UNIVERSITAT POLITÈCNICA DE CATALUNYA

Programa de Doctorat:

AUTOMÀTICA, ROBÒTICA I VISIÓ

Tesi Doctoral

# DETECTION AND DIAGNOSIS OF FAULTS AND DAMAGE IN WIND TURBINES

Bryan Joao Puruncajas Maza

Directors:
Dra. Yolanda Vidal
Dr. Christian Tutivén

Març de 2023

*La mejor forma de predecir el futuro es crearlo.*
*Peter Drucker*

# List of papers

This thesis is structured as a compendium of publications presented in Part II as appended documents. Contributions have first been separated in the areas of structural health monitoring (SHM) and condition monitoring (CM). Then, each contribution has been classified as article paper (AP), which is part of the PhD thesis compendium, or related paper (RP) as follows. Note that the first two contributions (AP-I, AP-II) belong to the SHM research field, whereas the latter two (AP-III, AP-IV) pertain to the CM field. Finally, note that the quartile ranking of each journal is given according to the journal citation reports (JCR).

- **AP-I:** Puruncajas, B., Vidal, Y., & Tutivén, C (2020). Vibration-response-only structural health monitoring for offshore wind turbine jacket foundations via convolutional neural networks. Sensors, 20(12), 3429.

  First quartile, Q1, `https://doi.org/10.3390/s20123429`.

  Number of article citations (Wos Core Collection): 23

- **AP-II:** Baquerizo, J., Tutivén, C., Puruncajas, B., Vidal, Y., & Sampietro, J. (2022). Siamese neural networks for damage detection and diagnosis of jacket-type Offshore wind turbine platforms. Mathematics, 10(7), 1131.

  First quartile, Q1, `https://doi.org/10.3390/math10071131`.

  Number of article citations (Wos Core Collection): 0

- **AP-III:** Encalada-Dávila, Á., Puruncajas, B., Tutivén, C., & Vidal, Y. (2021). Wind turbine main bearing fault prognosis based solely on SCADA data. Sensors, 21(6), 2228.

  First quartile, Q1, `https://doi.org/10.3390/s21062228`.

  Number of article citations (Wos Core Collection): 24

  This article has been awarded the editor's choice article distinction. Editor's choice articles are based on recommendations by the scientific editors that select a small number of articles that they believe will be important in the respective research area.

- **AP-IV:** Encalada-Dávila, Á., Moyón, L., Tutivén, C., Puruncajas, B., & Vidal, Y. (2022). Early fault detection in the main bearing of wind turbines based on gated

recurrent unit (GRU) neural networks and SCADA data. IEEE/ASME Transactions on Mechatronics.

First quartile, Q1, `https://ieeexplore.ieee.org/document/9823312`.

Number of article citations (Wos Core Collection): 3

- **RP-I:** Puruncajas, B., Vidal, Y., & Tutivén, C. Damage detection and diagnosis for offshore wind foundations. In proceedings of the 17th international conference on informatics in control, automation and robotics (ICINCO), Paris, France, July 7 - July 9, 2020.

- **RP-II:** Prime archives in sensors, book chapter: Puruncajas, B.; Vidal, Y.; Tutivén, C. Damage detection and identification in jacket-type support structures via deep learning and accelerometer sensors data, 2020, edited by Xin-Wei Yao.

- **RP-III:** Tutivén, C., Baquerizo, J., Vidal, Y., Puruncajas, B., & Sampietro, J. Offshore wind turbine jacket damage detection via a siamese neural network. In european workshop on structural health monitoring (EWSHM), Palermo, Italy, July 4 - July 7, 2022.

- **RP-IV:** Encalada-Dávila, Á., Tutivén, C., Puruncajas, B., & Vidal, Y. (2021). Wind turbine multi-fault detection based on SCADA data via an autoencoder. Renewable energy and power quality journal, 19, 487-492.

- **RP-V:** Puruncajas, B., Alava, W., Encalada-Dávila, Á., Tutivén, C., & Vidal, Y. (2021). Convolutional neural network for wind turbine failure classification based on SCADA data. Renewable energy and power quality journal, 19, 447-451.

- **RP-VI:** Encalada-Dávila, Á., Puruncajas, B., Tutivén, C., & Vidal, Y. Fault prognosis for wind turbines' main bearing based on SCADA data. In 2021 10th international conference on structural health monitoring of intelligent infrastructure: Advanced research and real-world application, Porto, Portugal, June 30 - July 2, 2021.

- **RP-VII:** Tutivén, C., Benalcazar–Parra, C., Encalada-Dávila, A., Vidal, Y., Puruncajas, B., & Fajardo, M. Wind turbine main bearing condition monitoring via convolutional autoencoder neural networks. In 2021 international conference on electrical, computer, communications and mechatronics engineering (ICECCME), Flic-en-Flac, Mauritius, October 7 - October 8, 2021.

- **RP-VIII:** Encalada-Dávila, A., Tutivén, C., Moyón, L., Puruncajas, B., & Vidal, Y (2022). Main bearing fault prognosis in wind turbines based on gated recurrent unit neural networks. Renewable energy and power quality journal, 20, 324 - 360.

# Preface

This doctoral thesis is focused on the detection and diagnosis of faults and damage in wind turbines. It was developed in the framework of the automatic control, robotics and vision doctoral programme of the Universitat Politècnica de Catalunya (UPC, Barcelona, Spain). The study was carried out between October 2019 and January 2023 in collaboration with the Escuela Superior Politécnica del Litoral (ESPOL, Guayaquil, Ecuador), thanks to the opportunity provided by my PhD advisors, Dr. Yolanda Vidal (UPC) and Dr. Christian Tutivén (ESPOL). The thesis is presented as a collection of articles.

# Acknowledgments

First of all, I would like to thank UPC for the opportunity of this Ph.D and ESPOL for the help and facilities provided throughout my doctoral study.

Secondly, my most sincere thanks to my thesis supervisors. Thank you for your time, patience, and commitment to the work carried out throughout the thesis. Although there are no written words that can express how grateful I am, thank you very much.

Also, I cannot dedicate enough words to my family, my father Raul Puruncajas, my mother Patricia Maza, my brother Patricio Puruncajas, my sisters Jhoanna Puruncajas and Narcisa Piedra, my uncles, Medardo Maza, Bernarda Maza and Armando Maza, who have supported me in these four years of study. And especially, this last year to my girlfriend Andrea Martinez, who has always supported me to finish my thesis, thank you very much.

Last but not least, I am grateful for all the privileges and luxuries, including this work, that has been a part of my life for the past four years.

Bryan Puruncajas.

# Abstract

In recent decades, large amounts of fossil fuels (coal, oil, gasoline, and natural gas) have been used in energy production, generating an environmental impact and global warming. This unsustainable situation has led to the search for alternative energy sources with low environmental impact. In recent years, wind power, which has been presented as a reliable and emission-free energy source, has become the fastest growing renewable source worldwide. According to the Global Wind Report 2021 of the Global Wind Energy Council (GWEC), the wind industry in 2020 was growing by 13% compared to the previous year. Furthermore, the recent war between Russia and Ukraine has increased the urgency of Europe to transform its energy system to end its dependence on fossil fuels, which prompted the European Commission (EC) to further increase its support for the development and deployment of renewable energy in general, and wind energy in particular (since the EC requires 50% of the energy mix of the European Union in 2050).

Wind turbines are huge structures (e.g., rotor diameters higher than 120 m are common), with technological and industrial development on land and at sea. These structures are located in remote areas of difficult access with extreme environmental conditions such as high and low temperatures, as well as stochastic wind speed, which cause different operating regions in the wind turbines. All of these factors greatly increase the cost of operation and maintenance (O&M). In addition, low reliability directly reduces the availability of wind power due to turbine downtime. For an onshore wind farm the O&M is estimated between 20-25% of the levelized cost of energy (LCOE), while for offshore ones it can account for up to 30% of the LCOE. Although larger turbines can reduce the O&M cost per unit of power, the cost of failure and damage increases. Therefore, it is crucial to create and apply strategies for early damage and failure detection. On the one hand, structural health monitoring (SHM) addresses wind turbine structural damage detection by monitoring parameters indicative of the condition of the examined structure. On the other hand, condition monitoring (CM) addresses the detection of faults in rotating components of the wind turbine, such as shafts, bearings, gearbox, or generator, among others. The main objective is to identify any abnormal change in condition or specific events that may indicate an incipient damage/failure. This thesis states different techniques for wind turbine SHM (first part of the thesis) and CM (second part of the thesis).

The first overall objective of this doctoral thesis is to propose and validate different SHM strategies for the detection and diagnosis of damage in jacket-type offshore support. The development of the SHM system uses a laboratory down-scaled jacket-type support, positioning this research at the technology readiness level (TRL) 4 (technology validated

in a lab). The main contributions in this area correspond to (i) the conversion of the vibration signals to multichannel grayscale images with as many channels as the number of sensors, (ii) the experimentation is carried out for all the operating regions of the wind turbines, (iii) the comparison between convolutional neural networks and Siamese networks is made, where in the first case an increase in data is presented, and for the second case a small amount of data is used, and (iv) its viability is demonstrated through different metrics.

It may be surprising that many wind turbines do not have specific CM systems. Currently, CM systems are installed on new wind turbines on the first day of operation. However, many older wind turbines do not have these types of system installed that involve high-frequency sensors specifically tailored to monitor the condition of the rotating machinery in the drive train. In wind turbines close to their life expectancy, a technological update (to install newly specific tailored sensors for condition monitoring) is not economically viable for the owners. This is relevant because 38 GW of wind farms in Europe are expected to reach their life expectancy in the next five years. Based on current trends, it is estimated that around 2.4 GW will be decommissioned for re-powering and 7 GW will be fully decommissioned. The remaining 29 GW will continue to operate and will be considered for lifetime extension services. With this background, the second overall objective of this thesis is to develop early fault prediction strategies based only on real SCADA data (available in all industrial-sized wind turbines). This is a cost-effective approach that does not require the installation of expensive sensors. This research is placed at the TRL 7 level (system prototype demonstration in an operational environment). The main contributions are: (i) Early detection is achieved using only SCADA data. (ii) The proposed methods only require healthy data to be developed; therefore, they can be applied to any wind farm, even where no faulty data have been recorded. (iii) Normality models that avoid the imbalance problem in the data are developed and contrasted using artificial neural networks and gated recurrent unit (GRU) neural networks. And (iv) the validity and performance of the established methodologies are demonstrated in a real in-production wind farm. The results obtained show that early detection strategies based solely on SCADA data can predict failures several months before fatal failure occurs and, thus, allow wind turbine operators to plan their maintenance operations.

# Resumen

En las últimas décadas se han utilizado grandes cantidades de combustibles fósiles (carbón, petróleo, gasolina y gas natural) en la producción de energía, generando un impacto ambiental y el calentamiento global. Esta situación insostenible ha llevado a la búsqueda de fuentes de energía alternativas con bajo impacto ambiental. En los últimos años, la energía eólica, se ha presentado como una fuente de energía fiable y libre de emisiones, convirtiendose en la fuente renovable de mayor crecimiento a nivel mundial. Según el Global Wind Report 2021 del Global Wind Energy Council (GWEC), la industria eólica en 2020 creció un 13% respecto al año anterior. Además, la reciente guerra entre Rusia y Ucrania ha aumentado la urgencia de Europa de transformar su sistema energético para acabar con su dependencia de los combustibles fósiles, lo que llevó a la Comisión Europea (CE) a aumentar aún más su apoyo al desarrollo y despliegue de energías renovables, y la energía eólica en particular (ya que la CE demanda el 50% del mix energético de la Unión Europea en 2050).

Las turbinas eólicas son estructuras enormes (por ejemplo, los diámetros de rotor superiores a 120 m son comunes), con un desarrollo tecnológico e industrial en tierra y mar. Estas estructuras están ubicadas en áreas remotas de difícil acceso, con condiciones ambientales extremas como altas y bajas temperaturas, así como velocidades estocásticas del viento, que provocan diferentes regiones de operación en los aerogeneradores. Todos estos factores aumentan en gran medida el costo de operación y mantenimiento (O&M). Además, la baja confiabilidad reduce directamente la disponibilidad de energía eólica debido al tiempo de inactividad de la turbina. Para un parque eólico terrestre, el O&M se estima entre un 20-25% del costo nivelado de la energía (LCOE), mientras que para los marinos puede suponer hasta un 30% del LCOE. Aunque las turbinas más grandes pueden reducir el costo de operación y mantenimiento por unidad de potencia, el costo de fallos y daños aumenta. Por lo tanto, es crucial crear y aplicar estrategias para la detección temprana de daños y fallos. Por un lado, el monitoreo de la salud estructural (SHM) aborda la detección de daños estructurales en aerogeneradores mediante el monitoreo de parámetros indicativos del estado de la estructura examinada. Por su parte, el monitoreo de la condición (CM) se ocupa de la detección de fallos en los componentes giratorios del aerogenerador, como ejes, rodamientos, caja de cambios o generador, entre otros. El objetivo principal es identificar cualquier cambio anormal en la condición o eventos específicos que puedan indicar un daño/fallo incipiente. Esta tesis establece diferentes técnicas para aerogeneradores en SHM (primera parte de la tesis) y CM (segunda parte de la tesis).

El primer objetivo general de esta tesis doctoral es proponer y validar diferentes es-

trategias de SHM para la detección y diagnóstico de daños en cimentaciones marinas tipo chaqueta. El desarrollo del sistema SHM utiliza una cimentación tipo chaqueta a escala en laboratorio, lo que posiciona esta investigación en el nivel de preparación tecnológica (TRL) 4 (tecnología validada en un laboratorio). Las principales contribuciones en esta área corresponden a: (i) La conversión de las señales de vibración a imágenes multicanal en escala de grises con tantos canales como sensores. (ii) La experimentación se realiza para todas las regiones de operación de los aerogeneradores. (iii) Se realiza la comparación entre redes neuronales convolucionales y redes neuronales siamesas, donde en el primer caso se presenta un aumento de datos, y para el segundo caso se utiliza una pequeña cantidad de datos. Y (iv) se demuestra su viabilidad a través de diferentes métricas.

Puede resultar sorprendente que muchas turbinas eólicas no tengan sistemas CM específicos. Actualmente, los sistemas CM se instalan en los aerogeneradores nuevos el primer día de funcionamiento. Sin embargo, muchas turbinas eólicas más antiguas no tienen instalados estos tipos de sistemas que involucran sensores de alta frecuencia diseñados específicamente para monitorear la condición de la maquinaria giratoria en el tren de transmisión. En turbinas eólicas cercanas a su vida útil, una actualización tecnológica (instalar nuevos sensores especificos personalizados para el monitoreo de condición) no es económicamente viable para los propietarios. Esto es relevante porque se espera que 38 GW de parques eólicos en Europa alcancen su esperanza de vida en los próximos cinco años. Según las tendencias actuales, se estima que alrededor de 2,4 GW se repotenciarán y 7 GW se desmantelarán por completo. Los 29 GW restantes continuarán operando y se considerarán para servicios de extensión de vida. Con estos antecedentes, el segundo objetivo general de esta tesis es el desarrollo de estrategias de predicción temprana de fallos basados únicamente en datos SCADA reales (disponibles en todos los aerogeneradores de tamaño industrial). Este es un enfoque rentable que no requiere la instalación de sensores costosos. Esta investigación se sitúa en el nivel TRL 7 (demostración del prototipo del sistema en un entorno operativo). Las principales contribuciones son: (i) La detección temprana se logra utilizando solo datos SCADA. (ii) Los métodos propuestos solo requieren que se desarrollen datos saludables; por lo tanto, se pueden aplicar a cualquier parque eólico, incluso donde no se hayan registrado datos defectuosos. (iii) Los modelos de normalidad que evitan el problema de desequilibrio en los datos se desarrollan y contrastan utilizando redes neuronales artificiales y redes neuronales de unidades recurrentes (GRU). Y (iv) se demuestra la validez y rendimiento de las metodologías establecidas en un parque eólico real en producción. Los resultados obtenidos muestran que las estrategias de detección temprana basadas únicamente en datos SCADA pueden predecir fallos varios meses antes de que ocurra una fallo fatal y, por lo tanto, permitir que los operadores de aerogeneradores planifiquen sus operaciones de mantenimiento.

# Table of contents

List of papers **v**

Preface **vii**

Abstract **ix**

Resumen **xi**

List of figures xv

List of tables xv

List of acronyms xix

Chapter 1 – Introduction 3
1.1 Motivation . . . 4
1.2 Aim and scope . . . 7
1.3 Outline . . . 11

Chapter 2 – Structural Health Monitoring 15
2.1 The Wind Turbine . . . 15
2.1.1 Structural parts in a wind turbine . . . 17
2.2 SHM state of the art for offshore wind turbine substructures . . . 20
2.2.1 Overview of sensor technologies for SHM in general . . . 24
2.2.2 Overview of sensor technologies for SHM in offshore structures . . . 26
2.2.3 Vibration-based SHM methodologies . . . 28
2.3 Comparison of the strategies stated in AP-I and AP-II . . . 36

Chapter 3 – Condition Monitoring 43
3.1 State of the art in wind turbine CM . . . 43
3.1.1 Specific CM strategies . . . 45
3.1.2 CM methodologies based on SCADA data . . . 48
3.2 Comparison of the strategies stated in AP-III and AP-IV . . . 50

Chapter 4 – Conclusions and future research 59
4.1 Conclusions . . . 59
4.2 Further research . . . 60

References 61

**Part II** **71**

Appended Paper I 73

Appended Paper II 95

Appended Paper III 117

Appended Paper IV 141

# List of figures

Figure 1.1 Thesis framework. AP-I refers to the first appended paper of the compendium of publications, AP-II to the second, and so on. . . . . 8
Figure 2.1 Basic structure and main components of the WT. . . . . 16
Figure 2.2 Wind farm classification . . . . . 16
Figure 2.3 Principal components inside the nacelle of a conventional WT. . . 17
Figure 2.4 Structural components of a fixed offshore WT. . . . . 19
Figure 2.5 Types of substructures-foundations of offshore wind farms. . . . . 21
Figure 2.6 SHM flowchart process. . . . . 22
Figure 2.7 Deep-Learning architectures. . . . . 33
Figure 2.8 Framework of AP-I and AP-II . . . . . 37
Figure 2.9 CNN architecture used in AP-I . . . . . 39
Figure 2.10 SNN architecture of a convolutional layer (model 1) used in AP-II 40
Figure 2.11 SNN architecture of two convolutional layers (model 2) used in AP-II 40
Figure 3.1 Structure of WT drive train with seven acceleration transducers installed for CM. . . . . 46
Figure 3.2 Evolution of gearbox damage in terms of wear particle size and concentration. . . . . 47
Figure 3.3 Framework of AP-III and AP-IV. . . . . 51
Figure 3.4 WT2 (WT number 2 in the wind farm) data for training and test of AP-III and AP-IV methodologies. . . . . 52
Figure 3.5 ANN of the AP-III method. . . . . 52
Figure 3.6 Many-to-one GRU architecture of the AP-IV method. . . . . 53
Figure 3.7 ANN indicator values (blue line) for test data, and threshold (red line) in the AP-III method. . . . . 54
Figure 3.8 EWMA on residual errors for the WT's test dataset (using 1008 spans), where the red line represents a fault warning and the green line indicates a definite fault. . . . . 57

# List of tables

Table 2.1 Number of images in AP-I and AP-II . . . . . 38
Table 2.2 Number of parameters in each layer of the CNN architecture of the AP-I. The neural network has a total of 2,176,308 parameters. . . . . 39
Table 2.3 Number of parameters in each layer of the SNN model 1 architecture of the AP-II. The neural network has a total of 81,283,488 parameters. . 39
Table 2.4 Number of parameters in each layer of the SNN model 2 architecture of the AP-II. The neural network has a total of 40,725,024 parameters. . 41
Table 2.5 Number of images for training and validation in AP-I . . . . . 41
Table 2.6 Number of images for training, validation and test in AP-II . . . . 41
Table 2.7 Metrics for each label of the diagnosis and detection problem and comparison between the datasets in AP-I . . . . . 41
Table 2.8 Metrics for the diagnosis and detection problem and comparison between models in AP-II . . . . . 42
Table 2.9 Accuracy, validation error, training error, and training time when using data augmentation and initial data set in AP-I . . . . . 42
Table 2.10 Accuracy, validation error, and training error when using model 1 and model 2 in AP-II . . . . . 42
Table 3.1 Selected SCADA variables used to develop the normality model in AP-III and AP-IV, its description, range of possible values and units. . . 51
Table 3.2 Number of parameters in each layer of the ANN architecture for AP-III. The neural network has a total of 1,153 parameters . . . . . 53
Table 3.3 Number of parameters in each layer of the GRU architecture for AP-IV. The neural network has a total of 52,353 parameters. . . . . 54
Table 3.4 Alarm activations over the test dataset for the AP-III methodology. 55
Table 3.5 False-positives alarms (x-mark) over the training and validation datasets in the AP-IV method. . . . . 56
Table 3.6 Alarm activations over the test dataset for the AP-IV methodology. 56
Table 3.7 Accuracy, precision, specificity, recall, F1-score and training time of the AP-III and AP-IV methodologies. . . . . 56

# List of acronyms

**AE** acoustic emission

**AI** artificial intelligence

**ANFIS** adaptive neuro-fuzzy inference system

**ANNs** artificial neural networks

**AP** article paper

**AR** auto regressive

**AU** acousto-ultrasonics

**CM** condition monitoring

**CMS** condition monitoring system

**CNN** convolutional neural network

**CNNs** convolutional neural networks

**DBNs** deep Belief Networks

**DIC** digital image correlation

**DL** deep learning

**EOC** environmental and operational conditions

**ER** electrical resistance

**EWMA** exponential weighted moving average

**FBG** fiber bragg gratings

**FEM** finite element method

**FNNs** fuzzy neural networks

**GANs** generative Adversarial Networks

**GPR** ground penetrating radar

**GRU** gated recurrent unit

**IRT** infrared thermography

**LCOE** levelized cost of energy

**LiDAR** light detection and ranging

**LMs** laser testing methods

**LSTMs** long short term neural networks

**LT** leak testing

**MEMS** micro-electromechanical systems

**MFL** magnetic flux leakage

**ML** machine learning

**MLPs** multilayer perceptrons

**MT** magnetic particle testing

**NDTE** non-destructive testing and evaluation

**OF** optical fiber

**PCA** principal component analysis

**PdM** predictive maintenance

**PNNs** probabilistic neural networks

**PT** fluid penetrant testing

**PZT** lead-zirconium-titanium oxide

**RNNs** recurrent Neural Networks

**RP** related paper

**RT** radiographic testing

**SCADA** supervisory control and data acquisition

**SHM** structural health monitoring

**SNN** siamese neural network

**SNNs** siamese neural networks

**SOMs** self-organizing maps

**SVD** singular value decomposition

**SVDD** support vector data description

**SVMs** support vector machines

**UPNs** Unsupervised Pretrained Networks

**UT** ultrasonic testing

**VT** visual testing

**WT** wind turbine

Part I

# Chapter 1

# Introduction

Economic growth and industrialization have led to an increase in the use of fossil fuels. The growing production and consumption of fossil fuels has had several adverse environmental impacts on the planet, including global warming, air pollution, and increased health risks. Fossil fuels are expected to continue to dominate the energy supply for the foreseeable future because of their higher energy density and the time it takes to innovate. Environmentally related technological progress can lead to significant reductions in energy consumption and increased energy efficiency. These technologies can help reduce the negative effects of energy use and encourage nations to rethink how energy is consumed in all activities [1]. Renewable energy is a type of energy technology that uses nature as an energy source and is found in almost all the planet, unlike fossil fuels, which are concentrated in specific countries. The use of renewable energy democratizes the use of energy and can help resolve complex situations such as Europe's current gas crisis with Russia due to the war in Ukraine. Additionally, it brings considerable benefits from a consumer, environmental, and economic perspective. According to WindEurope, the European Commission forecasts show that renewable energy-based electricity will be essential to achieve climate neutrality in Europe by 2050 [2].

Renewable energies favor environmental sustainability towards a more desirable nature-climate balance. They include solar energy, wind, gravity-falling water, heat from the earth (geothermal), plant materials (biomass), ocean waves and currents, temperature differences in the oceans, and tide energy [3]. Among these sources, wind as an essential source of global energy has taken a leading position among renewable sources. This source can be considered one of the cleanest forms of energy generation.

The wind energy sector has grown significantly in the last two decades due to its great opportunity for future electricity production. Consequently, installed wind capacity has increased by more than 30% per year [4]. However, two trends endanger the fulfillment of this global role. The first corresponds to the aging of existing wind turbine (WT)s, which have already reached their estimated useful life of 20 years, which requires additional maintenance [5]; and the second, the technological evolution to ever larger WTs in remote offshore locations, which poses new accessibility problems for inspection and maintenance

[6]. For wind power, the levelized cost of energy (LCOE) represents the sum of all costs of a fully operational wind power system over the life of the project, with financial flows discounted to a common year. The main components of the LCOE of wind energy systems include capital costs, operation, and maintenance costs, and expected annual energy production [7]. The development of effective strategies for the operation and maintenance (O&M) of WTs (both onshore and offshore wind farms) must be addressed to keep wind power financially viable, reducing LCOE costs.

## 1.1 Motivation

From an engineering point of view, WTs are made up of the following:

(i) static load components.

(ii) moving/rotating parts, necessary to harness the kinetic energy of the wind and convert it into electricity.

The elements in (i) are generally called support structures. The components of (ii) can be further divided into slow-rotating elements (blades) and high-speed rotating mechanisms. The latter are all included in the rotor-gondola assembly. These distinctions are essential, as blades and supporting structures are applications for structural health monitoring (SHM), while condition monitoring (CM), by the standard definition of the term, is a related term to machinery and rapidly moving components, such as gears, bearings in the gearbox, and generator [8]. Elements in (i) and (ii) are subject to natural use and consumption and therefore can develop long-term and short-term damage. Damage to the external structure (i) will cause collapse (partial and localized or global), while damage to internal mechanisms (ii) will cause faults, interruptions in energy production, and, above all, fires, explosions, oil leaks, or other events. In this thesis, the word damage will be used to refer to an issue in the structural part of the wind turbine (WT), while the term faults will be used for problems related to the internal nacelle components (bearings, shafts, gearbox, generator, etc.).

WT faults and damage are difficult to detect due to the nature of their evolution and variable operating conditions. Therefore, maintenance is in demand to reduce operation and costs. There are three types of maintenance: corrective, preventive, and predictive. In corrective maintenance, equipment is used to its limit, and repairs are made when components are damaged or failed. Preventive maintenance is performed at a regular rate to prevent damage or faults; the problem with this maintenance is that if maintenance is planned too early, it wastes valuable life, raising expenditures. If damage or faults can be early predicted (months before the fatal breakdown occurs), maintenance can be scheduled in advance, which would help managers bridge the gap between corrective and preventive maintenance by performing maintenance not too late and not too early, but just in time. This fits the definition of predictive maintenance. Predictive maintenance can help to detect problems in the equipment in advance, and help identify which parts need to be repaired [9].

A system that allows for a significant increase in profitability, reliability, and sustainability in WTs is SHM system. Some of the most well-known roots of structural damage are moisture absorption, fatigue, wind gusts, thermal stress, corrosion, fire, and lightning. SHM data can be used to prevent unnecessary replacement of components and unexpected catastrophic damage, minimize inspection times, keep the wind power supply chain running smoothly and reliably, support the further development of a WT, and provide remote side monitoring and remote diagnostics [10].

SHM techniques have been developed over many years. In general, vibration-based SHM methods are the most widely adopted. The goals of these SHM methods are to determine the existence of structural damage, identify its location and severity, assess the safety of the structure, predict the remaining service life of the structure, and make decisions about maintenance strategy, if possible. The vibration characteristics of a structure are a function of its physical parameters. Structural damage causes a change in the physical parameters of the structure, and therefore the change in the physical parameters can be used as an indicator of the state of health of the structure. Through the signal monitored by the sensors installed on the structure, the vibration characteristics can be extracted, and the corresponding change can be detected and analyzed. In addition, from the change of vibration characteristics, the change in the physical parameters of the structure can be further obtained to diagnose the health status of the structure. Recently, with the rapid development of modern computer technology, the progress of sensor technology, and signal processing technology, test signals can be accurately and quickly analyzed and processed [11].

Generally, vibration-based SHM techniques can be classified into three categories: frequency domain, time domain, and time-frequency domain approaches [12]. The time domain approach does not require as many computational resources for property extraction as frequency-related approaches do, and thus saves time. Among the dynamic responses of a structure in the time domain, acceleration is one of the most used, and it can also be used directly in SHM systems. However, the time history data set for acceleration is much larger than the frequency data set, so convergence is difficult to achieve [11]. In recent years, deep learning has proven to be a powerful tool for SHM applications, to the point that it has made inroads into almost every damage level of SHM. It has achieved structural damage detection (level 1), damage localization (level 2), damage quantification (level 3), and damage classification (level 4). The evaluation of structural integrity (level 5) remains a concept for SHM technologies to this day [13]. SHM has matured as a routine practice in aerospace structures for many years. The approach is still being considered for routine applications for civil infrastructure systems. A very small number of studies have reported on WT structures, and most of them have focused on blade control [14]. The field of WTs is more complicated due to the existence of input excitations (wind, waves, currents), which cannot be imposed. There are several challenges with the SHM of WTs, such as (1) difficulties in inspection and maintenance in the field due to the height of the structure, the operation of the blades, (2) the effect of environmental conditions, especially in high seas, and (3) the remote site of wind farms.

Deep learning can be applied as a supervised or unsupervised learning problem. In the

unsupervised approach, there is no need to collect damaged data, that is, the machine learning algorithm is trained with only healthy data. However, it is unfeasible to correctly identify different states of damage using only the data obtained during a healthy state. In this framework, detection is applicable, but not a classification according to the type of damage. On the other hand, the supervised approach addresses, in addition to detection, classification. Therefore, it requires data from the damaged structure for model training. In practice, this is achieved by computer models such as the finite element method (FEM) and the use of experimental scale models. Therefore, the first motivation of the thesis is to study supervised models for the detection and diagnosis (levels 1 and 4) of different structural damage to the supports of the WT jackets using a laboratory down-scaled model.

Over the past decade, there has been a rapid rise in autonomous CM systems to monitor the performance of equipment, including WTs. The CM strategy can be applied based on the vibration sensor system, which has vibration sensors, strain gauges, or oil particle counters matched to turbine subcomponents for localized monitoring. The problem with this CM strategy is the high cost and that commercial options can have a high false alarm rate [15]. On the other hand, modern WTs are equipped with sophisticated supervisory control and data acquisition (SCADA) systems, which broadcast a wealth of information every 10 minutes: details about wind flow and weather conditions, the alignment of the turbine with the wind, on the conversion of the kinetic energy of the wind into active power, on the vibratory and mechanical state of the machine, on the thermal conditions in the relevant parts of the turbines, etc. [16]. WT CM based on the use of SCADA data is considered a promising and cost-effective approach, as SCADA data are available at no additional cost [17]. Currently, it is unavoidable to intervene on-site, and it is inconceivable to do it in a base condition. Consequently, the rule is that WT reviews are periodic and, therefore, have little to do with the onset of possible breakdowns. Based on the above considerations, it would be worth developing SCADA data analysis methodologies that could be useful in assessing the health status of WTs. Therefore, the scientific and technological challenge is to address the complexity of the SCADA data flow and to process it into novel knowledge. It is estimated that a sudden failure of a 1.5 MW WT during winter causes a loss of production of around €50,000. This amount is up to five times greater than the production lost due to a wisely planned maintenance program [16]. This sheds powerful light on the economic impact of CM for the prediction of WT failures.

The most important challenges of the SCADA data-driven approach stem from the abundance of operational data and the wide variety of fault types that are predominantly observable in wind farms. The lack of data with observed failures can result in a strong class imbalance and potentially a lack of representativeness if supervised methods of fault detection based on machine learning are sought. Therefore, it is more feasible to rely on a representation of the normal behavior of the turbine and detect relevant deviations from this behavior.

In recent years, deep learning approaches have been widely used to solve complex problems in various disciplines. Given the powerful learning ability, neural networks can

be trained to obtain models with appreciable accuracy. Furthermore, it does not need expert knowledge of system evaluation, but instead requires a lot of data extracted from the real world system [18]. Given a converged deep learning model, the normality model can be completed instantly and achieve high-fidelity accuracy. Therefore, the second challenge in this thesis is to propose new deep learning-based models to capture the normal behavior of a WT, taking into account that WTs work in various regions of operations under a broad range of wind speed conditions and are affected by seasonality (e.g., temperature variation between different year seasons).

## 1.2 Aim and scope

The purpose of this doctoral study is two-fold. First, contribute solutions to the SHM problem in offshore WT supports, specifically in the jacket-type, for which models based on deep learning will be stated to detect and diagnose damage. Second, related to the area of CM, the challenge of early prediction of main bearing faults is faced using SCADA data. In both challenges, deep learning models are used. The state of the art of SHM in WT supports (Chapter 2, Section 2.2) and the state of the art of CM in WTs (Chapter 3, Section 3.1) allow us to place the contributions of this doctoral thesis in context. The two general objectives attained in the doctoral thesis are described below:

- Objective 1: Develop and validate methodologies for the detection and diagnosis of damage to the structure health of jacket-type supports of offshore WTs.

- Objective 2: Develop and validate early fault detection methodologies based on SCADA data from wind farms based on deep learning techniques. The fault of the main bearing of the WT will be analyzed, for which the methodology will be tested with SCADA data from a wind farm in operation.

The state of the art and the framework addressed in this thesis are shown in Figure 1.1. The state of the art are the gray boxes, while the contributions done in the investigation of SHM and CM are indicated in blue (recall that AP-I refers to the first appended paper, AP-II to the second, and so on). The evolution and order of the investigations are represented by arrows, which represent the timeline of the SHM and CM investigations. Within each step, a review of the literature is carried out to choose the best candidate method most suitable for SHM and CM applications. Subsequently, the feasibility and limitations of the chosen methods are investigated based on the documents attached to this thesis. Depending on the degree of development, an analysis of the feasibility and limitations of the developed methods is carried out. It is important to note that this thesis aims to investigate the feasibility of CM technology from SCADA data and SHM technology for damage identification and detection, to provide practical value. Therefore, the focus is on practical applications rather than theoretical developments. The following paragraphs state the limitations of each proposed methodology.

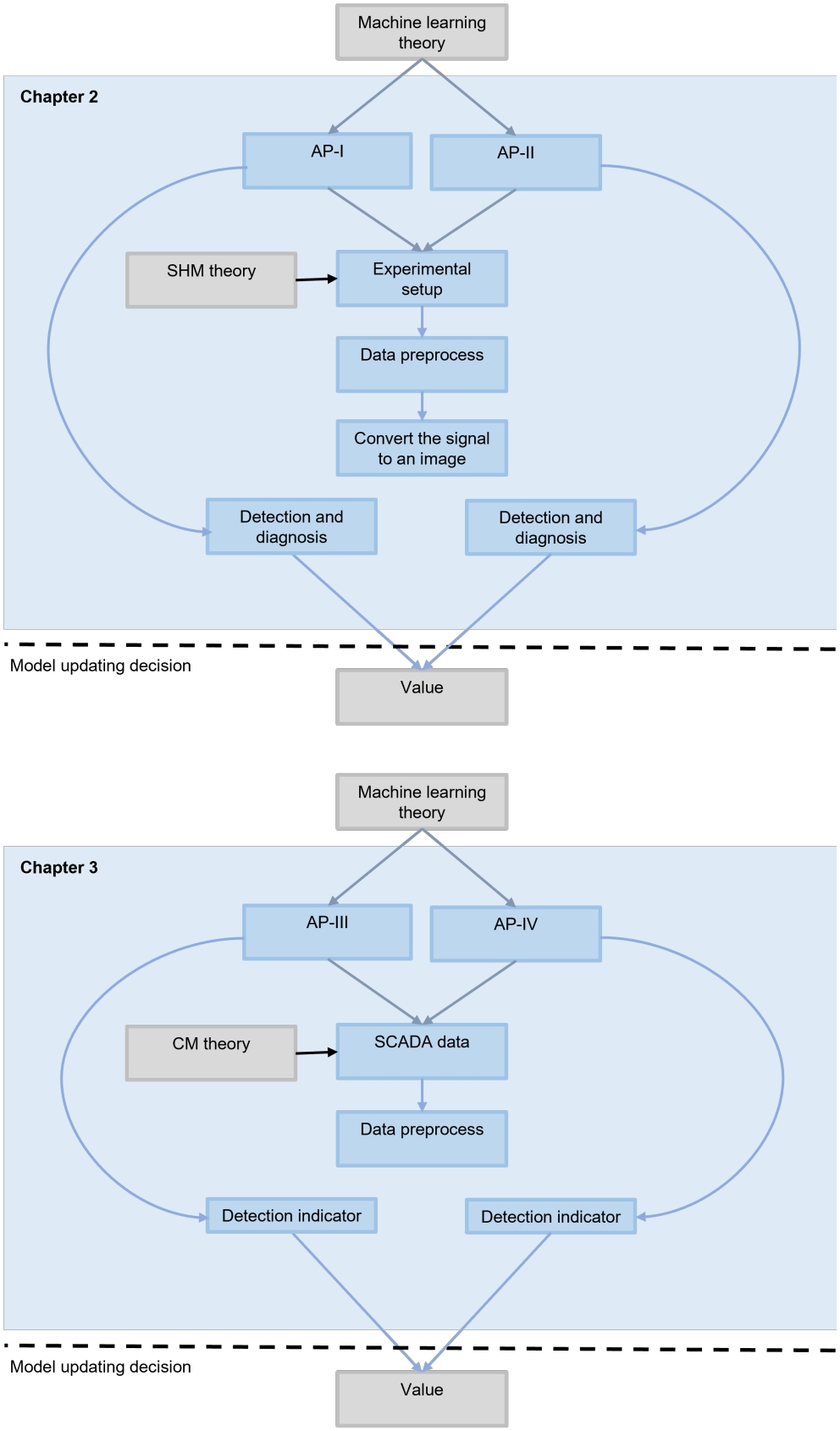



Figure 1.1: Thesis framework. AP-I refers to the first appended paper of the compendium of publications, AP-II to the second, and so on.

The framework proposed in the SHM research is applicable to WT supports in general, however, its validation is limited to jacket supports as they are widely used in practice and can be expected to be used even more in the future when sites are considered with greater water depths. In addition, jacket substructures provide some unique practical challenges that have not yet been adequately addressed in the literature. The SHM methodologies in jacket-type supports are validated and studied from experimentation in a jacket-type support model in a laboratory, so its development in the industrial environment would be pending.

On the other hand, the framework proposed for the investigation of CM from SCADA data is carried out with data from industrial WTs in operation (in particular from two wind farms), thus with a high technology readiness level of 7 (system prototype demonstration in an operational environment). The study is limited to faults of the main bearing, due to its high cost and long replacement time. Here, it is important to understand that the terms fault and failure have different meanings in the context of reliability engineering. A fault is a defect or imperfection in a system or component that may or may not cause a failure. On the other hand, a failure occurs when a system or component no longer performs its intended function. Thus, a fault may or may not lead to a failure, and a failure may or may not be preceded by a fault. In this thesis, early fault detection is addressed by using only SCADA data.

This thesis provides answers to several pertinent research questions that are not only crucial to achieving the objectives of SHM and CM as stated in the PhD thesis, but also to the particular data-based methodologies and artificial intelligence (AI) used to achieve them. The specific research questions addressed in this thesis are as follows:

- How can a deep learning model be created using minimal data?
- How can imbalanced datasets be effectively handled?
- How can time series data be efficiently processed?

The research questions are investigated in the four papers summarized and discussed in Chapters 2 to 4. The articles and their relationship to the specific boxes of the CM and SHM framework are indicated in Figure 1.1. Specifically, AP-I and AP-II documents are related to the detection and diagnosis of damage. The AP-III and AP-IV papers establish a methodology for the early detection of main bearing faults from SCADA data. In the following paragraphs, a brief summary of each AP paper is given to highlight its main contributions.

The AP-I titled *Vibration-Response-Only Structural Health Monitoring for Offshore WT Jacket Foundations via Convolutional Neural Networks*, uses damage data from a jacket-type support of offshore WTs. Four possible cases are addressed: replica bar, bar with loose bolt, bar with fracture, and healthy bar, relying solely on the vibrational response collected through accelerometers to monitor the structural health of the support. The vibratory excitation is obtained from white noise with four different amplitudes. The data collected from the sensors (accelerometers) are converted into multichannel grayscale images with as many channels as sensors. A 25200% data augmentation strategy

is presented to decrease the classification error of the deep convolutional neural network used to classify the images. The results show metrics higher than 99% in all metrics (accuracy, precision, recall, and F1-score), where 1196 minutes were needed to train the convolutional neural network.

The AP-II titled *Siamese Neural Networks for Damage Detection and Diagnosis of Jacket-Type Offshore WT Platforms* uses the same data as the AP-I study, but with a methodology based on two cascaded Siamese convolutional neural networks. The first Siamese network detects damage (discerns whether the structure is healthy or damaged). Then, in the case of damage, a second Siamese network determines the damage diagnosis (classifies the type of damage). The results show that this approach requires very little data to train. The investigation shows values of 100% in all metrics (accuracy, precision, recall, and F1 score) of both convolutional networks.

The AP-III titled *Wind turbine main bearing fault prognosis based solely on SCADA data* selects the SCADA data located near the main bearing, in addition to the ambient temperature, the generated power, and the speed of the rotor. The data are preprocessed to remove outliers, but at the same time, filled by the piecewise cubic Hermite interpolating polynomial because a time series without missing time gaps is needed. The normality model is implemented on the basis of an artificial neural network trained only with healthy data. The validity and performance of the established methodology is demonstrated by a fault indicator on a real wind farm composed of 12 WTs. The results show that the methodology based solely on SCADA data can detect main bearing failure in three and a half months in advance of the fatal breakdown. It is important to recognize that the terms prediction and prognosis have distinct meanings. Prognosis involves calculating the remaining useful life of a component and predicting its parameters for a future time horizon, while also evaluating the accuracy of the prediction and the prognostic horizon. On the other hand, early prediction, as used in the context of incipient damage detection and being the focus of this article, refers to detecting small signs of damage to trigger an alarm before a total or fatal breakdown occurs. Therefore, please, note that the term prediction would be more suitable than prognosis in this paper title.

The AP-IV titled *Early Fault Detection in the Main Bearing of Wind Turbines Based on gated recurrent unit (GRU) Neural Networks and SCADA Data*, continues the main bearing failure early detection investigation started in AP-III, with the differences that the failure indicator is based on the exponential weighted moving average (EWMA) and the normality model is based on a gated recurrent unit GRU neural network. The model is built entirely from healthy data and is robust against all operational and environmental variations. The strategy is trained, validated, and finally tested using data from a wind farm in production consisting of nine WTs. The results show that the system produces a minimum of false alarms and that the defect in question is predicted two months in advance.

## 1.3 Outline

The rest of the thesis consists of three chapters and four appendices that contain the main articles published during this thesis. This section provides a summary of the content of Chapters 2 through 4 (part I of the thesis) and appendices AP-I through IV (part II of the thesis). The thesis is constructed in a way that aims to describe the entire process for the stated SHM and CM methodologies according to the framework presented in Figure 1.1.

**Chapter 2** starts by reviewing the state of the art in SHM related to WTs. Then, a comparison between the methodologies proposed in AP-I and AP-II is carried out. Challenges and limitations associated with the experimentation and the applied method are indicated.

**Chapter 3** begins with a state of the art in CM for WTs. In particular, the use of only SCADA data for this purpose is extensively reviewed. Then, a comparison between the methodologies proposed in AP-III and AP-IV is performed.

**Chapter 4** concludes Part I of the thesis with a summary and discussion of the material presented. The main results achieved in the project are described and suggestions for future work are provided.

**AP-I** This appendix contains the paper: Puruncajas, B., Vidal, Y., & Tutivén, C (2020). *Vibration-response-only structural health monitoring for offshore wind turbine jacket foundations via convolutional neural networks.* Sensors, 20(12), 3429. First quartile, Q1. Number of article citations (Wos Core Collection): 23.

This paper proposes a new approach for monitoring the structural health of offshore wind turbine jacket foundations using vibration-response data and convolutional neural networks (CNNs).

The collective contribution of this paper is the development of a new monitoring approach that can accurately identify damage in offshore wind turbine jacket foundations by analyzing the vibration-response data from accelerometers. The proposed method is tested on a scaled model of a jacket foundation, and the results demonstrate that the CNN-based approach outperforms traditional damage detection techniques.

The individual author contributions are as follows: Yolanda Vidal and Christian Tutivén (supervisors of the PhD thesis) conceived the main conceptual ideas. Bryan Puruncajas performed the implementation and numerical computation as well as conceived the data augmentation strategy. All authors discussed the results and contributed to the final manuscript.

Bryan Puruncajas's specific role in the paper was critical in developing the methodology, conducting investigations, analyzing the data, and drafting the manuscript.

Additionally, he played an instrumental role in the conceptualization of the data augmentation strategy and contributed significantly to the writing process.

**AP-II** This appendix contais the paper: Baquerizo, J., Tutivén, C., Puruncajas, B., Vidal, Y., & Sampietro, J. (2022). *Siamese neural networks for damage detection and diagnosis of jacket-type Offshore wind turbine platforms.* Mathematics, 10(7), 1131. First quartile, Q1. Number of article citations (Wos Core Collection): 0.

The collective contribution of this paper is the development and validation of a methodology to detect and classify damage in jacket-type support structures for offshore wind turbines that uses vibration-response-only accelerometer measurements and is based on two in-cascade Siamese convolutional neural networks (that can learn from very little data compared to standard machine/deep learning algorithms). The first network detects whether the structure is healthy or damaged, and the second network classifies the type of damage. The methodology is validated in a scaled-down experimental laboratory setup and has high accuracy, precision, recall, and F1 score metrics, making it a feasible option for detecting and classifying damage in offshore wind turbine support structures.

The individual author contributions are as follows: Yolanda Vidal and Christian Tutivén conceived the main conceptual ideas and contributed to the writing-review and editing. Joseph Baquerizo (undergraduate student) performed the data curation and implementation of the methodology, Bryan Puruncajas contributed the formal analysis, investigation, validation, and drafting the original manuscript. José Sampietro (doctor) contributed in the formal analysis.

Bryan Puruncajas played a crucial role in the development of the methodology presented in the paper, as he not only contributed to the formal analysis, investigation, and validation but also acted as a supervisor to (at that time) undergraduate student Joseph Baquerizo, who performed the data curation and implementation of the methodology. Bryan's guidance and leadership skills were instrumental in ensuring the successful implementation of the methodology. Additionally, he was involved in drafting the original manuscript and contributed significantly to the writing-review and editing of the paper.

**AP-III** This appendix contains the paper: Encalada-Dávila, Á., Puruncajas, B., Tutivén, C., & Vidal, Y. (2021). *Wind turbine main bearing fault prognosis based solely on SCADA data. Sensors, 21(6), 2228.* First quartile, Q1. Number of article citations (Wos Core Collection): 24

The collective contribution of this paper is a data-based methodology for early fault prediction using SCADA data. The proposed method only requires healthy data to be collected and works under different operating and environmental conditions. The established methodology is demonstrated on a real in production wind farm consisting of 12 wind turbines, showing that advanced predictive maintenance systems based solely on SCADA data can predict faults several months prior to the

fatal breakdown of the component, allowing wind turbine operators to plan their operations.

The individual author contributions are as follows: Yolanda Vidal and Christian Tutivén conceived the main conceptual ideas and contributed to the supervision and writing-review and editing. Ángel Encalada (undergraduate student) and Bryan Puruncajas performed the data curation, investigation, validation, and drafting the original manuscript. Bryan Puruncajas also contributed to the methodology and supervision.

Bryan Puruncajas played a crucial role in this paper as he not only contributed to the methodology, data curation, investigation, validation, and drafting the original manuscript but also provided supervision to Ángel Encalada, at that time an undergraduate student who performed the data curation, investigation, and validation. His role as a supervisor to an undergraduate student highlights his expertise in the field and his ability to guide and mentor young researchers, which is essential in promoting the growth and development of future scientists.

**AP-IV** This appendix contains the paper: Encalada-Dávila, Á., Moyón, L., Tutivén, C., Puruncajas, B., & Vidal, Y. (2022). *Early fault detection in the main bearing of wind turbines based on gated recurrent unit (GRU) neural networks and SCADA data. IEEE/ASME Transactions on Mechatronics.* First quartile, Q1. Number of article citations (Wos Core Collection): 3.

The collective contribution of this paper is the development of a methodology to detect faults in wind turbine main bearings using a gated recurrent unit (GRU) neural network. The GRU can identify data in a time series that is crucial enough to preserve or forget, allowing for early diagnosis of faults in the bearings before they cause significant downtime and replacement costs. This methodology only requires healthy SCADA data and can be deployed to both old and new wind parks. The strategy was trained, validated, and tested using SCADA data from a nine-turbine wind park currently in production.

