent CIRCE data collection reported results are in line with the discussed hypothesis: 0.46 and 0.52 failures per turbine and year, for the population below and above 1 MW, respectively [60].

This higher reliability would lead to generally higher technical availability. Reinforcing this argument, Harman et al. [116] observed that the operational availability of sub-MW units is higher than that of the larger units. However, at the array level, they additionally observed that an increase in the availability of larger farms (with more than 40 units), was proportional to the number of turbines and is independent of the units' rating.

At a system and a subsystem level

While this last deduction cannot be verified yet, as the wind farm specific data are limited and incomplete, the trend of the failure rate with the power rating is investigated here at the system (Figure 2.3) and subsystem (Figure 2.4) levels. When summarising the WMEP survey main findings in [115], [117], [130], Faulstich et al. developed a detailed analysis at the subsystem level (Figure 2.3). Significant differences in the contribution of single components' failures to the total failure frequency were detected for small and large sized turbines.

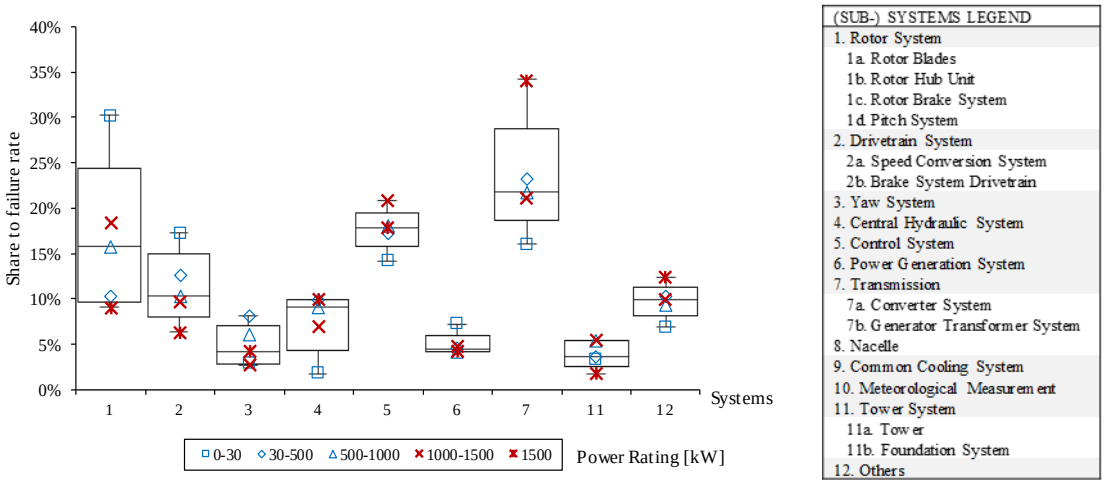


Figure 2.3: WMEP program final results in terms of share per subsystem, of the total average failure rate of each group of MW-class, adapted from [82].

A smaller scatter for the share of each system to the failure rate is generally observed for medium rating turbines (above 500 kW and below 1.5 MW), reaffirming what was

noticed by Dao et al. [106]. For the higher power class, the reduction in the percentage failure of the mechanical and structural components is balanced by an increase in the percentage of the electrical failures. As they are associated with an overall rise of the annual failure frequency [127], the electrical and control systems can be seen as the most critical components for the WMEP larger sized turbines. Although these results are in agreement with the project final average statistics [130], information on the distribution of the power rating in the final WMEP population is missing. Furthermore, it has to be noted that these results are mainly representative of technologies and layouts that are no longer adopted (type A0-A2) [78]. For these reasons, the results from other initiatives were analysed, seeking for a match with the WMEP trends.

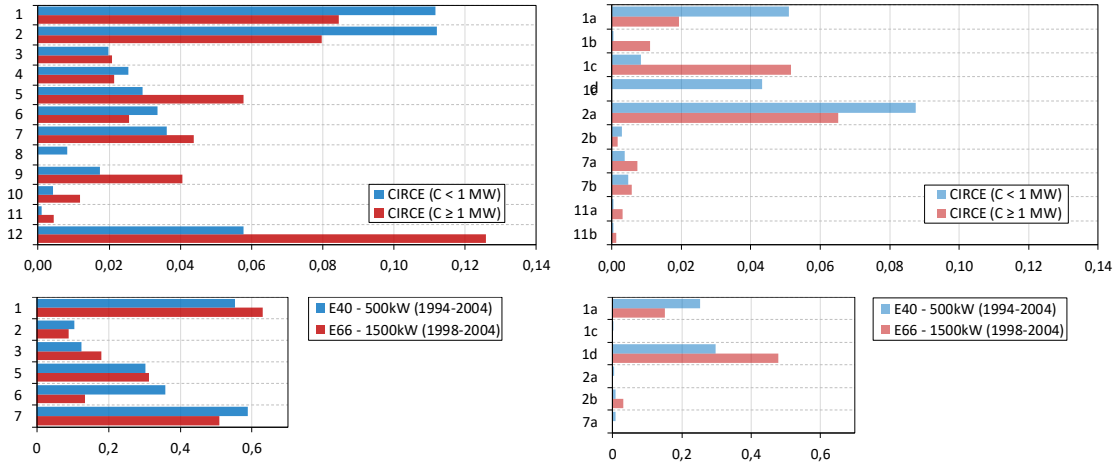


Figure 2.4: Comparison of the failure rates per systems (left) and subsystems (right) of different rating turbines. Statistics for type C turbines from CIRCE database (top) and type DDE turbines in LWK database (bottom).

The WMEP statistics are cross compared with those of the LWK and CIRCE surveys. To maintain the analysis as unbiased for the drivetrain configuration, a comparison is suggested among turbines of the same typology. The LWK results for the Enercon E40 (500 kW) and E66 (1.5 MW) gearless turbines [74], [113] are presented in Figure 2.4-bottom, to better understand the anomalous behaviour of LWK's DDE model's failures. Similarly, the failure rates derived by accessing CIRCE data [60] for the type C turbines, above and below 1 MW, are compared in Figure 2.4-top. The conclusions on the effect of the power class on the reliability figures per component, without the effect of multiple configurations' averaged results, are the following:

- The number of shutdowns caused by the rotor system (1) and the power generator system (6) – either EESG of LWK or the DFIGs of CIRCE – decrease when the turbine rating increases, in agreement with WMEP results.
- An opposite trend characterises the control (5) and transmission (7) systems of CIRCE turbines, in agreement with the WMEP observation. In contrast, the transmission failure rates for small sized DDE turbines are higher than for larger systems, as verified by [113] as well.
- The speed conversion (2a) and drivetrain brake (2b) subsystems of the CIRCE population seem to fail less frequently for higher power ratings, in line with WMEP results for the drivetrain system (2) and with what is reported by [144] for the type A1 turbines of the LWK survey.

- The pitch system (1d) – not identified as a separate subsystem in the WMEP results and likely to be integrated and averaged into the hydraulic system – follows the same trend for both LWK and CIRCE initiatives, increasing its number of failures with increasing turbine size.

Modern layouts statistics

With regard to the influence of drivetrain configurations on the statistics, the WMEP authors noticed a general decrease in the average reliability of the assets moving towards more advanced concepts, compared to the simple Danish concepts and standard variable speed [78]. Similarly, the analysis of van Bussel et al. [93], [94] identified the robust design – consisting of a two-bladed turbine, with no pitch control installed on a monopile foundation – as the best design solution for obtaining the highest availability in a large offshore wind farm project.

For the direct-drive configurations, the WMEP researchers observed that gearless layouts are not necessarily more reliable than geared ones [115]. Echavarria et al. [128] highlighted this aspect in more detail by analysing 10 years of WMEP time-trend results. They noticed that the direct-drive synchronous generator is not mainly responsible for the generally higher failure rate of these turbines compared to the geared alternative with an induction generator. In contrast, the failure events of power electronic components, in systems using synchronous generators, were significantly more frequent. This suggests that the statistics could possibly have been affected by the young age and novelty of the technology. Similarly, the LWK's authors [63], [73], [74] noticed that the aggregate failure of the generator and converters in direct-drive layouts (DDE) is greater than the aggregate failure rate of gearbox, generator and converters in indirect-drive ones. Indeed, the elimination of a gearbox resulted in a substantial increase in the failure rate of electrical-related subassemblies.

Table 2.7: Comparison of CIRCE type C and type DD reliability and maintainability statistics.

	λ [failure/turbine/year]		Downtime/Turbine [h]	
	CIRCE C (< 1 MW)	CIRCE DD (≤ 2 MW)	CIRCE C (0.3÷3 MW)	CIRCE DD (≤ 2 MW)
1. Rotor System	0.084	0.011	131.19	556.63
2. Drivetrain System	0.080	0.002	133.31	79.69
3. Yaw System	0.021	0.014	77.42	110.80
4. Central Hydraulic System	0.021	0	33.40	0.00
5. Control System	0.058	0.052	32.18	61.26
6. Power Generation System	0.025	0.011	305.70	508.41
7. Transmission	0.044	0.024	107.96	50.28
8. Nacelle	0	0		0.0
9. Common Cooling System	0.041	0.005	28.53	51.79
10. Meteorological Measurement	0.012	0.019	7.61	14.28
11. Tower System	0.004	0.002	44.34	52.00
12. Others	0.126	0.042	27.49	42.46

Some authors justified this tendency as being due to the immaturity of these technologies and the presence of new issues related to the new design [110] and larger dimensions [73] of the direct-drive generators. In agreement with this hypothesis, the more

recent data collection from Reder et al. [60] registered an overall decrease in the failure rate for the Spanish direct-drive population compared to the type C one: 0.19 vs. 0.49 failure per turbine and year, respectively. The updated RAM statistics for the newer typology of DDP published by Lin et al. [59], highlighted an increase in the availability, during the second year of operation, of direct-drive design compared to the C type. To complete the discussion investigating the variance in failure frequency and downtimes, per component, between type C and DD turbines, a comparison among CIRCE statistics [60] is suggested in Table 2.7. Analysing only the data for turbines above 1 MW, an unexpected higher failure of the rotor system is observed for the type C turbines. This behaviour could be affected by several factors. In contrast, the 50% reduction in the number of failures of the direct-drive generators, could be explained by a profound technology improvement of EESG and/or a switch to the PMSGs' direct-drive concept [133]. Nonetheless, as mentioned by [110], the lack of detailed information about the typology of the generator only allows us to speculate about the cause of this higher failure rate.

As far as the medium-speed configurations are concerned, it has been shown [145], [146] that they have the potential to offer a good compromise between reliable operations and cost optimisation. On the one side, Carroll et al. [133] compared the statistics of a population of 1,800 type C turbines with those of a group of 400 DImP. They observed that the full-rated converters are the most critical components in terms of failure rate, while the PMSG fails almost 40% less than a DFIG and its failure modes are for minor repairs of the auxiliary (lubrication and cooling) system. Nonetheless, due to the presence of the gearbox, they noticed that the hybrid layout fails nearly three times more often than the traditional type C configuration. On the other side, Lin et al. [59] reported a significantly higher technical availability of the hybrid configuration compared to both type C and DDP. This latter study is likely to be skewed by early-stage failures. Because of this contrast and due to the recent and sporadic installations, it is not yet possible to draw any conclusions on their robustness. Thus, more information needs to be collected and compared for these kinds of systems.

2.4.3 Trends with the deployment parameters

Although site-specific information is not yet available for supporting the observation of Harman et al. [116] on how the farm size influences its availability, some experience on the effect of the location and the environmental parameters on the turbine reliability figures can be found in the literature. Indeed, much research effort has been dedicated to identifying the critical meteorological parameters that influence the turbine failure behaviour negatively.

One of the first extensive analyses on the effects of weather on turbine reliability was presented by Hahn et al. [81] showing increased failure rates of certain components with rising average daily wind speeds. The electrical system subassemblies showed the strongest dependency on wind speed, followed by the control system, while a significantly weaker correlation was exhibited by the other main subassemblies. Tavner et al. [76] identified an annual periodicity in failure rates due to seasonal variation in weather conditions, by analysing the correlation between monthly averaged wind speed conditions and component failures. Following this first study, they extended their analysis in [75], [77] by cross-correlating the component failures with average monthly maximum and mean wind speed, maximum and minimum air temperature, and average daily mean relative humidity. They concluded that other weather conditions, rather than just wind speeds, can be closely

related to the turbine failures. Wilson et al. [147] used artificial neural networks to investigate if any relationship existed between maximum daily gust speed, average daily wind speed and temperature, and the turbine's failure rates. The gearbox, generator and hub were shown to be more likely to fail in variable wind conditions, with a high potential impact on the failure rates of these subassemblies offshore. They additionally noticed that gust speed is a key parameter of the number of failures. This observation reaffirms what was shown already by [76], who observed that malfunctions occur more frequently in the winter months where average daily wind speeds can be lower but maximum daily gust speeds are higher. Regarding the data collected for offshore wind turbines, Carroll et al. [62] noted that offshore turbines sited in areas with higher wind speeds experienced higher failure rates. This observation is in line with what was shown by Wilson and McMillan in [148] for onshore systems. Nonetheless, while this correlation appears to be rather weak onshore (linear regression slope is 0.08), the higher offshore wind speeds seem to have a higher impact on the failure rates (with 1.77 slope).

Finally, some studies on the possible effect of near-shore location were analysed. Examining the WMEP statistics, Faulstich et al. [78] observed that turbines located near the coast and in the highlands suffer higher failure rates. By analysing a more segregated population of turbines (type DImW, with sub-MW rating), Tavner et al. [75], [77], showed similarities between the results from the Krummhörn and Fehmarn wind farms, presumably because of their near sea locations, compared with Ormont farm, which is located inland. While the turbines from the first two locations are subject to humid conditions, Ormont failures cross-correlate with wind speed standard deviation, suggesting the influence of turbulence on failure rate.

2.5 Discussion of the RAM figures for offshore wind systems

The failure of offshore and near-shore deployments can differ in number and typology compared to those of onshore systems, due to the potential effect of certain environmental conditions (such as humidity, gust events and turbulence intensity). For this reason, in this final discussion, the influence of the studies' specific parameters (i.e., reliability statistics and power ratings) is analysed by estimating the lifetime operational availability of hypothetical wind farms installed at a typical offshore location. The normalised reliability and maintainability figures of onshore and offshore studies are eventually compared, to identify the possible sources of the discrepancies in the results.

2.5.1 Lifetime operational availability estimates and trends

The derivation of the lifetime operational availability for a set of offshore wind farms establishes a common ground for integrating the offshore statistics into the analysis and consistently comparing them with the existing onshore ones. A similar analysis has been already carried out for the DOWEC project, where van Bussel et al. [93], [94] estimated the availability and costs associated with the installation of different turbine technologies (drivetrain and foundations) for a fictional 500 MW offshore wind farm, erected at 35 km from the Dutch coast. Their findings concluded that there is a reduction in the availability of advanced layouts, as opposed to the traditional ones. Despite the fact that this observation supports the conclusions of Faulstich et al. in [78], the reliability figures used

are extracted from a specific population of coastal wind turbines and adapted to several types of designs and the offshore application, based on the authors' best knowledge.

In contrast, the analysis presented here focuses only on the impact of the implementation of the several failure statistics collected from the literature. The reliability studies selected and adapted as required for the implementation on openO&M are reported in Section B.2 of the Appendices. The systems and subsystems considered in the availability calculations are based on the taxonomy shown in Table 2.5 and are additionally subdivided into failure classes (*mr*, *Mr* and *MR*) as suggested in Section 2.3.3. For simplicity of comparison, the wind farm layout, the repair information, and the weather data inputs are kept the same for all case studies. The wind turbine power curves for the estimation of the operational uptime are assumed, based on commercial models at the average power rating of the population. This information, together with the estimated time-based availability, is reported in Table 2.8. The results are presented in terms of averaged value and standard deviation, due to the random selection of the failure modes (as explained in Section 2.3.3).

Table 2.8: Lifetime availability estimated for the selected surveys – see Section B.2 of the Appendices. In grey are the surveys from onshore failure statistics, with different shading according to the assumed wind turbine rating.

Reliability study	Mean rating* [MW]	Assumed rating [MW]	Turbine model for power curve	Mean availability [%]	Std [%]	Range per rating class [%]
LWK**	1	1	GEV HP 1000/62	99.0	+/- 0.05	3.01
WMEP**	1	1	GEV HP 1000/62	96.6	+/- 0.13	
Muppandal**	0.225	1	GEV HP 1000/62	98.6	+/- 0.08	
CIRCE DD	1.3	1	EWT DW61 1MW	99.4	+/- 0.08	
CIRCE GD > 1 MW	2	2	V80 2MW	99.1	+/- 0.05	3.03
Huadian**	2	2	V80 2MW	96.2	+/- 0.08	
Strath-Off	3	3	V90 3MW	84.1	+/- 0.10	4.04
OWEZ	3	3	V90 3MW	80.2	+/- 0.04	

* average between the highest and the lowest power rating from synoptic tables
** average reported by [106]

Although the same characteristics and maintenance strategy for each of the offshore wind farms scenarios are selected, a spread in the estimation of the availability associated with the several reliability datasets can be noticed. This indicates that the frequency of the specific failure events largely affects the availability calculations and consequently the maintenance decision making. In contrast, the low standard deviation of the estimated availability (the highest of +/-0.13% for the WMEP survey), implies a small impact of the randomness in using these failure data when supporting decisions for the operation of wind farms.

A generally higher availability is found for the onshore studies compared to offshore ones, independently from the turbine rating, supporting the hypothesis of a high correlation of the failures with environmental conditions (generally more unfavourable offshore). It is worth pointing out that the direct drive wind farm scenario achieves the highest lifetime availability, in line with the calculations of Carroll et al. in [146], [149], and supporting further the observation of Section 2.4.2 on the high potential of this technology. Among the onshore failure-based scenarios, the WMEP and the Huadian datasets resulted in the smallest estimated availability, as low as 96%. Nonetheless, these results seem reasonable when compared against the operational availabilities collected by [58], on average around 95% for onshore turbines, with the exception of the sub-MW population, and around 93% for the offshore systems in the SPARTA and the WInD-Pool projects. Therefore, the

generally high predicted values for the onshore-based results could suggest a general underestimation of the repair and logistic times as inputs of the maintenance module.

Independently from the sensitivity of availability to these parameters, some of the offshore-based results look surprising. On one side, the low availability, of about 80%, for the OWEZ wind farm can be related to the extraordinary maintenance activity during the time of the survey. This result agrees with the statistics for the UK offshore round I, affected as well by infant mortality events and the underdeveloped supply chain for O&M. On the other side, the 84% operational availability of the Strath-Off population is unexpected, considering that it is associated with the statistics for a population with similar characteristics (turbine configuration and power rating) to those of the recently published RAM database results [58]. The explanation of such behaviour can be associated with:

- the effect of site-specific environmental conditions, which significantly affect the failure behaviour and enhance the statistical uncertainty of the reliability figures – considering that these data are built upon averages from various sites;
- the effect of failure distribution on the estimated availability levels, which has been shown to have a potentially high impact [150], and it has been considered here as exponential only;
- the potential impact of preventive maintenance activities and condition monitoring systems, implemented in the real application, which could potentially avoid or prevent the failure
- the potential impact of maintenance activities performed vi either helicopter or other service operation vessels, which are not yet included in the tool, but they could increase the accessibility windows and reduce the downtime for mobilisation and logistics.

Speculating on the uncertainty surrounding the recently published availability statistics, the WInD-Pool project only reported vague averaged results in [58], while the SPARTA data are provided for 14 months only of recording [91]. Another year of availability figures from the SPARTA project have recently been published [151]. However, their results are reported in terms of production-based availability, giving only an overview on the goodness of the performance of the turbines compared to their power curves.

2.5.2 Influence of the offshore location on systems and subsystems

With the intent of understanding which systems and subsystems could mainly be affected by the offshore location, the reliability statistics from CIRCE [60], Strath-Off [62] and OWEZ offshore WF [90] (cf. Table 2.4) are compared in the bubble plot of Figure 2.5. While their populations are consistent – being representative of geared turbines with an induction generator installed and of comparable power rating –, the definition of failure and downtime differs among these initiatives. Thus, normalised failure rates and downtimes (in terms of h/turbine/year) for each assembly (horizontal axis), are respectively represented on the vertical axis (indicated by the centre of the bubble) and by the bubble size. It is first worth stating that, although adding value to the analysis, the statistics of the OWEZ farm can be affected by early failure events and are skewed by their derivation from the number of stops [152]. This is also the reason for the low share in frequency and repair time of "others" (12) unforeseen failures, while a higher ranking is in the Strath-Off and CIRCE statistics because of the use of a more detailed taxonomy for the collection and analysis of the data.

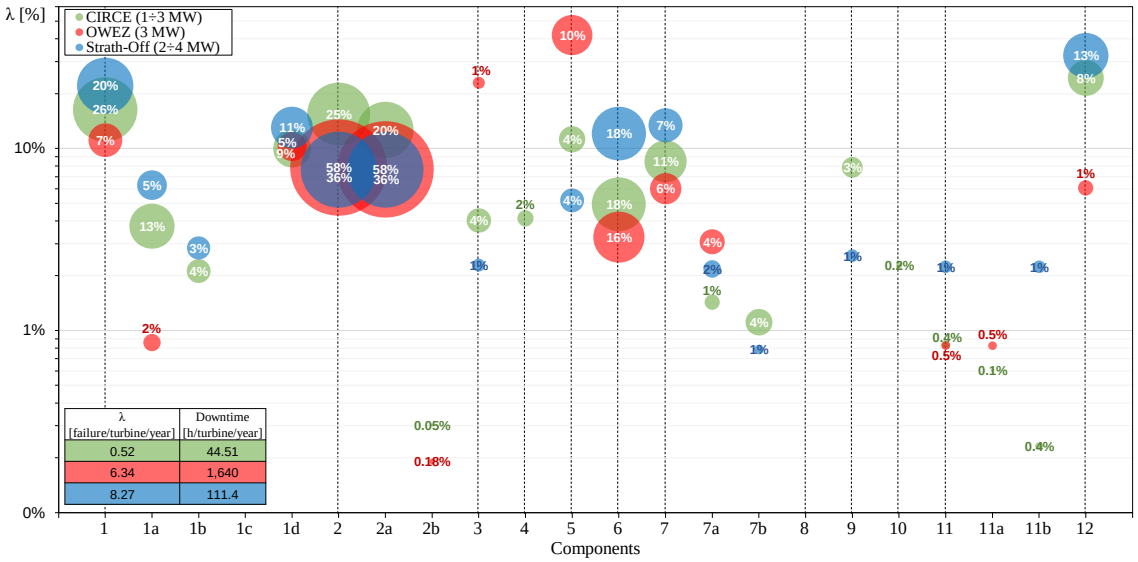


Figure 2.5: Comparison of offshore and onshore reliability statistics for geared turbines with an induction generator. In the x-axis are systems and subsystems, while on the y-axis are the normalised failure rates. The normalised downtime is represented by the bubble size.

The highest share in the failure of the OWEZ farm is in the control system (5). Differently, the Strath-Off (and SPARTA [91]) surveys, recorded more frequent malfunctioning of the pitch (1d) and its hydraulic system (if present). These observations are in line with the results from the onshore surveys. Recently published statistics from the CREW data collection ranked the rotor as the assembly with the highest contribution to the turbine unavailability [110]. In the CIRCE survey, the rotor (mainly for the blade adjustment system and its hydraulics) and control system failures are second only to those of the drivetrain system.

With respect to gearbox (2a), the two offshore studies show a similar percentage share of failure rates, even though the high failure rate of Strath-Off is mainly caused by the inclusion of minor malfunctions, while that of the OWEZ statistics is mainly affected by catastrophic infant mortality failures. It is interesting to observe that, independently of the class of maintenance performed, the gearbox's overall share of the downtime for the offshore population is higher than for all the other components. Therefore, the observed potential high criticality of this component justifies the consistent research effort put into the direct drive designs [110] and advanced monitoring systems and techniques [153]–[155].

Regarding the power generation system (6), it is noticeable that its unplanned maintenance actions have a similar impact on the total corrective downtime of the several initiatives. Nonetheless, the Strath-Off induction generator has a considerably higher share of the frequency of the repair, compared to the CIRCE statistics, suggesting again a potential high influence of the offshore condition on this system. Furthermore, as already shown by [62], the repair rate of offshore generators is significantly higher than that of onshore ones, mainly due to the frequent minor maintenance actions required. Of minor impact for frequency and repair time required are the converters (7a). However, it is worth noticing that both offshore studies recorded a higher share of this subsystem to the failure frequency than in the onshore dataset. Similarly, the repair rate of the converters is ranked fourth by the SPARTA monthly averaged statistics, after issues related to the rotor system. The cause

of their failure lies in the offshore environment and can be either associated with low temperature and/or thermal cycling, as intuitively proposed by [59], or related to other environmental factors (such as humidity) as shown by [37]. Because of the relatively high cost of repair and replacement, and the relatively long repair time [62], components in the transmission system (including the transformer) have a potentially high criticality.

With respect to the structural parts, the unpredicted failures of the rotor system (blades and hub) are shown (1) and known to be critical to the O&M cost of offshore wind projects. On the other hand, it is worth pointing out the potentially high criticality of the foundation systems failures (11b) [156], [157] as opposed to the onshore structure. Despite the little share of the failure frequency, maintenance activities for the inspection and repair of structures below the water level can be significant.

2.6 Conclusions

This chapter presented a comprehensive review and discussion of the identification of critical components of the currently installed and next generation of offshore wind turbines. A systematic review on the reliability, availability, and maintainability data of both offshore and onshore wind turbines is performed, by collecting the results from 24 initiatives, at system and subsystem level.

Trends based on the deployment parameters for the influence of design characteristics and environmental conditions on the onshore wind turbines' reliability and availability are investigated. The main findings from the literature on onshore systems are:

- at a turbine level, an observed generally lower reliability for increasing turbines' size, but a potentially inverse trend in the availability of the wind farms with the number of units installed, independently from their power rating; and
- at a system level, a reduction in the shutdown events linked to the unforeseen failure of the rotor, the drivetrain and the power generator systems are observed as opposed to an intensification in the corrective maintenance of the control, transmission, and blade adjustment systems for an increase in the turbine size; and
- at a turbine level, a higher number of failures is generally recorded for higher averaged wind speed, gust speed, and in presence of other environmental conditions (such as humidity and potentially high turbulence level), with a sensitivity of the drivetrain, power generation and transmission subsystems to these phenomena.

The estimation of the operational availability for a set of offshore wind farm scenarios allowed a comparison with the recently published performance statistics and the discussion of the integrity of the data available to date. The failure statistics of the systems deployed offshore are then discussed and compared to the onshore ones, with regard to their normalised results. The availability calculations supported the hypothesis of the negative impact of the offshore environmental conditions on the reliability figures. Nonetheless, similarities in the reliability figures of the blade adjustment system and the maintainability of the power generation and the control systems are outlined. Thus, with respect to offshore wind turbines, it is noticed that:

- Although generally higher failure frequencies are observed compared to the onshore projects, the recently collected statistics from the industry-led RAM databases show a significant improvement in the operational availability compared to the first generation of offshore turbines.

- A high share of the blade adjustment, drivetrain and transmission systems to the overall failure rate is common to all offshore studies, with the drivetrain and rotor systems being potentially the most critical due to being associated with longer downtime and cost of the repair.

This analysis adds to the existing body of knowledge on the identification of the trends at the turbine, its system and subsystem level, based on the specific design parameters and the deployment conditions – either onshore or offshore, as well as site-specific parameters. This can subsequently allow the development of a better understanding of the sensitivity of certain components and technologies to key design and environmental parameters, facilitating technology qualification of new alternatives and the reliability analysis of the units in a farm.

Chapter 3

Condition monitoring of drivetrains: case studies for model-based and data-driven approaches

Constancy is a great way of de-risking offshore wind, but unfortunately some failures are inherent. To detect the failure of the components of the drivetrain, the mainstream approaches rely on the availability of either a high-fidelity model for the representation of the physics of failure (aka model-based approach), or a representative and big set of data (aka data-driven approach). The first approach requires a deep understanding of the physics of the failure, to reduce the order of the model and to allow improvements of its computational speed – for their application to real-time monitoring task. As concerns the second approach, the limited accessibility to valuable databases of failure challenges the model's diagnostics performances. These two aspects are investigated in this chapter, by considering two case studies for the failure of offshore wind gearboxes. A first study investigates the goodness of some model-order reduction methods in view of the development of a digital twin of an offshore wind farm [1]. A second study is aimed at testing and comparing the effectiveness of a transfer learning model, as opposed to traditional deep learning model, when having a low availability of run-to-failure data [2].

3.1 Background

Offshore wind turbines are complex systems operating in a multi-phase environment, under various field conditions and loads. Their continuously improved and optimised designs can

[1] Z. Lin, **D. Cevasco**, and M. Collu, (2020) "A methodology to develop reduced-order models to support the operation and maintenance of offshore wind turbines," *Applied Energy*, vol. 259, no. 114228. https://doi.org/10.1016/j.apenergy.2019.114228

[2] I. M. Black, **D. Cevasco**, and A. Kolios, (2022) "Deep Neural Network Hard Parameter Multi-Task Learning for Condition Monitoring of an Offshore Wind Turbine," *IOP Conference Series: Journal of Physics*, vol. 2265, no. 3, p. 032091. https://doi.org/10.1088/1742-6596/2265/3/032091

fail suddenly for quality or stress related failures causing a potential great revenue loss. This is particularly true for the complex mechanical components of the drivetrain, which:

- are subject to varying environmental excitations, while coupling and being subject to the aero-, electro- and thermal-dynamic phenomena, and
- are undergoing remodelling and innovations to accommodate bigger power rating.

Several research activities have focused their effort on the investigation of monitoring systems and frameworks for assessing the health status of the offshore wind turbines drivetrain [3]. The CMS in combination with high frequency SCADA data have been extensively used to train data-driven models to learn the normal behaviour, predict the incipiency of a failure, and/or detect the failed status of components, such as the gearbox, the generator, and the drivetrain bearings [40], [41]. In [41], Stetco et al. documented the state-of-the-art ML algorithms and processes for the condition monitoring of wind turbines. Tautz-Weinert and Watson examined and discussed, in [40], the effectiveness of the several diagnostics methods for the condition monitoring of the turbines' mechanical and electrical components, based on the use of SCADA data only. However, it is evident from these literatures that the training of these data-drive models strongly relies on the availability of relatively big sets of data – generally at least for more than 3 turbine years.

In addition to their merely design purpose, the aero-hydro-servo-elastic (AHSE) codes, built to represent the multi-dynamics of the offshore wind systems, can be deployed for the diagnosis and prognosis of critical failure modes – e.g., [158]. The complexity of the dynamics of the different systems can be captured by several level of fidelity in these physics-based numerical models – cf. Paper 10. In the recent years, a considerable research effort has been directed toward the creation of the so-called "digital twin" models. As defined in Section 1.1.1, a digital twin model can build on a physical-based model of a large engineering system, by first applying model-order reduction methods. A digital twin of the torsional dynamic model for the monitoring of the drivetrain failures has been already suggested by Farid et al. [21].

3.2 Problem statement and aim

The components of the speed conversion system (i.e., gearbox) are among the ones with, potentially, the highest impact to the offshore wind farms' OPEX and availability [159], due to their generally low reliability, the high cost of their spare parts and their generally long time to rectify their functionality. The availability of their CMS data, and labelled data indicating their health status, is generally limited. If data are collected, these are generally collected in either short-term or localised (restricted to only few turbine positions) monitoring campaigns, to reduce the cost related to the hardware and the storage of data (aka "on-demand" monitoring). Therefore, the scope of this chapter is to investigate how to build models for the prognosis and diagnosis based on low run-to-failure availability of data. In the specific, the case studies of these chapter are the focused on developing and comparing model-based (reduced order model towards the creation of wind farm digital twin) and data-driven approaches to support the detection of the failure of offshore wind turbines' gearboxes.

The first study is aimed to apply model-order reduction techniques on a full-order nonlinear AHSE code, to capture the key degree-of-freedom (DOF) of an exemplificatory gearbox failure mode. The targeted signal simulated from the reduced-order models (ROMs) is compared against the one outputted from the nonlinear full-order model results.

The scope is to draw conclusions on which method is the most suitable to be applied in this case study, when searching for a good representation of the targeted DOF while reducing the simulations computational time.

The second work is aimed to apply a data-driven multi-task learning (MTL) approach for the monitoring of the health status of an offshore wind turbine gearbox. The purpose is to demonstrate how such a MTL hard parameter sharing approach can achieve greater results than a conventional machine learning (ML) one, when applied to a limited amount of training data. Therefore, the scope of this paper is to make a step forward the understanding of the suitable setup for the training of a satisfactory monitoring algorithms for an offshore case study based on a small set of representative failure data.

3.3 Model-order reduction of wind turbine's dynamics targeting a gearbox failure mechanism

This work is organised as follows. Section 3.3.1 briefly revise the literature on the topic. Section 3.3.2 introduces the methodology, including a description of the nonlinear AHSE model, and the linearisation and the modal truncation (MT) methods applied. In Section 3.3.3, the wind turbine, the failure mode of its gearbox, and a representative set of load cases are selected for this case study. Section 3.3.4 reports the results on the goodness of the ROMs, either linear or nonlinear, to capture the rotor torque signal is compared full-order models. Finally, Section 3.3.5 discusses the computational time reduction achieved by the several ROMs. The conclusions of this comparative study are compiled in Section 3.3.5.

3.3.1 Model-order reduction methods

Model-order reduction methods are well established and widely applied in a variety of areas – e.g. [160]–[162]. These methods include static reduction (e.g., [163]), dynamic reduction (e.g., [164]), and balanced methods (e.g., [165]) techniques among a lot of others. The majority of the ROMs rely on a linearisation of the nonlinear system, and the accuracy of the linear system is limited to a small perturbation around the operating point (OP).

The Guyan's condensation is a static reduction method which focuses on the reduction of stiffness and mass matrix [166]. This method does not lose the information on the structural complexity, and it results in a non-preserved eigenvalue-eigenvector problem, by the coupling of mass and stiffness matrix. An improvement to Guyan's approach is the Improved Reduced System (IRS) method, proposed by O'Callaghan in [167]. This method suggests introducing a correction to the static reduction transformation to account for some allowance of the inertia forces. In [161], Friswell et al. developed a dynamic and iterated version of the IRS method. This method was shown to have some challenges in capturing the low frequencies of the full model with a large centre frequency. Additionally, as for the other dynamic reduction method, the IRS method comes to the same additional computational cost.

The reduction methods mentioned above focus on the reduction of the system's mass and stiffness matrix, derived from the linearisation of the system equations of motion. Based on the linearisation method, the balanced truncation (BT) approach deals with the reduction of the state-space matrix. It transforms the state-space matrix into a 'balanced' form without losing the system's stability and passivity [168]. Based on different numerical

methods, BT can be further divided into Lyapunov balancing, stochastic balancing, bounded real balancing, and others [169].

Currently, the ROMs applied to offshore wind assets are mainly employed in the field of control theory, by focusing on the modelling of the electric and electronic dynamics aspects of the wind turbine. For instance, Ghosh and Senroy [165] applied the BT method to reduce the dynamics model of a farm, by obtaining a ROM that retains the dynamic relationships between the variation in wind speed and the power output variations, accurately over a wide range of frequencies. However, the wind turbine structural and mechanical dynamics are substantially simplified, as they are represented by only one degree of freedom (DOF): the drivetrain rotational speed. Therefore, such approaches do to include the impact of the stochastic nature of the wind speed and the effects of structural dynamics on the drivetrain, generator, and other electrical parts.

3.3.2 Methodology

This section introduces to the approach adopted for the comparison ROMs. It additionally summarises the main modules of dynamics of the AHSE code, and the model-order reduction techniques implemented to decrease the DOFs of the wind turbine system.

Suggested approach

The workflow of this first case study is shown in Figure 3.1, by giving an overview to the inputs, the methods and the processes adopted for this investigation. A nonlinear, full-order AHSE model, which accounts for the simplified dynamics of the modular drivetrain, is used as a basis for the linearisation and the derivation of the ROM (a).

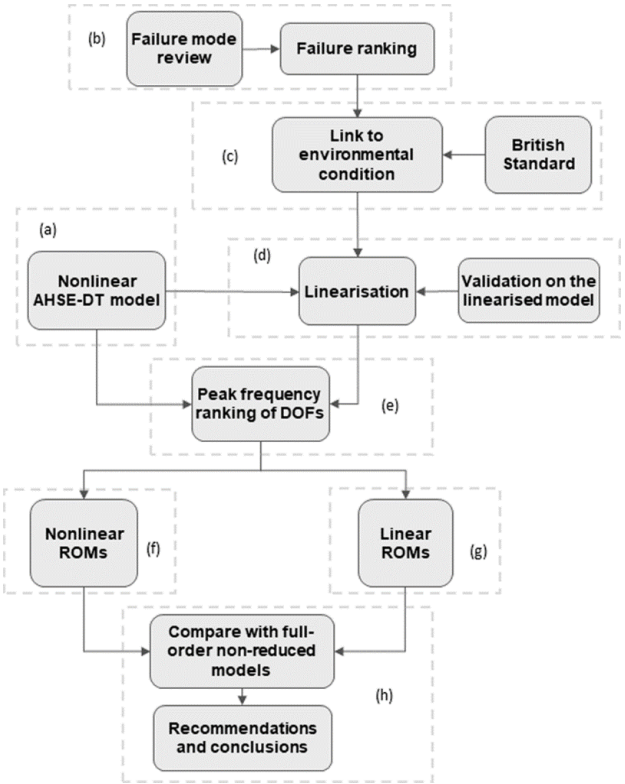


Figure 3.1: Workflow of the inputs, methods, and processes.

A potentially critical failure mechanism of an offshore wind turbine system is selected (b), by following identifying its representative load case(s) and the significant DOF(s) for its representation (c). A linearisation technique is then applied to the nonlinear model, at the operating point (OP) representative of the load case(s) identified (d). The time signals of all the system's DOFs are generated with the nonlinear, full-order model, and they are post-processed to derive and rank the peak frequencies of each DOF (e). A linear ROM is developed, to base on it the MT method (g). In addition to the linear MT ROMs, nonlinear MT ROMs is also investigated (f), to discuss the pros and cons of the linearization approach. The results are then compared by quantifying the deviation of the targeted DOF(s) from each ROM and from the signal(s) generated by the non-reduced model (h).

AHSE code

The FAST code (version 8) is implemented in this analysis for modelling the nonlinear AHSE dynamics of the offshore wind turbine systems, and this section briefly recaps on the modelling approaches for each of the module of dynamics [170].

The aerodynamic of the forces on the single wind turbine are evaluated through the blade element momentum (BEM) theory, widely used in calculating aerodynamics of wind turbines – refer to Paper 10. The BEM theory consists of two coupled sub-theories, the blade element theory, and the momentum theory. The main advantage of the BEM theory lies in transforming the 3D problem into a number of 2D problems, significantly reducing the numerical complexity, especially when compared with more advanced computational fluids methods. Further details on the BEM theory are summarised and reviewed in ***Paper 7***. The influence of the tower on the flow local to the blade is based on the downstream tower shadow model. The wind load on the tower and the tower shadow influence are calculated separately and superimposed.

Table 3.1: Definition of the DOFs for the FAST AHSE code of a bottom-fixed three-bladed offshore wind turbine.

System	DOF	Number
Drivetrain	Rotational	1
	Generator azimuth	1
	Nacelle yaw	1
Tower	1st fore-aft (FA) tower bending	1
	1st side-side (SS) tower bending	1
	2nd FA tower bending	1
	2nd SS tower bending	1
Blade	1st edgewise	3
	1st flapwise	3
	2nd flapwise	3
Rotor	furl	1
Tail	furl	1

A combined multi-body and modal analysis method is applied to determine the structural dynamic responses. The flexible elements, such as tower or blade, are modelled by a linear modal representation. In this full-order model of a three-bladed bottom-fixed offshore wind turbine, there are 18 DOFs – cf. Table 3.1. For each blade, there is one edgewise model and two flapwise modes. Tower flexibility modes are realised through the

fore-aft (FA) modes and the side-side (SS) bending modes, up to the second order. The remaining three DOFs are the drivetrain rotational DOF, the generator azimuth and the nacelle yaw motion DOF. The last two DOFs are for rotor- and tail-furl.

Incident wave kinematics is modelled using Airy wave theory, for both regular and random waves. For regular waves, the wave elevation is modelled by a sinusoid wave with a certain frequency and amplitude. For random waves, superposition theory is applied, for which the irregular wave elevation is represented by a linear summation of different sinusoid waves. Wave forces are calculated using potential flow theory. Considering non-rotational, inviscid, and incompressible problems, the wave-structural interaction problem can be determined by solving the Laplace equation, subject to a series of boundary conditions. The above-mentioned equations are solved numerically via the boundary element method in the frequency domain, considering different wave periods. Based on the Bernoulli's equation, wave diffraction forces are calculated in the frequency domain by an integral over the body surface using another solver, e.g., WAMIT [171]. Based on Cummins' impulse response function theory [172], the first and second-order wave transfer functions are transferred from frequency domain to time domain using the Fourier Transform.

Model linearization

A wind turbine multi-body system is highly nonlinear, and it includes over a dozen DOFs even for a single wind turbine – cf. Table 3.1. The linearisation is used to build a linear time invariant (LTI), input-output equivalent system. It is realised by applying a state-space representation of the nonlinear system at the OP [173]:

$$\begin{cases} \dot{\mathrm{x}} = A\mathrm{x} + B\mathrm{u} \\ \mathrm{y} = C\mathrm{x} + D\mathrm{u} \end{cases} \qquad \textbf{Equation 3.1}$$

The first step for linearisation is to determine an OP. The finding of an OP is important for the linearisation, as the linear system is valid only around a small perturbation from an OP. For the purpose of this study, the OP is determined in the nonlinear time-marching process through each sub-module. Further details on the linearisation for each sub-module and the final matrix assembly can be found in [173].

Modal truncation

There are two types of MT approaches. In the first type, it is assumed that the transient phase of the response of the states characterised by high-frequency dynamics vanishes more quickly than the transient phases of the states characterised by a lower-frequency dynamics. The second type of MT assumes that the derivatives of the high-frequency states are the states to be truncated [174].

As derived in [175], the state-space equation in a vector form can be written as:

$$\begin{cases} \begin{bmatrix} \dot{\mathbf{x}}_1 \\ \dot{\mathbf{x}}_2 \end{bmatrix} = \begin{bmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_2 \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} + \begin{bmatrix} \mathrm{B}_{\mathbf{x}_1} \\ \mathrm{B}_{\mathbf{x}_2} \end{bmatrix} \mathbf{u} \\ \mathbf{y} = [\mathrm{C}_{\mathbf{x}_1} \quad \mathrm{C}_{\mathbf{x}_2}] \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} + \mathrm{D}\mathbf{u} \end{cases} \qquad \textbf{Equation 3.2}$$

For a wind turbine system, it is difficult to justify, at a physical level, the truncation of only the derivative of a state. Therefore, the first type of MT is here adopted. Under this condition, the Equation 3.2 becomes:

$$\begin{cases} \dot{\mathbf{x}}_1 = \Lambda_1 \mathbf{x}_1 + \mathrm{B}_{\mathbf{x}_1} \mathbf{u} \\ \dot{\mathbf{x}}_2 = \mathrm{B}_{\mathbf{x}_2} \mathbf{u} \\ \mathbf{y} = \mathrm{C}_{\mathbf{x}_1} \mathbf{x}_1 + D\mathbf{u} \end{cases} \qquad \textbf{Equation 3.3}$$

3.3.3 Case study

This section reports the characteristics of the offshore wind turbine, taken as reference for this study. It additionally explains and justifies the selection of the failure mode, its related DOF, and the corresponding relevant environmental condition around which the proposed ROMs are developed.

Offshore NREL 5MW wind turbine on monopile substructure

The model of the 5 MW reference wind turbine from the National Renewable Energy Laboratory (NREL) – refer to [176] – is use in this analysis to run the fully coupled AHSE simulations. The main characteristics of this offshore wind turbine installed on monopile foundation are summarised in Table 3.2.

Table 3.2: Main characteristics of the NREL 5 MW wind turbine [176].

Parameters	Value
Rotor Orientation, Configuration	Upwind, 3 Blades
Control type	Variable Speed, Collective Pitch
Drivetrain type	High Speed, Multiple-Stage Gearbox
Rotor, Hub diameter	126 m, 3 m
Hub height	90 m
Cut-in, Rated, Cut-out speed	3 m/s, 11.4 m/s, 25 m/s
Cut-In, Rated Rotor speed	6.9 rpm, 12.1 rpm
Gearbox ratio	97:1

Table 3.3: Natural frequencies of the full-order model.

Mode	Frequency [Hz]
1st blade asymmetric flapwise yaw	0.6664
1st blade asymmetric flapwise pitch	0.6675
1st blade collective flap	0.6993
1st blade asymmetric edgewise pitch	1.0793
1st blade asymmetric edgewise yaw	1.0898
2nd blade asymmetric flapwise yaw	1.9337
2nd blade asymmetric flapwise pitch	1.9223
1st FA tower bending	0.3240
1st SS tower bending	0.3120
2nd FA tower bending	2.9003
2nd SS tower bending	2.9361
1st drivetrain torsion	0.6205

The soil profile is the one of the OC3 Phase II: a three-layered profile representing a medium-density sand. The soil-pile interaction is modelled with p–y curves, with an increasing stiffness proportionally with depth due to the effective stress increase from the weight of soil. To maximise power generation and regulate generator speed, the baseline collective blade-pitch control is deployed. Due to the linearisation approach implemented prior the model reduction, a simple variable-speed generator control is adopted. The generator has a rated speed of 1173.7 rpm and a rated torque of 43093.55 Nm. Further details on the baseline control for a 5 MW reference wind turbine can be found in [176].

For the validation and the comparison of the ROMs, the natural frequencies of the full system are provided in Table 3.3. The smallest and the biggest system frequencies are the one of the tower's bending modes: the 1st SS mode has a value of around 0.3 Hz, while the 2nd SS mode reaches about 3 Hz.

Failure mode

Previous studies showed that the components of the drivetrain have the longest downtime per failure [105]. In [177], Wang et al. claimed that the failures of the wind turbine gearbox are the main responsible of an increase in the downtime and a reduction of the reliability of the asset. This observation is supported by the analysis in Paper 8, which is based on the data collected by Carroll et al. in [62]. Therefore, the improvement of the gearbox reliability and identification of its failure inception have been among the main topics of research for the last 15 year. The Gearbox Reliability Collaborative (GRC) project [65], run by the NREL institute, states that *"most gearbox failures do not begin as gear failures or gear-tooth design deficiencies"*, but the failures *"… appear to initiate at several specific bearing locations under certain applications, which may later advance into the gear teeth as bearing debris and excess clearances cause surface wear and misalignments"*. By the combined use of full-scale dynamometer testing and high-fidelity simulations, the top bearing and gears failure are identified and ranked per criticality in [154].

The most severe failure identified and targeted in this work is the one related to the excessive tooth root bending stress deformation, which proportional to the transmitted contact force (F_t) [178]. The rationale behind this choice, over the possible failure of the rolling element of the bearing – as shown in [179] for the roller and ring cracks, and in [154] for their wear –, is:

- the highest severity of the consequences of the failure. The gear wear and corrosion for contact stress, although more likely to happen [32], leads only to higher vibration and noise [35]. The high bending stress on the tooth from either high cycling fatigue or extreme loads, leads to tooth deformation and cracking, up to its final fracture.
- the identification of the possible environmental causes and related DOFs is more straight-forward for gears than bearings.

To support the latter observation, it must be noticed that, to identify the fundamental DOFs driving the failure, simplified physics considerations are preferred to more complex modelling. The high complexity models – from multibody to finite element models – are generally used for the bearings dynamics and loads calculation [180]. By contrast, in [158], Nejad et al. showed that a simplified model can be sufficient to derive the main gear dynamics without the need for the multibody model. In these works, the simplified formulation for the transmitted contact force, expressed by assuming rigid bodies and contacts, zero damping gears, and neglecting their internal dynamics is:

$$F_t = \frac{2}{d}T \qquad \textbf{Equation 3.4}$$

The input torque (T) to the gearbox – here on referred to as "rotor torque" – is, thus, the DOF linked to the detection of the exceedance of the transmitted contact force, when considering the gear at the first stage of the gearbox – shown to be the one with higher criticality [158].