The individual author contributions are as follows: Bryan Puruncajas and Christian Tutivén conceived the main conceptual ideas, designed the study and played a supervisory role, Yolanda Vidal contributed the formal analysis, Ángel Encalada (undergraduate student) and Luis Moyón (undergraduate student) analyzed the data, performed the implementation and numerical computation. All authors contributed to writing and revising the manuscript.

Bryan Puruncajas played a key role in the success of the study by conceiving the main conceptual ideas, designing the study, and playing a supervisory role for the undergraduate students involved in the research. Bryan proved his management and mentorship abilities, which is a vital achievement in the academic and research communities.

# Chapter 2

# Structural Health Monitoring

In this chapter, the WT is first presented in Section 2.1. Then, an SHM state of the art for offshore WT substructures is addressed in 2.2. Next, the two stated and validated methodologies comprehensively exposed in AP-I and AP-II, for the detection and diagnosis of damage in jacket-type offshore supports, are compared to each other in Section 2.3.

## 2.1 The Wind Turbine

The WT is an electromechanical system that converts the kinetic energy developed in the blades into electrical energy. Energy conversion is done in two steps. First, the kinetic energy of the wind captured in the blades is converted into mechanical energy, and second, the conversion of mechanical energy into electrical energy by the WT generator [19]. The principle of capturing wind energy is the same as that used more than two centuries ago, with the difference that today aesthetics and technology have changed WTs.

The basic parts of a WT are the foundation, tower, nacelle, and rotor with three blades, see Figure 2.1.

Wind energy is collected in wind farms (set of WTs). Currently, these farms are found in two different types: onshore wind farms (see Figure 2.2), which are large WT installations located on land, and offshore wind farms, which are facilities located on the sea. Onshore wind farms are the most popular type of wind farm today, but there is growing interest in the development of offshore wind farms in developed countries [21]. A key difference between onshore and offshore wind farms is the consistency of the power they can generate. Offshore wind farms generate more electricity than their onshore counterparts, due to higher and more constant wind speeds offshore [22]. Another difference is the location of offshore wind farms that are much easier to find and have less visual and sound impact on people. In addition, restrictions in activity policies on and around offshore wind farms benefit replenishing fish populations and helping fish farming [23]. By comparison, onshore wind farms require careful analysis to ensure that wind speeds are sufficient and environmental impact is reduced. Some benefits of onshore

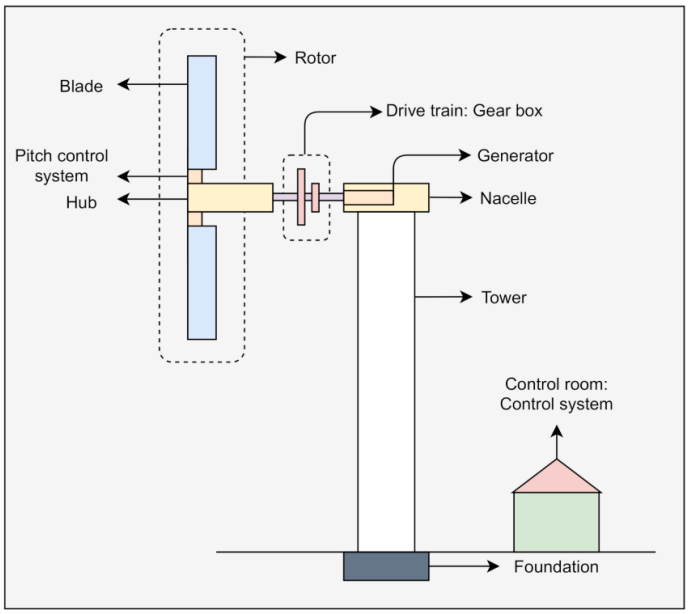



Figure 2.1: Basic structure and main components of the WT [20].

wind farms are their easier accessibility on land and their easy connection to local power grids.

(a)

(b)

Figure 2.2: Wind farm classification. (**a**) Offshore wind farm [24]. (**b**) Onshore wind farm [25].

Another important point for a meaningful comparison is the cost. Currently, onshore wind farms are much cheaper to build and maintain than offshore wind farms [26]. However, technologies are improving and it is likely that at some point the cost of offshore wind farms will be similar to or cheaper than onshore wind farms, at which point their larger capacity will likely make them more attractive.

In general, the main components of an industrial WT (see Figure 2.1) are: the torque-

generating blades connected to the hub, which is an integral part of the nacelle; the main shaft, gearbox, and generator are housed in the nacelle; electrical cables descend from the gondola through the tower to the foundation. The main difference between onshore and offshore WTs is how they are structurally supported. Offshore WTs, in addition to the tower and the foundation, have a substructure that connects the transition tower to the foundation at seabed level [23]. A further classification is shown in Section 2.2.

In more detail, the nacelle houses all the components that are essential to operate the turbine efficiently. It is fitted at the top of a tower and includes the gearbox, low- and high-speed shafts, generator, controller, and brakes. A wind speed anemometer and weather vane are mounted on the nacelle, among other possible sensors. Figure 2.3 shows the principal components of a conventional WT.

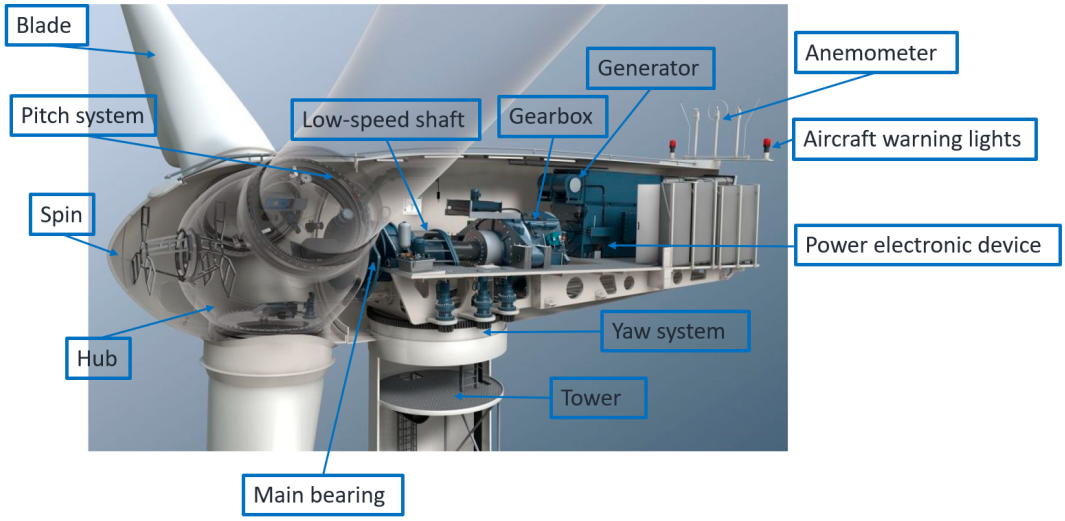



Figure 2.3: Principal components inside the nacelle of a conventional WT [27].

### 2.1.1 Structural parts in a wind turbine

The structural parts of the WT that suffer damage, as already defined in Section 2.1, and thus fall into the SHM objective are the following.

- **The blades.-** They are generally made of polymeric materials reinforced with fiber-glass or carbon. The blade design is quite complex, with several components and materials. This complexity makes them particularly susceptible to manufacturing defects, which account for 51% of all damage (with detachment and voids in the skin core being the most common defects at 20% and 18%, respectively). From a geometric perspective, the cross-sectional profile varies from root to tip, often rotating around its main axis. The structural integrity of the WT blades is of the

utmost importance due to the possible impacts of the totally or partially detached blades on the neighboring structures [28].

- **The rotor.-** The rotor is a crucial component of the WT. It is the most expensive mechanical component, up to about a fifth of the total cost. Its design is one of the most critical and delicate phases, especially in terms of expected performance. Therefore, the economic viability of the entire system depends on these aspects. The diameter of the rotor governs the height of the tower and thus the overall size of the structure, including the foundations. A larger rotor will result in a consequent increase in power output but, on the other hand, will require a wider tower cross section and more massive and/or deeper foundations, thus increasing construction costs[29].

- **Support structure.-** The basic function of the support structure is to hold the WT in place. This means that it must be built to withstand the loads of sea currents, waves, and wind, acting both on the supporting structure and on the turbine in operation. There are a variety of WTs available on the market, designed by different turbine manufacturers, in a range of power ratings. Each WT has different characteristics. The environmental conditions offshore can also vary from site to site. Therefore, the support structures are designed specifically for each case. It is not uncommon for an offshore site to have several variations of one type of support structure for one type of turbine [30].

  Normally, the support structure is divided into two main parts (the tower and the foundation) for onshore WTs and three main parts (the tower, the substructure, and the foundation (see Figure 2.4)) for offshore WTs:

  - **The tower.-** The tower is the main component of the supporting structure. The main parameter of the tower is its height. This is usually around 1.5 times the diameter of the rotor; it is usually never less than 20 meters and can reach up to 150 meters or more (for 10-12 MW outputs). In absolute terms, the taller the tower, the better the wind conditions in terms of strength and consistency. The diameter of the tower (maximum at its base and minimum at its top) increases with the height of the tower. For example, a typical 50 m tall WT will have a tower diameter ranging from 3.5 meters to 0.4 meters [31].
  - **The substructure.-** It is worth mentioning that offshore WTs, unlike their onshore counterparts, include an additional group of structural elements, which are located below the platform and above the seabed. These components are especially exposed to risks because of their location underwater or in the splash zone, immediately above or below the mean water level, and are highly exposed to corrosion. Furthermore, being submerged, they cannot be easily inspected visually without the use of divers or manned or unmanned underwater inspection robots. They are also subject to marine growth and other potentially damaging environmental conditions, such as waves, tides, and currents [8].

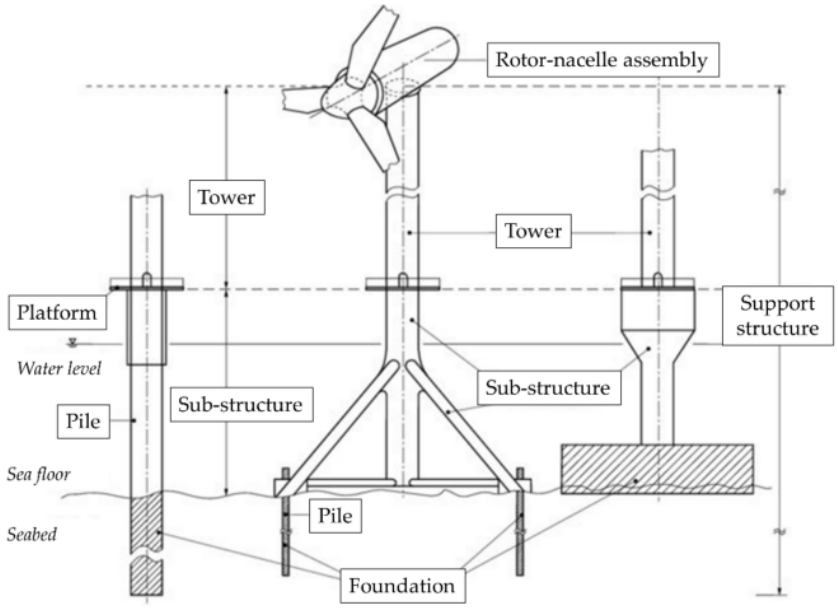



Figure 2.4: Structural components of a fixed offshore WT.

– **The foundation.-** Onshore and offshore foundations differ markedly. However, in both cases, the choice of a specific structural design depends on location and site conditions. For example, soil quality and strength are the main determinants that precipitate the size and shape of the onshore foundations, while the depth of water and distance from the shore are the key factors for offshore turbines [8].

Different types of foundations-substructures of offshore wind farms are shown in Figure 2.5. The function of the substructure and foundation is to direct the loads toward the seabed. Essentially, the following types are the most used:

- **Gravity base.-** Gravity base foundations are placed on the seabed, and their stability is ensured by dead weight. They are built almost exclusively as solid structures or shell structures, using reinforced or prestressed concrete. They are the most common type of foundation for onshore WTs and were therefore chosen for the first offshore WTs located in very shallow water. One of the drawbacks of gravity-based substructures is that they require soil preparation prior to installation, as well as extensive scour protection.

- **Suction bucket.-** The suction bucket foundation is a large cylindrical structure that is open at the bottom and closed by the bucket lid at the top. During installation, the bucket is lowered to the bottom of the sea, and the skirt slightly penetrates the ground due to its weight. Then, water is pumped out of the bucket, drawing a vacuum under the lid, causing the skirt to go deeper into the seabed until the bucket lid rests on it.

- **Monopile.-** It consists of a single, large-diameter steel pile that is driven into the seabed to provide a stable base for the WT tower. The monopile is typically installed using a pile driver, and can be up to several meters in diameter and over 100 meters in length, depending on the size of the WT. Monopile foundations are commonly used for offshore wind farms due to their relatively simple design and ease of installation, but may not be suitable for deeper waters or more challenging seabed conditions.

- **Tri- o tetrapod.-** The tri- or tetrapod type consists of three or four concrete legs, respectively, which extend from a common base and taper upwards to support the tower of the WT. This type of foundation is designed for shallow waters to withstand the environmental loads and support the weight of the WT while minimizing the environmental impact and installation costs.

- **Jacket.-** The Jacket-type base is a square network of steel rods, anchored at four points, and the entire steel construction can be assembled in a single piece. The use of a three-dimensional reinforcement such as the jacket base substantially increases the rigidity. Although more expensive than a monopile or gravity base, the jacket base is more cost-effective at greater depths.

- **Floating.-** Many proposed floating concepts use designs borrowed from the oil and gas industry. An example is the HyWind concept, installed in June 2009 by Siemens and StatoilHydro; this is the first megawatt floating turbine. It is designed for depths of 120-700 m and has a capacity of 2.3 MW. The structure consists of a steel float filled with ballast water and rocks. Extending 100 m below the surface, it is attached to the seabed by three steel anchor cables.

As shallower sites are developed, facilities will naturally move into deeper water. Increasing depth leads to an increase in overall offshore project costs [33]. The cost of an offshore structure at a specific site is a function of the depth of the water and the distance from the shore. For example, the total costs of foundations at water depths ranging from 40 to 50 m are 1.32 and 1.13 times higher than the cost for water depths of 20 - 30 and 30 - 40 m, respectively. Therefore, the right foundation plays an important role, as it can efficiently reduce the total cost of wind projects [34]. Consequently, this thesis focuses on SHM studies with a focus on jacket foundations because it is the validated option as offshore wind farms transition to deeper waters and before floating foundations become economically feasible.

## 2.2 SHM state of the art for offshore wind turbine substructures

SHM can generally be defined as a strategy for early detection of damage, deterioration, or boundary condition changes of a structure or system through pattern recognition of measured values or damage-sensitive parameters calculated from recorded sensor values.

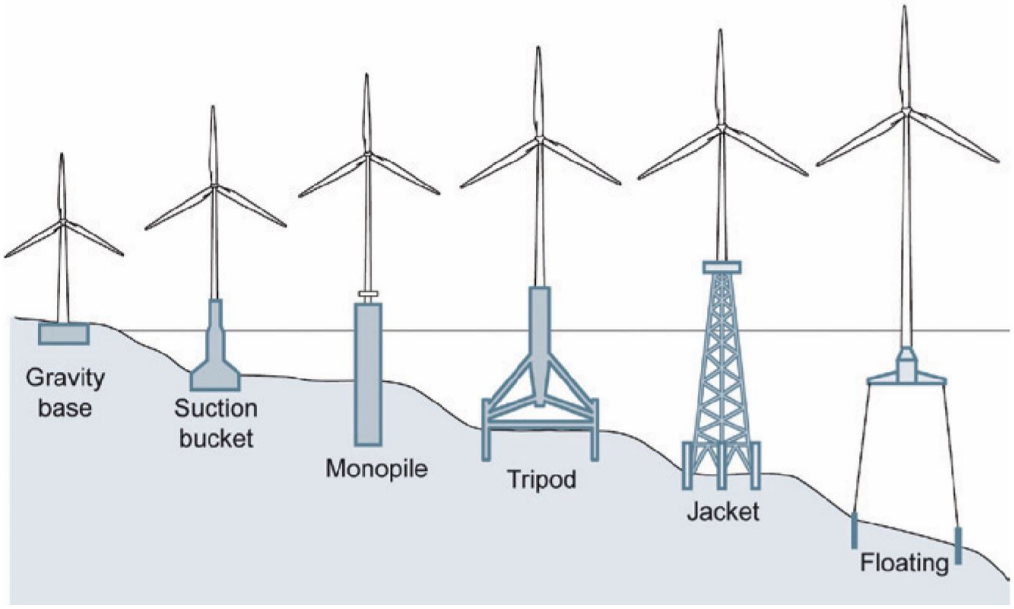



Figure 2.5: Types of substructures-foundations of offshore wind farms [32].

Figure 2.6 shows a general SHM process, which comprises operational evaluation, data acquisition, feature extraction, and diagnosis/prognosis [35].

Operational evaluation, being the first step in developing an SHM system, sets the boundaries by answering basic questions about the safety and economic importance of the planned SHM system, possible damage that are of concern, under which they set the environmental and operational conditions (EOC) to operate, and the effect of EOCs on limiting data collection. The second stage, data acquisition, deals with the selection of suitable sensors and data acquisition equipment based on the responses from the first step. The feature extraction step deals with the processing of the collected data by selecting an appropriate technique relevant to the particular application. Finally, the diagnosis and prognosis step uses modeling for feature discrimination, in which statistical models and/or machine learning (ML) approaches are used to separate features from damaged and undamaged states through pattern recognition [36].

The level of detail of the last diagnosis and prognosis step depends on the monitoring concept, the sensitivity of the monitoring system, and the data analysis technique used. The diagnosis of damage and quantification of its severity, as well as its effect on the remaining useful life of the structures, can be carried out through the damage identification levels. Each level requires that the previous levels be known, as they are hierarchically in increasing the knowledge about damage [37].

i. Damage detection (level 1): Is there any indication of damage?

ii. Damage localization (level 2): If there is evidence of damage, where is it located?

iii. Damage quantification (level 3): How severe is the indicated damage?

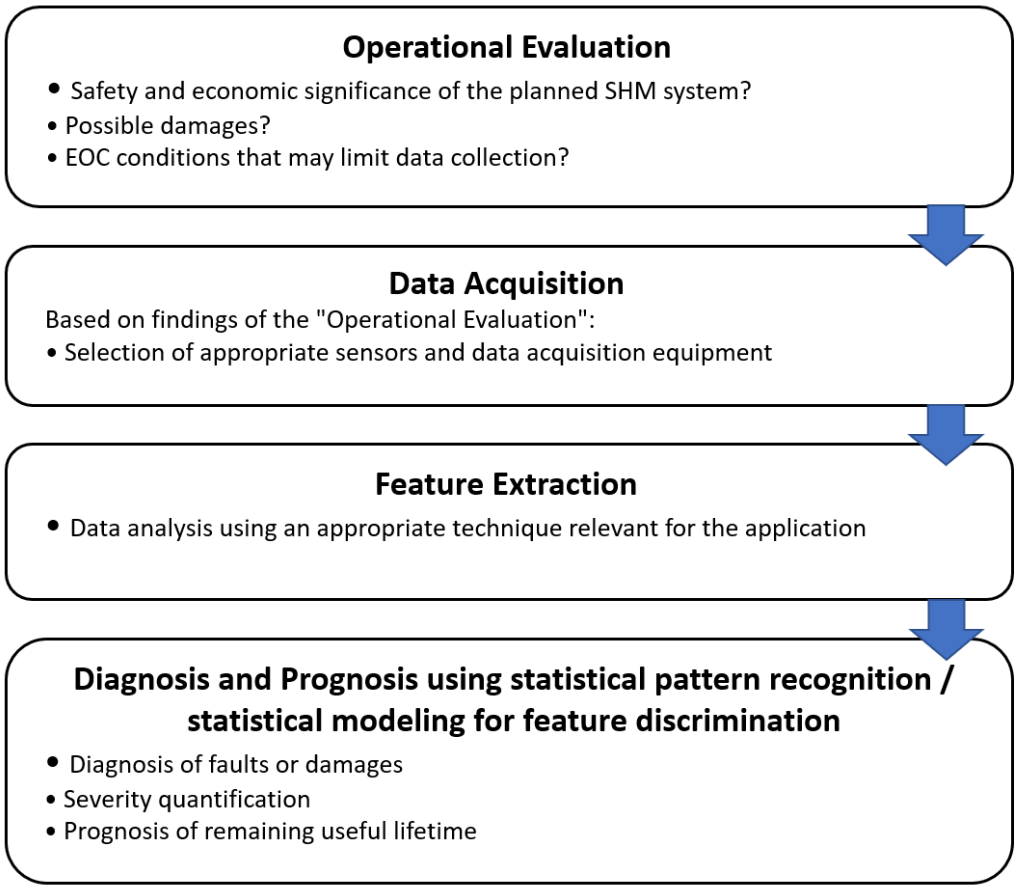



Figure 2.6: SHM flowchart process [35].

iv. Damage classification (level 4): What kind of damage is it?

v. Evaluation of structural integrity (level 5): How much service life is left?

For offshore WTs, due to their dynamic sensitivity, the rotating mass at the top, and the EOC to which they are exposed, the damage identification procedure is more complicated. In particular, the monitoring of substructures is challenging due to their variety and the installation of these systems during construction. Installing a sensor system during construction may require additional coordination between the builder (often local companies) and SHM system technicians. Finally, the durability and accessibility of the sensor are a concern. If such a sensor fails, it is impractical to access it for maintenance because the sensors must be submerged in salt water, which requires special packaging and diver-based maintenance [38].

In a global SHM concept, the sensors used in SHM are few and strategically located. They continuously record the response of the structures from which the damage-sensitive parameters are directly obtained or calculated. Especially in the marine environment, robust sensors suitable for harsh and corrosive environments are used. Additionally, as far

as possible, the sensors are placed in accessible places. Among the most commonly used structural (vibration) sensors for marine structures are acceleration sensors, displacement transducers, inclination sensors, and others.

SHM in WT substructures can be achieved by measuring signals from sensors and evaluating changes in the characteristics of the signal measurements. The vibration-based SHM technique, also known as global method, is common in monitoring offshore WT structures. This method uses the global vibration characteristics (dynamic properties) of a structure, where each structure has a unique vibration signature. If the healthy state vibration signature (dynamic property) is correctly determined, then the deviation from these healthy state references will serve as an indicator of damage or change in the boundary condition. For optimal results, the local and global methods should be combined. Global concepts with their scattered sensors need local non-destructive techniques during inspection to take a closer look at the location and assessment of damage before a decision to repair it can be made. Similarly, local methods need prior knowledge of potential damage. In our approach, offshore WT substructure damage can be obtained from global approaches [39].

With the rapid advancement in computing power and the availability of vast amounts of data, artificial intelligence plays an important role in the development of new technologies. Specifically, machine/deep learning, which is a part of the broad field of AI, uses historical data to develop and enable tools to learn, predict, and make decisions without being programmed to perform such specific tasks. Like most other disciplines, the SHM field has also been making use of machine/deep learning capabilities to improve damage detection and diagnosis strategies. Typically, damage-sensitive parameters or characteristics are extracted from monitoring data. For SHM vibration-based techniques, the basic idea of using automatic learning is to train the model using measured responses or extracted features for different damage scenarios and the undamaged case so that it knows how the structure responds to different damage and damage levels. For such training, measurement data, model experiments, numerical simulation, or a combination of these are used. Machine/deep learning can assist in the automation of the SHM process, in particular by mapping the damage-sensitive characteristics and/or parameters of offshore WT substructures in detection and diagnosis to minimize expert human intervention.

In general, monitoring the health of offshore WT substructures is currently limited to research activities and specific applications. As an example, Strainstall company installed SHM systems consisting of strain gauges, displacement sensors, and accelerometers. The efforts in the supports are monitored and the fatigue life of the critical areas is calculated. In addition, the natural frequencies of the tower are tracked, since a change in the tower frequencies could indicate a foundation problem, and some systems also use inclinometers to track the angle [40]. An extensive monitoring system has been installed at the Belwind offshore wind farm in the North Sea. The system includes foundation corrosion control using Zensor sensors, load and displacement control of grouted connections, and dynamic control of the overall structure [41]. Brincker and Ibsen instrumented a Vestas 3 MW turbine with accelerometers to help a FEM perform fatigue analysis on a new

offshore "cube" foundation design to aid in the design process and installation of the foundation [42]. Strainstall has a scour monitoring system that tracks the progress of a magnetic collar along a metal pipe embedded in the seabed, to detect excessive scour in the substructure-pile foundation. As the soil erodes, the collar slides down, triggering an alert each time one of the sensors along the tube passes [38].

In the following sections, a review of the different sensors used in SHM in general (Section 2.2.1) and the ones that can be used in the specific application of offshore structures (Section 2.2.2) will be carried out, as well as a classification of the different vibration-based SHM methodologies (Section 2.2.3).

### 2.2.1 Overview of sensor technologies for SHM in general

SHM damage detection methodologies can be classified according to the nature of the sensor used or according to the main physical parameter studied. Hereby, an overview of different sensor technologies used in general for SHM is given.

- **By fiber-optic sensors .-** Optical fiber (OF) itself is a conduit for light that transmits information, but it can also be sensitive to changes in the external environment surrounding the fiber. The main purpose of using fiber optics is usually for strain and temperature measurements. The common advantages of all types of fiber optic sensors arise from their small size and non-electrical nature, making them immune to electromagnetic interference and electrical noise [43]. For SHM, fiber bragg gratings (FBG) are widely used. FBGs are recognized as having many advantages compared with conventional electrical resistance foil gauges due to their corrosion resistance, immunity to electromagnetic interference, light weight and small physical dimensions facilitating embedded deployment of FBGs into structures for long-term strain monitoring purposes [44]. This is still an open field of research, and significant efforts are being made to analyze the behavior of FBGs when subjected to these phenomena and to obtain the actual strain field applied to the network by demodulating its spectral information.

- **By piezoelectric sensors .-** Piezoelectric sensors can be found on different principles. The first is based on acoustic emission (AE) while the other is based on acousto-ultrasonics (AU).

  AE is mainly used to study the physical parameters and damage mechanisms of a material, but it is also used as a non-destructive testing technique. The phenomenon is based on the release of energy in the form of transient elastic waves within a material that undergoes dynamic deformation processes. The waves, of different types and frequencies, propagate in the material and detect possible modifications before reaching the surface of the studied sample [45]. Once the system has detected and located the occurrence of damage, the system performs an acousto-ultrasonic test throughout the impact to rate the severity of the damage. Traditional AE parameter analysis allows for simple, fast, and cost-effective inspection or damage detection.

The AE described in the previous paragraph makes use of piezoelectric sensors attached or embedded in a structure passively. The same type of attached sensors can also be used to actively produce and detect high-frequency vibrations. The AU uses a transmitter to send diagnostic stress waves through the structure and a receiver is used to measure the signal, which has changed due to the presence of damage to the structure. This wave propagation approach is very effective in detecting damage in the form of geometric discontinuities [46]. AU is generally used for 2D type structures (plate structures for example), and although some work has been done for 3-D parts, volume waves are usually used in that case.

- **By electrical impedance .-** The goal of impedance-based damage detection is to measure overall electrical resistance. This seems to be a valuable technique to monitor fiber fractures in unidirectional materials and the delamination process connected with a modification of the resistive tracks in the laminates. The impedance based damage detection method uses electromechanical coupling of lead-zirconium-titanium oxide (PZT) patches. When a PZT transducer is attached to a structure, it forms a co-located sensor and actuator, often referred to as an auto-sensing actuator. When a voltage is applied across the PZT, the structure is displaced, and conversely, when the structure is displaced, a voltage is developed in the PZT transducer. Therefore, the PZT transducer can act on the surrounding area of the structure and detect the resulting structural response [47].

- **By low-frequency electromagnetic techniques .-** Using electromagnetic methods, a family of electromagnetic techniques has been developed that allows obtaining thoroughly information on the health of composite structures. These techniques consist of determining the state of health of a structure by measuring its main electrical parameters, electrical conductivity, and/or dielectric permittivity since damage induces locally significant variations in these two or three parameters. There are three possible techniques: the magnetic technique, the electrical technique, and the hybrid technique. Which of these methods is most suitable depends on the type of material [48].

- **By capacitive methods .-** The principle of the capacitive method consists in placing two (or more) electrodes on the outer surface of the samples and applying a voltage between them. This system forms a capacitor, and changes in capacitance are indicative of internal constituents (such as the nature of the materials or their moisture content). The capacitance value is found using a resonant circuit that provides an alternating voltage. The resonant frequency shift is obtained simply by using a frequency analyzer. This type of monitoring has been used primarily to monitor historic buildings.

- **Vibration-based .-** Vibration-based methods use vibration measurement sensors (piezoelectric, accelerometers,...) to detect vibrations in the structure. These methods assume that the vibration characteristics of the structure change due to damage, and by identifying the new characteristics and comparing them with those of the

healthy structure, the existence of damage can be detected. The reliance on vibration signals gives this method additional advantages, such as the ability to be automated and reasonably low-cost signal acquisition and measurement technologies [49].

In WTs, environmental and operational variations of WTs significantly alter the dynamic response characteristics, these include wind conditions, sea state, temperature, humidity, operating states, rotation speeds, yaw angles, etc. The vibration-based method fits with the aforementioned variations of WTs because this method does not require knowledge of the excitations of the structure, which is why is the one used in this doctoral thesis. Therefore, Section 2.2.3 is fully devoted to a comprehensive review of vibration-based techniques for SHM.

### 2.2.2 Overview of sensor technologies for SHM in offshore structures

To design an SHM system, one of the critical missions is to discover how to determine an appropriate type of sensor that can cope with operating and environmental conditions efficiently and also meet the scopes of the designed detection system. Recall that as the structure is submerged under the sea level, robust sensors suitable for harsh and corrosive environments are required. Therefore, this section aims to present a brief review of the different types of sensors for offshore structural health monitoring [50].

- **Accelerometer .-** There are several different types of accelerometers. Furthermore, accelerometers can measure in 1, 2, or 3 directions, and they have different properties with respect to sensitivity, resolution, and noise characteristics. Common types of accelerometers are forcebalanced, micro-electromechanical systems (MEMS), variable capacitance, piezo-electric or piezo-resistive.

  Limitations: The sensitivity, resolution, and noise characteristics of accelerometers must be considered with respect to the expected structural response and loading to produce usable data. When attached to structural members they are in principle both measuring the components and systems accelerations.

  Possible applications: System identification (natural frequencies, mode shapes and damping), measurements of global displacements, measured fatigue accumulation, model and load calibration, uncertainty assessment, detection of anomalies (damage).

- **AE sensor .-** The purpose of AE sensing is detection of anomalies in materials. In the offshore industry, the technique is primarily used to detect fatigue crack initiation and the monitoring of crack growth. Modern AE sensors and systems are accurate and can detect crack initiation and early stages of propagation.

  Limitations: The system can be prone to background noise which can affect its accuracy and reduce probability of detection. The system can be partially sensitive to any activity such as maintenance works in the local sensor vicinity.

Possible applications: Identification of crack initiation and crack growth.

- **Electrical resistance (ER) probe .-** ER probes and instruments determine metal loss from corrosion or erosion by the electrical resistance method.

  Limitations: The ER method allows only the measurement of uniform corrosion but cannot identify localised corrosion phenomena.

  Possible applications: Quantification of corrosion impact.

- **Inclinometer/tilt sensor .-** Inclinometers, also called tilt sensors, measure the slope/angle/tilt of objects based on gravity in various applications.

  Limitations: The position relative to other sensors must be known to provide meaningful results.

  Possible applications: Identification of platform tilt due to, e.g., subsidence.

- **Light detection and ranging (LiDAR) sensor .-** LiDAR sensing is a method for determining ranges (variable distance) by targeting an object with a laser and measuring the time for the reflected light to return to the receiver.

  Limitations: High costs. Unable to measure distances through heavy rain, fog and snow.

  Possible applications: Scanning of surfaces, identification of distances/contours.

- **Load cell .-** A load cell measures mechanical force, mainly the weight of objects.

  Limitations: Calibration/accuracy and maintenance (over time).

  Possible applications: Monitoring of topside weight and mooring tendons on tension-leg platform (TLP).

- **Scour sensor (acoustic) .-** Scour sensors are typically covered by echo sounders and sonars. Scour sensors can be used to perform scour depth measurements in an area around a bottomfixed structure. When excessive sediment transport occurs, the sensor will indicate a change in elevation.

  Limitations: See AE sensors.

  Possible applications: Identification of scour depth.

- **Strain gauge .-** A strain gauge is a sensor used to measure strains on an object, as the resistance in the sensor varies with applied force. The strain gauge converts force, pressure, tension, weight, etc., into a change in electrical resistance which can then be measured.

  Limitations: The performance of a strain gauge can be affected by the change in temperature and humidity. The sensor must be installed on a clean surface. The robustness of strain gauges can be limited.

  Possible applications: Measurements of strain/stress in structural members, strain levels in bolts and flanges.

- **Video camera .-** Video cameras are e.g., used for surveillance of the subsea system.

  Possible applications: Video surveillance.

- **Wind and humidity sensor .-** A wind sensor (anemometer) is an instrument used to measure the speed or velocity of gases in e.g., unconfined flows, such as atmospheric wind. Humidity sensors (hygrometers) work by detecting changes that alter electrical currents or temperature in the air.

  Limitations: Anemometers may get damage during strong winds.

  Possible applications: Identification of meteorological properties such as wind speeds and directions.

In this thesis, taking into account that vibration-based methodologies are the focus of the PhD research in the area of SHM, accelerometers are used due to their ability to measure mechanical vibrations and accelerations in structures with high accuracy and sensitivity.

### 2.2.3 Vibration-based SHM methodologies

Structural damage detection techniques can be broadly classified as either global or local methods [51]. While vibration-based methods are typically considered global, local methods focus on detecting and measuring damage on a smaller scale without relying on vibration responses. Since the detection range of local methods is relatively small, most non-destructive testing and evaluation (NDTE) methods are mostly considered local methods [52]. Local methods employ various tools such as ultrasonic testing (UT), AE, infrared thermography (IRT), radiographic testing (RT), magnetic flux leakage (MFL), magnetic particle testing (MT), digital image correlation (DIC), fluid penetrant testing (PT), laser testing methods (LMs), ground penetrating radar (GPR), leak testing (LT), visual testing (VT), and numerous optical methods to inspect, test, and evaluate structural components and assemblies in local infrastructure areas. It is important to note that while these methods are useful for detecting damage on a smaller scale, they are insufficient for SHM of large-scale civil structures on their own. To gain a comprehensive understanding of the structural condition of a large structure, an effective SHM system should integrate both local and global damage detection techniques [53].

In contrast, global (i.e., vibration-based) structural damage detection methods utilize the vibration response of the monitored structure to assess its overall condition. Damage detection methods through vibration response of structures has been an area of research over the decades. Researchers focused on the time, frequency and modal domains searching for the presence, location and the severity of damage on engineering structures [54]. These methods involve the placement of a network of accelerometers at strategic locations on the structure to capture its vibratory response. Subsequently, advanced algorithms are utilized to convert the recorded acceleration signals into indicators that indicate the presence of damage, identify its location, and estimate its severity [55]. Compared to local methods, vibration-based structural damage detection techniques provide several advantages, including [56]:

1. Vibration-based methods offer an advantage in that they do not require a dense network of sensors. Even with a limited number of accelerometers, dynamic properties can be accurately identified in large and complex structures.

2. Unlike local methods, vibration-based methods do not require prior knowledge of the location of the damage. This makes them particularly useful for detecting damage in structures with complex geometries.

3. Another advantage of vibration-based methods is that the equipment required for monitoring can be easily transported, handled, and attached to the monitored structure. This ease of use allows for efficient and cost-effective monitoring of large-scale structures.

The subsequent subsections present a categorization of vibration-based SHM techniques, organizing them into three distinct groups: AI-independent, ML-based, and deep learning (DL)-based methodologies.

#### 2.2.3.1 Vibration-based SHM methods AI-independent

Depending on the information extracted from the measured signals, vibration-based structural damage detection methods that are not based on ML or DL can be categorized into parametric and non-parametric methods [57].

- **Parametric vibration-based.-** The term "parametric" as an adjective refers to something that is related to or expressed in terms of parameters. In parametric vibration-based structural damage detection methods, vibration signals obtained from the detection interface are used to estimate unknown dynamic parameters of the structural system, including modal frequencies, modal mass, modal damping, stiffness, and mode shapes [58]. Parametric methods aim to detect structural damage by comparing the dynamic parameters of the undamaged structure with those of the damaged structure. Any changes in these parameters relative to a predefined reference condition can be used as indicators of the location, severity, and extent of the damage [59].

  Previous parametric methods have mainly relied on correlating structural damage with changes in modal characteristics. One approach involves applying a known input excitation to the undamaged structure and measuring the vibration response using a set of accelerometers. This can be achieved using a modal sledgehammer or an electro-dynamic shaker [60]. The input/output signals are then processed by a modal identification algorithm to determine the modal parameters of the undamaged structure. To evaluate the condition of the structure over time, the same procedure is repeated, and the extracted modal parameters are compared with those of the reference condition [61].

  It is not feasible to experimentally excite civil structures [62]. However, seismic excitation can serve as natural experimental data if appropriate identification methods

are used, despite the fact that the seismic input excitation is mostly unmeasurable and non-stationary due to structure interaction effects, making traditional output-only methods ineffective. The following shortcomings of structural damage detection methods based on parametric vibrations are identified below [63]:

1. Recent research suggests that changes in global dynamic characteristics, specifically those corresponding to the lowest vibration modes, cannot always be associated with certain types of structural damage. Local structural damage, on the other hand, primarily affects higher frequency modes, which can be challenging to identify using output-only methods.
2. Modal parameters can be influenced by various factors, including temperature, humidity, and measurement noise, in addition to damage. Therefore, changes in these parameters do not necessarily indicate the presence of actual structural damage.
3. Output-only parametric methods rely on sophisticated system identification algorithms to solve inverse problems, estimating modal parameters based on environmental and structural response. However, implementing such algorithms is often impractical, making them unsuitable for real-time damage detection applications.
4. Parametric methods of damage detection are centralized, requiring that all signals be transferred to a central processing unit for damage identification. However, when wireless sensor networks are used, the transfer and synchronization of a large number of measurements can be problematic. As a result, centralized methods are not resilient to sensor failures, as they require full operation of all sensors in the network.

- **Non-parametric vibration-based .-** Unlike parametric approaches based on system identification, non-parametric vibration-based structural damage detection methods use statistical means to detect structural damage directly from measured accelerations and can extract damage characteristics that cannot be easily attributed to physical changes to the structure. Non-parametric methods combine time-series modeling with statistical classification [64]. The first step is extracting damage-sensitive features from raw signals using a time-series modeling technique. The extracted features are then processed by a classifier or an outlier detector to assess the current state of health of the structure [53].

#### 2.2.3.2 Vibration-based SHM based on ML

This section will review vibration-based structural damage detection methods based on ML methods, particularly supervised ML algorithms. The field of AI has made remarkable progress in developing data-driven training techniques. Once trained, these methods offer a versatile and adaptable solution that is fine-tuned by the underlying training method for the specific task at hand [65]. Over time, as the system parameters change,

the model can be incrementally improved (retrained). Recent literature has introduced several approaches that achieve expert-level, or even superior, performance for a variety of problems. These methods are both speedy and cost-effective to implement. Moreover, a data-driven method originally trained for a specific problem can often be repurposed for other challenges, even unrelated ones, through a process known as transfer learning [66].

Traditional ML algorithms typically require data to be represented in a fixed number of features. Therefore, a feature extraction step is necessary, which involves preprocessing the data to extract relevant attributes that capture the most significant information. Together with data labels, processed data samples, expressed in terms of extracted features, are used to train the ML system using a specific training method, allowing it to learn how these features correlate with various data patterns [67]. Vibration-based structural damage detection methods based on ML can also be categorized as either parametric or non-parametric. The classification is explained in detail below [68].

- **ML methods for parametric vibration-based SHM.-** Various parametric SHM methods based on ML have been proposed, where the feature extraction process involves identifying key modal parameters of structural systems using input-output or output-only modal identification techniques. Once the modal parameters are extracted, a well-trained ML classifier processes them to evaluate structural integrity [69]. The most commonly used parametric ML approaches rely on modal features, such as natural frequencies and mode shapes, as the extracted features. Common ML classifiers used in these approaches include artificial neural networks (ANNs), multilayer perceptrons (MLPs), fuzzy neural networks (FNNs), Bayesian networks, support vector machines (SVMs), principal component analysis (PCA), probabilistic neural networks (PNNs), among others [53].

- **ML methods for non-parametric vibration-based SHM.-** ML algorithms are increasingly used to develop new non-parametric methods for vibration-based structural damage detection. Unlike traditional ML approaches, these methods employ signal processing techniques to extract damage-sensitive features without relying on modal identification [70]. The extracted features are then fed into an ML algorithm, which performs structural damage detection and localization. Among the commonly used ML classifiers in the literature for non-parametric vibration-based structural damage detection are ANNs, factor analysis, Mahalanobis distance, singular value decomposition (SVD), support vector data description (SVDD), K-Nearest Neighbors, SVMs, and adaptive neuro-fuzzy inference system (ANFIS), among other combinations. Feature extraction methods found in the literature include auto regressive (AR) modeling, maximum and variance acceleration signals, PCA, wavelet decomposition, interval modeling, self-organizing maps (SOMs), and others [53].

In the context of SHM, both parametric and non-parametric approaches rely on extracting a fixed set of hand-drawn (i.e. user-defined) features from vibration signals.

Therefore, the successful performance of ML-based structural damage detection methods is highly dependent on the quality of the extracted features and the choice of classifier [71].

Selecting appropriate features is critical to effectively capturing the characteristic information of the analyzed signals. In addition to feature selection, a suitable classifier is required to achieve accurate and robust damage detection and localization [72]. The literature presents various combinations of features and classifiers determined through trial and error, searching for the optimal combination capable of characterizing structural damage.

However, several issues plague currently available ML based parametric and non-parametric SHM methods, including [53]:

1. There is no guarantee that a specific set of features/classifiers will be optimal for all types of structural damage detection scenarios. In other words, a feature/classifier combination that works well for a particular type of structure may not be suitable for other structures.

2. There is no guarantee that a given set of features/classifiers will be optimal for all types of structural damage. For example, a feature/classifier combination that works well for detecting stiffness loss may not be effective in detecting changes in boundary conditions.

3. Using a fixed set of handcrafted features or inappropriate classifiers may result in poor performance of the feature/classifier combination.

4. Feature extraction techniques can be computationally expensive and complex, making real-time SHM challenging when using ML-based methods.

5. Most parametric and non-parametric ML-based methods, with few exceptions, are centralized. This means that all measured signals must be synchronized and transferred to a single processing unit before damage detection can be performed.

#### 2.2.3.3 Vibration-based SHM based on DL

Conventional ANNs typically consist of a single hidden layer or two, making them shallow [73]. In contrast, a neural network with more than three layers is considered a DL network. Specifically, a deep neural network contains multiple layers between the input and output [74]. DL is the most recent breakthrough in the field of ML and has emerged as an active area of research. DL has made significant contributions to our daily lives by addressing challenging problems that were once considered insurmountable.

The main limitation of conventional ML-based methods is their reliance on handcrafted features, which may not optimally represent the measured signal, resulting in unreliable structural damage detection outcomes [75]. In contrast, DL overcomes this limitation by utilizing multiple layers that learn abstract representations of the data at various levels of complexity. With DL, the computational model can learn the entire

feature extraction process and map raw inputs to final outputs without prior feature extraction. This ability to break down complex tasks into simpler components allows DL algorithms to excel in challenging tasks.

DL architectures can be broadly categorized into four main groups: Unsupervised Pretrained Networks (UPNs), convolutional neural networks (CNNs), recurrent Neural Networks (RNNs), and Recursive Neural Networks. Within UPNs, there are three additional types of architectures: Autoencoders (also known as Deep Autoencoders), deep Belief Networks (DBNs), and generative Adversarial Networks (GANs) [76]. These categories and their corresponding architectures are illustrated in the accompanying figure 2.7.

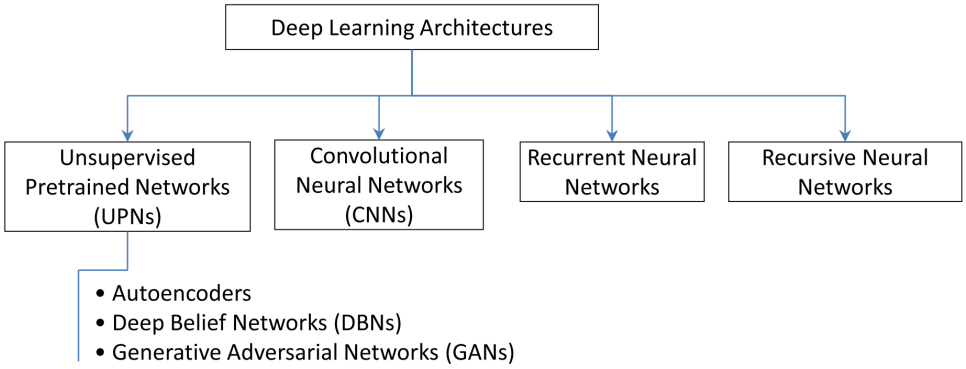



Figure 2.7: Deep-Learning architectures [76]

In particular, in this doctoral thesis, DL models are used to detect faults and damage in wind turbines. The questions posed in Section 1.2 are promptly addressed and answered below.

- How can a DL model be created using minimal data?

  Creating a DL model with minimal data can be a challenging task, as DL models typically require large amounts of data to learn meaningful patterns and generalize well to new, unseen data. However, several strategies can be used to address this problem:

  - **Transfer learning:** Transfer learning is a technique in which a previously trained model, trained on a large data set, is used as the starting point for a new model. The new model can achieve good performance with minimal data by fitting the previously trained model on a smaller data set [77].
  - **Data augmentation:** Data augmentation involves generating new training samples from existing data by applying transformations such as rotations, flips, and scales. This increases the size of the training set and helps prevent over fitting [78].

- **Bayesian optimization:** Bayesian optimization is a technique that can be used to optimize the hyper-parameters of a DL model with minimal data. By modeling the model's performance as a probabilistic function of the hyper-parameters, Bayesian optimization can efficiently explore the hyper-parameter space and find suitable configurations [79].
- **Siamese neural networks (SNNs):** SNNs is a type of DL model that can be used for tasks such as image recognition, similarity scoring, and text classification. They are particularly useful for scenarios where there is limited data available [80].

This thesis presents two different approaches to DL for SHM. The AP-I employs data augmentation by incorporating accelerometer vibrations in a serial manner while maintaining a constant sampling time. In contrast, the AP-II uses SNNs in its architecture, which have demonstrated high efficacy in developing DL models with limited data. Employing SNNs provides several benefits when working with minimal data, including:

- They are relatively simple and require minimal data preprocessing. Since the network is trained on pairs of examples, creating a large set of labeled data is unnecessary.
- They are easy to train and optimize. Siamese neural networks can be trained using standard optimization techniques such as Adam or stochastic gradient descent, and the similarity metric can be chosen based on the specific task.
- They can be easily adapted to new tasks. Siamese neural networks can be tuned for new tasks by adjusting the weights of the shared network or by adding new layers on top of the network.

Additionally, the bayesian optimization is explored in the context of CM, particularly in AP-III.

- How can imbalanced datasets be effectively handled?

  Unbalanced data sets can present a challenge for ML algorithms, as they tend to be biased towards the majority class and perform poorly in the minority class. However, there are several strategies that can be used to handle imbalanced data sets effectively:

  - **Resampling:** Resampling involves oversampling the minority class or sub sampling the majority class to balance the data set. Oversampling can be done by randomly duplicating examples from the minority class, generating synthetic examples, or both. Sub sampling can be done by randomly removing examples from the majority class or by selecting representative examples [81].

  - **Cost-sensitive learning:** Cost-sensitive learning involves modifying the algorithm to give more weight to the minority class or penalize more misclassifications of the minority class. This can be done by adjusting the loss function or by assigning different costs to different types of errors [82].
  - **Ensemble methods:** Ensemble methods involve combining multiple models to improve performance. In the case of unbalanced data sets, one approach is to train multiple models on different subsets of data, with each subset balanced in terms of class distribution. The models are then combined using techniques such as majority voting or weighted average [83].
  - **Anomaly detection:** Anomaly detection involves identifying the minority class as an anomaly or outlier and training a model to detect it. This approach can be useful when the minority class is inherently different from the majority class and can be identified using features that are not present in the majority class [84].
  - **Transfer learning:** Transfer learning involves using a pre-trained model on a large data set and fitting the unbalanced data set. This approach can be useful when the imbalanced data set is too small to train a model from scratch [77].