Load cases

A failure of the tooth root bending stress deformation in the gears of a wind turbine gearbox can be caused either by the high cycling fatigue or by the loads in extreme operating conditions. In [181], Nejad et al. observed that wind speeds near cut-out hold the highest contributions in extreme loads on gears in normal operations. Dabrowski et al. reached the same conclusion, while conducting a study on the implementation of a storm controller for the reduction of the probability of gear tooth failure in the first planetary stage of the NREL 5 MW wind turbine [182]. Independently on the type of controller, they observed that already at operating conditions, the transmitted force at the tooth has higher averaged values in the case of extreme turbulence model (ETM) employment. Dealing with a bottom-fixed structure, the influence of the wave excitation on the drivetrain load can be neglected [182]. Additionally, Nejad et al. showed that the load effect of the drivetrain components are mainly dependent on the axial force, the tower bending moment (a function of the thrust force), and the shaft torque (influenced by the power control system) [183].

For these reasons, a wind-only load cases with the wind turbine operating under ETM is selected for running 10-minute simulations. For the demonstrative purpose of this study, the set of possible wind speed is reduced to one load case only, for a wind speed at hub-height of 24 m/s. For the linearisation, steady wind conditions are run. The stochastic inflow turbulent wind files are generated with the IEC Kaimal wind spectra model by running TurbSim tool, a computer-aided pre-processing software by NREL [184].

3.3.4 Results

In this study, the full-order model of a three-bladed bottom-fixed offshore wind turbine, accounts for 16 DOFs, by excluding the rotor- and the tail-furl, generally used for modelling unusual bearing configurations – cf. Table 3.1. This section presents the models and the results of the simulations for deriving the drivetrain torque, by comparing the DOFs signals of the full-order nonlinear models to the one of the ROMs.

Linear ROMs

The first step for the linearisation is to determine an OP, which is a periodic steady-state condition of the system. At the OP, derivatives of the states are zero. It is important to set up the OP accurately as the linear representation of the nonlinear system is valid only for small perturbations around the OP. For setting up the OP, an initial rotor rotating speed equals 12.1 rpm is selected. As the steady horizontal speed is 24 m/s for ETM condition, the wind turbine operates in Region III [176]. By applying the baseline control, the corresponding blade pitch angle is 22.35 degree. After determining the OP for the full-order model, a single input single output (SISO) system with 32 states is set up; the input is the wind speed, and the output is rotor torque – cf. Figure 3.2. As the wind turbine blade was

rotating during operation condition, the linearisation is carried out every 36-azimuth angle, in a total number of 10 realisations.

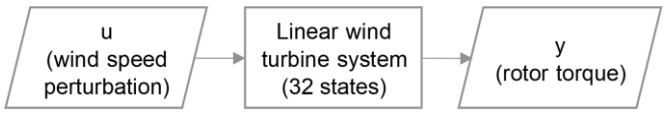


Figure 3.2: Linearised system input and output.

For the validation of the linearised model, a mono-direction step wind profile is simulated in a 650 s long run: the wind speed of 24 m/s is stepped up to 25 m/s at 300 s. The linear model shows a good agreement with the nonlinear results at steady wind condition for either time histories of the rotor torque and tower top fore-after deflection – cf. Figure 3.3. However, the linear results are confirmed only around the OP; thus, a larger discrepancy between the nonlinear and linear results is evident a is at 25 m/s steady wind speed – e.g., Figure 3.3 (b).

(a) Rotor torque **(b) Tower top FA deflection**

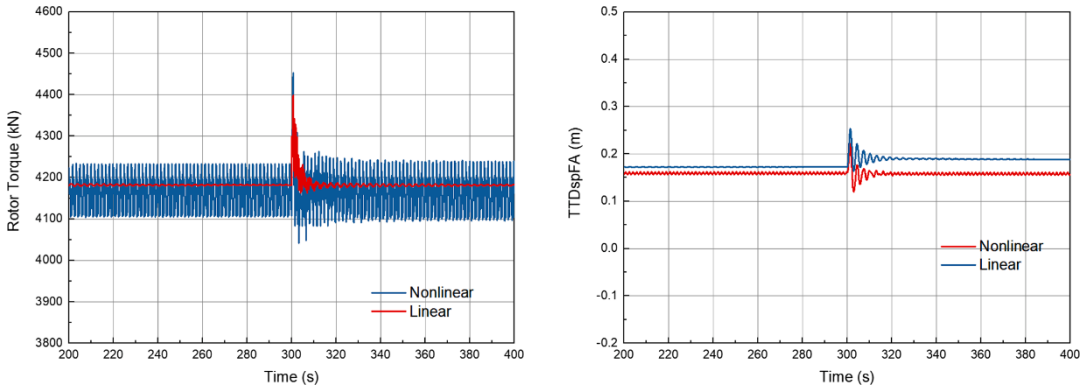


Figure 3.3: Comparison of (a) rotor torque, and (b) tower top FA deflection for the full-order nonlinear model and linear one, in response to a mono-direction step wind profile.

(a) Rotor torque **(b) Tower top FA deflection**

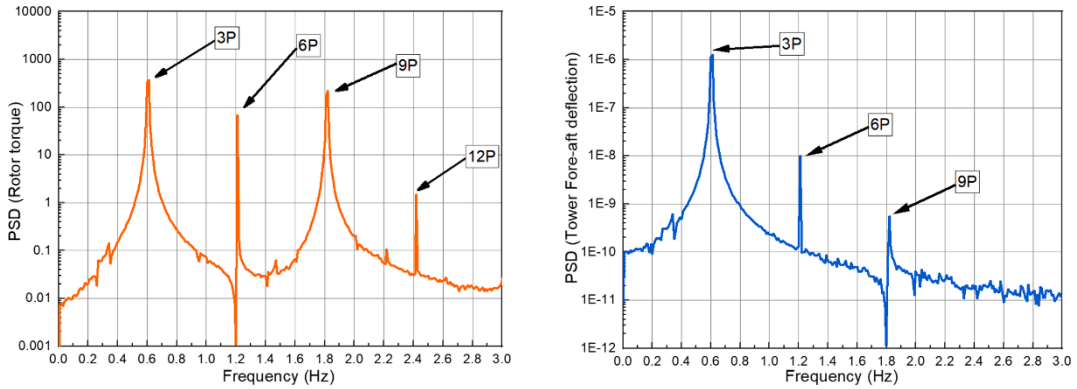


Figure 3.4: Spectra of (a) rotor torque, and (b) tower top FA deflection for the full-order nonlinear model at 24 m/s steady wind condition.

To speculate on these results, the spectra of these two signals are reported in Figure 3.4. These signals are obtained by running the nonlinear model with steady wind inflow at

24 m/s, and by applying the Fast Fourier transform (FFT) to derive their power spectral densities. A possible reason for the higher oscillation of the signals of the nonlinear model, is the resonant response with the frequency of the blade passing in front of the tower (i.e., 3P, 6P, 9P) – cf. Figure 3.4 (b). By contrast, this phenomenon cannot be well captured by the linear model.

ROM based on MT method

The workflow of the process to apply the MT method to the AHSE model to reduce its order is shown in Figure 3.5. The wind-only ETM load condition is used to run the full-order AHSE nonlinear model. The peak frequency of each of the output DOFs is identified through an FFT-spectral analysis. An example of how the peak frequency of 0.34 Hz is identified for the 1st SS tower bending mode, is shown in Figure 3.6, together with the ordered ranking of the peak frequency for each of the mode considered in this analysis.

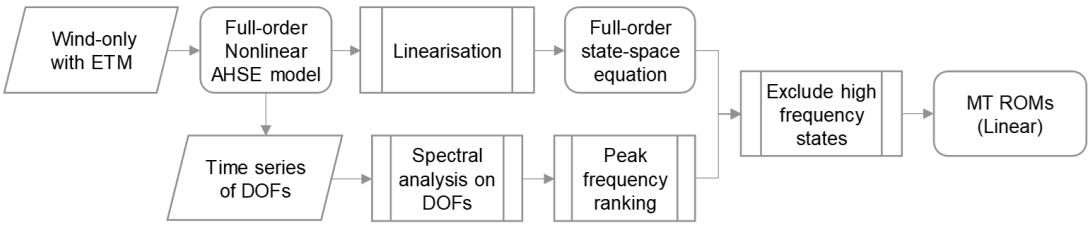


Figure 3.5: Workflow of linear MT ROM.

By applying the linear SISO model with 32 states, the MT ROMs are developed by excluding high-frequency states. Table 3.4 describes the corresponding DOFs enabled in MT1, MT2 and the full-order model. To comply with the aim of the study, and the need to model a variable speed generator and a drive train with a gearbox, the DOFs linked to the generator and the drivetrain are always included. Apart from the drivetrain rotational, the nacelle yaw, and the generator DOFs, the MT2 model includes the 2nd tower bending modes. On the other hand, the MT1 model includes all of the DOFs of the MT2 model and three blade modes.

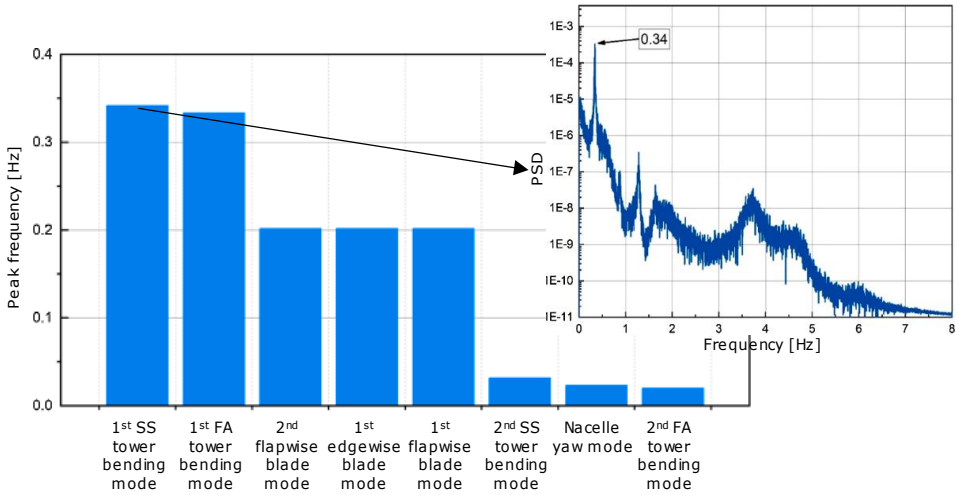


Figure 3.6: Definition on the peak frequency of each mode (explicit plot of the peak frequency picking for the 1st SS tower bending mode).

Table 3.4: List of DOFs considered in the models – i.e., full-order, and MT ROMs.

System	DOF	Model			
		Full-order	MT1	MT2	MT3
Drivetrain	Rotational	✓	✓	✓	✓
	Generator azimuth	✓	✓	✓	✓
	Nacelle yaw	✓	✓	✓	✓
Tower	1st FA bending	✓			
	1st SS bending	✓			
	2nd FA bending	✓	✓	✓	
	2nd SS bending	✓	✓	✓	
Blade	1st edgewise	✓	✓		✓
	1st flapwise	✓	✓		✓
	2nd flapwise	✓	✓		✓

Comparison against full-order models

Figure 3.7 visualise the results of the first two linear MT ROM (MT1 and MT2) against the linear and nonlinear full-order models. Due to the randomness in the numerical simulation of the turbulent wind time series, each load simulation includes five realisations. Each load case results are then reported by averaging the output across these five signals. The initial 50 s of the rotor torque time series is excluded from the analysis, representing the transient part of the numerical simulations. The rotor torque signal, as simulated by the ROMs including 14 DOFs (MT1) and 5 DOFs (MT2), is compared against the one simulated by the full-order models, with 16 DOFs (i.e., full-order nonlinear and full-order linear). Due to the periodic steady linearisation process, it is difficult to include blade pitch control into the linearisation study. For this reason, the comparison is performed by additionally distinguishing between the full-order nonlinear model results carried out with and without blade pitch control. As it can be observed in Figure 3.7, the torque signals simulated by the full-order linear model and the linear MT ROMs have a generally good agreement with the nonlinear full-order model.

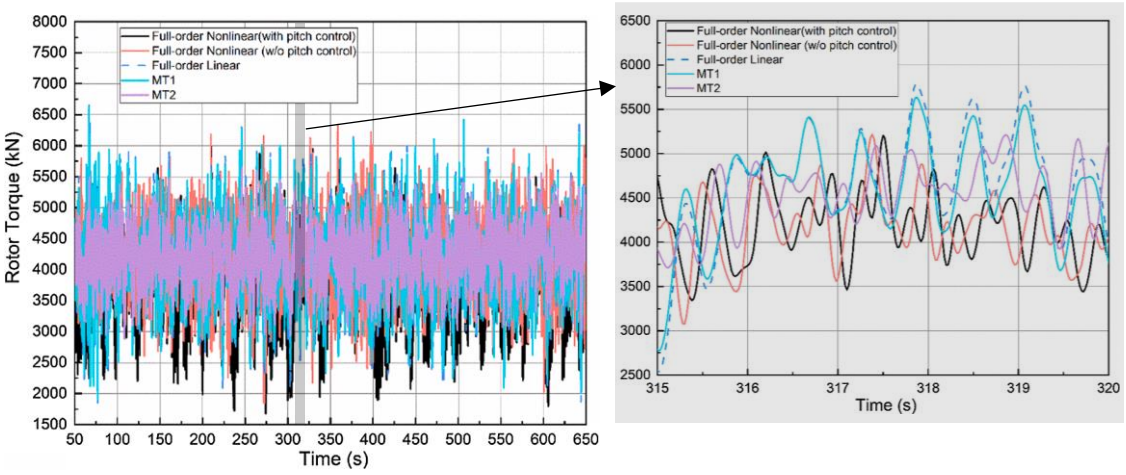


Figure 3.7: Comparison of rotor torque time series (one realisation only) across nonlinear full-order and linear models. On the right, a close up for 5 s of simulation.

To ease the comparison, the main statistics of the torque signal – i.e., minimum (Min), maximum (Max), standard deviation (StD), and mean value (Mean) – are plotted, in Figure 3.8, in a bar plot across the several models. By excluding the pitch control, the rotor torque Mean is 10% larger than the one of the full-order nonlinear model. Compared to the full-order nonlinear model with pitch control, the extreme values (Min and Max) of the linear MT1 model have only a little discrepancy. Furthermore, by comparing the linear MT2 model to the full-order linear and nonlinear models, it is evident that the removal of the 1st tower bending modes has no significant effects on rotor torque Max. In other words, when tackling the rotor torque Max, the 1st tower bending modes is not strongly coupled with other modes.

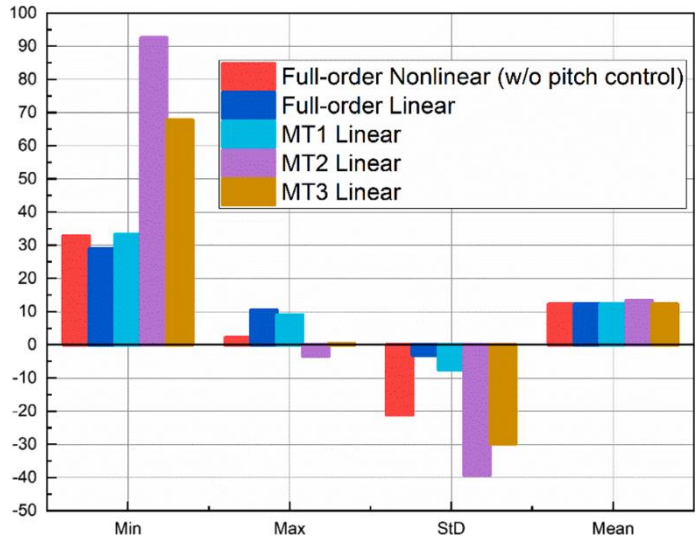


Figure 3.8: Comparison of statistical data of the models, normalised against the ones of the full-order nonlinear model (with pitch control enabled).

(a) Rotor torque **(b) PSD rotor torque**

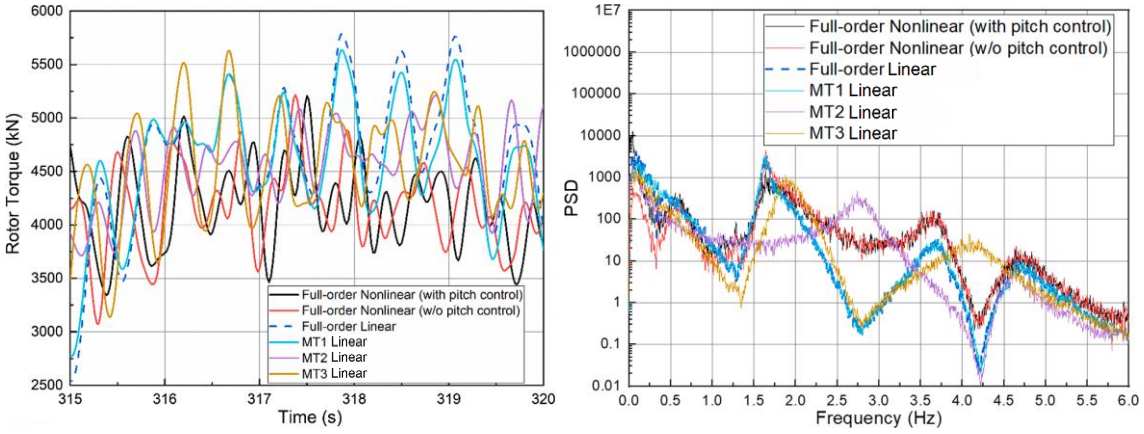


Figure 3.9: Comparison of rotor torque time series (one realisation only) across the nonlinear full-order and the linear models: (a) a close up for 5 s of simulation, (b) spectra of the signals.

The spectrum of the torque signal obtained with the linear MT1 model shows a good agreement with the one of the full-order linear model, except for frequencies below 0.5 Hz – cf. Figure 3.9. This is due to the fact that the linear MT1 model does not consider the 1st tower bending modes. Although the linear MT2 model has a relatively good agreement with the full-order nonlinear model's results in terms of Mean, its results are inaccurate for frequencies above 0.5 Hz. This discrepancies might be related to the inconsistency of the

wind input between linear and nonlinear approaches, and the fact that the MT2 model does not account for the blades DOFs. In the nonlinear models, the turbulent inflow wind is generated in TurbSim under ETM condition with the IEC Kaimal spectrum. Although the wind direction remains unchanged, the turbulence influences the blades motion, and eventually it influences the corresponding rotor torque signal.

The mean squared error (MSE) – as defined in Section C.1 of the Appendices – is used to further compare of the spectra of the rotor torque signal simulated by the several models. The PSD MSE metric, together with all the other signal statistic are reported in Table 3.5. The smallest is the MSE, the closest is the rotor torque to the one of the full-order nonlinear model, which is taken as term of comparison.

Table 3.5: Statistics of the rotor torque (kN) time series across the several models: nonlinear full-order models, linear full-order model, and linear MT ROMs. The MSE of the rotor torque spectrum is calculated by taking as reference the full-order nonlinear model's output.

	Nonlinear		Linear			
	Full-order with pitch control	Full-order without pitch control	Full-order	MT1	MT2	MT3
Min	1447.76	1992.60	1866.00	1928.39	2787.20	2428.94
Max	6010.20	6140.00	6636.13	6554.89	5802.28	6033.15
StD	671.52	529.95	650.44	621.77	407.73	471.38
Mean	3726.29	4181.41	4182.77	4183.03	4222.30	4182.17
MSE	-	0.27	1.29	1.64	4.99	91.26

Among the MT ROMs models, the MSE of the linear MT1 model is smallest one. The MSE then increased to 4.99 for the linear MT2 model. To verify the speculation on the link of this linear MT2 model underperformance to the exclusion of the blades' DOF, a linear MT3 model (another linear MT ROM) is built by excluding the tower DOFs, but by including the blades and the yaw DOFs – cf. Table 3.4. As it can be observed in Figure 3.9, the difference between the torque spectra obtained with the linear MT3 model and the one of linear MT1 model are relatively small. However, the output of linear MT3 model is characterised by a relatively higher MSE – cf. Table 3.5. This can be explained by the difference in response of the linear MT1 and MT3 models, for frequencies below 2.5 Hz, where the absolute value of the power is much higher than that the one of the other frequency ranges.

This study shows that not only the coupling effects between DOFs but also the linearisation process contribute to the differences between the linear ROMs and the full-order nonlinear models, as the linearisation is only valid for small perturbation of the OP. Because nonlinearities may influence the dynamics of wind turbine structures, especially in the low-frequency response, the next section investigates the nonlinear ROMs.

Nonlinear ROMs

As discussed previously, the advantage of a ROM based on the application of a MT method is its capability of selecting which state to include in, and/or exclude from, a state-space model. A state in the state-space equation corresponds to a DOF in the nonlinear motion equation. The MT approach followed excludes the high-frequency states from the nonlinear equation of motion, and thus the corresponding DOFs.

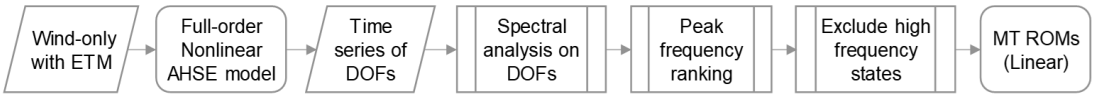


Figure 3.10: Workflow of nonlinear MT ROM.

The workflow in Figure 3.10 shows the process to derive the nonlinear MT ROMs. The signal of the rotor torque is generated by running a full-order nonlinear model simulation in a wind-only with ETM. By using the same peak frequency ranking, as showed in Figure 3.6, the nonlinear MT ROMs are developed. The DOFs enabled in each ROM are according Table 3.4. As opposed the linear model analysis, only the full-order nonlinear model enabling the pitch control is used as term of comparison.

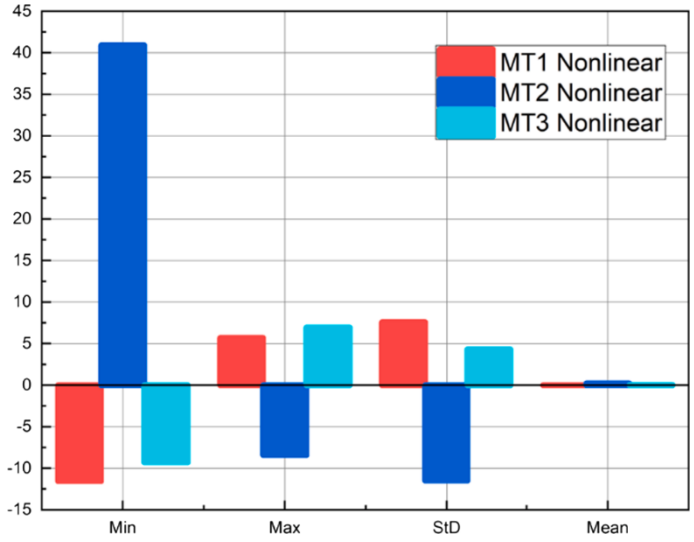


Figure 3.11: Comparison of statistical data of the models, normalised against the ones of the full-order nonlinear model (with pitch control enabled).

(a) Rotor torque **(b) PSD rotor torque**

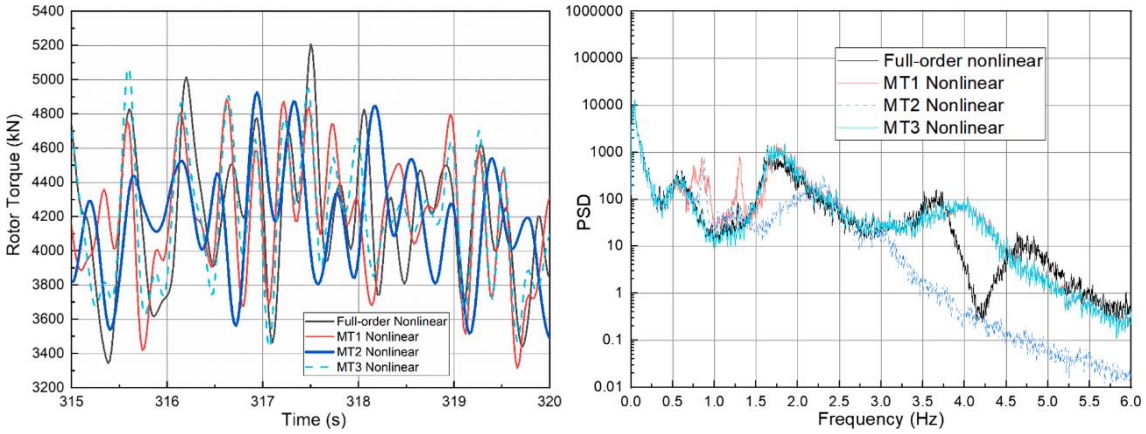


Figure 3.12: Comparison of rotor torque time series (one realisation only) across the nonlinear models: (a) a close up for 5 s of simulation, (b) spectra of the signals.

As it can be observed in Figure 3.11, the Mean of the rotor torque resulting from all the nonlinear MT ROMs are in very good agreement with the one of the full-order nonlinear model. The rotor torque time signal simulated with the nonlinear MT2 model, has the

highest discrepancy to the full-order nonlinear model – i.e., higher variation of the signal than any of the other nonlinear MT ROMs. This behaviour has been already observed in the linear MT2 model – cf. Figure 3.8. Therefore, it can be inferred that the exclusion of the blades modes has a high impact to the rotor torque response, independently on the model linearisation. However, the use of a nonlinear model prior the application of the MT method seems to be slightly beneficial for this application. As concerns the nonlinear MT3 model, it can be noticed that its simulated rotor torque is in better agreement with the one simulated with the full-order nonlinear model, than by using the nonlinear MT1 model – which additionally includes the tower's DOFs.

By looking at the spectra, it can be observed that the nonlinear MT3 model has the best agreement with the full-order nonlinear model. This is confirmed by the MSE in Table 3.6, where it is evident that the output from the nonlinear MT1 and MT2 model has a much different energy content than the one of the full-order nonlinear model. It is evident, from Figure 3.12, that the discrepant peaks of the nonlinear MT1 and MT2 model, in the frequency range between 0.5 and 1.5 Hz, have a greater contribution to the absolute value of the power. The spectrum of the nonlinear MT3 model shows instead discrepancies for frequency values above 3 Hz, which however have a smaller contribution to the MSE due to the much smaller value of the power spectra.

Table 3.6: Statistics of the rotor torque (kN) time series across the several models: nonlinear full-order models, and nonlinear MT ROMs. The MSE of the rotor torque spectrum is calculated by taking as reference the full-order nonlinear model's output.

	Full-order with pitch control	MT1	MT2	MT3
Min	1447.76	1280.66	2040.00	1313.36
Max	6010.20	6352.60	5503.80	6426.20
StD	671.52	722.40	594.44	700.49
Mean	3726.29	3726.34	3726.34	3727.57
MSE	-	217.27	188.99	2.03

3.3.5 Discussion on reduction of computational time

The computational-cost benefits of applying model-order reduction methods for this case study is presented in terms of simulation time in Figure 3.13. Without considering the time taken by the linearisation process – i.e., to generate the state-space matrix, which has to be done only once –, the simulation time required for the run of the linear models is only around 7% of the time required for the full-order nonlinear model. The simulations with the nonlinear MT ROMs take almost as long as the full-order AHSE code. The processing time for the nonlinear full-order model, on a Windows personal laptop (Intel ® Core i7-3770 CPU @ 3.40 GHz, 8 GB RAM), is 8.06 min for a 10.83-min real-time simulation.

The simulation time of the most performing among the MT ROMs – nonlinear MT3 model – is around 0.32 min shorter than the full-order, non-linear model. Although this might not seem a substantial reduction of the computational time, it should be considered that this study is relative to a single wind turbine, and a single load case only. The currently planned and future offshore wind project consist of hundreds of offshore wind turbines. Furthermore, a predictive model for the diagnostics and prognostics of an offshore wind turbine drivetrain requires for several load cases and coupled dynamics. To move towards

the development of a digital twin for the detection of the drivetrain failure at the wind farm level, even small-time savings can have a great impact.

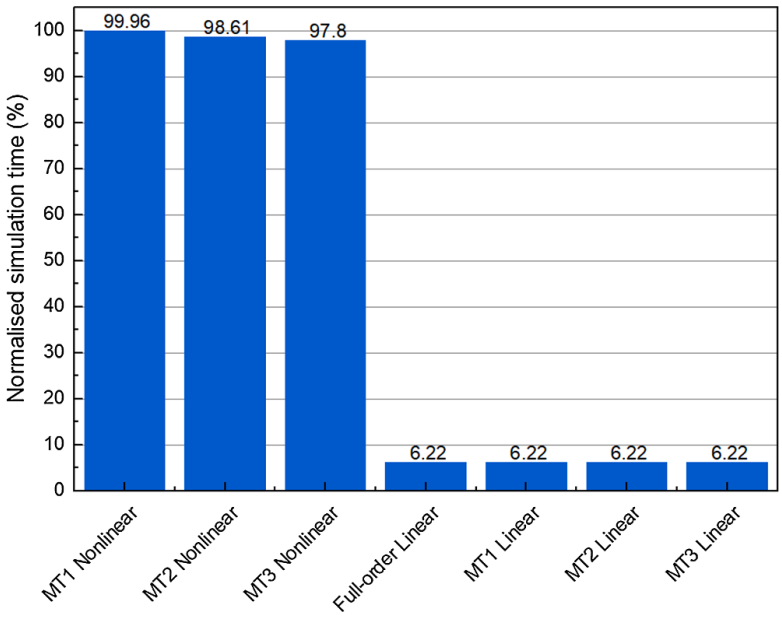


Figure 3.13: Normalised simulation time against the one of the full-order nonlinear model.

3.4 Data-driven detection of gearbox health status based on low availability of run-to-failure data

This work is organised as follows. Section 3.4.1 provides a brief overview of the deep ML models and the theoretical basis of the MTL approach. Section 3.4.2 presents the case study of this alarm-identification task, by transferring the knowledge between two offshore wind turbines (of 1 MW rating) from the same homogeneous population. In Section 3.4.3, the approach followed for the comparison of the performance of a MTL model against the conventional ML model is introduced. This section additionally provides details on the methods applied for the data pre-processing and the model training. Section 3.4.4 reports the regression and classification results of the models. Finally, in Section 3.4.5, the benefits and limitations of the MTL approach are discussed.

3.4.1 Multi-task learning

The machine learning algorithms that adopt shallow structures generally struggle to acquire abstract features minimizing the probability distributions among varying domains. Mathematically, there is a neutral network (NN) architecture function $f(x) = y$ that can map the attributes $x = x_1, x_2, ..., x_n$ to the output y. The Universal Approximation Theorem states that there is an architecture, no matter what $f(x)$, that can approximate y; however, the accuracy is dependent on the architecture and the dataset [185]. With the computers becoming more and more powerful, the large NN architectures – also called deep neural networks (DNN) –, and the convolutional neural networks (CNN) are showing significant promise as opposed to shallow networks. The main differences between shallow networks

and the DNN – or the CNN – is in the architectural shape and in an increased number of hidden layers.

A conventional ML typically involves optimising for a particular task $T = (y, f(x))$, where y is the output feature domain and $f(x)$ is the predicative function made up of X feature data. The model that is trained for a single task may achieve an acceptable performance for a single domain $D = (x, p(x))$ of marginal probability distribution $p(x)$. However, by focusing on one signal task, it might be that other information – which may help the model to perform better on other metrics related to that task – goes overlooked. In a multi-task learning (MTL) approach for DNNs, a model seeks to learn shared hidden representations on different datasets and for the different tasks. The most widely used approach for MTL with neural networks is hard parameter sharing [186], in which the models learns a common space representation on a source dataset by completely sharing weights and/or parameters between the different tasks. Task-specific layers are then used to model the different tasks, by learning independently for each tasks based on the same source dataset or other specific target datasets [187].

The MTL models come in a plethora of forms: joint learning, learning to learn, and learning with an auxiliary task, are among some of the names that have been allocated to its predictive assignment. To generalise the need for its application, it can be stated that a MTL approach is worth being investigated as soon as the problem requires optimising for more than one task. The MTL approaches have been employed across many fields of application requiring a supervised prediction of one or more classes. For instance, they have been deployed for image quality assessment [188], facial recognition applications [189], and deep reinforcement learning [190].

3.4.2 Case study

The analysis presented in this case study is built on time series data from 1 MW offshore wind turbines, in normal operation. The signals from the SCADA and CMS consist of 8 monitoring channels, recording the 10-minute mean. These channels include meteorological information, the operational data of the wind turbine, and the vibrational data from the gearbox, with the associated warning raised in case of anomalies. This latter provides the (binary) label targeted in the training of the classification models. For practicality reasons, these channels – the features of the ML models – are following referred to according to the below numbering convention:

(1) Power operating regime [-]
(2) Gearbox gear oil temperature [°C]
(3) Power [W]
(4) Generator speed [RPM]
(5) Wind speed [m/s]
(6) Gearbox vibration root mean square (RMS) [-]
(7) Gearbox vibration peak-to-peak (P2P) [-]
(8) Gearbox alarm (0 or 1) [-]

A dataset containing 31,804 time-steps (ca. 220 operational days) is collected from a turbine and it is referred to as source domain data. Another turbine from the same homogeneous population – whose distributions of the individual features are similar – is characterised by a smaller dataset of 8,141 time-steps (ca. 56 operational days), which is referred to as target domain data. The source domain data has alarm signatures for 19% of

the data in the set, and the target domain data has a similar a portion, with 18%. These two datasets are representative of an existing wind turbine and a newly installed turbine.

3.4.3 Methodology

The approach and the methods used for the training and the comparison of models predicting, in binary form, the status of an offshore wind turbine gearbox is presented in this section.

Suggested approach

Two different models and two sets of data are investigated in this study. Figure 3.14 reports the workflow of the data and their deployment for the training of the predictive models. On one side, a conventional model is trained and tested on the small set of monitoring data. On the other, a MTL approach is set up by merging a feature extractor – trained on the large set of monitoring data –, to a classifier which predicts, in binary form, the status of the gearbox.

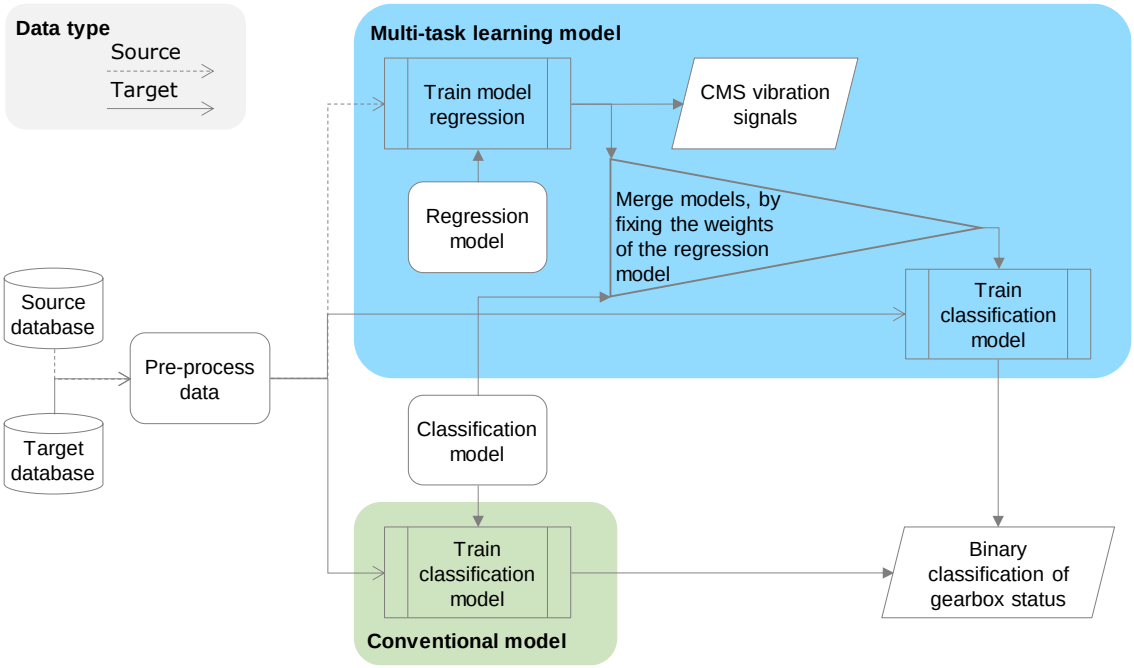


Figure 3.14: Workflow for the training of the conventional and the MTL models.

The MTL model is built up in two training stages, and by using both datasets. The knowledge acquired from the source domain data is transferred to the learning from the target domain dataset, to leverage the prediction of the gearbox status. At first, a feature extractor (i.e., regression model) is trained on the larger set of monitoring data. This model uses the SCADA data and makes predictions on the vibrational data from the gearbox. A classifier is then merged to this first model, by using a hard parameter sharing and connecting the last neurons from the artificial NN to the first layer of a CNN (i.e., classification model). The weights of the neurons of the regression model are fixed, and the entire model is trained using the target domain dataset, to determine whether the gearbox is running into a problem.

The conventional model takes the same architectural form as the artificial NN regressor and the convolutional neural network classifier together, but it is trained by using only the

data from the target domain dataset – i.e., the smaller dataset. This classification model receives the CMS and SCADA data in input, and it outputs the gearbox binary status.

Data pre-processing

The datasets used for the training and testing of a machine learning algorithm generally require some pre-processing to ensure good performances of the prediction model. For the application in this study, this consists of a two-step procedure.

First, a data cleaning process is performed for removing the outliers and the missing values. For the treatment of the missing values, time instances of the database are removed from the analysis if over the 50% of the data is missing. The outlier removal, this is processed by removing vectoral instances where the values should be scalar. For the remainder of the data, a K-nearest neighbours imputation method was applied [191].

The data, as pre-processed at the first step, are then split into training and testing sets. In the specific of this experiment, the 80% of the data is employed for the training of the models, of which the 40% of it to be used for the models' validation (parameter tuning). The remaining 20% is used for the final testing.

Machine learning algorithms and metrics

Three models are trained for the purpose of this comparative analysis: a feature extraction model, the classification model used for the MTL procedure, and a conventional classification model.

The feature extractor takes as input the meteorological data, the wind turbine operational data, and the gear oil temperature, and then outputs the gearbox vibration features. This model, having a DNN architecture, consists of 14 sequential layers, all implementing a rectified linear unit (RELu) activation function. The utilisation of a uniform variance scaling allows the NN to train extremely deep rectified models directly from scratch [192]. The optimiser used for the regression feature extraction is the so-called ' Adam', which is effective for noisy, nonlinear data and it is derived from the adaptive moment estimation [193].

The MTL with hard parameter sharing is carried out by merging the classification model to the feature extractor model. This first one has a CNN architecture, featuring one convolutional layers of width 64 and utilising the RELu activation function. This is followed by a drop out of 0.5 to further four layers of a one-dimensional convolution, which implements the sigmoid function - commonly used in classification models. The kernel initialisation of the weights for the CNN uses uniform variance scaling. A standard gradient descent with Nesterov Momentum is employed to improve the accuracy while dealing with noisy data from the vibration signals, with a learning rate of 0.1 and a momentum of 0.9. Lastly, the cross-binary entropy loss function is implemented to distinguish the gearbox status class.

For a consistent comparison, the conventional classification model has the same architecture of the feature extractor and the classification model together. However, the main differences to the hard parameter sharing model lies in:

- the binary cross entropy optimiser is applied to the whole model,
- the entire model is trained in one process and one dataset, and
- the model only has one output stream of information representing the gearbox status class.

3.4.4 Results

The correlation plot in Figure 3.15 reports the Pearson's correlation of the dataset signals. In machine learning applications, it is generally the case that, the higher is the correlation between the data, the better is the chances of the predicative character of the model on the targeted task. It can be observed that the first five features have no relation to the gearbox alarm (8). On the other hand, the correlation between the vibration data – features (6) and (7) – and the gearbox status is higher.

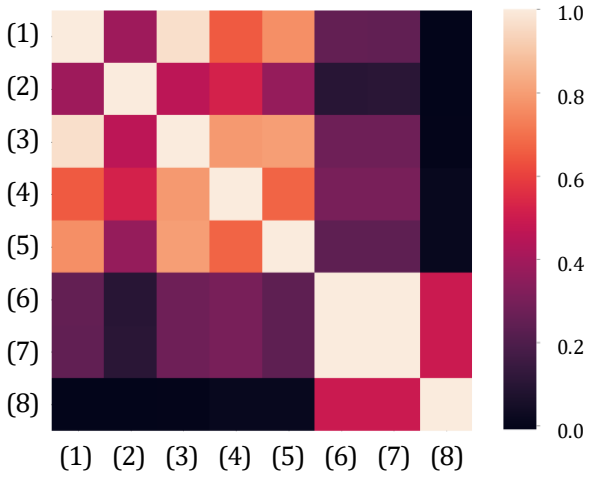


Figure 3.15: Linear correlation for all the features, the first five features are the inputs for both the regression model and the traditional model. The vibration signals are the outputs of the regression aspect hard parameter transfer model, and the Error is the output of the full hard parameter transfer model and the conventional model.

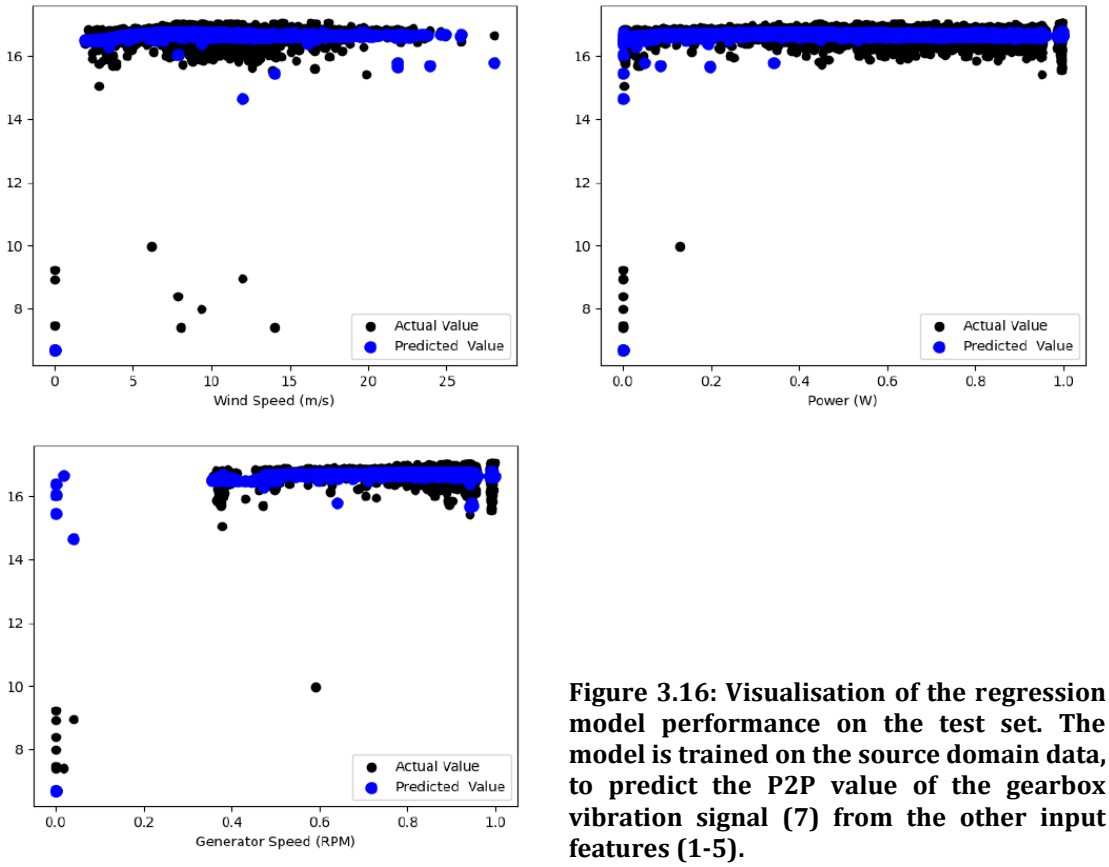


Figure 3.16: Visualisation of the regression model performance on the test set. The model is trained on the source domain data, to predict the P2P value of the gearbox vibration signal (7) from the other input features (1-5).

The regression model makes use of the SCADA data from the larger, source domain data set to predict on the CMS gearbox vibration – targeting channel (6) and (7). The training stage is carried out over 1,000 epochs having a total of 388,803 trainable parameters. To validate the prediction of the model new, unseen data from the wind turbine is fed into the model. In the validation phase the best performing model scores a *MAPE* of 27.00% and a *R2* of 68.61% – see Section C.1 of the Appendices, for the definition and explanation of these regression metrics. These scores are deemed sufficient for the purpose of this process: fix and merge the weights for the MTL hard parameter sharing. The optimisation of the models, and of the model for the prediction of the gearbox vibration, is not the scope of this comparative study. A visualisation of the predictions of the model on the test set is reported in Figure 3.16, for the targeted channel (7).

Table 3.7: Performance metrics of the MTL and conventional classifiers on the test set of the target domain data.

	acc	***F1-score***
Conventional classification	83.76%	57.56%
MTL hard parameter sharing	91.29%	69.54%

The performances of the two classification models are reported in Table 3.7 – see Section C.2 of the Appendices, for the definition and explanation of these classification metrics. The MTL classification model trains 83,329 parameters over 1,000 epochs. Compared to the feature extractor, this model uses significantly less data points for its training phase, and it scores an accuracy (*acc*) of 91.29%, and an *F1-score* of 69.54% on the test set of the target domain data. The conventional classification model undertakes the same training process as the MTL one, and it scores an *acc* of 83.76%, and an *F1-score* of 57.56%, on the test set of the target domain data.

3.4.5 Discussion and comparison of predictive models

It is evident, from Table 3.8, that the MTL model scores a better accuracy and *F1-score* than the conventional one on the test set of target domain data, despite being trained on the same dataset for the classification task.

The comparison continues, in Table 3.8, on the results related to the set of test data of the source domain dataset. This table additionally reports the false positive rate (*fa*) – see Section C.2 of the Appendices for the definition and explanation of the classification metrics. In this case, it can be observed that, overall – in terms of accuracy and false alarms rate –, the MTL model still performs better that the conventional one in the classification task. However, on the set of unforeseen but homogeneous data the MTL model forecast has a lower detection rate – cf. the *F1-score* – than the one of a conventional model.

Table 3.8: Performance metrics of the MTL and conventional classifiers on the test set of the source domain data.

	acc	***F1-score***	***fo***
Conventional classification	80.87%	46.17%	13.17%
MTL hard parameter sharing	86.68%	36.46%	9.93%

An important aspect worth mentioning, and that has not been investigated in this study, is the effect of the varying the sample size of the failure data in these unbalance datasets.

Furthermore, there might be a crossover point where a conventional model might outperform the model built with a MTL approach. The current understanding of a conventional machine learning model is that the larger are the data and the network, the greater is its accuracy.

3.5 Conclusions

To deal with two case studies related to the offshore wind turbine gearbox failures and the lack of a sufficient set of run-to-failure data, two approaches are suggested and tested in this chapter. The first analysis sets a framework for the reduction of the offshore wind turbine model for developing a digital twin model able to capture possible drivetrain anomalies. The second analysis investigates data-driven approaches. A multi-task hard parameter sharing approach is compared to a conventional neural network classifier, to outline the benefits of the transfer learning for small database of run-to-failure data.

The first study developed and compared modal truncated reduced-order models (MT ROMs), for a single offshore wind turbine, with the purpose of capturing a key DOF – the rotor torque – that can be linked to the presence of a critical failure mode – extreme load crack of a gear tooth in the first stage of the gearbox. The aim is to build a model able to reproduce the significant dynamics of the full-order, nonlinear AHSE model, at a reduced computational cost. Both linear and nonlinear MT ROMs are investigated, and their resulting torque signals are compared to the one of the full-order models. Based on the case study selected, the main findings are:

- The linear MT ROMs are recommended for evaluating the rotor torque mean value only, while they cannot represent well the extreme values of this signal.
- The rotor torque is more sensitive to blade modes than tower modes. The MT ROMs including all the blade modes, while excluding all the tower modes, show better results than the MT ROMs with only a selected number of tower and blade modes.
- The spectral analysis demonstrates that the nonlinear effects have an essential contribution to the rotor torque signal for frequencies below 0.5 Hz. If the failure is expected to be captured at a frequency below 0.5 Hz, then the nonlinear MT ROMs should be deployed, as the linear ones fail to accurately capture the torque signal.
- Linear ROMs provide an easier and faster way to study the coupling effects between DOFs, as it can save up to 93% of the simulation time.
- The implementation of the model-order reduction approach (i.e., modal truncation) on the nonlinear AHSE model results in the best representation of the rotor torque. However, this comes with only limited savings on the computational time (about 2.2% for the most promising model).
- Even this slight reduction of the computational time that might seem not significant for a single wind turbine, can save the simulation time for the order of months when applied to the simulations of a typical offshore wind farm – consisting of about 100 wind turbines and covering about 1,000 load cases.