  In this thesis, anomaly detection in the context of SHM has not been extensively explored using an unsupervised approach, and therefore is proposed as a potential avenue for future research. However, in the field of CM, fault detection is proposed using anomaly detection models for the AP-III and AP-IV.

- How can time series data be efficiently processed?

  Effectively processing time series data in DL requires careful consideration of the data's unique characteristics, including time dependencies, variable-length sequences, and noisy measurements. To this end, the following strategies can be employed to optimize the processing of time series data in DL:

  - **RNNs:** RNNs are a type of DL model that is specifically designed to handle sequential data. They have been widely used for time series data due to their ability to capture temporal dependencies between data points. RNNs process input data one-time step at a time and use a cache state to maintain a memory of previous inputs [85].
  - **Long short term neural networks (LSTMs):** LSTMs are a type of RNN that is designed to address the problem of gradient disappearance that can occur with traditional RNNs. They have a memory cell that can selectively remember or forget information, allowing them to handle long-term dependencies on time series data [86].
  - **CNNs:** CNNs are a type of DL model commonly used for image processing but also for time series data. CNNs can be used to process two-dimensional time series data by applying convolutions along the time dimension [87].

- **Attention mechanisms:** Attention mechanisms allow the model to focus on specific parts of the input sequence that are most relevant to the task. This can be useful for time-series data, where certain parts of the sequence may be more important than others [88].
- **Ensemble:** Ensemble combines multiple models' predictions to improve performance. This can be especially useful for time series data, where multiple models can capture different aspects of the data [89].

This thesis focuses on the SHM approach using CNNs, which are particularly efficient at working with time series data by extracting local features. The application of this approach is discussed in both AP-I and AP-II, where incipient damage is detected. Prior to feeding time-series data into a convolutional neural network (CNN), pre-processing may be required to convert the raw vibration measurements into 2D matrices. Specifically, multiple-channel matrices are used to represent the vibration data captured by multiple sensors. Additionally, recurrent neural networks are explored in the context of CM, particularly in AP-IV.

## 2.3 Comparison of the strategies stated in AP-I and AP-II

In this chapter, a comparison is conducted between the two strategies proposed in the two appended papers on SHM. The objective of the comparison is to evaluate the effectiveness of each strategy in the context of SHM. The comparison is based on a set of criteria relevant to SHM, including accuracy, efficiency, and robustness. The results of the comparison provide insights into the relative strengths and limitations of the two strategies, and may be useful for practitioners in selecting an appropriate strategy for a given SHM application.

To ensure a fair comparison between the two papers, the same experimental benchmark (down-scaled jacket-type support tower) and type of structural states (a healthy bar, a replica bar, a cracked bar, and a loose bolt) were used. This involved using the same initial data set for both papers, which ensured that any differences in performance were due to the specific strategies being evaluated, rather than differences in the underlying data. Finally, note that both papers used grayscale images as the primary input data for their respective approaches. These images were obtained by pre-processing the raw data in the same manner, that is same image dimension and number of channels ($16 \times 16 \times 24$) were used. By using the same experimental setup, raw data, and pre-process, it was possible to directly compare the effectiveness of the two different approaches in a controlled manner. Figure 2.8 shows the general framework of the AP-I and AP-II proposed strategies.

The first paper (AP-I) in this comparison focuses on exploring a data augmentation strategy. Data augmentation involves generating additional, synthetic data samples from existing data, and can be used to improve the performance of ML models by increasing

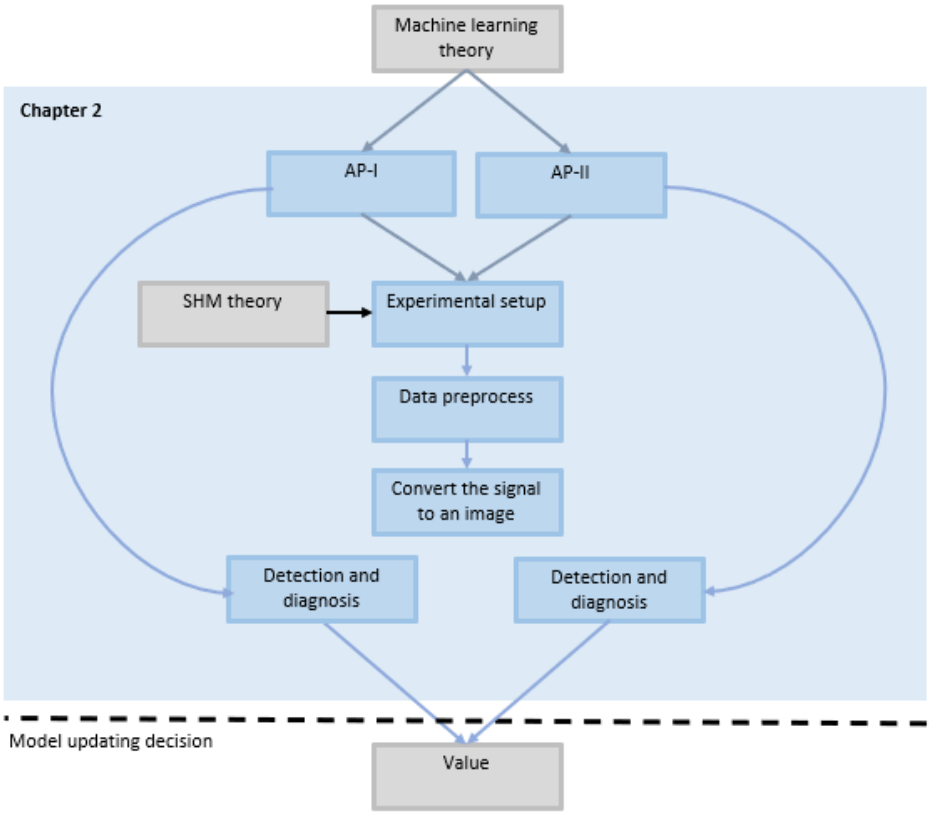



Figure 2.8: Framework of AP-I and AP-II

the size of the training dataset. By using a data augmentation approach, the accuracy and robustness of the strategy, based on a deep CNN, are improved. In this case, six convolutional layers and two fully connected layers at the output for damage detection and diagnosis are used. In contrast, the second paper (AP-II) focuses on obtaining the best results with only the available data, which is a common challenge in SHM due to the typically limited amount of data available. To address this challenge, a siamese neural network (SNN) is used. A SNN is a type of ML model that uses two (or more) identical neural networks to process inputs and compare their outputs. Siamese neural networks are a powerful tool for addressing the problem of having few data in comparison to other neural networks because they are designed to learn similarity between pairs of inputs, rather than simply classifying individual inputs. By learning to identify the similarity between pairs of data points, Siamese networks can effectively leverage the limited amount of data available to create accurate and robust models. The structure of an SNN is composed of two identical CNNs connected in parallel, where their outputs compare the similarity of responses with the euclidean distance. The first SNN (damage detection) is designed in its architecture with a convolutional layer and a fully connected layer; while the second SNN (damage diagnosis) comprises two convolutional layers and a fully connected layer. The reason for the difference in architecture between the first

SNN (damage detection) and the second SNN (classifying type of damage) is likely due to the differences in the tasks they are designed to perform. The first SNN is designed for damage detection, which involves identifying the presence or absence of damage in a structure. The use of a single convolutional layer and a fully connected layer in the architecture may be sufficient for this task since convolutional layers are effective at detecting features in images, and the fully connected layer can learn to map the extracted features to a binary output indicating the presence or absence of damage. On the other hand, the second SNN is designed to classify the type of damage present in a structure. This task is more complex than simple damage detection, as it involves identifying the specific type of damage present. Therefore, the architecture of the second SNN comprises two convolutional layers and a fully connected layer, which can extract more complex features from the input images and learn to map these features to multiple output classes representing the different types of damage. In summary, the differences in the architecture of the two SNNs are due to the differences in the complexity of the tasks they are designed to perform. The first SNN focuses on binary damage detection, while the second SNN requires the ability to classify multiple types of damage. Figure 2.9 shows the architecture used in AP-I and Table 2.2 shows the number of parameters. Figure 2.10 and Figure 2.11 show two models of the SNN architectures used in AP-II; Table 2.3 and Table 2.4 show the number of the SNN architectures. The use of SNNs represents a novel approach to SHM, and the results of this paper demonstrate the potential of this approach for solving SHM problems with limited data. Finally, Table 2.1 shows the number of images used in the AP-I and AP-II papers.

Table 2.1: Number of images in AP-I and AP-II

| | Number of images | |
|---|---|---|
| AP-I | 1,612,900 | Data augmentation |
| | 6,400 | Initial dataset |
| AP-II | 6,400 | Initial dataset |

In addition to using different architectures, the two papers in this comparison also partitioned the image dataset differently. In the first paper, a direct division was made between the training and validation sets, with 75% of the data being used for training and 25% being used for validation. This division was based on the data labels, with the goal of training and evaluating the ML model on distinct sets of data. In contrast, the second paper partitioned the data into training, validation, and test sets based on pairs of images. These pairs could be either positive (same labels) or negative (different labels), resulting in two classes of images (both in the SNN for damage detection and in the SNN for damage diagnosis). Table 2.5 shows the image division of the AP-I and Table 2.6 shows the image division of the AP-II.

As for the results obtained, table 2.7 shows the results of the AP-I, where the recovery,

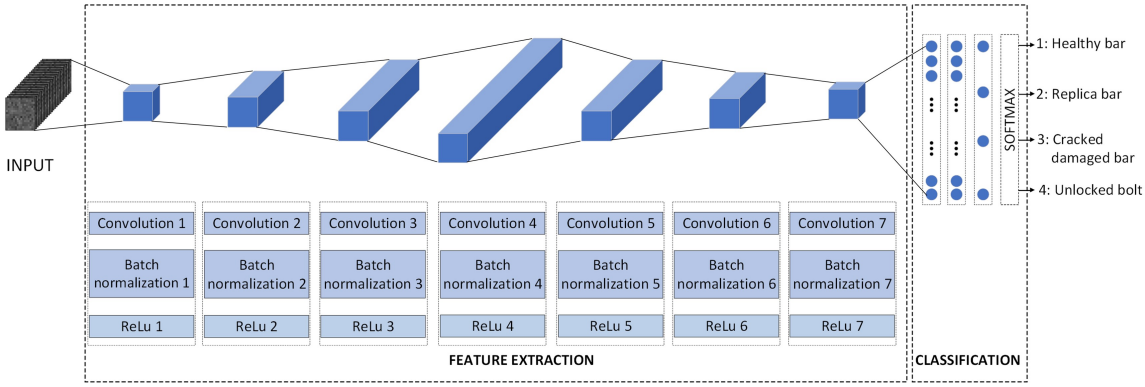



Figure 2.9: CNN architecture used in AP-I

Table 2.2: Number of parameters in each layer of the CNN architecture of the AP-I. The neural network has a total of 2,176,308 parameters.

| Layer | Output size | Number of parameters |
|---|---|---|
| Input (16×16×24) | 16×16×24 | - |
| Convolution#1 | 14×14×32 | 19,296 |
| Convolution#2 | 12×12×64 | 51,392 |
| Convolution#3 | 10×10×128 | 205,184 |
| Convolution#4 | 8×8×256 | 819,968 |
| Convolution#5 | 6×6×128 | 819,712 |
| Convolution#6 | 4×4×64 | 204,992 |
| Convolution#7 | 2×2×32 | 51,296 |
| Fully connected layer#1 | 32 | 4,128 |
| Fully connected layer#2 | 16 | 528 |
| Fully connected layer#3 | 4 | 68 |

Table 2.3: Number of parameters in each layer of the SNN model 1 architecture of the AP-II. The neural network has a total of 81,283,488 parameters.

| Layer | Output size | # of parameter |
|---|---|---|
| Input (16×16×24) | 16×16×24 | - |
| Convolution | 16×16×64 | 13,888 |
| Fully connected layer | 4,960 | 81,269,600 |

precision and F1 score are shown. The best results are obtained with data augmentation where the values of each metric are greater than 99.88%. This methodology shows better performance for large amounts of data, in this case data augmentation is larger than initial data set by 25,200%. While for the AP-II, the table 2.8 shows the data for the damage detection SNN and the damage diagnosis SNN, showing recovery, precision and

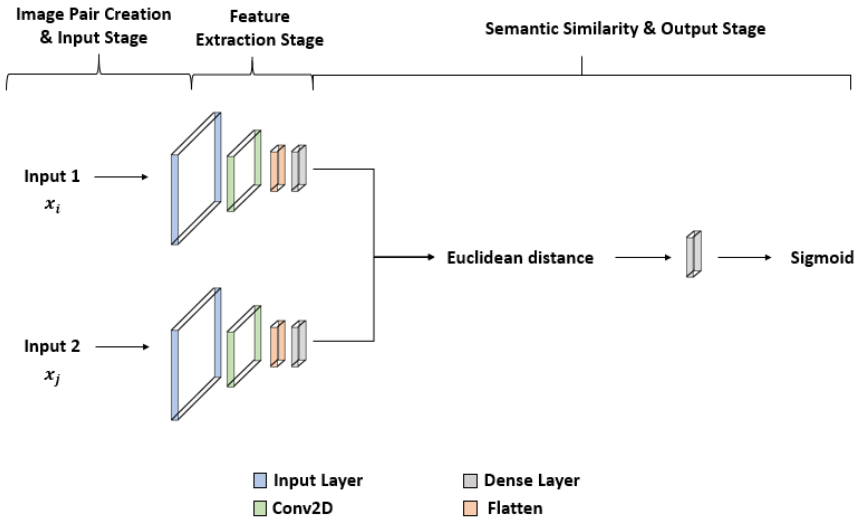



Figure 2.10: SNN architecture of a convolutional layer (model 1) used in AP-II

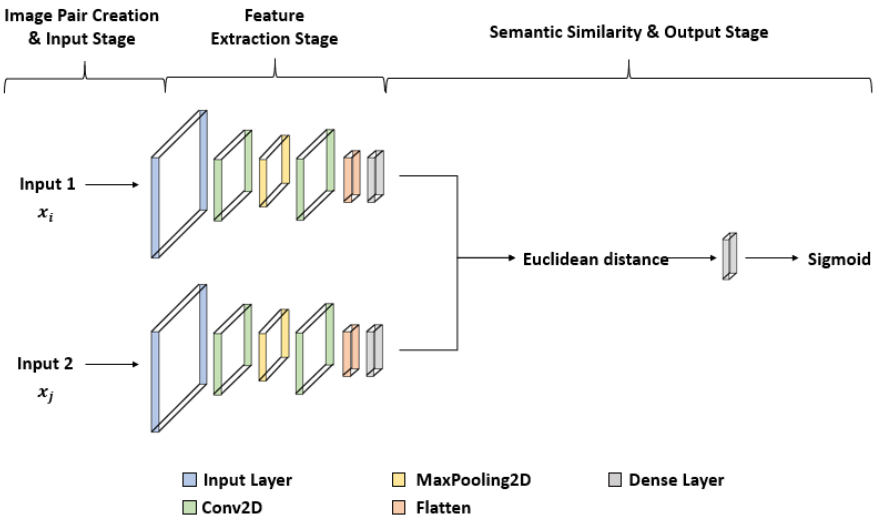



Figure 2.11: SNN architecture of two convolutional layers (model 2) used in AP-II

F1 score. In this case, the best results are obtained with model 2, where all the metrics are 100%.

About precision and training time, Table 2.9 shows the results of the AP-I, the drastic difference in training time between data augmentation and initial data set is evident, the use of a data set with a greater number of images leads to a greater load of the use of computational resources due to the number of parameters of the architecture (see Table 2.2) and therefore a greater training time. As expected, the data augmentation leads to better accuracy results than initial data set. The results of the AP-II are shown in Table

Table 2.4: Number of parameters in each layer of the SNN model 2 architecture of the AP-II. The neural network has a total of 40,725,024 parameters.

| Layer | Output size | # of parameter |
|---|---|---|
| Input (16×16×24) | 16×16×24 | - |
| Convolution#1 | 16×16×64 | 13,888 |
| Max pooling | - | - |
| Convolution#2 | 8×8×128 | 73,856 |
| Fully connected layer | 4,960 | 40,637,280 |

Table 2.5: Number of images for training and validation in AP-I

| | Training | Validation |
|---|---|---|
| # of images (data augmentation) | 1,209,675 | 403,225 |
| # of images (initial data set) | 4,800 | 1,600 |

Table 2.6: Number of images for training, validation and test in AP-II

| | | Training | Validation | Test |
|---|---|---|---|---|
| SNN for damage diagnosis | Positive pair images | 6,144 | 768 | 768 |
| | Negative pair images | 6,144 | 768 | 768 |
| SNN for damage detection | Positive pair images | 2,048 | 256 | 256 |
| | Negative pair images | 2,048 | 256 | 256 |

2.10, where model 2 has better accuracy than model 1. In addition, model 2 has a shorter training time due to the difference in parameters (see Table 2.3 and Table 2.4).

Table 2.7: Metrics for each label of the diagnosis and detection problem and comparison between the datasets in AP-I

| Dataset | label | Precision | Recall | F1-score |
|---|---|---|---|---|
| Data augmentation | 1: Healthy bar | 99.89 | 99.96 | 99.92 |
| | 2: Replica bar | 99.90 | 99.87 | 99.88 |
| | 3: Crack damaged bar | 99.94 | 99.86 | 99.90 |
| | 4: Unlocked bar | 99.90 | 99.86 | 99.88 |
| Initial data set | 1:Healthy bar | 97.97 | 94.14 | 96.02 |
| | 2: Replica bar | 90.31 | 94.75 | 92.48 |
| | 3: Crack damaged bar | 90.31 | 92.63 | 91.46 |
| | 4: Unlocked bolt | 92.50 | 93.38 | 92.94 |

Table 2.8: Metrics for the diagnosis and detection problem and comparison between models in AP-II

| | | Precision | Recall | F1-score |
|---|---|---|---|---|
| Damage detection | model 1 | 97.66 | 99.60 | 98.62 |
| | model 2 | 100 | 100 | 100 |
| Damage diagnosis | model 1 | 94.92 | 98.78 | 96.81 |
| | model 2 | 100 | 100 | 100 |

Table 2.9: Accuracy, validation error, training error, and training time when using data augmentation and initial data set in AP-I

| | Accuracy | Training time | # of images |
|---|---|---|---|
| Data augmentation | 99.90 | 1196 min | 1,612,800 |
| Initial data set | 93.81 | 11 min | 6,400 |

Table 2.10: Accuracy, validation error, and training error when using model 1 and model 2 in AP-II

| | | Accuracy | Training time | # of pair of images |
|---|---|---|---|---|
| Damage detection | model 1 | 98.63 | 39.40 min | 6,912 |
| | model 2 | 100 | 22.51 min | 6.912 |
| Damage diagnosis | model 1 | 96.87 | 27.45 min | 2,304 |
| | model 2 | 100 | 15.75 min | 2.304 |

Finally, a conclusion about which methodology is more suitable for damage detection will be given in section 4.1.

# Chapter 3

# Condition Monitoring

## 3.1 State of the art in wind turbine CM

CM and predictive maintenance (PdM) are related concepts, but they have distinct meanings and approaches, see [90]. CM is the process of monitoring the condition of a machine or system in real-time, typically through sensors and other data collection methods. The goal of CM is to identify potential problems or anomalies as soon as they occur, so that corrective action can be taken before significant damage or downtime occurs. CM is typically based on statistical methods and signal processing techniques, and does not require the use of machine learning or artificial intelligence. On the other hand, PdM is a broader approach to maintenance that incorporates CM as well as other techniques, such as machine learning and AI, see [91]. PdM uses historical data and advanced analytics to make predictions about when maintenance should be performed, based on the expected lifespan of components or the likelihood of failure. PdM is focused on maximizing the reliability and availability of assets, while minimizing maintenance costs and downtime.

- **Drivetrain.-** This system comprises the slow, multiplier, and high-speed axes. The low-speed slow shaft of the WT connects the rotor hub to the multiplier, inside the hydraulic and electrical system run conduits, which actuate the aerodynamic brakes. The multiplier, at its input, is the low-speed shaft, and through a gearing system, the high-speed output shaft rots at a higher frequency (between 80 and 50 times faster, depending on the turbine model). The high-speed shaft rotates at approximately 1,500 revolutions per minute (RPM), allowing the operation of the electric generator. It is equipped with an emergency mechanical disc brake.

- **Hydraulic system.-** This system comprises the pressure group, hydraulic lines, and control valves. The pressure group is responsible for providing hydraulic fluid at a certain pressure to allow the activation of capture, orientation, or transmission

systems. Hydraulic lines channel the fluid to the point of use. The control valves adapt the pressure and flow of the fluid based on the actuator to be actuated.

- **Cooling systems.-** This system is made up of fans and heat exchangers. The fans work at the request of the controller to create air circulation. The heat exchangers dissipate the heat from the component to be cooled (generator, multiplier, or hydraulic unit) towards the air current created by the fans.

- **Generation system.-** This system is made up of the generator and the transformer. The generator converts the mechanical energy produced by the rotor into electrical energy. In the transformer, the electrical power delivered by the generator is low voltage. Through the transformer, the voltage is raised to connect to the medium voltage distribution network.

- **Guidance system.-** This system is composed of an orientation and rotation system, an anemometer, and a wind vane. In the orientation and rotation system, the nacelle rotates on the top of the tower by means of an active yaw and rotation control system so that the rotor is always in a position transverse to the wind direction, the orientation of the WT changes based on wind conditions recorded by sensors located on the nacelle deck. The correct positioning signal is received from the turbine controller, based on the readings of the vane and the anemometer.

- **Control system.-** This system consists of the turbine controller, control sensors, and regulation signals. The turbine controller continuously monitors the conditions of the WT, collects statistics on its operation, and regulates switches, hydraulic pumps, valves, and other elements of the WT. Control sensors are used to measure the physical parameters of turbine operation and monitor turbine operation. The electronic controller uses the signals to connect the WT when the received signal is correct. To protect the turbine, the controller would stop the equipment automatically if the information received from the sensors is wrong. The control and regulation signals of the WT controller, based on the analysis of the sensor data, generate orders that affect the operation and functioning of the WT.

The most common faults of these systems are mechanical, electrical, or electronic. Faults can be divided into two large groups: minor corrective, which are faults that are typically solved in less than 24 hours and that cause small production losses. The second group is the major corrective, which are serious failures in the main elements, which can involve an insurance company or the manufacturer of the WT and cause significant production losses. Of the total problems in WTs, 40% are related to minor corrective actions, 30% to major corrective actions, 10% to preventive activities, and 20% to substation incidents [92].

CM in WTs involves analyzing the operating condition parameters or its components. The objective of CM is to identify any abnormal changes in the condition or events that may indicate a fault (failure in development) [93]. This type of maintenance is profitable, optimal, and less expensive than maintenance based on service time. Monitoring of the

state of the WT is carried out online or offline. In the case of offline CM, it implies periodic inspections where the WT must stop, resulting in a loss of energy production, not to mention the additional costs incurred during offline inspections. Additionally, offline CM falls short of detecting and reporting faults between inspection intervals. The above deficiencies become increasingly severe, especially in offshore wind farms, where inspection intervals are more extended than in onshore wind farms [94]. On the other hand, online CM provides a real-time view of the health of the WT or its main components. It is based on data acquisition systems to obtain continuous measurements of WT components. Real-time CM is highly customizable due to the wide variety of sensors and data acquisition systems available for WT components. Factors such as wind farm environment, turbine design specifications, type and characteristics, technology limitations, and costs affect real-time (online) condition monitoring system (CMS) [95].

For successful maintenance management, information on the status of the turbine is essential. Traditionally, information was acquired through manual on-site inspections. However, with the increasing number of turbines installed at remote sites, frequent inspections are becoming more challenging and expensive. Therefore, new CM strategies are developed, combining new sensor technology with online and offline data analysis.

### 3.1.1 Specific CM strategies

In the following paragraphs, a review of specific CM techniques used in the context of WTs is comprehensively stated.

- **Vibration signal analysis .-** Vibration analysis is the most widely used technique to control the state of rotating equipment. Different sensors are required for different frequencies: position transducers are used for the low-frequency range, velocity sensors in the middle-frequency and accelerometers in the high-frequency range, and spectral emitted energy sensors for very high frequencies [96]. In the context of WT application, vibration monitoring covers the overall drive train of the turbine where rotating machinery.

  Figure 3.1 illustrates a typical accelerometer design for vibration monitoring purposes in a WT drive train with seven acceleration transducers. The collected vibration measurements are typically subjected to frequency-based analysis and the spectrum obtained provides an explicit indication of a component condition that facilitates fault diagnosis for specific components.

  In case of quality degradation within the component, large harmonics could appear in the spectrum or increase the energy contained in the sideband of the spectrum. Very detailed knowledge about the drive train parameters is required, including the dimensions for each subassembly within the gearbox and the number of gear teeth for all stages of the gearbox, in order to gain better understanding of which part of the spectral signal corresponds to normal operation and which part is caused by the deterioration of the component [98]. In addition to frequency domain-based techniques, anomaly detection can also be carried out using time series analysis

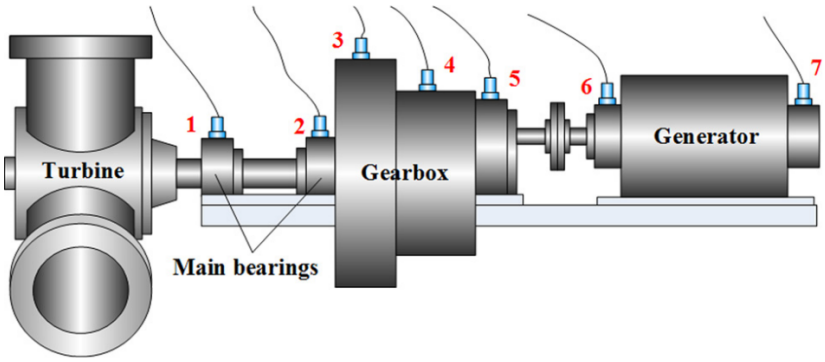



Figure 3.1: Structure of WT drive train with seven acceleration transducers installed for CM [97].

of vibration signals, where the development of imminent gearbox bearing fault is indicated. through an increasing trend over time.

- **Oil debris analysis .-** Oil debris analysis is also a potentially useful method of monitoring gearbox health. Oil analysis is mostly carried out offline by taking samples (thus, it has a very low sampling rate), despite the fact that online sensors have (for years) been available at an acceptable price level for monitoring oil temperature, contamination, and moisture [96]. Sometimes, it is integrated with vibration analysis to provide a complete gearbox CMS that can achieve a better and more accurate detection of gearbox faults [99].

  Debris found in gearbox lubricating oil can be considered an indication of wear or damage to gearbox components, where particles of different sizes and materials imply different types and positions of damage. A commonly used device, the induction sensor, is used to check the size and amount of ferrous and non-ferrous debris in lubricating oil. After particle count, oil is typically pumped into the filtration system to remove debris, before returning to the gearbox [100].

  The evolution of gearbox damage in terms of wear particle size is illustrated in Figure 3.2, from which it can be seen that the detection of large ferrous wear particles (with a size greater than 100 micron) can provide an early indication of possible gearbox subassembly wear. However, for small and medium particle size ranges, the rate of increase of the oil count particles will be more informative than the absolute value of the cumulant.

- **Strain and optical monitoring .-** Recently, strain measurement and optical fiber monitoring for WT structures has received increasing attention, as the fatigue loads to which the turbine is exposed can be estimated. The measurements of strain gauges, which can be placed randomly on the structure, are processed with the help of a finite element method to monitor the effects of high dynamic loads. However,

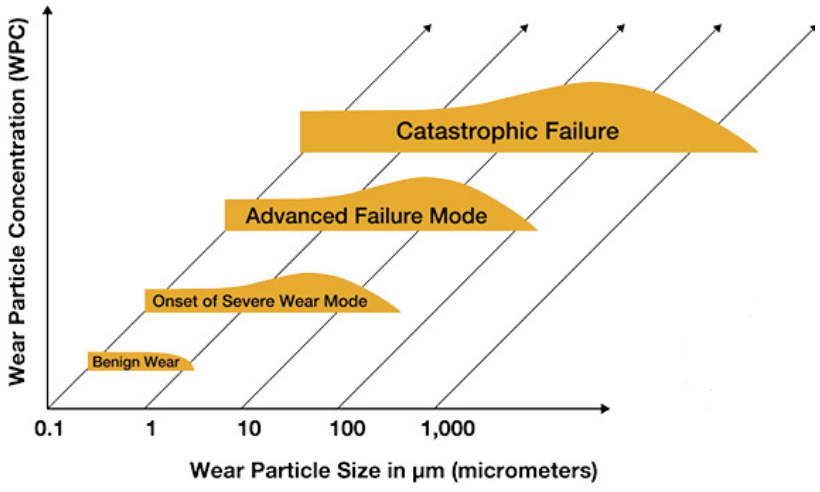



Figure 3.2: Evolution of gearbox damage in terms of wear particle size and concentration [99].

strain gauges are not very durable, and these techniques require expensive measurement equipment. New approaches try to connect available SCADA-data measurements and short-term strain measurements to extrapolate strain estimations. Such applications might help the technology reach a broader application in the future [101].

- **Temperature monitoring .-** Temperature monitoring is one of the most popular CM tools applied in WT. This can be done with thermometers and infrared thermography. Since each component has a maximum operating temperature that is generally exceeded only in the event of abnormally high friction, it is a reliable criterion for fault detection. Also, the temperatures are measurements that change slowly due to the thermal inertia of the components. This can be an advantage when analyzing data with a low sampling rate, for example, 10-minute average values stored in a SCADA system. For temperature, this may be sufficient resolution for CM. On the other hand, slowly changing measurements are of limited value in early fault prediction because they simply indicate a fault too late. However, temperatures are often used as a secondary criterion in the case the vibration control shows an alarm [101].

- **Acoustic emission .-** Acoustic emission can detect faults characterized by high-frequency vibrations ranging from 50 kHz to 1 MHz. Piezoelectric transducers and optic fiber displacement sensors are often employed in this approach [102]. Acoustic monitoring consists of two types: passive, where the excitation is produced by the component itself; and a second type, where the excitation is applied externally. Acoustic emission monitoring is closely related to vibration monitoring, and some

recent systems combine the two techniques to obtain a more accurate condition indicator. Acoustic emission signals are also used to identify potential blade fatigue in critical areas, such as the blade root [98].

The technologies presented in the previous paragraphs are mainly used to monitor a specific subsystem within the turbine. Other approaches widen the balance limits and aim to monitor the global WT system. Different mechanical and electrical faults for example, lead to disturbances in the mechanical as well as the electrical energy flow. Consequently, mechanical torque oscillation can also be detected on the electrical side of the power train through power signal analysis. That way, blade, or rotor imbalances can be detected. A comparably simple method is the monitoring of process parameters. There, the values and relationships of temperatures, power, wind, and rotor speed or blade angles are compared with specifications and limits determined by manufacturers. For this kind of analysis, SCADA signals can be used.

The importance of CM is expected to increase further in the future due to the developments mentioned above in the wind industry. The more mature the new techniques become, the cheaper their application becomes. Also, the cost of condition monitoring can be compensated with lower premiums for insurance that rewards such systems. Developing more reliable, cost-effective, integrated, and smart solutions, condition monitoring is about to become an integral part of modern maintenance strategies.

### 3.1.2 CM methodologies based on SCADA data

Large-scale WTs operate through a SCADA system. The system usually samples data at low frequency (typically 1 Hz) with standard practice to store 10-minute averaged values of the parameters characterizing the operating and environmental conditions. The number of channels available varies considerably between manufacturers and SCADA services providers, although the minimum set typically includes wind speed and direction, active and reactive power, rotational speed, pitch and yaw angles, and ambient temperature. In the following, CM methodologies based on SCADA data are reviewed.

- **Power curve-based modeling-** The relationship between the power output of the WT and the corresponding wind speed experienced by the turbine rotor defines a WT power curve, and this provides a fundamental but important metric to identify the operational health of the WT. A key attraction for CM purposes is that power curves can be calculated with an acceptable degree of accuracy using already available SCADA data from wind farms [103]. To date, power curves have been used primarily by WT owners (purchasers of the turbines) to ensure that turbines supplied by original equipment manufacturers meet their specifications, of which the power curve of the manufacturer is an essential part. With the help of proper techniques, regularly updated power curves can also provide a convenient means of identifying whether operating WTs continue to function well.

- **Trending-** A simple way to determine the health of the WT is to track the trends of various parameters over time to see if any obvious changes can be observed visually.

The thinking behind this is that any obvious change in a trend may indicate a change in the internal physical state of some turbine components and therefore warrants further investigation to diagnose the source of the change.

In its most basic form, turbine operators can visually observe trends and use domain knowledge and experience to spot performance deviations or diagnose faults. For example, the diagnosis of obvious changes in the shape of the power curve during different periods can indicate problems in the control system. As mentioned above, this is widely practiced in the industry. However, subtle changes in shape can be challenging to detect, so a more analytical approach may yield better results.

Looking at general trends in turbine data can show indications of faults before the fatal failure occurs. These trends can be easy to implement, however, the downside of such approaches is that they rely on human visual interpretation and domain knowledge. In addition, specific trends related to different types of faults must be tracked. This leads to a heavy workload for the operator, with little scope for automatic alarm generation or automation [104].

- **Normality behavior modeling-** Different techniques are used to detect anomalies, ranging from simple threshold checks to complex statistical analysis. A common approach is the application of normal behavior models. On the basis of the inputs extracted from the SCADA data, the model should be able to predict a target parameter under normal operating conditions. The real-time signal is compared to the estimated model output for anomaly detection. The accuracy of the developed model determines the success of the approach. Here, artificial intelligence methods have proven to be a sufficient tool for modeling complex systems [101].

Since SCADA systems were not initially designed for CM purposes, the performance provided is limited, mainly due to the following concerns and shortcomings [94].

- **1.** Although SCADA data provides relatively extensive information that can be useful in identifying abnormal WTs in a wind farm, the data often does not include all the information necessary for detailed and complete monitoring of WT subsystems and components.

- **2.** The distribution of SCADA data is generally unbalanced, and the extraction of anomalous data is often insufficient. This means that the amount of normal data is often much larger than that of anomalous data, which could lead to poor health monitoring performance, as data-driven models tend to be biased towards the majority class.

- **3.** SCADA data sampling rate is much slower than required for some CM techniques.

- **4.** Data quality is often a concern. For example, 10-minute records of SCADA data are often affected by problems such as missing values, "NULL" entries, statistical

outliers, large blocks of consecutively identical values, bad data format, etc. Consequently, any necessary corrections made to the data may lower the required level of precision.

- **5.** Given the variability in operating conditions of modern WTs, the rapid fluctuations in environmental conditions (wind speed and direction, air density, turbulence, etc.), and the low sampling frequency of typical SCADA systems, it is difficult to detect, diagnose and forecast incipient faults promptly.

Due to the limitations listed above, SCADA-based CM cannot replace a specifically purpose-build CMS, which is specially designed for WT CM purposes. Compared to SCADA systems, the measured signals are sampled at higher frequencies, providing richer information and allowing greater understanding of the operating state of the WT. This comes at a higher cost, which depends on the measurement accuracy, sampling rate, system functionality, and application environment. However, CM based on SCADA data is a cost-effective solution for WTs that are close to their useful life and do not have CMS, as its implementation does not imply adaptations and sensors, which makes it attractive for such situations.

## 3.2 Comparison of the strategies stated in AP-III and AP-IV

This chapter compares the strategies proposed in the two appended papers related to CM. The objective of the comparison is to evaluate the effectiveness of each proposal by comparing performance and efficiency. The results provide information on the limitations and strengths of each proposal, which is useful for the selection of appropriate strategies by professionals. To ensure a fair comparison of the two papers, the same SCADA variables from the same operating wind farm consisting of 12 WTs, the same fault of interest (the main bearing fault), the same data preprocessing, and the same splitting of training and test (see figure 3.4) are used. Figure 3.3 shows the general framework of the AP-III and AP-IV strategies. Table 3.1 shows the selected SCADA data variables.

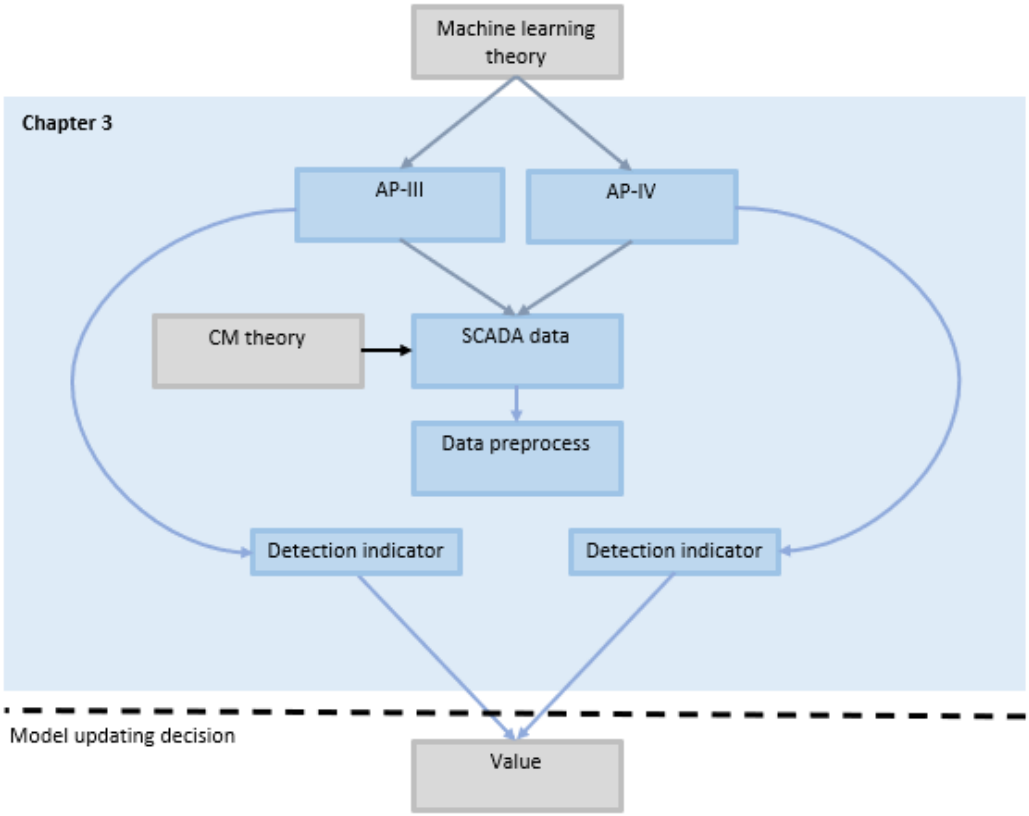



Figure 3.3: Framework of AP-III and AP-IV.

Table 3.1: Selected SCADA variables used to develop the normality model in AP-III and AP-IV, its description, range of possible values and units.

| Variable | Description | Range | Units |
|---|---|---|---|
| Pot | Generated real power | [0, 2000] | kW |
| TempAmb | Ambient temperature | [−5, 40] | °C |
| TempCojLA | Bearing coupling side temperature | [0, 120] | °C |
| TempCojLOA | Bearing non-coupling side temperature | [0, 120] | °C |
| TempEjeLento | Low-speed shaft temperature | [0, 120] | °C |
| TempGen | Generator temperature | [0, 175] | °C |
| TempRodamMultip | Gearbox temperature | [0, 120] | °C |
| VelRotor | Rotor speed | [0, 50] | rpm |

The paper (AP-III) proposes a normality model based on an artificial neural network (ANN). The stated ANN is shown in Figure 3.5, where the output of the ANN is the low-speed shaft temperature at time $t$, and the inputs are the remaining seven variables at time $t-1$ and $t$. In total, the model comprises 14 inputs and 1 output, with a hidden

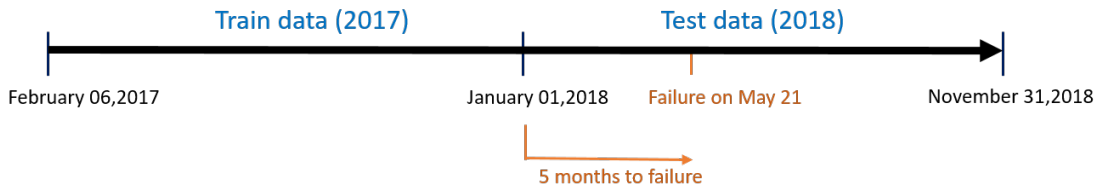



Figure 3.4: WT2 (WT number 2 in the wind farm) data for training and test of AP-III and AP-IV methodologies.

layer made up of 72 neurons. A weekly fault indicator with a range from 0 to 1 is proposed, where a function counts the times a threshold is exceeded based on the mean and standard deviation obtained from the absolute value of the training residuals that comprise the prediction of the model and the real value. Figure 3.2 shows the number of parameters of the ANN architecture, it should be noted that this model was trained with Bayesian regularization, which indicates the number of parameters sufficient for training.

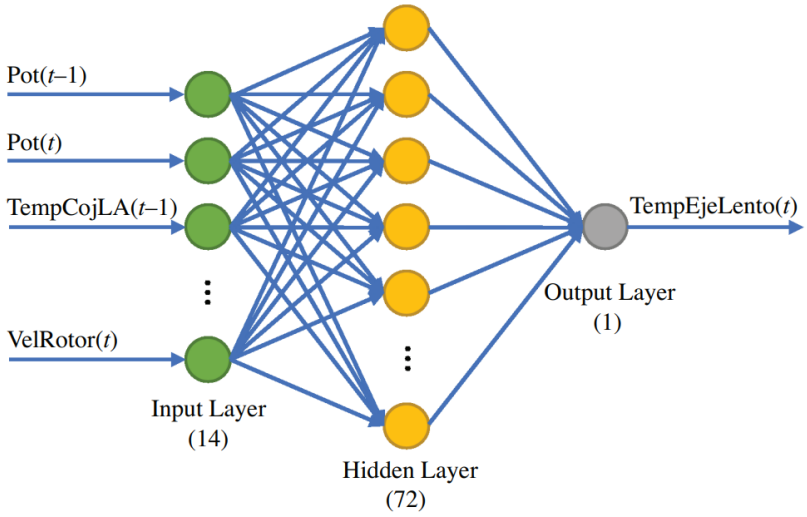



Figure 3.5: ANN of the AP-III method.

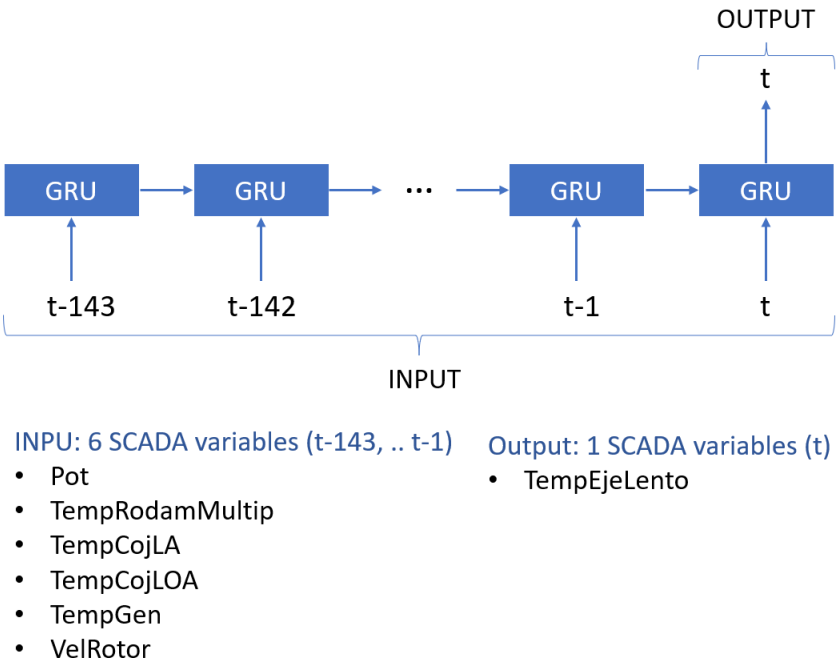
In contrast, the paper (APIV) proposes a normality model based on a GRU network, where the output corresponds to the low-speed shaft temperature at time $t$ and the inputs are 6 variables detailed in Figure 3.6( note that the temperatures are seasonally adjusted with the ambient temperature) in 144-time steps, that is, $(t-143, t-142, ..t-1, t)$. In total, the model has 6 inputs and one output with a hidden layer made up of 128 neurons. The proposed indicator uses the EWMA to predict the trend at time $t+1$. It is based on the residuals of the model prediction and the real value, squared. The alarm is

Table 3.2: Number of parameters in each layer of the ANN architecture for AP-III. The neural network has a total of 1,153 parameters

| Layer | Output size | # of parameter |
|---|---|---|
| Input (1×14) | 1×14 | - |
| Linear 1 | 1×72 | 1,080 |
| Linear 2 | 1×1 | 73 |

obtained based on the standard deviation and the mean obtained from the EWMA. The alarm is tuned based on the value of the standard deviation and the number of spans of the EWMA. Figure 3.6 shows the GRU model, and Table 3.3 shows the number of parameters.

Figure 3.6: Many-to-one GRU architecture of the AP-IV method.




Regarding the results, Figure 3.7 shows the indicator and residues obtained from AP-III in the test data set. The first activation corresponds to WT2, this alarm is a true positive. It was activated on February 4, with the main bearing fault being detected three and a half months in advance. The second activation shown occurs with WT8. In this turbine, the normality model detects a maintenance that is bein carried to replace the gearbox. In a real situation, the park manager knows that this WT is already under

Table 3.3: Number of parameters in each layer of the GRU architecture for AP-IV. The neural network has a total of 52,353 parameters.

| Layer | Output size | # of parameter |
|---|---|---|
| Input (1×144×6) | 1×144×6 | - |
| GRU | 1×144×128 | 52,244 |
| Layer | 144×1 | 129 |

maintenance. Therefore, it should not be counted as a false positive. Finally, WT9 has an active alarm, a false positive of the method, as this WT had no major work orders during 2018. The rest of the park's WTs were correctly classified as healthy across all test data sets. In summary, Table 3.4 shows the activated alarms.

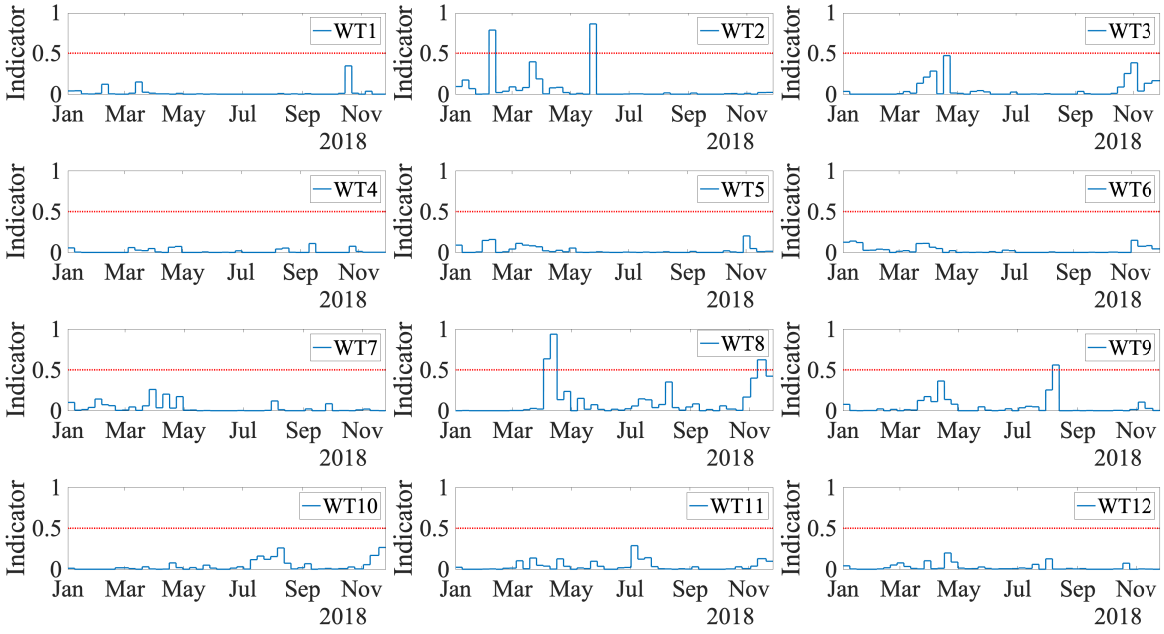



Figure 3.7: ANN indicator values (blue line) for test data, and threshold (red line) in the AP-III method.