The second study demonstrated the successful deployment of multi-task learning (MTL) approach for a neural network (NN) model with hard parameter sharing. This work has highlighted how the MTL accelerates the accuracy of data-driven condition monitoring for a limited amount of failure data. This is achieved by implementing the two different cost functions, so that the model is better suited at extracting the features correlated to the alarm, and at classifying with a reduced noise and overfitting phenomenon. The power of

the MTL learning lies on the maximisation of the usable information from the data: a (regression) model learns the relationships across the signals, which are then passed as features of the (classification) model for the failure detection. By contrast, a classical NN (classification) model learns only based on one task, for which, however, the amount and quality of data play a fundamental role. Based on the case study selected, the main findings are:

- The MTL hard parameter sharing approach, which underwent a training on ca. 220 days to learn the hidden patterns among the features for the classification task, outperforms a conventional classification model on the detection of failure events on a smaller dataset of ca. 56 days.
- There is a trade-off in the benefit of MTL, and future works should include the investigation of the impact of varying database size and ratios of representative run-to-failure data.

Chapter 4

Numerical models for the structural health monitoring and sensitivity to structural failure mechanisms

Structural failures of offshore wind substructures might be less likely than failures of other equipment's of the offshore wind turbines, but they pose a high risk due to the possibility of catastrophic consequences. The availability of numerical models, able to replicate the structural dynamics of turbine, unlocks new opportunities to investigate scenario for which no, or only a limited amount of, data are available. The combination of the knowledge from the data with the power of a simulation tool builds the digital twin technology. This chapter introduce to data collected, the virtual prototypes, and the digital twin of an offshore wind turbine on jacket substructure – case study of the remainder of this thesis. It additionally presents the methods deployed to support the damage detection frameworks developed in following chapters. Finally, several damage scenarios are implemented, analysed, and discussed for their impact to the global structural response and fatigue accumulation on critical areas of the jacket substructure. The results have been generated to support some of the analysis performed for the ROMEO project [1], by publishing a summary of the main findings in an open-access report [2].

4.1 Background

Although most of structural damage events are not likely to reach a safety-critical level, their late detection can lead to critical consequences which will result to high cost of mitigation

[1] C. S. Wendelborn, **D. Cevasco**, et al., (2020), "D4.4 Implementation of low-cost monitoring algorithms", *ROMEO deliverable*.

[2] A. Kolios, **D. Cevasco**, C. S. Wendelborn, and C. Y. Baonza, (2022), "D4.6 Final report on best practice guidelines for future wind farm structural condition monitoring using low-cost monitoring", *ROMEO deliverable*.

actions [47]. Furthermore, the technical assessment and knowledge of the status of the turbine's structures are necessary to prove that operating assets can maintain the required safety levels during lifetime extensions [156]. To address these issues, the digital twin technology – as introduced in Chapter 1, Section 1.1.1 – can be applied to continuously monitoring the condition of the offshore wind turbine structures.

4.1.1 True digital twin technology

In [17], Augustyn extended the research on the "True Digital Twin" (TDT) to estimate the fatigue damage accumulation for joints and other fatigue-driven structural components of offshore wind jacket substructures. Firstly, he investigated a two-step calibration approach to tune the parameters of the virtual prototype – Finite Element Model (FEM) –, available from design phase, to the measurements from the vibration based SHM sensors installed on the WT structure [194], [195]. Then he suggested a probabilistic framework for updating the structural reliability of wind turbine substructures based on the digital twin, in view of the decision-making process for performing inspections [196].

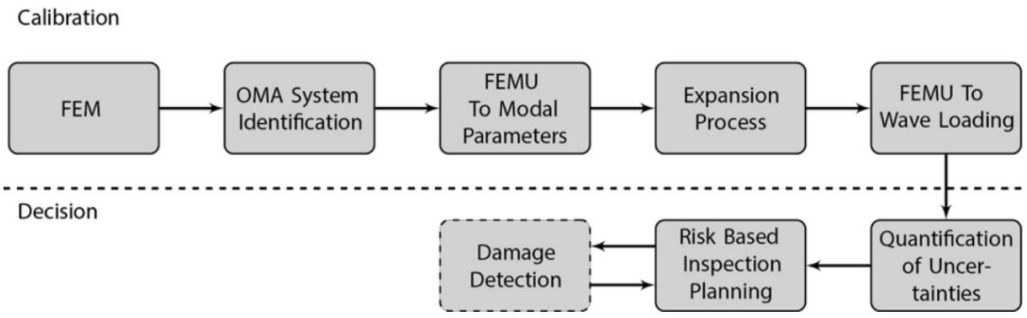


Figure 4.1: Digital twin framework for improved decision models in an offshore oil and gas application, from [15], [16].

The first model calibration is a data-driven process performing a FEM model updating (FEMU), to find a match between the model eigenfrequencies and measured modal properties [194]. This calibration procedure consists of the following steps,

1. An engineering-, and expert-, based updated of the virtual prototype is performed to integrate the offline information on the structural health status of the offshore wind turbine.
2. The acquired SHMS data are processed: (i) by extracting the global modal properties from the vibrational signal via system identification techniques, (ii) by clustering (manually or with data-driven techniques) these modal properties based on the environmental and operating conditions from the SCADA system data.
3. A data-driven model updating of the wind turbine support structure is performed by tuning a selection of the FEM parameters according to the pre-set confidence levels.

The updated numerical model to the modal properties can be then used to deterministically reassess the fatigue life of the assets. This procedure has been shown to have the potential to increase the lifetime of the offshore wind asset installed on jacket substructures [197].

To further reduce the modelling uncertainties for the estimation of the fatigue damage cumulated in the structural elements of the substructure, a FEMU to the wave loading can

be implemented – as demonstrated in the ROMEO's deliverable D4.5 [198]. The scope is to calibrate the wave load model such that the predicted stress distribution corresponds to the measured distribution from the modal decomposition and expansion (MDE) algorithm. For this purpose, the MDE algorithm utilises vibration and strain sensor data, as well as the FEM model updated to the modal parameters [194]. The reader is referred to [195] for more information on the MDE process.

4.1.2 Calibration based on modal parameters

In [194], Augustyn deploys the reference wind turbine of the ROMEO project, from the Wikinger wind farm – following illustrated in Figure 4.3 and described in Section 4.3 – as demonstrator to the first data-driven model updating procedure. This approach which is used to find a match between the model eigenfrequencies and measured modal properties, is briefly summarised in Section 4.4.1. Due to the characteristics of the wind turbine support structure in his analysis, no distinction is made between the fore-aft (FA) and side-side (SS) vibrational motion. Furthermore, by neglecting the measurements associated to the misaligned wind and wave excitation, the 1st torsional mode – potentially in the same frequency range – is not extracted and thus unused in the calibration process. More information on the models and the naming adopted in the Table 4.1 is given in Section 4.4.1.

Table 4.1: Global modal frequencies as extracted from the measurements and as numerically represented in the FEMs, and comparison of the modelled.

	Frequency [Hz]				MAC			
	1st FA	**1st SS**	**2nd FA**	**2nd SS**	**1st FA**	**1st SS**	**2nd FA**	**2nd SS**
Measured/Extracted	0.32	0.32	1.96	1.99	-	-	-	-
As-designed FEM	0.30	0.30	1.37	1.40	0.96	0.98	0.85	0.93
As-built FEM	0.31	0.31	1.57	1.63	0.97	0.99	0.98	0.99
As-installed FEM (digital twin)	0.32	0.32	1.94	2.00	0.96	0.98	0.99	0.99

The discrepancy between the measured (as installed) and the numerically modelled (as built) global frequencies is reduced from 3.1% to 0.3% for the 1st tower bending modes, and from a maximum of 20% to 1.0% for the 2nd tower modes – cf. Table 4.1. Furthermore, the consistency of the model shape to the measured one (i.e., MAC values) generally improved after the update, reaching a value of 0.99 for the 2nd FA tower mode from the initial 0.85 of the design. To visualise the status of the WT after the calibrating procedure, a sketch of the Campbell diagram of the WT numerical models is presented in Figure 4.2. The diagram is drafted by employing the maximum rotor speed of the Adwen AD-5MW turbine [199], and by assuming for the minimum and rated rotor speed – by using NREL 5MW turbine [176] as reference.

It is evident that the turbine has a soft-stiff design, with the 1st natural frequencies of the tower bending modes falling between the 1P and the 3P. These frequencies are only slightly affected by the calibrating procedure – note that these lines overlap in Figure 4.2 –, by reaching a maximum difference of 3.2% between the as-built and the as-installed models. By contrast, as a consequence of the model updating procedure, the 2nd tower bending mode – in the FA direction more than in the SS – is about 23% stiffer than in the as-build model. It additionally gains some reserved on the 6P and 9P resonant response, which can lead to some saving on the structures' fatigue damage accumulation.

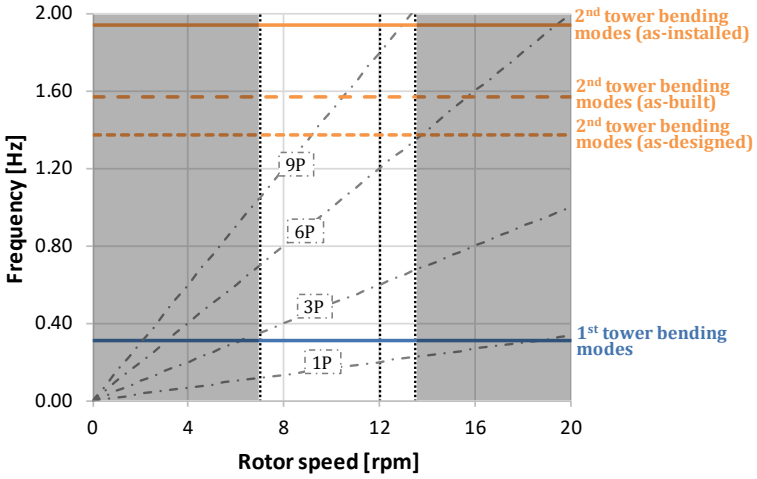


Figure 4.2: Sketch of the Campbell diagram of the WK64 turbine FEMs – cf. Section 4.4.1 – with the 1st tower bending frequency (overlapping for FA and SS, and for all numerical models), and 2nd tower bending frequencies (only in the FA direction).

4.2 Problem statement and aim

The margin between the 2nd tower bending modes – cf. Figure 4.2 – of the as-designed and the as-built models, is due to the implementation of environmental phenomena and structural changes – e.g., corrosion, scour, marine growth – up to the design allowances. It is usually the case that such phenomena, and critical structural failure events, make the structure "softer", decreasing the frequencies of the global modes and potentially varying their mode shapes. This change in the dynamic behaviour rarely cause catastrophic failure, but it can have a considerable impact of the expected life.

The decisions taken in the design phase – i.e., design safety factors –, together with the planned of scheduled maintenance activities, are generally sufficient to cope with these degradation events. As a consequence of the current increase of the investment in bigger (in size and in number) offshore wind projects and of the relatively small global supply chain, the future offshore wind assets are more likely to be characterised by inherent manufacturing defects and loadings miscalculations caused by human-error. Therefore, the scope of this chapter is to investigate what and how the failure scenarios can affect the response of the WK64, and if their impact can be critical for the turbine's operability and fatigue life.

4.3 Case study

An Adwen AD-5MW wind turbine generator installed on a 4-legged jacket substructure structure is taken as a reference for the studies conducted in this chapter, and in Chapter 5 and 6. This structure sits in the 350 MW Wikinger wind farm, which is in the German exclusive economic zone of the Baltic Sea, and it consists of other 70 offshore wind turbines of the same type. The water depth at the reference turbine location (referred to with the acronym of WK64) is approximately 38 m at the mean sea level. The pile penetration depth is approximately 30 m beneath the seabed, which mainly consists of cohesive clay. This wind turbine belongs to the 10% of the structures with a SHM system installed. The setup of the monitoring systems, and the naming convention for the sensors and the structural elements' levels are illustrated in Figure 4.3.

As is generally recommended for offshore turbine structures [200], the accelerometers are installed at 3-4 levels, including the tower-top (section A-A), around the mid-tower (section B-B), the tower base (section C-C), and jacket base (section D-D). These consist of a set of four triaxial accelerometers, at each leg of the transition piece, and three biaxial accelerometers, along the tower sections. This sensor configuration aims to extract the first five global dynamic modes of the structures, consisting of the 1st and 2nd global bending modes of the tower, both in FA and SS directions, and the 1st torsional mode of the support structure. Other installed high-resolution monitoring systems comprise: an inclinometer at the tower base, and a set of temperature sensors and strain gauges on the jacket legs.

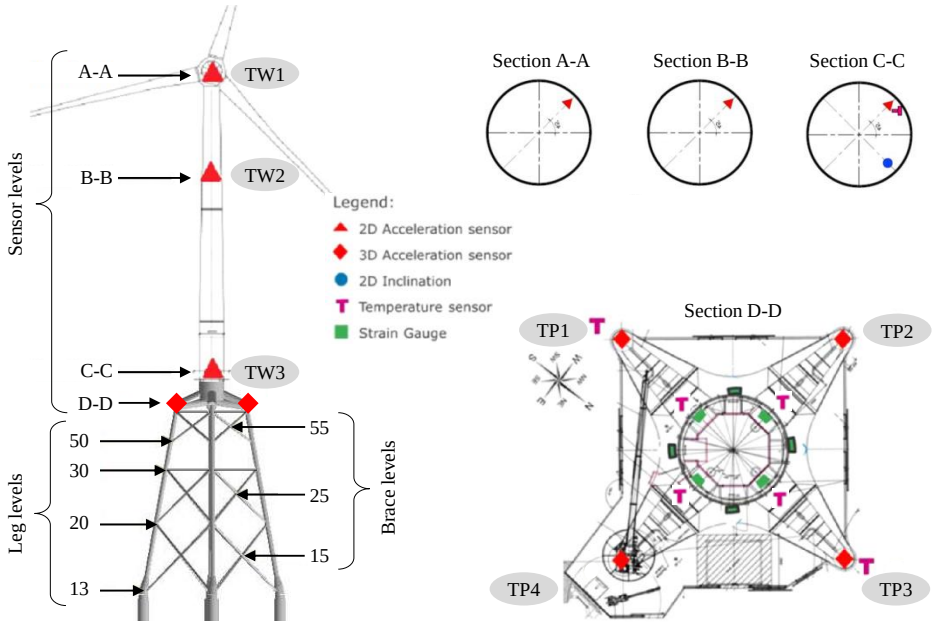


Figure 4.3: Wind turbine geometry and SHM system. The x and y axes are oriented along north and west directions, respectively.

4.4 Methods

This section introduces to the models, the methods and the metrics deployed for the analysis of this chapter, and for the generation of the simulation data for Chapter 5 and 6.

The numerical models of the virtual prototype and the digital twin of the WK64 offshore wind turbine are described in Section 4.4.1. A simplified representation of the WK64 turbine and the beam model of the jacket substructure are modelled by using the Ramboll's in-house software namely ROSAP (Ramboll Offshore Structural Analysis Package) [201]. This section summarises the assumption for the modelling of the as-designed and the as-built representation of the WK64 asset, and for the calibration of the model to represent the as-installed condition.

Sets failure scenario, which are deemed worth being investigated due to their potential criticality to the system dynamics and fatigue life, are integrated in the as-built virtual prototype, as explained in Section 4.4.2. These are differentiated, based on their failure mechanisms, into the extreme (abnormal) site conditions and the structural damage events.

Finally, Section 4.4.3 introduces to the numerical analysis run to generate the results of the sensitivity, of the modal properties and the structural components' fatigue life, on the

failure scenarios. The general-purpose finite element software of ROSAP, version 53, is used to perform static and dynamic analysis of offshore structures. The AHSE code of LACFlex is deployed to run, in a semi-coupled manner, the dynamic analysis of a representation of the Adwen AD-5MW offshore wind turbine. Thus, these numerical simulations capture the operating states and aerodynamic responses of the WK64 turbine.

4.4.1 Modelling of WK64 wind turbine

Two numerical representations of the wind turbine structure are used in this thesis:

- the as-built model, which represent the structural systems according to the design specifications, excluding the maximum allowances of the extreme site conditions, and
- the digital twin (also referred to as-installed) model, which is obtained by modifying a set of parameters of the as-designed model to better represent the as-installed physical twin dynamics.

In the following subsections, a brief recap on the assumptions taken for the representation of the turbine WK64 in the Ramboll's in-house finite element software (ROSAP) is provided, together with a summary of the FE-updating modifications implemented for the calibration of the virtual prototype with the one of its physical twin.

As-designed FEM

The FEM of the as-designed structure is built on the information available from the design reports. The soil-structure interaction is modelled by the use of a Winkler model [202], and the springs, which represent the interface to the surrounding, are linearized according to the API method [203]. The jacket substructure is modelled with 3D Bernoulli-Euler beam elements, with a joint stiffness modified according to Buitrago et al. approach [204]. The complex structure of the transition piece (TP) connecting the foundation and the turbine tower (TOW) bottom, is modelled in a simplistic manner by a representative 3D Bernoulli-Euler beam system. Similarly, the representation of the rotor-nacelle-assembly (RNA) – comprising the nacelle, hub, and blades – is simplified into a two-point mass system, by including a mass moment of inertia tensor.

As-built FEM

The as-built FEM is the results of a first "manual" updating based on the as-built reports. These reports include some information on the as-built weight of the structure and information from pile driving (e.g., pile stick-up). The as-design FEM is thus modified by assuming that at this stage no scour, corrosion, or marine growth phenomena are developed. As reported in Table 4.1, these modifications lead to a reduction of the discrepancies to the real system measurements – about 3% and 20% for the frequency of the 1st and 2nd tower bending modes frequencies and with an improvement to their MAC values to 0.97 and 0.98, respectively.

As-installed FEM

Finally, the as-installed FEM model, or digital twin of the real system, is obtained by applying the data-driven updating procedure presented in [194]. In particular, the FEM parameters are updated as reported in Table 4.2.

Table 4.2: Updated values (in terms of normalised variation) of the FEM parameters after model-updating to the modal properties.

Parameter	Normalised variation
Parameter	0.99
Jacket stiffness	1.83
TOW stiffness	0.36
TP stiffness	1.02
Pile stiffness	0.84
Added mass	1.01
Hub mass	1.02
Joint stiffness	0.13
Lateral soil stiffness	0.07
Axial soil stiffness	0.66
Pile tip stiffness	0.99

4.4.2 Modelling of failure scenarios

This section explains the way of modelling and implementing a set of potentially critical failure modes in the ROSAP FEM of an offshore wind turbine. These failure scenarios are selected based on Failure Mode Effects and Criticality Analysis (FMECA) reported in [47], by following the recommendations of Scheu et al. and integrating their findings with the conclusions from the feasibility analysis of Richmond et al. [53].

Extreme site conditions

As regards the exceedance of the site conditions, the impact of extreme scour, global corrosion, and marine growth phenomena is assessed. The following subsections describe, for each of these phenomena, their modelling assumptions, and the design allowances for the WK64 turbine.

Scour phenomena

The scour phenomenon is the process of the removal of sediment from around the turbine foundations. The seabed being carried away by the hydrodynamic action can either result in the reduction of the soil around the whole area of the foundation (global scour) or of the solely area at the structure piles (local scour).

The finite element model includes details of multiple soil layers and so when a scour depth is introduced, the topsoil layers are removed. The removal of the weight of the top layers is associated to a variation of the soil stiffness. However, by updating the force-displacements curve, it was noticed that this variation has only minor effect on the structure dynamics, affecting mainly the loading in z direction. Thus, it is considered a good assumption for this analysis to refer to the force-displacement curves used for design. The global scour phenomenon, which would impact the most the soil stiffness, is generally considered in case of non-cohesive soil. The cohesive clay soil of the Wikinger wind farm, is primarily affected by a local scour phenomenon. A design limit of 2.2 m scour erosion is given for the turbine in position WK64.

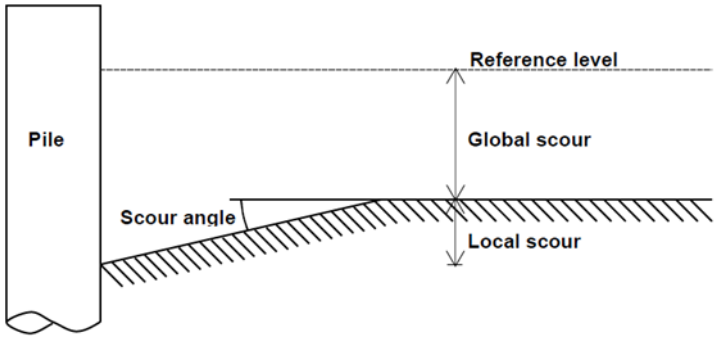


Figure 4.4: Definition of scour.

Global corrosion profile

The design of offshore steel jackets must allow for global corrosion phenomena, both internal and external to the jacket members. It must be modelled in the jacket splash zone and can optionally have a profile at other water levels. It is implemented in ROSAP as the reduction of the cross-sectional area of the jacket beam elements by controlling the reduction of their wall thickness, leading to twice the reduction of the outside and the inside diameter. The program then updates the stiffness and stress calculations. The mass matrix can be updated by removing the mass of the corroded material; however, it was tested that greatest global mode frequency and shape discrepancies are obtained by leaving the mass matrix unaltered.

Marine growth profile

A profile for the modelling of the growth marine vegetation on the submerged elements of the jacket structure can be included in ROSAP similarly to profile for the global corrosion phenomena. The additional mass and the increased surface roughness are modelled via a thickness and a roughness height parameter. For the analysis here presented both parameters are scaled by a multiplicative factor. A factor of 1 represents the design allowance and it is chosen as reference scenario to normalise the results.

Structural integrity loss

With respect to the simulation of the loss of integrity of the structural elements of the jacket substructure, the focus is on the analysis of the influence of local weakening phenomena. While the global corrosion – as described above – is controlled by varying the thickness of the substructure element with the purpose of modifying the stiffness matrix, the localised structural damage is implemented by varying the Young's modulus (E) of elements and sub-elements. The damage types that could reduce the stiffness of the WTG jacket structural member are corrosion, material softening due to cyclic loading, and loosening of the connection between elements.

This way of implementing the damage is realistic for leg and braces elements. On the contrary, it does not apply for the realistic modelling of the grout failure and the bolt loosening. In the given design FEM, the only possible implementation of grout failure is by varying the E of the elements in its proximity of the grout. As concerns the simulation of bolt loosening, the E of the node at the location of the bolted joint is reduced. Richmond et al.

[53] observed that neither of these damage implementations lead to significant variation of the global mode shapes and/or their frequency. However, it could be worth to investigate further the phenomenon of the bolt loosening at the TP and tower connection. This type of failure, indeed, would not appear as a gradual change but rather as a sudden loss of connection as soon as the load exceeds the clamping strength.

4.4.3 Numerical analyses and metrics

The ROSAP software is deployed for the natural frequency analysis of the WK64 with a simplified representation of the Adwen AD-5MW turbine – refer to Section 4.4.1 for more details on the modelling assumptions. The ROSAP FEM is also used for performing the dynamic analysis of the WK64 turbine by accounting for the wave loading on the substructure.

The AHSE code of LACFlex employs a modal-based representation of the turbine (including the tower, rotor, and blades), while the substructure is imported as a Craig-Bampton superelement – extracted by the ROSAP software. This setup allows capturing the aero-servo dynamics of the turbine by performing the numerical simulations in a sequentially coupled manner – cf. Figure 4.6. The key steps of these analysis and the extracted metrics are briefly summarized in the following subsections.

Natural frequency analysis

The study of the dynamics of the turbine in its healthy conditions, and in response to the system's failure mechanism, is obtained via NFA run of the ROSAP FEMs. This implementation does not allow the integration of moving and/or rotating parts, and thus represents only the turbine dynamics in idling condition. The eigenfrequencies and their mode shape vectors are derived for several integrity scenarios of the jacket substructure and for varying environmental conditions.

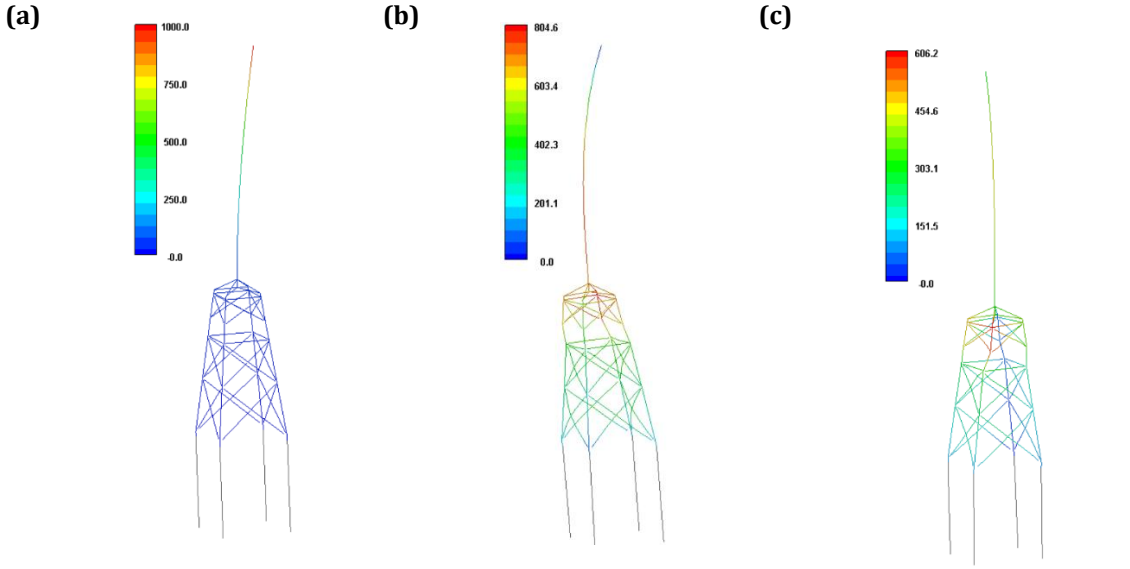


Figure 4.5: Visualisation of WT global modes, (a) 1st tower bending either in fore-aft (FA) or side-side (SS) direction (b) 2nd tower bending either FA or SS direction, and (c) torsional.

Although the 1st torsional mode of the jacket was not used for this model calibration, it will be included in the following analysis to investigate whether its monitoring is beneficial

for detection purposes. The following metrics can be calculated with respect reference (health status) scenario.

Frequency relative difference

The percentage change between the frequency of the system in analysis (Σ) and the reference system (Σ_{ref}) is defined as in Equation 4.1, for each global mode of vibration:

$$freq_{diff} = 100\ \frac{freq_{\Sigma} - freq_{\Sigma_{ref}}}{freq_{\Sigma_{ref}}} \qquad \textbf{Equation 4.1}$$

Modal assurance criterion

The modal assurance criterion (MAC) provides a comparative value between two vectors, giving a measure of their level of consistency. A value closer to 1 means that the vectors are consistent and a value at or close to 0 means that the vectors are inconsistent. The MAC between two real-value modal vectors (e.g. $\{\varphi_r\}$ and $\{\varphi_s\}$), is calculated according to the following equation [205]:

$$MAC(\{\varphi_r\}, \{\varphi_s\}) = \frac{\left|\{\varphi_r\}^T\{\varphi_s\}\right|^2}{\left(\{\varphi_r\}^T\{\varphi_r\}\right)\left(\{\varphi_s\}^T\{\varphi_s\}\right)} \qquad \textbf{Equation 4.2}$$

Modal flexibility variation

The modal flexibility matrix ([F]) is derived according to Equation 4.3, with n being the number of measured modes. In the calculation of the [F], the mode shape matrix [Φ], with $\{\varphi_i\}$ being the mode shape vector of the i-th mode, is weighted by the diagonal matrix of rigidity [Ω], corresponding to $[\omega_i^2]$ where ω_i is the i-th frequency. Each column of [F] represents the displacement pattern of the structure associated with a unit force applied to each degree of freedom (DOF) of the structure.

$$[F]=[\Phi]\,[\Omega]^{-1}[\Phi]^T = \sum_{i=1}^{n} \frac{1}{\omega_i^2}\{\varphi_i\}\{\varphi_i\}^T \qquad \textbf{Equation 4.3}$$

The residual matrix of the modal flexibility ([ΔF]) is measured by calculating the flexibility matrices before and after the damage ([F*]) and by subtracting them, as reported in Equation 4.4.

$$[\Delta F]=[F^*]-[F] \qquad \textbf{Equation 4.4}$$

The absolute maximum of each j-th column of the [ΔF] – as defined in Equation 4.5 – is the modal flexibility variation in each DOF ($\overline{\delta_j}$), with n being the total number of monitored DOFs. It indicates where the maximum variation in flexibility is produced. This quantity has been historically used to estimate changes in the static behaviour of the structure from the dynamically measured modal properties of the system [206].

$$\overline{\delta_j} = \max\left|\delta_{ij}\right| \qquad \text{for } i,j = [1, n] \qquad \textbf{Equation 4.5}$$

Fatigue loads simulations

The generation of time series of the turbine loading conditions, and the calculation the fatigue cumulated into the several structural elements, is achieved by running, in a semi-

coupled manner, the ROSAP model of the foundation with an aero-servo-elastic software of the wind turbine – cf. Figure 4.6. This latter calculates the wind turbine response in time domain, by modelling of the tower and the rotor-nacelle assembly, together with the control strategies in response to the stochastic turbulent wind loads. The aero-servo-elastic code used for the analyses of this thesis is the LACFlex (based on Flex5) software [207]. Not having access to details from the manufacturer, the simulations are run by using a generic representation of a 5 MW turbine.

The Jonswap spectrum for normal sea state conditions is used for the estimation of the wave parameters. Additionally, wind-generated sea current is derived according to the normal current model and considered aligned with the wave. A package of the ROSAP software is used to reduce the model of the substructure into Craig-Bempton superelements [208], by approximating its dynamics to include only a limited number of deformation modes – 30 modes are considered into the analyses of this thesis. It is worth mentioning that the coupling of the foundation via the superelement innerly holds some small discrepancies with the full model of dynamics, for frequencies above 10 Hz [208]. Furthermore, the damping of the tower needs to be calibrated – in the aero-servo-elastic code –, to match the system damping specified and/or targeted in the design phase.

It should be noticed that the implementation of the failure scenarios – of Section 4.4.2, in the FLS simulations of Section 4.5.2) – is created by neglecting their possible effect of the structural damping. This assumption is judged acceptable since the aim of the following investigations (in Chapter 5and 6) is to demonstrate only the feasibility of the detection.

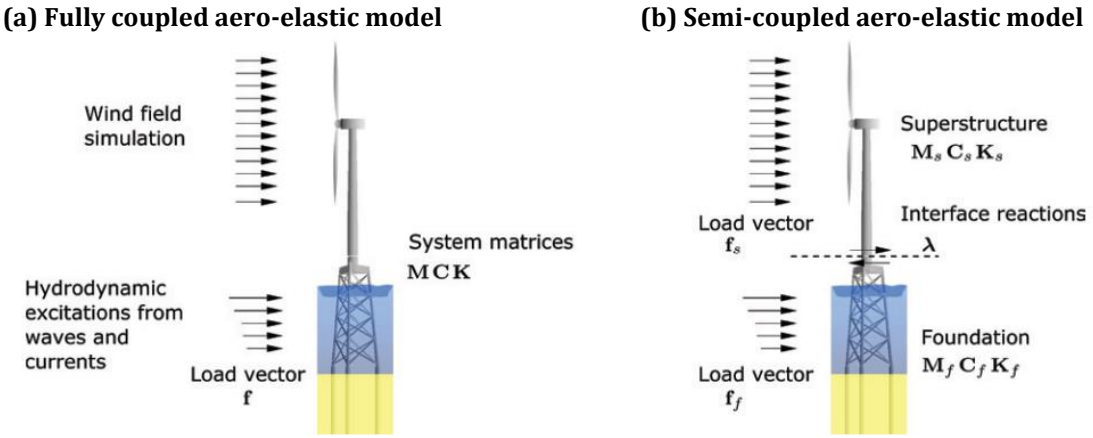


Figure 4.6: Visualisation of (a) fully coupled aero-elastic model, and (b) sub-structured model with interface reaction forces (λ), from [208].

The superelement files derived in ROSAP are coupled into the wind turbine model in LACFlex, and the time simulations are run for each combination of wind and wave loadings. Several short-term stochastic time histories of turbulent wind inflow and irregular wave (i.e., realisations) are realised. The load combination and the number of simulations run per wind turbine status are different for the feasibility and the applicability studies, and thus reported in the respective sections.

The damage equivalent loads (DEL) and the fatigue accumulated in the structures of the turbine can be derived by postprocessing the obtained time series of loadings in LACFlex. A force-controlled recovery run is performed in ROSAP based on the load calculated at the interface – between the substructure and the turbine tower –, to derive the fatigue and consumption rates of the elements of the foundation.

4.5 Results and discussion of structural response sensitivity on failure scenarios

This section assesses the impact of the extreme site conditions and the structural failures on the turbine dynamics in terms of variation of the modal properties, and reduction of the lifetime of the structural components, for the jacket support structure of the WK64 turbine.

4.5.1 Impact to modal properties

The ROSAP parameters representative of the failure mechanism – e.g., the meters of soil to remove for the local scour phenomenon –, are varied around their expected value, to an extreme value falling out of the design allowance. The results are presented in terms of frequency percentage error, modes consistency (i.e., MAC), and modal flexibility variations (i.e., $\overline{\delta}_j$) with respect to a reference scenario, as defined in Section 4.4.3. The reader is referred to the coloured version of the following plots for accessing the full set of information.

Extreme site conditions

In Figure 4.7, the scour parameter is ranged from 0 to a critical value of 3.2 m – i.e., 1 m more than the 2.2 m scour design allowance –, by keeping a constant angle of 18 degree. The results are normalised with respect to the design scour allowance. It is evident that the tower 2nd bending modes are the one mainly affected by the scour depth variation. In the event of extreme scour phenomena, at 3.2 m scour depth, their mode shapes seem unaffected by the change (within a MAC of 0.9 for the torsional mode). On contrary, the frequency of the 2nd modes drops by a further 3% with respect to the design allowance. A total reduction of ca. 7% of the 2nd tower FA frequencies is measured with respect to the no scour scenario. In addition, it is worth pointing out that the tower top sensors are the one recording most of the variation.

In Figure 4.8, the design global corrosion profile allowance is varied with a multiplicative factor from 0.25 to 1.25, including in the calculation the mass of the corroded material. The results are normalised with respect to the 0.5 scenario. The multiplicative factor 1 refers to the design allowance profile. Quickly after this extreme value, the dynamic properties of the structure change drastically, due to the presence of a through-thickness corrosion for some of the jacket elements. Once again, the 2nd tower modes and the jacket torsional mode are the one mainly affected by the phenomena.

Although the 1st torsional mode shape seems to be quite affected by the global corrosion phenomena, its extraction is rather seldom due to the lack of loads exciting this mode. The 2nd FA mode frequency is characterised by a 10% reduction with respect to the no corrosion scenario. The tower top accelerometers have the potential to record most of the variation in the WT structure dynamics as before, but with a modal flexibility variation four times larger than for the scour phenomenon.

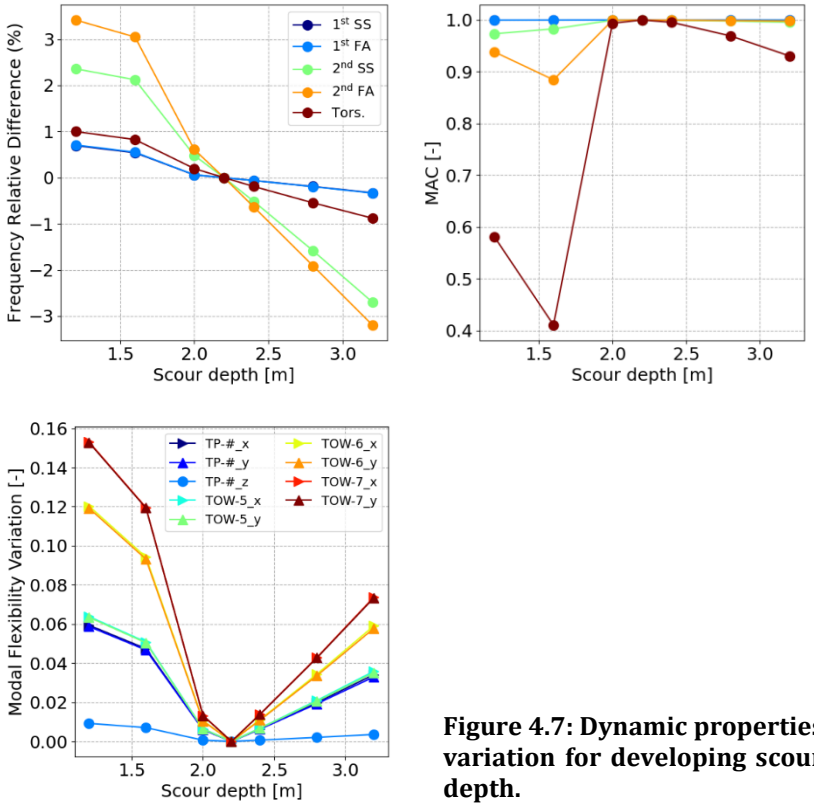


Figure 4.7: Dynamic properties variation for developing scour depth.

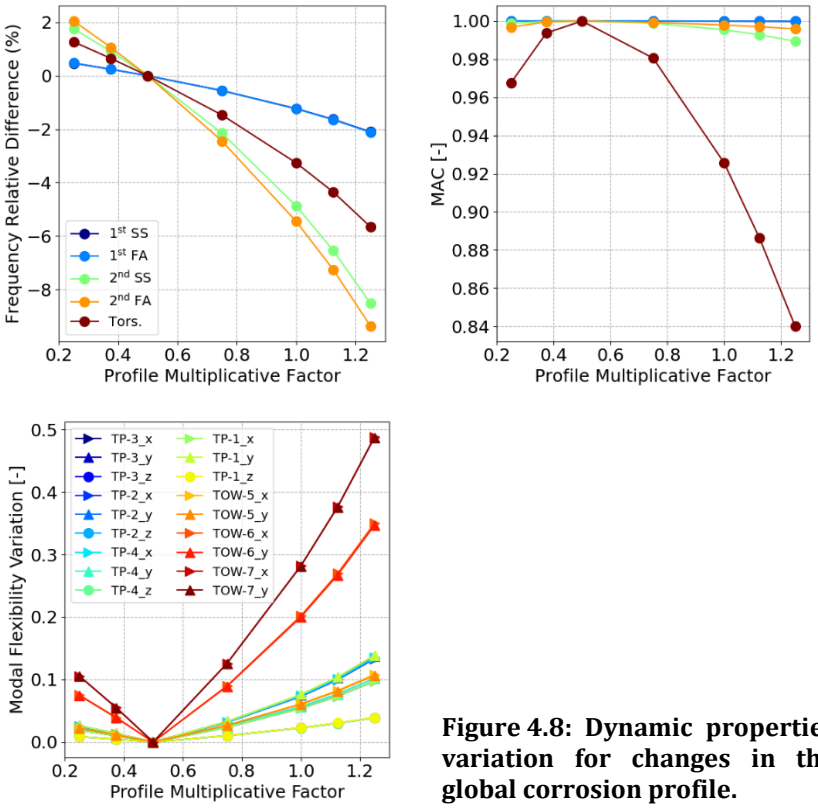


Figure 4.8: Dynamic properties variation for changes in the global corrosion profile.

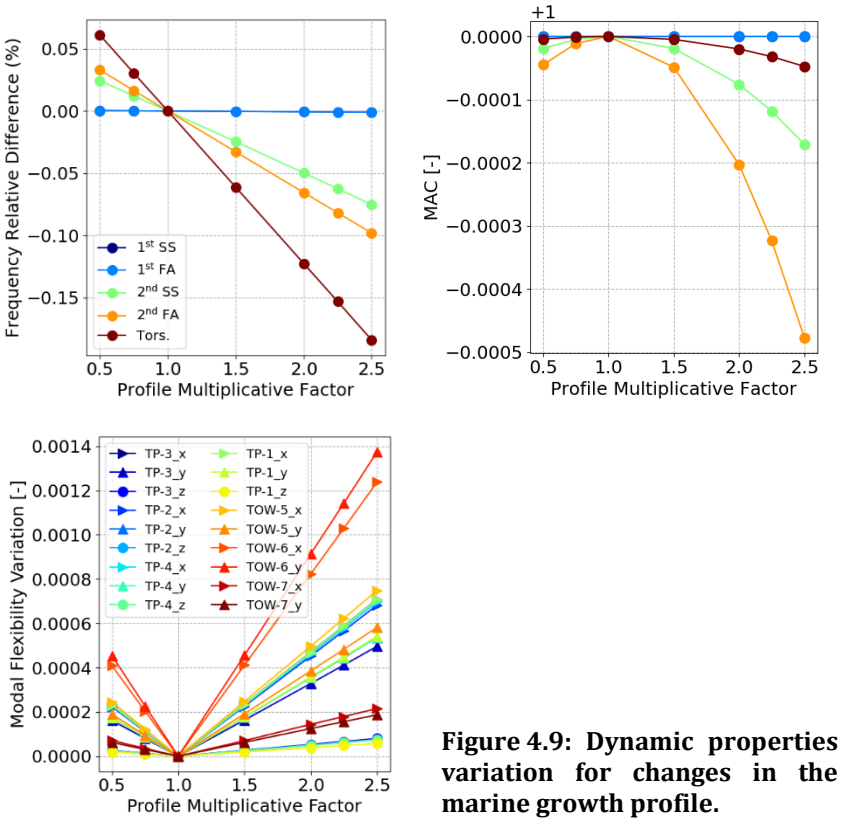


Figure 4.9: Dynamic properties variation for changes in the marine growth profile.

In Figure 4.9, it is evident that the marine growth phenomenon has a very little impact on the global modal properties of the structure. The biggest variation is in the frequency of the torsional mode, which is reduced by only 0.1% by doubling the thickness and roughness of the marine formation.

Structural integrity loss

The simulation of the full loss of the WTG leg would not lead to representative results, since either the full or the partial integrity of the WTG leg is required for the substructure's survivability. On the contrary, it was observed that the substructure might survive the presence of a disconnection damage to braces. Thus, the E of the brace elements is reduced to 99% of its value to simulate a single and localised disconnection. Figure 4.11 shows the results for brace member losses along the structure from level 13 to level 55, representing bottom to top of the jacket – cf. Figure 4.3.

It is observable that, once again, the 2nd tower modes and the torsional mode are the one mainly affected by the presence of the structural damage. For loss of integrity at low water levels (from 13 to 15), the frequency of the 2nd SS mode can drop by more than 4%. In addition, the shape of the 2nd tower bending modes is almost not consistent anymore with the one of the healthy structure – the MAC values reach a value of 0.6. With regards to the modal flexibility variation, it is observed that, although the deviation is relatively small, it is recorded at most by the sensors on TP section.

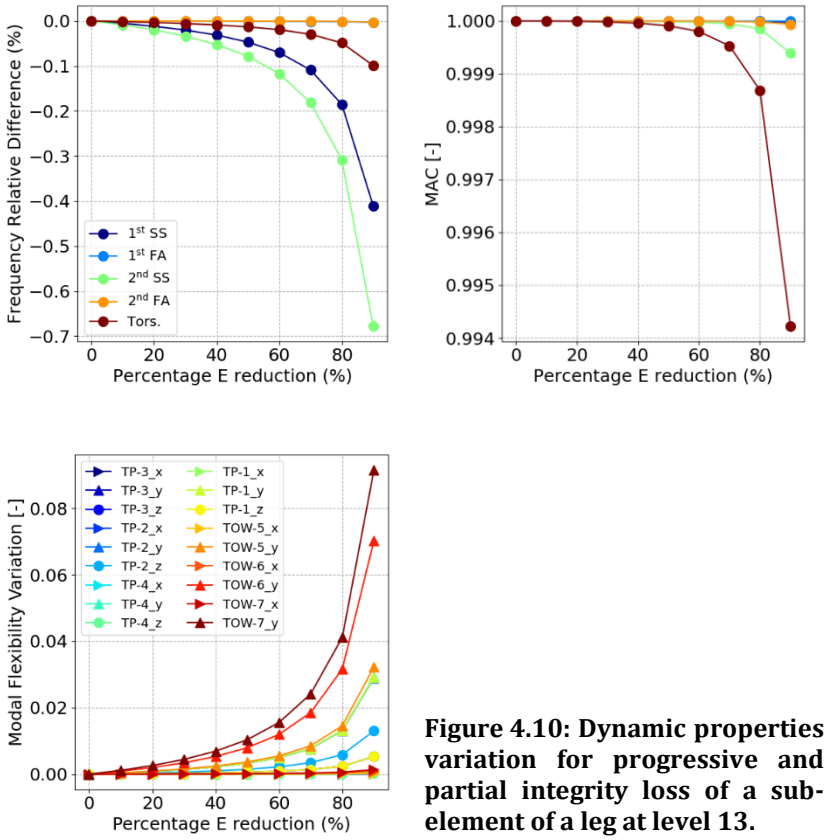
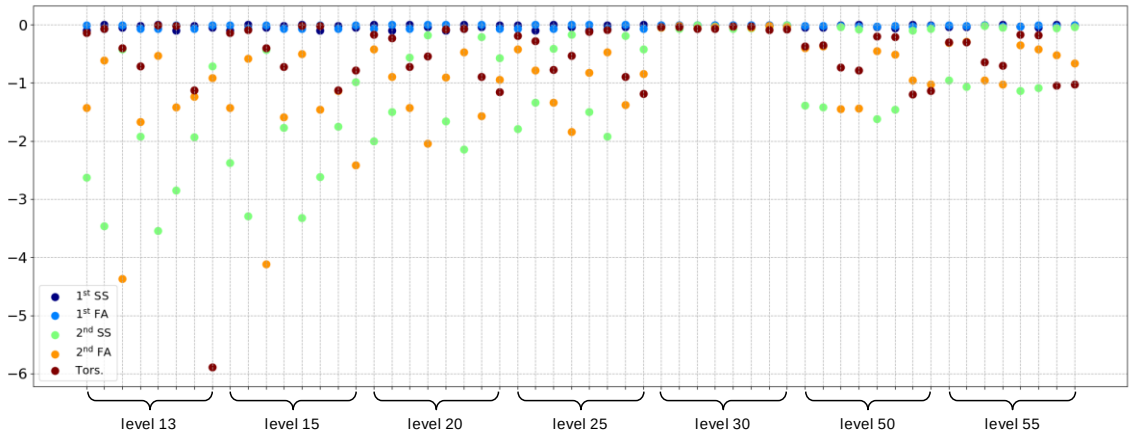


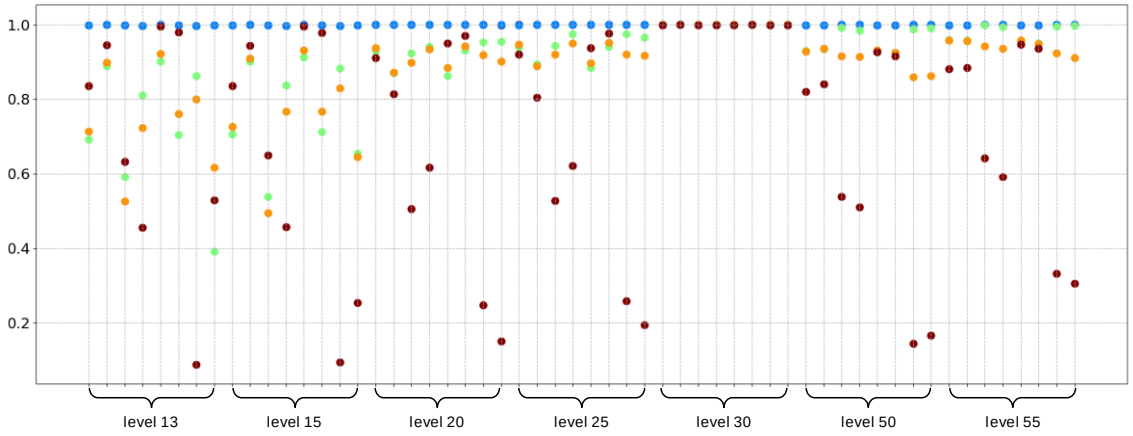
Figure 4.10: Dynamic properties variation for progressive and partial integrity loss of a sub-element of a leg at level 13.