With the results of AP-IV, Table 3.5 shows the tuning of the indicator. The aim is to propose a threshold of the form $\mu + \kappa\sigma$. To select the value for $\kappa$ it is proposed that the false-positive alarms over the training and validation dataset for each WT are minimized. Note that the selected value for $\kappa$ is adjusted based only on the observation of the training and validation dataset where the WTs are healthy. The value of $\kappa$ is set to minimize the number of false alarms over these datasets. Therefore, there is no information from the test set (or from the knowledge of the occurred fault on the test set) used to decide the $\kappa$ value. Finally, 1008 spans are chosen because the alarms have

Table 3.4: Alarm activations over the test dataset for the AP-III methodology.

| WT | Indicator |
|---|---|
| WT1 | |
| WT2 | x |
| WT3 | |
| WT4 | |
| WT5 | |
| WT6 | |
| WT7 | |
| WT8 | x |
| WT9 | x |
| WT10 | |
| WT11 | |
| WT12 | |

zero activation for $(\mu + 6\sigma)$ and $(\mu + 9\sigma)$, which means that there are no false positives over the training dataset. It should be noted that this tuning is done with only training (healthy) data. Figure 3.8 shows the results over the test dataset in the entire wind farm. Table 3.6 shows the activations for warning $(\mu + 6\sigma)$ and fault alarm $(\mu + 9\sigma)$. The first activation corresponds to WT2, which is a true positive, the warning is activated on March 18, detecting the failure of the main bearing two months in advance. As with the AP-III, the WT8 presents a warning activation due to gearbox maintenance. Finally, WT11 presents a warning activation, this is a false positive. The rest of the WTs in the park were correctly classified as healthy.

Up to this point, the AP-III and AP-IV methodologies present similar results, with a slight advantage of the AP-III strategy, since it detected the failure of the main bearing a month and a half earlier than the AP-IV one. Table 3.7 shows the same results regarding accuracy, precision, specificity, recall and F1-score. However, when comparing the training time, the AP-IV method shows a much longer training time because the GRU neural network has much more parameters than the ANN (52,353 parameters versus 1153).

Finally, a conclusion of which methodology is more suitable for the early detection of main bearing faults will be given in Section 4.1.

Table 3.5: False-positives alarms (x-mark) over the training and validation datasets in the AP-IV method.

| WT | 144 spans | | | 1008 spans | | |
|---|---|---|---|---|---|---|
| | $\mu + 3\sigma$ | $\mu + 6\sigma$ | $\mu + 9\sigma$ | $\mu + 3\sigma$ | $\mu + 6\sigma$ | $\mu + 9\sigma$ |
| WT1 | x | x | | x | | |
| WT2 | x | | | | | |
| WT3 | x | | | x | | |
| WT4 | x | x | | x | | |
| WT5 | x | | | x | | |
| WT6 | x | x | | x | | |
| WT7 | x | | | x | | |
| WT8 | x | | | x | | |
| WT9 | x | x | | x | | |
| WT10 | x | | | x | | |
| WT11 | x | | | x | | |
| WT12 | x | | | x | | |

Table 3.6: Alarm activations over the test dataset for the AP-IV methodology.

| WT | 1008 spans | |
|---|---|---|
| | $\mu + 6\sigma$ | $\mu + 9\sigma$ |
| WT1 | | |
| WT2 | x | x |
| WT3 | | |
| WT4 | | |
| WT5 | | |
| WT6 | | |
| WT7 | | |
| WT8 | x | |
| WT9 | | |
| WT10 | | |
| WT11 | x | |
| WT12 | | |

Table 3.7: Accuracy, precision, specificity, recall, F1-score and training time of the AP-III and AP-IV methodologies.

| | Accuracy | Precision | Specificty | Recall | F1-score | Training time |
|---|---|---|---|---|---|---|
| AP-III method | 91.67 | 100 | 100 | 66.67 | 80 | 25 min |
| AP-IV method | 91.67 | 100 | 100 | 66.67 | 80 | 47 min |

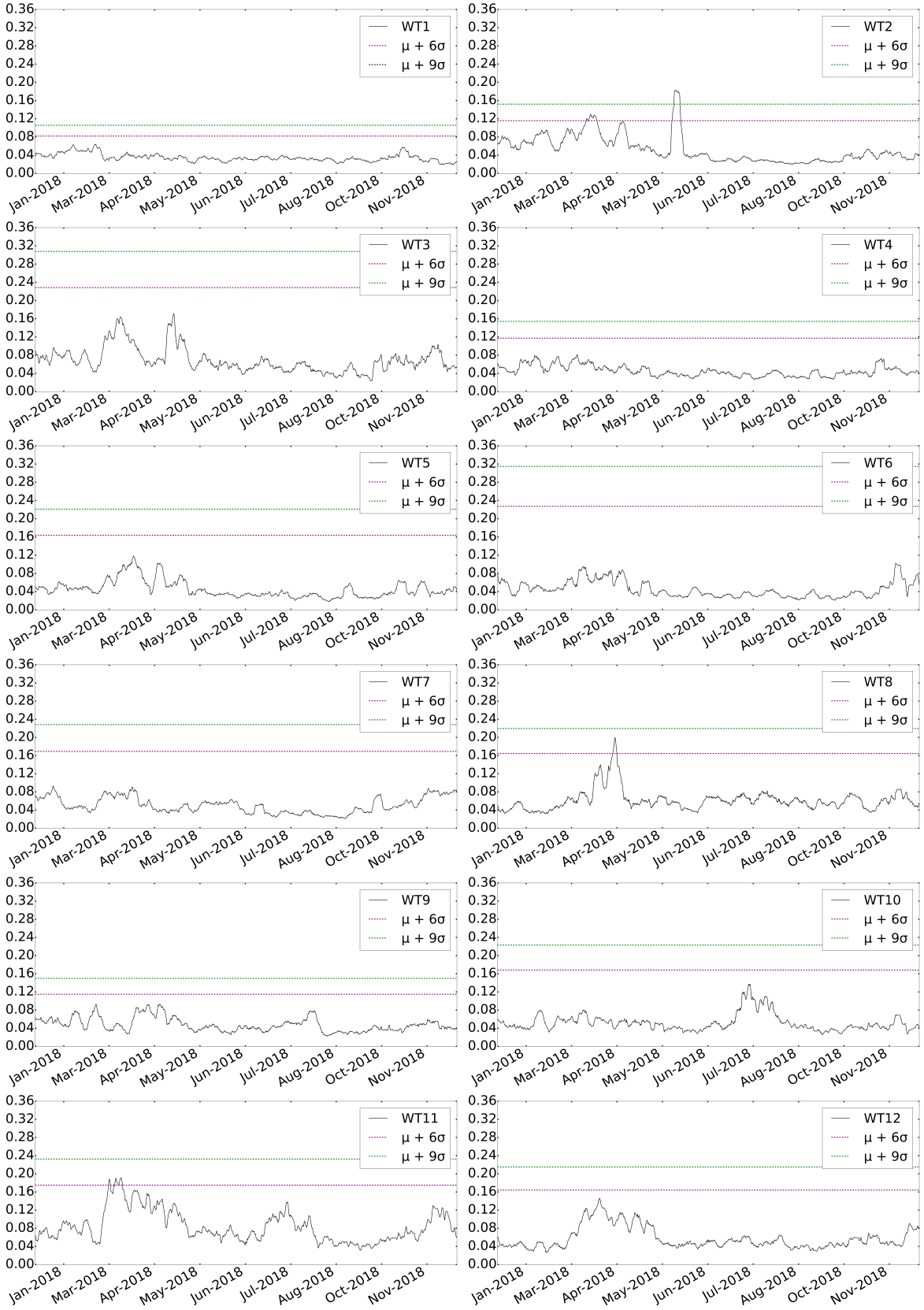



Figure 3.8: EWMA on residual errors for the WT's test dataset (using 1008 spans), where the red line represents a fault warning and the green line indicates a definite fault.

# Chapter 4

# Conclusions and future research

## 4.1 Conclusions

In conclusion, this doctoral study encompasses two main areas of research: structural health monitoring (SHM) and condition monitoring (CM) of wind turbines.

In regard to the SHM area of research, the results of this PhD provide comprehensive insights into the SHM of jacket supports of offshore WTs and demonstrate the potential of vibration response-based monitoring methods. In particular, the comparison of two strategies (AP-I and AP-II) based on the vibration response of the structural health of the support of offshore WTs is demonstrated. The overall contribution of this SHM study of offshore WT jacket supports is: *i*) how three-dimensional data (derived from different times, sensors, and experiments) are pre-processed (collected, scaled, reshaped, augmented, and converted into gray-scale images with as many channels as sensors), *ii*) the design of neural networks (CNN for the API and SNN for the APII), whose architecture and hyperparameters play a key role in the specific application at hand, the detection and diagnosis of damage, and *iii*) the comparison of the results obtained from the AP-I and AP-II. In particular, the comparison of the CNN and the SNN found that the SNN results have exceptional performance for all the metrics and, in addition, SNN only needs few data, on the contrary to the case of CNN, where the computational load is compromised and translates into a greater training time.

In a nutshell, the results of this study suggest that SNN may be an effective tool for SHM in marine WT supports and offer a recommendation for future research to explore other machine learning techniques, including CNN, when sufficient data are available. The study provides valuable insights into the preprocessing and neural network design for SHM applications and highlights the potential of SNN to perform well even with limited data.

Regarding the CM area of research, this PhD contributes to the advancement of the field of wind turbine monitoring and provides valuable insights into early detection of main bearing faults based on SCADA data. In particular, the two methodologies proposed

in AP-III and AP-IV, based on normality models from SCADA data, are compared. The main contributions of this CM study are: *i*) the use of SCADA data focused on capturing changes in temperature of the main bearing under different environmental conditions and operating conditions, *ii*) the preprocessing of SCADA data, *iii*) the design of the neural networks ANN and GRU, and *iv*) the validation and performance of the AP-III and AP-IV methodologies established in a real in production wind farm composed of 12 WT were demonstrated. The results show that the main bearing fault can be detected early, several months in advance, thus giving the plant operator time to schedule maintenance. In particular, the ANN and GRU showed the same indicators, except for the difference in detection time. The ANN detected the fault a month and a half earlier than the GRU. In addition, the ANN showed a lower computational load.

In a nutshell, the findings indicate that the use of neural networks to process SCADA data is a promising approach that can effectively detect faults in wind turbines. This study represents a novel contribution to the field of CM and demonstrates the potential of using neural networks for this purpose in real-world applications. The results of this study provide valuable insights into the design and implementation of future neural network-based models for CM of wind turbines.

## 4.2 Further research

Future work in the area of SHM will focus on improving the ability to detect and locate different types of damage to the structure. One challenge in this regard is to incorporate the effects of environmental conditions, such as wave excitation, in the analysis of scale jacket-type supports. To address this challenge, the utilization of large-scale simulation through finite element analysis is proposed as a potential solution. This will enable the design and validation of SHM methodologies and provide valuable insights into the impact of different types of damage on the overall structural health. By addressing these challenges and pursuing these research directions, it is expected to achieve significant advancements in the field of SHM.

As a continuation of this study, future research in the field of CM of wind turbines could focus on several promising avenues. First, it is suggested to explore the use of multiple models simultaneously (ensemble methods) to improve the reliability and accuracy of CM. By using two or more models that detect a fault in the same turbine, it is possible to decrease the probability of false alarms and increase the confidence in the results. Additionally, it is recommended to explore other deep learning techniques, such as long-short term memory (LSTM) networks, to further enhance the performance and capabilities of CM models. This research could also involve the integration of specific CM data and SCADA data for early fault detection through the use of normality models. Finally, another important aspect for future research is to expand the study to cover other types of faults in wind turbines beyond the main bearing one. This will require a comprehensive and systematic approach to identify and assess various failure modes in wind turbines.

# References


[1] S. R. Paramati, U. Shahzad, and B. Doğan, "The role of environmental technology for energy demand and energy efficiency: Evidence from oecd countries," *Renewable and Sustainable Energy Reviews*, vol. 153, p. 111735, 2022.

[2] D. Fraile, A. Vandenberghe, V. Konari, L. Ramirez, I. Pineda, P. Tardieu, B. Malvault, and I. Komusanac, "Getting fit for 55 and set for 2050: Electrifying europe with wind energy," *ETIPWind—WindEurope Report. June*, 2021.

[3] A. Hepbasli, "A key review on exergetic analysis and assessment of renewable energy resources for a sustainable future," *Renewable and sustainable energy reviews*, vol. 12, no. 3, pp. 593–661, 2008.

[4] P. Scarabaggio, S. Grammatico, R. Carli, and M. Dotoli, "Distributed demand side management with stochastic wind power forecasting," *IEEE Transactions on Control Systems Technology*, vol. 30, no. 1, pp. 97–112, 2021.

[5] L. Ziegler, E. Gonzalez, T. Rubert, U. Smolka, and J. J. Melero, "Lifetime extension of onshore wind turbines: A review covering germany, spain, denmark, and the uk," *Renewable and Sustainable Energy Reviews*, vol. 82, pp. 1261–1271, 2018.

[6] Z. Wang, Y. Guo, and H. Wang, "Review on monitoring and operation-maintenance technology of far-reaching sea smart wind farms," *Journal of Marine Science and Engineering*, vol. 10, no. 6, p. 820, 2022.

[7] Á. M. Costa, J. A. Orosa, D. Vergara, and P. Fernández-Arias, "New tendencies in wind energy operation and maintenance," *Applied Sciences*, vol. 11, no. 4, p. 1386, 2021.

[8] M. Civera and C. Surace, "Non-destructive techniques for the condition and structural health monitoring of wind turbines: A literature review of the last 20 years," *Sensors*, vol. 22, no. 4, p. 1627, 2022.

[9] W. Udo and Y. Muhammad, "Data-driven predictive maintenance of wind turbine based on scada data," *IEEE Access*, vol. 9, pp. 162370–162388, 2021.

[10] 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.

[11] Y. Yang, Y. Zhang, and X. Tan, "Review on vibration-based structural health monitoring techniques and technical codes," *Symmetry*, vol. 13, no. 11, p. 1998, 2021.

[12] C. P. Fritzen, "Vibration-based structural health monitoring–concepts and applications," *Key Engineering Materials*, vol. 293, pp. 3–20, 2005.

[13] C. Kralovec and M. Schagerl, "Review of structural health monitoring methods regarding a multi-sensor approach for damage assessment of metal and composite structures," *Sensors*, vol. 20, no. 3, p. 826, 2020.

[14] T.-C. Le, T.-H.-T. Luu, H.-P. Nguyen, T.-H. Nguyen, D.-D. Ho, and T.-C. Huynh, "Piezoelectric impedance-based structural health monitoring of wind turbine structures: Current status and future perspectives," *Energies*, vol. 15, no. 15, p. 5459, 2022.

[15] E. Gonzalez, B. Stephen, D. Infield, and J. Melero, "On the use of high-frequency scada data for improved wind turbine performance monitoring," in *Journal of Physics: Conference Series*, vol. 926, p. 012009, IOP Publishing, 2017.

[16] F. Castellani, D. Astolfi, P. Sdringola, S. Proietti, and L. Terzi, "Analyzing wind turbine directional behavior: Scada data mining techniques for efficiency and power assessment," *Applied Energy*, vol. 185, pp. 1076–1086, 2017.

[17] W. Yang, P. J. Tavner, C. J. Crabtree, Y. Feng, and Y. Qiu, "Wind turbine condition monitoring: technical and commercial challenges," *Wind energy*, vol. 17, no. 5, pp. 673–693, 2014.

[18] S. Dargan, M. Kumar, M. R. Ayyagari, and G. Kumar, "A survey of deep learning and its applications: a new paradigm to machine learning," *Archives of Computational Methods in Engineering*, vol. 27, no. 4, pp. 1071–1092, 2020.

[19] N. A. Patil, S. A. Arvikar, O. S. Shahane, and S. G. Rajasekharan, "Estimation of dynamic characteristics of a wind turbine blade," *Materials Today: Proceedings*, 2022.

[20] M. Sawant, S. Thakare, A. P. Rao, A. E. Feijóo-Lorenzo, and N. D. Bokde, "A review on state-of-the-art reviews in wind-turbine-and wind-farm-related topics," *Energies*, vol. 14, no. 8, p. 2041, 2021.

[21] C. W. Zheng, C. Y. Li, J. Pan, M. Y. Liu, and L. L. Xia, "An overview of global ocean wind energy resource evaluations," *Renewable and Sustainable Energy Reviews*, vol. 53, pp. 1240–1251, 2016.

[22] J. Kaldellis and M. Kapsali, "Shifting towards offshore wind energy—recent activity and future development," *Energy policy*, vol. 53, pp. 136–148, 2013.

[23] J. C. Svendsen, B. Ibanez-Erquiaga, E. Savina, and T. Wilms, "Effects of operational off-shore wind farms on fishes and fisheries. review report," 2022.

[24] X. Guo, Y. Zhang, J. Yan, Y. Zhou, S. Yan, W. Shi, and X. Li, "Integrated dynamics response analysis for iea 10-mw spar floating offshore wind turbine," *Journal of Marine Science and Engineering*, vol. 10, no. 4, p. 542, 2022.

[25] J.-S. Lai, Y.-H. Tsai, M.-J. Chang, J.-Y. Huang, and C.-M. Chi, "A technical and operational perspective on quality analysis of stitching images with multi-row panorama and multimedia sources for visualizing the tourism site of onshore wind farm," *ISPRS International Journal of Geo-Information*, vol. 11, no. 7, p. 362, 2022.

[26] M. Blanco, "The economics of wind energy, renewable and sustainable energy reviews," *Elsevier*, vol. 13, pp. 1372–1382, 2009.

[27] P. Zhang and D. Lu, "A survey of condition monitoring and fault diagnosis toward integrated o&m for wind turbines," *Energies*, vol. 12, no. 14, p. 2801, 2019.

[28] O. T. Thomsen, "Sandwich materials for wind turbine blades—present and future," *Journal of Sandwich Structures & Materials*, vol. 11, no. 1, pp. 7–26, 2009.

[29] T. K. Barlas and G. A. van Kuik, "Review of state of the art in smart rotor control research for wind turbines," *Progress in Aerospace Sciences*, vol. 46, no. 1, pp. 1–27, 2010.

[30] J. V. D. Tempel, "Design of support structures for offshore wind turbines," *Holland: Offskore Engineering and Bind Energy Sections, Technische Universiteit Delft*, pp. 51–58, 2006.

[31] L. Guo, C.-M. Uang, A. Elgamal, I. Prowell, and S. Zhang, "Pushover analysis of a 53 m high wind turbine tower," *Advanced Science Letters*, vol. 4, no. 3, pp. 656–662, 2011.

[32] W. S. Association *et al.*, "Steel solutions in the green economy-wind turbines," *Brussels, Belgium*, 2012.

[33] W. Musial and S. Butterfield, "Future for offshore wind energy in the united states," tech. rep., National Renewable Energy Lab., Golden, CO (US), 2004.

[34] A. Jadali, A. Ioannou, and A. Kolios, "A multi-attribute review toward effective planning of end-of-life strategies for offshore wind farms," *Energy Sources, Part B: Economics, Planning, and Policy*, vol. 16, no. 6, pp. 584–602, 2021.

[35] R. Yan, X. Chen, and S. C. Mukhopadhyay, "Advanced signal processing for structural health monitoring," in *Structural Health Monitoring*, pp. 1–11, Springer, 2017.

[36] C. R. Farrar and K. Worden, *Structural health monitoring: a machine learning perspective*. John Wiley & Sons, 2012.

[37] S. G. Tewolde, "Life cycle management and risk based inspection planning of offshore wind turbine support structures using structural health monitoring," 2022.

[38] M. L. Wymore, J. E. Van Dam, H. Ceylan, and D. Qiao, "A survey of health monitoring systems for wind turbines," *Renewable and Sustainable Energy Reviews*, vol. 52, pp. 976–990, 2015.

[39] N. Johnson, *Global and Local Structural Health Monitoring Methods Based on Wireless Telemetry and Boundary-based Thermography*. PhD thesis, 2017.

[40] P. Faulkner, P. Cutter, and A. Owens, "Structural health monitoring systems in difficult environments—offshore wind turbines," in *6th European workshop on structural health monitoring*, pp. 1–7, 2012.

[41] G. De Sitter, C. Devriendt, and P. Jordaens, "Dynamic monitoring of offshore wind farms," *Study day on reliability and efficiency in renewable energy sources. Bruges, Belgium*, pp. 1–33, 2013.

[42] L. B. Ibsen and R. Brincker, "Design of a new foundation for offshore wind turbines," in *Proceedings of IMAC-22: A Conference on Structural Dynamics, January 26-29, 2004, Hyatt Regency Dearborn, Dearborn, Michigan, USA*, pp. 359–366, Society for Experimental Mechanics, 2004.

[43] C. Pendão and I. Silva, "Optical fiber sensors and sensing networks: Overview of the main principles and applications," *Sensors*, vol. 22, no. 19, p. 7554, 2022.

[44] N. Zhang, C. Davis, W. K. Chiu, T. Boilard, and M. Bernier, "Fatigue performance of type i fibre bragg grating strain sensors," *Sensors*, vol. 19, no. 16, p. 3524, 2019.

[45] S. K. Al-Jumaili, M. R. Pearson, K. M. Holford, M. J. Eaton, and R. Pullin, "Acoustic emission source location in complex structures using full automatic delta t mapping technique," *Mechanical Systems and Signal Processing*, vol. 72, pp. 513–524, 2016.

[46] A. Aabid, B. Parveez, M. A. Raheman, Y. E. Ibrahim, A. Anjum, M. Hrairi, N. Parveen, and J. Mohammed Zayan, "A review of piezoelectric material-based structural control and health monitoring techniques for engineering structures: Challenges and opportunities," in *Actuators*, vol. 10, p. 101, Multidisciplinary Digital Publishing Institute, 2021.

[47] Y. Yang, Y. Hu, and Y. Lu, "Sensitivity of pzt impedance sensors for damage detection of concrete structures," *Sensors*, vol. 8, no. 1, pp. 327–346, 2008.

[48] M. Lemistre, "Low frequency electromagnetic techniques," *Structural Health Monitoring*, pp. 411–461, 2006.

[49] S. D. Fassois and J. S. Sakellariou, "Time-series methods for fault detection and identification in vibrating structures," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 365, no. 1851, pp. 411–448, 2007.

[50] P. S. A. Norway, "The use of digital solutions and structural health monitoring for integrity management of offshore structures," *Energy*, p. 1100046947, 2022.

[51] W. Fan and P. Qiao, "Vibration-based damage identification methods: a review and comparative study," *Structural health monitoring*, vol. 10, no. 1, pp. 83–111, 2011.

[52] R. Hou and Y. Xia, "Review on the new development of vibration-based damage identification for civil engineering structures: 2010–2019," *Journal of Sound and Vibration*, vol. 491, p. 115741, 2021.

[53] 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, p. 107077, 2021.

[54] X. Zhao, *New methods for structural health monitoring and damage localization*. PhD thesis, University of Sheffield, 2015.

[55] O. Avci and O. Abdeljaber, "Self-organizing maps for structural damage detection: a novel unsupervised vibration-based algorithm," *Journal of Performance of Constructed Facilities*, vol. 30, no. 3, p. 04015043, 2016.

[56] H.-P. Chen, "Structural health monitoring of large civil engineering structures," 2018.

[57] F. P. Kopsaftopoulos and S. D. Fassois, "Vibration based health monitoring for a lightweight truss structure: experimental assessment of several statistical time series methods," *Mechanical Systems and Signal Processing*, vol. 24, no. 7, pp. 1977–1997, 2010.

[58] M. Radzieński, M. Krawczuk, and M. Palacz, "Improvement of damage detection methods based on experimental modal parameters," *Mechanical Systems and Signal Processing*, vol. 25, no. 6, pp. 2169–2190, 2011.

[59] W.-J. Yan, M.-Y. Zhao, Q. Sun, and W.-X. Ren, "Transmissibility-based system identification for structural health monitoring: Fundamentals, approaches, and applications," *Mechanical Systems and Signal Processing*, vol. 117, pp. 453–482, 2019.

[60] O. Abdeljaber, O. Avci, S. Kiranyaz, M. Gabbouj, and D. J. Inman, "Real-time vibration-based structural damage detection using one-dimensional convolutional neural networks," *Journal of Sound and Vibration*, vol. 388, pp. 154–170, 2017.

[61] S. Alabbasi, M. Hussein, O. Abdeljaber, and O. Avci, "A numerical and experimental investigation of a special type of floating-slab tracks," *Engineering Structures*, vol. 215, p. 110734, 2020.

[62] R. Cantieni, "Experimental methods used in system identification of civil engineering structures," in *Proceedings of the International Operational Modal Analysis Conference (IOMAC)*, pp. 249–260, Citeseer, 2005.

[63] C. R. Farrar and K. Worden, "An introduction to structural health monitoring," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 365, no. 1851, pp. 303–315, 2007.

[64] Y. Ou, E. N. Chatzi, V. K. Dertimanis, and M. D. Spiridonakos, "Vibration-based experimental damage detection of a small-scale wind turbine blade," *Structural Health Monitoring*, vol. 16, no. 1, pp. 79–96, 2017.

[65] F. Strieth-Kalthoff, F. Sandfort, M. H. Segler, and F. Glorius, "Machine learning the ropes: principles, applications and directions in synthetic chemistry," *Chemical Society Reviews*, vol. 49, no. 17, pp. 6154–6168, 2020.

[66] R. Akerkar and P. Sajja, *Knowledge-based systems*. Jones & Bartlett Publishers, 2009.

[67] S. Khalid, T. Khalil, and S. Nasreen, "A survey of feature selection and feature extraction techniques in machine learning," in *2014 science and information conference*, pp. 372–378, IEEE, 2014.

[68] X. Kong, C.-S. Cai, and J. Hu, "The state-of-the-art on framework of vibration-based structural damage identification for decision making," *Applied Sciences*, vol. 7, no. 5, p. 497, 2017.

[69] S. Hakim and H. A. Razak, "Modal parameters based structural damage detection using artificial neural networks-a review," *Smart Structures and Systems*, vol. 14, no. 2, pp. 159–189, 2014.

[70] S. Sony, "Bridge damage identification using deep learning-based convolutional neural networks (cnns)," 2021.

[71] A. Gomez-Cabrera and P. J. Escamilla-Ambrosio, "Review of machine-learning techniques applied to structural health monitoring systems for building and bridge structures," *Applied Sciences*, vol. 12, no. 21, p. 10754, 2022.

[72] R.-C. Chen, C. Dewi, S.-W. Huang, and R. E. Caraka, "Selecting critical features for data classification based on machine learning methods," *Journal of Big Data*, vol. 7, no. 1, p. 52, 2020.

[73] C. C. Aggarwal and C. C. Aggarwal, "Machine learning with shallow neural networks," *Neural Networks and Deep Learning: A Textbook*, pp. 53–104, 2018.

[74] Y. Xu, Y. Zhou, P. Sekula, and L. Ding, "Machine learning in construction: From shallow to deep learning," *Developments in the built environment*, vol. 6, p. 100045, 2021.

[75] A. S. Mohamed, S. Sassi, and M. Roshun Paurobally, "Model-based analysis of spur gears' dynamic behavior in the presence of multiple cracks," *Shock and Vibration*, vol. 2018, 2018.

[76] J. Patterson and A. Gibson, *Deep learning: A practitioner's approach.* " O'Reilly Media, Inc.", 2017.

[77] L. Torrey and J. Shavlik, "Transfer learning," in *Handbook of research on machine learning applications and trends: algorithms, methods, and techniques*, pp. 242–264, IGI global, 2010.

[78] S. Yang, W. Xiao, M. Zhang, S. Guo, J. Zhao, and F. Shen, "Image data augmentation for deep learning: A survey," *arXiv preprint arXiv:2204.08610*, 2022.

[79] J. Snoek, H. Larochelle, and R. P. Adams, "Practical bayesian optimization of machine learning algorithms," *Advances in neural information processing systems*, vol. 25, 2012.

[80] D. Chicco, "Siamese neural networks: An overview," *Artificial neural networks*, pp. 73–94, 2021.

[81] O. Arbelaitz, I. Gurrutxaga, J. Muguerza, and J. M. Pérez, "Applying resampling methods for imbalanced datasets to not so imbalanced datasets," in *Advances in Artificial Intelligence: 15th Conference of the Spanish Association for Artificial Intelligence, CAEPIA 2013, Madrid, Spain, September 17-20, 2013. Proceedings 15*, pp. 111–120, Springer, 2013.

[82] A. Fernández, S. García, M. Galar, R. C. Prati, B. Krawczyk, F. Herrera, A. Fernández, S. García, M. Galar, R. C. Prati, *et al.*, "Cost-sensitive learning," *Learning from Imbalanced Data Sets*, pp. 63–78, 2018.

[83] O. Sagi and L. Rokach, "Ensemble learning: A survey," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 8, no. 4, p. e1249, 2018.

[84] G. Pang, C. Shen, L. Cao, and A. V. D. Hengel, "Deep learning for anomaly detection: A review," *ACM computing surveys (CSUR)*, vol. 54, no. 2, pp. 1–38, 2021.

[85] D. Mandic and J. Chambers, *Recurrent neural networks for prediction: learning algorithms, architectures and stability.* Wiley, 2001.

[86] R. C. Staudemeyer and E. R. Morris, "Understanding lstm–a tutorial into long short-term memory recurrent neural networks," *arXiv preprint arXiv:1909.09586*, 2019.

[87] J. Koushik, "Understanding convolutional neural networks," *arXiv preprint arXiv:1605.09081*, 2016.

[88] Z. Niu, G. Zhong, and H. Yu, "A review on the attention mechanism of deep learning," *Neurocomputing*, vol. 452, pp. 48–62, 2021.

[89] M. A. Ganaie, M. Hu, A. Malik, M. Tanveer, and P. Suganthan, "Ensemble deep learning: A review," *Engineering Applications of Artificial Intelligence*, vol. 115, p. 105151, 2022.

[90] D. Goyal, A. Saini, S. Dhami, B. Pabla, *et al.*, "Intelligent predictive maintenance of dynamic systems using condition monitoring and signal processing techniques—a review," in *2016 international conference on advances in computing, communication, & automation (ICACCA)(Spring)*, pp. 1–6, IEEE, 2016.

[91] N. Sakib and T. Wuest, "Challenges and opportunities of condition-based predictive maintenance: a review," *Procedia Cirp*, vol. 78, pp. 267–272, 2018.

[92] "major-and-minor-faults-in-wind-turbines." `http://www.renovetec.com/irim/14-revista-irim-6/272-averias-mayores-y-menores-en-aerogeneradores`, 2022. Accessed: 2022-07-30.

[93] K. Fischer and D. Coronado, "Condition monitoring of wind turbines: state of the art, user experience and recommendations," *Fraunhofer-IWES, Bremerhaven*, pp. 1–89, 2015.

[94] H. Badihi, Y. Zhang, B. Jiang, P. Pillay, and S. Rakheja, "A comprehensive review on signal-based and model-based condition monitoring of wind turbines: Fault diagnosis and lifetime prognosis," *Proceedings of the IEEE*, 2022.

[95] S. Sharma and D. Mahto, "Condition monitoring of wind turbines: a review," *Global Journal of Researches in Engineering, Mechanical and Mechanics Engineering*, vol. 13, no. 6, 2013.

[96] F. P. G. Márquez, A. M. Tobias, J. M. P. Pérez, and M. Papaelias, "Condition monitoring of wind turbines: Techniques and methods," *Renewable energy*, vol. 46, pp. 169–178, 2012.

[97] Z. Ma, Y. Liu, D. Wang, W. Teng, and A. Kusiak, "Cyclostationary analysis of a faulty bearing in the wind turbine," *Journal of Solar Energy Engineering*, vol. 139, no. 3, 2017.

[98] Z. Hameed, Y. Hong, Y. Cho, S. Ahn, and C. Song, "Condition monitoring and fault detection of wind turbines and related algorithms: A review," *Renewable and Sustainable energy reviews*, vol. 13, no. 1, pp. 1–39, 2009.

[99] D. Coronado and J. Wenske, "Monitoring the oil of wind-turbine gearboxes: Main degradation indicators and detection methods," *Machines*, vol. 6, no. 2, p. 25, 2018.

[100] A. Hamilton and F. Quail, "Detailed state of the art review for the different online/inline oil analysis techniques in context of wind turbine gearboxes," 2011.

[101] S. Letzgus, "Scada-data analysis for condition monitoring of wind turbines," Master's thesis, 2015.

[102] P. Tchakoua, R. Wamkeue, M. Ouhrouche, F. Slaoui-Hasnaoui, T. A. Tameghe, and G. Ekemb, "Wind turbine condition monitoring: State-of-the-art review, new trends, and future challenges," *Energies*, vol. 7, no. 4, pp. 2595–2630, 2014.

[103] International-Electrotechnical-Commission, "Wind turbines-part 12-1: Power performance measurements of electricity producing wind turbines," *IEC 61400-12-1*, 2005.

[104] J. Tautz-Weinert, *Improved wind turbine monitoring using operational data*. PhD thesis, Loughborough University, 2018.

# Part II

# Appended Paper I

# Vibration-Response-Only Structural Health Monitoring for Offshore Wind Turbine Jacket Foundations via Convolutional Neural Networks

**Authors:**
Bryan Puruncajas, Yolanda Vidal and Christian Tutivén

**Paper published in:**
*Sensors* 2021, 21(6), 2228
DOI: 10.3390/s21062228
Ranking JCR: Q1 (2021)

**Paper history:**
Submitted: 04-03-2021
Accepted: 20-03-2021
Published: 23-03-2021

sensors

MDPI

Article

# Vibration-Response-Only Structural Health Monitoring for Offshore Wind Turbine Jacket Foundations via Convolutional Neural Networks

**Bryan Puruncajas [1,2], Yolanda Vidal [1,*] and Christian Tutivén [2]**

[1] Control, Modeling, Identification and Applications (CoDAlab), Department of Mathematics, Escola d'Enginyeria de Barcelona Est (EEBE), Universitat Politècnica de Catalunya (UPC), Campus Diagonal-Besós (CDB), Eduard Maristany, 16, 08019 Barcelona, Spain; bpurunca@espol.edu.ec

[2] Mechatronics Engineering, Faculty of Mechanical Engineering and Production Science (FIMCP), Escuela Superior Politécnica del Litoral (ESPOL), Guayaquil 09-01-5863, Ecuador; cjtutive@espol.edu.ec

* Correspondence: yolanda.vidal@upc.edu; Tel.: +34-934-137-309

Received: 25 May 2020; Accepted: 16 June 2020; Published: 17 June 2020


**Abstract:** This work deals with structural health monitoring for jacket-type foundations of offshore wind turbines. In particular, a vibration-response-only methodology is proposed based on accelerometer data and deep convolutional neural networks. The main contribution of this article is twofold: (i) a signal-to-image conversion of the accelerometer data into gray scale multichannel images with as many channels as the number of sensors in the condition monitoring system, and (ii) a data augmentation strategy to diminish the test set error of the deep convolutional neural network used to classify the images. The performance of the proposed method is analyzed using real measurements from a steel jacket-type offshore wind turbine laboratory experiment undergoing different damage scenarios. The results, with a classification accuracy over 99%, demonstrate that the stated methodology is promising to be utilized for damage detection and identification in jacket-type support structures.

**Keywords:** structural health monitoring; damage detection; damage identification; offshore wind turbine foundation; jacket; signal-to-image conversion; convolutional neural network

## 1. Introduction

Globally, wind power generation capacity has increased exponentially since the early 1990s, and as of the end of 2019, this capacity amounted to 650 GW [1]. Whereas onshore wind turbines (WTs) have dominated new wind installations in the past, the growth of offshore WTs is poised to become the new leader because of steadier wind, in addition to vast regions where its installation is possible. In regard to the global offshore market, the cumulative installations have now reached 23 GW, representing 4% of total cumulative installations. Unfortunately, offshore WTs are placed in a harsh environment that originates from the wind and sea conditions [2]. As a consequence, offshore WTs require rigorous safety measures because it is extremely complicated to do operation and corrective work on these huge WTs placed in remote locations. Given that approaches centered on enhancing component reliability are likely to increase capital expenditures, system design optimization research and development activities should instead focus on minimizing and, if possible, even eliminating unexpected failures. In other words, the wind industry must abandon corrective maintenance (remedy failures) and move toward predictive maintenance (repair immediately before failure) to achieve maximum availability. Thus, the development of a structural health monitoring (SHM) strategy is particularly necessary to achieve this goal.

Onshore and offshore fixed WTs differ mainly in the structure of their respective foundations. Several types of offshore foundations are used, with foundation choice depending on the water depth. The most common foundations are shown in Figure 1, see [3]. Note that jacket foundations, which are the object of study of this work, are preferred for greater depths (usually, between 30 to 90 m).

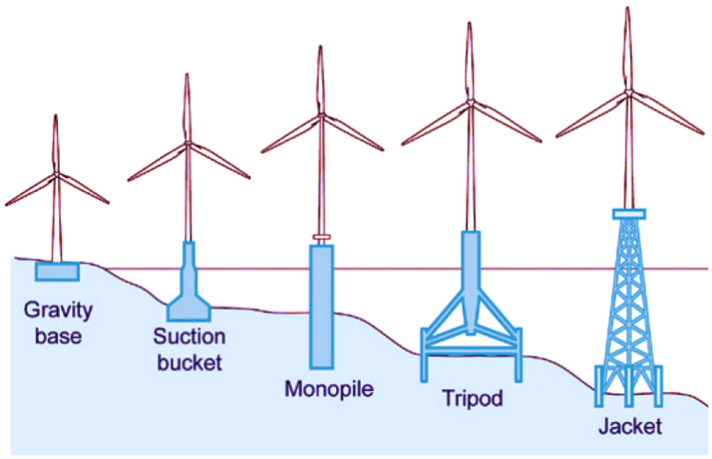



**Figure 1.** Fixed bottom wind turbine foundations [3].

The detection of early-stage damage in the foundation is of great importance to avoid the possible future collapse of the entire structure. As stated in "Long-term research challenges in wind energy—a research agenda by the European Academy of Wind Energy" [4]:

> A defining marine environment main characteristic is that structures are always subject to excitations. Techniques for structural health monitoring, vibration and data analysis must be capable of coping with such ambient excitations. As the input is typically not known, a normal input-output formalism cannot be used.

Thus, to overcome this challenge—which is posed by the fact that the input is typically not known—in this work, a structural health monitoring strategy for jacket-type foundations is developed based on a vibration response-only methodology. This is a challenge by itself as many of the works in the SHM field are based on the principle of guided elastic waves with a given (known) input excitation. See, for example, the overview of SHM systems for various WT components presented by Liu et al. [5]. In contrast, in this work, a new paradigm is introduced in which a predetermined excitation in the structure is no longer forced, but rather, the incident wind and waves serve as the exciting forces in the structure. In this way, the classic pattern recognition paradigm with identical excitation (e.g., [6]) becomes a problem of pattern recognition with variable excitation. Consequently, the new paradigm implies greater complexity in the damage detection process. Furthermore, until recently, few contributions have targeted offshore WT foundations. Notably, work by Weijtjens et al. [7] was focused on a real WT foundation and contributed an SHM strategy based on the resonance frequencies of the foundation. However, the results only proved some increased stiffness of the structure and could not give a final diagnosis about damage detection. Similarly, Oliveira et al. [8] introduced the main aspects in the development of a vibration-based monitoring system for an onshore 2.0-MW wind turbine based on identification of the modal properties of the most important vibration modes, in which detailed attention was given to the statistical procedure based on regression models that was used to minimize the influence of operational and environmental effects over the features considered to detect structural changes in the WT. However, only damage detection was pursued with a single type of damage. Noteworthily, the work by Zugasti [9] used damage estimators to detect damage in an experimental offshore tower similar to that employed in this work. Nevertheless, only damage detection was attained. In this work, in contrast to the aforementioned references, several types of damage are studied, and not only damage detection but also its classification is achieved.

It is important to note that the SHM standard approach for the problem at hand is usually an unsupervised one. That is, as no one would purposely damage their assets to train a SHM tool, only healthy data from the real structure is used. However, it is unfeasible to correctly identify different damage states using solely data obtained during what is assumed to be a healthy state. In this framework, detection can be accomplished by using a model of normality or unsupervised models, but not classification on the type of damage. The approach proposed in this work is the opposite, that is, a supervised approach. Thus, data from the damaged structure are required to train the model. In practice, this will be accomplished by means of computer models, as the finite element method (FEM). The FEM model should be validated with a downscaled experimental tower (as the one proposed in this work). Then, the full-scale finite element model would be used to generate healthy (to validate with the real asset) and damage samples. Finally, the stated supervised methodology proposed in this work can be used. In this work, a satisfactory experimental proof of concept has been conducted with the proposed strategy and a laboratory downscaled WT. However, future work is needed to validate the technology in a full-scale and more realistic environment. Some examples of this type of approach are given in [10], where bridge damage detection is accomplished by a neural network considering errors in baseline finite element models, and [11] where the stated SHM method for an oil offshore structure is capable to cope with several types of damage based on a finite element model.

On the one hand, it has been shown that traditional machine learning requires complex feature extraction processes and specialized knowledge, especially for a complex problem such as WT condition monitoring [12–14]. Moreover, extracting features with classic machine learning methods faces the classic bias-variance dilemma from inference theory. The bias–variance trade-off implies that a model should balance under-fitting and over-fitting; that is, the model should be rich enough to express underlying structure in the data but simple enough to avoid fitting spurious patterns, respectively. On the other hand, in the modern practice of deep learning, very rich models are trained to precisely fit (i.e., interpolate) the data. Classically, such models would be considered over-fit, and yet they often obtain high accuracy on test data. Thus, this paper proposes to use deep convolutional neural networks (CNN) for pattern recognition (classification), avoiding the aforementioned usual problems in the literature—e.g., [12–14]—related to feature extraction and bias–variance trade-off. In particular, we develop a novel damage diagnosis method for WT offshore foundations based on transforming condition monitoring multi-vibration-signals into images (with as many channels as sensors) to be processed afterward using deep CNN.

The paper is organized in the following manner. First, in Section 2, the experimental setup is introduced. It consists of a steel jacket-type offshore WT laboratory structure undergoing different damage scenarios. Then, in Section 3, the proposed SHM strategy is described in detail. The approach can be summarized by the following steps: (i) accelerometer data is gathered, (ii) a preprocess is designed to extract the maximum amount of information and to obtain a dataset of 24 (that is, the same number as accelerometer sensors) channel gray-scale images, (iii) 24-channel-input deep CNN is designed and trained for classification of the different structural states. In Section 4, the obtained results are conferred, showing an exceptional performance with all considered metrics giving results greater than 99%. Lastly, the main conclusions are given in Section 5 as well as future work research directions.

## 2. Experimental Setup

The laboratory experimental setup is described in the following. First, a function generator (GW INSTEK AF-2005 model) is employed to generate a white noise signal. Then, this signal is amplified and applied to a modal shaker (GW-IV47 from Data Physics) that induces the vibration into the structure. The general overview of the experimental setup is shown in Figure 2 (left). The structure is 2.7 m tall and composed of three parts:

1. The top beam ($1 \times 0.6$ m), where the modal shaker is attached to simulate a nacelle mass and the effects of wind excitation;
2. The tower with three tubular sections connected with bolts;

3. The jacket, which includes a pyramidal structure made up by 32 bars (S275JR steel) of different lengths, sheets (DC01 LFR steel), and other elements such as bolts and nuts.

It should be noted that different wind speeds are considered by modifying the white noise signal amplitude (i.e., scaling the amplitude by 0.5, 1, 2, and 3).

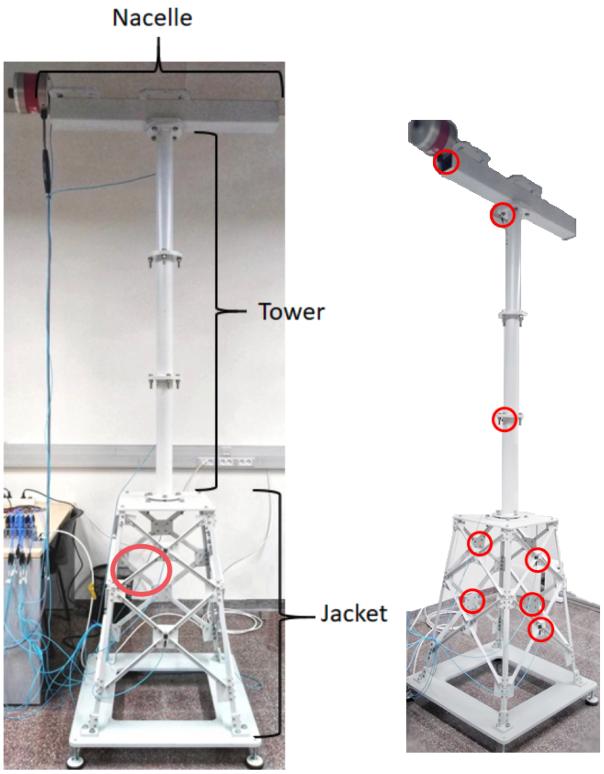



**Figure 2.** The experimental setup (**left**) detailing the location of the damaged bar (red circle). Location of the sensors on the overall structure (**right**).

To measure vibration, eight triaxial accelerometers (PCB® Piezotronic, model 356A17) are placed on the structure, see Figure 2 (right). The optimal number and placement of the sensors is determined according to [9]. The accelerometers are connected to six National Instruments™ cartridges (NI 9234 model) that are inserted into the National Instruments chassis cDAQ-9188. Finally, the Data Acquisition Toolbox™ is employed to configure the data acquisition hardware and read the data into MATLAB®.

The studied damage states are related to one of the jacket bars, see Figure 3. The damage states include a 5-mm bar crack and loosening one of the jacket bolts. Furthermore, a pristine replica bar is also considered.

**Figure 3.** Different structural state scenarios studied in this work. Replica (healthy) bar (**left**). Crack damage, where $L$ is the length of the bar, $d = 5$ mm is the crack size, and $X = L/3$ is the location of the crack in the bar (**center**). Missing bolt (**right**).

Finally, note that the purpose of the paper is to verify that the conceived methodology has practical potential. The laboratory tower is a simplified model, but it is valid for this preliminary study because it is similar to the laboratory towers used, for example, in [9], where damage detection is accomplished (but not localization or identification) via damage indicators; in [15,16], where statistical time series are employed to detect damage; and in [17,18], where damage detection is accomplished through principal component analysis and support vector machines.

## 3. Structural Health Monitoring Proposed Methodology

The proposed SHM strategy follows the steps detailed here. First, the raw time series data are collected. Second, the data are preprocessed to obtain a dataset of 24-channel gray-scale images. Third, a 24-channel-input CNN is designed and trained for classification of the different structural states. The following subsections describe the aforementioned procedure in detail.

### *3.1. Data Gathering*

The data were gathered in different experiments with a sampling rate of 275.27 Hz and a duration of 60 s each. Table 1 shows the total number of realized experiments for the corresponding structural state (with its corresponding label) and white noise amplitude. A total of $K = 100$ experiments were conducted. Given the $k$-th experiment, where $k$ is varied from 1 to $K = 100$, the raw data were then saved in the matrix $\mathbf{X}^{(k)} \in \mathcal{M}_{16517\times 24}(\mathbb{R})$

$$\mathbf{X}^{(k)} = \begin{pmatrix} x_{1,1}^{(k)} & x_{1,2}^{(k)} & \cdots & x_{1,24}^{(k)} \\ x_{2,1}^{(k)} & x_{2,2}^{(k)} & \cdots & x_{2,24}^{(k)} \\ \vdots & \vdots & \ddots & \vdots \\ x_{16517,1}^{(k)} & x_{16517,2}^{(k)} & \cdots & x_{16517,24}^{(k)} \end{pmatrix}. \tag{1}$$

Note that there are as many rows as the number of measurements in each experiment—that is, $I = 16,517$—and as many columns as the number of sensors, $J = 24$ (because each column is related to one sensor). Ultimately, the overall data matrix $\mathbf{X} \in \mathcal{M}_{1651700\times 24}(\mathbb{R})$ is constructed by stacking the matrices that arise from each different experiment,

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}^{(1)} \\ \vdots \\ \mathbf{X}^{(k)} \\ \vdots \\ \mathbf{X}^{(100)} \end{pmatrix}. \tag{2}$$

**Table 1.** Total number of experimental tests for the different white noise (WN) amplitudes and for each structural state.

| Label | Structural State | 0.5 WN | 1 WN | 2 WN | 3 WN |
|---|---|---|---|---|---|
| 1 | Healthy bar | 10 tests | 10 tests | 10 tests | 10 tests |
| 2 | Replica bar | 5 tests | 5 tests | 5 tests | 5 tests |
| 3 | Crack damaged bar | 5 tests | 5 tests | 5 tests | 5 tests |
| 4 | Unlocked bolt | 5 tests | 5 tests | 5 tests | 5 tests |

### *3.2. Data Preprocessing: Scaling, Reshaping, Augmentation, and Signal-To-Image Conversion*

Data preprocessing is both the initial step and a critical step in machine learning. In this work, data reshaping is employed to guarantee that each sample includes multiple measurements from each sensor and thus has sufficient information to make a diagnosis regarding the state of the structure. Furthermore, a data-augmentation strategy is proposed to improve the final test set error of the prediction model. It is clear that the signal-to-image conversion as well as the architecture and hyperparameters of the deep CNN will play a key role in the damage detection methodology. However, the manner in which these data are scaled, augmented, and reshaped will significantly impact the overall performance of the strategy [19].