The simulation of a progressive and partial loss of a leg member at level 13 is presented Figure 4.10. The E of the material of one of the leg's sub-elements is reduced up to 90% of its value. It is noticed that the impact of a leg partial damage to the global modes is it generally very little. For instance, the frequency of the 1st and 2nd fore-aft tower modes is unaffected (i.e., see the overlapping lines in the plot) by the integrity loss, due to the directionality of the damaged element with these modes of vibration. However, it is interesting to observe that the frequency of both 1st and 2nd side-side tower modes are the one mainly affected, varying to 0.4% and 0.7% respectively. As regards the modal flexibility variation, the tower sensors (TOW-6 and TOW-7) record most of the deviation.

(a) Frequency relative difference (%)



(b) MAC



(c) Modal flexibility variation

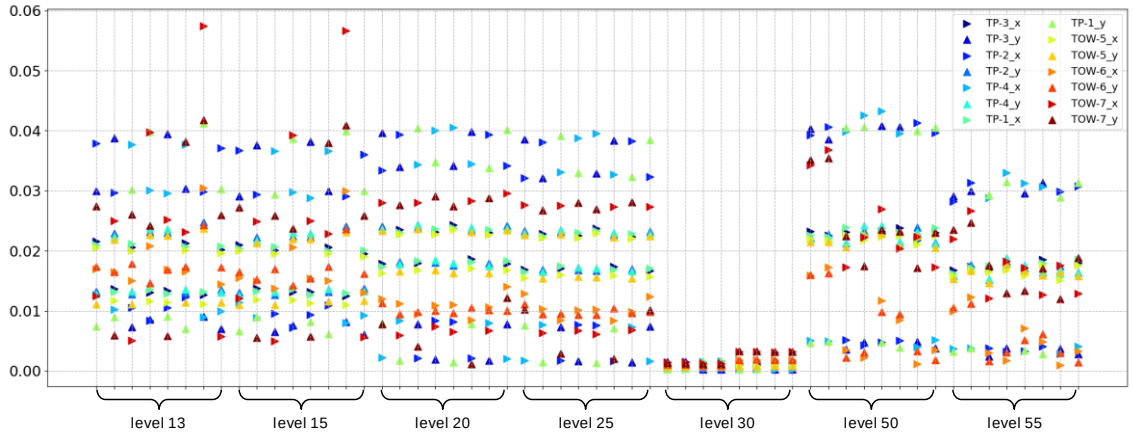


Figure 4.11: Dynamic properties variation for the full integrity loss of single brace members.

4.5.2 Impact to fatigue life

To assess the impact of these failure mechanism on the fatigue life of the elements of the WK64 jacket substructure, FLS loads simulations are run as explained in Section 4.4.3. The normal production load cases – i.e., DLC1.2 of IEC 61400-3:2009 [209]– for the WK64 are setup as in [20], by reducing the load case set to exclude the load combinations with low probability of occurrence. More details on this fatigue damage calculations are reported in the ROMEO deliverable D4.3 [19].

The results of this sensitivity study are assessed by analysing the variation of the fatigue damage cumulated in a year by the legs, the braces, and the joints at the several levels of the jacket structure. The yearly fatigue damage of the (undamaged) as-built model is taken as reference. This yearly damage is normalised based on the foundation design lifetime (i.e., 25 years) and it is reported in percentage in Figure 4.12. The 100% represents the structural failure of the element, and thus a 4% yearly damage is the limit to achieve the 25 years expected lifetime. The results are visualised for the several details of the substructure, at several levels.

In Figure 4.12, and as a convention for the following results, the element with the lowest fatigue life for the whole substructure is crossed in “red”, while in “black” is the element

with the lowest fatigue life with respect to its class of details (i.e., leg, joints, and braces). It is evident that a join at level 30 is the most critical element, being associated with the 4% yearly damage threshold. In comparison, the braces and the legs elements are associated to a higher fatigue life – of the order of hundreds of years –, and they are critical at level 15 and 13, respectively.

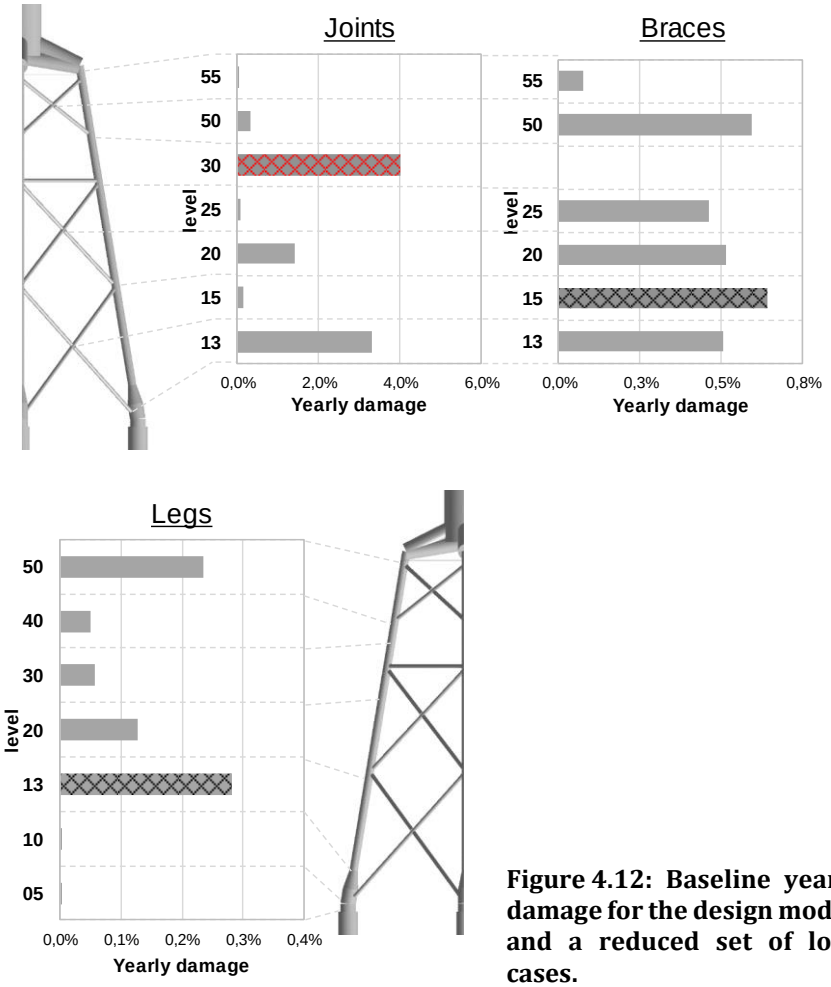


Figure 4.12: Baseline yearly damage for the design model, and a reduced set of load cases.

Extreme site conditions

Among the phenomena introduced in Section 4.4.2, only the case for exceedance of the scour and the global corrosion allowances are further discussed. The marine growth phenomenon was shown to have too small of an impact on the global modes and, thus, intuitively on the fatigue life.

The exceedance of the scour design-allowance does not seem to have a big impact on the fatigue life of the WTG structure. In this case study, the presence of scour leads to a slight decrease of the fatigue damage of the components at the higher levels. The critical yearly damage of the joint at level 30 decreases of 0.9% (corresponding to ca. +7 years fatigue life). However, a small increase of the yearly damage of joints and braces at lower level is observable. It should be noted that this phenomenon cannot be generalised but is a result of the specific combination of assumptions and structural dynamics in this study. In general, scouring might reduce fatigue life and poses a risk for the overall integrity.

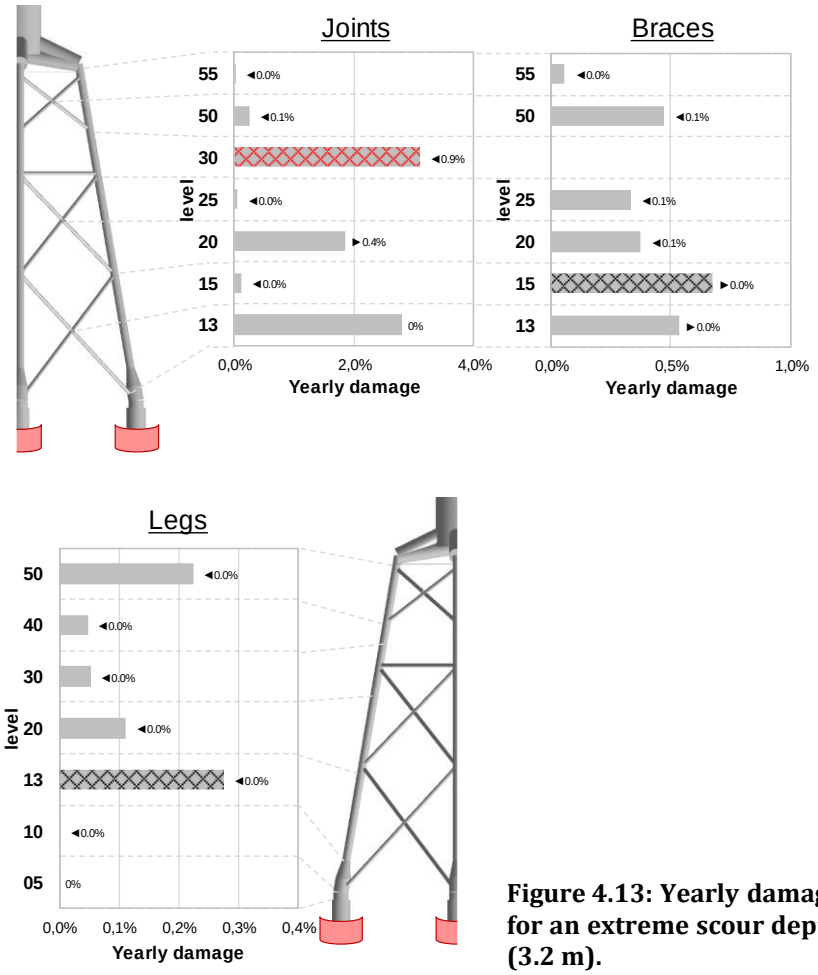
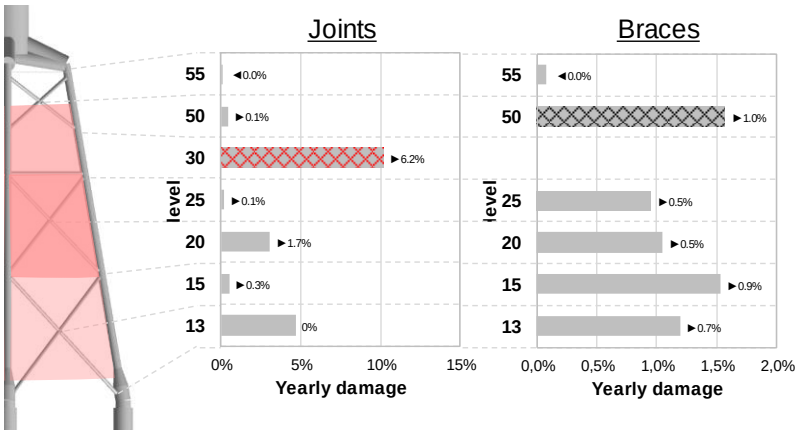


Figure 4.13: Yearly damage for an extreme scour depth (3.2 m).

As regards the global corrosion phenomenon, an overall increase in the yearly damage can be observed in Figure 4.14. The failure of the critical joint of level 30 shortens the critical fatigue life to 10 years. In addition, the yearly fatigue damage increases for the elements (legs and braces) of the higher levels (50 and 55, respectively). Consequently, the exceedance of the global corrosion allowance has the potential to be tracked and/or detected by the fatigue prediction models – further discussed in Chapter 6.



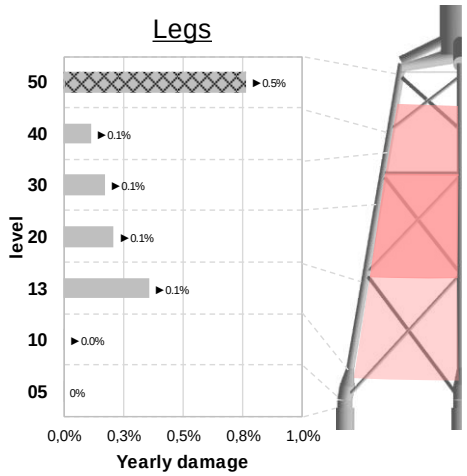


Figure 4.14: Yearly damage for maximum global corrosion allowance.

Structural integrity loss

As a consequence of the full loss of integrity of a brace member, the deviation of the expected lifetime of the structural element of the WK64 jacket substructure can be observed in Figure 4.16, for the joints and the braces, and in Figure 4.15, for the legs.

The yearly cumulated fatigue damages are reported in Figure 4.16. The structural failure of a brace at level 13 initiates the failure of another brace element at the same level, which in turns fails in less than a year of operation. Similar effects are obtained by the full integrity loss of a brace at level 20; the critical fatigue life of the joint at level 30 decrease to 1 year, followed by the catastrophic failure of another brace element at level 20. It can be additionally observed that these localised damages cause an overall reduction of the fatigue lives of the WK64 substructure for the several details.

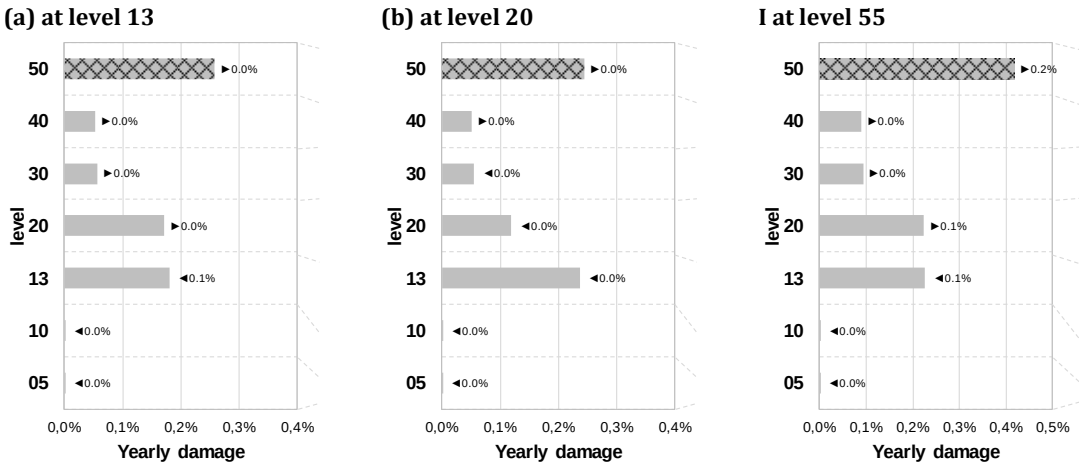
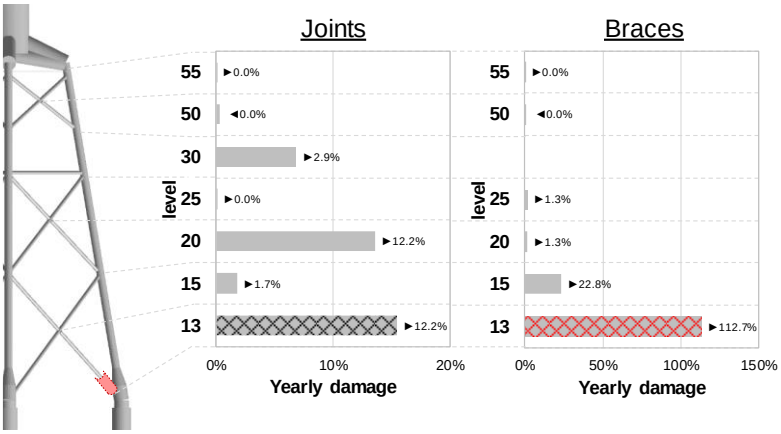


Figure 4.15: Yearly damage of the legs for the full loss (disconnection) of a single brace member at varying levels.

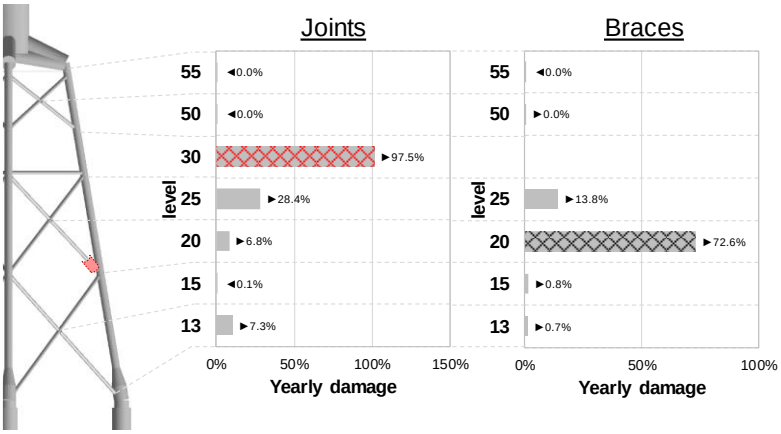
Although these structural damages are likely to lead to a catastrophic sequence of structural collapses – which should be further investigated in survivability analyses –, it seems that they have only a minor impact on the fatigue damage which is measured at the higher levels, making the monitoring of these events complex. On the contrary, the structural failure of a brace at level 55 induces bigger fatigue damage at the higher levels,

potentially allowing to track this damage via indirect monitoring – see Section 1.1.2. However, the full integrity loss of a brace at level 55 might not be as critical as the one of level 13 and 20, as it causes the failure of other structural element only after further 10 years of operation.

(a) at level 13



(b) at level 20



(c) at level 55

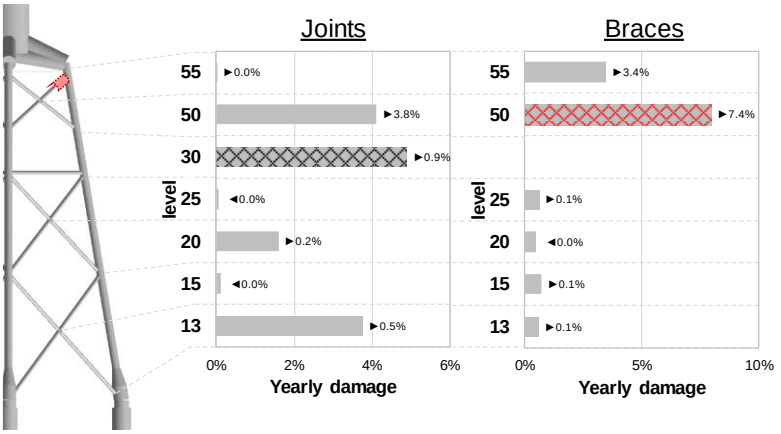


Figure 4.16: Yearly damage of joints and braces for the full loss (disconnection) of a single brace member at varying levels.

4.6 Conclusions

Potential failure mechanism, like the loss of structural integrity of an element of a jacket substructure, or an excessive scour phenomenon, call for the establishment of continuous condition monitoring scheme to ensure the structural health of the asset and the meeting of its expected lifetime. As such extreme events rarely occur, the feasibility and applicability studies had to rely on the data from the simulations, generated by various representation of the real turbine – i.e., level of model calibration.

This chapter describes the models built for the WK64 turbine, and the numerical analysis executed to support the structural health monitoring studies of the remainder of this thesis. Sensitivity studies are conducted to investigate the impact of the failure scenarios on the turbine dynamics. It has been shown that jacket brace member losses, scour and corrosion mainly affect the 2nd tower bending modes, while having little impact on the displacements, rotations, and accelerations at the interface between transition piece and tower. The fatigue cumulated in the foundation can be affected by these failure mechanisms; however, their impact might be local only, and thus difficult to be monitored if not relying on direct sensing. Accordingly, the indirect detection of a structural damages through the fatigue monitoring might not be successful, while the failure mechanisms still pose a risk to the structural integrity.

Chapter 5

Vibration-based diagnostics in an offshore wind jacket substructure

This chapter investigates the feasibility of a level II detection scheme – in terms of damage identification and location of the damage –, for a set of failure mechanism of a jacket support structure. The WK64 turbine is used for this analysis – see Chapter 4 – and the prediction of its structural health status is based on the deployment of true digital twin technology. This as-installed model is used to simulate the local scour phenomenon and the loss of integrity of some of the jacket's brace members. The detection model is trained to identify the correct failure mechanism based on three types of modal indicators, namely, natural frequency, the modal assurance criterion between mode shapes, and the modal flexibility variation. The soft and hard classification results prove that it is feasible to detect and locate the simulated scenarios via the global monitoring of an offshore wind jacket structure. The material of this chapter has been peer reviewed and published in a journal publication [1].

5.1 Background

Global monitoring and parametric-based detection methods have been already extensively applied to some offshore wind full-scale case studies – cf. Section 1.1.2. Prior reviewing the literature is however important to define and clarify on the possible level of the detection. A structural damage can be detected according to several levels, as defined in [210], and this terminology and ranking is adopted for the remainder of the thesis:

- the "detection" (level I) provides the qualitative indication that a damage might be present in the system and potentially identify the damage type,
- the "localisation" (level II) is to identify the probable location of the damage,

[1] **D. Cevasco**, J. Tautz-Weinert, M. Richmond, A. Sobey, A.J. Kolios, (2022), "A damage detection and location scheme for offshore wind turbine jacket structures based on global modal properties", *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering*, vol. 259, no. 2. https://doi.org/10.1115/1.4053659

- the "assessment" (level III) quantifies the severity (i.e., the size) of the damage, and
- the "consequence" (level IV) informs on its progression and degradation process, to advise on the safety of the system.

Weijtjens et al. proposed a data-driven SHM framework in [26], [49]. Their approach consists of tracking the system modal properties according to the different operational conditions (i.e., parked, and operating case). These are then sub-grouped according to shaft rotations per minute and pitch ranges, and the variability of each mode of vibration is normalised by using a non-linear regression model. In [30], they observed that some of the modes correlate to environmental effects such as the sea water temperature – especially for low order modes –, the tidal level, and the water height, besides being dependent on the yaw angle for asymmetry of the structure. In [26] they additionally recognised changes in the vibration levels of the tower top accelerometer depending on the wind farm wake effect, for two of the turbines in the C-Power wind farm. Finally, by detrending the OMA results based on the environmental fluctuations, they managed to recognise the stiffening of the response of a turbine installed on a monopile foundation [49]. A similar approach was followed by Oliveira [211], and validated for an onshore wind turbine case study [32], to identify the presence of numerically simulated damage in the foundations (crack and scour), and blades (uniform deterioration due to operation). These monitoring methods, based on the use of the control chart theory [212], allow for the damage detection only (level I). If interested in identifying the location of the damage and assessing it (levels II and III, respectively), more advanced machine and/or deep learning approaches might be preferred [25].

However, data-driven methods are conditioned by the significance of historical data. These methods can be applied only if the set of data contains information on the variation of the physical properties as a consequence of the damage. For this reason, several authors have made use of models to simulate representative datasets for the healthy and damaged system dynamics. In [51], Nguyen et al. numerically investigated the feasibility of vibration-based damage assessment for an offshore wind turbine with gravity-based foundation, excited by various waves loads. In [53], Richmond et al. conducted a sensitivity study on the changes in the dynamic response of an offshore wind jacket structure, for several classes of anomalies and by ranging their severity. In [52], Nguyen et al. then proposed the use of a vibration-based artificial neural network for the estimation of location and severity of simulated structural damage in onshore wind towers. The main findings outlined that a detection algorithm trained on frequencies only performed better for the assessment of the severity.

Extensive research was also conducted for the detection of structural failures in offshore (oil and gas) jacket platforms. Liu et al. [213] suggested a modal flexibility-based method using a FEM updating technique. Modal flexibility detection approaches belong to a family of traditional vibration-based methods, together with frequency-based and mode shape (and its derivatives) approaches [214]. In [213], Liu et al. applied a gradient-based method for the minimisation of the Frobenius norm of the matrix representing the residual of the flexibility between the damage condition and a healthy reference (cf. [25] and Section 4.4.3). Another model updating approach applied to an offshore truss structure was proposed by Malekzehtab and Golafshani in [215]. For their updating procedure they applied a genetic algorithm to optimise a cost function defined as the sum of the distances between the frequencies (relative difference) and of the mode shapes (level of consistency, i.e., MAC values) between the reference and the damaged conditions. In both [213] and

[215], the algorithms successfully identified the damages' location and severity, additionally accommodating measurements uncertainties and several noise levels. As regards the application of machine learning methods, Xu et al. [216] used a residual strain energy method and employed principal component analysis to remove the influence of the environmental temperature variation. Concerning the non-parametric vibration-based methods, it is worth mentioning the work of Diao et al. [217] and Cheng and Wang [218] as applied to offshore platform structures.

5.2 Problem statement and aim

Despite the extensive research effort, to the knowledge of the author, no thorough investigation on the vibration-based monitoring of offshore wind turbine jacket substructures has been performed yet to distinguish between several failure mechanism. Therefore, this works is aimed to study the feasibility of a health monitoring scheme for a wind jacket substructure, to detect and locate the damage. A monitoring scheme is deemed suitable for industrial needs if the following criteria are achieved: (i) detecting and distinguishing anomalies of different causes, (ii) use of low-cost measurement technologies, (iii) transparency of reasoning process – versus black-box models –, (iv) use of a probabilistic approach for decision making, (v) enabling real-time monitoring.

The workflow of this monitoring strategy is presented in Section 5.3.1. The core concept of the suggested approach is to employ (modal) simulated data from the as-installed FEM of the WK64 turbine – described in Section 4.4.1 and briefly summarised in Section 5.4 – to train the machine learning algorithm to detect and locate the damage cause by environmental and structural failures. Starting from the findings of [53], six types of localised structural damage and several levels of scour are simulated – as presented in Section 5.5. The feasibility of the detection is discussed in Section 5.6, together with a consideration of the set of features to derive and track. Finally, in Section 5.7, the challenges and limitations of the current approach, prior a field application, are outlined.

5.3 Methodology

The flow of data and processes for the field application of this monitoring approach is given in Section 5.3.1. In Section 5.3.2, details are provided on the simulation setup, to extract the modal properties of the structure and calculate their deviations. The machine learning algorithms and the processes for their training, testing and validation are presented in Section 5.3.3.

5.3.1 Suggested approach

An overview of the workflow for the suggested health monitoring of the WTG structure and damage detection is given in Figure 5.1. The pillar of this strategy is the "true digital twin" technology – refer to Section 4.1.1. Passing through the screening and diagnostics of the structure (level 1), the FEM updating procedures (level 2 and 3), and the quantification of the uncertainties (level 4), the digital twin can be employed to continuously monitor the accumulated fatigue damage in the hot spots of the structure (level 5).

For this feasibility study, a level 2 digital twin is deemed sufficient. Augustyn's data-driven FEM updating – summarised in Section 4.1.2 –, is used for this scope (highlighted in

blue in Figure 5.1). The calibrated FEM matches the structure's measured global modal properties (in *green*) as close as possible. The modal properties of the system in its normal behaviour – depending on the environmental and operational conditions (EOC) – and in its damaged status, are then retrieved from natural frequency analyses – cf. Section 4.4.3. By comparing the derived modal properties with those of a reference healthy scenario, modal indicators such as the modal assurance criteria (MAC) and the modal flexibility variation ($\overline{\delta}_j$) are calculated and employed to track the system's dynamics evolution and deviation – cf. Section 4.4.3. This set of information, stored into a database, is used for instructing detection models to identify and locate (detection levels I and II) the system's anomalies (in *pink*). The instructed algorithms can then be used on the modal data calculated from the field vibrational data, to raise alarms if the pattern of one of the simulated damage scenarios is recognised (in *orange*).

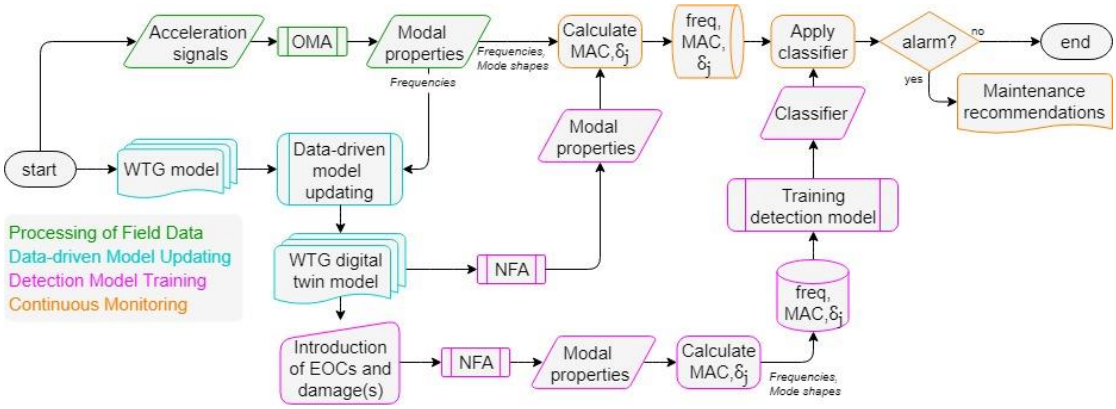


Figure 5.1: Processes and data flow for the suggested monitoring strategy.

In contrast to the approaches of [213] and [215] – where a further model updating is used to detect location and severity (level III) of the damage –, a classification-based detection is here suggested. The reason for this decision is related to the fact that, for offshore wind applications, several types of damage are potentially critical for the structure. The approach proposed can easily be extended to set up the detection of anomalies of every type, while the studies in [213], [215] find their application for the detection of damage controlled by a single model parameter – e.g. elements' and joints' stiffness. Additionally, the uncertainty of the simulated data is introduced by a slight variation of the environmental parameters (e.g., scour and tidal phenomena), rather that adding several levels of white noise to the signals.

5.3.2 Data generation

The study of the dynamics of the WK64 support structure, in its healthy conditions and in response to anomalies in the system, is obtained via batch run of the NFA in the Ramboll's in-house ROSAP software – cf. Section 4.4.3. This implementation does not allow the integration of moving and/or rotating parts, and thus represents only the turbine dynamics in idling condition. The rotor-nacelle-assembly is, however, integrated for its contribution in terms of mass and inertia. It should be noted that the methodology suggested can be applied independently of the type of software used for the analysis. It is worth mentioning that, if the modes extraction is carried out via an AHSE tool, it would be possible to extend

the analyses to account for the interference of the rotor dynamics and the effect of the different turbine operating regimes.

The eigenfrequencies and their corresponding mode shape vectors are derived for several integrity scenarios of the jacket substructure and for varying environmental conditions. After selecting a healthy reference scenario, the indicators in Section 4.4.3 can be calculated.

5.3.3 Machine learning processes and algorithms

This section introduces to the selected machine learning algorithms and features, by explaining the criteria and the decision-making process behind their selection. Finally, the processes for the training and testing of the detection models and their validation is detailed.

Features

For the investigation of this paper, three sets of features are analysed: a set of frequencies only, a set consisting of frequencies and MAC values, and a set including frequencies and $\overline{\delta_j}$. Frequencies, rather than relative difference of frequencies, are selected and normalised, together with other possible features, during the pre-processing phase. By adding the MAC values, the algorithm is also informed on the deviation of the shapes of vibration. However, being used as a mode-goodness indicator during the post-processing of the OMA results, it can happen that modes with a low MAC value get filtered out of the analysis, although potentially signalling the presence of damage. An alternative measure of the modes' deviation can be given by the $\overline{\delta_j}$, which additionally provides a higher level of information because of being not only sensitive to the mode, but also to the sensor location.

Algorithms

Allowing data samples to belong to two or more class types, with different levels of membership, is the main criterion for the selection of the detection algorithms to be tested in this feasibility analysis. This requirement reduces the choice to soft-classification models, to explicitly estimate the class conditional probabilities, and discard the more complex-to-interpret deep learning and tree-based models. Either fuzzy- or Bayesian-based models inform on the degree of membership of each data sample to the given classes. Targeting a multi-class classification, a predictive model based on the linear discriminant analysis (LDA) theory is deemed more suitable than setting-up multi-class logistic regression models. However, the particular set of data in analysis violates the LDA assumptions of normally distributed data and identical covariance matrices for every class [219]. For this reason, fuzzy-based models only are investigated further.

Fuzzy logic is an organised and mathematical approach, able to handle inherently imprecise concepts by using membership functions. In their simplest application, then, these functions are manually defined setting the truth values and a set of fuzzy rules is given to describe how one or more fuzzy variables relate to another. Although such a transparent approach would be preferred, the mapping by hand of variables and rules is not straightforward for this detection application. Therefore, the fuzzy logic principles are automatically applied to cluster the multidimensional data, according to the so-called fuzzy c-means (FCM) method [220]. The algorithm works by assigning membership to each data point corresponding to each cluster centre on the basis of distance between the cluster and

the data point. This unsupervised method is controlled by specifying the number of clusters to identify, the fuzzy exponent and a termination tolerance [221]. Specifically, the Python open-source version of [222] is used for the purpose of this work.

Training, testing and validation processes

A sketch of the flow of data and processes for the training, testing and validation of the FCM model is given in Figure 5.2. The cross-validation procedure [223] is employed to verify the independency of the prediction on the WTG operating condition (yaw angle rotation). Although the supervised LDA was not considered suitable for the application, Fisher's LDA reduction technique is applied, transforming the set of features while maximising the separability of the classes. As investigated by Li et al. [219], LDA for dimensionality reduction can also work reasonably well if those assumptions are violated. Based on the optimal rotation and the reduced features found in the training set, the datasets for testing and validation are consequently transformed. Multiple FCM models are then trained during the cross-validation process, assigning each data sample a membership to each cluster centre. Only the best performing models – one for each set of feature combinations – are tested on the validation dataset. To discuss the detection capability of the algorithms, hard-threshold metrics are used – as explained in Section 5.6 –, by allocating to each data sample the label according to the highest membership predicted.

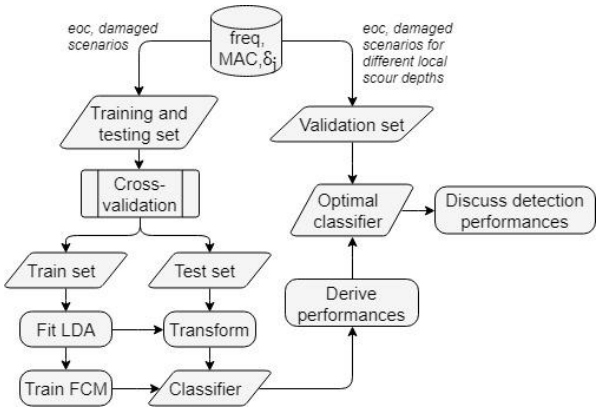


Figure 5.2: Processes and data flow for training, testing, and validating the detection algorithms.

5.4 Case study

The WK64 turbine and its numerical representation in the as-installed condition – cf. Section 4.4.1 –, is deployed to generate the databased of healthy and damaged scenarios – refer to Section 5.5. The selection of the damage scenarios to be implemented and simulated is based on the findings of Section 4.5, and on the following considerations.

As the design of this jacket sub-structure is fatigue-driven, the extreme scour events is generally of concern for the fatigue damage [224]. This does not seem the case of the WK64 turbine, by looking at the FLS results of Section 4.5.2. However, an undetected scour phenomenon still holds a risk for the ultimate and the serviceability limit states. On the contrary, the full integrity loss of the brace members has been seen to bring the utilisation of the jacket structure outside the design assumptions in the FLS setup.

Therefore, in this case study, on the detectability and discernability of different mechanism of failure – all potentially leading to the presence of subsea cracks in the elements of the jacket substructure – the local scour phenomenon, and the loss of integrity of the jacket's brace members are simulated.

5.5 Database and preliminary analysis

The database for this feasibility study is built by collecting the metrics of the variation of the first five global modes, with respect to the healthy reference scenario, for varying environmental conditions and structural health status. The detectability of these failure mechanisms is discussed for the as-installed structure, as it has been already done in the sensitivity analysis of Section 4.5, for the as-built representation of the WK64 turbine.

5.5.1 Scenario simulation

Based on the design specification, the site accounts for only a few centimetres of variation in the water level for tidal phenomena, the measurements' uncertainty is introduced in the simulated dataset solely by varying the local scour depth within the design limit.

Regarding the operating regimes, the idling-only conditions are mimicked, by accounting for the impact of the rotor-nacelle-assembly yawing on this asymmetric system inertia, and thus on its modal properties. Furthermore, although the 1st torsional mode of the jacket was not used for this model calibration, it is included in the following analysis to investigate whether its monitoring is beneficial for detection purposes.

Local scour phenomenon

As for the sensitivity in Section 4.5.1, the local scour depth is ranged from 0 to a critical value of 3.2 m, uniformly for all the legs, and by keeping the scour angle constant at 18 degrees – as assumed from the design. The design limit value is reported in Figure 5.3 with a cross symbol; this rightly identifies the maximum fluctuation of the estimated metrics in the normal state of the structure for a fixed nacelle angle. The reference for estimating the metrics of Figure 5.3 is taken with respect to the no-scour scenario.

It is evident that the tower's 2nd bending modes are the ones mainly affected by the scour depth increase. However, the variation mainly affects their frequencies – with a drop of about 4% for the 2nd FA mode at 3.2 m scour –, while their mode shapes are almost unchanged – with a MAC value higher than 0.99. In addition, it is worth pointing out that, in the case of the presence of local scour, a higher $\overline{\delta_j}$ is recorded by the tower's top sensors.

Loss of structural integrity

In Figure 5.4, the E of each of the brace elements of the jacket structure connecting to a leg element is reduced to 1% of the design value. The results are reported with respect to the several levels of Figure 4.3, and in terms of the relative difference in frequency, the MAC values, and the $\overline{\delta_j}$. For each brace level, eight values are reported, corresponding to the eight brace-to-leg connections, two per leg, of this 4-legged jacket structure. Because of this, and due to the fact that the results in Figure 5.4 are relative only to a single rotor-nacelle-assembly position, it is possible to observe some analogies in the results at each level, with slight differences that are caused by the damage locations and system's asymmetry between the legs. By reporting the thresholds identified in Figure 5.3 to Figure 5.4 (light grey shaded

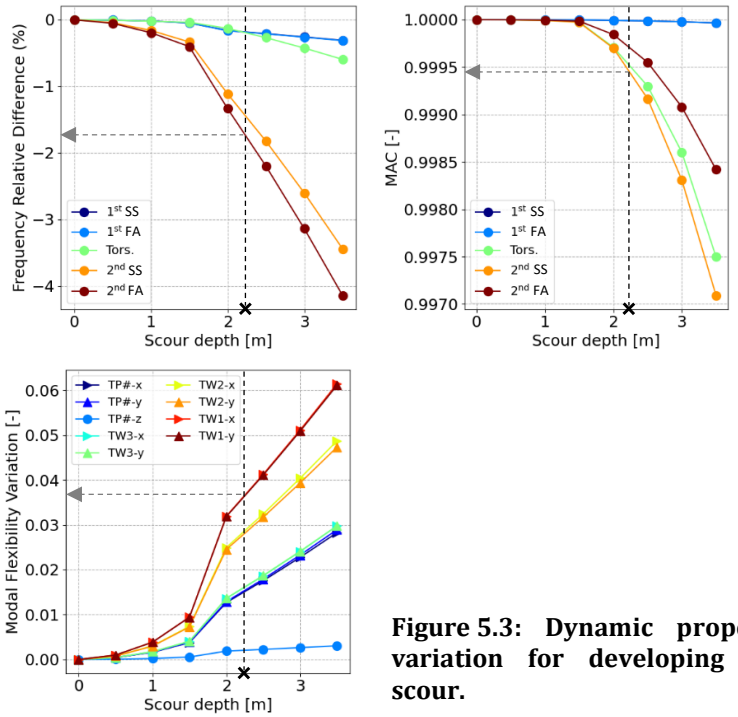


Figure 5.3: Dynamic properties variation for developing local scour.

areas), it is evident that the ranges of variability due to the EOC overlap with the deviations caused by the structural failures, emphasizing how this poses a challenge on damage detection and location.

In general, it can be observed in Figure 5.4 (a), that the frequencies of the 1st torsional mode and the 2nd tower modes are mainly affected by the presence of the disconnection. These modes grow in difference with respect to the reference healthy scenario – almost up to 2.5% for the 1st torsional mode – if the damage is closer to the splash-zone, and thus to the sensorized area of the WTG jacket – cf. Figure 5.4 (b). The changes in the mode shapes, shown in Figure 5.4 (b), concern mainly the 2nd SS mode for the lower brace level. The 1st torsional mode consistency declines in the case of damage to the higher brace levels, scoring a MAC value as low as 0.88 at level 55. Concerning the $\overline{\delta_j}$ of Figure 5.4 (c), it can be noticed that the sensors located in the transition piece are the ones recording the highest variation, with alternating direction and sensors positioning, depending on the leg where the damage is implemented. Only a few damage locations, and for heights above level 20, impact the flexibility recorded by the sensor located at the tower top. It is finally interesting to observe that the presence of structural damage in the horizontal elements of the jacket (level 30) cannot be detected by any of the modal global modes. For this reason, the detection of these damage locations will be excluded from the following analysis. Instead, the detection of disconnections at one location per level (13-15-20-25-50-55) is investigated, implementing for simplicity all damages on the same leg and on the same leg-side.

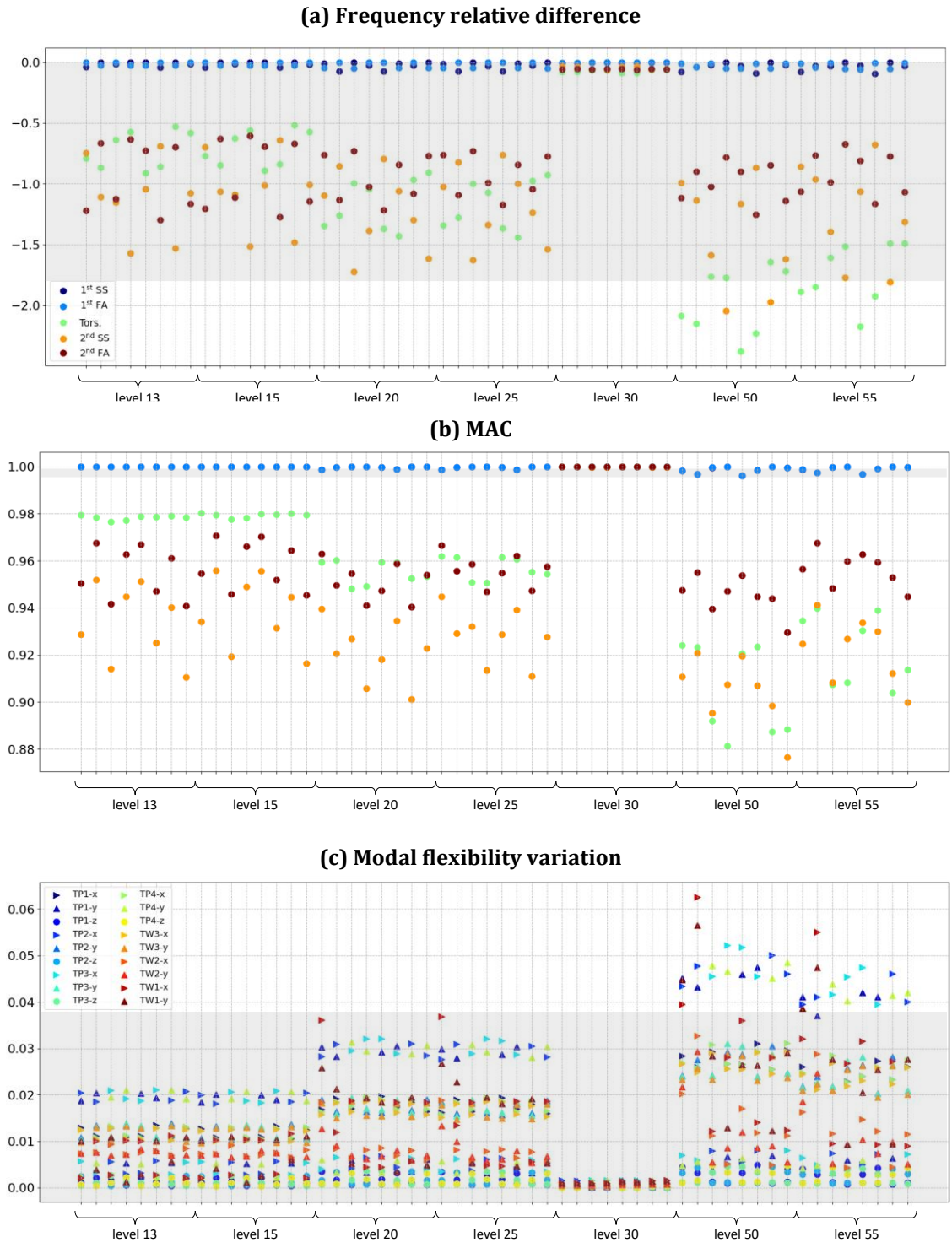


Figure 5.4: Dynamic properties variation for the full integrity loss of single brace members.

5.5.2 Training, testing, and validation datasets

The data samples for training, testing, and validating the detection algorithms are simulated by introducing the anomalies described in Section 5.5.1. The detection algorithms then try

to cluster the data into as many clusters as the number of simulated WTG status, corresponding to the following labels:

- *eoc*, reproducing the structure behaviour for local scour depth up to the design threshold,
- *scour*, modelling the local scour phenomenon over the design allowance,
- *D55*, *D50*, *D25*, *D20*, *D15*, and *D13,* mimicking the integrity loss of brace members at the leg K or Y joints, for the levels from 55 to 13, as explained in Section 4.4.1.

The training and testing phases are used to verify the effectiveness of the clustering of the WTG status for variation in the nacelle position. The training set contains data samples at every yaw angle from 0 to 359 degrees, with a 1-degree step, for each of the simulated labels. A ten-fold cross-validation – as presented in Section 5.3.3 – is implemented to split in training and testing. The testing results are used to select the best set-up of the hyperparameters, in terms of LDA components, and to verify the goodness of the clustering via the fuzzy partition coefficient [222].

The validation of the algorithms on data samples for variation of the scour level is used to validate the algorithm's performance for unseen data and recommend the best set of modal indicators to be used for the detection task. The validation set contains all simulated labels at every yaw angle from 0 to 359 degrees, as for training and testing. However, for each of the simulated labels, the environmental conditions - implemented via the scour depth parameter - are varied. Two different validation sets are investigated, one for a slight variation of the scour depths and one for scour depths approaching the design allowance, respectively.

5.6 Results and discussion

This section presents and discusses the results from the LDA transformation and the fuzzy clustering model which are used to investigate the detectability of anomalies in the system. Although the fuzzy clustering belongs to the so-called soft-clustering methods, the results are mainly reported in terms of threshold metrics [225]. Therefore, a label is assigned to each data sample based on the highest membership predicted by the fuzzy clustering model.

5.6.1 Training and testing on datasets of operational variations

The focus of this preliminary analysis is on the identification of the optimal number of the LDA transformed component, for the reduction and rotation of the modal indicators into features separating at best the eight classes. The cross-validated estimates are reported in terms of macro averages of the accuracy and of the F1-score – defined and explained in Section C.2 of the Appendices. For multiclass classification, the macro average (arithmetic) of these metrics can be calculated by aggregating the contributions of all classes [225], with c being the number of classes, as indicated in Equation 5.1.

$$\text{metric}_{\text{macro}} = \frac{1}{c}\sum_{i=1}^{c} \text{metric}_i \qquad \textbf{Equation 5.1}$$

The fuzzy clustering models are trained to identify all eight centres on uniform, but randomly selected, subsets of the training set – according to the cross-validation process explained in Section 5.3.3. In Figure 5.5, the box plots of the fuzzy clustering results, for the three feature combinations, are presented for varying numbers of the LDA components. The

random selection of the subsets for this training and testing phase is the reason for the predictions' variance.

It can be observed, in Figure 5.5 (a), that the detection based on the tracking of frequency only has already quite satisfactory performances. The macro accuracy and F1-score reach median values above 93% and 74%, respectively, by selecting the first two LDA components. Slight improvements – with a median macro accuracy of about 94% and a median macro F1-score of 76% – are achieved by including the features relative to the MAC values of the modes. As shown in Figure 5.5 (b), this is achieved by additionally extending the number of LDA components (from three to five). It is finally evident, in Figure 5.5 (c), that the fuzzy clustering models correctly classify the majority of the WTG status, by including the information on the modal flexibility variations.