### 3.2.1. Data Scaling

The importance of preprocessing techniques for image classification by CNN is well known [20]. The main reason for data scaling is to enhance the efficiency of the neural network training process, significantly decreasing the number of epochs required for the network to learn, and thus leading to a better predictor. In particular, here, the data are scaled column-wise to fall within the specific range $[0, 255]$. This range is selected to later allow for easy conversion into gray-scale images. In particular, the range is computed as follows. Assuming that there are $K$ experimental tests, $I$ samples per experiment, and $J$ sensors,

$$M_j = \max\left(x_{ij}^{(k)}\right),\ i = 1, \ldots, I,\ k = 1, \ldots, K, \tag{3}$$

$$m_j = \min\left(x_{ij}^{(k)}\right),\ i = 1, \ldots, I,\ k = 1, \ldots, K, \tag{4}$$

where $M_j$ and $m_j$ are the maximum and the minimum values, respectively, of all the measures at column $j$, where $j = 1, \ldots, J$. Accordingly, the elements of matrix $\mathbf{X}$ are scaled

$$y_{ij}^{(k)} := \left(x_{ij}^{(k)} - m_j\right) \frac{255}{M_j - m_j},\ i = 1, \ldots, I,\ \ j = 1, \ldots, J,\ \ k = 1, \ldots, K, \tag{5}$$

to create a new matrix $\mathbf{Y}$ as

$$\mathbf{Y} = \left(\begin{array}{ccc|c} y_{1,1}^{(1)} & y_{1,2}^{(1)} & \cdots & y_{1,24}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ y_{16517,1}^{(1)} & y_{16517,2}^{(1)} & \cdots & y_{16517,24}^{(1)} \\ \hline y_{1,1}^{(2)} & y_{1,2}^{(2)} & \cdots & y_{1,24}^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ y_{16517,1}^{(2)} & y_{16517,2}^{(2)} & \cdots & y_{16517,24}^{(2)} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline y_{1,1}^{(100)} & y_{1,2}^{(100)} & \cdots & y_{1,24}^{(100)} \\ \vdots & \vdots & \ddots & \vdots \\ y_{16517,1}^{(100)} & y_{16517,2}^{(100)} & \cdots & y_{16517,24}^{(100)} \end{array}\right) = \left(\begin{array}{c} \mathbf{Y}^{(1)} \\ \hline \mathbf{Y}^{(2)} \\ \hline \vdots \\ \hline \mathbf{Y}^{(100)} \end{array}\right). \tag{6}$$

### 3.2.2. Data Reshaping

In this section, data reshaping is employed to guarantee that each sample has multiple measurements from each sensor and thus has sufficient information to diagnose the state of the structure. In particular, matrix (6) is reshaped to matrix $\mathbf{Z} \in \mathcal{M}_{(6400)\times(256\cdot 24)}$, as given in Table 2. It should be noted that the data in the first 256 columns are related to sensor 1 and define the first submatrix block, denoted as $\mathbf{Z}_1$. Then, the data in columns 257 to 512 are related to sensor 2 and define the second submatrix block $\mathbf{Z}_2$. Next, the columns 513 to 768 are related to sensor 3 and define the third submatrix block $\mathbf{Z}_3$, and so on and so forth, until the last sensor related to $\mathbf{Z}_{24}$ has been accounted for.

**Table 2.** Data reshaping. On the one hand, this process can be viewed as the vertical stacking of $K = 100$ matrices $\mathbf{Z}^{(k)}$, $k = 1, \ldots, K$, where each matrix is associated with a different experiment. On the other hand, this process can also be viewed as the horizontal concatenation of $J = 24$ matrices, $\mathbf{Z}_j$, $j = 1, \ldots, J$, where each matrix is associated with a different sensor.

| | **Sensor 1** | | | **...** | | | **Sensor 24** | | |
|---|---|---|---|---|---|---|---|---|---|
| $\mathbf{Z} =$ | $y^{(1)}_{1,1}$ | $\cdots$ | $y^{(1)}_{256,1}$ | $\cdots$ | $y^{(1)}_{1,24}$ | $\cdots$ | $y^{(1)}_{256,24}$ | $=$ | $\mathbf{Z}^{(1)}$ |
| | $y^{(1)}_{257,1}$ | $\cdots$ | $y^{(1)}_{512,1}$ | | $y^{(1)}_{257,24}$ | $\cdots$ | $y^{(1)}_{512,24}$ | | |
| | $\vdots$ | $\ddots$ | $\vdots$ | | $\vdots$ | $\ddots$ | $\vdots$ | | |
| | $y^{(1)}_{16129,1}$ | $\cdots$ | $y^{(1)}_{16384,1}$ | | $y^{(1)}_{16129,24}$ | $\cdots$ | $y^{(1)}_{16384,24}$ | | |
| | $\vdots$ | $\ddots$ | $\vdots$ | | $\vdots$ | $\ddots$ | $\vdots$ | | $\vdots$ |
| | $y^{(k)}_{1,1}$ | $\cdots$ | $y^{(k)}_{256,1}$ | | $y^{(k)}_{1,24}$ | $\cdots$ | $y^{(k)}_{256,24}$ | | $\mathbf{Z}^{(k)}$ |
| | $y^{(k)}_{257,1}$ | $\cdots$ | $y^{(k)}_{512,1}$ | | $y^{(k)}_{257,24}$ | $\cdots$ | $y^{(k)}_{512,24}$ | | |
| | $\vdots$ | $\ddots$ | $\vdots$ | | $\vdots$ | $\ddots$ | $\vdots$ | | |
| | $y^{(k)}_{16129,1}$ | $\cdots$ | $y^{(k)}_{16384,1}$ | | $y^{(k)}_{16129,24}$ | $\cdots$ | $y^{(k)}_{16384,24}$ | | |
| | $\vdots$ | $\ddots$ | $\vdots$ | | $\vdots$ | $\ddots$ | $\vdots$ | | $\vdots$ |
| | $y^{(100)}_{1,1}$ | $\cdots$ | $y^{(100)}_{256,1}$ | | $y^{(100)}_{1,24}$ | $\cdots$ | $y^{(100)}_{256,24}$ | | $\mathbf{Z}^{(100)}$ |
| | $y^{(100)}_{257,1}$ | $\cdots$ | $y^{(100)}_{512,1}$ | | $y^{(100)}_{257,24}$ | $\cdots$ | $y^{(100)}_{512,24}$ | | |
| | $\vdots$ | $\ddots$ | $\vdots$ | | $\vdots$ | $\ddots$ | $\vdots$ | | |
| | $y^{(100)}_{16129,1}$ | $\cdots$ | $y^{(100)}_{16384,1}$ | | $y^{(100)}_{16129,24}$ | $\cdots$ | $y^{(100)}_{16384,24}$ | | |

$$= \left( \begin{array}{c|c|c} \mathbf{Z}_1 & \cdots & \mathbf{Z}_{24} \end{array} \right)$$

It should be noted that each row of matrix $\mathbf{Z}$ contains the information of one sample of our SHM strategy. Notice that to diagnose a WT, the trained model requires at least one sample. Based on the aforementioned reshaping process, the expected sample now contains 256 time stamps from each sensor. In this manner, less than 1 second is required to gather the necessary data when the sampling frequency is 275.27 Hz. Thus, this process leads to a faster detection time (amount of time that elapses between fault occurrence and detection). The intuition behind the proposed data reshape is twofold: (*i*) it supplies more information to each sample, and (*ii*) it simplifies the signal-to-image conversion, as stated in Section 3.2.4, because 256 is a perfect square.

Finally, observe that from matrices $\mathbf{Y}^{(k)}$, $k = 1, \ldots, K$ in Equation (6), the last samples $y^{(k)}_{i,j}$ from $i = 16385, \cdots, 16517$, are discarded to reshape the data in the aforementioned new matrices $\mathbf{Z}^{(k)}$, $k = 1, \ldots, K$.

### 3.2.3. Data Augmentation

Deep convolutional neural networks rely heavily on big data to avoid overfitting, see [21]. Unfortunately, many application domains lack access to big data. In this work, to build a better deep CNN model, a data augmentation strategy is proposed that artificially expands the size of the training dataset without actually collecting new data.

The method consists of using each time stamp as the beginning of a new sample (and using the subsequent 255 measures to complete the sample), as shown in Table 3. Accordingly, instead of the previously defined matrices (see Table 2) $\mathbf{Z}^{(k)} \in \mathcal{M}_{(64)\times(256\cdot24)}$, $k = 1,\ldots,K$, augmented matrices with the same number of columns but more rows are obtained, namely, $\mathbf{D}^{(k)} \in \mathcal{M}_{(16129)\times(256\cdot24)}$, $k = 1,\ldots,K$. Thus, from the initial 64 samples per experiment, we increased to 16,129 samples per experiment. This is an increment of 25,200% in the total number of samples in the dataset.

**Table 3.** Synthetic data augmentation for experiment $k$, $k = 1,\ldots,K$.

| | **Signal 1** | | | **Signal 2** | | | **...** | **Signal 24** | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $\mathbf{D}^{(k)} =$ | $y^{(k)}_{1,1}$ | $\cdots$ | $y^{(k)}_{256,1}$ | $y^{(k)}_{1,2}$ | $\cdots$ | $y^{(k)}_{256,2}$ | | $y^{(k)}_{1,24}$ | $\cdots$ | $y^{(k)}_{256,24}$ |
| | $y^{(k)}_{2,1}$ | $\cdots$ | $y^{(k)}_{257,1}$ | $y^{(k)}_{2,2}$ | $\cdots$ | $y^{(k)}_{257,2}$ | | $y^{(k)}_{2,24}$ | $\cdots$ | $y^{(k)}_{257,24}$ |
| | $y^{(k)}_{3,1}$ | $\cdots$ | $y^{(k)}_{258,1}$ | $y^{(k)}_{3,2}$ | $\cdots$ | $y^{(k)}_{258,2}$ | $\cdots$ | $y^{(k)}_{3,24}$ | $\cdots$ | $y^{(k)}_{258,24}$ |
| | $y^{(k)}_{4,1}$ | $\cdots$ | $y^{(k)}_{259,1}$ | $y^{(k)}_{4,2}$ | $\cdots$ | $y^{(k)}_{259,2}$ | | $y^{(k)}_{4,24}$ | $\cdots$ | $y^{(k)}_{259,24}$ |
| | $\vdots$ | $\ddots$ | $\vdots$ | $\vdots$ | $\ddots$ | $\vdots$ | | $\vdots$ | $\ddots$ | $\vdots$ |
| | $y^{(k)}_{16129,1}$ | $\cdots$ | $y^{(k)}_{16384,1}$ | $y^{(k)}_{16129,2}$ | $\cdots$ | $y^{(k)}_{16384,2}$ | | $y^{(k)}_{16129,24}$ | $\cdots$ | $y^{(k)}_{16384,24}$ |

Finally, the data matrix $\mathbf{D} \in \mathcal{M}_{1612900\times(256\cdot24)}(\mathbb{R})$—which contains the scaled, reshaped, and augmented data from all of the experiments—is defined by stacking the data matrices derived from each different experiment (recall that $K = 100$),

$$\mathbf{D} = \begin{pmatrix} \mathbf{D}^{(1)} \\ \vdots \\ \mathbf{D}^{(k)} \\ \vdots \\ \mathbf{D}^{(100)} \end{pmatrix}. \tag{7}$$

#### 3.2.4. Signal-To-Image Conversion

The fault diagnosis method converts time-domain signals from the 24 measured variables into 2D gray-scale images to exploit texture information from the converted images. The data conversion process was inspired based on reference [13], although the process is enhanced here by using multichannel images.

The image size used for signal-to-image conversion is $16 \times 16$ (256 pixels) with 24 channels, constructed as follows. Each row of matrix $\mathbf{D}$, see Equation (7), is converted to one image of size $16 \times 16$ with 24 channels (one channel per sensor), similar to a standard RGB image with 3 channels. It should be noted that because the sampling time is $1/257$ seconds, each image contains approximately one second of data from each sensor, which is sufficient to capture all of the system dynamics. The total number of images in the dataset is 1,612,900, because 16,129 images are obtained from each of the 100 experiments. Figure 4 shows one example of such a multichannel image.

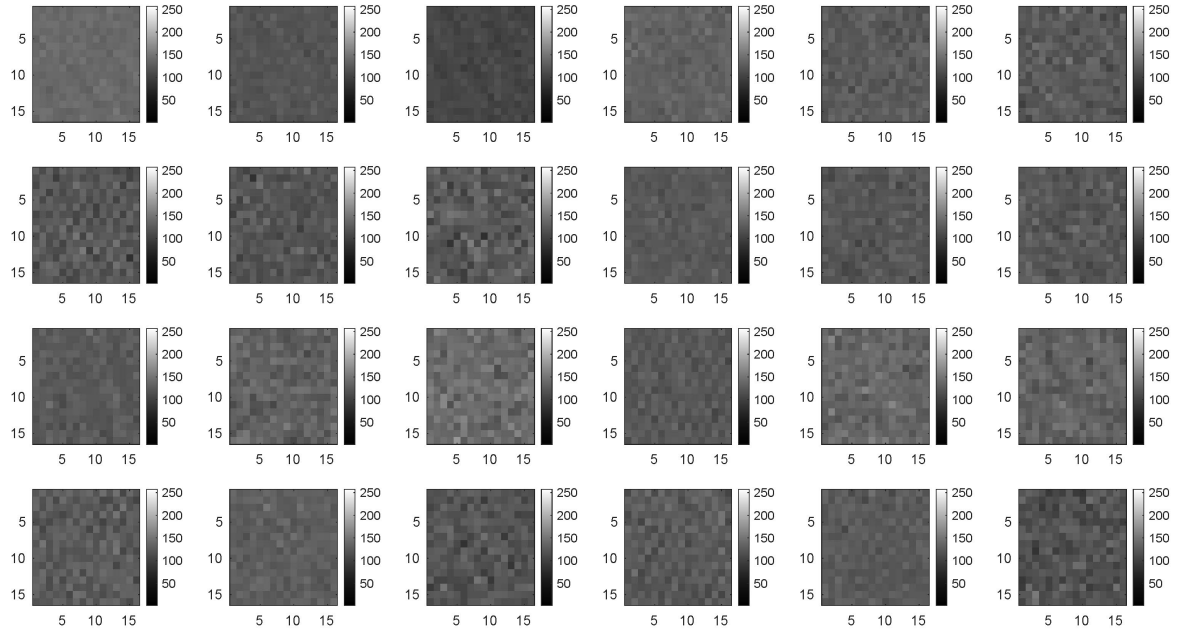


**Figure 4.** Multichannel gray-scale image corresponding to the 24 sensors (size 16 × 16).

### *3.3. Deep Convolutional Neural Network*

CNNs are feed-forward artificial neural networks that use the convolution operation instead of matrix multiplication. The preprocessing required in a CNN is significantly less than that required by other classification algorithms because features are not hand-engineered but learned. Typically, there are three kinds of layers: convolution, fully-connected, and softmax. The main aspects pf the convolution layer are its sparse local connectivity and filters, which significantly diminish the number of network parameters while simultaneously increasing its performance. The convolution layer's last step is to apply the so-called activation function, which is a nonlinear function. Fully-connected layers are normal neural network layers in which all the outputs from the previous layer are connected to all the nodes in the next layer. Normally, these layers go towards the end of the network. Finally, a softmax layer assigns probabilities to each class and connects to the final output layer that will have the same number of neurons as classes.

To construct a deep CNN for a particular application is a complex task. In comparison to the wealth of research related to color images, very little work has been carried out for gray-scale images. In this work, a CNN is designed for the detection of different structural damage states based on 24-channel gray-scale images.

#### 3.3.1. Data Split: Training Set and Validation Set

To develop the classification model, deep learning methods divide the available data into training and validation sets. The training dataset is the actual dataset used to train the model (weights and biases in a CNN). In other words, the training dataset is the sample of data used to fit the model. In contrast, the validation dataset is the sample of data used to provide an unbiased evaluation of the model fit on the training dataset while tuning the model hyperparameters.

In this work, the following dataset split ratio has been used: 75% of the whole dataset is assigned to the training set, and 25% is assigned to the validation set. That is, 1,209,675 images with data augmentation, or 4800 without data augmentation, are used to train the CNN. Then, 403,225 images with data augmentation, or 1600 without data augmentation, are used to validate the model.

### 3.3.2. Network Architecture

The network presented in Figure 5 was designed in this work.

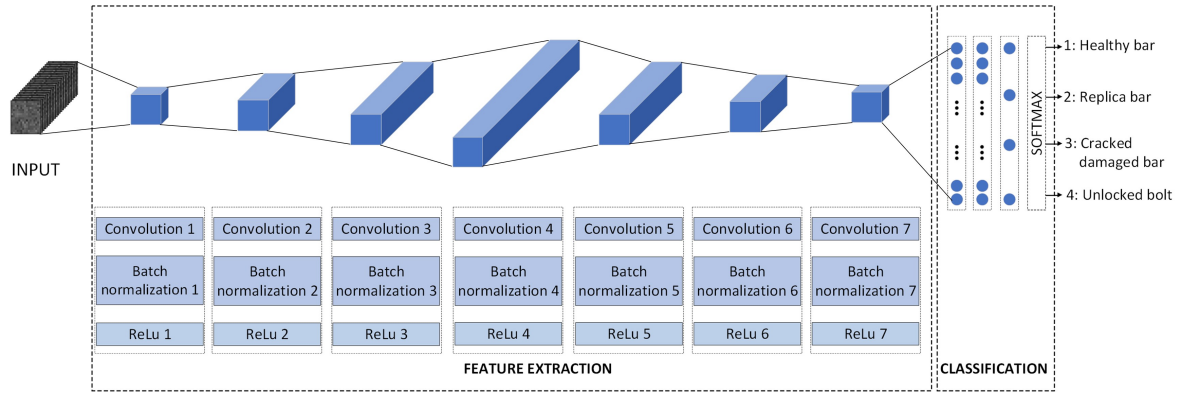



**Figure 5.** Architecture of the developed convolutional neural network (CNN).

The input is a $16 \times 16$ image with 24 channels, all of which are gray-scale. Figure 6 shows an example of one image in the dataset that was obtained after the preprocess procedure stated in Section 3.2.

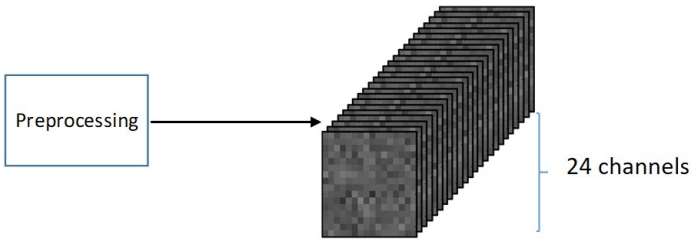



**Figure 6.** Example of one image in the dataset (24 channels) used as the CNN input.

The input is convoluted by a series of 7 convolutional layers. Each convolution layer is followed by a batch normalization step, which is used to improve the speed, performance, and stability of the CNN [22]; and a ReLU (Rectified Linear Unit) activation function ($f(x) = max(0, x)$), because this approach has been shown to speed up the training process in comparison to the classic sigmoid alternative. The final layers of the network are three fully connected layers and a softmax block, used to squash the 4-dimensional output into a categorical probability distribution: (1) original healthy bar, (2) replica bar, (3) crack damaged bar, and (4) unlocked bolt.

The most significant characteristics of the CNN architecture are summarized in Table 4.

It should be noted that the convolutions with a maximum number of parameters are the intermediate case (convolutions 4 and 5), whereas those with the minimum number of parameters correspond to the first and last convolutions. Finally, the three fully connected layers have sizes 32, 16, and 4, respectively, and are followed by the softmax function with four outputs.

**Table 4.** Characteristics of the designed CNN. The neural network has a total of 2,176,308 parameters. The number of output channels in each layer is highlighted in boldface font.

| Layer | Ouput size | Parameters | # of Parameters |
|---|---|---|---|
| Input<br>16 × 16 × 24 images | 16 × 16 × **24** | - | 0 |
| Convolution#1<br>32 filters of size 5 × 5 × 24 with padding [1 1 1 1] | 14 × 14 × **32** | Weight 5 × 5 × 24 × 32<br>Bias 1 × 1 × 32 | 19,232 |
| Batch Normalization#1 | 14 × 14 × **32** | Offset 1 × 1 × 32<br>Scale 1 × 1 × 32 | 64 |
| ReLu#1 | 14 × 14 × **32** | - | 0 |
| Convolution#2<br>64 filters of size 5 × 5 × 24 with padding [1 1 1 1] | 12 × 12 × **64** | Weight 5 × 5 × 32 × 64<br>Bias 1 × 1 × 64 | 51,264 |
| Batch Normalization#2 | 12 × 12 × **64** | Offset 1 × 1 × 64<br>Scale 1 × 1 × 64 | 128 |
| ReLu#2 | 12 × 12 × **64** | - | 0 |
| Convolution#3<br>128 filters of size 5 × 5 × 24 with padding [1 1 1 1] | 10 × 10 × **128** | Weight 5 × 5 × 64 × 128<br>Bias 1 × 1 × 128 | 204,928 |
| Batch Normalization#3 | 10 × 10 × **128** | Offset 1 × 1 × 128<br>Scale 1 × 1 × 128 | 256 |
| ReLu#3 | 10 × 10 × **128** | - | 0 |
| Convolution#4<br>256 filters of size 5 × 5 × 24 with padding [1 1 1 1] | 8 × 8 × **256** | Weight 5 × 5 × 128 × 256<br>Bias 1 × 1 × 256 | 819,456 |
| Batch Normalization#4 | 8 × 8 × **256** | Offset 1 × 1 × 256<br>Scale 1 × 1 × 256 | 512 |
| ReLu#4 | 8 × 8 × **256** | - | 0 |
| Convolution#5<br>128 filters of size 5 × 5 × 24 with padding [1 1 1 1] | 6 × 6 × **128** | Weight 5 × 5 × 256 × 128<br>Bias 1 × 1 × 128 | 819,456 |
| Batch Normalization#5 | 6 × 6 × **128** | Offset 1 × 1 × 128<br>Scale 1 × 1 × 128 | 256 |
| ReLu#5 | 6 × 6 × **128** | - | 0 |
| Convolution#6<br>64 filters of size 5 × 5 × 24 with padding [1 1 1 1] | 4 × 4 × **64** | Weight 5 × 5 × 128 × 64<br>Bias 1 × 1 × 64 | 204,864 |
| Batch Normalization#6 | 4 × 4 × **64** | Offset 1 × 1 × 64<br>Scale 1 × 1 × 64 | 128 |
| ReLu#6 | 4 × 4 × **64** | - | 0 |
| Convolution#7<br>32 filters of size 5 × 5 × 24 with padding [1 1 1 1] | 2 × 2 × **32** | Weight 5 × 5 × 64 × 32<br>Bias 1 × 1 × 32 | 51,232 |
| Batch Normalization#7 | 2 × 2 × **32** | Offset 1 × 1 × 32<br>Scale 1 × 1 × 32 | 64 |
| ReLu#7 | 2 × 2 × **32** | - | 0 |
| Fully connected layer#1 | 1 × 1 × **32** | Weight 32 × 128<br>Bias 32 × 1 | 4128 |
| Fully connected layer#2 | 1 × 1 × **16** | Weight 16 × 32<br>Bias 16 × 1 | 528 |
| Fully connected layer#3 | 1 × 1 × **4** | Weight 4 × 16<br>Bias 4 × 1 | 68 |
| Softmax | - | - | 0 |
| classoutput | - | - | 0 |

It should also be noted that each convolution employs a padding of 1. The main intuition behind this selection is that normally, the filter is applied by superimposing it on the image from the upper left edge. Then, a columnar translation is applied until the filter is superimposed with its right edge on the right edge of the image. This usual way of proceeding has a problem, the edge pixels are never subjected to the central part of the filter. This is sometimes known as the border effect problem and can be solved by incorporating so-called padding [23]. That is to apply the filter beginning from outside the image frame as well as ending also outside the image, in such a manner that edge pixels reach also the center part of the filter. In this work, a padding of 1 is used to enhance the texture features

extracted by the CNN for all of the data in the image, regardless of whether the data are located in the image. Table 5 compares different metrics (see Section 4.1, where a definition of these metrics is given) with and without padding (without data augmentation). It can be observed that when using padding, better results are attained.

**Table 5.** Metrics for different CNN architectures without data augmentation. The best metric results are highlighted in boldface font.

| Strategy | Accuracy | Precision | Recall | F1 Score | Specificity |
|---|---|---|---|---|---|
| ReLu-Padding-L2 regularization | **93.81** | **92.77** | **93.73** | **93.22** | **97.98** |
| Relu-No padding-L2 regularization | 93.69 | 92.73 | 93.44 | 93.07 | 97.92 |
| Relu-Padding-No L2 regularization | 93.63 | 92.73 | 93.82 | 93.25 | 97.89 |

#### 3.3.3. Network Training

The training of the CNN consists of the minimization of a loss function by means of a numerical optimization algorithm. In this work, the Adam optimizer [24] is employed to minimize the categorical cross entropy [25]. The Adam algorithm combines two versions of speeding up gradient descent: (i) gradient descent with momentum, where the basic idea is to compute an exponentially weighted average of the gradients, and (ii) root mean square propagation (RMSProp), that makes use of the gradient second moments. Specifically, the Adam numerical method puts together the exponential moving average of the gradient and the squared gradient (second moment), and hyperparameters $\beta_1$ and $\beta_2$ handle their decrease rates, respectively. In this work, the Adam optimizer has been tuned and thus employs an initial learning rate of $\alpha_0 = 0.01$, and values $\beta_1 = 0.9$, $\beta_2 = 0.992$, and $\varepsilon = 10^{-7}$ to avoid divisions by zero. Furthermore, here, the learning rate is decreased every 2 epochs by multiplying with factor 0.5.

Convolutional layer initialization is carried out by the so-called Xavier initializer [26]. Mini-batches of size 75 in the initial dataset and 590 for the augmented dataset are used to update the weights.

Finally, $L2$ regularization with $\lambda = 10^{-6}$ is employed. Table 5 compares the different metrics (see Section 4.1 for a definition of these metrics) with and without $L2$ regularization (without data augmentation). It can be observed that when using regularization, better results are obtained because regularization reduces high variance in the validation set.

#### 3.3.4. Network Architecture and Hyperparameter Tuning

To select the best architecture and to tune the different hyperparameters usually requires significant computational resources. As one of the most critical aspects of computational cost is the dataset size, in this paper, following the results presented in [27,28], the small dataset (without augmentation) is used to define the CNN architecture and quickly (coarse) tune the hyperparameters. Next, the obtained optimal hyperparameters for the small dataset are used as initial values to fine-tune the hyperparameters with the large dataset (with data augmentation).

#### 3.3.5. Network Implementation

The stated methodology is coded in MATLAB® using its Deep Learning Toolbox™ on a laptop running the Windows® 10 operating system, with an Intel Core i7-9750H processor, 16 GB of RAM, and an Nvidia GeForce RTX™2060 graphics card that has 6 GB of dedicated VRAM.

## 4. Results and Discussion

### *4.1. Metrics to Evaluate the Classification Model*

To measure classification performance, several metrics can be computed from a confusion matrix such as that shown in Table 6. Normally, these metrics evaluate binary classification problems.

Note that true positive (TP) is the number of positive samples that are correctly predicted as such, false positive (FP) is the number of negative samples that are incorrectly predicted, true negative (TN) is the number of negative samples that are correctly predicted, and false negative (FN) is the number of positive samples that are incorrectly predicted.

**Table 6.** Binary confusion matrix.

| | | **Predicted Class** | |
|---|---|---|---|
| | | **Positive** | **Negative** |
| **Actual class** | **Positive** | True positive (TP) | False negative (FN) |
| | **Negative** | False positive (FP) | True negative (TN) |

The most common metrics for binary classification problems are the following.

- Accuracy: proportion of true results (both true positives and true negatives) among the total number of cases examined.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}}$$

- Precision: proportion of positive results that are true positive.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

- Recall: proportion of actual positives that are correctly identified as such.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

- Specificity: proportion of actual negatives that are correctly identified as such.

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}}$$

- F1-score: harmonic mean of the precision and recall.

$$\text{F1} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

In a multiclass classification problem, such as that considered in this work, these metrics are also applicable using a one-vs.-all approach to compute each metric for each class, see [29]. This is, essentially, to compute the different metrics for each label as if the problem has been reduced to a binary 'label *X*' versus 'not label *X*' situation.

### *4.2. Results of the CNN Classification Method*

To evaluate the developed methodology, this section presents the results obtained from the proposed SHM strategy. A flowchart of the proposed approach is given in Figure 7. When a WT must be diagnosed, the accelerometer data are scaled, reshaped, and converted into gray-scale images that are fed into the already trained CNN, and a classification is obtained to predict the structural state condition.

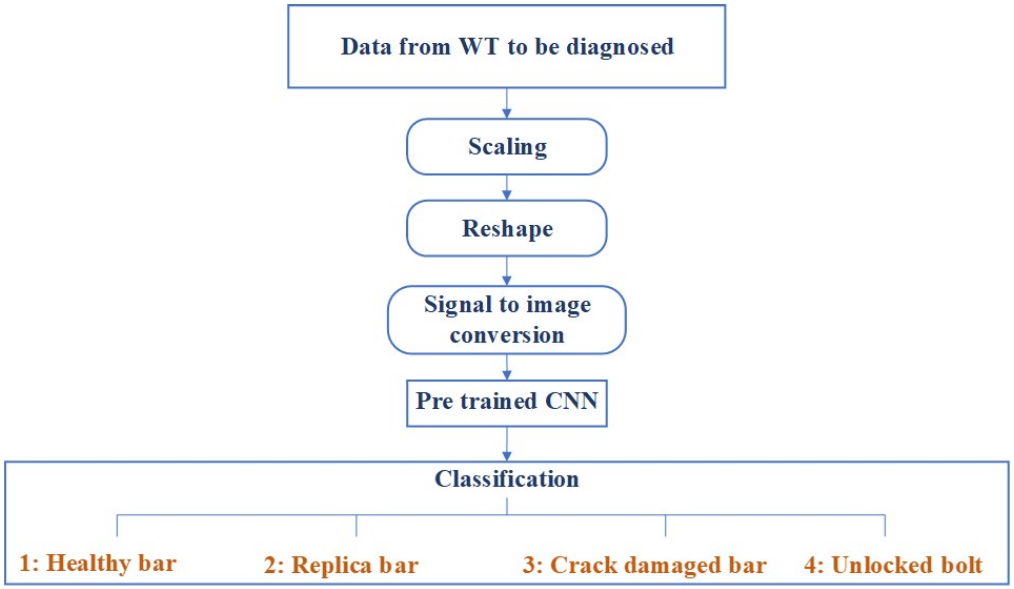



**Figure 7.** Flowchart to illustrate how the proposed structural health monitoring (SHM) strategy is applied when a wind turbine (WT) must be diagnosed.

To thoroughly test the functional characteristics of the algorithm, the datasets with and without data augmentation are considered, as well as comparison with two other methodologies given in [9,17] that make use of the same laboratory structure. The first methodology, given in [17], is based on principal component analysis and support vector machines. The second methodology, given in [9] (page 67), is based on the well-known damage indicators: covariance matrix estimate and scalar covariance.

Figures 8 and 9 illustrate the confusion matrices for the validation dataset without and with data augmentation, respectively. The rows represent the true class, whereas columns represent the predicted class. The precision and false discovery rate are given in the rightmost columns. Finally, the recall and false negative rate are given at the bottom rows. An examination of both confusion matrices reveals that some misclassifications come from the model confounding the healthy and replica bars (labels 1 and 2). However, this level of misclassification is acceptable because both bars are in a healthy state. In contrast, some errors are derived from the model misclassifying the crack and unlocked bolt damages (labels 3 and 4), which will not correctly detect the type of damage but would at least lead to a damage alert. Finally, it should be noted that very few damaged samples (labels 3 and 4) are classified as healthy or replica bar (labels 1 and 2).

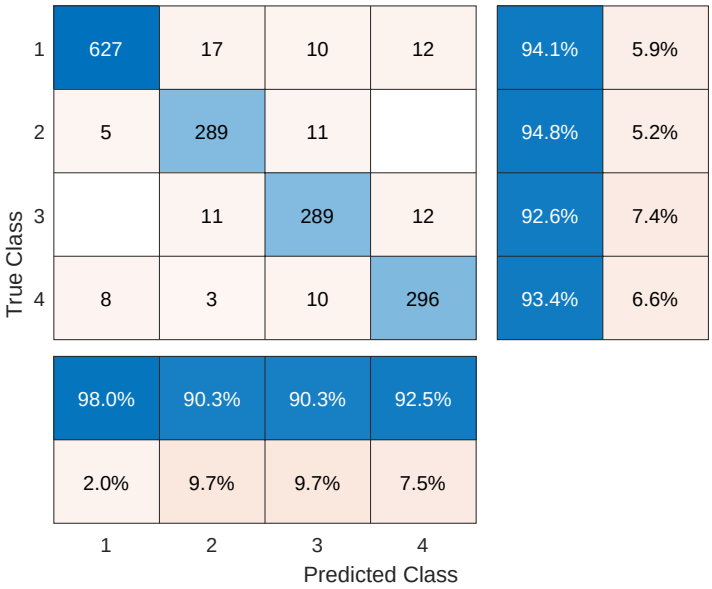



**Figure 8.** Confusion matrix for the validation dataset without data augmentation.

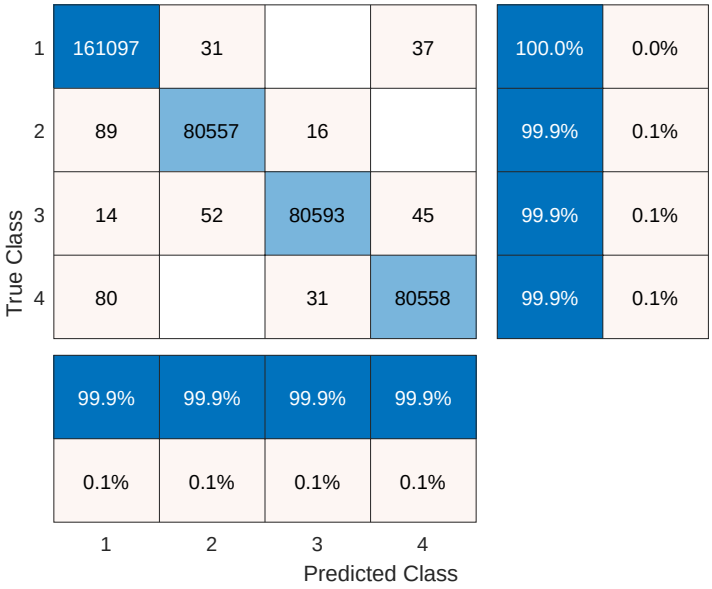



**Figure 9.** Confusion matrix for the validation dataset with data augmentation.

From the confusion matrices, the different metrics to evaluate the classification model, see Section 4.1, are computed and presented in Table 7.

**Table 7.** Metrics for each label of the multiclassification problem and comparison between the datasets without and with data augmentation.

| Dataset | Label | Precision | Recall | F1-Score | Specificity |
|---|---|---|---|---|---|
| Without data augmentation | 1: Healthy bar | 97.97 | 94.14 | 96.02 | 98.61 |
| | 2: Replica bar | 90.31 | 94.75 | 92.48 | 97.61 |
| | 3: Crack damaged bar | 90.31 | 92.63 | 91.46 | 97.59 |
| | 4: Unlocked bolt | 92.50 | 93.38 | 92.94 | 98.13 |
| With data augmentation | 1: Healthy bar | 99.89 | 99.96 | 99.92 | 99.92 |
| | 2: Replica bar | 99.90 | 99.87 | 99.88 | 99.97 |
| | 3: Crack damaged bar | 99.94 | 99.86 | 99.90 | 99.99 |
| | 4: Unlocked bolt | 99.90 | 99.86 | 99.88 | 99.97 |

The impact of the data augmentation strategy can clearly be seen. Although no new experimental data were collected, nonetheless the metrics were significantly improved. It should be noted that all of the metrics (precision, recall, F1-score, and specificity) are higher than or equal to 99.86% for each label when using the augmented dataset in comparison to values between 90.31% and 98.61% for the initial dataset. Despite all metrics being relevant, considering the specific problem at hand, the most important metric is recall, which is the proportion of actual damaged cases that are correctly identified as such. It can be observed that the crack damage and the unlocked bolt, even without data augmentation, obtain recall values of 92.63% and 93.38%, respectively. When data augmentation is used, the recall values are all higher than or equal to 99.86% for all of the studied classes. The results associated with the precision metric are also satisfactory. When the initial dataset is used, precision values are between 90.31 and 97.97, but with the augmented dataset, such values are all higher than or equal to 99.89. Finally, it should be noted that the specificity metric is that which experiences less improvement when using the augmented dataset.

As already mentioned before, here, a comparison is made between our obtained results and two other methodologies. On the one hand, when using the first approach stated in [17], the crack damaged bar has a recall of 96.08%, and is thus inferior to the one obtained with the proposed strategy in this work which attained a value of 99.86%. Note that the crack damage is the most challenging. In fact, the second approach stated in [9] (page 82) was not capable of detecting this type of incipient damage when using the scalar covariance or mean residual damage indicators. On the other hand, the first approach obtains a recall of 99.02% for the unlocked bold damage; whereas with the proposed strategy, a slightly higher value of 99.86% is obtained. Finally, note that the unlocked bold damage is not studied in the second approach

The proposed CNN exhibits low bias and variance for both datasets, because the training and validation errors are small (low bias), as well as the difference between them (low variance), as shown in Table 8. In particular, when using the initial dataset, the training error is equal to 0.1167 and the validation error is quite close to this same value, being equal to 0.1692. When using the augmented dataset, the training error diminishes to 0.0026 and the validation error is only slightly greater at 0.0044. From this table, the significantly increased training time (1196 min) of the augmented dataset in comparison to that of the initial dataset (11 min) can be seen, which is easily understood due to the size of each dataset. That is, there are 1,612,800 images in the augmented dataset and only 6400 images in the initial.

**Table 8.** Comparison of obtained accuracy, validation error, training error, and training time when using data augmentation with respect to the original dataset.

| | **Accuracy** | **Validation Error** | **Training Error** | **Training Time** | **# of Images** |
|---|---|---|---|---|---|
| Whitout data augmentation | 93.81 | 0.1692 | 0.1167 | 11 min | 6400 |
| With data augmentation | 99.90 | 0.0044 | 0.0026 | 1196 min | 1,612,800 |

Finally, Figure 10 shows the accuracy and loss curves during training and validation (black dotted lines) when using the augmented dataset. It should be noted that after 5 epochs, the CNN obtains an accuracy of 99.90% and a final validation loss of 0.0044, as shown in Table 8.

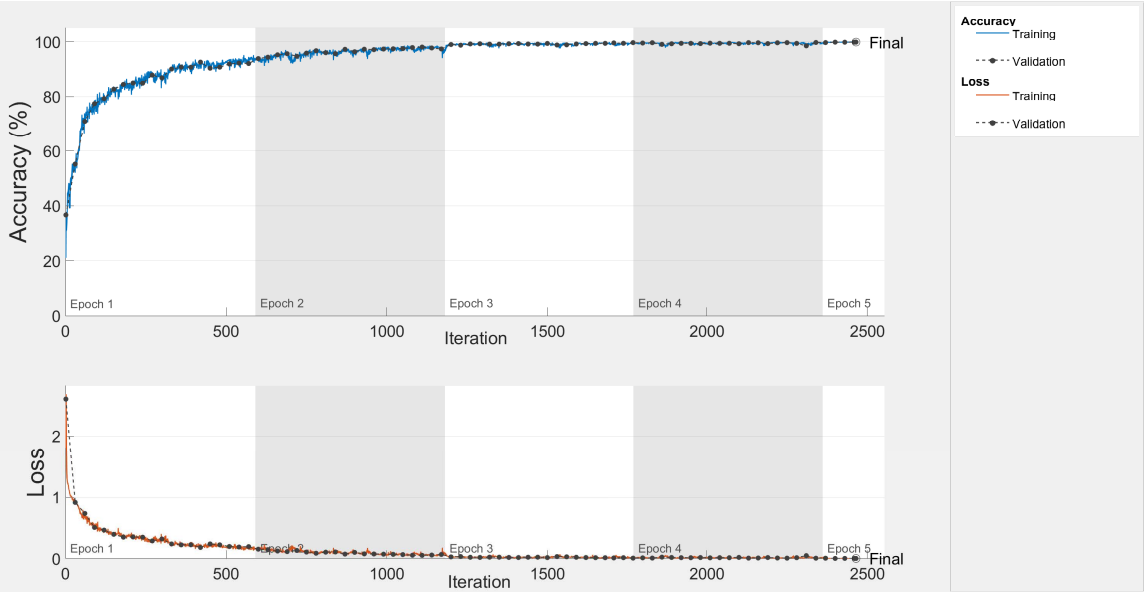



**Figure 10.** Accuracy and loss curve for the augmented dataset.

## 5. Conclusions and Future Work

In this work, a strategy based solely on vibration response was demonstrated for the structural health monitoring of offshore WT foundations. The approach was tested on a laboratory setup, for which four different structural states for a jacket bar were studied: healthy bar, replica, crack damage, and an unlocked bolt.

The contribution of this work is twofold: (i) how three-dimensional data (derived from different time, sensors, and experiments) are preprocessed (collected, scaled, reshaped, augmented, and converted into gray-scale images with as many channels as sensors), and (ii) the design of a deep CNN, the architecture and hyperparameters of which play a key role in the specific application that concerns us—damage diagnosis. Furthermore, the proposed method does not require hand-designed features beforehand because the CNN learns features automatically.

The conceived SHM methodology with data augmentation shows exceptional performance, with all considered metrics (accuracy, precision, recall, F1-score, and specificity) giving results greater than 99.8%. In particular, a noteworthy overall accuracy of 99.90% is obtained with data augmentation. These results show that large (deep) CNNs are promising for the development of SHM strategies for WT offshore foundations.

Future work will focus on three main areas. First, based on open set domain adaptation [30], research to render capability of separating unknown damage from known targeted types of damage will be conducted. Second, not only detection and classification but also the localization of the damage will be attempted by designing an ensemble of deep CNNs, the main idea being to take advantage of individual information from each sensor signal. Lastly, to deal with the validation of the proposed strategy in a more realistic environment, a water tank facility will be used, in which the laboratory tower will be placed and subjected to the action of regular and irregular waves.

**Author Contributions:** Y.V. and C.T. conceived the main conceptual ideas. B.P. performed the implementation and numerical computation as well as conceived the data augmentation strategy. All authors discussed the results and contributed to the final manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was partially funded by the Spanish Agencia Estatal de Investigación (AEI)—Ministerio de Economía, Industria y Competitividad (MINECO), and the Fondo Europeo de Desarrollo Regional (FEDER) through research project DPI2017-82930-C2-1-R; and by the Generalitat de Catalunya through research project 2017 SGR 388. We gratefully acknowledge the support of NVIDIA Corporation with the donation of the Titan XP GPU used for this research.

**Acknowledgments:** We thank the three anonymous reviewers for their careful reading of our manuscript and their many insightful comments and suggestions.

**Conflicts of Interest:** The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; nor in the decision to publish the results.

## References

1. Ohlenforst, K.; Backwell, B.; Council, G.W.E. Global Wind Report 2018. Available online: https://gwec.net/global-wind-report-2018/ (accessed on 15 June 2020).
2. Lai, W.J.; Lin, C.Y.; Huang, C.C.; Lee, R.M. Dynamic analysis of Jacket Substructure for offshore wind turbine generators under extreme environmental conditions. *Appl. Sci.* **2016**, *6*, 307. [CrossRef]
3. Moulas, D.; Shafiee, M.; Mehmanparast, A. Damage analysis of ship collisions with offshore wind turbine foundations. *Ocean. Eng.* **2017**, *143*, 149–162. [CrossRef]
4. Van Kuik, G.; Peinke, J. *Long-Term Research Challenges in Wind Energy-A Research Agenda by the European Academy of Wind Energy*; Springer: Berlin/Heidelberg, Germany, 2016; Volume 6.
5. Liu, W.; Tang, B.; Han, J.; Lu, X.; Hu, N.; He, Z. The structure healthy condition monitoring and fault diagnosis methods in wind turbines: A review. *Renew. Sustain. Energy Rev.* **2015**, *44*, 466–472. [CrossRef]
6. Qing, X.; Li, W.; Wang, Y.; Sun, H. Piezoelectric transducer-based structural health monitoring for aircraft applications. *Sensors* **2019**, *19*, 545. [CrossRef] [PubMed]
7. Weijtjens, W.; Verbelen, T.; De Sitter, G.; Devriendt, C. Foundation structural health monitoring of an offshore wind turbine: A full-scale case study. *Struct. Health Monit.* **2016**, *15*, 389–402. [CrossRef]
8. Oliveira, G.; Magalhães, F.; Cunha, Á.; Caetano, E. Vibration-based damage detection in a wind turbine using 1 year of data. *Struct. Control. Health Monit.* **2018**, *25*, e2238. [CrossRef]
9. Zugasti Uriguen, E. Design and validation of a methodology for wind energy structures health monitoring. Ph.D. Thesis, Universitat Politècnica de Catalunya, Barcelona, Spain, 2014.
10. Lee, J.J.; Lee, J.W.; Yi, J.H.; Yun, C.B.; Jung, H.Y. Neural networks-based damage detection for bridges considering errors in baseline finite element models. *J. Sound Vib.* **2005**, *280*, 555–578. [CrossRef]
11. Kim, B.; Min, C.; Kim, H.; Cho, S.; Oh, J.; Ha, S.H.; Yi, J.H. Structural health monitoring with sensor data and cosine similarity for multi-damages. *Sensors* **2019**, *19*, 3047. [CrossRef] [PubMed]
12. Stetco, A.; Dinmohammadi, F.; Zhao, X.; Robu, V.; Flynn, D.; Barnes, M.; Keane, J.; Nenadic, G. Machine learning methods for wind turbine condition monitoring: A review. *Renew. Energy* **2019**, *133*, 620–635. [CrossRef]
13. Ruiz, M.; Mujica, L.E.; Alferez, S.; Acho, L.; Tutiven, C.; Vidal, Y.; Rodellar, J.; Pozo, F. Wind turbine fault detection and classification by means of image texture analysis. *Mech. Syst. Signal Process.* **2018**, *107*, 149–167. [CrossRef]
14. Vidal, Y.; Pozo, F.; Tutivén, C. Wind turbine multi-fault detection and classification based on SCADA data. *Energies* **2018**, *11*, 3018. [CrossRef]
15. Spanos, N.I.; Sakellariou, J.S.; Fassois, S.D. Exploring the limits of the Truncated SPRT method for vibration-response-only damage diagnosis in a lab-scale wind turbine jacket foundation structure. *Procedia Eng.* **2017**, *199*, 2066–2071. [CrossRef]
16. Spanos, N.A.; Sakellariou, J.S.; Fassois, S.D. Vibration-response-only statistical time series structural health monitoring methods: A comprehensive assessment via a scale jacket structure. *Struct. Health Monit.* **2019**, *19*, 736–750. [CrossRef]
17. Vidal Seguí, Y.; Rubias, J.L.; Pozo Montero, F. Wind turbine health monitoring based on accelerometer data. In Proceedings of the 9th ECCOMAS Thematic Conference on Smart Structures and Materials, Paris, France, 8–11 July 2019; pp. 1604–1611.
18. Vidal, Y.; Aquino, G.; Pozo, F.; Gutiérrez-Arias, J.E.M. Structural Health Monitoring for Jacket-Type Offshore Wind Turbines: Experimental Proof of Concept. *Sensors* **2020**, *20*, 1835. [CrossRef] [PubMed]
19. Pozo, F.; Vidal, Y.; Serrahima, J. On real-time fault detection in wind turbines: Sensor selection algorithm and detection time reduction analysis. *Energies* **2016**, *9*, 520. [CrossRef]

20. Pal, K.K.; Sudeep, K. Preprocessing for image classification by convolutional neural networks. In Proceedings of the 2016 IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), Bangalore, India, 20–21 May 2016; pp. 1778–1781.
21. Chen, X.W.; Lin, X. Big data deep learning: challenges and perspectives. *IEEE Access* **2014**, *2*, 514–525. [CrossRef]
22. Santurkar, S.; Tsipras, D.; Ilyas, A.; Madry, A. How does batch normalization help optimization? In Proceedings of the conference on Advances in Neural Information Processing Systems, Montreal, QC, Canada, 3–8 December 2018; pp. 2483–2493.
23. Albawi, S.; Mohammed, T.A.; Al-Zawi, S. Understanding of a convolutional neural network. In Proceedings of the 2017 International Conference on Engineering and Technology (ICET), Antalya, Turkey, 21–23 August 2017; pp. 1–6.
24. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. In Proceedings of the International Conference on Learning Representations, San Diego, CA, USA, 7–9 May 2015.
25. Rusiecki, A. Trimmed categorical cross-entropy for deep learning with label noise. *Electron. Lett.* **2019**, *55*, 319–320. [CrossRef]
26. Glorot, X.; Bengio, Y. Understanding the difficulty of training deep feedforward neural networks. In Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, Sardinia, Italy, 13–15 May 2010; pp. 249–256.
27. DeCastro-García, N.; Muñoz Castañeda, Á.L.; Escudero García, D.; Carriegos, M.V. Effect of the Sampling of a Dataset in the Hyperparameter Optimization Phase over the Efficiency of a Machine Learning Algorithm. *Complexity* **2019**, *2019*, 6278908. [CrossRef]
28. Swersky, K.; Snoek, J.; Adams, R.P. Multi-task bayesian optimization. In Proceedings of the Advances in Neural Information Processing Systems, Stateline, Nevada, USA, 5–10 December 2013; pp. 2004–2012.
29. Hossin, M.; Sulaiman, M. A review on evaluation metrics for data classification evaluations. *Int. J. Data Min. Knowl. Manag. Process. (IJDKP)* **2015**, *5*, 1–11.
30. Saito, K.; Yamamoto, S.; Ushiku, Y.; Harada, T. Open set domain adaptation by backpropagation. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 153–168.


# Appended Paper II

# Siamese Neural Networks for Damage Detection and Diagnosis of Jacket-Type Offshore Wind Turbine Platforms

**Authors:**
Joseph Baquerizo, Christian Tutivén, Bryan Puruncajas, Yolanda Vidal and José Sampietro

**Paper published in:**
*Mathematics* 2022, 10(7), 1131
DOI: 10.3390/math10071131
Ranking JCR: Q1 (2022)

**Paper history:**
Submitted: 22-02-2022
Accepted: 29-03-2022
Published: 01-04-2022

*mathematics*

MDPI

*Article*

# Siamese Neural Networks for Damage Detection and Diagnosis of Jacket-Type Offshore Wind Turbine Platforms

**Joseph Baquerizo [1], Christian Tutivén [1], Bryan Puruncajas [1,2], Yolanda Vidal [2,3,*] and José Sampietro [4]**

[1] Mechatronics Engineering, Faculty of Mechanical Engineering and Production Science (FIMCP), Escuela Superior Politécnica del Litoral (ESPOL), Campus Gustavo Galindo Km. 30.5 Vía Perimetral, Guayaquil P.O. Box 09-01-5863, Ecuador; josanbaq@espol.edu.ec (J.B.); cjtutive@espol.edu.ec (C.T.); bpurunca@espol.edu.ec (B.P.)
[2] Control, Modeling, Identification and Applications (CoDAlab), Department of Mathematics, Escola d'Enginyeria de Barcelona Est (EEBE), Campus Diagonal-Besós (CDB), Universitat Politècnica de Catalunya (UPC), Eduard Maristany 16, 08019 Barcelona, Spain
[3] Institute of Mathematics (IMTech), Universitat Politècnica de Catalunya (UPC), Pau Gargallo 14, 08028 Barcelona, Spain
[4] Facultad de Ingenierías, Universidad ECOTEC, Km. 13.5 Vía a Samborondón, Guayaquil 092302, Ecuador; joseluis.sampietro05@gmail.com
* Correspondence: yolanda.vidal@upc.edu

**Abstract:** Offshore wind energy is increasingly being realized at deeper ocean depths where jacket foundations are now the greatest choice for dealing with the hostile environment. The structural stability of these undersea constructions is critical. This paper states a methodology to detect and classify damage in a jacket-type support structure for offshore wind turbines. Because of the existence of unknown external disturbances (wind and waves), standard structural health monitoring technologies, such as guided waves, cannot be used directly in this application. Therefore, using vibration-response-only accelerometer measurements, a methodology based on two in-cascade Siamese convolutional neural networks is proposed. The first Siamese network detects the damage (discerns whether the structure is healthy or damaged). Then, in case damage is detected, a second Siamese network determines the damage diagnosis (classifies the type of damage). The main results and claims of the proposed methodology are the following ones: (i) It is solely dependent on accelerometer sensor output vibration data, (ii) it detects damage and classifies the type of damage, (iii) it operates in all wind turbine regions of operation, (iv) it requires less data to train since it is built on Siamese convolutional neural networks, which can learn from very little data compared to standard machine/deep learning algorithms, (v) it is validated in a scaled-down experimental laboratory setup, and (vi) its feasibility is demonstrated as all computed metrics (accuracy, precision, recall, and F1 score) for the obtained results remain above 96%.

**Keywords:** offshore fixed wind turbine; jacket structure; damage detection; damage diagnosis; vibration-based SHM; data-driven; Siamese neural network; convolutional neural network

**MSC:** 00A69; 68T07



**Citation:** Baquerizo, J.; Tutivén, C.; Puruncajas, B.; Vidal, Y.; Sampietro, J. Siamese Neural Networks for Damage Detection and Diagnosis of Jacket-Type Offshore Wind Turbine Platforms. *Mathematics* **2022**, *10*, 1131. https://doi.org/10.3390/math10071131

Academic Editor: Ripon Kumar Chakrabortty

Received: 22 February 2022
Accepted: 29 March 2022
Published: 1 April 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



## 1. Introduction

The world energy system is undoubtedly in transition. The widespread adoption and use of renewable energy is key to fighting climate change and ensuring a sustainable future. According to WindEurope [1], the European Commission's forecasts demonstrate that renewable-based electricity will be critical to achieving climate neutrality in Europe by 2050. This will need wind accounting for 50% of the EU's power mix, with renewables accounting for 81%. To accomplish this target, offshore wind is a crucial component as it offers higher and steadier wind speeds and vast possibilities for their placement (easy to find new locations compared to on-shore).

The crux of the matter in the advancement of the offshore wind industry is the reduction in the levelized cost of energy, and a key factor to accomplishing it is through the optimization of the inspection and maintenance strategies. In particular, structural health monitoring (SHM) of the support structure has an added value, estimated to be between 1.93–18.11 M€, see [2]. An effective structural integrity management of offshore wind farm assets could reduce the number of inspections they need during their lifetime. SHM has been widely applied to civil infrastructures [3], such as bridges, but there is a research gap in the area when being applied to hazardous environments, such as the ones where support structures of offshore wind turbines (WT) are installed. On the one hand, in this complex environment, the data suffer from noise measurement that must be removed in deep learning methods. Noise reduction in the data are crucial and can be accomplished, for example, by a Savitzky–Golay filter and wavelet decomposition, e.g., [4]. On the other hand, WTs are placed in a marine environment, subject to potential extreme winds, waves, and currents, which can change rapidly and are initially unknown. Related to the main challenges of SHM in offshore foundations, as stated in [5]: "A defining marine environment main characteristic is that structures are always subject to excitations. Techniques for structural health monitoring, vibration, and data analysis must be capable of coping with such ambient excitations. As the input is typically not known, a normal input–output formalism cannot be used". That is, the standard SHM approach based on guided waves (where the input excitation is known and imposed to the structure and then the output vibration is measured), widely used in many areas such as aeronautics [6], cannot be applied in a straightforward manner to offshore WTs; as the excitation is not known (wind, waves, currents), neither can be imposed. Thus, an output-only approach is imperative.

A new paradigm, a vibration-response-only methodology, must be developed that assumes unknown input excitations and that only the vibration response is measurable by means of different sensors (accelerometers or fiber Bragg grating, for instance). In recent years, interest in this type of methodology has grown. For example, in [7], parametric reduced order models for cracked shells are developed and applied to crack detection problems, and an output-only scheme is adopted based on transmissibility functions. It is also noteworthy that the vibration-response-only approach for a jacket structure in [8] where a comprehensive and critical assessment of the diagnostic performance of five prominent response-only methods is presented based on incipient, 'minor' to 'mild', damages on a lab-scale wind turbine jacket structure. In [9], an SHM method for floating offshore WTs was tested using operational modal analysis. The results showed that the curvature mode shape was the most effective modal property to detect damage location and intensity. Likewise, Ref. [10] contributed an SHM system for real tripod WT supports based on fiber Bragg grating (FBG) sensors to detect and localize the damage. A meaningful work was presented in [11], where a time–frequency analysis is proposed based on single mode function decomposition to overcome the mode-mixing problem. Finally, some publications using the same test bench as the one used in this study are: [12], where the SHM for jacket foundations is stated via a signal-to-image conversion of the accelerometer data into multichannel images and convolutional neural networks, combined with synthetic data augmentation; Ref. [13] that proposes the fractal dimension as a suitable feature to identify and classify different types of damage; and Ref. [14], where structural damage classification is achieved by using principal component analysis and extreme gradient boosting. It is noteworthy that, in contrast to all aforementioned references, where large datasets with faulty data are available (or synthetic data need to be generated); in this work, Siamese neural networks (SNNs) are used, taking advantage of their ability to learn from very little data. Furthermore, most of the aforementioned references only detect one specific type of damage but do not face the challenge of detecting and classifying different types of damage, which is accomplished in this study.

SNNs are made up of two identical artificial neural networks that function in parallel and compare their outputs at the end, typically using a distance metric. The output of

the SNNs execution may be thought of as a semantic similarity between the projected representations of the two input vectors. The ability to learn from very little data has made SNNs more popular in recent years, being applied in a wide variety of applications. For example, in [15], SNNs are employed in differential lung diagnoses with CT scans as a key element of the proposed approach to facilitate the implementation of explainable artificial intelligence systems. In [16], they are proposed to identify cyber-physical attacks dealing with the problem of few labeled data and to alleviate the over-fitting issue while enhancing accuracy. In [17], robust and discriminative gait features for human identification are automatically extracted based on SNNs and limited training data. However, to the best of the authors' knowledge, SNNs have not yet been used in the area of damage detection and/or diagnosis. In this work, two in cascade Siamese convolutional neural networks are proposed to detect and classify the faults under study.

This work contributes a vibration-response-only SHM methodology for jacket type support structures in offshore WTs, based on convolutional SNNs. In contrast to standard SNNs, which are feedforward neural networks [18], this is proposed to introduce convolutional layers. To avoid the traditional complex feature extraction processes that appear in machine learning approaches [19], this study advises utilizing deep convolutional SNNs. Thus, the initial raw accelerometer data will be converted into gray-scale multichannel images, and then features will be automatically extract by the deep convolutional SNNs. The methodology then follows the subsequent steps: (i) Vibration data are acquired, (ii) conversion of the dataset to multichannel gray-scale images, (iii) a first convolutional SNN discerns between healthy and damaged structural states, and a second convolutional SNN classifies the samples, detected as damaged by the first network, between crack or unlocked bolt types of damage. In a nutshell, the contributions of the proposed methodology that should be highlighted are:

- It is based only on the output vibration data gathered by accelerometer sensors (the excitation given by the wind is assumed to be unknown). Thus, it is a vibration-response-only methodology.
- It achieves damage detection and, in case damage is detected, damage type classification based on two in-cascade Siamese convolutional neural networks.
- It works under all regions of operation of the wind turbine.
- It needs little data to be trained, as it is based on Siamese convolutional neural networks that have the ability to learn from very little data in comparison to standard machine learning approaches.
- It is tested in a downscaled experimental laboratory structure.
- The performance indicators show all results above 96%.

The following is the paper's outline. The experimental down-scaled setup is introduced in Section 2. Section 3 details the proposed strategy. Finally, findings are discussed in Section 4, and conclusions are derived in Section 5.

## 2. Laboratory Setup

The configuration of the experimental test bench is detailed in Figure 1. The process begins with the white noise signal obtained by the function generator model GW INSTEK AF-2005. The generated signal is amplified and enters the inertial shaker model Data Physics GW-IV47, located in the upper side of the scale turbine structure, which simulates gusts of wind. The vibrations produced by the wind gust simulation are directly related to the amplitude of white noise, which has factors of 0.5, 1, 2, and 3. Vibration monitoring is carried out using eight triaxial accelerometers (PCB R Piezotronic, model 356A17), positioned as shown in Figure 2 (right); there are 24 vibration signals. The sensors are linked up to six National InstrumentsTM cartridges (model NI 9234) that are attached to the National Instruments cDAQ-9188 chassis.

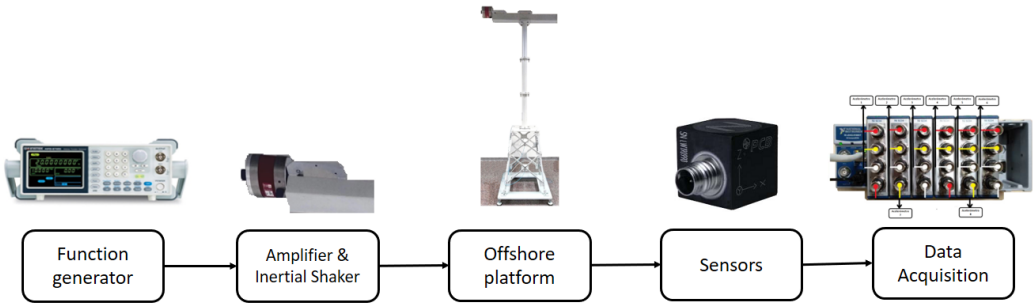



**Figure 1.** Experimental test bed.

The structure itself is 2.7 m high and divided into three sections (see Figure 2 (left)) shown below.

- The upper section is composed of a bar one meter long and 0.6 m wide; here, the wind turbine nacelle and the wind speed are simulated using the agitator and different excitation signals.
- The central section is made up of a tower divided into three parts and bolted together.
- Finally, at the bottom is the jacket section, which is made up of 32 S275JR steel bars, DC01 LFR steel sheets, and components like screws and nuts. All sections are screwed in with a torque of 12 Nm.

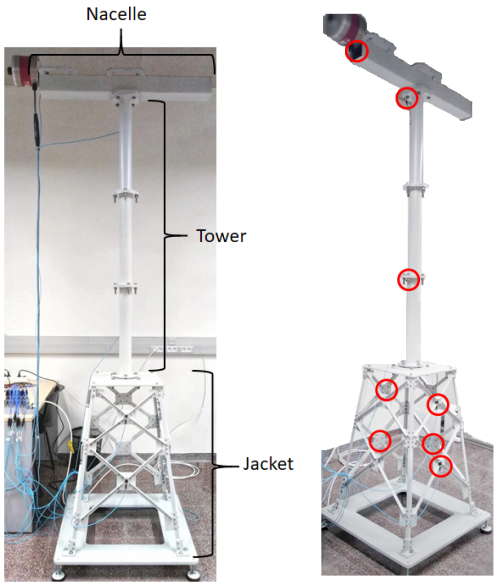



**Figure 2.** Components of the scale offshore jacket-type support laboratory model (**left**); sensor location on the wind turbine down-scaled model (**right**).

The approach of the proposed strategy is that it must be able to detect and classify the types of damage studied, as well as be robust enough to replace a bar with a new healthy one (avoiding false alarms). The following structural states are presented:

- Bar damaged by a 5 mm crack;
- Bar with unlocked bolt;
- Replica bar.

Figure 3 shows the different structural states in detail.

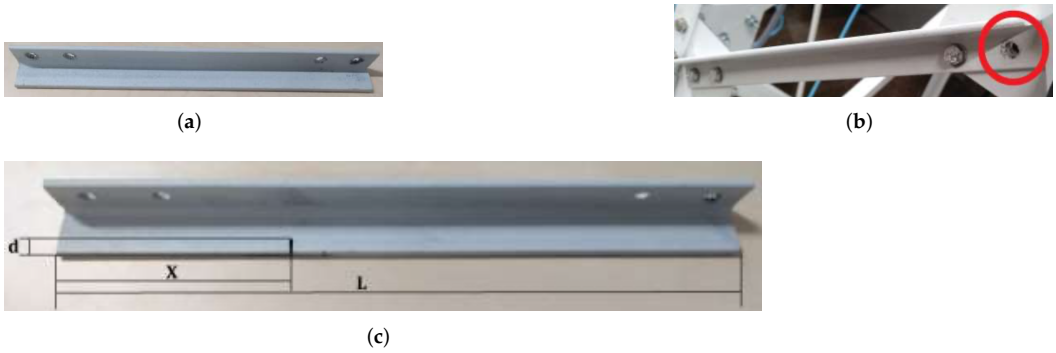



**Figure 3.** Different structural states studied: (**a**) replica bar (Healthy); (**b**) unlocked bolt; (**c**) details of the crack damage. L is the bar length, d = 5 mm is the crack size, and X = L/3 specifies the position of the crack.

The down-scaled laboratory structure is a simplified but valid model for the practical study of this work, which aims to be a proof of concept for the detection and diagnosis of damaged bars in jacket-type platforms. This is demonstrated by the fact that similar laboratory structures have been previously used in the literature for this aim, such as [8] and [12].

## 3. Methodology

The suggested methodology's stages are all listed below. First, the raw data from the sensors are collected. Second, an exploratory data analysis process is carried out to validate the hypothesis of this research. Then, the data are pre-processed to obtain a 24-channel image dataset. Subsequently, the data are reshaped and divided to be later entered into the first convolutional SNN. Then, the images classified as damaged are introduced to a second convolutional SNN to classify the damage between the crack or unlocked bolt types. Next, subsections comprehensively draw the above-mentioned different stages.

### *3.1. Data Acquisition*

Each experimental test lasts 60 s with an approximate sampling frequency of 275.27 Hz. As a result, 16,516 measurements were obtained from each of the 24 sensors (24 vibration signals). Twenty-five experimental tests were carried out for each of the white noise (WN) amplitudes (0.5, 1, 2, and 3), obtaining a total of 100 experiments. The experiments performed for each amplitude are detailed below:

- 10 tests with the original bar;
- 5 tests with the replica bar;
- 5 tests with a bar damaged by a 5 mm crack;
- 5 tests with an unlocked bolt damage.

Table 1 presents the number of experiments for each structural state and associated white noise amplitude factor.

**Table 1.** Number of experiments for each structural state and white noise (WN) amplitude factor.

| Label | Structural State | 0.5 WN | 1 WN | 2 WN | 3 WN |
|---|---|---|---|---|---|
| 1 | Healthy bar | 10 tests | 10 tests | 10 tests | 10 tests |
| 2 | Replica bar | 5 tests | 5 tests | 5 tests | 5 tests |
| 3 | Crack damaged bar | 5 tests | 5 tests | 5 tests | 5 tests |
| 4 | Unlocked bolt | 5 tests | 5 tests | 5 tests | 5 tests |

Table 2 shows the data obtained from each experimental test. The number of timestamps (16516) determine the number of rows, and the number of columns reflect the number of sensors. Take note that the data in the first column is connected to sensor A, the data in the second column is related to sensor B, and so on and so forth, until all 24 available sensors are covered.

**Table 2.** Data in each experimental test.

| **0** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | ... | **23** |
|---|---|---|---|---|---|---|---|---|---|
| A1 | B1 | C1 | D1 | E1 | F1 | G1 | H1 | ... | X1 |
| A2 | B2 | C2 | D2 | E2 | F2 | G2 | H2 | ... | X2 |
| A3 | B3 | C3 | D3 | E3 | F3 | G3 | H3 | ... | X3 |
| A4 | B4 | C4 | D4 | E4 | F4 | G4 | H4 | ... | X4 |
| A5 | B5 | C5 | D5 | E5 | F5 | G5 | H5 | ... | X5 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| A16516 | B16516 | C16516 | D16516 | E16516 | F16516 | G16516 | H16516 | ... | X16516 |

### *3.2. Exploratory Data Analysis*

Exploratory data analysis is essential to fully comprehend the information available from the data, and it is very useful when it comes to obtaining insights from the available data. The role of this process is to explore all the data to answer questions that help validate the hypotheses raised [20].

In this work, the following concern is stated: How does the distribution of each sensor signal, associated with a specific state, perform with different white noise amplitude factors? For this, after the data collection process is implemented, data visualization is created via a histogram.

As can be observed in Figure 4, there are four plots. On each of these plots, the statistical distributions for each state (healthy, replica, crack and unlocked bolt) are shown. On the upper left plot, where the white noise amplitude factor is small (0.5), it can be noted that the statistical distributions of the default state and the replica state (henceforth called healthy pair) are similar, as their centers tend to the plot's right side. Moreover, the statistical distributions of the crack state and the unlocked bolt state (henceforth called faulty pair) are similar to each other, but their centers tend to the plot's left side.

In the upper right of the plot, where the white noise amplitude factor is mid-low (1), it can be seen that the healthy pair and faulty pair statistical distributions still maintain a similarity to each other. However, the similarity differs among the pairs, where the centers of the healthy pair and the faulty pair are close to the right and left of the plot, respectively. At the lower left plot, where the white noise amplitude factor is mid-high (2), it can be observed that the pairs of distributions are almost similar to each other—however, with a slight level of difference. Finally, for the lower right plot, where the white noise amplitude factor is high (3), it can be noted that the distribution pairs, healthy and faulty, are almost indistinguishable due to the overlay behavior of each individual distribution. Taking into account that the different applied white noise amplitude factors represent different wind-speed regions of operation of the wind turbine, it may be stated that, at greater wind speeds (region where WTs intended to run the majority of the time), distinguishing whether a sample belongs to a given structural state becomes more challenging.

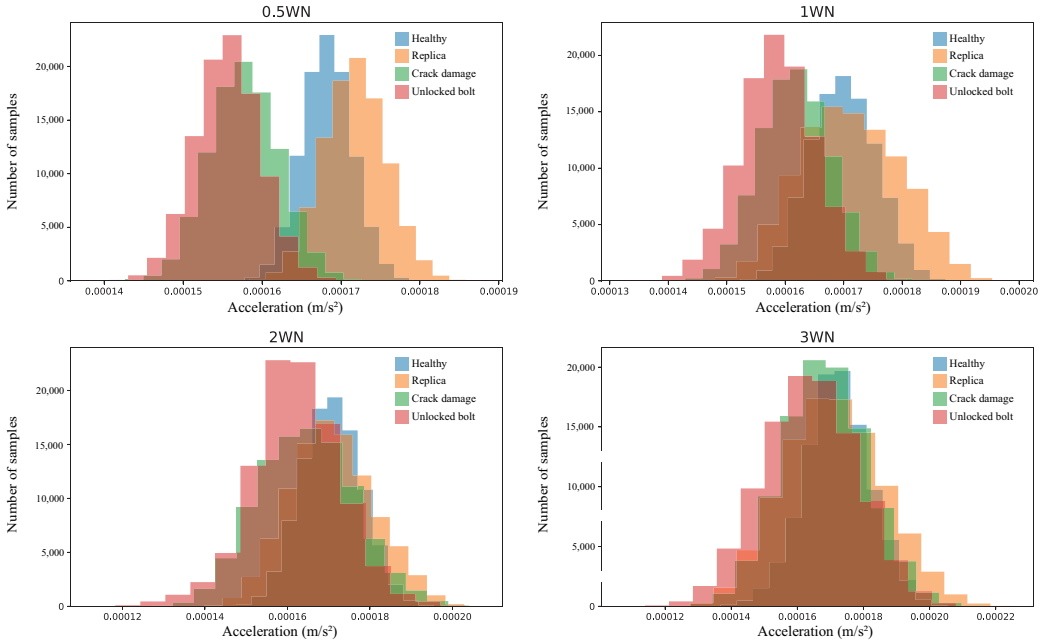



**Figure 4.** Histograms for each structural state (blue stands for the healthy state, orange for the replica bar, green represents the crack damage, and pink is related to the unlocked bolt state) and for each different white noise (WN) amplitude factor.

### *3.3. Data Preprocessing: Reshape*

In this section, a feature engineering technique, data reshaping, is applied to ensure that each one of the samples to be processed by the SNNs have sufficient information from each sensor to determine the state of the structure. Initially, each experiment had 24 columns and 16,516 rows, representing the source of the data (sensors) and the data acquired over time, respectively (see Table 2).

In this study, images (samples) that contain the information of approximately one second of data are created. Recall that the sample rate is 275.27 Hz. Thus, the first 256 values from each column (approximately one second of data) were reshaped into 16 × 16 matrices. Then, the next 256 values from each column were reshaped into 16 × 16 matrices, and so on, as can be seen in Figures 5 and 6. At the end, the final target images' shapes were 16 × 16 for each sensor. Note that, in Figures 5 and 6, the values A*i*, B*i*, ⋯, X*i* are measurements of the different sensors that correspond to the same time step—that is, acquired at the same time instant. In other words, in general, A*i*, B*i*, . . . , and X*i* all correspond to the same time step *i*.

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | ... | 23 |
|---|---|---|---|---|---|---|---|---|---|---|
| 256 values | A1 | B1 | C1 | D1 | E1 | F1 | G1 | H1 | ... | X1 |
| | A2 | B2 | C2 | D2 | E2 | F2 | G2 | H2 | ... | X2 |
| | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| | A256 | B256 | C256 | D256 | E256 | F256 | G256 | H256 | ... | X256 |
| 256 values | A257 | B257 | C257 | D257 | E257 | F257 | G257 | H257 | ... | X257 |
| | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| | A511 | B512 | C512 | D512 | E512 | F512 | G512 | H512 | ... | X512 |
| | A512 | B513 | C513 | C513 | E513 | F513 | G513 | H513 | ... | X513 |
| | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| | A16516 | B16516 | C16516 | D16516 | E16516 | F16516 | G16516 | H16516 | ... | X16516 |

**Figure 5.** Selection of 256 sequential values for each row.

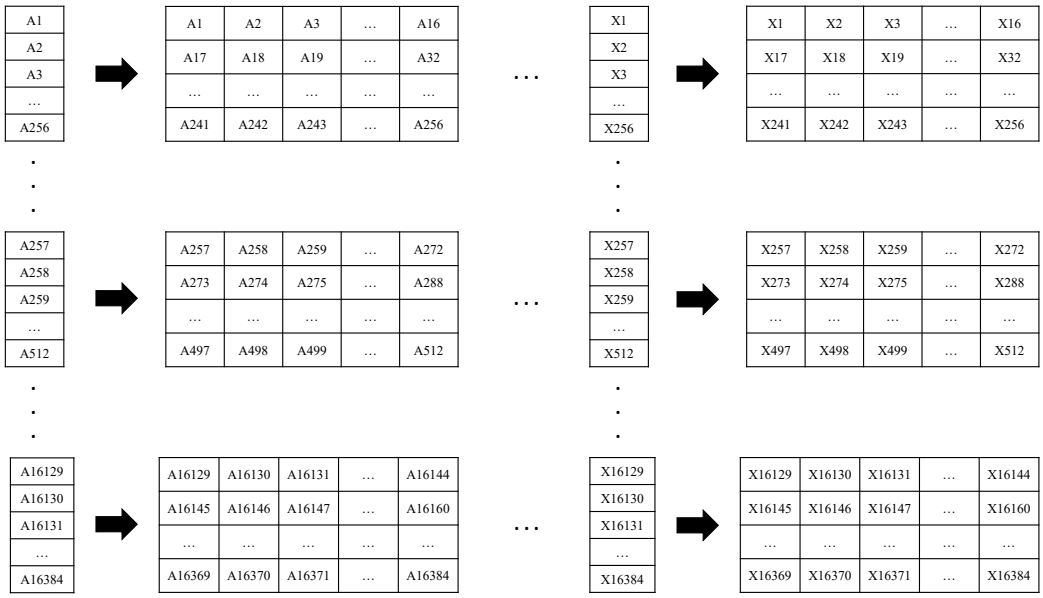



**Figure 6.** Reshape of the 256 values into 16 × 16 size matrices.

The last 132 values from each column were not considered because they could not complete the 256 values required to build a (16 × 16) matrix. Once the two-dimensional matrices were obtained, the process was continued with the creation of the 16 × 16 × 24 size images. For each experiment, the first matrices formed by the 256 first values of their respective columns were time-related. Meaning that each value from a specific position on each matrix was sampled at the same time as the values occupying the same position on the rest of the 23 matrices. Basically, the first 24 matrices were grouped together to maintain this relation provided by this new feature. The same was applied to the second group of matrices, and so on and so forth, until 64 groups of matrices were obtained from each experiment (see Figure 7).

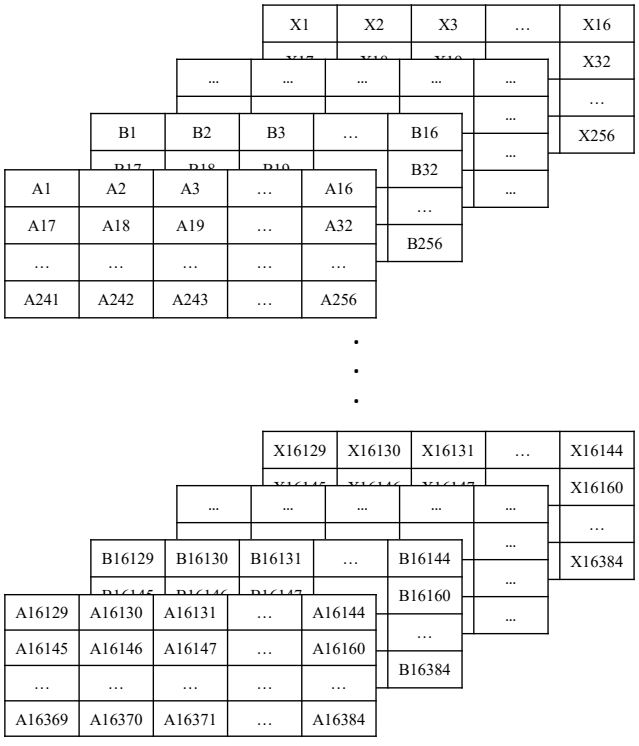



**Figure 7.** Shape of the images obtained from experiments.

At the end of the process, 1280 images were obtained from each of the replica bar, cracked, and unlocked bolt bar experiments, while 2560 images were obtained from the healthy bar experiments.

### *3.4. Data Split: Train, Validation, and Test Sets*

For the data split process, 80 percent of the data was considered for the training set, 10 percent of the data for the validation set, and the remaining 10 percent was put on hold to be used only for testing. This applies to both the damage detection and the damage diagnosis SNN models. For the damage detection SNN, the data were grouped in such a way that the healthy and the replica structure state images were grouped into the first class (healthy), while the crack and the unlocked bolt structure state images were grouped into the second class (faulty). On the other hand, for the damage diagnosis SNN, the data were grouped into following two classes: crack structural state and unlocked bolt structural state.

Once the images were in their assigned classes, various subsets were created in order to separate the images, as per the structural state and white noise level. After this process, the following subsets were obtained per state: white noise level of 0.5, white noise level of 1, white noise level of 2, and white noise level of 3. Finally, recall that 80, 10, and the remaining 10 percent of each one of the obtained subsets are used in the training, validation, and test set, respectively. Thus, a data balance is ensured per structural state and white noise level, as can be seen in Tables 3 and 4.

**Table 3.** Number of images used in the train, validation, and test sets per white noise (WN) amplitude factor and per studied class for damage detection.

| | | **0.5 WN** | **1 WN** | **2 WN** | **3 WN** |
|---|---|---|---|---|---|
| Train | Healthy & Replica | 768 | 768 | 768 | 768 |
| | Crack & Unlocked Bolt | 512 | 512 | 512 | 512 |
| Validation | Healthy & Replica | 96 | 96 | 96 | 96 |
| | Crack & Unlocked Bolt | 64 | 64 | 64 | 64 |
| Test | Healthy & Replica | 96 | 96 | 96 | 96 |
| | Crack & Unlocked Bolt | 64 | 64 | 64 | 64 |

**Table 4.** Number of images used in the train, validation, and test sets per white noise (WN) amplitude factor and per studied class for damage diagnosis.

| | | **0.5 WN** | **1 WN** | **2 WN** | **3 WN** |
|---|---|---|---|---|---|
| Train | Crack | 256 | 256 | 256 | 256 |
| | Unlocked Bolt | 256 | 256 | 256 | 256 |
| Validation | Crack | 32 | 32 | 32 | 32 |
| | Unlocked Bolt | 32 | 32 | 32 | 32 |
| Test | Crack | 32 | 32 | 32 | 32 |
| | Unlocked Bolt | 32 | 32 | 32 | 32 |

### *3.5. Siamese Neural Network (SNN)*

The SNN algorithm was developed by Bromley et al. [21] in 1994 to verify signatures written on a touch-sensitive pad. The SNNs consist of two identical neural network architectures capable of learning and extracting the hidden representation of their respective inputs [18].

In this work, the two neural networks are both convolutional neural networks [22] and employ back-propagation during training [23]. The basic idea of this methodology is that both networks work in parallel and finally compare their outputs. The comparison function

is the Euclidean distance. The output generated by an SNN execution can be considered the semantic similarity between the projected representation of the two input matrices [24].

Two models were deployed for the damage detection and diagnosis problems to compare their performance. The first model (model 1) consisted of an SNN with a feature extraction stage of one convolutional layer (see Figure 8), while the second model (model 2) implemented two convolutional layers in its feature extraction stage (see Figure 9). The rest of this section details each step in the methodology:

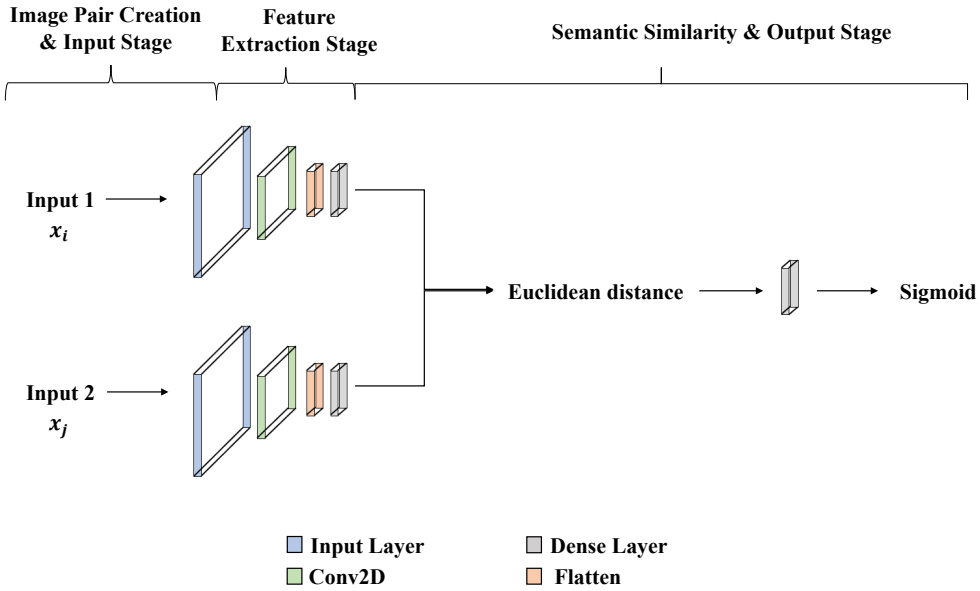



**Figure 8.** One convolutional layer SNN's architecture.

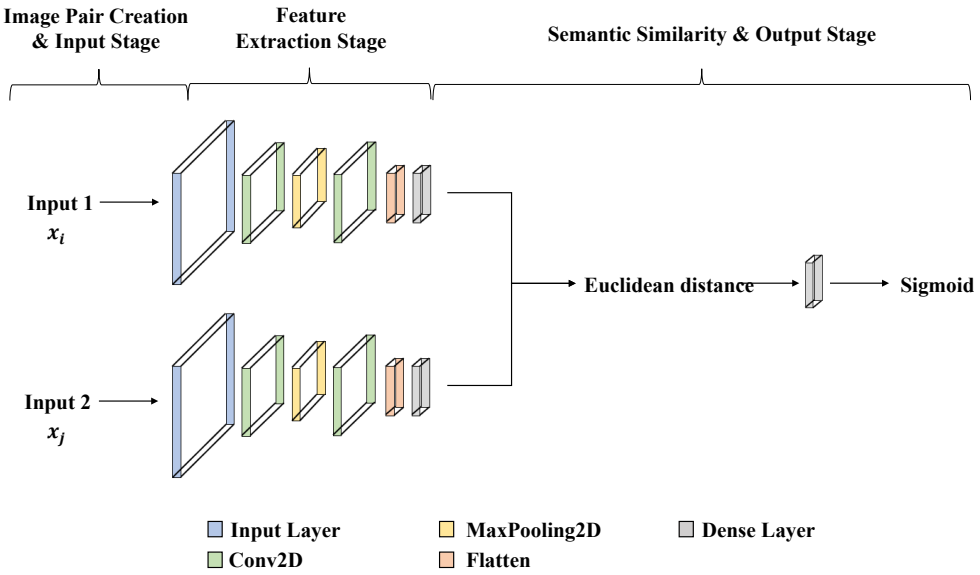



**Figure 9.** Two convolutional layer SNN's architecture.

1. **Image pair creation Stage:** The SNN has a pair of images input, where this pair can be positive or negative. A positive pair consists of two images that belong to the same class, while a negative pair consists of two images from different classes [25]. For the training, validation, and test sets, positive and negative pairs were created. For explanation, the two classes are noted as class A and class B. First, an empty array for the pair of images is set as well as an empty array of labels, which help to indicate by index if a pair of images (from the pair of images array) is positive (1) or

negative (0). Next, an iterative process is performed through the set of images that belong to class A. For each image, a random image is selected from the same class. Next, a positive pair is created by the image that is being iterated, and the random image selected from the same class (this pair is added to the image pair array). Then, a label with the value of one is added to the labels array previously created, so the image pair added, and the label are related by position through both of the arrays. This process can be observed in Figure 10.

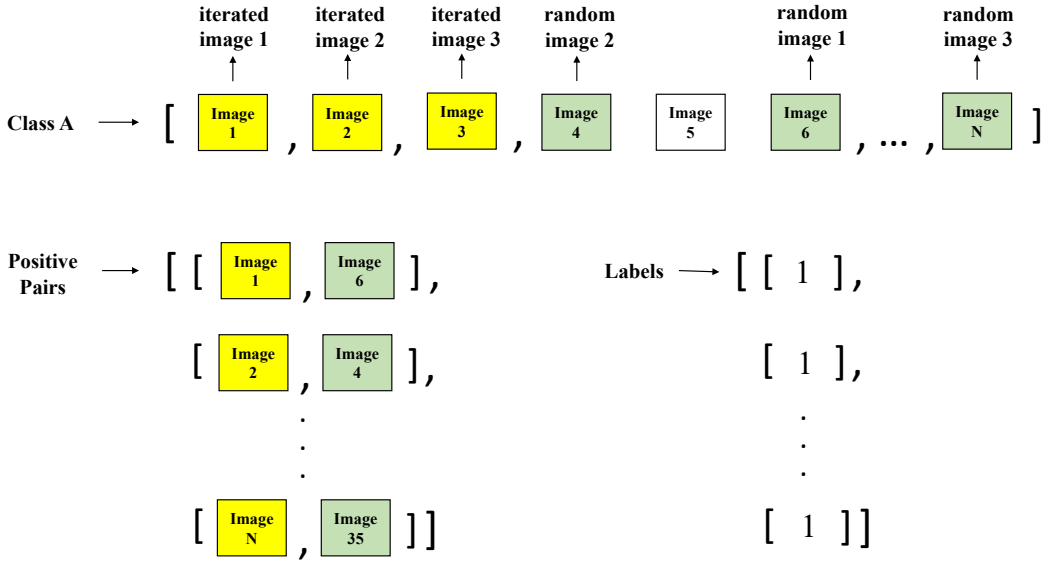



**Figure 10.** Positive pairs' creation process.

For the creation of negative pairs, a similar process is carried out, but the random image is chosen from class B. Then, a label with the value of 0 is added to the labels array previously created in order to maintain the index relation between the pair and the label. This iterative process is carried out also for class B (the negative image comes from class A and positive image comes from class B), so it can be ensured that all the images are used. This process can be observed in Figure 11.

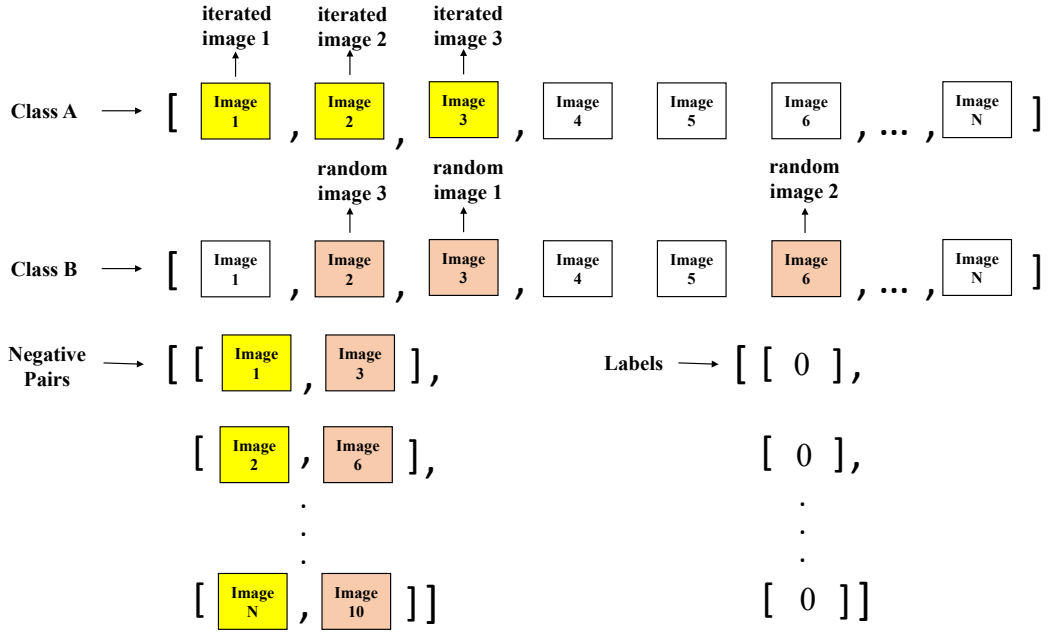



**Figure 11.** Negative pairs' creation process.

If the sets of images are not of the same length, then the difference is resolved by iterating a quantity of images equal to this difference on the lower length set. At the end, a set $X$ is composed by $N$ sample pairs $(x_i, x_j)$, which are two images from the same or different classes. Remember, for both creations' processes (positive and negative), both states' groups in each case are used. For example, for the damage detection stage, first, the healthy and replica images are used as class A and positive pairs are created. Then, the other group class (crack and unlocked bolt) is used as class A, and the other positive pairs are created. The same process is used for the negative pairs creation and the damage diagnosis stage. Tables 5 and 6 detail the number of image pairs used for training, validation, and test datasets.

**Table 5.** Number of pairs of images used for training, validation, and test in the damage detection model.

| | Training | Validation | Testing |
|---|---|---|---|
| Positive pair images | 6144 | 768 | 768 |
| Negative pair images | 6144 | 768 | 768 |

**Table 6.** Number of pairs of images used for training, validation, and test in the damage diagnosis model.

| | Training | Validation | Testing |
|---|---|---|---|
| Positive pair images | 2048 | 256 | 256 |
| Negative pair images | 2048 | 256 | 256 |

2. **Input Stage:** In the proposed methodology, two two-dimensional CNNs are used to extract hidden representation (spatial feature vectors), so the input is the matrix mentioned in Section 3.3, in the shape of $16 \times 16 \times 24$.
3. **Feature extraction:** A Siamese network architecture is employed at this stage to extract features from the input sample pairs. The Siamese network is made up of two identical CNNs with the same network topology and one fully connected layer at the end. Tables 7 and 8 indicate the different feature extraction layers, their configurations, and their dimensions for the two studied models noted as model 1 (shallow NN with only one CNN layer) and model 2 (NN with two CNN layers).

**Table 7.** One convolutional layer SNN (model 1).

| Layer | Kernel Size | Stride | Padding | Filters | Output Size |
|---|---|---|---|---|---|
| Convolutional 2D | 3 | 1 | 1 | 64 | $16 \times 16 \times 64$ |
| Flatten | - | - | - | - | $1 \times 16{,}384$ |
| Fully connected | 4960 | - | - | 1 | $1 \times 4960$ |

**Table 8.** Two convolutional layers SNN (model 2).

| Layer | Kernel Size | Stride | Padding | Filters | Output Size |
|---|---|---|---|---|---|
| Convolutional 2D | 3 | 1 | 1 | 64 | $16 \times 16 \times 64$ |
| MaxPooling 2D | - | - | - | - | $8 \times 8 \times 64$ |
| Convolutional 2D | 3 | 1 | 1 | 128 | $8 \times 8 \times 128$ |
| Flatten | - | - | - | - | $1 \times 8192$ |
| Fully connected | 4960 | - | - | 1 | $1 \times 4960$ |

4. **Similarity measurement:** The output vectors obtained from both fully connected layers are introduced to a new function layer to compute the similarity (distance) between them. This process can be carried out by metrics such as Euclidean distance, cosine distance, or Manhattan distance [26] because it highlights the geometric differences

between two elements. In this work, the Euclidean distance is used as the similarity matrix. The similarity between the input vectors is calculated by the formula

$$S_{i,j}^k = \sqrt{F(x_i^k) - F\left(x_j^k\right)}, \tag{1}$$

where $F(x)$ represents the feature vector obtained by one of the CNNs, $x$ refers to the input, and $k$ denotes the $k$-th sample for a pair $(x_i, x_j)$.

5. **Output stage:** After the similarity $S_{i,j}^k$ is calculated, this value enters a last fully connected layer to convert it to a similarity scalar value $O_{i,j}^k$. Because the idea is to calculate a similarity probability between 0 and 1, a sigmoid function is used as activation function

$$P_{i,j}^k = \sigma\left(O_{i,j}^k\right), \tag{2}$$

where $\sigma$ is the sigmoid function [27]

$$\sigma(x) = \frac{1}{1+e^{-x}}. \tag{3}$$

Note that $P_{i,j}^k$ is a value between 0 and 1. The closer this value is to 1, the greater the probability that the two matrices are of the same class. Likewise, the closer this value is to zero, the lower the probability that the two matrices are of the same class.

6. **Hyperparameters:** The SNNs are configured with the following hyperparameters' selection. The Adams optimizer with learning rate 0.05, $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 10^{-7}$ is used. The used cost function is the binary cross entropy, and the batch size is set to 32. Hyperparameter tuning did not change the obtained results much, except for the value of the learning rate, where lower values improved the accuracy.

## 4. Results

To recognize whether a model is overfitting, the loss curves of the training set and the validation set are first presented. Overfitting implies that the model is too closely aligned with a limited set of data points (training data) [28], thus reducing its predictive power. In Figure 12 (left), it can be observed that, for the damage detection stage, model 1 is overfitting (from epoch 2, the validation loss starts to increase while the training loss continues to decrease). On the other hand, Figure 12 (right) shows that model 2 has an appropriate fitting.

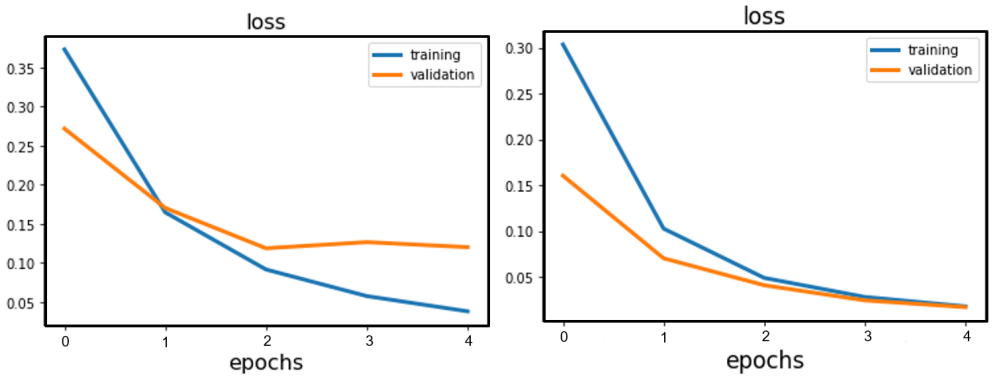



**Figure 12.** Loss curves for damage detection models 1 (**left**) and 2 (**right**).

The same performance is observed in the damage diagnosis stage. As can be seen in Figure 13, model 1 (left) is overfitting, while model 2 (right) is able to adapt properly to previously unseen data.