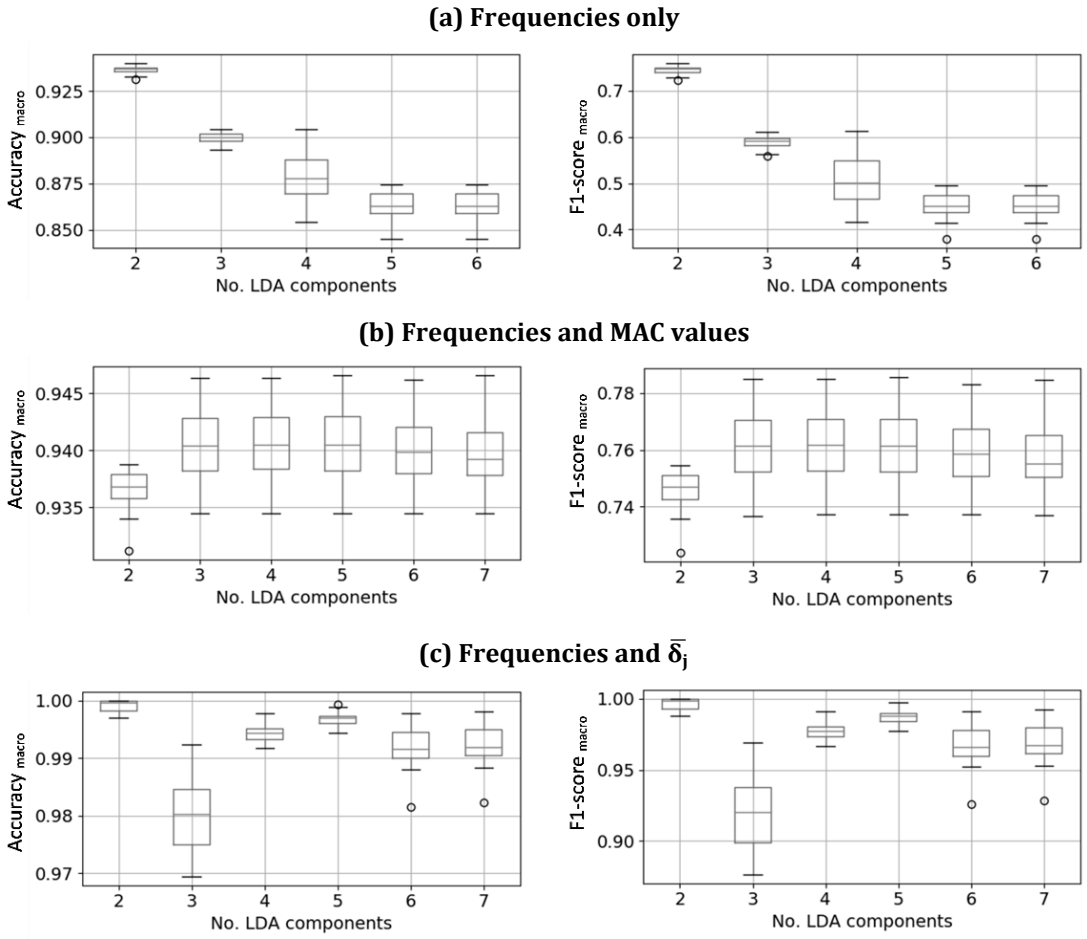


Figure 5.5: Macro average of the metrics from the hard-threshold clustering results for a varying number of LDA components. The titles report the features deployed.

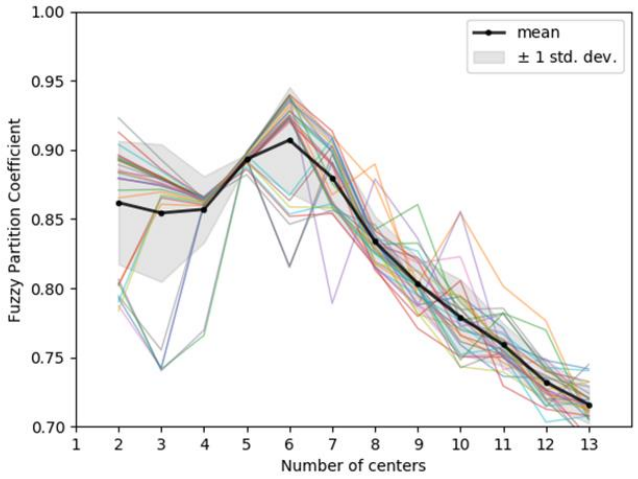
The generally low variability of the prediction indicates the independency of the models of the nacelle position. By accessing the first two LDA components only, macro accuracies and macro F1-scores close to 100% are obtained. For each of the features' combinations in analysis, the identified best hyperparameter, the corresponding estimated metrics, and their 95% confidence intervals, are reported in Table 5.1.

Table 5.1: Summary of the optimal number of LDA-transformed features and estimated metrics on the test set.

Input	Number of features	Optimal no. of LDA features	Accuracy $_{macro}$ (95% CI)	F1-score $_{macro}$ (95% CI)
Frequencies	5	2	93.6% +/- 0.3%	74.4% +/- 1.4%
Frequencies + MAC values	10	5	94.1% +/- 0.5%	76.4% +/- 1.8%
Frequencies + $\overline{\delta_j}$	23	2	99.8% +/- 0.1%	99.5% +/- 0.5%

The goodness of the detection, based on the features' combination including the frequencies and the $\overline{\delta_j}$, can also be observed in the generally high values of the fuzzy partition coefficient of Figure 5.6 (b). The coefficient represents how cleanly the data are separated into the selected number of clusters [222]: it ranges from 0 to 1, with 1 being the best separation. The scatter in the results is again given by the training and testing of several models on subsets of the dataset. Although the models do not reach their best performances by clustering into eight groups of behaviour – but into six –, it is observable that the detection achieved by using the $\overline{\delta_j}$ generally outperforms the one by using the MAC values, in Figure 5.6 (a).

(a) Frequencies and MAC values



(b) Frequencies and the $\overline{\delta_j}$

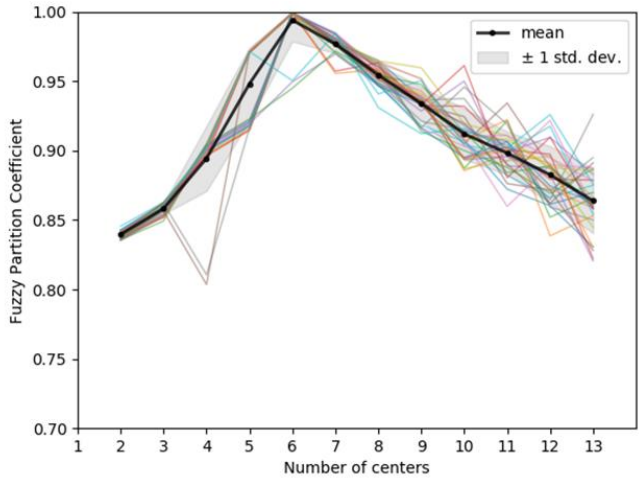


Figure 5.6: Plotting of the fuzzy partition coefficient for a varying number of cluster centres, training the model on either (a) or (b).

5.6.2 Validation for variation of environmental conditions

Figure 5.7 shows the performances of the tuned fuzzy clustering models of Section 5.6.1 on the first validation datasets with slightly increased scour depths. The results are presented in the form of a confusion matrix. A grey colour scale is used to indicate the density of the data sample for each pair of true-predicted labels, with dark grey being 1 (or 100% data samples) and white being 0. The diagonal of the confusion matrix, in dashed lines, represents the correctly labelled predictions, which is supposed to be populated by density 1 (thus in a dark grey colour). To ease the interpretability of the results, the validation outcomes are presented here only for the best performing models on the training and testing sets – cf. Figure 5.5.

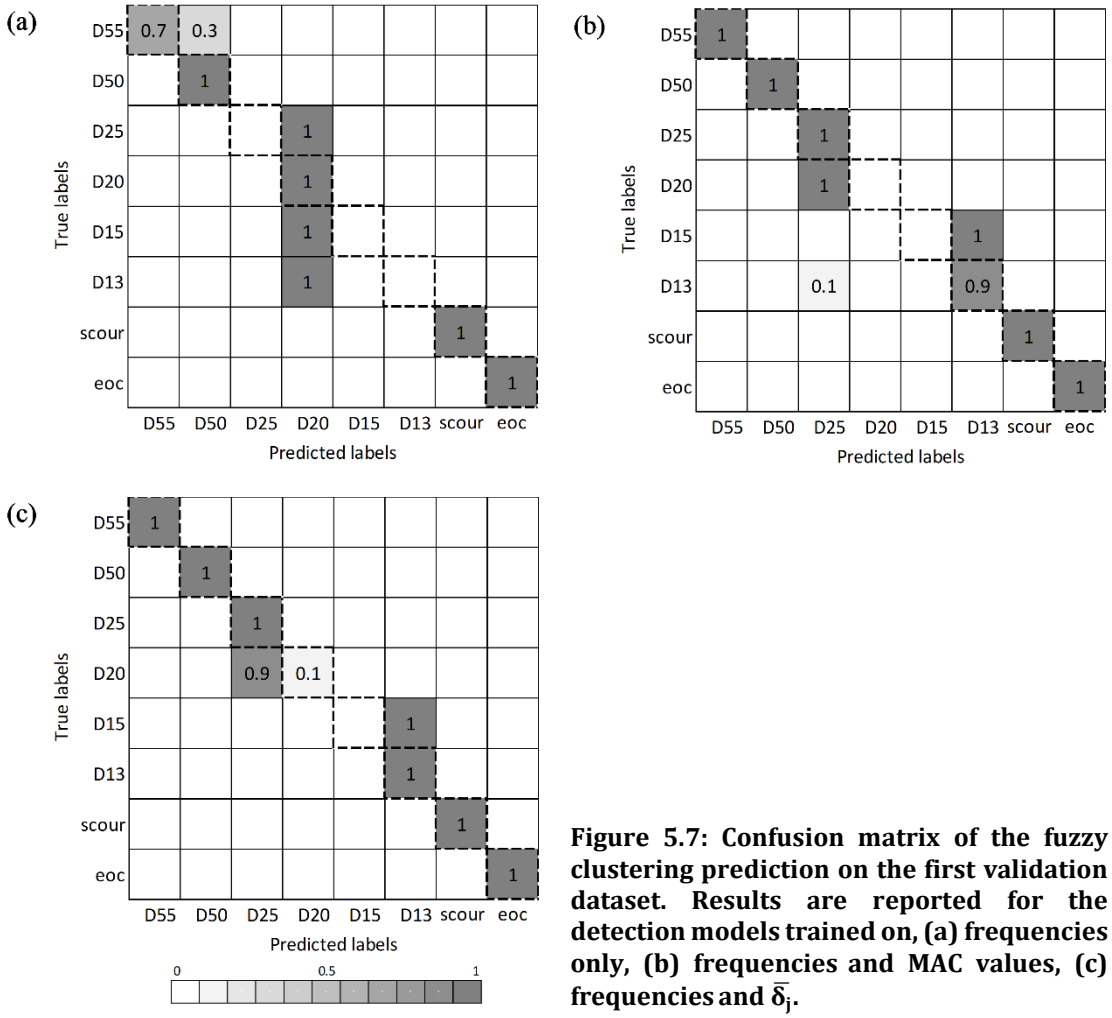


Figure 5.7: Confusion matrix of the fuzzy clustering prediction on the first validation dataset. Results are reported for the detection models trained on, (a) frequencies only, (b) frequencies and MAC values, (c) frequencies and $\overline{\delta}_j$.

In Figure 5.8, the algorithms trained on MAC values and $\overline{\delta}_j$ are further validated on the second validation dataset replicating the brace disconnections, together with higher depths of local scour, yet within the scour allowance. In Figure 5.8 (a), it can be observed that the algorithm relying on frequencies and MAC values transfers all its predictions to the extreme scour scenario. In contrast, the algorithm trained on frequencies and $\overline{\delta}_j$ correctly identified the disconnection damages as such. Nonetheless, as shown in Figure 5.8 (b), the addition of extra scour caused the mislocation of the damage for levels 55, 20, 13 and 15.

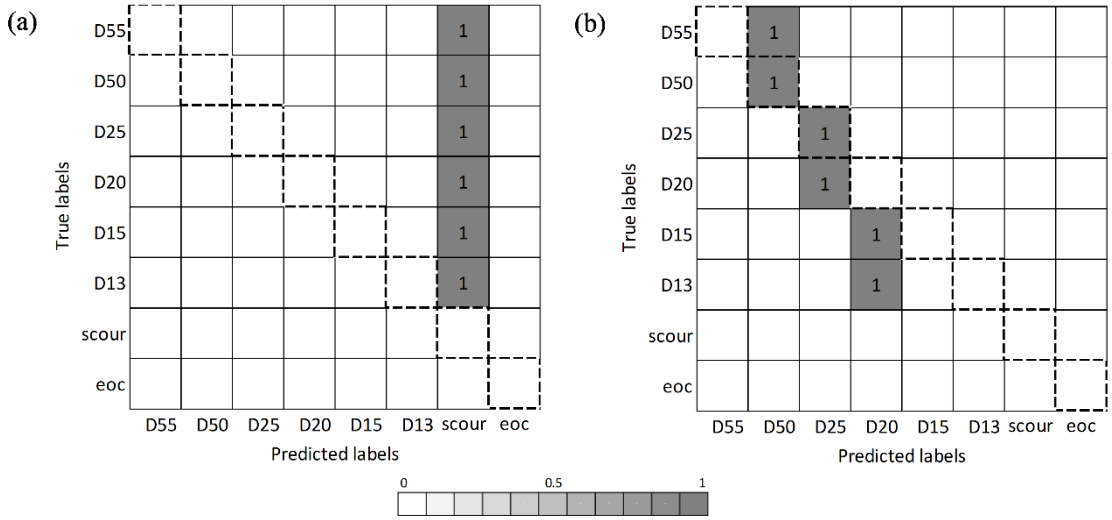


Figure 5.8: Confusion matrix of the fuzzy clustering prediction of the dataset for the combination of single brace damage and scour depth approaching the design allowance. Results are reported for the detection models trained on, (a) frequencies and MAC values, (b) frequencies and $\overline{\delta}_j$.

To show the benefit of the probabilistic monitoring approach, the membership predictions of the fuzzy clustering are reported in Figure 5.9 as histograms of probabilities for case D50 of Figure 5.8. It is evident, in Figure 5.9 (a), that the misclassification into extreme scour scenario for the algorithm trained on frequencies and MAC values, is associated to generally low membership to any of the simulated labels. Even if the assigned label – i.e., "scour" – has clearly the highest predicted probability, its value is below 0.3. Concerning the detection via frequencies and $\overline{\delta}_j$, it can be observed, in Figure 5.9 (b), that although the damage scenario D50 has the highest probability, the true scenario label D55 shows a probability higher than the remaining scenarios.

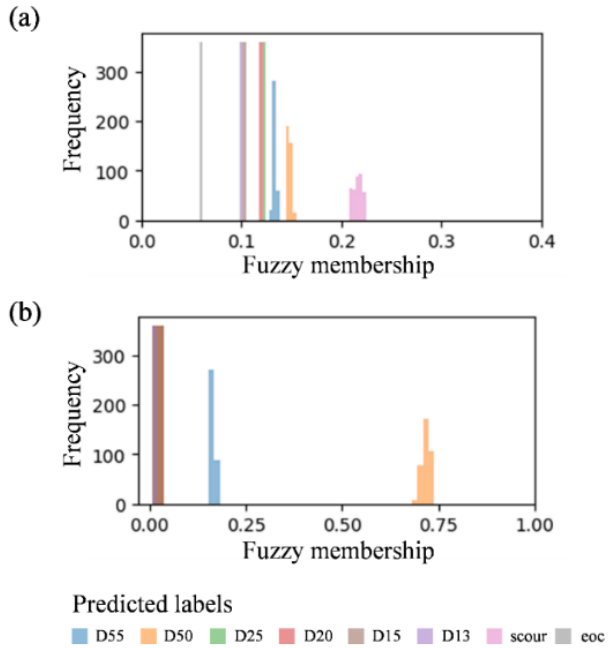


Figure 5.9: Histogram of soft clustering membership predictions on the dataset for the combination of D50 and scour depth approaching the design allowance. Results are reported for the detection models trained on, (a) frequencies and MAC values, (b) frequencies and $\overline{\delta}_j$.

5.7 Discussion of the challenges and limitations towards field applicability

The results, shown and discussed in Section 5.6, prove of the feasibility of the suggested approach for the detection and location of failure events in the jacket sub-structure of the offshore wind turbine in analysis. The monitoring strategy outlined in Figure 1 is achieved to the extent of the "Detection Model Training" – based on the simulated data. This detection algorithm fulfils the criteria of (i) diagnostic capability, (ii) low-cost – as opposed to any other ad-hoc monitoring system and field inspections –, and (iii) transparency of reasoning process, as required for the industrial needs delineated in Section 5.3.1. As concerns the eventual use of this probabilistic models for decision making (iv) of maintenance actions, the fuzzy clustering method allows to judge the prediction for the membership of the data to all the possible classes. However, this will be not as easy to interpret for the real-time data and the raising of alarms. Instead, it should be considered to make some engineering judgment on the evolution of the predictions in time. The implementation for real-time field monitoring (v) requires, as a next step, to verify the accuracy of its predictions to a set of data from the real structure.

Some of the challenges of dealing with modal data extracted from field measurements – especially in the case of offshore wind structures – come from their scatter and fluctuation in time caused by complex loads and rotating mass. As a first step one could apply the detection algorithm on only the data from the idling turbine – as here suggested –, when the excitations and the inoperability of the turbine have less of an impact on the methods for the extraction of the modal properties. Alternatively, the extracted modal properties can be pre-processed, by filtering mode shapes that do not satisfactorily match with the analytical modes of the FE-updated. By setting a suitable threshold on the distance – either in terms of MAC value or as combination of frequency difference and MAC value – between the extracted and the analytical modes, some of the scatter in the data would be removed without losing important information for the detection algorithm.

However, it must be noted that this filtering procedure, as well the lack of excitation and thus poor OMA performance, can quite often lead to a lack of some required modes. In this respect, multiple detection models should be setup for adapting the prediction to the varying number of features available. This approach and the likely drop in accuracy caused by the removal of modes in the training phase must be yet investigated.

5.8 Conclusions

This study demonstrated the feasibility of the identification of damage scenarios and their location based on the tracking of the modal properties for an offshore wind jacket structure. The approach suggested is based on the training of an unsupervised fuzzy clustering algorithm, after having applied a supervised features transformation technique (i.e., LDA), on a reduced set of data, for obtaining the maximum separability of the clusters. The detection scheme fulfils the identified needs in low-cost equipment, transparency, probabilistic output and low computational effort for real-time monitoring and decision support.

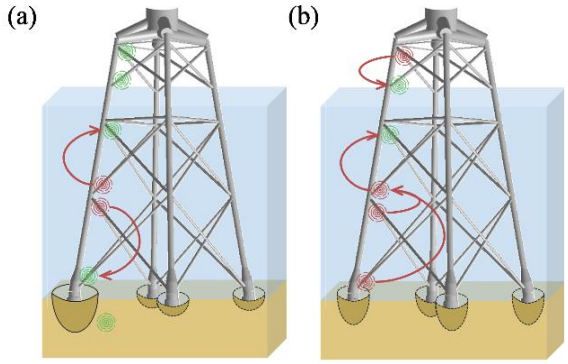


Figure 5.10: Illustration of the detection algorithm capability – trained on frequencies and the $\overline{\delta_j}$ –, to locate (a) each anomaly scenarios for slight variations of scour depth, and (b) the integrity loss of brace members for local scour depths close to the design allowance.

The results from applying the trained algorithm on the validation datasets showed the correct detection of all anomalies with promising capabilities to identify the location of the brace integrity loss. The healthy status and extreme scour scenarios were always classified correctly. Additionally, the brace disconnection-damage was always classified as such. Best damage location capabilities were seen by combining frequencies and $\overline{\delta_j}$ as training features, followed by the combination of MAC and frequencies. The frequencies-only detection showed the most mistaken results as concerns the location of the anomalies. A summary of the damage location capability of the best feature combination is further visualised in Figure 5.10. Each anomaly is indicated with a circular sign, coloured in green if the location is correctly identified, and in red otherwise. The arrows are used to point to the mistaken location of the damage. A good identification of the location of the brace integrity loss is possible for small variations of scour depth – Figure 5.10 (a). Mistaken identification of the damage location is likely for higher scour depth variations – Figure 5.10 (b). Yet, it is worth noticing that the algorithm correctly distinguishes the damage locations between above and below the water level.

Chapter 6

Detection of structural damage in offshore wind jacket substructures based on low-resolution data

This chapter discusses the feasibility and applicability of machine learning techniques to the detection of structural damage in the jacket substructure of an offshore wind turbine, by employing standard 10-minute SCADA data. Two frameworks are possible, depending on whether the damage equivalent load (DEL) varies, or not, with the presence of a damage in the structure. If the DEL is a good indicator of the presence of a failure mechanism, the scientific literature has shown that the SCADA data have the potential to quantify it and its deviation. If not, the signals of the SCADA data and of any other low-resolution measurements should be investigated for the detection purpose. The feasibility of this detection approach is proven by using well-known binary classifiers for the detection of the damaged status of the turbine [1]. However, these supervised models strongly rely on the quality and type of simulated data, and thus on the virtual prototype deployed for the generation of the data. This limitation, and the potential solutions offered by the unsupervised models, is discussed in an applicability study [2].

6.1 Background

The accuracy of the fatigue loads estimated from direct monitoring is generally affected by missing data and noise in the acquired signals [24]. Data-driven approaches have been

[1] **D. Cevasco**, J. Tautz-Weinert, U. Smolka, A.J. Kolios, (2020), "Feasibility of machine learning algorithms for classifying damaged offshore jacket structures using SCADA data", *IOP Conference Series: Journal of Physics*. https://doi.org/10.1088/1742-6596/1669/1/012021

[2] **D. Cevasco**, J. Tautz-Weinert, A.J. Kolios, U. Smolka, (2020), "Applicability of machine learning approaches for structural damage detection of offshore wind jacket structures based on low resolution data", *IOP Conference Series: Journal of Physics.* https://doi.org/10.1088/1742-6596/1618/2/022063

shown capable of predicting fatigue damage equivalent loads (DELs) at several locations of an offshore wind turbine (bladed and substructure) by using low resolution SCADA data only [226]. Therefore, to avoid relying on the installation of strain gauges and dealing with data management issues, several authors proposed and validated the use of ten-minute statistics of available operational SCADA signals for the prediction of structural fatigue loads from trained neural network models [56], [227]–[229]. In [226], Cosack tackled for the first time the estimation of fatigue loads from standard signals from the operating turbine and proved its feasibility. Neural networks algorithms were used for predicting the main fatigue loads on the blades and the tower of a 5 MW prototype (Multibrid M5000) and a 2.5 MW commercial (Nordex N80) onshore wind turbine, excluding wake operation. Additionally, he studied in detail the impact of environmental and operational disturbances on the input and targeted signals.

Based on these findings, Smolka et al. [55], [56] extended the analysis to offshore wind structures. In [55], they investigated, on simulated data, the sensor layout necessary to replace the need for sea state measurements for predicting the fatigue loads on the tower and the support structure. By training and testing neural networks on two operational conditions (idling and power production condition), they deduced that, for a 4-legged jacket structure, the employment of a either a tower top or tower base accelerometer can improve the accuracy of the estimation. The main findings of [56], for the estimation of the tower base bending fatigue loads of a 5 MW wind turbine on a tripod foundation (AREVA M5000), can be summarised as follow:

- The results from the correlation analysis showed that total set of 64 inputs, among standard SCADA data and their statistics, can be potentially reduced to 28 inputs for power production conditions. A configuration with only four standard data signals – namely the generator speed, the blades pitching angle, and two-dimensional acceleration of the tower top – to be measured and the computation their statistics, for a total of 10 inputs to the neural network, achieves a mean estimation error of 0.5% on the test set, only slightly higher than the full set one.
- It is necessary to capture a broad range of operational conditions and turbulence intensity level for achieving low estimation error during training and testing. For this reason, although around 50 and 100 samples per wind are shown to be optimal – for a total of half a month of data for 24 wind speeds –, the site-specific probability distributions of the environmental conditions dictate the length of the measurement campaign for the training of the neural network for fatigue monitoring.

Finally, they concluded that the combination of several specialised neural networks should be used to track the loading history of the turbine. The prediction on the validation dataset is just slightly underestimated (mean error -0.3%) for loading conditions like the one experienced by the algorithm training. However, they observed that the cumulated error tendency is to drift away towards a persistent underestimation.

Vera-Tuleda et al. [228], [229] focused their analysis on the prediction of blades fatigue loads of the turbine in the EnBW Baltic 1 offshore wind farm. Their predictions were accurate throughout all WTG operating cases, with an average overprediction error below 1.5%. In wake flow conditions only slightly deteriorated the estimation of the DEL of the blade root bending moment. Therefore, they concluded that a reasonable monitoring system can be based on the training of a neural network model without the need to distinguish between inflow conditions.

6.2 Problem statement and aim

The literatures' findings, summarised in Section 6.1, have shown that the SCADA system data holds the information and the relationships for the prediction of the fatigue of the structures of offshore wind turbine. Therefore, if a failure mechanism has an impact on the fatigue cumulated in the wind turbine's structure, it can be that the SCADA signals are good at predicting the presence of the damage. The aim of this work is to propose an alternative to the current practice for the identification of a damage in an offshore wind turbine substructure – to-date mainly based on either on-site practical assessments, or data-driven vibration-based methods [230].

The focus of the analysis is on the feasibility of the detection of a structural damage by capturing abnormal variations of the ten-minute statistics of the SCADA system data. As opposed to the SHMS – installed on no more than 10% of the turbines across a farm –, the SCADA system collects operating and environmental signals the throughout the service life of all assets. Therefore, the proposed approach has the potential to be applied to all units across a farm. Furthermore, the requirement for low-frequency statistics of the SCADA data is less of a burden from a data storage and handling point of view.

6.3 Methodology

Section 6.3.1 introduces to the approaches adopted for proving the feasibility and discussing the applicability of a SHM scheme based on low-resolution data and non-parametric approaches. In Section 6.3.2, the setup of the EOC for the FLS semi-coupled simulations – with the FEM in healthy and damaged status – is described. Finally, Section 6.3.3 briefly introduces to the main concepts of the traditional machine learning models used for the classification and the normal behaviour task.

6.3.1 Suggested approach

For the feasibility study, the data are generated from the as-designed numerical representation of the WK64 turbine. Simulated data is used to train and test classifiers for damage detection in a binary form (damaged/healthy). The tests with unseen conditions for stochasticity of the loading conditions and variability of wind inflow parameters are conducted. The detection capability of logistic regression, support vector machine, k-nearest neighbour, random forest, and Gaussian naïve Bayes based classifiers is tested in this feasibility study. Each algorithm's best performance is obtained through an iterative process across the dataset for variation of EOC.

Once the best classifier is identified to perform the binary-classification on the integrity status of the turbine's foundations, the capability of the classifiers to accommodate the uncertainties associated with the model used for the training of the algorithm is investigated into an applicability study. The concept behind this process is illustrated in Figure 6.1. The model of the structure used for the design purpose is thought to give a good representation of the behaviour of the real structure. However, this model is aimed to meet required safety levels. In contrast, the real structure differs from the designed structure (described with system uncertainty Δ_3) due to manufacturing tolerances, deviation in environmental parameters and changes during lifetime as e.g., corrosion and scouring. For this reason, the FEM of the structure is updated to mimic the behaviour of the real structure

based on measurements [194]. The FEMU can be described with an uncertainty Δ_1 with respect to the virgin model. Representing the as-installed status of the turbine, the system uncertainty to the real structure (Δ_2) should be small compared to the one from the virgin model ($\Delta_2 \ll \Delta_1$). The objective of this study is to verify if the algorithms trained on the data from the as-designed model of the jacket are successful in classifying the status of the turbine when tested on the datasets from the FEMU. If this is achieved, then it will prove that the virgin model, and the algorithms trained on it, have the potential of being directly employed for the damage detection task of the real structure (since $\Delta_3 \sim \Delta_1$).

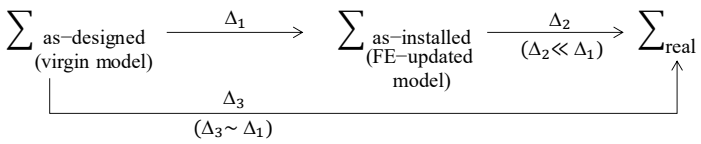


Figure 6.1: Relationship between the modelled as-design, as-installed and real systems (∑), and the uncertainties (Δ) among these.

6.3.2 Data generation

It is very unlikely to have access to real run-to-failure data of the wind turbine substructure, as these events rarely happens, and such data would be unlikely to be shared for research purposes due to confidentiality reasons. Therefore, the collection of data representing the dynamics of the wind turbine structure in its healthy conditions, and in response to anomalies, is obtained via FLS simulations run in the LACFlex software, as explained in Section 4.4.3. A set of representative FLS load combinations is investigated – refer to Section 6.5.1 and 6.6.1.

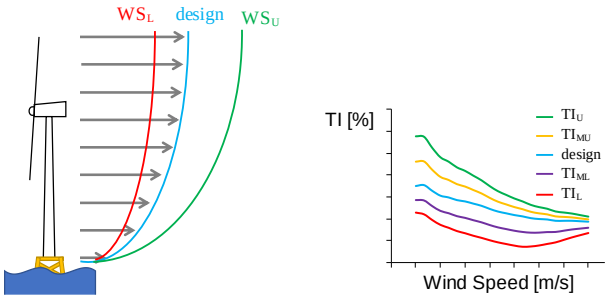


Figure 6.2: Visualisation of the uncertainties considered (upper and lower bounds) in the wind farm flow condition of the fatigue design simulations: wind shear (left), and TI (right).

To account for the uncertainty associated to the real operational conditions of the turbine, the healthy and damaged structural responses are derived for changes in the wind farm flow conditions as illustrated in Figure 6.2. The wind shear exponent (WS) – which is potentially correlated to multiple factors [231] – is varied from the design base specification to a minimum of 0.08 (WS_L) and a maximum of 0.3 (WS_U), changing the distribution of the normal wind profile. The 90th percentile of the effective turbulence intensity (TI) is used as reference, a value considered representative for the fatigue design calculations [232]. Based on the experience from a similar farm, an upper (TI_U) bound curve is defined to represent the missing extreme cases. Similarly, a lower (TI_L) bound curve is drawn corresponding to

the 10th percentile of the effective turbulence intensity. These are then implemented in a Mann turbulence model to represent the fluctuating wind field.

6.3.3 Machine learning algorithms

Several detection algorithms are experimented to identify the possible best predictors of the damage. These algorithms range from supervised to unsupervised approaches, and they are implemented via the Python machine learning package (skikit) [233].

For the training of the ***supervised algorithms*** the "grid search" method is followed [223]. This implies that various combinations of configurations of the hyperparameters are tried during the training phase. The combination giving the overall best performance on the folds of the training set is selected – following the so called "cross-validation" method. These folds, i.e. subsets, are selected by applying a stratified k-fold approach [223], which divides the training set in homogeneous splits of healthy and damaged data samples. The algorithms tuned with the optimal set of hyperparameters, are then fitted to the full set of training data. As concerns the ***unsupervised algorithm,*** the hyperparameters selection is done manually because of the targeting satisfactory results of multiple performance metrics – i.e., precision, recall and false alarm rate (refer to Section C.2 of the Appendices) – on both the training and the test sets.

The classifiers are trained on balanced datasets containing both the healthy and damage samples. Each dataset for varying environmental loadings of the structure has a 50:50 ratio of healthy and damaged conditions. In contrast, the anomaly detection algorithm is instructed to build a decision function for recognizing the healthy data only and considering all data samples do not experience during the training phase as damaged.

6.4 Case study

As damage case study, the full loss of a cross-member of the jacket structure is simulated. The stiffness of a brace close to the seabed which connects diagonally two of the legs of the jacket is reduced to a value close to zero. This failure location is selected because, as shown in Section 4.5, it is associated to a high deviation of the global natural frequencies and a potentially high impact on the jacket structure fatigue life.

6.5 Feasibility study

Figure 6.3 shows the workflow for the demonstration of the feasibility of the detection approach. The dynamics analysis – cf. Section 4.4.3 – of the as-designed structure is setup, to record its response to the EOC and as a consequence of a structural failure mechanism. The time histories of loading conditions, and the relative structural responses, are post-processed into ten-minute statistics, similarly to SCADA data. The data are collected into a database, which is sequentially accessed to construct the datasets for the training and testing of machine learning algorithms. The features, i.e., the set of independent variables used for the prediction, of each dataset are standardised before being used in the models. The dataset in analysis is then divided into subsets, named "Tr" for the tuning/training subsets and "Te" for testing subsets.

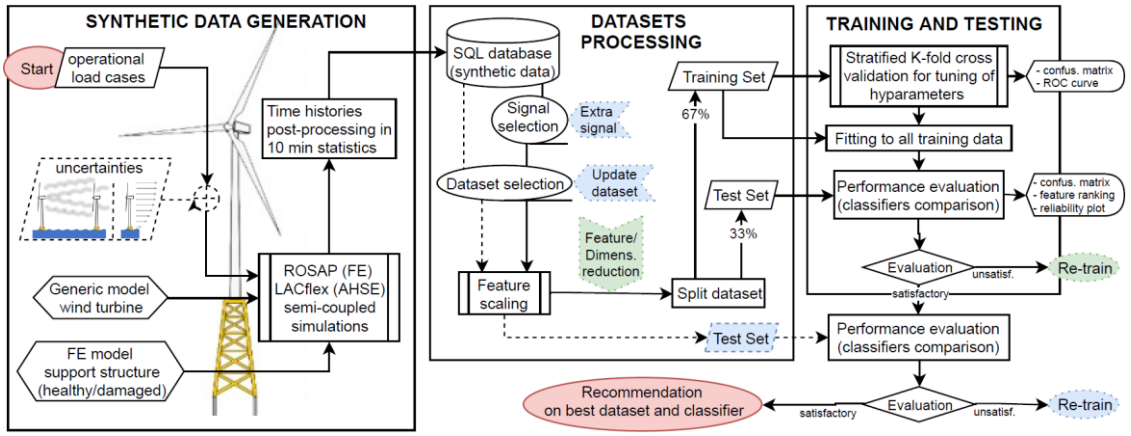


Figure 6.3: Workflow for the data generation, the datasets processing and the algorithms training and testing. In "dotted lines" are the iterative processes for the re-training of an algorithm.

The tuning, training, and testing methods are applied in the same manner for all datasets considered and all algorithms investigated. If a classifier performs in an unsatisfactory way on the test set for stochastic variation of the loadings, feature and dimension reduction techniques are applied (see Figure 6.3, in green). The most promising algorithms are then tested on the subsets for variations of the environmental parameters (see Figure 6.3, with "dashed" lines in the block for "synthetic data generation" and "data processing"). If insufficient detection performance is exhibited at this stage, the subset for the training of the classifier is extended to include additional features and/or the data from the misclassified cases. Tuning, fitting, and testing are then repeated (see Figure 6.3, in blue). Eventually, a recommendation of the best training set and algorithm for the damage detection task is given.

6.5.1 Databases and preliminary analysis

Due to the symmetries of the jacket substructure, and the location chosen for the implementation of the structural damage, a combination of four wind directions and twelve wave directions is deemed sufficient. The average wind speed at the hub height is simulated for six values, to represent three below and three above rated conditions. Nine realizations of the wind and wave time histories are processed for each loading combination, to guarantee the capability of the detection algorithms to distinguish the response due to load stochasticity from one of the damaged status. Therefore, a total of 2,592 simulations per turbine status, and TI value are performed. The load case settings are reported in Table 6.1.

Table 6.1. Load combinations simulated per jacket status (healthy/damaged) and jacket model (as-designed/as-installed).

Load case	Wind speed	Wind direction	Wave direction	Seeds	Total
Settings	(a) 14, 18, 22 m/s (b) 4, 6, 8 m/s	0°, 90°, 180°, 270°	0°, 30°, 60°, 90°, 120°, 150°, 180°, 210°, 240°, 270°, 300°, 330°	9	–
Number	6	4	12	9	2,592

As outlined in [22], any structural health monitoring method employed in the detection task must be able to distinguish between signal variations related to EOC, as opposed to the ones corresponding to a structural anomaly. The damage implemented was observed to lead to significant changes of the second modes and their natural frequencies with respect to the healthy status. It is then reasonable to expect that this has an impact on the loads at the tower base, building the interface between the turbine and the jacket structure. However, variations of the environmental conditions could affect the global response in a similar manner. Therefore, a pre-analysis for a reduced number of simulations is performed to investigate the influence of wind flow parameters (upper- and lower-bound values) on the structural response.

The time histories, output from the AHSE code with 50 Hz frequency, are post processed into ten-minute minimum (min), maximum (max), average (mean), and standard deviation (std) values. At first, all the measurable signals are collected into the database of operational conditions, potentially being meaningful indicators of the structural failure. The potential predictors of the damage are then investigated by quantifying the deviation of the statistics, given the scatter for the stochastic variation of the loading, for the healthy and damaged status data.

Effect of environmental parameters variations

The sensitivity of the loads at the tower interface on the WS and the TI parameters is presented in Figure 6.4 and Figure 6.5, respectively, with respect to a design load case at a below rated condition.

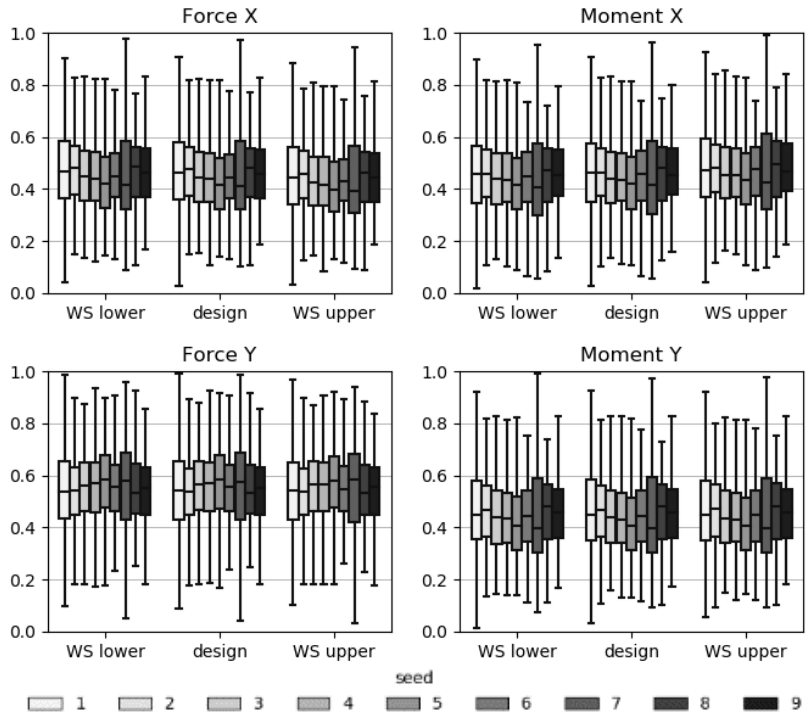


Figure 6.4: Box plots of the interface load of the as-designed WK64 turbine, for lower-, design-, and upper-bound of the wind shear parameter, at a below rated condition.

The box plots of the time histories of the interface forces and moments, in the fore-aft (y) and the side-side (x) direction, are reported for the nine stochastic variations of wind and wave loadings.

In Figure 6.4, it can be noticed that the wind shear exponent does not seem to significantly affect any of the loads. By contrast, a high level of turbulence intensity is associated to higher load ranges and standard deviations compared to the design base scenario – cf. Figure 6.5. Opposite behaviour is then observed for the low-turbulence level. Consequently, it is deemed necessary to feed the datasets for the varying TI parameter into the machine learning training phase. On the other hand, it is assumed that the variation of the wind shear that can be fully captured in the structural dynamics of the design load case.

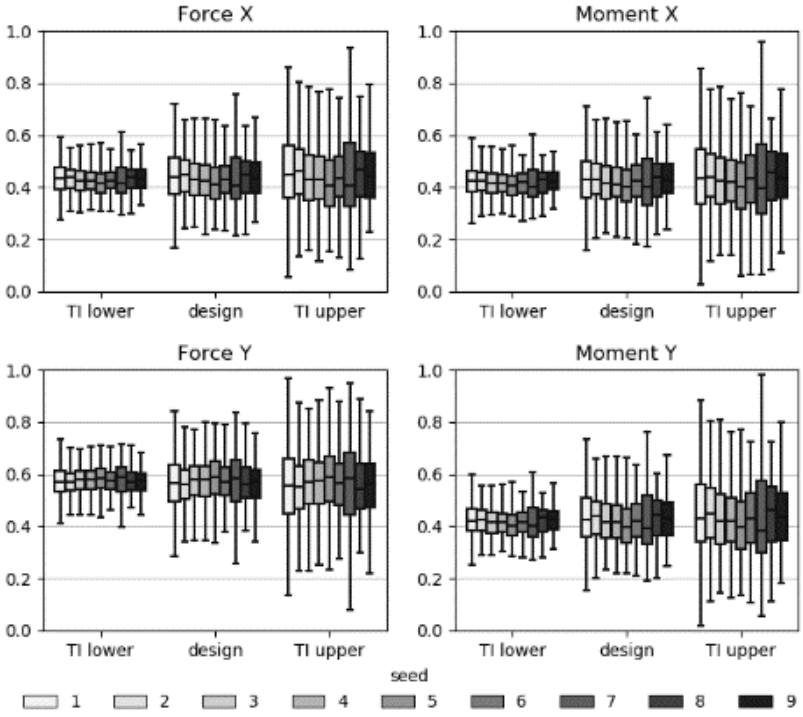


Figure 6.5: Box plots of the interface load of the as-designed WK64 turbine, for lower-, design-, and upper-bound of the wind shear parameter, at a below rated condition.

Detectability and signals deviation

A visual representation of a preliminary analysis of some of the channel's statistics is given in Figure 6.6. This plot reports the mean of the tower top accelerations and the DEL of tower bending moments at the interface. It can be observed that, with respect to the DEL of the bending moments, only the *MyF0*, relative to the side-side motion, deviates slightly from the healthy status for some of the angles of the wind-wave misalignment. By contrast, it has been noticed that the DEL of the *MxF0* of the tower base, in the fore-aft direction, as well as the DELs of the moments at the foundation base, and at the blade roots remained mainly unaffected. On the other hand, it has been observed that the tower top acceleration channels (*AxTT* and *AyTT*), the power output range (for the extreme turbulence case), the tower bottom acceleration, the rotations, and some of the forces and moments in the drivetrain (mainly at the main bearing) captures the discrepancies between the healthy and the structural damaged-status, throughout the load combinations considered.

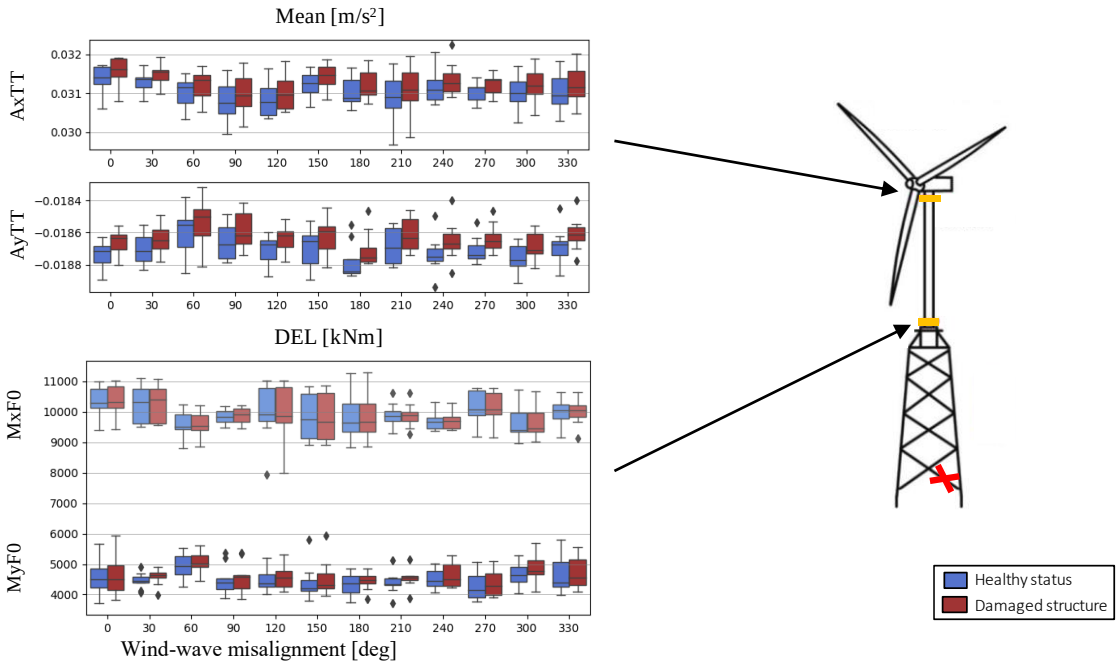


Figure 6.6: Box plots of the ten-minute mean tower top acceleration (A•TT) and interface bending moments DEL (M•F0) against wind-wave misalignments for an above rated case load combination.

Sets of data for training and testing

The datasets built for training and testing of the classifiers are reported in Table 6.2, while the details on the several sets of sensors tested are reported in Table 6.4. Table 6.3 records the characteristics of the subsets for used for the training, the testing, and the validation of the classifiers.

A detection through standard SCADA signals (sensor setup S0) is preferred and attempted at first. Initially, the investigation of the feasibility of the status classification is performed on the dataset of the design load combinations (dataset D0). Then the data derived for the different TI levels are added to this base scenario, expecting significant variations in the loadings and response of the structure (datasets D1, D2 and D3).

Table 6.2: Datasets for the feasibility analysis.

Datasets	Effective TI	No. simulations
D0	design	5,184
D1	design + TI_U	10,368
D2	design + TI_L	10,368
D3	design + TI_U + TI_L	15,552

Each of the training subsets (Tr#) consists of 67% of the data of the current set, by randomly selecting six out of the nine realizations for each load case. The remaining set of three realizations per load case, consisting of the 33% of the data in the set, is collected in the subset Te33 and used for testing. Additionally, the mid -upper and -lower TI curves (Figure 6.2) are derived, and the statistics associated to these loadings are collected into the Te3 and Te4 test sets, over that the one for upper and lower bound values (in the test sets Te1 and Te2).

Table 6.3: Subset of data for training, testing and validation.

Set	Acronym	Effective TI	No. simulations
Training	Tr#	–	67% D#
Testing	Te	–	33% D#
Validation	Va1	TI_U	5,184
	Va2	TI_L	5,184
	Va3	TI_{MU}*	5,184
	Va4	TI_{ML}*	5,184

* mid-upper and -lower TI curves selected as shown in Figure 6.2

While pursuing the best classification results, the sensor setup is redefined during the iterative process of the classifiers' training and testing – cf. Table 6.4. With respect to the preferred baseline scenario – deploying the sensor setup S0 –, the benefits of the installation of an inclinometer and a set of strain gauges, at the tower base, is tested in the setup S1 and S2, respectively. The reason for the selection of these sensors' signals, and location, is: 1) in avoiding installing monitoring devices below the water level and, 2) to maintain the analysis as independent as possible from measurement from the drivetrain, which are highly related to the specific control strategies and optimization. These sensors allow for the indirect and direct measurement of strain at the tower interface. While the inclinometers' channels are added to the datasets (i.e., any of the D#) in terms of min, max and std, the channels of tower bending moments are additionally represented by their DEL.

Table 6.4: Sets of sensors and signals considered for the detection purpose.

Monitoring System	Measurement	Signal acronym	Unit	Sensor set up				
				S0	**S1**	**S2**	**S3**	**S4**
	Nacelle direction	YawPos	[deg]	x	x	x	x	
	Wind direction	WDir	[deg]	x	x	x	x	
	Yaw angle (misalign. error)	YawErr	[deg]	x	x	x	x	
SCADA	Wind speed	Whub	[m/s]	x	x	x	x	
	Power	Pow	[kW]	x	x	x	x	
	Rotor speed	RotSpd	[rpm]	x	x	x	x	
	(Collective) Pitch angle	PiPos1	[deg]	x	x	x	x	
	Tower top acceleration	AxTT, AyTT	[m/s²]	x	x	x		
Inclinometer	Tower bottom rotation	UrxF, UryF	[deg]		x	x	x	x
Strain gauge	Tower bending moment *	MxF0, MyF0	[kNm]			x		

* at the (turbine) tower bottom, at interface between the foundation and the turbine

6.5.2 Results

These section reports the results of the binary classification performed by the different models on the several datasets of Table 6.2 and the subsets of Table 6.3. The classification metrics used are the one defined in Section C.2. of the Appendices.

Algorithms' selection

A preliminary selection of the most promising models is done on the data from D0 and the sensor setup S0, already at the tuning and training stage.