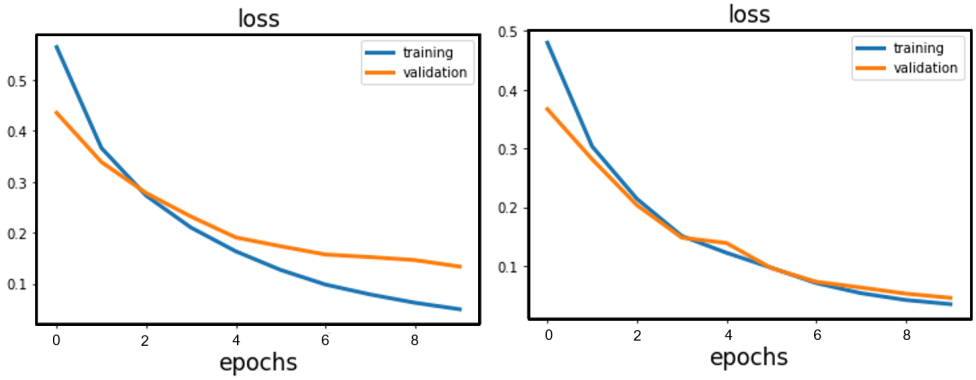



**Figure 13.** Loss curves for damage diagnosis models 1 (**left**) and 2 (**right**).

Additionally, to measure the performance of detection and diagnostic models, the results of a confusion matrix can be used to determine the accuracy, precision, recall, and F1 score of the predictions made on the test data set [29]. As can be seen in Figure 14, a confusion matrix is an array that shows the predictions of the true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) made by a classification model [30]. Note that, in this work, the output of the model is a probability of similarity between two images. In other words, the algorithm states whether the two samples are similar (labeled as 0) or not (labeled as 1). For the binary problem at hand, a TP occurs when both samples in the pair (images) are similar (positive pair), and the algorithm predicts accordingly. A TN occurs when both images in the pair are not similar (negative pair) and the algorithm predicts correctly. An FP results when the samples are similar, but the algorithm predicts that they are not similar. Finally, an FN occurs when both samples are not similar, but the algorithm predicts the opposite. The aforementioned metrics can be obtained from Equations (4)–(7).

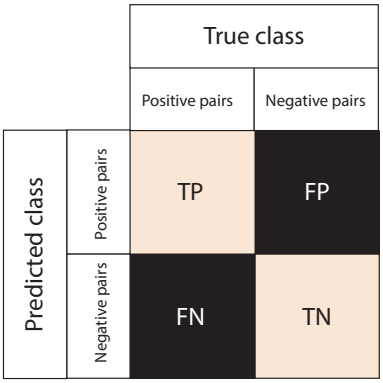



**Figure 14.** Confusion matrix.

- Accuracy: proportion of true results among the total number of results:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}. \tag{4}$$

- Precision: positive predictive value:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}. \tag{5}$$

- Recall: proportion of true positive predictions made out of all positive predictions that could have been made:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \tag{6}$$

- F1-score: harmonic mean of precision and recall:

$$\mathrm{F1} = 2 \cdot \frac{\mathrm{Precision} \cdot \mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}}. \tag{7}$$

The confusion matrices obtained from the implementation of the two models, model 1 and model 2, in both stages (damage detection and damage diagnosis) are shown in Figures 15 and 16. For the first model, the failed predictions consist of 18 FPs and 3 FNs for the detection stage, and 13 FPs and 3 FNs for the diagnosis stage. As it can be observed, adding an extra convolutional layer to the SNN feature extraction stage can improve the performance of the model in both stages, since the confusion matrices show neither FPs nor FNs. Figures 17 and 18 demonstrate the same confusion matrices, but using a 70%, 15%, and 15% data split (for the training, validation, and testing sets, respectively). The results are shown to be comparable to those obtained with the 80%, 10%, and 10% data split.

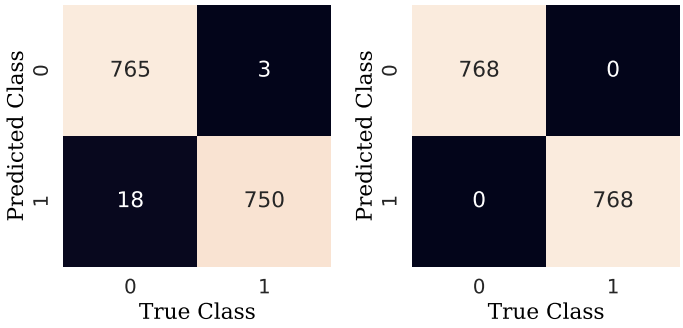



**Figure 15.** Confusion matrices for damage detection models 1 (**left**) and 2 (**right**).

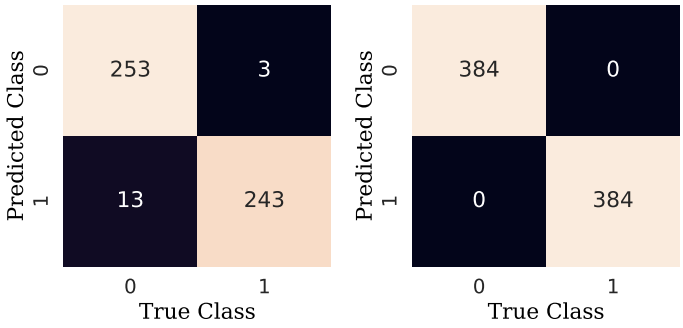



**Figure 16.** Confusion matrices for damage diagnosis models 1 (**left**) and 2 (**right**).

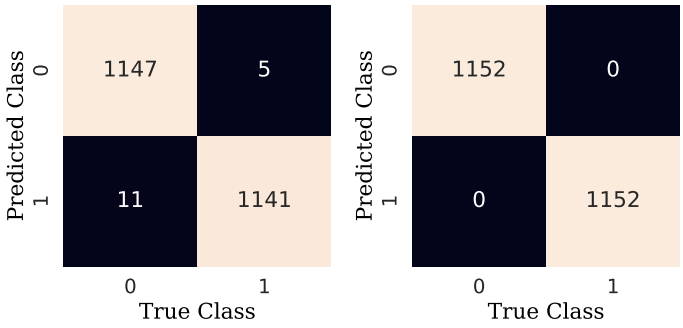



**Figure 17.** Confusion matrices for damage detection models 1 (**left**) and 2 (**right**) when using a data split of 70% training, 15% validation, and 15% for testing.

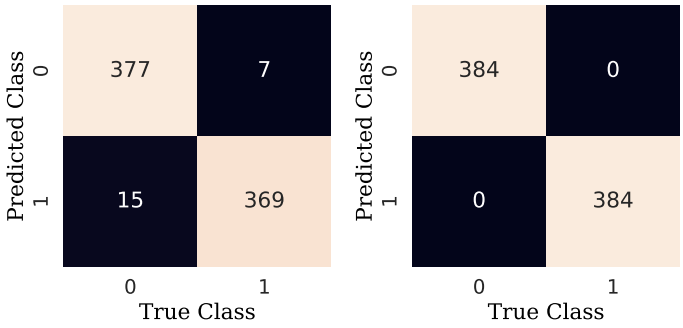



**Figure 18.** Confusion matrices for damage diagnosis models 1 (**left**) and 2 (**right**) when using a data split of 70% training, 15% validation, and 15% for testing.

A follow-up was conducted on the performance of model 1, according to the different levels of WN amplitude factors (0.5, 1, 2, 3) for the initial 80%, 10%, and 10% data split. The purpose was to find how many failed predictions are made for each of the different WN amplitudes. Figure 19 (left) shows the results for the damage detection stage. Specifically, it can be seen that, for the case of WN with amplitude factor 0.5, there are no wrong predictions. Similarly, for the WN amplitude factor 1, there is only one incorrect prediction that is at least identified as belonging to amplitude 1. In the case of WN amplitude factor 2, there are seven incorrect predictions that are assigned as similar to amplitude factor 0.5. Finally, for WN with factor 3, there are 10 incorrect predictions that map to WN with factor 0.5, and two incorrect predictions where the model assimilates them to samples in WN factor 1. Figure 19 (right) shows similar results but for the damage diagnosis stage. These results can be summarized in the following manner. When an image comes from a WT operating at higher wind speeds (simulated in the experimental tower with a higher white noise amplitude factor), the model has far more failed predictions, which is in good agreement with the insight obtained in the exploratory data analysis performed in Section 3.2.

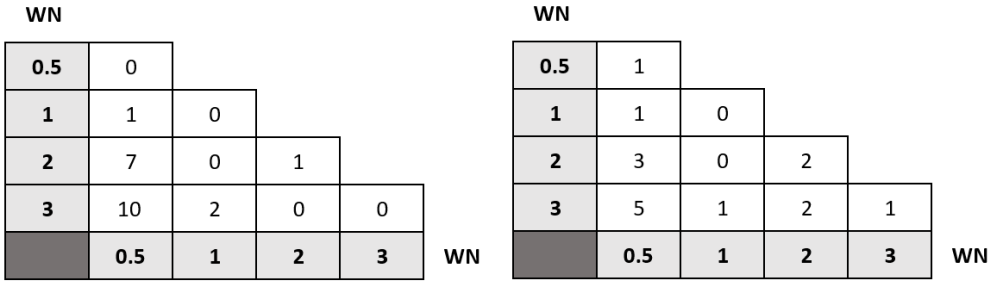



**Figure 19.** Damage detection (**left**) and damage diagnosis (**right**) number of failed predictions per WN combination for model 1. Rows are related to the true amplitude factor and columns to the predicted amplitude factor.

In Tables 9–12, a deeper exploration of the results is shown to gain an insight on how many FP or FN outputs are obtained, according to the structural state, among the images by model 1. It is shown that the predictions given by model 1 (one convolutional layer) for both cases, damage detection and damage diagnosis, output a higher number of FNs than FPs. Furthermore, model 2 (two convolutional layers) outperformed model 1 with no false predictions.

**Table 9.** False positives for damage detection (model 1).

| Class 1 | Class 2 | False Positives |
|---|---|---|
| healthy | unlocked bolt | 1 |
| healthy | crack | 2 |

**Table 10.** False negatives for damage detection (model 1).

| Class 1 | Class 2 | False Negatives |
|---|---|---|
| healthy | replica | 2 |
| healthy | healthy | 4 |
| crack | unlocked bolt | 4 |
| crack | crack | 5 |
| unlocked bolt | unlocked bolt | 3 |

**Table 11.** False positives for damage diagnosis (model 1).

| Class 1 | Class 2 | False Positives |
|---|---|---|
| crack | unlocked bolt | 3 |

**Table 12.** False negatives for damage diagnosis (model 1).

| Class 1 | Class 2 | False Negatives |
|---|---|---|
| crack | crack | 10 |
| unlocked bolt | loose bolt | 3 |

Eventually, the accuracy, precision, recall, and F1 score for both stages and models are detailed in Tables 13 and 14. The results show that the damage detection and damage diagnosis results are promising, as they achieved great performance on different structural state samples.

**Table 13.** Performance of damage detection models.

| Models | Accuracy | Precision | Recall | F1 Score |
|---|---|---|---|---|
| 1 conv | 98.63 | 97.66 | 99.60 | 98.62 |
| 2 conv | 100.00 | 100.00 | 100.00 | 100.00 |

**Table 14.** Performance of damage diagnosis models.

| Models | Accuracy | Precision | Recall | F1 Score |
|---|---|---|---|---|
| 1 conv | 96.88 | 94.92 | 98.78 | 96.81 |
| 2 conv | 100.00 | 100.00 | 100.00 | 100.00 |

Finally, to thoroughly test the functional characteristics of the algorithm, a comparison is made with four other methodologies given in [31], [32], [13], and [12] that use the same laboratory structure. The first methodology, given in [31], is based on principal component analysis and support vector machines. The second methodology, given in [32] (page 67), is based on the well-known damage indicators: covariance matrix estimate and scalar covariance. The third methodology, given in [13], is based on machine learning methods and the fractal dimension feature. The last methodology, given in [12], utilizes a signal-to-image conversion of the accelerometer data into multichannel images and convolutional neural networks (CNN), combined with synthetic data augmentation. First, when using the first approach stated in [31], the crack damaged bar has a recall of 96.08% and is therefore inferior to the one obtained with the strategy proposed in this work, which reached a value of 100%. Note that the crack damage is the most challenging. In fact, the second approach stated in [32] (page 82) was unable to detect this type of incipient damage when using scalar covariance or mean residual damage indicators. Furthermore, the first approach obtains a recall of 99.02% for the unlocked bold damage, while, with the proposed strategy, a slightly higher value of 100% is obtained. Note that the unlocked bold damage is not studied in the second approach. The third approach [13] requires hand-made feature extraction,

and the performance metrics obtained are inferior to those obtained in the present study. Furthermore, the machine learning methods proposed in [13] need a large data set to achieve good performance, while Siamese neural networks have the ability to learn from very little data, which is crucial in the specific application faced in this work. Finally, the fourth approach [12] requires a deep CNN as well as a data augmentation of 25,200% in the total number of samples to achieve 99% accuracy, while, with the proposed strategy, a better accuracy is obtained using much fewer data and a much simpler neural network architecture.

## 5. Conclusions

In this work, the proposed test bench consists of a WN generator, connected to an amplifier that simulates different wind speed regions of the operation of the WT. Subsequently, triaxial accelerometers are connected to obtain the vibration signals. Different simulations were carried out, taking into account four types of structural states, such as the healthy bar, the replica bar, the crack damaged bar, and the unlocked bolt. It was concluded that, when wind speeds are higher (regions where turbines operate or are desired to operate most of the time), it is more challenging to distinguish when a sample belongs to a specific state. The main contribution of this work is informing the use of SNNs for the damage detection and damage diagnosis stages. Increasing the number of convolutional layers to extract features from the data increased the performance of the model. The conceived SHM methodology with two convolutional layers showed exceptional performance, demonstrating results of 100% for all considered metrics. These findings indicate that SNNs are promising for developing SHM techniques for offshore platforms.

Note that this study is a proof-of-concept contribution, as the data were obtained in a controlled laboratory environment. Therefore, as future work, it is proposed to incorporate other environmental conditions, such as the wave excitation, by placing the experiment in a water tank facility to simulate the effect of regular and irregular waves. Finally, it is important to note that environmental and operational conditions (EOC) play an important role when dealing with long-term monitoring because they can complicate damage detection. Large variations in EOCs make EOC monitoring almost as important as structural monitoring itself. Therefore, its influence should be compensated. Several methods for EOC compensation for WTs have been developed to make SHM possible. For example, in [33], affinity propagation clustering is used to delineate data into WT groups of similar EOC. In [34], covariance-driven stochastic subspace identification is used. Finally, in [35,36], fuzzy classification techniques are used for EOC compensation. However, as noted previously, this work is an experimental proof of concept, and EOC compensation is left as future work using pattern recognition techniques in a more realistic environment.

**Author Contributions:** Conceptualization, Y.V. and C.T.; methodology, J.B.; software, J.B.; validation, B.P. and J.S.; formal analysis, Y.V.; investigation, J.B. and B.P.; resources, Y.V.; data curation, J.B.; writing—original draft preparation, J.B. and C.T.; writing—review and editing, Y.V.; visualization, J.S.; supervision, Y.V. and C.T.; project administration, C.T.; funding acquisition, Y.V. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work has been partially funded by the Spanish Agencia Estatal de Investigación (AEI)—Ministerio de Economía, Industria y Competitividad (MINECO), and the Fondo Europeo de Desarrollo Regional (FEDER) through research project DPI2017-82930-C2-1-R, and by the Generalitat de Catalunya through research project 2017 SGR 388.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this work are available on request from the corresponding author.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

| | |
|---|---|
| SHM | Structural Health Monitoring |
| SNN | Siamese Neural Network |
| WN | White Noise |
| TP | True Positive |
| TN | True Negative |
| FP | False Positive |
| FN | False Negative |

## References


1. Fraile, D.; Vandenberghe, A.; Konari, V.; Ramirez, L.; Pineda, I.; Tardieu, P.; Malvault, B.; Komusanac, I. *Getting Fit for 55 and Set for 2050: Electrifying Europe with Wind Energy*; Technical Report; WindEurope: Brussels, Belgium, 2021.
2. Martinez-Luengo, M.; Shafiee, M. Guidelines and cost-benefit analysis of the structural health monitoring implementation in offshore wind turbine support structures. *Energies* **2019**, *12*, 1176. [CrossRef]
3. An, Y.; Chatzi, E.; Sim, S.H.; Laflamme, S.; Blachowski, B.; Ou, J. Recent progress and future trends on damage identification methods for bridge structures. *Struct. Control Health Monit.* **2019**, *26*, e2416. [CrossRef]
4. Bi, J.; Yuan, H.; Zhou, M. Temporal prediction of multiapplication consolidated workloads in distributed clouds. *IEEE Trans. Autom. Sci. Eng.* **2019**, *16*, 1763–1773. [CrossRef]
5. Van Kuik, G.; Peinke, J.; Nijssen, R.; Lekou, D.; Mann, J.; Sørensen, J.N.; Ferreira, C.; van Wingerden, J.W.; Schlipf, D.; Gebraad, P.; et al. Long-term research challenges in wind energy—A research agenda by the European Academy of Wind Energy. *Wind. Energy Sci.* **2016**, *1*, 1–39. [CrossRef]
6. Gianneo, A.; Carboni, M.; Giglio, M. Feasibility study of a multi-parameter probability of detection formulation for a Lamb waves–based structural health monitoring approach to light alloy aeronautical plates. *Struct. Health Monit.* **2017**, *16*, 225–249. [CrossRef]
7. Agathos, K.; Tatsis, K.E.; Vlachas, K.; Chatzi, E. Parametric reduced order models for output-only vibration-based crack detection in shell structures. *Mech. Syst. Signal Process.* **2022**, *162*, 108051. [CrossRef]
8. Spanos, N.A.; Sakellariou, J.S.; Fassois, S.D. Vibration-response-only statistical time series structural health monitoring methods: A comprehensive assessment via a scale jacket structure. *Struct. Health Monit.* **2020**, *19*, 736–750. [CrossRef]
9. Kim, H.C.; Kim, M.H.; Choe, D.E. Structural health monitoring of towers and blades for floating offshore wind turbines using operational modal analysis and modal properties with numerical-sensor signals. *Ocean. Eng.* **2019**, *188*, 106226. [CrossRef]
10. Mieloszyk, M.; Ostachowicz, W. An application of Structural Health Monitoring system based on FBG sensors to offshore wind turbine support structure model. *Mar. Struct.* **2017**, *51*, 65–86. [CrossRef]
11. Liu, F.; Gao, S.; Tian, Z.; Liu, D. A new time-frequency analysis method based on single mode function decomposition for offshore wind turbines. *Mar. Struct.* **2020**, *72*, 102782. [CrossRef]
12. Puruncajas, B.; Vidal, Y.; Tutivén, C. Vibration-Response-Only Structural Health Monitoring for Offshore Wind Turbine Jacket Foundations via Convolutional Neural Networks. *Sensors* **2020**, *20*, 3429. [CrossRef] [PubMed]
13. Hoxha, E.; Vidal, Y.; Pozo, F. Damage Diagnosis for Offshore Wind Turbine Foundations Based on the Fractal Dimension. *Appl. Sci.* **2020**, *10*, 6972. [CrossRef]
14. Leon-Medina, J.X.; Anaya, M.; Parés, N.; Tibaduiza, D.A.; Pozo, F. Structural damage classification in a Jacket-type wind-turbine foundation using principal component analysis and extreme gradient boosting. *Sensors* **2021**, *21*, 2748. [CrossRef] [PubMed]
15. Meldo, A.; Utkin, L. A new approach to differential lung diagnosis with ct scans based on the siamese neural network. *J. Phys. Conf. Ser.* **2019**, *1236*, 012058. [CrossRef]
16. Zhou, X.; Liang, W.; Shimizu, S.; Ma, J.; Jin, Q. Siamese neural network based few-shot learning for anomaly detection in industrial cyber-physical systems. *IEEE Trans. Ind. Inform.* **2020**, *17*, 5790–5798. [CrossRef]
17. Zhang, C.; Liu, W.; Ma, H.; Fu, H. Siamese neural network based gait recognition for human identification. In Proceedings of the 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Shanghai, China, 20–25 March 2016; pp. 2832–2836.
18. Chicco, D. Siamese neural networks: An overview. In *Artificial Neural Networks*; Humana: New York, NY, USA, 2021; pp. 73–94.
19. Ruiz, M.; Mujica, L.E.; Alferez, S.; Acho, L.; Tutiven, C.; Vidal, Y.; Rodellar, J.; Pozo, F. Wind turbine fault detection and classification by means of image texture analysis. *Mech. Syst. Signal Process.* **2018**, *107*, 149–167. [CrossRef]
20. Jebb, A.; Parrigon, S.; Woo, S.E. Exploratory data analysis as a foundation of inductive research. *Hum. Resour. Manag. Rev.* **2016**, *27*, 265–276. [CrossRef]
21. Bromley, J.; Bentz, J.W.; Bottou, L.; Guyon, I.; LeCun, Y.; Moore, C.; Säckinger, E.; Shah, R. Signature verification using a "siamese" time delay neural network. *Int. J. Pattern Recognit. Artif. Intell.* **1993**, *7*, 669–688. [CrossRef]
22. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.* **2012**, *25*, 1097–1105. [CrossRef]

23. Rumelhart, D.E.; Durbin, R.; Golden, R.; Chauvin, Y. Backpropagation: The basic theory. In *Backpropagation: Theory, Architectures and Applications*; Lawrence Erlbaum Associates, Inc.: Mahwah, NJ, USA, 1995; pp. 1–34.
24. Koch, G.; Zemel, R.; Salakhutdinov, R. Siamese Neural Networks for One-shot Image Recognition. In Proceedings of the 32nd International Conference on Machine Learning, Lille, France, 6–11 July 2015; Volume 37, p. 20.
25. Melekhov, I.; Kannala, J.; Rahtu, E. Siamese network features for image matching. In Proceedings of the 2016 23rd International Conference on Pattern Recognition (ICPR), Cancun, Mexico, 4–8 December 2016; pp. 378–383.
26. Homayouni, R.; Heinrich, K.; Wei, L.; Berry, M.W. Gene clustering by latent semantic indexing of MEDLINE abstracts. *Bioinformatics* **2005**, *21*, 104–115. [CrossRef]
27. Han, J.; Moraga, C. The influence of the sigmoid function parameters on the speed of backpropagation learning. In Proceedings of the International Workshop on Artificial Neural Networks, Malaga-Torremolinos, Spain, 7–9 June 1995; Springer: Berlin/Heidelberg, Germany, 1995; pp. 195–201.
28. Hawkins, D.M. The problem of overfitting. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 1–12. [CrossRef] [PubMed]
29. Forman, G. An extensive empirical study of feature selection metrics for text classification. *J. Mach. Learn. Res.* **2003**, *3*, 1289–1305.
30. Kang, M.; Jordan Jameson, N. Machine Learning: Fundamentals. In *Prognosis and Health Management of Electronics: Fundamentals, Machine Learning and the Internet of Things*; John Wiley & Sons: Hoboken, NJ, USA, 2018; p. 20.
31. Vidal Seguí, Y.; Rubias, J.L.; Pozo Montero, F. Wind turbine health monitoring based on accelerometer data. In Proceedings of the 9th ECCOMAS Thematic Conference on Smart Structures and Materials, Paris, France, 8–11 July 2019; pp. 1604–1611.
32. Zugasti Uriguen, E. Design and Validation of a Methodology for Wind Energy Structures Health Monitoring. Ph.D. Thesis, Universitat Politècnica de Catalunya, Barcelona, Spain, 2014.
33. Häckell, M.W.; Rolfes, R.; Kane, M.B.; Lynch, J.P. Three-tier modular structural health monitoring framework using environmental and operational condition clustering for data normalization: Validation on an operational wind turbine system. *Proc. IEEE* **2016**, *104*, 1632–1646. [CrossRef]
34. Kraemer, P.; Friedmann, H.; Ebert, C.; Mahowald, J.; Wölfel, B. Experimental validation of stochastic subspace algorithms for structural health monitoring of offshore wind turbine towers and foundations. In Proceedings of the 8th European Workshop on Structural Health Monitoring, Bilbao, Spain, 5–8 July 2016; pp. 5–8.
35. Fritzen, C.P.; Kraemer, P.; Buethe, I. Vibration-based damage detection under changing environmental and operational conditions. *Adv. Sci. Technol. Trans Tech Publ.* **2013**, *83*, 95–104.
36. Ostachowicz, W.; Güemes, A. *New Trends in Structural Health Monitoring*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2013; Volume 542.

# Appended Paper III

# Wind Turbine Main Bearing Fault Prognosis Based Solely on SCADA Data

**Authors:**
Ángel Encalada-Dávila, Bryan Puruncajas, Christian Tutivén, and Yolanda Vidal

**Paper published in:**
*Sensors* 2021, 21(6), 2228
DOI: 0.3390/s21062228
Ranking JCR: Q1 (2021)

**Paper history:**
Submitted: 04-03-2021
Accepted: 20-03-2021
Published: 23-03-2021

sensors

MDPI

*Article*

# Wind Turbine Main Bearing Fault Prognosis Based Solely on SCADA Data

**Ángel Encalada-Dávila [1], Bryan Puruncajas [1,2], Christian Tutivén [1,2,3] and Yolanda Vidal [2,4,*]**

1 Mechatronics Engineering, Faculty of Mechanical Engineering and Production Science (FIMCP), Campus Gustavo Galindol, ESPOL Polytechnic University, Escuela Superior Politécnica del Litoral, ESPOL, Km. 30.5 Vía Perimetral, Guayaquil 090112, Ecuador; angaenca@espol.edu.ec (Á.E.-D.); bpurunca@espol.edu.ec (B.P.); cjtutive@espol.edu.ec (C.T.)
2 Control, Modeling, Identification and Applications (CoDAlab), Department of Mathematics, Escola d'Enginyeria de Barcelona Est (EEBE), Campus Diagonal-Besós (CDB), Universitat Politècnica de Catalunya (UPC), Eduard Maristany, 16, 08019 Barcelona, Spain
3 Facultad de Ingenierías, Universidad ECOTEC, Km. 13.5 Vía a Samborondón, Guayaquil 092302, Ecuador
4 Institute of Mathematics (IMTech), Universitat Politècnica de Catalunya (UPC), Pau Gargallo 14, 08028 Barcelona, Spain
* Correspondence: yolanda.vidal@upc.edu; Tel.: +34-934-137-309

**Abstract:** As stated by the European Academy of Wind Energy (EAWE), the wind industry has identified main bearing failures as a critical issue in terms of increasing wind turbine reliability and availability. This is owing to major repairs with high replacement costs and long downtime periods associated with main bearing failures. Thus, the main bearing fault prognosis has become an economically relevant topic and is a technical challenge. In this work, a data-based methodology for fault prognosis is presented. The main contributions of this work are as follows: (i) Prognosis is achieved by using only supervisory control and data acquisition (SCADA) data, which is already available in all industrial-sized wind turbines; thus, no extra sensors that are designed for a specific purpose need to be installed. (ii) The proposed method only requires healthy data to be collected; thus, it can be applied to any wind farm even when no faulty data has been recorded. (iii) The proposed algorithm works under different and varying operating and environmental conditions. (iv) The validity and performance of the established methodology is demonstrated on a real underproduction wind farm consisting of 12 wind turbines. The obtained results show that advanced prognostic systems based solely on SCADA data can predict failures several months prior to their occurrence and allow wind turbine operators to plan their operations.

**Keywords:** fault prognosis; wind turbine; main bearing; normality model; real SCADA data



**Citation:** Encalada-Dávila, A.; Puruncajas, B.; Tutivén, C., Vidal, Y. Wind Turbine Main Bearing Fault Prognosis Based Solely on SCADA Data. *Sensors* **2021**, *21*, 2228. https://doi.org/10.3390/s21062228

Academic Editor: Steven Chatterton

Received: 4 March 2021
Accepted: 20 March 2021
Published: 23 March 2021



## 1. Introduction

Energy is a key pillar of human evolution. Currently, the challenge of obtaining energy while minimizing costs, and pollution is a matter of concern owing to climate change and global warming, as well as the need to democratize the extraction of energy worldwide. In this regard, renewable energy, i.e., energy collected from renewable resources, is an excellent option, as they are clean and exist over a wide geographical area, unlike fossil-fuel energy sources, which are air pollutants and are concentrated in a limited number of countries. Among renewable energy sources, the wind-energy sector has grown significantly in the last two decades. In 2019, wind energy generated enough electricity to meet 15% of Europe's electricity demand [1] and was the leading source of new capacity in Europe, the U.S., and Canada, as well as the second largest in China. However, unleashing the massive potential of wind energy is crucial for reducing the levelized cost of electricity (LCOE) [2]. Increasing the size of wind turbines (WTs) and moving offshore, where steadier and higher wind speeds are available, are the two key factors in decreasing the LCOE. However, these two factors have accelerated the need for better condition monitoring strategies.

Condition monitoring is the process of monitoring a signal indicating the condition of machinery (vibration, temperature, etc.) to identify a deviation from the normal operation behavior, which is indicative of a developing fault. Condition monitoring is the crux of the matter to move from time-based preventive maintenance, which is still the current mainstream practice for WTs, to predictive maintenance, as it relies on the actual condition of the equipment rather than the average or expected life statistics. Because a failure is a process rather than an event, the earlier the process is detected, the more the flexibility that exists to manage it. Fault detection strategies usually warn about the appearance of a fault too late, and the fault is already mature when it is detected, which prevents proper planning of the maintenance operation [3–5]. Meanwhile, prognosis strategies provide a predictive maintenance option that gives the decision-maker the flexibility to decide whether and when to act before the subsystem or turbine fails. Thus, WT downtime is minimized, and the component lifetime is maximized. A significant amount of research on fault prognosis for WTs exists, some of which are reviewed in Reference [6]; the main subsystems that the majority of the research focus on include: blades (e.g., References [7,8]), gearboxes (e.g., References [9,10]), and bearings (e.g., References [11–13]). These aforementioned studies use data from specific and costly condition monitoring sensors, as they are mainly based on high-frequency vibration analysis, acoustic emission signals, or oil analysis sensors. In contrast, in this work, the proposed predictive maintenance strategy is achieved without needing to invest in additional hardware; it only requires the already available supervisory control and data acquisition (SCADA) data.

Although SCADA data have not been developed specifically for the purpose of condition monitoring, being able to extract relevant information from it could result in rapid deployment and modest set-up costs. SCADA data have been collected for long, but owing to the lack of appropriate data interpretation tools, they have not been considered for condition monitoring purposes. SCADA data are highly variable owing to the constantly changing operational conditions caused by the fluctuation of environmental conditions (such as wind speed and direction, turbulence intensity, ambient temperature, etc.), which are affected by seasonality. Furthermore, these parameters have a lower sampling rate (usually once every 10 min) compared to the kHz frequency of traditional condition monitoring strategies, they are rarely standardized, and the description of work orders is generally not clear. Thus, it is challenging for researchers to contextualize SCADA data for fault prognosis [14]. However, recent research has focused on this approach, and there are some success stories about using only SCADA data from real WTs for condition monitoring. It is important to note that using just SCADA data means that no extra sensors are used; however, some information from the SCADA alarm logs might be used but with limited fault detection and diagnosis, and no prognosis or functionality. Additionally, work orders could also be used for data labeling. For example, in Reference [15], the diagnosis and prediction of WT faults from SCADA data were accomplished using support vector machines (SVM), and, in Reference [16], a fault prognosis procedure was proposed using an a priori knowledge-based adaptive neuro-fuzzy inference system. By using a priori knowledge about faults (the data of six known WT pitch faults are used to train the system), the proposed system improves fault diagnosis. In Reference [17], a framework for automatically identifying periods of faulty operation using rules applied to the turbine alarm system are presented and applied to perform fault classification.

The aforementioned studies used SCADA data and validated their approach on real WTs; however, all of them required faulty data (historical fault data). Historical SCADA data must be accurately labeled with the periods when the turbines are down due to a fault, as well as with the cause of the fault. However, this is time-consuming, error-prone, and likely to result in a set of labeled vectors with an unbalanced number of classes. In contrast, in this work, there is no need for historical fault data; thus, the proposed strategy can be applied to any wind farm, even when no faulty data have been recorded. In this work, a normal behavior model is proposed, i.e., the model is built using normal (healthy) operation data. Heretofore, this introduction has focused on wind turbine fault diagnosis

related references, but it is also important to note that analogous concepts to the ones proposed in this work are used in other areas of application. For example, error estimation and accuracy of machine learning methodologies have been performed on real datasets in different applications, e.g., Reference [18] (vessels). Likewise, the use of solely healthy data to diagnose non-previously occurred failures is used for cross-domain fault diagnosis problems in rotating machines, e.g., Reference [19].

This work deals with the main WT bearing faults. There are two main reasons for this selection. First, as stated by the European Academy of Wind Energy (EAWE) [20], the wind industry has identified main bearing failures as a critical issue in terms of increasing WT reliability and availability, as they lead to major repairs with high replacement costs and long downtime periods. Second, the authors had access to two years of SCADA data from a real underproduction wind farm (composed of 12 WTs), where main bearing failure occurred in one of the WTs; thus, it could be used to verify the performance of the proposed strategy in a real-life situation.

Most models in literature use simulated SCADA (e.g., Reference [21]) or experimental data but rarely real data. Furthermore, when dealing with real data, only one or two WTs are tested. For example, in Reference [22], an ensemble approach was proposed to detect anomalies and diagnose faults; however, it was only tested on two WTs and the warning was only given less than a week before failure, thus not leaving adequate time to plan the operation and maintenance action. In Reference [23], data from one WT in the South-East of Ireland were used to validate the proposed strategy; fault and alarm data were filtered and analyzed in conjunction with the power curve to identify periods of nominal and faulty operation. In contrast, this study used a wind farm composed of 12 WTs to test the proposed strategy.

The remainder of this paper is organized as follows: A brief description of the wind farm is provided in Section 2. The main bearing types and their faults are reviewed in Section 3. A description of the SCADA data is provided in Section 4. The proposed fault prognosis methodology is described in Section 5. The obtained results are presented and discussed in Section 6. Finally, the conclusions and future work are presented in Section 7.

## 2. Brief Wind Farm Description

The wind farm is located in Spain and was commissioned in 2006. The WTs can generate 1500 kW of power and have a diameter of 77 m. Figure 1 shows the major components of these WTs. These are pitch-controlled WTs that not only use the pitch mechanism to brake, but also have an independent fail-safe piston accumulator on the blades. Additionally, they were equipped with a mechanical brake on a high-speed shaft. These turbines can also brake electromechanically using a generator to stop the rotation. Power production starts at a wind speed of 3.5 m/s. At 25 m/s, an automatic stopping occurs. The optimal performance can be achieved at a comparatively low wind speed of 11.1 m/s. This plant was certified according to IEC IIa. A summary of the technical specifications of the WTs is given in Table 1.

It is noteworthy that these WTs use a double-spherical main roller bearing. These types of bearings are suitable for large radial loads and low to medium speeds, thus compensating for misalignment. Spherical roller bearings have two rows of symmetrical rollers, a common spherical outer ring raceway and two spherical inner ring raceways inclined at an angle toward the bearing axis. The center point of the sphere in the outer ring raceway is at the bearing axis. Figure 2 shows a spherical roller main bearing used in the WTs. As the main bearing is the component of interest in this work, a brief review on the main bearing types and their faults is given in the next section.

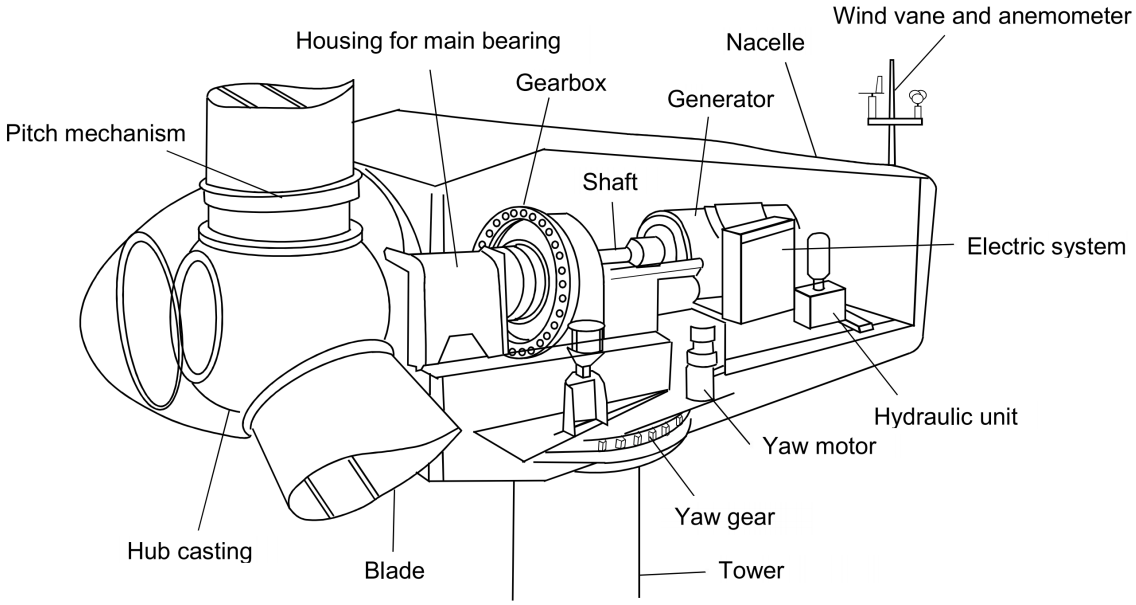



**Figure 1.** Main components of the wind turbine (WT) [24].

**Table 1.** Technical specifications of the WTs in the park.

| | |
|---|---|
| Number of blades | 3 |
| Nominal power | 1500 kW |
| Rotor diameter | 77 m |
| Wind class | IEC IIa |
| Swept area | 4657 $m^2$ |
| Nominal rotation speed | 18.3 rpm |
| Cut-in wind speed | 3.5 m/s |
| Cut-out wind speed | 25 m/s |
| Bearings | Double spherical roller bearings |
| Power regulation | Independent pitch (variable speed) |

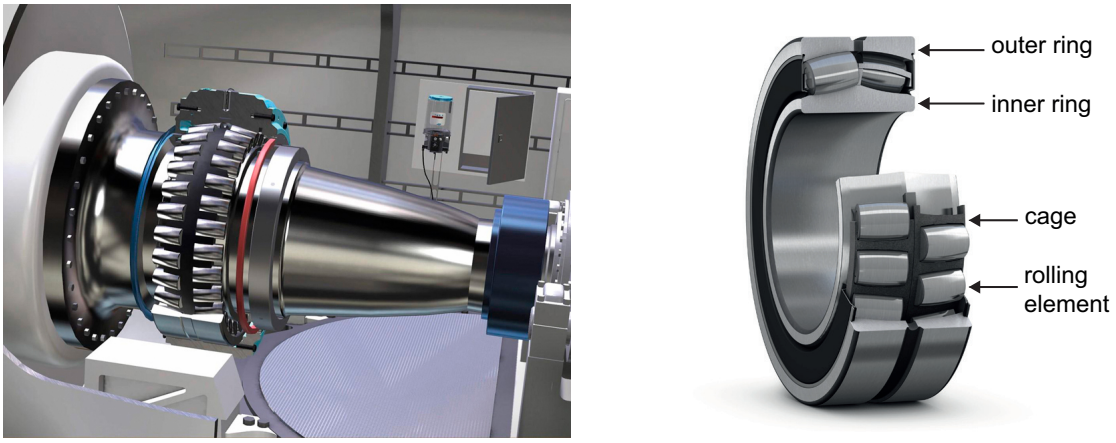



**Figure 2.** Spherical roller main bearing used in WTs. Courtesy of SKF.

## 3. Main Bearing Faults

The main bearing is a large component inside a WT, and it can be damaged in a variety of ways. In this section, the main parts of this component and their different and various failure modes are discussed. The objective of this section is to show that there is no single pattern to predict a fault in this component, but rather a large number of possible patterns. This supports the idea of using unsupervised normality-based methods to predict the main

bearing faults owing to the inability of supervised methods to predict failures outside their training dataset. However, should a supervised method be chosen, it would require a great variety of faulty data covering all failure types. Rolling bearings are composed of machine elements that permit the rotary motion of shafts for a wide range of applications, such as electric motors, aircraft gas turbines, gyroscopes, power transmissions, and WTs [25]. A typical rolling bearing consists of four elements: an inner race, an outer race, rolling elements that are in contact under heavy dynamic loads and relatively high speeds, and a cage around these rolling elements, as shown in Figure 2. Any of these parts are at risk of failure [26].

The SKF company classified the different failure modes of bearings by considering the ISO 15243 standard. This classification introduces the following failure modes [27]: (i) fatigue, (ii) wear, (iii) corrosion, (iv) electrical erosion, (v) plastic deformation, and (vi) fracture and cracking. These modes have different causes and behaviors, causing stress, deformation, micro-geometry destruction, cracking or spalling, shallow depressions, greyish-black patches, craters, and fractures. The objective of this work is to predict failures several months in advance, and, because all these failures develop progressively (i.e., develop through slow degradation), each one is briefly explained.

First, there are two types of fatigue: subsurface-initiated fatigue and surface-initiated fatigue. Subsurface-initiated fatigue occurs beneath the contact surfaces of the raceways and rolling elements (see Figure 3, left). In contrast, surface-initiated fatigue occurs due to damage to the rolling contact surface asperities, which is generally caused by inadequate lubrication (see Figure 3, right).

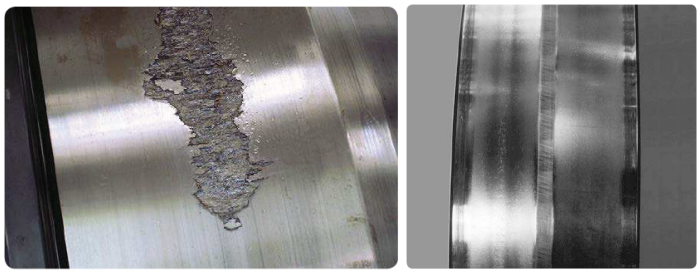


**Figure 3.** Fatigue failure. Subsurface-initiated (**left**) and surface-initiated (**right**) [27]. Courtesy of SKF.

Second, wear failure can be divided into two types: abrasive wear and adhesive wear. Abrasive wear is a degenerative process with the progressive removal of material, as shown in Figure 4 (left). In contrast, adhesive wear is a type of lubricant-related damage that occurs between two mating surfaces sliding relative to each other. It is characterized by the transfer of material from one surface to the other (called smearing). It is typically accompanied by frictional heat, which can sometimes temper or reharden the mating surfaces (see Figure 4, right).

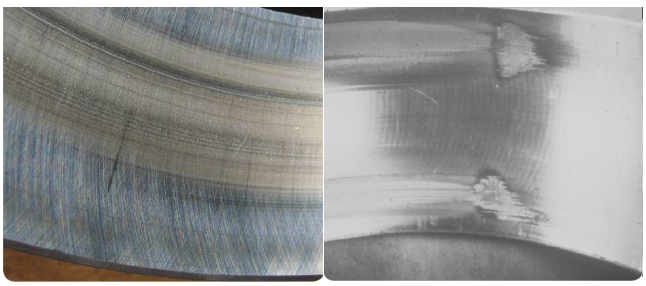


**Figure 4.** Wear failure. Abrasive wear (**left**) and adhesive wear (**right**) [27]. Courtesy of SKF.

Third, corrosion failures are divided into moisture corrosion, fretting corrosion, and false brinelling. Moisture corrosion occurs when a machine bearing is exposed to the

ingress of water and other liquids as part of the operational process, resulting in greyish-black patches coinciding with the rolling element pitch, as shown in Figure 5 (left). Fretting corrosion occurs when there is relative movement between a bearing ring and its seat on a shaft or in a housing (see Figure 5, middle). Finally, false brinelling occurs in the contact area owing to the micromovements and/or resilience of the elastic contact under cyclic vibrations. The root cause is vibration during standstill (see Figure 5, right).

**Figure 5.** Corrosion failures. Moisture (**left**), fretting (**middle**), and brinelling (**right**) [27]. Courtesy of SKF.

Fourth, there are two types of electrical erosion: excessive current erosion and current leakage erosion. Excessive current erosion occurs when an electric current flows from one ring to the other via the rolling elements, causing damage. At the contact surfaces, the process is similar to that of electric arc welding (high current density over a small contact surface). The material is heated to temperatures ranging from tempering to melting levels, as shown in Figure 6 (left). However, in the initial stage of current-leakage erosion damage, the surface is typically damaged by shallow craters that are closely positioned and smaller in diameter compared to those from the damage from excessive current erosion (see Figure 6, right).

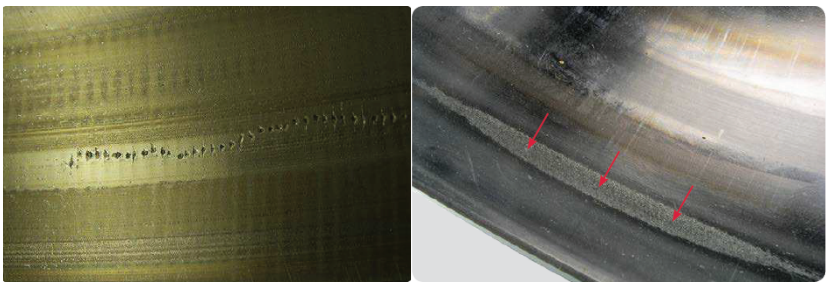


**Figure 6.** Electrical erosion failures. Excessive current (**left**) and current leakage (**right**) [27]. Courtesy of SKF.

Fifth, plastic deformation can occur due to an overload or indentations from debris. Overload deformation can be caused by static overloading, shock loads, or improper handling, as shown in Figure 7 (left). In the case of indentations from the debris failure type, solid contaminants are introduced into a bearing via the seals or lubricant. They can also be the result of wear or damage to an adjacent component, such as a gear (see Figure 7, right).

Finally, a bearing can be affected by forced fracture, fatigue fracture, or thermal cracking. A forced fracture occurs when the stress concentration exceeds the tensile strength of the material (see Figure 8, left). In contrast, a fatigue fracture starts when the fatigue strength of a material is exceeded under cyclic bending, as shown in Figure 8 (right). Finally, a thermal crack can occur when two surfaces slide against each other and generate frictional heat.

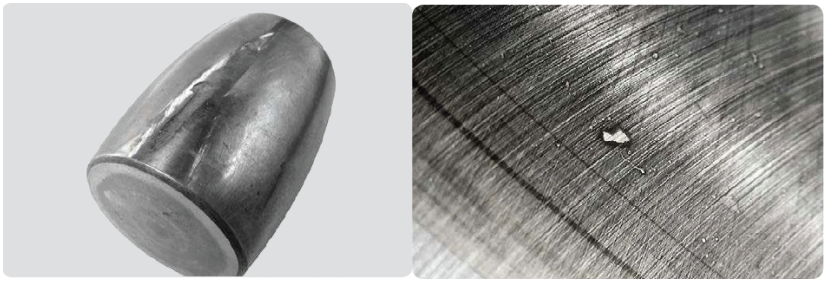


**Figure 7.** Plastic deformation failure. Overload (**left**) and indentation (**right**) [27]. Courtesy of SKF.

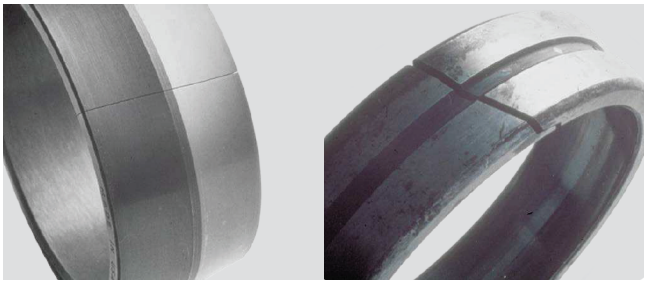


**Figure 8.** Fracture failure. Forced (**left**) and fatigue (**right**) [27]. Courtesy of SKF.

## 4. Real SCADA Data Description

The SCADA data used in this work were obtained from 12 operational WTs that can generate 1.5 MW. The continuous operational data were collected from the beginning of February 2017 till the end of November 2018. The wind farm SCADA datasets contain different variables that can be broadly grouped into environmental, electrical, component temperature, hydraulic, and control variables (see Tables 2–6). The mean, maximum, minimum, and standard deviation of the 10 min averaging period of 1 Hz sampled values are available for these variables.

Table 2 shows all the environmental related variables of the SCADA data, for example, ambient temperature, which affects the temperatures of all subsystems (the temperature of bearings changes significantly between winter and summer). It should be pointed out that the wind speed, which defines the different operating regions of the WT, is the most important exogenous variable related to the WT owing to its direct effect on the operation of the WT [28].