Unsatisfactory performances are shown by the following classifiers:

- **Gaussian naïve Bayes**
  Based on the strong (naïve) independence assumptions between the features, it is here implemented by selecting the first 9 principal components (eigenvalues)

corresponding already to more than the 90% of all the variance (as explained in [234]). Yet, inacceptable performance is given, probably because of violation of the normal distribution assumption for the numerical predictors.

- **K-Nearest Neighbour**
  Despite of the broad and small stepped range of K given during the cross-validation tuning and fitting, this algorithm, given its implementation in [233], fails in finding boundaries separating the two classes.

By contrast, as it can be observed in Figure 6.7, the results obtained by deploying the following classifiers are worthy of further investigation:

- **Logistic Regression**
  Logistic regression (LR) is a linear method that models a binary dependent variable, where the predictions are transformed using the logistic (sigmoid) function. As for linear regression, the model can overfit if there are multiple correlated inputs [235]. Here, it does not seem to happen, despite the high dependency of some of the features.
- **Support Vector Machine**
  The support vector machine (SVM) approach aims to find a line, surface or hypersurface for the separation of the classes. When applied to the data, it fails in finding a linear hyperplane for a correct classification. On the other hand, by projecting the data into a higher-dimensional space defined by polynomials (poly) and Gaussian (radial) basis functions (rbf), the models manage to capture the nonlinearity of the classification problem yet requiring a higher computational time.
- **Random Forest**
  The random forest (RF) method fits a number of decision tree classifiers on various subsamples of the learning datasets, providing as output the average among the single trees' predictions. This 'trick', together with the limitation in the number of features for the trees and their depth, aims to control the over-fitting.

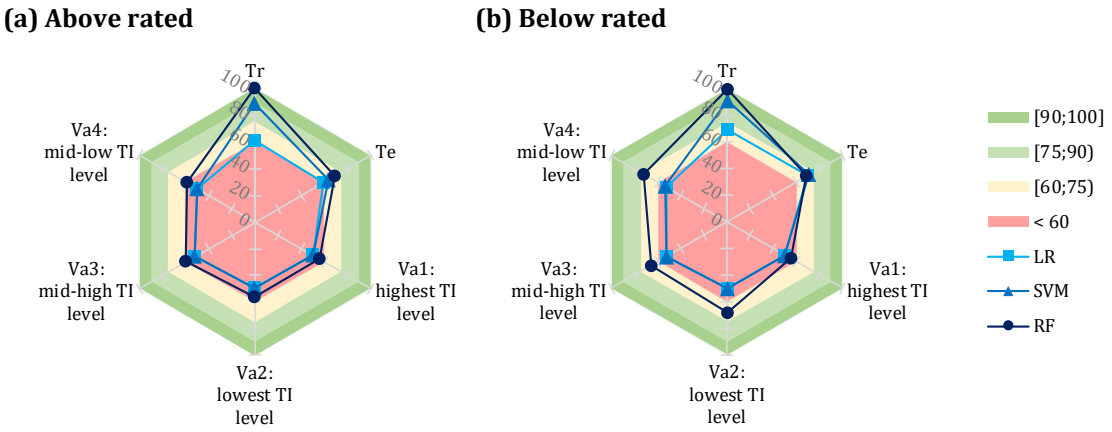


Figure 6.7: Accuracy of the most promising classifiers on the dataset D0 (thus, trained on Tr0), with the sensor setup S0, for the (a) above rated, and (b) below rated conditions. On the right, the colour legend – the targeted performances are in green.

By further testing the LR, the SVM and the RF classifiers on the test set for stochastic variations of the environmental conditions (Te), it can be observed that generally satisfactory performances are obtained for the below rated conditions; the models successfully distinguish the normal operating conditions from the damage status – cf. Figure

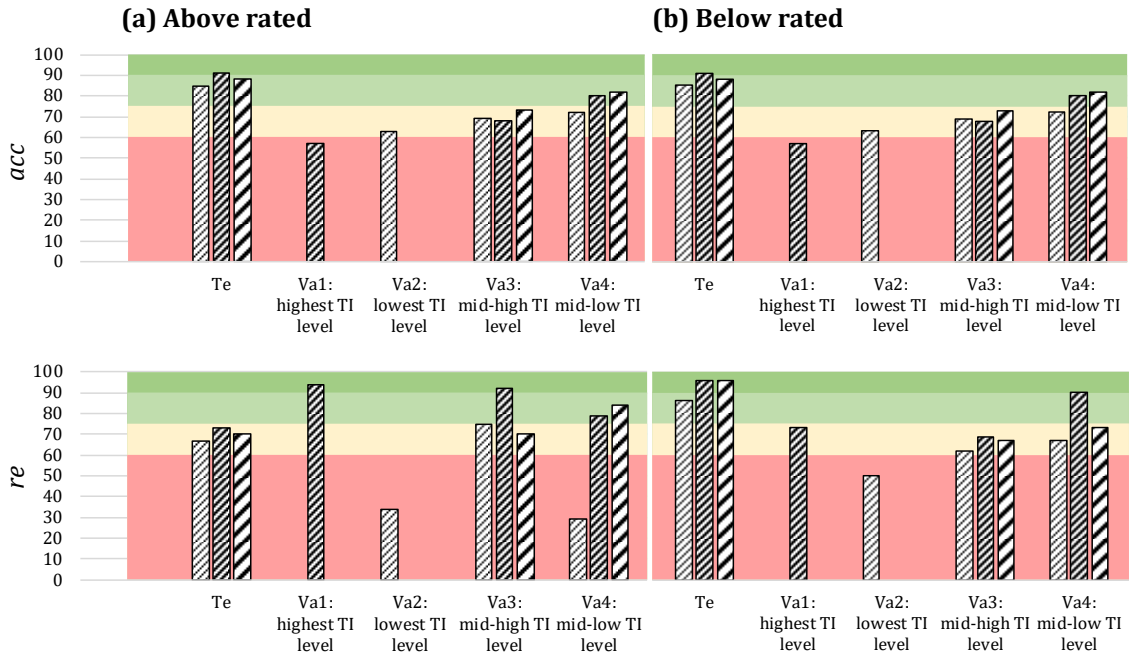
6.6 (b). Slightly more substandard performances are obtained, for all models, in the above rated conditions – cf. Figure 6.6 (a). This could be explained by the higher fluctuation of the tower top acceleration in above rated operating conditions.

The algorithms are then tested on the subsets of data corresponding to the response of the structure to variation of the TI, according to the curves in Figure 6.2 – i.e., test subsets Va1 to Va4. It is evident, from the radar plots in Figure 6.6, that none of the models can perform such generalization. Consequently, a re-training iteration is carried out by updating the training subset by either increasing the amount of data samples considered (varying the D# datasets) or changing the amount and/or type of sensors employed (varying the S# sensor set up per dataset). For brevity, the results are following reported only for the RF classifier. The complete results of the LR and the SVM models are given in a tabular format in Section D.2 of the Appendices, together with the tabular metrics for the RF classifier in the several D#-S# combinations.

Varying training datasets

First, the classification models are re-trained by adding the data samples corresponding to the varying turbulence levels to the dataset of the design load case (datasets D1, D2 and D3). Acceptable results are achieved for the RF classifier only, while the LR and the SVM do not exceed 60% accuracy – cf. Section D.2 of the Appendices. The classification results of the RF model are reported in Figure 6.8, in terms of accuracy (*acc*), recall (*re*), and specificity (*sp*), and by differentiating the results for the above and the below rated conditions.

As it can be noticed by looking at Figure 6.8, the targeted *acc* and *re* are generally achieved for the predictions on the test set (Te) and the set of mid-low TI conditions (Va4), when adding either the low-TI or both the extreme-TI data samples to the training set (D2 and D3, respectively). The performances are yet slightly better for the below rated case than for the above rated one. However, the improvement obtained by adding only further data sample is associated to a high false positive rate (defined as 1–*sp*), which is shown to reach a value up to 45% for above rated conditions.



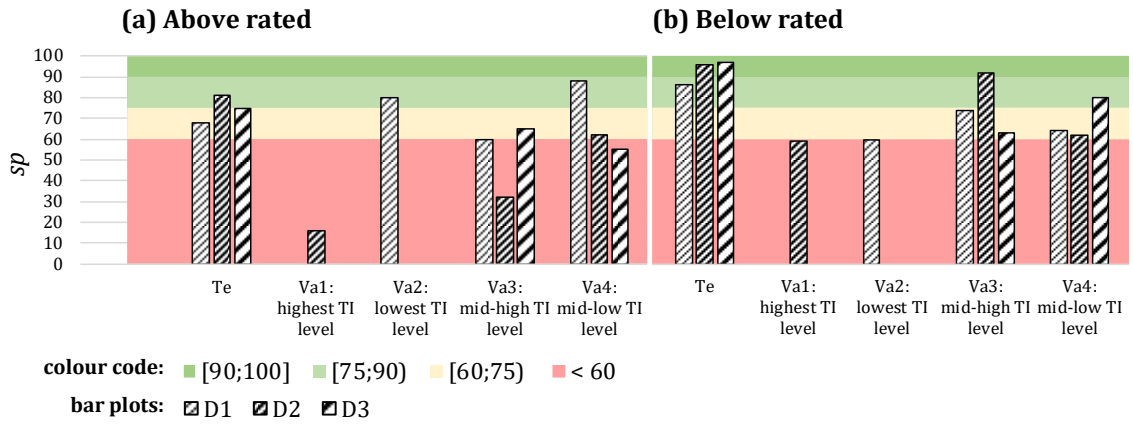
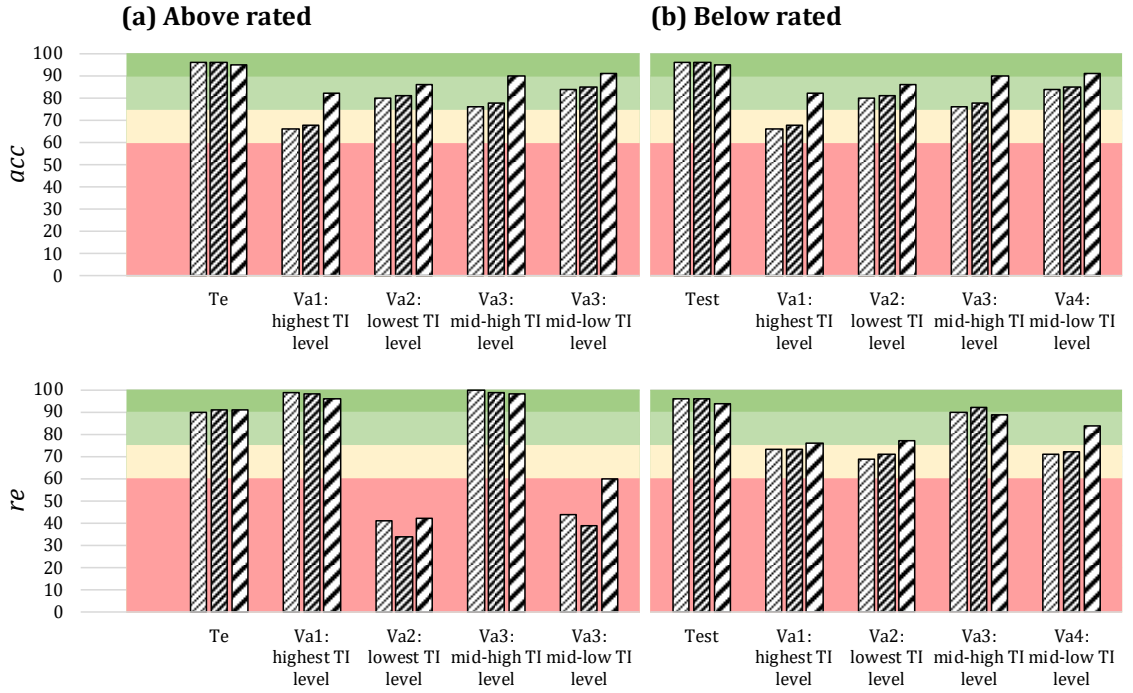


Figure 6.8: Classification performance of the RF on the test and validation sets, for varying size and characteristics of the training dataset (i.e., D1, D2, and D3 of Table 6.2), considering the (a) above rated, and (b) below rated conditions.

Varying sensor type

Figure 6.9 reports the RF classifier performances when trained on D0, but with varying type and number of sensors deployed for the detection. By adding the ten-minute statistics of the time signals from the tower base inclinometer to the SCADA data (sensor set up S1), significant improvements are achieved by the RF classifier. Due to the yet high number of false alarms (*fo*), for TI level above the 90th percentile curve (test sets Va1 and Va2), the classifier is re-trained by adding the statistics of tower base strain gauge signals (sensor set up S2), which includes the DELs as additional features.



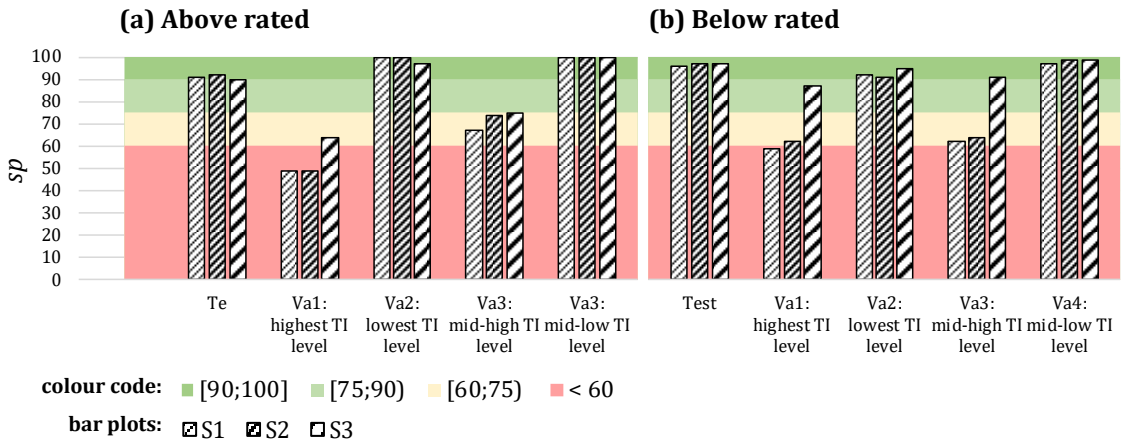
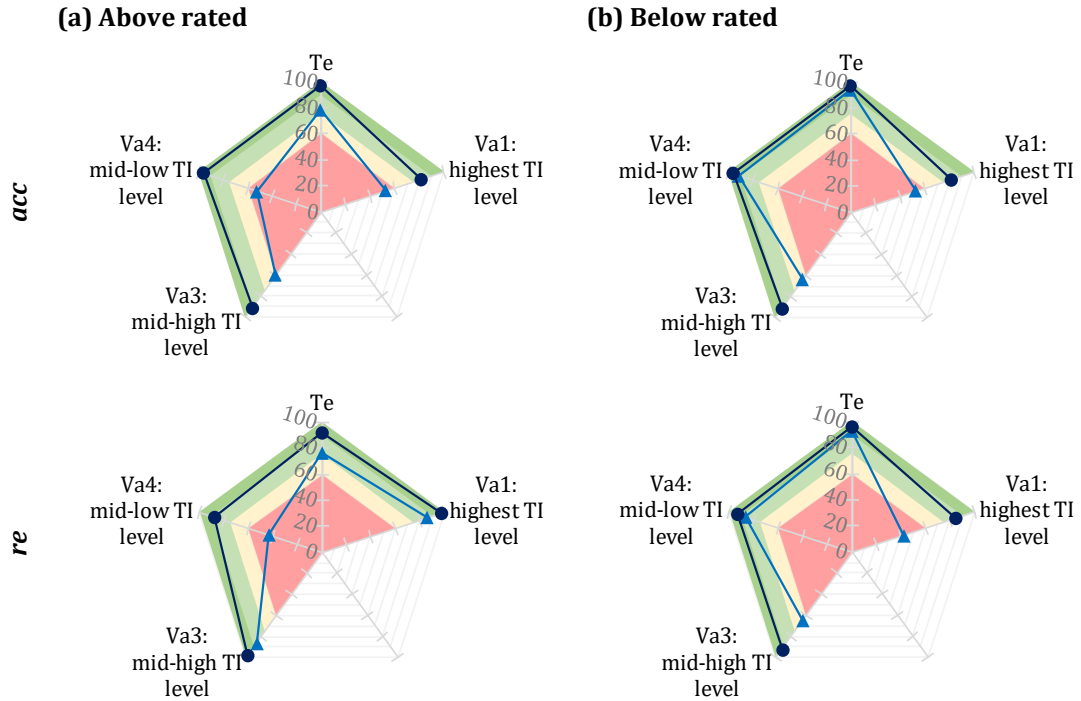


Figure 6.9: Classification performance of the RF on the test and validation sets, for varying sensors setup of the training dataset D0 (i.e., S1, S2, and S3 of Table 6.4, considering the (a) above rated, and (b) below rated conditions.

While only a slight improvement is recorded again for RF at below rated condition mainly – cf. Figure 6.9(b) in terms of *re* –, it is interesting to observe that LR and linear SVM accomplish exceptionally good results – cf. tables in Section D.2 of the Appendices. This is identified as symptom of the instability of the models due to the addition of collinearity into the analysis. The rotation and bending moment signals are, indeed, strongly correlated. The effect of this phenomenon is somewhat reduced in the RF classification thanks to random selection of a reduced number of features at each node.



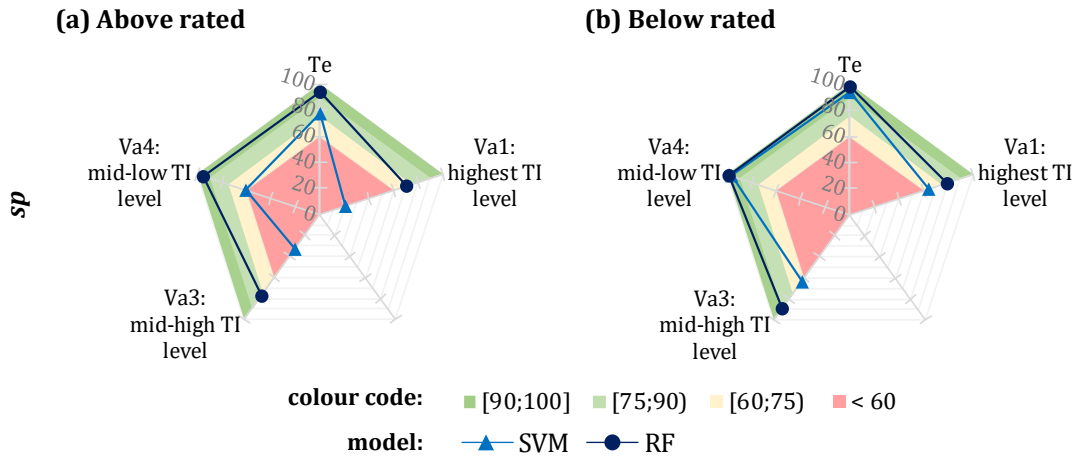


Figure 6.10: Accuracy of the most promising classifiers on the dataset D0 (thus, trained on Tr0), with the sensor setup S0, for the (a) above rated, and (b) below rated conditions. On the right, the colour legend – the targeted performances are in green.

In line with this logic, and by recognizing a generally high correlation of the acceleration signals with the power, the shaft rotational speed, and the wind speed – with *R2* values above 0.9 –, a re-training based on the use of the inclinometer statistics instead of the tower top accelerometer is attempted (sensor setup S3). The results, in the tables of Section D.2 and D.3 of the Appendices, confirm the hypothesis, by showing improved performance for the RF classifier, and unsatisfactory results for the LR and SVM classifiers.

Finally, by considering this sensor setup (S3) and by extending the training subset with data from the lowest TI level (D2), it can be observed, in Figure 6.10, that the targeted performance is achieved on all the test and the validation sets (Te, Va1, Va3 and Va4) for the RF classifier, at both below and above rated conditions. Furthermore, it is worth noting that also the SVM classifier achieves generally satisfactory performance for this dataset-sensors combination (D2-S3). Acceptable *re* and low number of false alarms are achieved on the test set Te and the validation set Va4, at below rated conditions, by employing a rbf kernel transformation.

6.5.3 Discussion on predictions' confidence

The confidence of the RF models in their prediction is discussed based on the reliability curve of Figure 6.11. The predicted probabilities for the damaged class are divided into bins – along the x-axis. The number of predicted damaged-status events are then counted for each bin and normalized on the y axis (observed relative frequency).

A well calibrated binary classifier should classify the samples such that, for instance, among the samples to which is associated a probability of 0.9, approximately 90% of the cases are classified as damaged. Therefore, the more reliable a forecast is, the closer the points will appear along the main diagonal ("perfectly calibrated"). For points of the curve below the diagonal, the model has over-forecasted the probability, while above the diagonal the probability forecasted is too small.

The reliability curves of the RF classifiers present their typical sigmoid shape [236]. This means that the algorithm is overconfident on small, predicted probability and underconfident for big, predicted probabilities. This behaviour is common for the RF classifier because the average predictions from the base-level trees can have high variance

due to feature sub-setting. A slightly better confidence of the RF prediction is given when trained on D3-S0 and tested on medium-high TI levels (Va3) – cf. Figure 6.11 (left).

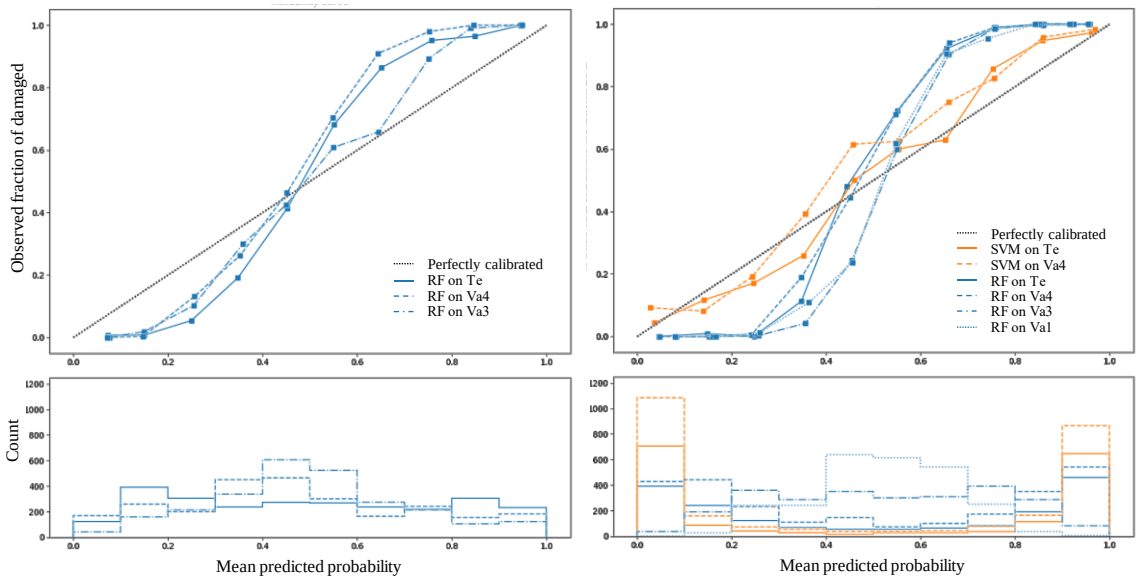


Figure 6.11: Reliability plots of the predictions from SVM (orange) and RF (blue) models, for below rated conditions. On the left, the RF model is trained and tested on the D3-S0 combination. On the right, the models are trained and tested on the D2-S3 combination.

When trained and tested on the D2-S3 combination, it can be noticed – cf. Figure 6.11 (right) – that the RF model is able to extend their prediction to all operating conditions, even for significant variation of the TI levels. Although the SVM model does not exhibits as good performance as the ones of the RF classifiers on this database-sensor combination, it can be observed that the rbf-SVM classifier is very confident in its predictions. The under-confidence of the RF model predictions is shown by the histogram peaks moving further away from 0 and 1 – cf. Figure 6.11 (right). To improve the performance of the RF classifier, a re-calibration activity is suggested for the predictions associated to high-TI levels [236]. However, the derivation of the optimal model is not in the scope of this work.

6.6 Applicability study

At first, the best performing model that, is built by using the as-designed data, is tested on the set of data from the FEMU simulations (as-installed status). Then, a hybrid training is attempted by including the dataset of healthy status condition from the as-installed structure dynamics into the training.

Finally, the potential implementation of anomaly detection algorithms is briefly studied. The unsupervised algorithms are used based on healthy data only from the as-installed system. The algorithm is instructed to recognize the normal behaviour of the structure and to spot anomalies when fed with simulated data associated to the damaged structure. A flowchart of the approach is presented in Figure 6.12.

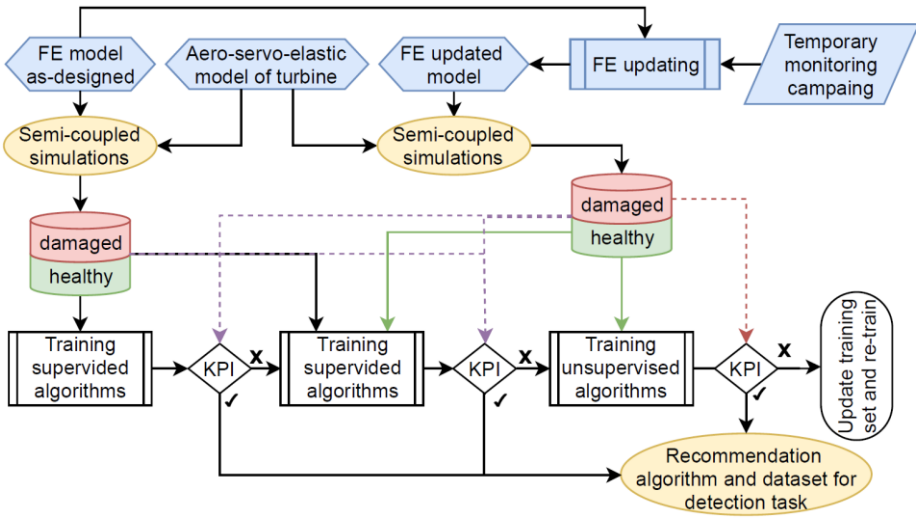


Figure 6.12: Workflow for the applicability of machine learning for the structural damage detection by using simulated data. The dashed lines are used to indicate the data flow for the testing of the algorithms.

6.6.1 Databases and preliminary analysis

The investigation of the applicability of this low-resolution, non-parametric approach, for the detection of the turbine's damaged status, continues with the setup and the findings of Section 6.5. The load combination and the number of simulations run per wind turbine status are reported in Table 6.5. The setup of the sensors' is varied according to Table 6.4. As for the previous analysis, the features for the training of the algorithms are the minimum, mean, maximum, and standard deviation statistics – usually collected in the SCADA data –, and the DEL of the tower bottom bending channels.

Table 6.5: Load combinations simulated per jacket status (healthy/damaged) and jacket model (as-designed/as-installed).

Load case	Wind speed	Wind direction	Wind-wave misalignment	Effective TI	Seeds	Total
Settings	(a) 14, 18, 22 m/s (b) 4, 6, 8 m/s	0°, 60°, 120°, 180°, 260°, 320°	-60°, 0°, 60°	TI_U, TI_{design} TI_L	6	–
Number	6	6	3	3	6	1,944

Of the total amount of simulation run, the dataset splits and the samples number are, hierarchically:

- 3,888 for each structural health condition (healthy and damaged), of which
- 1,944 for each represented status (as-designed and as-installed), of which
- 648 for each of the three TI levels, of which
- 108 for each of the six seedings.

The deployment of each dataset for the training and testing of the detection algorithms, is reported in Table 6.6. Details are given on the use of each dataset for the supervised and the unsupervised training approaches, according to the varying TI levels.

Table 6.6: Overview of the data type and their utilisation for the training and testing of the supervised and unsupervised algorithms.

Status→	Healthy			Damaged		
Model↓	**TI_L**	**TI_{design}**	**TI_U**	**TI_L**	**TI_{design}**	**TI_U**
As-designed	S: train/test SH: train/test	S: test SH: test	S: train/test SH: train/test	S: train/test SH: train/test U: test	S: test SH: test U: test	S: train/test SH: train/test U: test
As-installed	S: train/test SH: train/test U: train/test	S: test SH: test U: train/test	S: train/test SH: train/test U: train/test	S: test U: test	S: test SH: test U: test	S: test U: test

U: unsupervised algorithms

"train/test" refer to a training vs testing seed ratio of 4:2

6.6.2 Results

The algorithms to be trained on the each of the datasets of Table 6.6 are:

- the SVM and the RF, for supervised approach, based on the most promising classifiers identified in Section 6.5,
- a one-class support vector machine (OCSVM), for the unsupervised approach; the selection of this algorithm is guided by considerations on characteristics of the datasets generated by the as-installed FEM, and the typology of the detection wanted – explained more in details in Section 6.6.3.

For the training of these algorithms a "grid search" method is followed [223]. This implies that various combinations of configurations of the hyperparameters are tried during the training phase. For the supervised approaches, the best set is then selected as the one giving the best accuracy on sub-sets of the training dataset, so-called cross-validation (CV). On the other hand, the hyperparameters selection is done manually for the unsupervised algorithm, targeting satisfactory results of *pr*, *recall* and *fo* on both the training and the test sets – refer to Section C.2. of the Appendices for the definition of these metrics.

The hyperparameters are varied as follows [233]:

- SVM: kernel = {linear, polynomial, radial-based function (rbf)}, C = {0.01:1000}, gamma = {0.01:1000}
- RF: number of estimators = {10:100}, node split criterium = {"gini", "entropy"}, maximum number of features = {"sqrt", "log2", "dataset features"}, maximum depth = {5:30}
- OCSVM: kernel = {sigmoid, polynomial, rbf}, nu = {0.01:1}, gamma = {0.01:1000}

The classifiers are trained on balanced datasets containing both the healthy and damage samples. In contrast, the anomaly detection algorithm is instructed to build a decision function for recognizing the healthy data only and considering all data samples do not experience during the training phase as damaged.

The work of Section 6.5 addressed the training, testing and validation of supervised detection algorithms on the as-designed model. The SVM and the RF model showed satisfactory performances in predicting the correct label, healthy or damaged, on unseen sets of data associated to:

- the variation of the level of TI: upper bound (T_U), 90th percentile (T_d), lower bound (T_L),

- the stochasticity (seedings) of the wind and wave loadings, and
- the presence of a damage in the jacket structure.

The best set of data for the generalization of the prediction over a broad set of inflow wind conditions has been found by including the extreme (upper and lower bound) turbulence intensity levels per wind speed. The optimal sensor setup for the detection task has been identified by substituting the tower top accelerometer with a 2D inclinometer installed at the bottom of the turbine tower (set S1).

Testing on the as installed structure

The applicability of these classification models – trained on the as-designed model –, to datasets of the as-installed structure is verified in this section. Table 6.7 reports the training and testing performances for the below rated operating condition, by having obtained similar but slightly worse results for the above rated condition. It can be observed that both classifiers have a random guess on test sets derived from the as-installed model, for all levels of turbulence intensity. The inclusion of the DEL from the tower bottom strain gauge – having the potential to a good predictor of this type of failure mechanism [237] –, to the training set (set S2) leads only to minor improvements.

Table 6.7: Performances of the classification models – trained on the datasets generated by the as-designed FEM – on the variation of the EOC for the as-designed FEM, and on the dataset generated by the as-installed FEM.

			As-designed									As-installed		
Models and tuned hyperparameters		**Set**	**CV**	**Seedings**				**TI_d**				**TI_d**	**TI_U**	**TI_L**
			acc	***acc***	***pr***	***re***	***fo***	***acc***	***pr***	***re***	***fo***	***acc***	***acc***	***acc***
SVM	kernel: rbf C = 1000 gamma: scale	S1	93.3	100	100	100	0	84.0	81.3	88.3	20.4	51.9	50.0	53.7
		S2	91.4	100	100	100	0	83.8	83.5	84.3	16.7	50.9	50.5	54.6
RF	criterion: entropy max depth = 15 estimators = 80	S1	95.6	100	100	100	0	91.7	94.1	88.9	5.6	52.8	55.6	49.5
	criterion: entropy max depth = 20 estimators = 100	S2	95.4	100	100	100	0	89.8	93.9	85.2	5.6	53.2	54.6	49.5

Hybrid training on as-designed and as-installed

Consequently, a hybrid training approach is pursued, providing the classifiers with the datasets of the loadings and responses of the as-installed structure in its healthy status. This approach could close the gap for the real application of the detection algorithms, by screening at first the potential detectability of the structural damages on the as-designed model of the structure, to then extend the training set with the data collected from the real structure. The results are reported in Table 6.8, for below rated cases only.

It can be noticed that overall unsatisfactory performances on the as-installed dataset are obtained for both investigated supervised algorithms. Nonetheless, it seems that some improvements could be further obtained for the SVM, when trained on the sensor setup S1. Despite of the low *acc* associated to the low *re* (only about 10%), this classification model raises a low percentage of false alarm – see *fo* of about 6% –, while detecting a promising number of damaged cases per number of alarms – see *pr* of about 65%.

Table 6.8: Performances of the classification models when trained on the datasets generated by the as-designed FEM, and the healthy-status data generated by the as-installed FEM, by including the T_L and T_U conditions.

Models and tuned hyperparameters		Set	CV	Seed.	As-designed at T_d				As-installed at T_d			
			acc	*acc*	*acc*	*pr*	*re*	*fo*	*acc*	*pr*	*re*	*fo*
SVM	kernel: rbf C = 1000 gamma: 0.1	S1	70.7	100	70.8	80.0	55.6	13.9	52.3	64.7	10.2	5.6
		S2	71.5	100	86.1	89.0	82.4	10.2	50.0	0	0	0
RF	criterion: gini max depth = 20 estimators = 90	S1	67.1	100	67.6	88.0	40.7	5.6	50.0	0	0	0
	criterion: gini max depth = 15 estimators = 100	S2	63.7	100	82.4	85.0	78.7	13.9	50.0	0	0	0

6.6.3 Discussion of alternative approach for anomaly detection

The classification approach trained on the as-designed data failed to identify healthy and damaged status on datasets from the as-installed condition. An adapted training set including the variety in dynamics from the as-installed healthy condition seems to have the potential to overcome this issue. For what concerns the SVM classification, some dataset balancing techniques should be applied during the training phase to improve the damage detection performances. In addition, it should be noted that the difference of the as-designed and as-installed systems (Δ_1) is likely to be greater than the uncertainty that the detection approach will face if applied to reality, cp. Figure 6.1 ($\Delta_2 \ll \Delta_1$) when the algorithms are trained on the simulated data from the FE-updated model. Nonetheless, questions towards the capability of this approach to sufficiently accommodate uncertainty remain.

As further study for demonstration of its applicability, the number of samples for the training of these algorithms should be extended with filtering for the conditions leading to misclassification.

Unsupervised alternatives

Although the applicability of a supervised method has not been fully disproved yet, an anomaly detection approach on the as-installed datasets is attempted as an alternative solution for the detection task. Several algorithms can be applied for this purpose, from a simple principal component analysis (PCA) applied in a semi-supervised manner to more complex normal behaviour models based on deep learning methods.

To guide the selection of a suitable approach, a normal PCA is conducted to understand and visualize the distribution of the healthy and damaged datasets for the as-installed structure. In Figure 6.13 a low separability of the healthy and damaged data from the PCA can be observed. This indicates the need for a method which does not apply linear separation techniques (e.g., linear SVM, PCA, etc.), but that either can map the data to a higher dimensional space (i.e., via the use of kernel transformations) or can handle non-linearities (e.g., tree-based algorithms and neural networks).

For the attempt presented, the OCSVM algorithm is selected, due to its advantage in learning a decision function. Thus, as opposed to other normal behaviour models, either semi-supervised PCA or neural network models, it does not require an arbitrary selection

of a threshold. In addition, preliminary tests showed superior results of OCSVM when compared to other algorithms which aim to isolate the anomalies (such as isolation forest and local outlier factor). This behaviour was excepted because of the good compatibility of the datasets to the SVC and the observed low separability between the two classes.

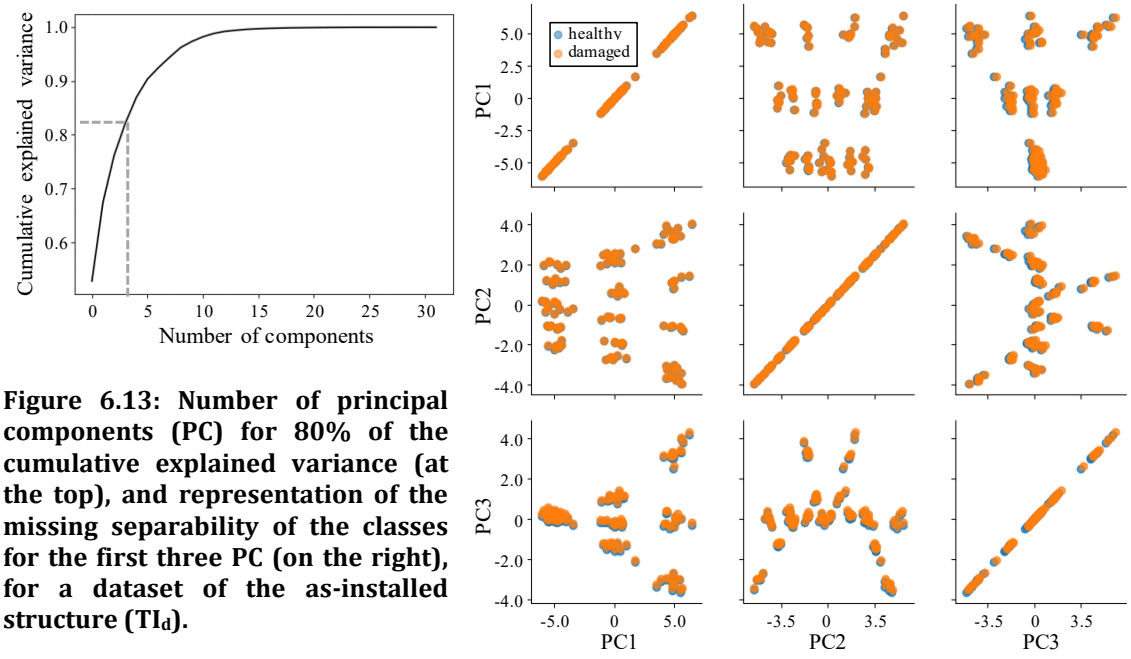


Figure 6.13: Number of principal components (PC) for 80% of the cumulative explained variance (at the top), and representation of the missing separability of the classes for the first three PC (on the right), for a dataset of the as-installed structure (TI_d).

OCSVM detection

The results from the unsupervised detection are shown in Table 6.9 for the hyperparameters combination leading to the best performances on the dataset of the as-installed structure in the healthy status. It is observed that the trained algorithm makes correct decisions only on the data for stochastic variations of the environmental conditions, with a relatively small deviation from the data experienced in the training phase. Thus, it is unable to generalize the prediction for different levels of turbulence intensity. For this reason, all the datasets from the healthy as-installed condition (cf. Table 6.6) had to be fed in during the training. Nonetheless, this approach shows potentially satisfactory results even when the algorithm is instructed on the information the tower bottom inclinometer only (set S3). Though, the more complete dataset - with additional statistics from the other SCADA signals - leads to improvements in the learned decision function. Task of a future analysis is the identification of the best set of signals, dependent on their reliability and availability.

Applicability of anomaly detection

The main criterion for the selection of the best tuning parameters is the successful fitting of the model to the healthy conditions in the training set, by targeting thus a low number of false alarms. Yet, the testing on the unseen damaged cases is necessary for the selection of a model with satisfactory detection performances. Therefore, even if the data from the normal operating turbine are available for the training task, the FE-updated model of the jacket structure should be setup for the collection of the reference damage state database. Future studies could investigate whether damage data statistics from the virgin model are

sufficient for this purpose, or if the variation of the modal properties between as-installed and as-designed systems is too big for this assumption.

Additionally, the high value of the kernel function parameter (*gamma*) and of the finetuning parameter (*nu*) are indicators of potential overfitting of the model to the dataset. Thus, the ability of the model to accommodate other slight variations in the environmental conditions should be tested as well.

Table 6.9: Performances of the OCSVM model – trained on the healthy-status data generated by the as-installed FEM – on the training dataset (Tr), on the testing dataset (Te) to identify the damaged-status, and on the variation of the EOC conditions (seeds). Finally, verification of its prediction on the datasets generated by the as-designed model.

Set	Hyperparameter			As-installed						As-designed			
				Tr	Te	Seedings							
	nu	kernel	*gamma*	*fa*	*re*	*acc*	*pr*	*re*	*fo*	*acc*	*pr*	*re*	*fo*
S1	0.04	rbf	100-1000	7	100	96-95	92-93	100	8-9	50	50	100	100
S3	0.06	rbf	1000	16	100	91	85	100	17	50	50	100	100

6.7 Conclusions

The analysis carried out in this chapter investigated the feasibility and applicability of supervised and unsupervised learning techniques for the detection of a damage in an offshore wind jacket structure. Differently from the approaches that utilise the high-frequency vibrational data, the methodology suggested employs only low-frequency aggregates from the standard collected SCADA data and, if necessary, few additional SHM sensors. A brace full-integrity loss is the test case for this experiment. However, it should be noticed that the analysis proposed is extendable to any other type of failure mechanism, once the set of environmental and operating condition, and the main indicators (i.e., predictors) for its detection are identified.

The supervised models are showed to be successful on the prediction of the brace integrity based on data simulated from the model of the as-designed structure. The trained classifiers are tested on the datasets derived from to the model of the as-installed structure, to verify their ability to cope with uncertainties related to the updated model dynamics. While the random forest-based model shows overall unsatisfactory results, the support vector model has room for improvements. For instance, a pre-processing data filtering could be put in place to perform the detection task only on the set of environmental and operating conditions with the smallest number of misclassified cases.

To give an alternative to these supervised methods, the feasibility of a novelty detection approach, based on only healthy status data, is verified on the database of simulated data from the as-installed structure. Among the algorithms of common use for this task, a one class support vector-based model is selected for its suitability to the characteristics of the as-installed datasets – i.e., having little separability between healthy and damaged status. Although the number of damage case detected and false alarm raised are acceptable, this algorithm can only perform the detection for small variations of the training set, being unable to correctly generalise the prediction on several levels on turbulence intensity. Nonetheless, its ability to cope with a good prediction based on the minimum set of signals – consisting of the inclinometer statistics only –, make it interesting for a real application where a high availability and reliability of the data is not always possible.

Chapter 7

Discussions

The goal of this thesis is to investigate the feasibility and applicability of several diagnostics methods for the condition monitoring of offshore wind turbines when only a limited amount of run-to-failure data is available. The scope is related to the study of data-driven and model-based approaches for the development and use of the digital twin technology.

This investigation is inherently a multidisciplinary task, as multidisciplinary is the dynamics of the offshore wind turbine systems. To include the different aspects of the monitoring of the health status of the drivetrain components and the structural elements of the foundations, the research has been divided into topics relative to the condition monitoring of offshore wind drivetrains, or the structural health monitoring of offshore wind jacket substructures SHMS studies. This separation in reflected in the sections of this discussion chapter.

7.1 Literature review

Chapter 2 analyses the RAM statistics from the scientific literature, collected from single initiatives and summary reports. As the type of data collected is heterogeneous, a terminology is introduced for describing and classifying the findings from the several databases. Standardised definitions of reliability, maintainability, availability, and performance indicators are summarised and are used to describe the data repositories throughout the analysis. The RDS-PP® designation is adopted to establish a uniform and easy comparison of the results among the several datasets. The quality of the collected statistics is discussed based on the type of records used for assessing failure of the turbine. Trends are identified from the comparison of the RAM statistics, based on the design and deployment parameters of the wind turbines. Finally, the RAM figures of offshore wind assets are discussed, by additionally simulating the operational availability for a set of reference offshore wind farms. This helped to quantify the impact of statistical uncertainties and unveil the challenges for the offshore wind industry in the collection of representative databases.

From these analyses it is possible to speculate on the potentially high criticality components for the next generation of offshore wind turbines. Moving towards larger systems and/or direct drive designs, there could be a reduction, or at least a steady trend, in the failure of the drivetrain and rotor systems, and of power generation systems, when

switching to a synchronous (permanent magnet) generator type. On the other hand, higher costs for corrective maintenance should be expected for the failures of the transmission systems and the tower structures. This should push research in looking for improvements at the design stage and/or for the implementation of monitoring systems on these assets.

7.2 Condition monitoring case studies

Chapter 3 introduced to case studies for employment of a model-based and a data-driven approaches for the condition monitoring of offshore wind turbines' gearboxes. More in general, these works perform some experiments and comparative studies in the attempt to address some of the challenges of a damage detection framework for offshore wind drivetrains:

- The excessive computational effort required for the holistic representation of the dynamics of a wind turbine, or even the whole wind farm, for the purpose of applying the digital twin technology to the drivetrain components, to support their condition monitoring.
- The limited access to a set of representative failure data for the new drivetrain technologies, and for most of the turbines in the offshore wind farm.

7.2.1 Investigation of model-order reduction: towards a digital twin targeting the monitoring of a gearbox failure mode

The development of a digital twin model of a wind farm by simply using and linking several advanced, single wind turbine models of dynamics can be computationally prohibitive. To this end, the first case study developed in Section 3.3 aimed at creating a reduced-order model (ROM). This model has the requirement to be able to capture the relevant dynamics of the offshore wind turbine system for a specific failure, having a lower computational cost and therefore more easily scalable up to a wind farm level. First, an aero-hydro-servo-elastic (AHSE) model is setup to derive the time domain response of the wind turbine in all its degrees of freedom (DOFs). One of the most critical failure modes of the wind turbine gearbox is then selected, and its relevant DOF – i.e., the rotor torque signal - is identified. The set of load cases for which the gearbox failure is likely to occur is recognised.

A linearisation of the AHSE model is then carried out, to subsequently derive a number of ROMs of the full-order system. The high-frequency states are excluded by using the modal truncation (MT) method. The load cases are simulated, and the torque signal is monitored. The results from the linear ROMs showed that the blade modes are important to capture not only extreme values of the DOF, but also its high-frequency responses (above 1.5 Hz). The results from nonlinear ROMs showed that when eliminating all the tower modes (rigid tower) the DOF is satisfactorily captured at low-frequency response (below 0.5 Hz), having almost the same spectral responses of the full-order nonlinear model.

7.2.2 Data-driven digital twin for the detection of gearbox anomalies on small datasets

Breaking the curse of small data is yet one of the major challenges of the machine learning models deployed for prognosis and diagnosis. When aiming to build detection models for the health monitoring of the offshore wind system, this issue emerges: (i) for the monitoring

newly installed assets, and (ii) when reducing the number of sensors installed for limiting the investment on the CMS.

To provide a possible solution to these applications, the aim of the case study in Section 3.4 is to discuss how the size of a dataset impacts the conventional machine learning methods, and to mitigate this issue by applying a multi-task learning approach with hard parameter sharing. Two datasets, consisting of SCADA and CMS data, are used for this experiment. The data come from two operating, homogeneous offshore wind turbine; the dataset relative to one of the two turbines contains about 220 days of operation, while the other (from the other wind turbine) accessed to only ca. 55 days of operation. Both datasets contain the same share of failure data (about 19%).

Considering that the accuracy of conventional machine learning methods suffers from the availability of a limited amount of data, the use of a hard-parameter multi-task learning (MTL) is tested and demonstrated to be an advantageous alternative. The study compares the results from the MTL approach – based on a hard parameter sharing of the deep neural network (DNN) for the prediction of the features to the CNN for the identification of the gearbox status – to the ones from a conventional deep neural network model – which instead only focuses on the classification task. The results are discussed outlining the benefits of the MTL approach, but also recognising the necessary future works to prove its validity to future field applications.

7.3 Structural health monitoring case studies

Based on the tools, analyses, findings, and discussions reported in Chapter 4, the Chapter 5 and 6 investigated low-cost monitoring and diagnostic approaches. The solutions suggested make use of the data that is already required for operational purposes (i.e., SCADA), and/or data generally collected for the purpose of the fatigue monitoring (e.g., accelerations). These are following described, by grouping them according to the data type and the nature of the approach deployed (parametric based versus non-parametric detection):

- **Vibration-based diagnostics**
  The monitoring of the modal properties showed good capabilities to not only detect the presence of a failure, but also identify the mechanism and the location of the failure if combined with a database of damage scenarios – simulated via the digital twin technology –, and when supported by data-driven clustering algorithms.
- **Non-parametric diagnostics based low-resolution data**
  Detecting damages based on 10-minute SCADA statistics with machine learning approaches showed to be generally feasible. However, deviations of the system used for training from the system used for monitoring, i.e., simulation model and real structure, respectively, can significantly reduce the reliability of the detection. Normal behaviour modelling approaches showed to be slightly more robust to uncertainties than classification approaches.

7.3.1 Vibration-based diagnostics of failure scenarios

This work investigates the feasibility of a two-level detection, in terms of damage identification and location, in the jacket support structure of an offshore wind turbine. A two-level detection scheme is suggested in this chapter by basing the detection on the digital twin technology on a database of simulated modal properties of the structure for different

failure scenarios. The detection model identifies the correct anomaly based on three types of modal indicators, namely, natural frequency, the modal assurance criterion between mode shapes, and the modal flexibility variation. The supervised Fisher's linear discriminant analysis is applied to transform the modal indicators to maximize the separability of several scenarios. A fuzzy clustering algorithm is then trained to predict the membership of new data to each of the scenarios in the database. In a case study, extreme scour phenomena and jacket members' integrity loss are simulated, together with variations of the structural dynamics for environmental and operating conditions. Cross-validation is used to select the best hyperparameters, and the effectiveness of the clustering is validated with slight variations of the environmental conditions. The results prove that it is feasible to detect and locate the simulated scenarios via the global monitoring of an offshore wind jacket structure.

The results have proved the feasibility of the suggested approach for the detection and location of failure events in the jacket sub-structure of the offshore wind turbine in analysis. The monitoring strategy outlined in Figure 1 is achieved to the extent of the “Detection Model Training – based on the simulated data. This detection algorithm fulfils the criteria of (i) diagnostic capability, (ii) low-cost – as opposed to any other ad-hoc monitoring system and field inspections –, and (iii) transparency of reasoning process, as required for the industrial needs. As concerns the eventual use of this probabilistic models for decision making (iv) of maintenance actions, the fuzzy clustering method allows to judge the prediction for the membership of the data to all the possible classes. However, this will be not as easy to interpret for the real-time data and the raising of alarms. Instead, it should be considered to make some engineering judgment on the evolution of the predictions in time. The implementation for real-time field monitoring (v) requires, as a next step, to verify the accuracy of its predictions to a set of data from the real structure.

7.3.2 Detection of a structural damage based on a non-parametric approach on standard, low-resolution data

The analysis carried out in Chapter 6 prove the feasibility of an approach for the indirect monitoring of a structural failure through the low-frequency statistics of the operational data from the offshore wind turbine. The supervised algorithms are trained on ten-minute SCADA data, which are derived from simulating the turbine dynamics. Databases associated to the healthy-status jacket structure, and a full-integrity-loss of one of jacket substructure brace members are thus generated. The overfitting of the algorithms is controlled by applying a cross-validation approach and by extensively testing their performances on subsets of the data, for unforeseen variation of the channels’ statistics due to the stochastic representation of wind and wave loadings, and the effective turbulence intensity at the turbine.

It is observed that, although the tower top accelerometer can give indications on the presence of the structural damage, its signals are highly affected by variations in the environmental conditions, making the classification task harder. Acceptable performance in the accuracy and detection rate of tree-based classifiers are obtained mainly for below rated conditions. However, the dataset used for the training of the models must be extended with additional data samples for correctly classifying the structural integrity status through a wide range of turbulence intensity levels. Furthermore, the high number of false alarms recorded can be reduced either by prior information on the turbulence intensity level (installing a specific sensor for this purpose) or better, by replacing the information from

the tower top accelerometer with the one from an inclinometer, positioned at the tower base. In this manner, the random forest model significantly improves its detection skills throughout at all the operating conditions. Similarly, also a support vector model shows satisfactory results at below rated conditions.

It should be noted that the training of the supervised and the unsupervised models, for the damage detection, depends on either the availability of data associated to the damaged structure, or their simulation through a true digital twin model of the structure. Furthermore, a broad set of isolated ten-minute SCADA data is required for the off-line training and testing of the algorithms, to ensure satisfactory detection performances among the several operating and environmental conditions. The practical value of this approach for the detection of the status of an operational system can be achieved by extending analysis to multiple damaged conditions, either by increasing the number of labels classified, or the number of classifiers.

Chapter 8

Conclusions

8.1 Summary and key findings

Besides the monitoring of the fatigue, for the assessment of the remaining useful life (prognostics), the structural integrity management of an offshore wind turbine structure, and its component should account for a framework to detect the presence of potentially critical failures (diagnostics). The catastrophic consequences related to these failures are almost always avoided by the decisions taken in the design phase. For instance, the in-situ monitoring of critical welds and joints can be expected to be planned in the list of the scheduled maintenance tasks, for the preventive detection of cracks. However, these inspections might not be the more cost-effective solution to the problem, and their cost can considerably increase when special procedures are required, such as non-destructive testing, visual examination at height, rope access and subsea inspections.

Therefore, the scope of this thesis has been to propose, to prove the feasibility, and to discuss the field applicability, of low-cost monitoring frameworks, which deploy model-based and data-driven approaches for the detection of systems' failures and the identification of the presence of evolving failure mechanisms. Due to the lack of representative sets of run-to-failure data from the field measurements, the digital twin technology is extensively deployed to support the creation of the databases on the failure scenarios. These scenarios have not only been limited to the modelling of structural damages, like cracks in the steel structures, but they also represented any significant deviation of the system due to extreme site conditions, such as scour, corrosion, and marine growth phenomena.

The key takeaway of this research is that due to the combination of continuously evolving technologies – the bigger sizes and ratings –, together with the lack of extensive investments for the collection of valuable field data, the digital twin technology and innovative data-driven solutions must be investigated for the strategic planning of the offshore wind asset management.

8.2 Recommendations for future work

To the knowledge of the author, this thesis is the first of its kind. This research project has been dedicated to the development of frameworks for the monitoring and the diagnostics

of the offshore wind turbines' systems, when state-of-the-art approaches cannot be applied due to the lack of a sufficient set of run-to-failure data. Several issues have been investigated and presented throughout the chapters and as summarised in Chapter 7. However, much research is left to be done. Among the recommendations for future work are:

1. **Tackle the issue of statistical uncertainty of offshore wind failure statistics:**
   As far as the availability estimation of offshore wind farms is concerned, the discrepancy between the predicted results and reference values from the literature suggests that a higher level of detail is needed and should be fed into the tool for obtaining conclusive results. This analysis, complemented with a cost analysis, is fundamental to a wide range of stakeholders in the offshore wind industry to achieve improvements of the profitability targets of current and future projects. To tackle the statistical uncertainty associated with the input failure statistics, information on the variation in time and among the turbines in the array (subjected to varying environmental conditions) should be provided. Finally, reference availability statistics for a longer period should be collected and made available for validation.
2. **Validate the vibration-based detection model for measurement and parametrization uncertainty:**
   The challenges for the field application of the vibration-based detection approach for SHM will come from the statistical uncertainty of the measured data, associated with the OMA and the violation of the method's assumptions for the extraction of modal parameters [33], [34], [194]. When the aim is to track the evolution and deviation of the modal properties, it is necessary to verify whether the deviation due to the presence of an anomaly is bigger than scatter in the data. Some of the challenges of dealing with modal data extracted from field measurements – especially in the case of offshore wind structures – come from their scatter and fluctuation in time caused by complex loads and rotating masses.
   As a first step one could apply the detection algorithm on only the data from the idling turbine – as done in Chapter 5 –, when the excitations and the inoperability of the turbine have less of an impact on the methods for the extraction of the modal properties. Alternatively, the extracted modal properties can be pre-processed, by filtering mode shapes that do not satisfactorily match with the analytical modes of the FEMU. By setting a suitable threshold on the distance – either in terms of MAC value or as combination of frequency difference and MAC value – between the extracted and the analytical modes, some of the scatter in the data would be removed without losing important information for the detection algorithm.
   However, it must be noted that this filtering procedure, as well the lack of excitation, and thus poor OMA performance, can often lead to a lack of some of the required modes. In this respect, multiple detection models should be setup for adapting the prediction to the varying number of features available. This approach and the likely drop in accuracy caused by the removal of modes in the training phase must be yet investigated.
3. **Cope with the natural uncertainty of the offshore wind diagnostic approaches:**
   The natural uncertainty is an inherent property of the underlying system and is the result of both spatial and temporal heterogeneity. The model calibration procedures are aimed at closing this gap. However, when deploying the digital twin technology

for the modelling of the failure of mechanisms, and their detection, the following considerations should be regarded:

- the detection approach based on low-resolution data, which is based on data simulated from the AHSE code semi-coupled to the FEMU of the wind turbine foundations, carries some uncertainties: the coupling of the foundation dynamics via the superelement (reduced order model) innerly holds some small discrepancies with the substructure dynamics (until 10 Hz), with respect to the full-order model [208].
- the global damping of the structure is assumed to be the same as the one defined at the design phase, neglecting thus the possible effect of the structural failure simulated.

These modelling uncertainties are here acknowledged but are not tackled in this thesis. The analysis is here harmonized by the fact that the first assumption is extended to the healthy model as well. The assumption on the structural damping is judged acceptable, since the aim is to demonstrate the detection feasibility and to discuss towards its applicability at most, without targeting the setup of a test filed demonstrator.

Furthermore, the diagnostic problems (at any of the detection levels) have been set to detect mutually exclusive failure events, by applying a multi-class classification approach. This decision is justified by the relatively young age of the test case structure (WK64). In the first years of life of the offshore structure, such failures can be caused either by the misjudgement of the field conditions or by unexpected events. This method could be replaced by a multi-label setup in the long run, when the structure would be more likely to be affected by multiple failure events (e.g., by reaching the scour allowance, while corrosion progressively develops on the jacket braces).

4. **Investigate the potential and the limitations of data-driven transfer learning:**
   By experimenting with a data-driven approach for failure events detection in the gearbox of an offshore wind turbine, the multi-task learning concept has been tested on a case study making use of the data from two operating wind turbines. Although out of the scope of the analysis, whose main objective was to learn on multiple tasks to improve the prediction of the neural network model, this study has showed the potential of a transfer of the knowledge between the two homogenous assets.
   The transfer learning theory, which is fundamentally different from the multi-task learning, should be the focus of further research on offshore wind application with limited availability of representative data.

8.3 Contribution to knowledge

This research contributed to scientific knowledge in a way which is novel and scientifically sound, and by providing value to one or more groups of stakeholders. In Table 8.1, it is reported and demonstrated how the initial objectives are met, what is the novelty of the research conducted on their matter, and who are the potential stakeholders and how they could benefit from the findings of this thesis.

Table 8.1: Contribution to the knowledge of this thesis' research

Objective	Novelty	Scientific soundness	Value / Stakeholders
(1.a) Review and classify, in a comprehensive study, the reliability, availability, and maintainability (RAM) data from onshore and offshore wind turbines, to identify the systems holding the highest criticalities and trends based on deployment parameters and environmental conditions of offshore wind turbines.	- Comprehensive review of existing published data related to reliability, availability, and maintenance of the state-of-the-art turbine with a view to critically discuss commonalities and distinguish correlation aspects between offshore and onshore assets. - Investigation of the sensitivity of certain components and technologies to key environmental and design parameters, facilitating technology qualification of new alternatives. - Highly informative for the qualification of new alternatives in the next generation of wind turbines.	- Systematic literature review based on exhaustive review of the literature and comparison with similar studies. - Transparent review process and presentation of the data.	- High level of disclosed information. - Relevant to the community of researchers and scientists approaching at first the issue reliability and availability increase for offshore wind structures. - Valuable for practitioners and operators to facilitate the technology qualification and/or maintenance strategies of new alternatives/generation wind turbines.
(1.b) Provide a comprehensive review on the numerical models for offshore wind turbines in an array.	- One of the first comprehensive review of this kind, for wind turbine and wind array models. - Classification of the numerical models for offshore wind turbines into the several areas of competency. - Analysis of the integration of these models for the array and review of the to date application. - Case based guidance for the model selection.	- Extensive literature review study showing the state-of-the-art of the models and their application. - Critical comparison of the numerical models for their area of competency.	Relevant to the community of researchers and scientists for guiding them in a critical selection of the most suitable numerical model of dynamics, suiting the different application.

Table 8.1: Contribution to the knowledge of this thesis' research

Objective	Novelty	Scientific soundness	Value / Stakeholders
(2) Investigate model-reduction techniques for a numerical model of wind turbine dynamics by targeting a drivetrain failure mode, to support the future development of digital twin technology for its remote monitoring.	- Investigation of the framework for model-order reduction for the purpose of damage detection. - Comparative analysis of the impact of different full-order model representation (linearised and non-linear) on the modal truncation effect on the targeted signal timeseries (rotor torque).	- Use of a well-know, and open-access, tool for the representation of the wind turbine dynamics (FAST v.8). - Systematic comparison and documentation of the effect of the modelling and reduction techniques on the rotor torque representation.	First step towards the developing digital twin models of the offshore wind farms with the purpose and focus on systems' monitoring and diagnostics.
(3) Experiment a data-driven approach for the detection of anomalies in a component of the drivetrain system with low availability of run-to-failure data.	- Testing of a multi-task learning model on the operating and the condition monitoring data from two wind turbine, one of which with limited data availability. - Comparative study of the prediction from this model against the one of a conventional neural network model.	- Use of state-of-the-art supervised learning algorithms for detection problem. - Transparent process in the preparation/analysis of the data.	Relevant to researchers and operators to make most of the use with the data currently collected and thus booster the information contained in the generally small run-to-failure databases.
(4) Demonstrate the feasibility of vibration-based detection of several damage scenarios in the substructure of an offshore wind turbine, with the support of the digital twin technology, and develop a framework for the future field application.	- Definition of the framework for the diagnostics of failure events in the substructure of an offshore wind turbine. - Use of the digital twin technology to generate a representative dataset of failure data on which train the detection algorithms. - Setup of an unsupervised model for the identification of the presence of the damage and the of the most likely location and failure mechanism of involved.	- Use of state-of-the-art supervised learning algorithms for detection problem. - Use of state-of-the-art unsupervised learning algorithms for extraction modal properties from real data. - Transparent process in the preparation/analysis of the data.	- Guidance for researchers and operators on level and type of damage possible to detect with current set up of SHMS. - First step towards guiding operators to the decision-making process for the best sensor placement for detection purpose.

Table 8.1: Contribution to the knowledge of this thesis' research

Objective	Novelty	Scientific soundness	Value / Stakeholders
(5.a) Investigate the feasibility of the detection of a structural failure via low-resolution data, …	- Feasibility of damage identification based on low-frequency statistics of standard collected data (SCADA). - Novel application of classification-based algorithm for detection of the structural status of the offshore wind turbine. - Feasibility of detection for environmental uncertainties.	- Use of state-of-the-art learning supervised algorithms for a classification problem. - Transparent process in the preparation of the data and training/testing of the algorithms. - Ranking of the optimal models based on known performance indicators. - Validation of the models and their performances for a representative and broad selection of loadings on the structure.	- Guidance for researchers and operators on the application of a non-parametric approach based on low-resolution data for the detection of specific damages, by making the most of the currently collected information from the operating turbine. - Potential reduction of the number of sensors deployed, while improving quality of the data collected (making the data collection more informative on the targeted failure mechanism), as opposed to the current best practise techniques for damage detection.
(5.b) … and develop a framework for the future field application with the support of the digital twin technology.	- Applicability of the damage identification based on low-frequency statistics of standard collected data (SCADA). - Feasibility of detection for system uncertainties: verification of the applicability of algorithms trained on simulated data for the detection of damages on the real structure. - Suggestion of anomaly detection approach for real application to damage identification.	- Verification of detection task on a model of the structure having the modal properties validated against the real structure. - Use of state-of-the-art supervised and unsupervised learning algorithms for detection problem. - Validation of the models and their performances for a representative and broad selection of loadings on the structure.	

Bibliography

[1] GL Garrad Hassan, "A guide to UK offshore wind operations and maintenance," 2013.

[2] CEN/TC 319 - Maintenance, *EN 13306:2017: Maintenance - Maintenance terminology*. 2017.

[3] K. Fischer and D. Coronado, "Condition Monitoring of Wind Turbines: State of the Art, User Experience and Recommendations," 2015.

[4] Deutsches Institut für Normung, *DIN EN 13306: Maintenance - Maintenance Terminology*, vol. 02. 2018.

[5] S. A. Asmai, B. Hussin, and M. Mohd Yusof, "A framework of an intelligent maintenance prognosis tool," in *2nd International Conference on Computer Research and Development, ICCRD 2010*, 2010, no. January, pp. 241–245.

[6] Z. Zhao, J. Wu, T. Li, C. Sun, R. Yan, and X. Chen, "Challenges and Opportunities of AI-Enabled Monitoring, Diagnosis & Prognosis: A Review," *Chinese Journal of Mechanical Engineering*, vol. 34, no. 1, 2021.

[7] M. A. Schwabacher, "A survey of data-driven prognostics," in *InfoTech at Aerospace: Advancing Contemporary Aerospace Technologies and Their Integration*, 2005, vol. 2, no. May, pp. 887–891.

[8] H. Zhang, R. Kang, and M. Pecht, "A hybrid prognostics and health management approach for condition-based maintenance," in *IEEM 2009 - IEEE International Conference on Industrial Engineering and Engineering Management*, 2009, pp. 1165–1169.

[9] M. Barnes *et al.*, "Technology Drivers in Windfarm Asset Management," 2018.

[10] Z. Lin, D. Cevasco, and M. Collu, "Progress on the development of a holistic coupled model of dynamics for offshore wind farms, phase I: Aero-Hydro-Servo-Elastic Model, with drive train model, for a single wind turbine," in *Proceedings of the ASME 37th International Conference on Ocean, Offshore and Arctic Engineering (OMAE2018)*, 2018, pp. 1–9.

[11] D. S. Herd and D. Flynn, "Prognostic Health Management techniques for intelligent condition monitoring of offshore renewable generation assets," in *IET Conference Publications*, 2013, vol. 2013, no. 623 CP.

[12] L. Wright and S. Davidson, "How to tell the difference between a model and a digital twin," *Advanced Modeling and Simulation in Engineering Sciences*, 2020.

[13] D. J. Wagg, K. Worden, R. J. Barthorpe, and P. Gardner, "Digital Twins: State-of-The-Art and Future Directions for Modeling and Simulation in Engineering Dynamics Applications," *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering*, vol. 6, no. 3, pp. 1–17, 2020.

[14] K. Worden, E. J. Cross, R. J. Barthorpe, D. J. Wagg, and P. Gardner, "On Digital Twins, Mirrors, and Virtualizations: Frameworks for Model Verification and Validation," *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering*, vol. 6, no. 3, pp. 1–9, 2020.

[15] U. T. Tygesen, M. S. Jepsen, J. Vestermark, N. Dollerup, and A. Pedersen, "The true digital twin concept for fatigue re-assessment of marine structures," *Proceedings of the International Conference on Offshore Mechanics and Arctic Engineering - OMAE*,

vol. 1, no. June, 2018.

[16] U. T. Tygesen, K. Worden, T. Rogers, G. Manson, and E. J. Cross, "State-of-the-Art and Future Directions for Predictive Modelling of Offshore Structure Dynamics Using Machine Learning," in *Pakzad S. (eds) Dynamics of Civil Structures, Volume 2. Conference Proceedings of the Society for Experimental Mechanics Series*, Springer, Cham, 2019.

[17] D. Augustyn, "Towards offshore wind digital twins . Application to jacket substructures . Towards offshore wind digital twins Application to jacket substructures," Faculty of Engineering and Science, Aalborg University, 2021.

[18] S. Siedler *et al.*, "D4.2 Report on design validation of WTG jacket and substation jacket based on FE updating," 2019.

[19] J. Tautz-Weinert *et al.*, "D4.3 Report on lifetime extension potential and remaining uncertainty," 2020.

[20] C. Wendelborn *et al.*, "D4.4 Report on the implementation of low-cost monitoring algorithms," 2020.

[21] F. K. Moghadam and A. R. Nejad, "Online condition monitoring of floating wind turbines drivetrain by means of digital twin," *Mechanical Systems and Signal Processing*, vol. 162, no. April 2021, p. 108087, 2022.

[22] M. Martinez-Luengo, A. Kolios, and L. Wang, "Structural health monitoring of offshore wind turbines: A review through the Statistical Pattern Recognition Paradigm," *Renewable and Sustainable Energy Reviews*, vol. 64, pp. 91–105, 2016.

[23] J. Lian, O. Cai, X. Dong, Q. Jiang, and Y. Zhao, "Health monitoring and safety evaluation of the offshore wind turbine structure: A review and discussion of future development," *Sustainability (Switzerland)*, vol. 11, no. 2, pp. 1–29, 2019.

[24] M. Martinez-luengo, M. Shafiee, and A. Kolios, "Data management for structural integrity assessment of offshore wind turbine support structures : data cleansing and missing data imputation," *Ocean Engineering*, vol. 173, no. November 2018, pp. 867–883, 2019.

[25] O. Avci, O. Abdeljaber, S. Kiranyaz, M. Hussein, M. Gabbouj, and D. J. Inman, "A review of vibration-based damage detection in civil structures : From traditional methods to Machine Learning and Deep Learning applications," *Mechanical Systems and Signal Processing*, vol. 147, 2021.

[26] W. Weijtjens, T. Verbelen, E. Capello, and C. Devriendt, "Vibration based structural health monitoring of the substructures of five offshore wind turbines," *Procedia Engineering*, vol. 199, pp. 2294–2299, 2017.

[27] D. Tcherniak, S. Chauhan, and M. H. Hansen, "Applicability limits of operational modal analysis to operational wind turbines," in *Structural Dynamics and Renewable Energy*, vol. 1, no. July, 2011, pp. 317–327.

[28] E. Di Lorenzo, S. Manzato, J. Houben, F. Vanhollebeke, S. Goris, and B. Peeters, "Wind turbine gearbox dynamic characterization using operational modal analysis," in *Conference Proceedings of the Society for Experimental Mechanics Series*, 2014, vol. 7, no. motor 1, pp. 41–52.

[29] M. Martinez-Luengo and M. Shafiee, "Guidelines and cost-benefit analysis of the Structural Health Monitoring implementation in offshore wind turbine support structures," *Energies*, vol. 12, no. 6, pp. 1–26, 2019.

[30] M. Richmond, S. Siedler, M. Häckell, U. Smolka, and A. Kolios, "Impact of accelerometer placement on modal extraction of offshore wind structures," in *Ocean Marine and Arctic Engineering (OMAE)*, 2020.

[31] D. Tcherniak, C. E. Carcangiu, and M. Rossetti, "Application of OMA to operational wind turbine: Methods for cleaning up the Campbell diagram," *International Conference on Noise and Vibration Engineering 2012, ISMA 2012, including USD 2012: International Conference on Uncertainty in Structure Dynamics*, vol. 6, pp. 4435–4446, 2012.

[32] G. Oliveira, F. Magalhães, Á. Cunha, and E. Caetano, "Vibration-based damage detection in a wind turbine using 1 year of data," *Structural Control and Health Monitoring*, vol. 25, no. 11, pp. 1–22, 2018.

[33] X. Dong, J. Lian, M. Yang, and H. Wang, "Operational modal identification of offshore wind turbine structure based on modified stochastic subspace identification method considering harmonic interference," *Journal of Renewable and Sustainable Energy*, vol. 6, no. 3, 2014.

[34] X. Dong, J. Lian, H. Wang, T. Yu, and Y. Zhao, "Structural vibration monitoring and operational modal analysis of offshore wind turbine structure," *Ocean Engineering*, vol. 150, no. 92, pp. 280–297, 2018.

[35] D. Tcherniak and G. C. Larsen, "Application of OMA to an operating wind turbine: Now including vibration data from the blades," *5th International Operational Modal Analysis Conference, IOMAC 2013*, pp. 1–12, 2013.

[36] D. Cevasco, S. Koukoura, and A. J. Kolios, "Reliability, availability, maintainability data review for the identification of trends in offshore wind energy applications," *Renewable and Sustainable Energy Reviews*, vol. 136, 2021.

[37] K. Fischer *et al.*, "Exploring the Causes of Power-Converter Failure in Wind Turbines based on Comprehensive Field-Data and Damage Analysis," *Energies*, vol. 12, no. 4, 2019.

[38] H. Luo, "Physics-based data analysis for wind turbine condition monitoring," *Clean Energy*, vol. 1, no. 1, pp. 4–22, 2017.

[39] R. K. Pandit and D. Infield, "SCADA based nonparametric models for condition monitoring of a wind turbine," in *The 7th International Conference on Renewable Power Generation (RPG 2018)*, 2019, vol. 2019, no. 18, pp. 4723–4727.

[40] J. Tautz-Weinert and S. J. Watson, "Using SCADA data for wind turbine condition monitoring - A review," *IET Renewable Power Generation*, vol. 11, no. 4, pp. 382–394, 2016.

[41] A. Stetco *et al.*, "Machine learning methods for wind turbine condition monitoring: A review," *Renewable Energy*, vol. 133, pp. 620–635, 2019.

[42] A. R. Nejad, P. F. Odgaard, Z. Gao, and T. Moan, "A prognostic method for fault detection in wind turbine drivetrains," *Engineering Failure Analysis*, vol. 42, pp. 324–336, 2014.

[43] R. Herring, K. Dyer, F. Martin, and C. Ward, "The increasing importance of leading edge erosion and a review of existing protection solutions," *Renewable and Sustainable Energy Reviews*, vol. 115, no. September, 2019.

[44] E. Di Lorenzo, G. Petrone, S. Manzato, B. Peeters, W. Desmet, and F. Marulo, "Damage detection in wind turbine blades by using operational modal analysis," *Structural Health Monitoring*, vol. 15, no. 3, pp. 289–301, 2016.

[45] N. Dervilis *et al.*, "On damage diagnosis for a wind turbine blade using pattern recognition," *Journal of Sound and Vibration*, vol. 333, no. 6, pp. 1833–1850, 2014.

[46] K. Chandrasekhar, N. Stevanovic, E. J. Cross, N. Dervilis, and K. Worden, "Damage detection in operational wind turbine blades using a new approach based on machine learning," *Renewable Energy*, vol. 168, pp. 1249–1264, 2021.

[47] M. N. Scheu, L. Tremps, U. Smolka, A. Kolios, and F. Brennan, "A systematic Failure Mode Effects and Criticality Analysis for offshore wind turbine systems towards integrated condition based maintenance strategies," *Ocean Engineering*, vol. 176, no. October 2018, pp. 118–133, 2019.

[48] T. Rubert *et al.*, "Wind turbine lifetime extension decision-making based on structural health monitoring," *Renewable Energy*, vol. 143, no. May, pp. 611–621, 2019.

[49] W. Weijtjens, T. Verbelen, G. De Sitter, and C. Devriendt, "Foundation structural health monitoring of an offshore wind turbine—a full-scale case study," *Structural Health Monitoring*, vol. 15, no. 4, pp. 389–402, 2016.

[50] W.-H. Hu, S. Thöns, R. G. Rohrmann, S. Said, and W. Rücker, "Vibration-based structural health monitoring of a wind turbine system. Part I: Resonance phenomenon," *Engineering Structures*, vol. 89, pp. 260–272, 2015.

[51] C. U. Nguyen, S. Y. Lee, H. T. Kim, and J. T. Kim, "Vibration-based damage assessment in gravity-based wind turbine tower under various waves," *Shock and Vibration*, vol. 2019, 2019.

[52] C. Nguyen, T. Huynh, and J. Kim, "Vibration-based damage detection in wind turbine towers using artificial neural networks," *Structural Monitoring and Maintenance*, vol. 5, no. 4, pp. 507–519, 2018.

[53] M. Richmond, U. Smolka, and A. Kolios, "Feasibility for Damage Identification in Offshore Wind Jacket Structures through Monitoring of Global Structural Dynamics," *Energies*, vol. 13, no. 21, p. 5791, 2020.

[54] C. Devriendt, P. J. Jordaens, G. De Sitter, and P. Guillaume, "Damping estimation of an offshore wind turbine on a monopile foundation," in *European Wind Energy Conference and Exhibition 2012, EWEC 2012*, 2012, vol. 2, pp. 1112–1121.

[55] U. Smolka, D. Kaufer, and P. W. Cheng, "Are sea state measurements required for fatigue load monitoring of offshore wind turbines?," *Journal of Physics: Conference Series*, vol. 555, no. 1, pp. 0–10, 2014.

[56] U. Smolka and P. W. Cheng, "On the Design of Measurement Campaigns for Fatigue Life Monitoring of Offshore Wind Turbines," *The Twenty-third International Offshore and Polar Engineering Conference*, vol. 9, pp. 408–413, 2013.

[57] S. Calvert, P. Goldman, E. Demeo, C. Mcgowin, B. Smith, and K. Tromly, "The EPRI/DOE Wind Turbine Performance Verification Program," in *Presented at Solar Energy Forum*, 1997, no. April 25-30.

[58] S. Pfaffel, S. Faulstich, and K. Rohrig, "Performance and reliability of wind turbines: A review," *Energies*, vol. 10, no. 11, pp. 1–27, 2017.

[59] Y. Lin, L. Tu, H. Liu, and W. Li, "Fault analysis of wind turbines in China," *Renewable and Sustainable Energy Reviews*, vol. 55, pp. 482–490, 2016.

[60] M. D. Reder, E. Gonzalez, and J. J. Melero, "Wind Turbine Failures - Tackling current Problems in Failure Data Analysis," in *The Science of Making Torque from Wind (TORQUE 2016)*, 2016, vol. 753, no. 7.

[61] P. Tavner, *Offshore Wind Turbines: Reliability, Availability and Maintenance*. London, United Kingdom: The Institution of Engineering and Technology, 2012.

[62] J. Carroll, A. McDonald, and D. McMillan, "Failure rate, repair time and unscheduled O&M cost analysis of offshore wind turbines," *Wind Energy*, vol. 19, pp. 1107–1119, 2016.

[63] F. Spinato, P. J. Tavner, G. J. W. van Bussel, and E. Koutoulakos, "Reliability of wind

turbine subassemblies," *IET Renewable Power Generation*, vol. 3, no. 4, p. 387, 2009.

[64] J. Bueno Gayo, "Final Publishable Summary of Results of Project ReliaWind," no. 14 March, pp. 1–62, 2011.

[65] W. Musial, "Wind Turbine Gearbox Reliability," 2006.

[66] M. A. Rumsey and J. A. Paquette, "Structural health monitoring of wind turbine blades," p. 69330E, 2008.

[67] K. Branner and A. Ghadirian, "Database about blade faults," 2014.

[68] R. Lynette, "Status of the U.S. wind power industry," *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 27, pp. 327–336, 1988.

[69] G. J. W. Van Bussel and M. Zaaijer, "Estimation of Turbine Reliability Figures within the DOWEC Project," 2003.

[70] J. Ribrant, "Reliability performance and maintenance - A survey of failures in wind power systems," KTH Royal Institute of Technology, 2006.

[71] J. Ribrant and L. M. Bertling, "Survey of failures in wind power systems with focus on Swedish wind power plants during 1997-2005," *IEEE Transactions on Energy Conversion*, vol. 22, no. 1, pp. 167–173, 2007.

[72] P. J. Tavner, G. J. W. Van Bussel, and F. Spinato, "Machine and Converter Reliabilities in Wind Turbines," in *International Conference on Power Electronics, Machines and Drives*, 2006, no. April, pp. 1–4.

[73] P. J. Tavner, J. Xiang, and F. Spinato, "Reliability analysis for wind turbines," *Wind Energy*, vol. 10, no. 1, pp. 1–18, 2007.

[74] P. J. Tavner and F. Spinato, "Reliability of different wind turbine concepts with relevance to offshore application," in *European Wind Energy Conference 2008*, 2008, no. April.

[75] P. J. Tavner, D. M. Greenwood, M. W. G. Whittle, R. Gindele, S. Faulstich, and B. Hahn, "Study of effects of weather and location on wind turbine failure rates," in *European Wind Energy Conference (EWEC 2010)*, 2010, no. April 20-23.

[76] P. Tavner, C. Edwards, A. Brinkman, and F. Spinato, "Influence of Wind Speed on Wind Turbine Reliability," *Wind Engineering*, vol. 30, no. 1, pp. 55–72, 2006.

[77] P. J. Tavner, D. M. Greenwood, M. W. G. Whittle, R. Gindele, S. Faulstich, and B. Hahn, "Study of weather and location effects on wind turbine failure rates," *Wind Energy*, vol. 16, no. May 2012, pp. 175–187, 2013.

[78] S. Faulstich, B. Hahn, and P. J. Tavner, "Wind turbine downtime and its importance for offshore deployment," *Wind Energy*, vol. 14, no. July 2010, pp. 327–337, 2011.

[79] S. Faulstich, "Component reliability ranking with respect to WT concept and external environmental conditions," Kassel, Germany, 2010.

[80] S. Faulstich, P. Lyding, and P. J. Tavner, "Effects of Wind Speed on Wind Turbine Availability," *Proceedings of European Wind Energy Conference & Exhibition 2011 (EWEC 2011)*, 2011.

[81] B. Hahn, "Zeitlicher Zusammenhang von Schadenshäufigkeit und Windgeschwindigkeit," Kassel, Germany, 1997.

[82] H. Jung, "Erhöhung der Verfügbarkeit von Windenergieanlagen EVW-Phase 2," p. 170, 2014.

[83] IZP, Fraunhofer IWES, SAG GmbH, and ENERTRAG AG, "Abschlussbericht für das Verbundprojekt 'Erhöhung der Verfügbarkeit von Windkraftanlagen,'" 2009.

[84] C. Su, Y. Yang, X. Wang, and Z. Hu, "Failures analysis of wind turbines: Case study of a Chinese wind farm," in *Proceedings of 2016 Prognostics and System Health Management Conference*, 2017, pp. 1–6.

[85] G. M. J. Herbert, S. Iniyan, and R. Goic, "Performance, reliability and failure analysis of wind farm in a developing Country," *Renewable Energy*, vol. 35, no. 12, pp. 2739–2751, 2010.

[86] Z. Ma, G. An, X. Sun, and J. Chai, "A study of fault statistical analysis and maintenance policy of wind turbine system," in *International Conference on Renewable Power Generation (RPG)*, 2015, p. 4.

[87] C. Carter, B. Karlson, S. Martin, and C. Westergaard, "Continuous Reliability Enhancement for Wind (CREW) Program Update," 2016.

[88] Y. Feng, P. J. Tavner, and H. Long, "Early experiences with UK Round 1 offshore wind farms," *Proceedings of the Institution of Civil Engineers : Eergy*, vol. 163, no. 4, pp. 167–181, 2010.

[89] C. J. Crabtree, D. Zappalá, and S. I. Hogg, "Wind energy: UK experiences and offshore operational challenges," *Proceedings of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy*, vol. 229, no. 7, pp. 727–746, 2015.

[90] NoordzeeWind, "Operations Report 2007-2009," 2008.

[91] A. Fox and C. Hill, "System Performance, Availability and Reliability Trend Analysis: Portfolio Review 2016," 2017.

[92] V. Berkhout *et al.*, "Wind Energy Report Germany 2013," 2013.

[93] G. J. W. van Bussel and M. B. Zaaijer, "Reliability, Availability and Maintenance aspects of large-scale offshore wind farms, a concepts study," in *International Conference on Marine Renewable Energy (MAREC 2001)*, 2001.

[94] G. J. W. van Bussel and M. B. Zaaijer, "DOWEC Concepts Study, Reliability, Availability and Maintenance Aspects," in *Proceedings of the European Wind Energy Conference*, 2001, pp. 557–560.

[95] A. P. W. M. Curvers and L. W. M. M. Rademakers, "RECOFF: Operation and Maintenance Task 1 : Standardisation of collecting failures and maintenance data," 2004.

[96] J. A. Andrawus, J. Watson, and M. Kishk, "Wind Turbine Maintenance Optimisation: principles of quantitative maintenance optimisation," *Wind Engineering*, vol. 31, no. 2, pp. 101–110, 2009.

[97] Z. Hameed, J. Vatn, and J. Heggset, "Challenges in the reliability and maintainability data collection for offshore wind turbines," *Renewable Energy*, vol. 36, no. 8, pp. 2154–2165, 2011.

[98] B. Hahn *et al.*, "Recommended practices for wind farm data collection and reliability assessment for O&M optimization," in *14th Deep Sea Offshore Wind R&D Conference, EERA DeepWind'2017*, 2017, vol. 137, pp. 358–365.

[99] CATAPULT Offshore Renewable Energy, "SPARTA: A UK offshore wind performance data benchmarking platform, transforming wind turbine operational performance," 2018. [Online]. Available: https://ore.catapult.org.uk/casestudies/sparta/. [Accessed: 10-Apr-2018].

[100] Fraunhofer IEE and Dresden IZP, "WInD-Pool (Windenergie-Informations-Datenpool)," 2018. [Online]. Available: https://wind-pool.iee.fraunhofer.de/wind_pool_de/WInD-Pool/. [Accessed: 10-Apr-2018].

[101] S. Faulstich, S. Pfaffel, P. Kühn, and P. Lyding, "Monitoring offshore wind energy use

in Europe - Offshore~WMEP," *Energy Procedia*, vol. 24, no. January, pp. 322–327, 2012.

[102] I. W. Es and I. Zp, "ABSCHLUSSBERICHT OFFSHORE~WMEP: Windenergieforschung am Offshore-Testfeld (WIFO) – Monitoring der Offshore-Windenergienutzung in Deutschland – Konzipierungsphase," 2012.

[103] S. Pfaffel, S. Faulstich, B. Hahn, J. Hirsch, V. Berkhout, and J. Harald, "OFFSHORE~WMEP: Wissenschaftliches Monitoring- und Evaluierungsprogramm zur Offshore-Windenergienutzung – Durchführungsphase I," 2016.

[104] S. (Shawn) Sheng and W. Yang, "Wind Turbine Drivetrain Condition Monitoring - An Overview," 2013.

[105] E. Artigao, S. Martín-Martínez, A. Honrubia-Escribano, and E. Gómez-Lázaro, "Wind turbine reliability: A comprehensive review towards effective condition monitoring development," *Applied Energy*, vol. 228, no. February, pp. 1569–1583, 2018.

[106] C. Dao, B. Kazemtabrizi, and C. Crabtree, "Wind turbine reliability data review and impacts on levelised cost of energy," *Wind Energy*, vol. 22, no. 12, pp. 1848–1871, 2019.

[107] K. Leahy, C. Gallagher, P. O'Donovan, and D. T. J. O'Sullivan, "Issues with data quality for wind turbine condition monitoring and reliability analyses," *Energies*, vol. 12, no. 2, pp. 1–22, 2019.

[108] S. Sheng and R. O'Connor, *Reliability of Wind Turbines - Chapter 15*. Elsevier Inc., 2017.

[109] M. Leimeister and A. Kolios, "A review of reliability-based methods for risk analysis and their application in the offshore wind industry," *Renewable and Sustainable Energy Reviews*, vol. 91, no. April, pp. 1065–1076, 2018.

[110] N. Kumar, D. Rogers, T. D. Burnett, E. Sullivan, and M. Gascon, "Reliability, Availability, Maintainability (RAM) for Wind Turbines," in *Proceedings of the ASME 2017 Power Conference Joint With ICOPE-17*, 2017, pp. 1–6.

[111] A. Kolios and M. Martinez-luengo, "The end of the line for today ' s wind turbines," *Renewable Energy Focus*, vol. 17, no. 3, pp. 109–111, 2016.

[112] A. Barabadi, O. Tobias, and J. Barabady, "RAMS data collection under Arctic conditions," *Reliability Engineering and System Safety*, vol. 135, pp. 92–99, 2015.

[113] J. M. Pinar Pérez, F. P. García Márquez, A. Tobias, and M. Papaelias, "Wind turbine reliability analysis," *Renewable and Sustainable Energy Reviews*, vol. 23, pp. 463–472, 2013.

[114] Z. Q. Zhu and J. Hu, "Electrical machines and power-electronic systems for high-power wind energy generation applications. Part I – market penetration, current technology and advanced machine systems," *The International Journal for Computation and Mathematics in Electrical and Electronic Engineering*, vol. 32, no. 1, pp. 7–33, 2013.

[115] S. Faulstich, B. Hahn, P. Lyding, and P. J. Tavner, "Reliability of offshore turbines – identifying risks by onshore experience," in *European Offshore Wind 2009 Conference*, 2009, pp. 1–10.

[116] K. Harman, R. Walker, and M. Wilkinson, "Availability trends observed at operational wind farms," in *European Wind Energy Conference (EWEC 2008)*, 2008, p. 8.

[117] S. Faulstich, M. Durstewitz, B. Hahn, K. Knorr, and K. Rohrig, "Windenergy Report Germany 2008: written within the research project 'Deutscher Windmonitor,'" 2009.

[118] K. Ma, *Power Electronics for the Next Generation Wind Turbine System*, Volume 5. Aalborg, Denmark: Springer, 2015.

[119] D. Cevasco and M. Collu, "Dataset for European Installed Offshore Wind Turbines (until year end 2017)," 2018. [Online]. Available: https://doi.org/10.17862/cranfield.rd.6133673.

[120] C. Kaidis, B. Uzunoglu, and F. Amoiralis, "Wind turbine reliability estimation for different assemblies and failure severity categories," *IET Renewable Power Generation*, vol. 9, no. 8, pp. 892–899, 2015.

[121] Suomen Tuulivoimayhdisty, "Wind power in Finland: Statistics." [Online]. Available: http://www.tuulivoimayhdistys.fi/en/wind-power-in-finland/statistics. [Accessed: 08-Apr-2018].

[122] F. Carlsson, E. Eriksson, and M. Dahlberg, "Damage preventing measures for wind turbines. Phase 1 - Reliability data," 2010.

[123] S. Sheng, "Report on Wind Turbine Subsystem Reliability - A Survey of Various Databases," 2013.

[124] A. Stenberg, "Analysis of wind turbine statistics in Finland," AALTO-UNIVERSITETET, 2010.

[125] "Wind Power Forecasting Sweden." [Online]. Available: http://vindstat.com/. [Accessed: 08-Apr-2018].

[126] F. Spinato, "The Reliability Wind of Turbines," Durham University, 2008.

[127] Hahn B, Durstewitz M, and Rohrig K., "Reliability of wind turbines: experiences of 15 years with 1,500 WTs," Kassel, Germany, 2006.

[128] E. Echavarria, B. Hahn, G. J. W. van Bussel, and T. Tomiyama, "Reliability of Wind Turbine Technology Through Time," *Journal of Solar Energy Engineering*, vol. 130, no. 3, p. 031005, 2008.

[129] S. Pfaffel *et al.*, "Wind Energy Report Germany 2011," Kassel, Germany, 2011.

[130] S. Faulstich, B. Hahn, J. H., K. Rafik, and A. Ringhandt, "Appropriate failure statistics and reliability characteristics," in *Proceedings of 9th German Wind Energy Conference, DEWEC 2008*, 2008, vol. 136, no. 1, pp. 23–42.

[131] C. Kaidis, "Wind turbine reliability prediction a SCADA data processing & reliability estimation tool," Uppsala University, 2013.

[132] J. Andrawus, "Maintenance optimisation for wind turbines," The Robert Gordon University, 2008.

[133] J. Carroll, A. McDonald, and D. McMillan, "Reliability Comparison of Wind Turbines With DFIG and PMG Drive Trains," *IEEE Transactions on Energy Conversion*, vol. 30, no. 2, pp. 663–670, 2015.

[134] E. Gonzalez, M. Reder, and J. J. Melero, "SCADA alarms processing for wind turbine component failure detection," in *The Science of Making Torque from Wind (TORQUE 2016)*, 2016, vol. 753, no. 7.

[135] M. Wilkinson *et al.*, "Methodology and Results of the Reliawind Reliability Field Study," in *European Wind Energy Conference (EWEC 2010)*, 2010, no. EWEC.

[136] V. A. Peters, A. B. Ogilvie, and C. R. Bond, "Continuous Reliability Enhancement for Wind (CREW) Database: Wind Plant Reliability Benchmark," 2012.

[137] V. Hines, A. Ogilvie, and C. Bond, "Continuous Reliability Enhancement for Wind (CREW) Database - Wind Plant Reliability Benchmark," Chicago, IL, USA, 2013.

[138] C. Bangalore, Pramod Boussion *et al.*, "IEA Wind TCP RP17. Expert group report on recommended practise: Wind farm data collection and reliabilty assessment for O&M optimisation," 2017.

[139] VGB PowerTech, "VGB-Standard RDS-PP Application specification, Part 32: Wind energy (VGB-S-823-T32;2012-04-EN)," 2012.

[140] A. Kolios, J. Walgern, S. Koukoura, R. Pandit, and J. Chiachio-Ruano, "openO&M: Robust O&M open access tool for improving operation and maintenance of offshore wind turbines," *Proceedings of the 29th European Safety and Reliability Conference (ESREL 2019)*, pp. 452–459, 2019.

[141] A. Pliego Marugán, F. P. García Márquez, and J. M. Pinar Pérez, "Optimal Maintenance Management of Offshore Wind Farms," *Energies*, pp. 1–20, 2016.

[142] G. Rinaldi, P. R. Thies, and L. Johanning, "A computational tool for the pro-active management of offshore wind farms," 2016.

[143] T. Bak *et al.*, "Baseline Layout and Design of a 0.8 GW Reference Wind Farm in the North Sea," *Wind Energy*, pp. 1665–1683, 2017.

[144] E. Koutoulakos, "Wind turbine reliability characteristics and offshore availability assessment," TU Delf, 2008.

[145] Z. Zhang, A. Matveev, S. Øvrebø, R. Nilssen, and A. Nysveen, "State of the art in generator technology for offshore wind energy conversion systems," in *2011 IEEE International Electric Machines & Drives Conference (IEMDC)*, 2011, pp. 1131–1136.

[146] J. Carroll, A. McDonald, I. Dinwoodie, D. McMillan, M. Revie, and I. Lazakis, "Availability, operation and maintenance costs of offshore wind turbines with different drive train configuration," *Wind Energy*, vol. 20, no. July 2016, pp. 361–378, 2017.

[147] G. Wilson, G. Ault, and D. McMillan, "Modelling the effects of the environment on wind turbine failure modes using neural networks," *International Conference on Sustainable Power Generation and Supply (SUPERGEN 2012)*, pp. 42–42, 2012.

[148] G. Wilson and D. McMillan, "Quantifying the impact of wind speed on wind turbine component failure rates," in *Eurpean Wind Energy Conference (EWEA) 2014*, 2014.

[149] J. Carroll and A. Mcdonald, "Drivetrain Availability in Offshore Wind Turbines," *EWEA 2014 Annual Event*, pp. 1–5, 2013.

[150] M. N. Scheu, A. Kolios, T. Fischer, and F. Brennan, "Influence of statistical uncertainty of component reliability estimations on offshore wind farm availability," *Reliability Engineering and System Safety*, vol. 168, no. May, pp. 28–39, 2017.

[151] A. Fox and C. Hill, "System Performance Availability and Reliability Trend Analysis – SPARTA: 2018/19 Portfolio Review," 2018.

[152] A. May, "Operational expenditure optimisation utilising condition monitoring for offshore wind parks," University of Strathclyde, 2016.

[153] A. Romero, Y. Lage, S. Soua, B. Wang, and T. H. Gan, "Vestas V90-3MW wind turbine gearbox health assessment using a vibration-based condition monitoring system," *Shock and Vibration*, no. July, pp. 1–18, 2016.

[154] H. Link *et al.*, "Gearbox Reliability Collaborative Project Report : Findings from Phase 1 and Phase 2 Testing," no. June, p. 85, 2011.

[155] J. Igba, K. Alemzadeh, C. Durugbo, and K. Henningsen, "Performance assessment of wind turbine gearboxes using in-service data : Current approaches and future trends," *Renewable and Sustainable Energy Reviews*, vol. 50, pp. 144–159, 2015.

[156] L. Ziegler, N. Cosack, A. Kolios, and M. Muskulus, "Structural monitoring for lifetime extension of offshore wind monopiles: Verification of strain-based load extrapolation algorithm," *Marine Structures*, vol. 66, no. March, pp. 154–163, 2019.

[157] A. J. Kolios, U. Smolka, and R. Wind, "Risk-based Maintenance Strategies for Offshore Wind Energy Assets," in *The 66th Annual Reliability & Maintainability Symposium (RAMS 2020)*.

[158] A. R. Nejad, Z. Gao, and T. Moan, "On long-term fatigue damage and reliability analysis of gears under wind loads in offshore wind turbine drivetrains," *International Journal of Fatigue*, vol. 61, pp. 116–128, 2014.

[159] D. Cevasco, M. 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 *Global Wind Summit (WindEurope) 2018*, 2018, vol. 1102, p. 12.

[160] I. T. H. Al-Iedani, "Model Order Reduction and Optimal Control of Wind Energy Conversion Systems Model Order Reduction and Optimal Control of Wind Energy Conversion," The State University of New Jersey, 2021.

[161] M. I. Friswell, S. D. Garvey, and J. E. T. Penny, "Model reduction using dynamic and iterated IRS techniques," *Journal of Sound and Vibration*, vol. 186, no. 2, pp. 311–323, 1995.

[162] F. Lemmer, W. Yu, H. Steinacker, D. Skandali, and S. Raach, "Advances on reduced-order modeling of floating offshore wind turbines," in *Proceedings of the International Conference on Offshore Mechanics and Arctic Engineering - OMAE*, 2021, vol. 9.

[163] M. M. Rezaei, M. Behzad, H. Haddadpour, and H. Moradi, "Development of a reduced order model for nonlinear analysis of the wind turbine blade dynamics," *Renewable Energy*, vol. 76, pp. 264–282, 2015.

[164] C. C. Flanigan, "Model reduction using Guyan, IRS, and Dynamic methods," *Proceedings of the International Modal Analysis Conference IMAC*, vol. 1, pp. 172–176, 1998.

[165] S. Ghosh and N. Senroy, "Balanced truncation based reduced order modeling of wind farm," *International Journal of Electrical Power and Energy Systems*, vol. 53, no. 1, pp. 649–655, 2013.

[166] R. J. GUYAN, "Reduction of stiffness and mass matrices," *AIAA Journal*, vol. 3, no. 2, pp. 380–380, 1965.

[167] J. C. O'Callahan, "A Procedure for an Improved Reduced System (IRS) Model," in *Proceedings of the 7th International Modal Analysis Conference*, 1989, pp. 17–21.

[168] S. Gugercin and A. C. Antoulas, "A survey of model reduction by balanced truncation and some new results," *International Journal of Control*, vol. 77, no. 8, pp. 748–766, 2004.

[169] A. C. Antoulas, "A survey of model reduction methods for large-scale systems," *American Mathematical Society*, vol. 280, pp. 1–28, 2006.

[170] B. Jonkman and J. Jonkman, "FAST v8.15.00a-bjj," 2016.

[171] WAMIT INC, "WAMIT: The state of the art in wave interaction analysis." [Online]. Available: https://www.wamit.com/. [Accessed: 25-Aug-2022].

[172] W. E. Cummins, "The Impulse Response Function and Ship Motions," 1962.

[173] J. M. Jonkman and B. J. Jonkman, "FAST modularization framework for wind turbine simulation: Full-system linearization," *Journal of Physics: Conference Series*, vol. 753, no. 8, 2016.

[174] Y. Lee, B. Seo, and E.-T. Lee, "Application of model reduction techniques to jacket structures," *International Journal of Steel Structures*, vol. 15, no. 1, pp. 1–6, 2015.

[175] C.-T. Chen, *Linear System Theory and Design*, Third. Oxford University Press, 1999.

[176] J. Jonkman, S. Butterfield, W. Musial, and G. Scott, "Definition of a 5-MW Reference Wind Turbine for Offshore System Development," 2009.

[177] X. Wang, P. Guo, and X. Huang, "A review of wind power forecasting models," in *ICSGCE 2011*, 2011, vol. 12, pp. 770–778.

[178] A. Salem, A. Abu-Siada, and S. Islam, "Application of Order Analysis to Diagnose Fatigue within Wind Turbine Gearbox," *Technol Econ Smart Grids Sustain Energy*, no. January 20th, pp. 3–7, 2017.

[179] D. Herr and D. Heidenreich, "Understanding the Root Causes fo Axial Cracking in Wind Turbine Gearbox Bearings," *issue 03*, pp. 38–45, Feb-2015.

[180] A. R. Nejad, Z. Gao, and T. Moan, "Fatigue reliability-based inspection and maintenance planning of gearbox components in wind turbine drivetrains," *Energy Procedia*, vol. 53, no. C, pp. 248–257, 2014.

[181] A. R. Nejad, Z. Gao, and T. Moan, "Long-term analysis of gear loads in fixed offshore wind turbines considering ultimate operational loadings," *Energy Procedia*, vol. 35, no. 1876, pp. 187–197, 2013.

[182] D. Dabrowski and A. Natarajan, "Assessment of gearbox operational loads and reliability under high mean wind speeds," *Energy Procedia*, vol. 80, pp. 38–46, 2015.

[183] A. R. Nejad, E. E. Bachynski, L. Li, and T. Moan, "Correlation between Acceleration and Drivetrain Load Effects for Monopile Offshore Wind Turbines," *Energy Procedia*, vol. 94, no. 1876, pp. 487–496, 2016.

[184] B. Jonkman and L. Kilcher, "TurbSim User's Guide: Version 1.06.00," 2012.

[185] A. Pinkus, "Approximation theory of the MLP model in neural networks," *Acta Numerica*, vol. 8, pp. 143–195, 1999.

[186] R. A. Caruana, "Multitask Learning: A Knowledge-Based Source of Inductive Bias," in *Machine Learning Proceedings 1993*, Elsevier, 1993, pp. 41–48.

[187] E. S. Olivas, J. D. M. Guerrero, M. Martinez-Sober, J. R. Magdalena-Benedito, and A. J. S. López, Eds., *Handbook of Research on Machine Learning Applications and Trends*. {IGI} Global, 2010.

[188] Z. Lin *et al.*, "Multi-task learning for quality assessment of fetal head ultrasound images," *Medical Image Analysis*, vol. 58, p. 101548, 2019.

[189] W. Xiaohua, P. Muzi, P. Lijuan, H. Min, J. Chunhua, and R. Fuji, "Two-level attention with two-stage multi-task learning for facial emotion recognition," *Journal of Visual Communication and Image Representation*, vol. 62, pp. 217–225, Jul. 2019.

[190] L. Huang, X. Feng, C. Zhang, L. Qian, and Y. Wu, "Deep reinforcement learning-based joint task offloading and bandwidth allocation for multi-user mobile edge computing," *Digital Communications and Networks*, vol. 5, no. 1, pp. 10–17, Feb. 2019.

[191] I. H. Witten and E. Frank, "Data mining: practical machine learning tools and techniques with Java implementations," *Acm Sigmod Record*, vol. 31, no. 1, pp. 76–77, 2002.

[192] K. He, X. Zhang, S. Ren, and J. Sun, "Delving Deep into Rectifiers: Surpassing Human-Level Performance on {ImageNet} Classification," in *2015 EEE International Conference on Computer Vision (ICCV)*, 2015.

[193] K. Marti, "Stochastic Optimization Methods," in *Stochastic Optimization Methods*, Springer Berlin Heidelberg, 2015, pp. 1–35.

[194] D. Augustyn, U. Smolka, U. T. Tygesen, M. D. Ulriksen, and J. D. Sørensen, "Data-driven

model updating of an offshore wind jacket substructure," *Applied Ocean Research*, vol. 104, no. May, p. 102366, 2020.

[195] D. J. Augustyn, R. Pedersen, M. D. Ulriksen, and J. D. Sørensen, "Feasibility of modal expansion for virtual sensing in offshore wind jacket substructures," *submitted to Marine Structures*, 2020.

[196] D. J. Augustyn, M. D. Ulriksen, and J. D. Sørensen, "Reliability Updating of Offshore Wind Substructures by Use of Digital Twin Information," *Energies*, vol. 14, no. 5859, pp. 1–22, 2021.

[197] Newsroom, "Ramboll's True Digital Twin technology has the potential to increase lifetime of offshore wind farms," *January 11, 2021*. [Online]. Available: https://ceenergynews.com/innovation/rambolls-true-digital-twin-technology-has-the-potential-to-increase-lifetime-of-offshore-wind-farms/. [Accessed: 25-May-2022].

[198] S. S. Siedler, I. M. Black, C. Wendelborn, M. W. Häckell, J. Tautz-Weinert, and D. Cevasco, "D4.5 Update of report D4.3 considering long term findings," 2022.

[199] L. Bauer and S. Matysik, "Adwen AD 5-135," *wind-turbine-models.com*. [Online]. Available: https://en.wind-turbine-models.com/turbines/1122-adwen-ad-5-135. [Accessed: 23-Oct-2020].

[200] NGI group, "BSEE Offshore Wind Recommendations - Guidlines for Structural Health Monitoring for Offshore Wind Turbine Towers and Foundations," 2017.

[201] Ramboll, "ROSAP 5.3 - Ramboll Offshore Structural Analysis Program Package.," 2019.

[202] R. D. Cook and others, *Concepts and applications of finite element analysis*. John wiley \& sons, 2007.

[203] American Petroleum Institute. Division of Production, *API Recommended Practice for Planning, Designing, and Constructing Fixed Offshore Platforms*. American Petroleum Institute, Production Department, 1974.

[204] J. Buitrago, B. E. Healy, and T. Y. Chang, "Local joint flexibility of tubular joints," 1993.

[205] M. Pastor, M. Binda, and T. Harčarik, "Modal Assurance Criterion," in *MMaMS*, 2012, vol. 48, no. 48, pp. 543–548.

[206] A. K. Pandey and M. Biswas, "Damage detection in structures using changes in flexibility," *Journal of Sound and Vibration*, vol. 169, no. 1, pp. 3–17, 1994.

[207] Ramboll, "LACFlex aero-servo-elestic tool v1.9.2.0," 2019.

[208] M. B. Nielsen, J. F. Jensen, D. Augustyn, and R. R. Pedersen, "Efficient response recovery procedures for detailed design of jacket foundations," in *Insights and Innovations in Structural Engineering, Mechanics and Computation - Proceedings of the 6th International Conference on Structural Engineering, Mechanics and Computation, SEMC 2016*, 2016, no. January 2018, pp. 2060–2065.

[209] *BS EN 61400-3:2009 - Wind turbines - Part 3: Design requirements for offshore wind turbines*. 2009.

[210] A. Rytter, "Vibrational Based Inspection of Civil Engineering Structures," Aalborg University, 1993.

[211] G. Oliveira, "Structural Health Monitoring of Wind Turbines," Faculdade de engenharia, Universidade do Porto, 2016.

[212] R. A. Johnson and D. W. Wichern, *Applied Multivariate Statistical Analysis*. 2013.

[213] K. Liu, R. J. Yan, and C. Guedes Soares, "Damage identification in offshore jacket

structures based on modal flexibility," *Ocean Engineering*, vol. 170, no. October, pp. 171–185, 2018.

[214] A. Alvandi and C. Cremona, "Assessment of vibration-based damage identification techniques," *Journal of Sound and Vibration*, vol. 292, no. 1–2, pp. 179–202, 2006.

[215] H. Malekzehtab and A. A. Golafshani, "Damage detection in an offshore jacket platform using genetic algorithm based finite element model updating with noisy modal data," in *2nd International Conference on Rehabilitation and Maintenance in Civil Engineering*, 2013, vol. 54, pp. 480–490.

[216] M. Xu, S. Wang, and H. Li, "A residual strain energy based damage localisation method for offshore platforms under environmental variations," *Ships and Offshore Structures*, vol. 14, no. 7, pp. 747–754, 2018.

[217] Y. Diao, X. Men, Z. Sun, K. Guo, and Y. Wang, "Structural damage identification based on the transmissibility function and support vector machine," *Shock and Vibration*, vol. 2018, 2018.

[218] Y. S. Cheng and Z. Wang, "Detecting damage to offshore platform structures using the time-domain data," *Journal of Marine Science and Application*, vol. 7, no. 1, pp. 7–14, 2008.

[219] T. Li, S. Zhu, and M. Ogihara, "Using discriminant analysis for multi-class classification: an experimental investigation," *Knowledge and Information Systems*, vol. 10, no. 4, pp. 453–472, 2006.

[220] R. Suganya and R. Shanthi, "Fuzzy C-Means Algorithm-A Review," *International Journal of Scientific and Research Publications*, vol. 2, no. 11, pp. 2250–3153, 2012.

[221] "Fuzzy c-Means," *2015*. [Online]. Available: http://researchhubs.com/post/ai/fundamentals/fuzzy-c-means.html#:~:text=Advantages%3A,more than one cluster center. [Accessed: 11-Mar-2021].

[222] J. D. Warner and et al., "Fuzzy Logic SciKit (Toolkit for SciPy)." [Online]. Available: https://github.com/scikit-fuzzy/scikit-fuzzy.

[223] "Scikit-learn 0.22 Online Guide - 3.1. Cross-validation: evaluating estimator performance - 3.1.2.2.1. Stratified k-fold."

[224] J. Weinert, U. Smolka, B. Schümann, and P. W. Cheng, "Detecting Critical Scour Developments at Monopile Foundations Under Operating Conditions," *Proceedings of the European Wind Energy Association Annual Event, EWEA 2015*, pp. 135–139, 2015.

[225] C. Ferri, J. Hernández-Orallo, and R. Modroiu, "An experimental comparison of performance measures for classification," *Pattern Recognition Letters*, vol. 30, no. 1, pp. 27–38, 2009.

[226] N. Cosack, "Fatigue Load Monitoring with Standard Wind Turbine Signals," Universität Stuttgart, 2010.

[227] N. Noppe, A. Iliopoulos, W. Weijtjens, and C. Devriendt, "Full load estimation of an offshore wind turbine based on SCADA and accelerometer data," *Journal of Physics: Conference Series*, vol. 753, no. 7, 2016.

[228] L. Vera-Tudela and M. Kühn, "Evaluation of a wind turbine fatigue load monitoring system based on standard SCADA signals in different wind farm flow conditions," *Dewek 2015*, vol. 49, no. 441, 2015.

[229] L. Vera-Tudela and M. Kühn, "Analysing wind turbine fatigue load prediction: The impact of wind farm flow conditions," in *Renewable Energy*, 2017, vol. 107, pp. 352–360.

[230] L. Ziegler, "Assessment of monopiles for lifetime extension of offshore wind turbines," Norwegian University of Science and Technology, 2018.

[231] A. Pena Diaz *et al.*, "Offshore vertical wind shear: Final report on NORSEWInD's work task 3.1.," 2012.

[232] *IEC 61400-1 Edition 3 Amendment 1*. 2010.

[233] F. Pedregosa *et al.*, "Scikit-learn: Machine Learning in Python," *JMLR*, 2011. [Online]. Available: http://jmlr.csail.mit.edu/papers/v12/pedregosa11a.html.

[234] "Scikit-learn 0.22 Online Guide - 2.5. Decomposing signals in components (matrix factorization problems) - 2.5.1. Principal component analysis (PCA)."

[235] J. Brownlee, "Logistic Regression for Machine Learning," *Last Updated on August 12, 2019*. .

[236] A. Niculescu-Mizil and R. Caruana, "Predicting Good Probabilities With Supervised Learning," in *Appearing in Proceedings of the 22 nd International Conference on Machine Learning*, 2005.

[237] D. Cevasco, U. Smolka, and A. Kolios, "Feasibility of machine learning algorithms for classifying damaged offshore jacket structures using SCADA data," in *IOP Conference Series, presented at EERA DeepWind2020*, 2020.

[238] International Organisation for Standardization, "Petroleum , petrochemical and natural gas industries - Reliability modelling and calculation of safety systems (PD CEN ISO/TR 12489:2016)," Geneva, Switzerland, 2016.

[239] International Organisation for Standardization, "Petroleum, petrochemical and natural gas industries - Collection and exchange of reliability and maintenance data for equipment (BS EN ISO 14224:2016)," Geneva, Switzerland, 2016.

[240] International Electrotechnical Commission, "Wind Energy Generation Systems - Part 26-1: Availability for Wind Energy Generating Systems (IEC CDV 61400-26-1)," Geneva, Switzerland, 2018.

[241] International Organisation for Standardization, "Functional safety of electrical/electronic/programmable electronic safety related systems - Part 4: Definitions and abbreviations (BS EN 61508:2010)," Geneva, Switzerland, 2010.

[242] E. Gonzalez *et al.*, "Key Performance Indicators for Wind Farm Operation and Maintenance," in *14th Deep Sea Offshore Wind R&D Conference, EERA DeepWind'2017*, 2017, vol. 137, pp. 559–570.

[243] A. de Myttenaere, B. Golden, B. Le Grand, and F. Rossi, "Mean Absolute Percentage Error for regression models," *Neurocomputing*, vol. 192, pp. 38–48, Jun. 2016.

[244] N. R. Draper and H. Smith, *Applied regression analysis*. Wiley, 1998.

[245] "Scikit-learn 0.22 Online Guide - 3.3. Metrics and scoring: quantifying the quality of predictions - 3.3.2.14. Receiver operating characteristic (ROC)." .

[246] J. Brownlee, "How and When to Use a Calibrated Classification Model with scikit-learn," *Last Updated on September 25, 2019*. [Online]. Available: https://machinelearningmastery.com/calibrated-classification-model-in-scikit-learn/.

Appendices

Appendix A

Further research activities

A.1 Other publications and collaborative works

Paper 12 Z. Lin, **D. Cevasco**, and M. Collu, "Progress on the development of a holistic coupled model of dynamics for offshore wind farms, phase I: Aero-Hydro-Servo-Elastic Model, with drive train model, for a single wind turbine," in *Proceedings of the ASME 37th International Conference on Ocean, Offshore and Arctic Engineering* (OMAE2018), 2018, pp. 1–9.

Paper 13 M. Leimeister **et al.**, "Human-free offshore lifting solutions," in *IOP Conference Series*, presented at Wind Summit: WindEurope 2018, 2018, vol. 1102, no. 1. https://doi.org/10.1088/1742-6596/1102/1/012030

Paper 14 M. Richmond **et al.**, "Multi-criteria decision analysis for benchmarking human-free lifting solutions in the offshore wind energy environment," *Energies*, vol. 11, no. 5, 2018. https://doi.org/10.3390/en11051175

Paper 15 **D. Cevasco**, M. Collu, C. M. Rizzo, and M. Hall, "On mooring line tension and fatigue prediction for offshore vertical axis wind turbines: A comparison of lumped mass and quasi-static approaches," Wind Engineering, vol. 42, no. 2, 2018. https://doi.org/10.1177/0309524X18756962

A.2 Oral and poster presentations

Presentation 1 Oral presentation of Paper 15 at the OWEMES 2018 (Bari, Italy).

Poster 1 Poster presentation of Paper 13 at the Wind Summit (WindEurope) 2018 (Hamburg, Germany).

Poster 2 Poster presentation of Paper 8 at the Wind Summit (WindEurope) 2018 (Hamburg, Germany).

Presentation 2 Oral presentation of Paper 3 at the DeepWind 2020 (Trondheim, Norway).

Poster 3 Poster and oral presentation of Paper 4***Paper 13*** at the Torque 2021 (Delft, Netherlands).

A.3 Authors' contribution matrices

The scientific papers, which have been published on journals and conferences, for the partial fulfilment of the requirements for the degree of Doctor of Engineering, were co-authored by several academic and industrial colleagues. As some chapters of this thesis are

based on these papers, this section is meant to clarify the authors contributions and roles, according to the CRediT standards. This standard characterises the authors' contributions according to the following classes – numbered to ease their referencing:

1.	*Conceptualization*	Ideas, formulation, or evolution of overarching research goals and aims.
2.	*Methodology*	Development or design of methodology, creation of models.
3.	*Software*	Programming, software development, designing computer programs, implementation of the computer code and supporting algorithms, testing of existing code components.
4.	*Validation*	Verification, whether as a part of the activity or separate, of the overall replication/reproducibility of results/experiments and other research outputs.
5.	*Formal Analysis*	Application of statistical, mathematical, computational, or other formal techniques to analyse or synthesize study data.
6.	*Investigation*	Conducting a research and investigation process, specifically performing the experiments, or data/evidence collection.
7.	*Resources*	Provision of study materials, reagents, materials, patients, laboratory samples, animals, instrumentation, computing resources, or other analysis tools.
8.	*Data Curation*	Management activities to annotate (produce metadata), scrub data, and maintain research data (including software code, where it is necessary for interpreting the data itself) for initial use and later reuse.
9.	*Original Draft Preparation (Writing)*	Creation and/or presentation of the published work, specifically writing the initial draft (including substantive translation).
10.	*Review and Editing (Writing)*	Preparation, creation and/or presentation of the published work by those from the original research group, specifically critical review, commentary, or revision – including prepublication or post publication stages.
11.	*Visualization*	Preparation, creation and/or presentation of the published work, specifically visualization/data presentation.
12.	*Supervision*	Oversight and leadership responsibility for the research activity planning and execution, including mentorship external to the core team.
13.	*Project Administration*	Management and coordination responsibility for the research activity planning and execution.
14.	*Funding Acquisition*	Acquisition of the financial support for the project leading to this publication.

The following tables make use of these classes to define the roles and contributions of the (first) author and the co-authors to the scientific publications (Paper 1-7) which are used for the preparation of some of the chapters of this thesis manuscript (Chapter 1, 2, 3, 5 and 6) – please refer to the chapters description in Section 1.4. A grey-scale coding is used to indicate the level of contribution of the authors: *dark grey* indicates a leading role, while *light grey* is for authors acting as support to the lead. When the authors share the same colour code for the same role, it indicates an equal contribution and distribution of the tasks.

Table A.1: Authors' contribution matrix to *Paper 1*.

	class 1	*class 2*	*class 3*	*class 4*	*class 5*	*class 6*	*class 7*	*class 8*	*class 9*	*class 10*	*class 11*	*class 12*	*class 13*	*class 14*
D. Cevasco														
S. Koukoura														
A. J. Kolios														

Table A.2: Authors' contribution matrix to *Paper 2*.

	class 1	*class 2*	*class 3*	*class 4*	*class 5*	*class 6*	*class 7*	*class 8*	*class 9*	*class 10*	*class 11*	*class 12*	*class 13*	*class 14*
Z. Lin														
D. Cevasco														
M. Collu														

Table A.3: Authors' contribution matrix to *Paper 3*.

	class 1	*class 2*	*class 3*	*class 4*	*class 5*	*class 6*	*class 7*	*class 8*	*class 9*	*class 10*	*class 11*	*class 12*	*class 13*	*class 14*
D. Cevasco														
J. Tautz-Weinert														
U. Smolka														
A.J. Kolios														

Table A.4: Authors contribution matrix to *Paper 4*.

	class 1	*class 2*	*class 3*	*class 4*	*class 5*	*class 6*	*class 7*	*class 8*	*class 9*	*class 10*	*class 11*	*class 12*	*class 13*	*class 14*
D. Cevasco														
J. Tautz-Weinert														
A.J. Kolios														
U. Smolka														

Table A.5: Authors' contribution matrix to *Paper 5*.

	class 1	*class 2*	*class 3*	*class 4*	*class 5*	*class 6*	*class 7*	*class 8*	*class 9*	*class 10*	*class 11*	*class 12*	*class 13*	*class 14*
D. Cevasco														
J. Tautz-Weinert														
M. Richmond														
A. Sobey														
A.J. Kolios														

Table A.6: Authors' contribution matrix to *Paper 6*.

	class 1	*class 2*	*class 3*	*class 4*	*class 5*	*class 6*	*class 7*	*class 8*	*class 9*	*class 10*	*class 11*	*class 12*	*class 13*	*class 14*
I. M. Black	■	■	■	■	■		■	■	■	■	■			
D. Cevasco	■			■	■	■			■	■	■		■	
A. J. Kolios							■					■	■	■

Table A.7: Authors' contribution matrix to *Paper 7*.

	class 1	*class 2*	*class 3*	*class 4*	*class 5*	*class 6*	*class 7*	*class 8*	*class 9*	*class 10*	*class 11*	*class 12*	*class 13*	*class 14*
M. Richmond	■	■			■	■	■	■	■	■	■			
D. Cevasco	■	■			■	■	■		■	■	■			
A. J. Kolios	■								■	■		■	■	■

Appendix B

RAM and performance metrics and failure statistics for availability calculations

B.1 RAM definitions and performance metrics

As the data collected in Chapter 2 is heterogeneous, a terminology is introduced for describing and classifying the results in the several databases. Standardised definitions of reliability, maintainability, and key performance indicators (KPIs) are summarised in the following section.

B.1.1 Reliability and maintainability terminology

In the reliability analysis, an indication of the frequency of the failure and/or the time elapsing until the system is restored is generally given. The frequency parameter is usually represented by the failure rate (λ), which is the likelihood of a system to fail within a specific period. Unlike a probability, however, it can reach values greater than 1. Focusing on the constant failure rate region of an asset, the indicator λ of a WT, consisting of K components, is averaged over the *i-th* recording periods according to [19], [122]. The I is the number of intervals for surveys of length T_i and n_i failures per interval. This Power Law Process is commonly used in the reliability analysis of repairable systems [18]. As the data collected are from many turbines, the data are generally normalised by the number of units in the population N_i, as well as providing, for instance, information on the number of failures per turbine, per year (i.e. [failure/turbine/year]).

$$\lambda = \frac{\sum_{i=0}^{I} \sum_{k=1}^{K} n_{i,k} / N_i}{\sum_{i=0}^{I} T_i / 8760} \qquad \textbf{Equation B-1}$$

According to the relevant ISO standards [238], [239], a distinction between a repairable and a non-repairable system is required, as well as distinguishing whether the maintenance action consists of a repair or a replacement measure. Having only had access to limited information, this quantity is considered in this paper as a general failure, or maintenance event rate. Furthermore, λ depends on the definition of failure itself. A "failure" is the loss of ability to perform as required; however, the qualitative judgement of the authors can vary when interpreting the term "loss of ability".

With respect to the quantity characterising the "failure time" several definitions are relevant, and the most recurring statistics analysed are reported in Table A.1. These measures are used as representative of the maintainability characteristics of the failure. They have the dimension of time, with varying resolution, and they can be normalised by either the number of turbines or the failures in the time interval considered. More

specifically, the definition of "downtime" varies from the minimum time to perform the repair (MART), to the total time expected from when the system fails to its restoration (MTTR or MTTRes). The more commonly collected mean downtime (MDT) differs from the MTTR in reporting shutdown events due to grid restrictions, weather conditions, and other causes (for more information refer to table 4 of reference [239]).

Table B.1: Acronyms and definition for the "downtime" terminology.

Acronym	Definition (cp. [238], [239])	Comments and Notes
MART	***Mean Active Repair Time*** *"Expected active repair time" or "expected effective time to repair."*	With "active", the ISOs define the effective time to achieve repair of an item. This accounts for: fault localization, correction, and checkout time. This definition is in agreement with the one of the International Electrotechnical Commission (IEC) [240] defined, and commonly called, "repair time". (Refer to *Figure 4* of [239], and *Figures 5 and 6* of [238]).
MTTR (MTTRes)	***Mean Time to Repair/Restoration*** *"Expected time to achieve the following actions:* - *time to detect the failure;* - *time spent before starting the repair (with administrative, logistics and technical delays);* - *effective repair time (MART);* - *time before the component is available to be put back into operation (possible other administrative delays)."*	The ISOs defined MTTRes (mean time to restoration) wants to be an elucidation to the MTTR (mean time to repair) from IEC [241]. In the latter, the fault detection time is not considered. Thus, MTTRes is defined as MTTRes = MRT + MFDT where MRT and MFDT are, respectively: - the time elapsing from the actual occurrence of the failure of an item to its detection, and - time elapsing from the detection of the failure of an item to the restoration of its function. (Refer to *Figure 5* of [238]).
MDT	***Mean Downtime*** *"Expectation of the downtime."* The downtime is the time interval during which an item is in a down state, and thus *"unavailable"*. (Refer to *Figures 3* and *4* of [239])	• It can be either for planned or unplanned maintenance actions. However, only the latter (corrective O&M actions) is considered in this work. • The downtime includes all the delays between the item failure and the restoration of its service. • It differs from MTTR accounting also for "other unplanned outages"; among these are: operational problems, restrictions, and machinery shutdown (trip* and manual).

* Defined as the shutdown of a piece of machinery (activated automatically by the control/monitoring system) from normal operating conditions to full stop. It can either be "real", if the result of exceedance (monitored or calculated) of a pre-set limit, or "spurious" when an unexpected shutdown caused by a failure.

B.1.2 Performance indicators

Production factors are used by several authors as an indication of the averaged performance of wind turbines and farms (e.g. [59], [88]). The KPIs used in these papers are explained and reported in and are in line with those adopted by [58], [242].

Among them, the technical availability (A_T) is the most meaningful one for the understanding of unexpected failures. It is defined as the ratio between uptime and downtime of the turbine, and, by considering for the latter only the downtime for corrective

maintenance (excluding scheduled actions), it gives combined information on the frequency of and restoring time for the failure. The time-based, or operational (A_O), and energetic (A_E) availability are employed as a measure of the actual performance of turbines and/or farms. The A_O is the availability derived by the lifecycle assessment tool of Section 2.3.3. As regards the A_E, being the estimation of the potential power output of a complex process, and subject to high uncertainties, the capacity factor (CF) is often encountered instead.

Table B.2: Symbols and definitions of the KPIs.

Symbol	Expansion, Reference, and Description	Formulae	Formulae terms
A_T	***Technical Availability*** [240] Fraction of a given period in which a turbine is operating according to its design specifications	$\frac{t_{available}}{t_{available} + t_{unavailable}}$	$t_{available}$ - time of full and partial performance - technical standby and requested shutdown - downtime due to environment and grid $t_{unavailable}$ time of corrective actions and force outage (excluding missing data and scheduled maintenance)
A_O	***Operational Availability*** [240] *(or **Time-based Availability**)* Share of the time when the system is operating and/or able to operate, compared to the total time		$t_{available}$ time of full and partial performance (considering low wind as well) $t_{unavailable}$ time of all the other cases (excluding missing data)
A_E	***Energetic Availability*** [58] Amount of energy produced by the system compared to its potential energy production	$\frac{\overline{P}_{actual}}{\overline{P}_{potential}}$	$\overline{P}_{actual}$: Average actual power output $\overline{P}_{potential}$: Average potential power output (excluding missing data)
CF	***Capacity Factor*** [240] Ratio between the amount of energy actually produced by the system and what it can theoretically produce	$\frac{\overline{P}_{actual}}{P}$	P: Rated power output

B.2 Failure statistics for availability calculations

In Table B.3, the failure rate statistics used as an input for the availability calculation in the openO&M tool are reported. The Strath-Off and onshore data are integrated by making assumptions on the time required for performing offshore maintenance. The OWEZ statistics are used for comparison, being the only other offshore survey reporting sufficient information on its reliability and maintainability figures. Additionally, with respect to the onshore studies, only the most complete and consistent studies identified from Section 2.3.1 are considered, further subdividing the surveys depending on turbine configuration and/or power rating where possible.

Table B.3: Repositories for the calculation of offshore wind farms' availability.

	CIRCE DD			CIRCE C > 1MW			Huadian			LWK		
	mr	Mr	MR	mr	Mr	MR	mr	Mr	MR	mr	Mr	MR
1. Rotor System			0.012		0.084			0.141		0.321		
1a. Rotor Blades			0.005			0.019			0.026	0.194		
1b. Rotor Hub Unit			0.005		0.011							
1c. Rotor Brake System										0.040		
1d. Pitch System			0.003		0.051			0.115		0.088		
2. Drive Train System	0.003				0.081			0.088			0.226	
2a. Speed Conversion System					0.065			0.062			0.142	
2b. Brake System Drivetrain				0.003					0.018	0.053		
3. Yaw System		0.014			0.021				0.026	0.115		
4. Central Hydraulic System				0.021						0.134		
5. Control System	0.052			0.058				0.106		0.222		
6. Power Generation System			0.011			0.025			0.150		0.140	
7. Transmission	0.026				0.044			0.291		0.323		
7a. Converter System	0.003				0.007			0.229		0.005		
7b. Generator Transformer System		0.008				0.006			0.018			
8. Nacelle												
9. Common Cooling System	0.005			0.041								
10. Meteorological Measurement	0.019			0.012						0.061		
11. Tower System	0.003			0.006								
11a. Tower				0.003								
11b. Foundation System	0.003				0.003							
12. Others	0.045			0.126			0.044			0.312		

Table B.3: Repositories for the calculation of offshore wind farms' availability (cont.).

	Muppandal			Strath Off			WMEP			OWEZ		
	mr	Mr	MR	mr	Mr	MR	mr	Mr	MR	mr	Mr	mr
1. Rotor System	0.027	0.133	0.027	1.462	0.227	0.003		0.522			0.704	
1a. Rotor Blades	0.027	0.133	0.027	0.456	0.010	0.001		0.113				0.054
1b. Rotor Hub Unit				0.182	0.038	0.001		0.171				
1c. Rotor Brake System												
1d. Pitch System				0.824	0.179	0.001	0.238				0.649	
2. Drive Train System	0.280			0.395	0.038	0.154		0.291				0.498
2a. Speed Conversion System	0.173			0.395	0.038	0.154		0.106				0.486
2b. Brake System Drivetrain	0.107						0.130					0.012
3. Yaw System	0.160			0.162	0.006	0.001	0.177			1.456		
4. Central Hydraulic System	0.173						0.225					
5. Control System		0.080		0.355	0.054	0.001	0.403			2.659		
6. Power Generation System	0.067			0.485	0.321	0.095			0.100			0.206
7. Transmission	0.040			0.812	0.154	0.010	0.548					0.381
7a. Converter System	0.040			0.076	0.081	0.005						0.195
7b. Generator Transformer System				0.052	0.003	0.001						
8. Nacelle								0.094				
9. Common Cooling System				0.190	0.007							
10. Meteorological Measurement	0.027											
11. Tower System				0.092	0.089						0.052	
11a. Tower											0.052	
11b. Foundation System				0.092	0.089							
12. Others				2.225	0.166	0.001	0.245			0.385		

Appendix C

Prediction metrics

C.1 Time series and regression metrics

Time series can be compared by the means of the mean squared error (*MSE*). As defined in Equation C-1 , the *MSE* measures the average of the squares of the errors – defined as the difference between the estimated and the actual value. – with Y being the vector of values and n being number of points in the sample. The smallest is the *MSE*, the closest is the rotor torque to the one of the full-order nonlinear model, which is taken as term of comparison.

$$\mathrm{MSE} = \frac{1}{\mathrm{n}}\sum_{\mathrm{i}=1}^{\mathrm{n}}(\mathrm{Y}_{\mathrm{actual}} - \mathrm{Y}_{\mathrm{estimated}})^2 \qquad \textbf{Equation C-1}$$

The performance of a regression model is generally evaluated by using the mean absolute percentage error (*MAPE*), and the coefficient of determination (referred to "R squared" and denoted as R2). The *MAPE* represents the average absolute percentage across the time periods. It is thus defined as in Equation C-2 where $\hat{y}_i$ is the predicted value, y_i is the actual value, and n is the number of time period samples [243].

$$\mathrm{MAPE} = \frac{1}{\mathrm{n}}\sum_{\mathrm{i}=1}^{\mathrm{n}}\left|\frac{\hat{\mathrm{y}}_{\mathrm{i}}\text{-}\mathrm{y}_{\mathrm{i}}}{\mathrm{y}_{\mathrm{i}}}\right| \qquad \textbf{Equation C-2}$$

Another common metric, so called *R2*, measures of how well the observed outcomes are replicated by the model, based on the proportion of total variation of outcomes explained by the model [244]. This metric can be more intuitive and informative than the *MAPE* for the purpose of the regression analysis evaluation, as the former can be expressed as a percentage, whereas the latter measures have arbitrary ranges.

C.2 Classification metrics

The confusion matrix indicators, visualised in the confusion matrix of Table C.1, allow defying the hard threshold metrics of binary-class classification problem. The *tn* and *tp* refer to the correct prediction of healthy and damaged cases, respectively, while *fp* and *fn* give the false alarms and miss, respectively, as common in the confusion matrix.

The classification models presented in this thesis aim to identify the damage status of the assets. Therefore, the “positive” class it is here referred to as “damaged”. The hard-threshold metrics, that are used to quantify the performance of supervised or unsupervised machine learning approaches, are defined as in [225], by adapting the terminology to fit the damage-detection purpose:

Table C.1: Confusion matrix for hard-threshold metrics.

		Predicted	
		negative (0 or healthy)	**positive (1 or damaged)**
Actual	**negative (0 or healthy)**	true negative (tn)	false positive (fp)
	positive (1 or damaged)	false negative (fn)	true positive (tp)

- Accuracy (*acc*) gives an indication of the total amount of correct predictions over the total amount of samples in the dataset tested.

$$\text{acc} = \frac{\text{tp} + \text{tn}}{\text{Total samples}} \qquad \textbf{Equation C-3}$$

- Precision (*pr*) is defined as the percentage of correctly detected damaged case with respect to the total amount of cases predicted damaged.

$$\text{pr} = \frac{\text{tp}}{\text{fp} + \text{tp}} \qquad \textbf{Equation C-4}$$

- Recall (*re*) is defined as the percentage of correctly detected damaged case with respect to the total amount of damaged cases in the dataset tested.

$$\text{re} = \frac{\text{tp}}{\text{fn} + \text{tp}} \qquad \textbf{Equation C-5}$$

- False positive rate, or fall-out (*fo*), gives an indication of the percentage of false alarm raised (the classifier predicts damaged status although the structure is healthy) with respect to the total amount of healthy cases in the dataset.

$$\text{fo} = \frac{\text{fp}}{\text{tn} + \text{fp}} \qquad \textbf{Equation C-6}$$

- *F1-score* is the harmonic mean of the precision and recall.

$$\text{F1} - \text{score} = 2\frac{\text{pr} \cdot \text{re}}{\text{pr} + \text{re}} \qquad \textbf{Equation C-7}$$

- Specificity (*sp*), defined as 1 – *fo*, and thus refers to the probability of a positive test, conditioned on truly being positive.

While the *acc* gives an overall indication on the goodness of the classifier, the *re* refers to the models' ability to detect the damage. Satisfactory *acc*, *pr* and *re* are generally assumed for values above 70%. The *fa*, instead, gives an indication of the percentage of false alarms raised by the classifier. A satisfactory *fo* is generally considered if below 30%.

Receiver operating characteristic (ROC) curves [245], and feature ranking plots are employed during the tuning and training phases to track the performance and validity of the prediction. Additionally, when performing classification, it is advised to investigate the probability corresponding to the predicted category. This probability gives a measure of the confidence on the prediction and is presented in the so-called reliability curve [246].

Appendix D

Tabular results of damage detection in an offshore wind jacket structure based on low-resolution

D.1 Training and testing

The complete classification results of the training and testing of the data-driven models, built for the purpose of the non-parametric detection of a brace full-integrity loss in the jacket substructure of the WK64 turbine, are presented in Table D.1. The metrics, and the test sets are as the ones of Section C.2 and Section 6.5, respectively, however the reader is referred to Paper 3 for their naming convention.

Table D.1: At the top, the performance of the best classifiers for the dataset D0 (trained on Tr0), sensor set up S0. At the bottom, the colour legend (targeted performances in green).

	Classifier	Tr0			Te33			Te1	Te2	Te3	Te4
		acc	*TDR*	*FDR*	*acc*	*TDR*	*FDR*	*acc*	*acc*	*acc*	*acc*
Below rated	LR	69%	67%	28%	70%	67%	27%	50%	50%	52%	52%
	SVM (poly)	91%	92%	9%	71%	72%	30%	50%	50%	53%	54%
	RF	100%	100%	0%	86%	86%	14%	55%	68%	66%	72%
Above rated	LR	61%	60%	39%	59%	59%	41%	50%	50%	52%	50%
	SVM (rbf)	89%	86%	9%	64%	66%	37%	50%	50%	52%	50%
	RF	100%	100%	0%	69%	66%	28%	56%	56%	60%	59%

	acc/TDR	*FDR*
	below 60	above 40
	[60;75)	(30;40]
	[75;90)	(10;30]
	[90;100]	[0;10]

D.2 LR and SVM

The results of the models based on the logistic regression (LR) and the support vector machine (SVM) algorithms are presented in Table D.2 and in Table D.3. For brevity, the performances on the test sets are reported in terms of accuracy (*acc*) only. However, the confusion matrix (*CM*) of each classification test is given to allow deriving the other performance indicators. The test sets are as the ones of Section 6.5, however the reader is referred to Paper 3 for their naming convention.

Overfitting is identified when a strong collinearity is introduced among the classification predictors, thus by implementing both the inclinometer and strain gauge measurements (sensor setup S2). Indicators of this phenomenon are:

- the improvement of the accuracy for both LR and SVM classifiers,
- the switch to a linear separation function for the SVM classifier, while more complex kernel transformations are employed for the all the other sensor combinations.

Table D.2: LR classifiers performance on the dataset D0, sensor set up S0, S1, S2 and S3.

	Sensor	Tr0		Te33		Te1		Te2		Te3		Te4	
		CM	*acc*	*CM*	*acc*	*CM*	*acc*	*CM*	*acc*	*CM*	*acc*	*CM*	*acc*
Below rated	S0	[[622, 242], [288, 576]]	69%	[[314, 118], [143, 289]]	70%	[[8, 1288], [2, 1294]]	50%	[[1296, 0], [1293, 3]]	50%	[[54, 1242], [15, 1281]]	52%	[[1267, 29], [1206, 90]]	52%
	S1	[[630, 234], [245, 619]]	72%	[[308, 124], [136, 296]]	70%	[[12, 1284], [6, 1290]]	50%	[[1294, 2], [1290, 6]]	50%	[[87, 1209], [16, 1280]]	53%	[[1247, 49], [1184, 112]]	52%
	S2	[[864, 0], [0, 864]]	100%	[[432, 0], [0, 432]]	100%	[[1273, 23], [2, 1294]]	99%	[[1255, 41], [122, 1174]]	94%	[[1294, 2], [0, 1296]]	100%	[[1293, 3], [18, 1278]]	99%
	S3	[[538, 326], [332, 532]]	62%	[[266, 166], [175, 257]]	61%	[[327, 969], [163, 1133]]	56%	[[1269, 27], [1192, 104]]	53%	[[435, 861], [227, 1069]]	58%	[[1144, 152], [902, 394]]	59%
Above rated	S0	[[529, 335], [342, 522]]	61%	[[256, 176], [175, 257]]	59%	[[7, 1289], [5, 1291]]	50%	[[1296, 0], [1296, 0]]	50%	[[107, 1189], [48, 1248]]	52%	[[1296, 0], [1296, 0]]	50%
	S1	[[549, 315], [310, 554]]	64%	[[261, 171], [138, 294]]	64%	[[539, 757], [276, 1020]]	60%	[[1295, 1], [1294, 2]]	50%	[[622, 674], [325, 971]]	61%	[[1261, 35], [1120, 176]]	55%
	S2	[[864, 0], [0, 864]]	100%	[[431, 1], [0, 432]]	100%	[[1277, 19], [0, 1296]]	99%	[[1296, 0], [882, 414]]	66%	[[1292, 4], [0, 1296]]	100%	[[1295, 1], [184, 1112]]	93%
	S3	[[538, 326], [334, 530]]	62%	[[262, 170], [146, 286]]	63%	[[18, 1278], [10, 1286]]	50%	[[1296, 0], [1296, 0]]	50%	[[90, 1206], [33, 1263]]	52%	[[1296, 0], [1296, 0]]	50%

Table D.3: SVM classifiers performance on the dataset D0, sensor set up S0, S1, S2 and S3.

	Kernel	Sensor	Tr0		Te33		Te1		Te2		Te3		Te4	
			CM	*acc*	*CM*	*acc*	*CM*	*acc*	*CM*	*acc*	*CM*	*acc*	*CM*	*acc*
Below rated	poly (4)	S0	[[787, 77], [73, 791]]	91%	[[303, 129], [121, 311]]	71%	[[916, 380], [922, 374]]	50%	[[358, 938], [346, 950]]	50%	[[707, 589], [632, 664]]	53%	[[683, 613], [574, 722]]	54%
	poly (2)	S1	[[815, 49], [67, 797]]	93%	[[378, 54], [84, 348]]	84%	[[1191, 105], [1140, 156]]	52%	[[1285, 11], [1074, 222]]	58%	[[1107, 189], [995, 301]]	54%	[[1169, 127], [574, 722]]	73%
	linear	S2	[[864, 0], [0, 864]]	100%	[[431, 1], [1, 431]]	100%	[[1266, 30], [3, 1293]]	99%	[[1294, 2], [122, 1174]]	95%	[[1287, 9], [1, 1295]]	100%	[[1296, 0], [25, 1271]]	99%
	rbf	S3	[[860, 4], [6, 858]]	99%	[[374, 58], [49, 383]]	88%	[[0, 1296], [0, 1296]]	50%	[[0, 1296], [0, 1296]]	50%	[[130, 1166], [84, 1212]]	52%	[[503, 793], [181, 1115]]	62%
Above rated	rbf	S0	[[789, 75], [123, 741]]	89%	[[272, 160], [149, 283]]	64%	[[0, 1296], [0, 1296]]	50%	[[0, 1296], [0, 1296]]	50%	[[119, 1177], [61, 1235]]	52%	[[0, 1296], [0, 1296]]	50%
	poly (2)	S1	[[740, 124], [126, 738]]	86%	[[328, 104], [117, 315]]	74%	[[1287, 9], [1271, 25]]	51%	[[1296, 0], [1296, 0]]	50%	[[1200, 96], [1030, 266]]	57%	[[1293, 3], [1293, 3]]	50%
	linear	S2	[[862, 2], [0, 864]]	100%	[[431, 1], [0, 432]]	100%	[[1294, 2], [2, 1294]]	100%	[[1228, 68], [177, 1119]]	91%	[[1294, 2], [0, 1296]]	100%	[[1285, 11], [16, 1280]]	99%
	rbf	S3	[[864, 0], [0, 864]]	100%	[[252, 180], [126, 306]]	65%	[[0, 1296], [0, 1296]]	50%	[[0, 1296], [0, 1296]]	50%	[[76, 1220], [33, 1263]]	52%	[[0, 1296], [0, 1296]]	50%

Finally, the performance of the SVM-based classifiers on the combination D2-S3 are reported in Table D.4. The metrics, and the test sets are as the ones of Section C.2 and Section 6.5, respectively, however the reader is referred to Paper 3 for their naming convention. At below rated conditions, the SVM model implementing a rfb kernel scores satisfactory accuracy (above 90%), detection rate (of about 90%) and alarm rate (below 10%) on the Te33 and Te4 tests set.

Table D.4: SVM classifiers performance on the dataset D2, sensor set up S3.

	Te33			Te1			Te2			Te3			Te4		
	acc	*TDR*	*FDR*	*acc*	*TDR*	*FDR*	*acc*	*TDR*	*FDR*	*acc*	*TDR*	*FDR*	*acc*	*TDR*	*FDR*
Below rated	94%	93%	6%	53%	42%	36%	-	-	-	64%	65%	37%	91%	87%	4%
Above rated	78%	77%	22%	53%	86%	79%	-	-	-	60%	86%	67%	52%	43%	39%

D.3 Random forest

The results of the random forest (RF) classifier are reported in Table D.5. The metrics, and the test sets are as the ones of Section C.2 and Section 6.5, respectively, however the reader is referred to Paper 3 for their naming convention.

Table D.5: RF classifier performance for the datasets D1, D2 and D3, in combination with the sensor S0, S1, S2 and S3.

	Dataset	Sensor	Te33			Te1			Te2			Te3			Te4	
			acc	*TDR*	*FDR*	*acc*	*TDR*	*FDR*	*acc*	*TDR*	*FDR*	*acc*	*TDR*	*FDR*	*acc*	*TDR*
Below rated	D1	S0	85%	86%	14%	-	-	-	63%	50%	40%	69%	62%	26%	72%	67%
	D2	S0	91%	96%	4%	57%	73%	41%	-	-	-	68%	69%	8%	80%	90%
	D3	S0	88%	96%	3%	-	-	-	-	-	-	73%	67%	37%	82%	73%
	D0	S1	96%	96%	4%	66%	73%	41%	80%	69%	8%	76%	90%	38%	84%	71%
		S2	96%	96%	3%	68%	73%	38%	81%	71%	9%	78%	92%	36%	85%	72%
		S3	95%	94%	3%	82%	76%	13%	86%	77%	5%	90%	89%	9%	91%	84%
	D2	S3	97%	96%	2%	82%	85%	21%	-	-	-	91%	92%	11%	96%	93%
Above rated	D1	S0	85%	67%	32%	-	-	-	63%	34%	20%	69%	75%	40%	72%	29%
	D2	S0	91%	73%	19%	57%	94%	84%	-	-	-	68%	92%	68%	80%	79%
	D3	S0	88%	70%	25%	-	-	-	-	-	-	73%	70%	35%	82%	84%
	D0	S1	96%	90%	9%	66%	99%	51%	80%	41%	0%	76%	100%	33%	84%	44%
		S2	96%	91%	8%	68%	98%	51%	81%	34%	0%	78%	99%	26%	85%	39%
		S3	95%	91%	10%	82%	96%	36%	86%	42%	3%	90%	98%	25%	91%	60%
	D2	S3	97%	92%	6%	82%	97%	29%	-	-	-	91%	98%	23%	96%	88%