**Table 2.** Environmental variables.

| Variable | Description | Units |
|---|---|---|
| TempAmb | Ambient temperature | °C |
| TempGond | Nacelle temperature | °C |
| VelViento | Wind speed | m/s |
| IndTurbul | Turbulence index | - |

Table 3 shows the electrical related variables, such as the active power, which is sensitive to wind variations. The electrical energy is measured before it enters the distribution network to consider the consumption of the WT; therefore, it is considered as power delivered to the network. Electrical network frequency measurements and phase voltage measurements are also obtained to monitor possible fluctuations. Measurements of the power factor and reactive power are also collected to make adjustments, using capacitors, in the electrical system.

**Table 3.** Electrical variables.

| Variable | Description | Units |
|---|---|---|
| Pot | Active power | kW |
| TensRed | Phase voltage | V |
| CosPhi | Power factor | - |
| TotPotReact | Reactive power | kW |
| FrecRed | Electric network frequency | Hz |

Table 4 shows the temperature related variables. As stated previously, this work focuses on the main bearing fault. Thus, it is important to note the low-speed shaft temperature, as this component is close to the main bearing, as well as the bearing temperatures at the coupling and non-coupling sides. In addition, note the variables related to the gearbox; the temperature readings were taken from the upper and lower gearbox radiators, and the lubrication of the gearbox was monitored using the hydraulic group oil temperature. Additionally, the generator temperature is another relevant variable.

**Table 4.** Component temperature variables.

| Variable | Description | Units |
|---|---|---|
| TempAceiteGH | Hydraulic group oil temperature | °C |
| TempAceiteMultip | Gearbox oil temperature | °C |
| TempRodamMultip | Gearbox bearing temperature | °C |
| TempGen | Generator temperature | °C |
| TempRodamTrasero | Rear bearing temperature | °C |
| TempCojLA | Bearing coupling side temperature | °C |
| TempCojLOA | Bearing non-coupling side temperature | °C |
| TempRadSup | Upper gearbox radiator temperature | °C |
| TempRadInf | Lower gearbox radiator temperature | °C |
| TempEjeLento_1 | Low-speed shaft temperature | °C |
| TempTrafo1 | Transformer 1 temperature | °C |
| TempTrafo2 | Transformer 2 temperature | °C |
| TempTrafo3 | Transformer 3 temperature | °C |

The hydraulic variables obtained from the SCADA system are listed in Table 5. They include the pressure measurement of the general accumulator, the hydraulic group pressure, brake pressure, and general accumulated pressure of the blades. Each blade has an independent actuator with an accumulator to position the blade according to the mode of operation.

**Table 5.** Hydraulic variables.

| Variable | Description | Units |
|---|---|---|
| AcumGralPala1 | General accumulator blade 1 | bar |
| AcumGralPala2 | General accumulator blade 2 | bar |
| AcumGralPala3 | General accumulator blade 3 | bar |
| PresAcumGral | Accumulated general pressure | bar |
| PresFreno | Brake pressure | bar |
| PresGH | Hydraulic group pressure | bar |

Table 6 shows the control related variables. The WTs are equipped with blade pitch control, which adjusts the blade's angle of inclination to control the rotor speed and can execute the rotor brake in the feathered position. Another important control system is the yaw controller, which ensures that the nacelle is oriented correctly. Additionally, the rotor and generator speeds are crucial variables to control the WT operation.

**Table 6.** Control variables.

| Variable | Description | Units |
|---|---|---|
| Pitch1 | Blade pitch angle 1 | °C |
| Pitch2 | Blade pitch angle 2 | °C |
| Pitch3 | Blade pitch angle 3 | °C |
| Yaw | Yaw angle | °C |
| VelRotor | Rotor speed | rpm |
| VelGen | Generator speed | rpm |
| SPPitch | Pitch system parameter | - |
| ContEnerActiva | Active energy counter | - |
| NivOscil | Oscillation level | Hz |
| NivVibra | Vibration level | Hz |
| date_time | Date and time of the sample | - |
| ld_id | Wind turbine ID | - |

Note that, apart from the SCADA data, information regarding maintenance and repair actions (work orders) were also available. This data provided information on the failures that occurred when they occurred, when the work was carried out, and information about the subsystem that was repaired or replaced. From this information, WT number 2 in the wind farm (WT2, from now on) had a main bearing fault on 21 May 2018. This information was used in this work to test whether the proposed methodology is capable of predicting the appearance of this fault months in advance.

## 5. Fault Prognosis Methodology

In this section, the proposed methodology is comprehensively described. First, the data preprocessing, which is performed to deal with real data that contains outliers and missing data, is thoroughly explained in Section 5.1. Second, the data split into training and test sets is given in Section 5.2. In this section, emphasis is placed on why the usual strategy of data shuffle must never be used in the context of time-series data. It will also explain how it is ensured that the anomalies detected are not just a change in seasonality. Furthermore, it is urged that almost one whole year of data is used for the training set. Then, in Section 5.3, the normality model for each WT is constructed based on an ANN. The ANN was trained to obtain a model that, from the input variables (at different time steps), can estimate the value of the main shaft temperature (at a specific time step), as a virtual sensor, when healthy data are provided. Fourth, in Section 5.4, the specific details of the ANN architecture are provided. Fifth, in Section 5.5, the application of Bayesian regularization, where the weights are regularized to improves the generalization without requiring a validation set, is discussed. Sixth, Section 5.6 discusses how to ensure that the data used to construct the normality model are healthy. Finally, a fault prognosis indicator is introduced in Section 5.7 to minimize the number of false positives (false alarms).

### *5.1. Data Preprocess*

In Section 3, the diverse main bearing failure modes are described, stating that, in many cases, they lead to an increase in temperature. For this reason, to build the normality model, the temperatures of the components located close to the main bearing are selected together with the ambient temperature, as it affects the temperatures of all subsystems. Additionally, the generated power and rotor speed provide information about the region of operation of the WT. The selected variables are shown in Figure 9 and are detailed in Table 7. Note that these variables are filtered through a range of realistic values for each sensor. Table 7 lists the specific ranges used for each sensor.

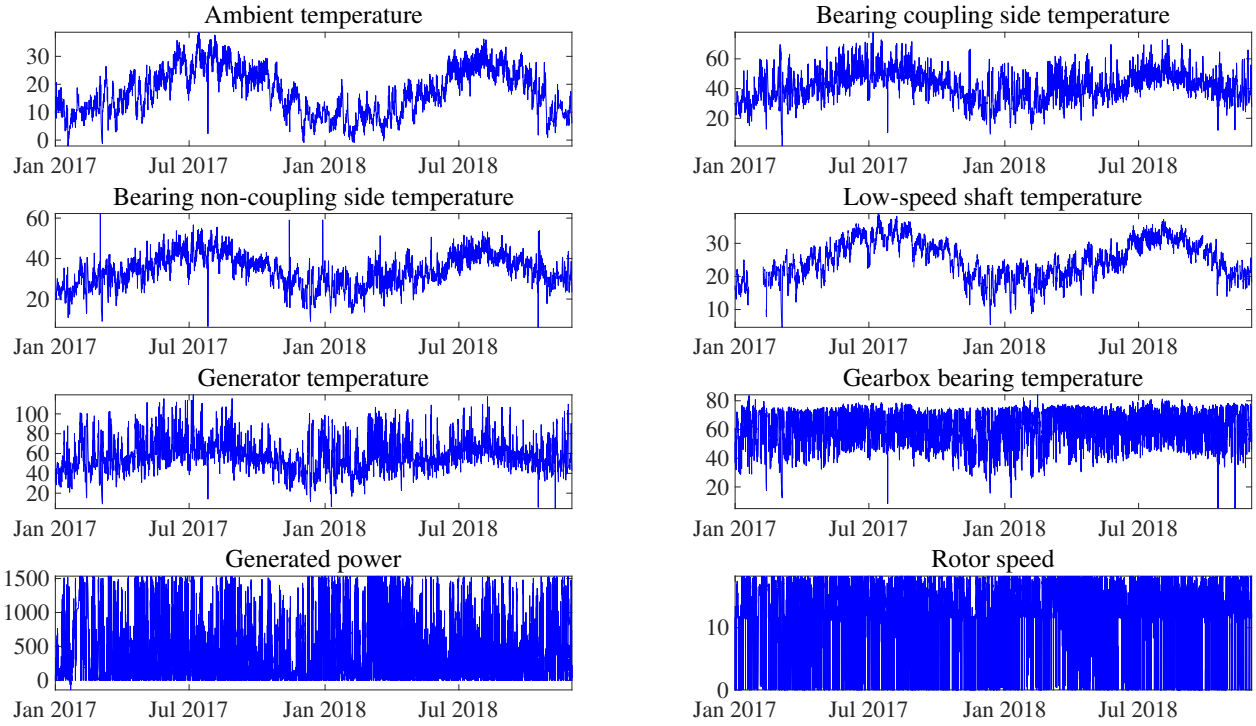



**Figure 9.** Plot example of the selected supervisory control and data acquisition (SCADA) variables used to develop the normality model. All of them are related to the mean value over a 10-min period.

**Table 7.** Selected SCADA variables used to develop the normality model, its description, range of possible values, and units. All of them are related to the mean value over a 10-min period.

| Variable | Description | Range | Units |
|---|---|---|---|
| Pot | Generated real power | [0, 2000] | kW |
| TempAmb | ambient temperature | [−5, 40] | °C |
| TempCojLA | Bearing coupling side temperature | [0, 120] | °C |
| TempCojLOA | Bearing non-coupling side temperature | [0, 120] | °C |
| TempEjeLento | Low-speed shaft temperature | [0, 120] | °C |
| TempGen | generator temperature | [0, 175] | °C |
| TempRodamMultip | Gearbox temperature | [0, 120] | °C |
| VelRotor | Rotor speed | [0, 50] | rpm |

Data cleaning is an important step prior to the application of data analysis because noisy data is removed that could cause interference in the study. Furthermore, when the study is based on real data, missing data and outliers are unavoidable (unlike when the work is based on simulated/synthetic data). In this work, extreme values (outliers) are not systematically removed because, as stated in Reference [29], it could lead to loss of information related to fault prediction. Conversely, the use of manually defined ranges based on realistic values that can be obtained by different sensors could be a better strategy. In this work, out-of-range values are first set as a missing value and then filled using the same strategy as that used for the original missing values. Figure 10 shows an example of the values outside the range for the low-speed shaft temperature.

As out-of-range values are removed, the number of missing values is increased; thus, there is a need for a data imputation strategy. Imputations with mean, median, and mode are simple techniques; however, this can introduce a bias in the mean and deviation [30]. In this work, a single imputation, while avoiding complex mathematical calculations, is proposed by using the piecewise cubic Hermite interpolating polynomial (pchip) [31]. This polynomial works for the given data points with specified slopes at the interpolation points. A meaningful property of this strategy is that the obtained

polynomial preserves the shape of the data and respects monotonicity, and guarantees that at least the first derivative is continuous. This interpolation strategy has been used in a variety of applications, such as calculating the signal-to-noise ratio in scanning electron microscopy (SEM) images [32], sampling smoothly predefined kinematic grids in high-energy particle collision problems [33], and decomposing nonlinear and non-stationary electromagnetic interference signals [34]. Figure 11 shows the original and imputed data, where the polynomial is computed and traced between the data points, considering the shape and continuity of the curve. Note that, for missing values that are at the beginning or at the end of the dataset, the closest value after or before the missing values is used.

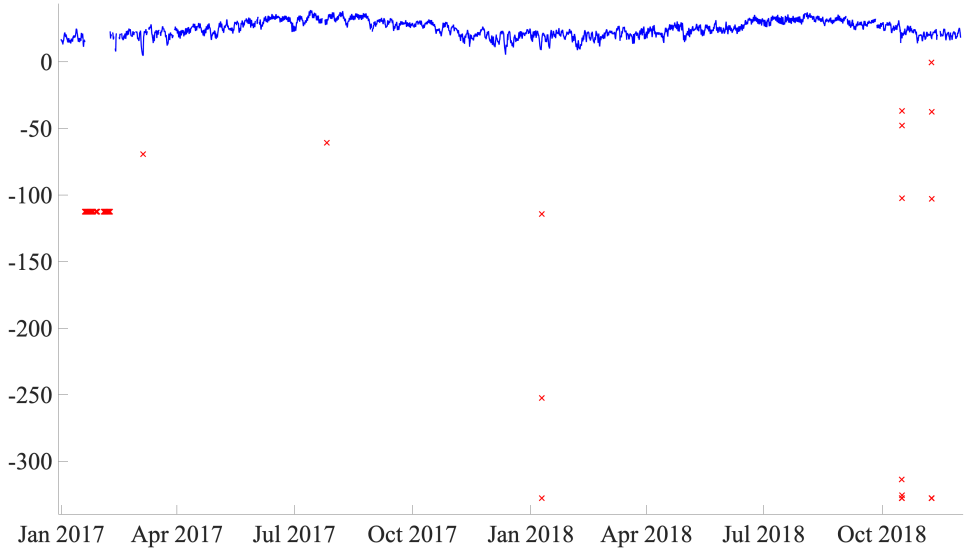



**Figure 10.** Out-of-range values are detected as outliers (red crosses) and assigned as missing values from the raw signal.

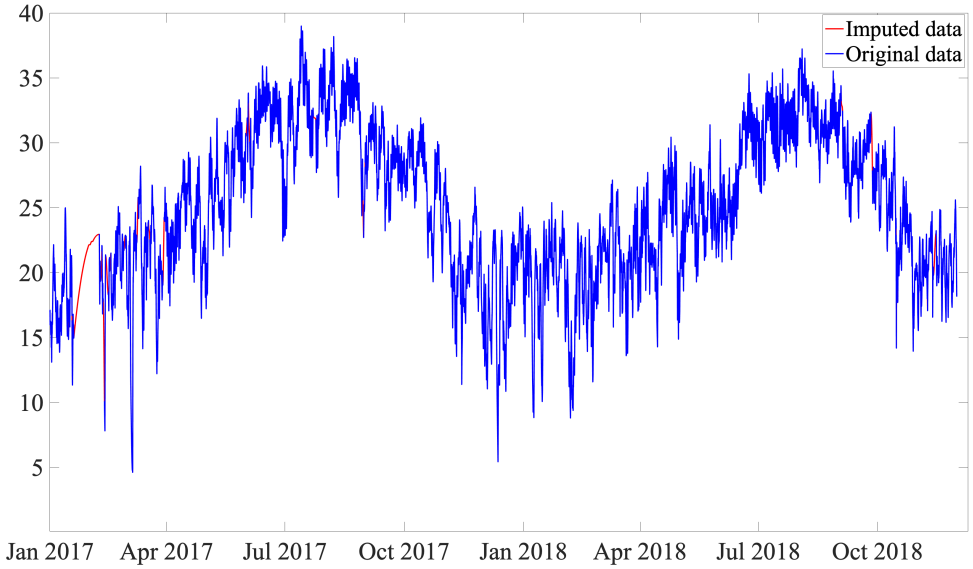



**Figure 11.** *Cont.*

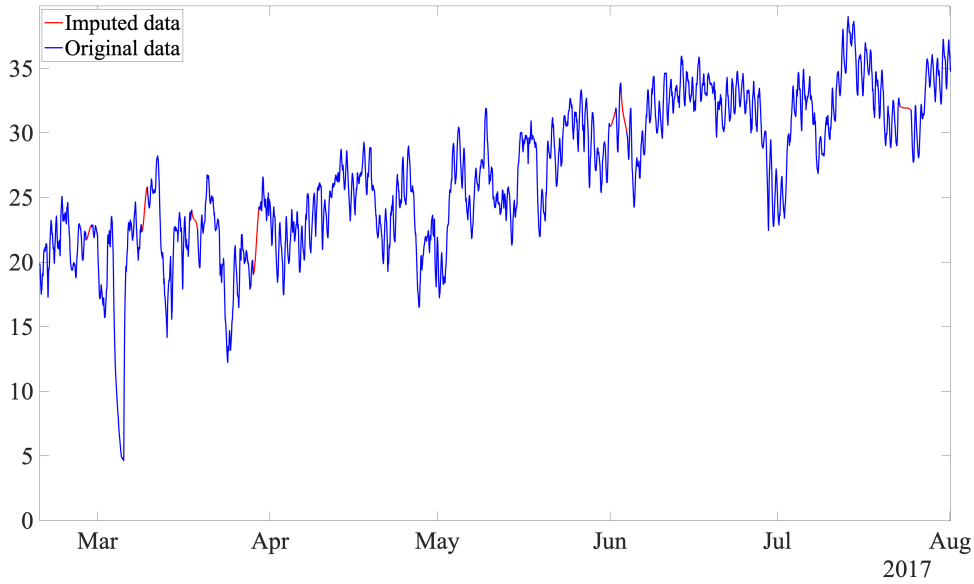



**Figure 11.** Low-speed shaft temperature raw data (without outliers) versus imputed data (**top**) and zoom in of the imputed data (**bottom**).

Finally, data from selected variables have different sources; therefore, the values have different orders of magnitude. Thus, it is highly recommended to scale the data prior to use in any machine learning approach. In this work, max-min scaling is selected to scale the data. Considering that the Bayesian regularization algorithm is used (see Section 5.5), the best results are obtained if the training data are first mapped into the range $[-1, 1]$ (or some similar region) [35]. Thus, max-min normalization was selected instead of the standard Z-score. Max-min normalization guarantees the data into the range $[0, 1]$. It is a simple technique, and its only disadvantage is coping with outliers, which has already been solved by using range filtering of the data.

### *5.2. Data Split: Train and Test Sets*

The basic steps in all neural network based models are: (1) divide the data into training and test datasets, (2) use the training dataset to train the neural network, and (3) evaluate the model using the test dataset to determine how well it predicts (generalizes). In this section, we focus on the first step.

How available data are split into training and test sets plays a fundamental role in the construction of ANNs and has a significant impact on the obtained model. In this work, a fault prognosis methodology that is insensitive to both operating and environmental conditions is sought; therefore, the training and test datasets must have data from all the working conditions.

It is noteworthy that, in this work, the training and test datasets have not been shuffled, as this can cause data leakage owing to the presence of strong time-series effects in the data [36]. Consequently, the training and test datasets were split in such a way that each set had one year of data. This approach ensures that the detected anomalies are not due to seasonality [37], and the model can cope with various operating and environmental conditions. Therefore, the available SCADA data were divided as follows: data corresponding to 2017 (47,232 samples) were used for training, and data from 2018 (43,920 samples) were used for testing. This data split was carried out for the entire wind farm. For example, Figure 12 shows the training and test data associated with WT2 (which had a main bearing fault on 21 May 2018).

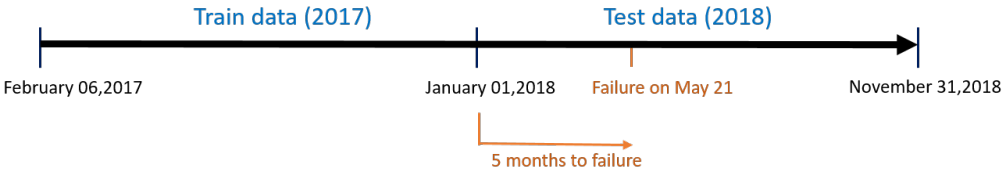



**Figure 12.** WT2 (WT number 2 in the wind farm) data for training and test.

As can be seen, in this work, there is no validation set because Bayesian regularization is used to train the ANN, as shown in Section 5.5.

*5.3. Normality Model Based on an Artificial Neural Network (ANN)*

The ANN model structure is proposed in this section and is based on the eight selected variables shown in Table 7. The output of the ANN is considered to be the temperature of the low-speed shaft (variable of interest) at time $t$, and the inputs are the remaining seven variables shown in Table 7 at time $t-1$ and $t$. Thus, referring to the structure of the ANN, there are 14 inputs and 1 output with a hidden layer comprising 72 neurons. Figure 13 shows the ANN structure.

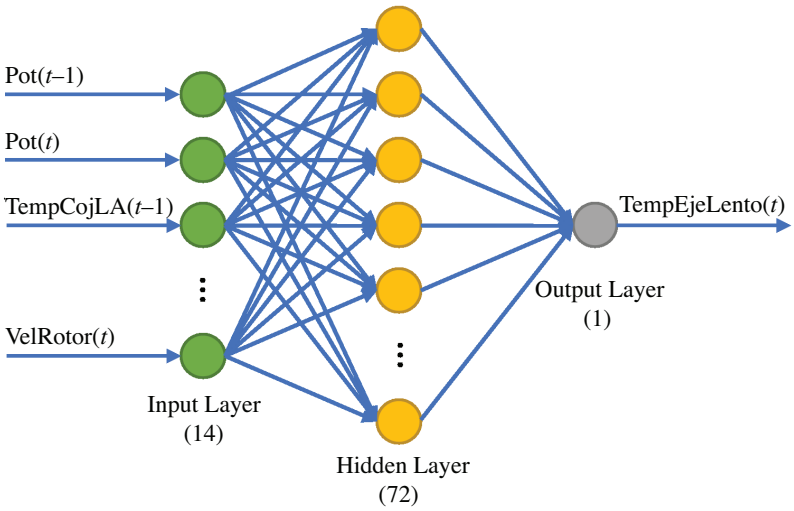



**Figure 13.** ANN model with 14 inputs, 72 neurons in the hidden layer, and 1 output.

The next sections provide a detailed explanation of the optimization method, the regularization method, and the selection of the number of neurons in the hidden layer.

*5.4. Setup of the Proposed ANN*

To provide a comprehensive reasoning of the hyper-parameter setup of the proposed ANN, a brief review of the Levenberg-Marquardt (LM) optimization method is given here to introduce the notation used.

First, note that the problem to be solved is

$$\underset{\beta}{\operatorname{argmin}} F(\beta) = \frac{1}{n}\sum_{i=1}^{n}\left(T_i - \widehat{T}_i(\beta)\right)^2,$$

where $\beta$ is the vector of parameters (in this work, it includes the weights and biases of the ANN, i.e., $\beta = (w, b)$), $n$ is the total number of samples in the training dataset, $T_i$ is the temperature value of the low-speed shaft given by the SCADA data for sample $i$, and $\widehat{T}_i$

is its estimation given by the output of the ANN. In other words, the problem consists of minimizing the mean squared error, which can be rewritten as

$$\underset{\beta}{\operatorname{argmin}} F(\beta) = \frac{1}{n}\sum_{i=1}^{n} r_i(\beta)^2,$$

where $r_i(\beta) = T_i - \widehat{T}_i(\beta)$ are the residuals. Numerical optimization algorithms are usually used to address this problem. They are iterative procedures that update the parameters at each iteration as follows:

$$\beta_{k+1} = \beta_k + \delta_k,$$

where $k$ is related to the $k$th iteration of the algorithm, and $\delta_k$ is the increment to be determined by the specifically selected algorithm. The gradient descent (GD) algorithm adopted the following increment:

$$\delta_k^{GD} = -\mu J^T r(\beta_k),$$

where $\mu$ is the learning rate, $J^T$ is the Jacobian matrix transpose of the objective function $F$, and $r$ is a column vector containing the residuals at each sample, that is, $r = (r_1, r_2, \cdots, r_n)^T$. In contrast, the Gaussian Newton (GN) algorithm uses as an increment:

$$\delta_k^{GN} = -H^{-1} J^T r(\beta_k),$$

where $H$ is the Hessian matrix of the objective function $F$. Alternatively, the LM algorithm applies an increment:

$$\delta_k^{LM} = -(H + \mu I)^{-1} J^T r(\beta_k),$$

where $I$ is the identity matrix, and $\mu$ is a nonnegative scalar parameter, usually called the damping parameter. Note that, when $\mu = 0$, the Gauss-Newton method is obtained, and, when $\mu$ is large, the method resembles gradient descent [38,39]. The fundamental idea behind the LM algorithm is to accomplish a performance similar to gradient descent when far away from the optimum, and to attain a performance similar to the Gauss-Newton method when close to the optimal value (to achieve fast convergence when being at the minimum neighborhood). To obtain this behavior, the damping parameter $\mu$ is adjusted at each iteration of the algorithm. In this work, it is raised by a factor of 10 if the current step fails to reduce the objective function, and it is decreased by a factor of 0.1 otherwise. The assigned initial value was $\mu = 0.005$. In this work, the LM algorithm terminates when at least one of the following conditions are met:

- A maximum number of 1000 epochs (using mini batches of size 128) is reached.
- The magnitude of the gradient, $J^T r(\beta_k)$, drops below the threshold $\varepsilon = 10^{-7}$.
- The damping parameter, $\mu$, exceeds its maximum possible value that has been set to $10^{10}$.

Note that the network used rectified linear unit (ReLU) activation functions, and initialization was performed using the Xavier initializer. Finally, to prevent overfitting, $L2$ regularization was introduced into the neural network using Bayesian regularization. Details are provided in the next section.

### *5.5. Bayesian Regularization*

In this study, Bayesian regularization was employed to train the ANN [35]. This regularization can be applied to multi-layer feed forward ANNs that are used for nonlinear regression, which is the case at hand. MacKay [40] comprehensively contributed to the utilization of Bayes' rule for NN training and regularization. First, in the Bayesian scheme, the ANN weights are considered as random variables, and their density functions are

updated according to Bayes' rule. Second, the training aims to minimize the objective function. Recall that, in this study, the mean squared error is minimized:

$$E(w,b) = \frac{1}{n}\sum_{i=1}^{n}\left(T_i - \widehat{T}_i(w,b)\right)^2,$$

where $w$ and $b$ are the parameters (weights and biases, respectively) of the ANN. When L2-regularization is used, an additional term is added to the objective function [41].

$$E_R(w) = \sum_{j=1}^{m} w_j^2,$$

where $m$ is the total number of weights in the ANN. Thus, the objective function $F$ becomes:

$$F(w,b) = \alpha E(w,b) + \lambda E_R(w),$$

where $\alpha$ and $\lambda$ are parameters whose relative values rule the priority for training and/or regularization, respectively. For instance, when $\alpha >> \lambda$, the training optimization algorithm minimizes errors on the training dataset, but it may lead to overfitting. However, when $\alpha << \lambda$, the training optimization algorithm will give priority to weight size curtailment (in exchange for some errors in the training dataset), thereby generating a smoother model. The main challenge in adding regularization is to set proper values for these parameters. To handle this problem, Bayesian regularization considers the application of Bayes' rule to optimize their values at each iteration of the numerical optimization [35]. A disadvantage is that this optimization requires the computation of the Hessian matrix of the objective function $F$. However, this can be approximated using the Gauss-Newton approximation, which, as noted in Reference [35], is readily available when the LM optimization algorithm is used [42] for training. A brief review of this method is given in Section 5.4 to introduce the notation used and thoroughly describe the hyper-parameter setup of the proposed ANN.

One benefit of Bayesian regularization is that it provides the so-called effective number of parameters, $\gamma$, which is a measure of how many network parameters (weights and biases) are effectively used to reduce the objective function [35,40]. If the final effective number of parameters is very close to the actual total number of parameters in the network, then the neural network may not be sufficiently large. In this case, more hidden layer neurons should be added and retrained; however, if the larger network has the same final $\gamma$ value, then the smaller network is sufficiently large. Otherwise, more hidden layer neurons may need to be added. Finally, when a sufficiently large network has been trained for a sufficient number of iterations to ensure convergence, $\gamma$ remains approximately the same, regardless of the total number of parameters in the network. That is, if an even larger network was tried, the network response would never overfit the data. This greatly simplifies the hyper-parameter tuning required to determine the optimum network size. In this study, as shown in greater detail in Section 6, a value of $\gamma = 1058$ was obtained from a total of 1153 parameters in the proposed network (number of weights and biases).

For the sake of completeness, note that the formulas to compute at each iteration, $k$, the effective number of parameters, $\gamma$, and the objective function parameters $\alpha$ and $\lambda$ are as follows [35]:

$$\gamma_{k+1} = n - 2\alpha_k tr(H)^{-1}, \qquad \alpha_{k+1} = \frac{n-\gamma_{k+1}}{2E(w,b)}, \qquad \lambda_{k+1} = \frac{\gamma_{k+1}}{2E_R(w)},$$

where $tr(\cdot)$ stands for the trace operator.

Finally, note that Bayesian regularization regularizes the weights and improves the generalization of the neural network; thus, a validation set is not required (as its main purpose is to accomplish regularization and generalization of the model).

In summary, in this work, the LM optimization algorithm was used for training optimization, and Bayesian regularization was introduced to regularize the weights and improve the generalization of the constructed model without requiring a validation set.

### 5.6. Discussion on How to Ensure That Data Used to Construct Normality Model Is Healthy

On the one hand, the proposed normal behavior model relies on the fact that healthy data are used to train it. From a pure definition point of view, this model is not completely unsupervised but is semi-supervised, as it is required to ensure that the training data are healthy. On the other hand, having complete assurance that the data used to train is healthy is not an easy task, as the absence of work orders does not guarantee that the data is completely healthy.

At the studied wind farm, there is an extra wind turbine, WT13, which is not included in the results of our methodology as it had a fault during 2017 (year used for training). Thus, it is not possible to construct a normality (healthy) model for a WT, as there are faulty data during the period decided to be used as training. However, in this section, we show the training error output when a model is built using this WT. The result is a huge training error of 30.4969, which is clearly inadmissible. Table 8 shows the training error for the rest of the WTs in the park that have healthy data during the training period (2017).

**Table 8.** Mean squared error (MSE) of the trained models for each WT. In all cases, the training time is close to 25 min ($\pm$30 s) on a 3.2 GHz 6-Core Intel Core i7 processor.

| WT | Train Error (MSE) |
|---|---|
| WT1 | 0.6984 |
| WT2 | 1.3104 |
| WT3 | 6.1074 |
| WT4 | 0.7227 |
| WT5 | 5.9989 |
| WT6 | 5.2275 |
| WT7 | 0.9214 |
| WT8 | 1.8503 |
| WT9 | 0.7373 |
| WT10 | 3.8815 |
| WT11 | 4.6074 |
| WT12 | 4.0173 |

In summary, the final train mean square error provides an estimate of the validity of the model. That is, when a WT model is harder to obtain (a much higher train mean square error is obtained), this could be a signal that the data has some anomalous behavior (it is not healthy) even if no work orders are reported. In this case, it is highly recommended to double check whether this turbine had issues during the year used as training.

### 5.7. Fault Prognosis Indicator

Typically, fault-detection indicators are defined using residuals and establishing a detection threshold. When a sample has a residual higher than the detection threshold, an alarm is triggered. However, in this case, if the residual, $|T - \widehat{T}|$, was used directly to establish a threshold above which it is decided to give the alarm signal, this would lead to a non-assumable number of false positives (false alarms) that would render the method useless. This fact is further explained in the results section. However, in this section, an indicator to overcome this problem is introduced.

As already mentioned, if the indicator is based on the residual of a single sample, there would be an excessive number of false alarms. Thus, it is important to define an indicator that considers the persistence of consecutive samples above a specified threshold. In particular, a threshold was first defined based only on the training data residuals. The

mean $\mu$ and standard deviation $\sigma$ of the residuals over all training data are computed. Then, the threshold is defined as follows:

$$\text{threshold} = \mu + 6\sigma. \tag{1}$$

Next, for each week in the test dataset, the number of samples that had a residual value greater than the threshold was counted and denoted as $n_{over}$. It is desired that the indicator has a range between 0 and 1. Thus, a minimum function is used to define the indicator in such a way that its value is 0 when no sample exceeds the threshold in that week, and a value of 1 when 504 samples (remembering that a week has 1008 samples) or more exceed the threshold in that week. Thus, the implemented weekly indicator is given as follows:

$$\text{indicator} = \min\left(1, \frac{n_{over}}{504}\right). \tag{2}$$

## 6. Results on a Real Wind Farm

The results of the proposed fault prognosis methodology for the entire wind farm is presented and discussed in this section.

First, Figure 14 (left) shows the minimization of the MSE, $E(w, b)$, during training for WT1. The best performance was 0.6984, and it is reached at the last epoch (1000). Recall that the target value is the low-speed shaft temperature; thus, the MSE has a direct physical interpretation in degrees Celsius. In addition, Figure 14 (right) shows a histogram (with 20 bins) of the final training error over all training samples for WT1. Note that the four bins with more counts have an error smaller than or equal to one degree Celsius. Furthermore, Figure 15 shows the parameter values related to the LM optimization algorithm and Bayesian regularization for WT1 training. Note that at epoch 1000, the following values were obtained: The gradients were $J^T r(\beta_{1000}) = 0.0421$, the damping parameter was $\mu = 0.5$, and the final number of effective parameters used by the ANN was $\gamma = 1058$. A larger network with an increased number of neurons in the hidden layer led to the same number for the parameter $\gamma$, thus proving that the size of the presented ANN (with 72 neurons in the hidden layer) was large enough.

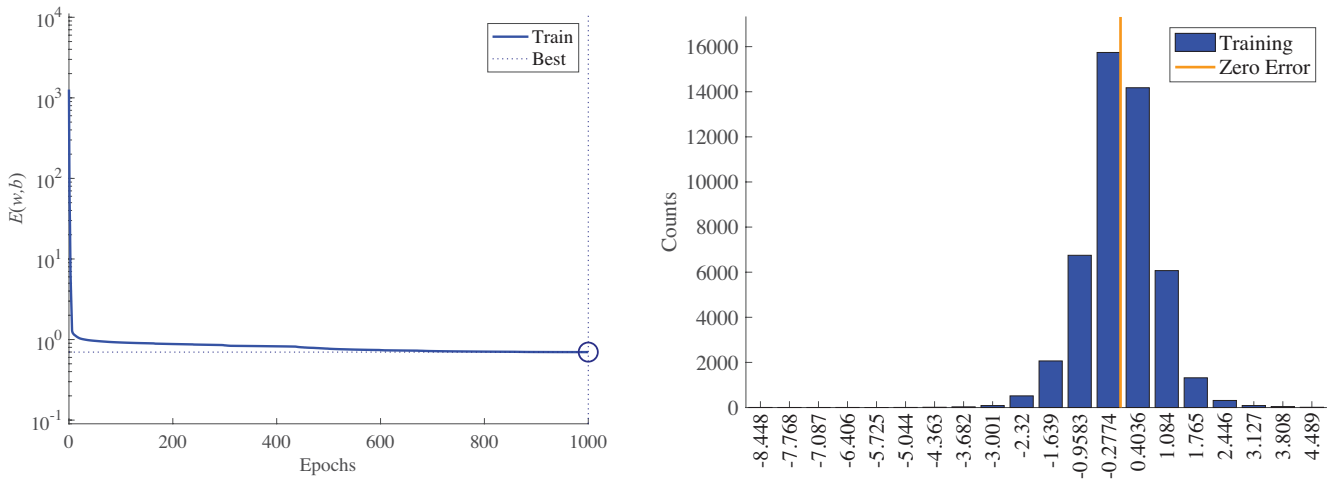



**Figure 14.** Minimization of the MSE, $E(w, b)$, during training of WT1 (**left**). Error histogram with 20 bins of final training error over all training samples for WT1 (**right**).

Next, recall that Table 8 in Section 5.6 shows the final training error for each WT in the park. There is some variability among the different wind turbines, with minimum and maximum values of 0.6984 and 6.1074, respectively. These are acceptable values, related to the real-life variability among WTs due to, for example, different locations in the park. However, it is important to note that, when the values of the training error are much higher,

then it should be considered that the training data might not be completely healthy. Most likely, if the ANN was not capable of significantly reducing the MSE for that specific WT, it is because the training data have some kind of anomaly. Recall that our main purpose is to construct a normality (healthy) model; thus, it is essential to have at disposal normal (healthy) training data.

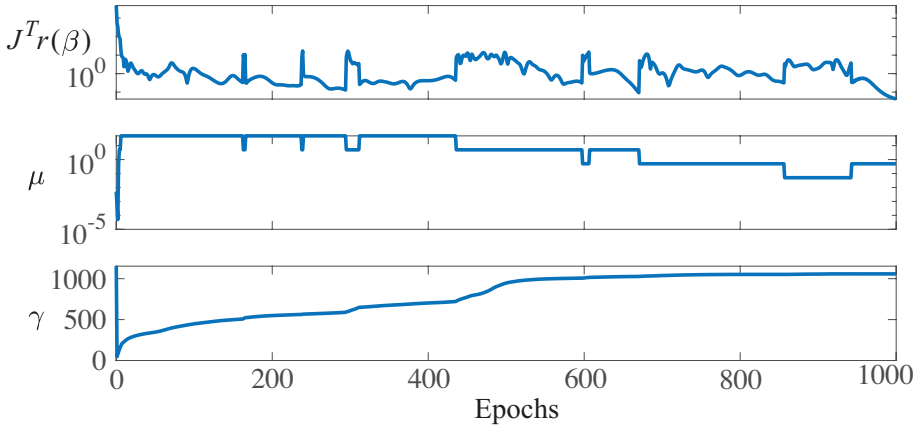



**Figure 15.** Values at each training epoch iteration for the gradient, $J^T r(\beta)$, damping parameter, $\mu$, and effective number of parameters, $\gamma$, for WT1.

Figure 16a,b show the ANN predicted value ($\widehat{T}$) and target ($T$) value for WT1 over the train and test dataset, respectively. Recall that this WT is healthy in both the training and test datasets. The prediction is close to the target in both figures, and only some samples on the test dataset have disparate values. This performance is shown in Figure 16c,d, where the absolute difference value between prediction and estimation, $|T - \widehat{T}|$, also called residual, is shown for WT1 over the training and test datasets, respectively. This residual has similar values over the training and test sets, and only a few samples have peak values. As already mentioned in Section 5.6, if the residual, $|T - \widehat{T}|$, was used directly to establish a threshold above which it is decided to give the alarm signal, this would lead to a non-assumable amount of false positives. Thus, the importance of defining an indicator, see Section 5.7, which considers the persistence of consecutive samples above a specified threshold.

In contrast, Figure 16e,f present the ANN predicted values ($\widehat{T}$) and target ($T$) values for WT2 over the training and test datasets, respectively. Recall that this WT had a main bearing fault on 21 May 2018. These figures show that the prediction over the test set has an overall performance that is different from that over the training set. Note that Figure 16g,h, which represent the residuals on the training and test datasets, also contribute to observing this change in performance over the two sets. It is noteworthy that after the failure (on 21 May 2018), the residual is no longer affected and has a similar performance to the one on the train set, with only a few isolated peaks.

Finally, the results obtained with the indicator proposed in Section 5.7 are shown in Figure 17 for the test dataset (2018) over the entire wind park. An alarm is triggered only when the indicator reaches a value greater than 0.5. The first WT with a triggered alarm was WT2, which is a true positive. The alarm was activated on 4 February; thus, three and a half months in advance of the actual breakdown reported on 21 May, where the low-speed shaft had to be replaced with a new one. The alarm stayed until 12 February and was then set off. This is because the possible heat created from an initial failure mode is detected by the fault prognosis methodology, but its appearance is not continuous over time until the final breakdown. In contrast, when the failure mode advances, for example, when a crack propagates, the generated heat appears. When the crack remains still, no further heat is generated; thus, the alarm is set off. However, cracks are already present and can advance at any time, leading to the possible failure of the component. Thus, in this methodology, whenever the alarm is on (even when it is set off after a few weeks), it is

highly recommended to check the specific WT. The second WT with a triggered alarm was the WT8. This WT has no main bearing damage, but the work orders show that the gearbox was broken and replaced from 22 March to 11 April. The proposed indicator triggered an alarm on 1 April. Thus, the method detects this maintenance operation on the gearbox as an anomaly. In a real situation, the wind park manager knows that this WT is already under maintenance. Thus, it was not a false positive. The final WT with a triggered alarm was WT9. This is a false positive of the method, as this WT had no important work orders during the year 2018. The rest of the WTs in the park were correctly classified as healthy over all the test datasets. To summarize, the results lead to a precision of 50% and a recall of 100% in the wind farm under study.

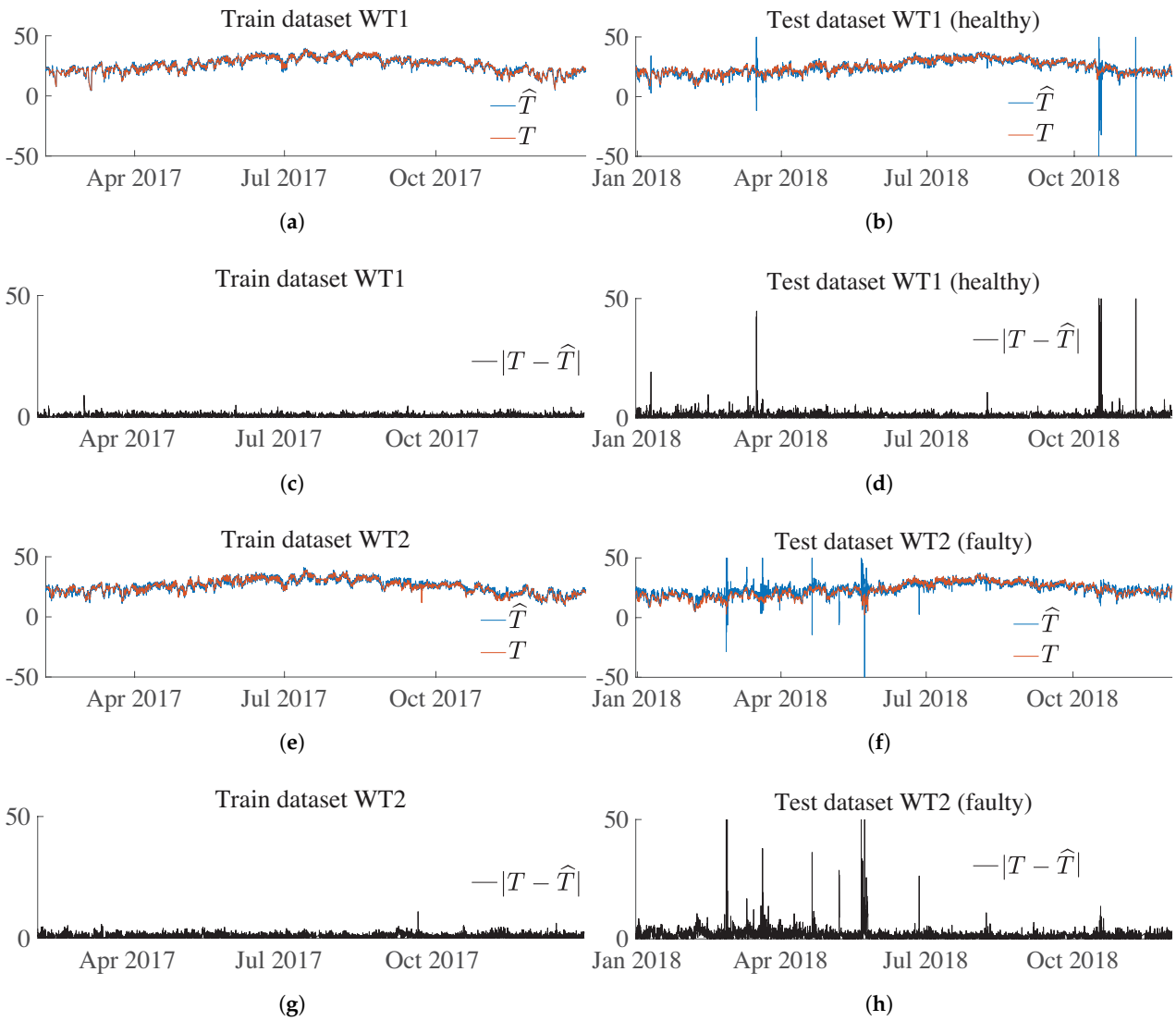



**Figure 16.** (**a**) ANN predicted value ($\widehat{T}$) and target ($T$) value for WT1 over the train dataset. (**b**) ANN predicted value ($\widehat{T}$) and target ($T$) value for WT1 over the test dataset. (**c**) Absolute difference value between the prediction and estimation, $|T-\widehat{T}|$, for WT1 over the train dataset. (**d**) Absolute difference value between the prediction and estimation, $|T-\widehat{T}|$, for WT1 over the test dataset. (**e**) ANN predicted value ($\widehat{T}$) and target ($T$) value for WT2 over the train dataset. (**f**) ANN predicted value ($\widehat{T}$) and target ($T$) value for WT2 over the test dataset. (**g**) Absolute difference value between the prediction and estimation, $|T-\widehat{T}|$, for WT2 over the train dataset. (**h**) Absolute difference value between the prediction and estimation, $|T-\widehat{T}|$, for WT2 over the test dataset.

Finally, the obtained results are compared to other approaches in the literature coping with the same proposed problem. The methodologies stated in References [15,17,23]

achieved promising results for the diagnosis and prediction of WT faults from SCADA data. However, the obtained results in this present work surpass the ones given in the aforementioned references for two main reasons: (i) the prognosis is accomplished months in advance instead of only hours in advance of the fault, and (ii) as the proposed approach is unsupervised it does not need previous faulty data to be trained neither has to deal with the problem of the highly imbalanced nature of fault data (with the no-fault class having an overwhelming majority of samples).

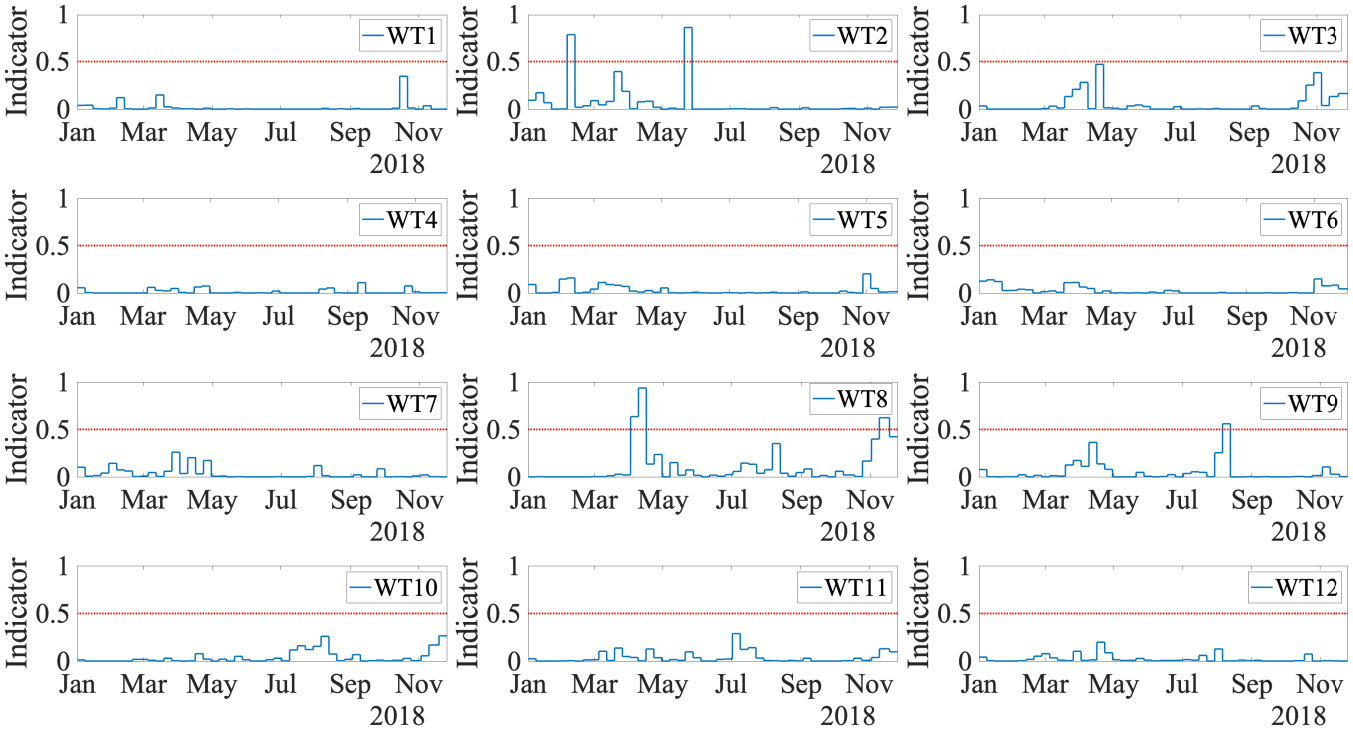



**Figure 17.** ANN indicator values (blue line) for test data, and threshold (red line).

## 7. Conclusions

In this work, an advanced prognostic system was proposed and proven to predict the main bearing failure before it occurs and let turbine operators plan their operations. In particular, a fault prognosis methodology that uses solely SCADA data requires only healthy data to be deployed. Furthermore, the stated strategy works under different operating and environmental conditions to which WTs are subject. Finally, the validity and performance of the established methodology were demonstrated on a real underproduction wind farm composed of 12 WTs. The results show that the time that the early prognosis can be generated is several months in advance, thus giving time to the plant operator to schedule maintenance. However, the studied wind farm has only one case with a failure of interest, which is not sufficient for statistical analysis. To set the expected predictive time and its confidence level is a future work that can be assessed when more cases appear in this wind farm, or data from other wind farms with more cases related to this fault are available. Furthermore, note that the work orders available to the authors of this work only contain those related to important systems substitutions, such as gearbox, generator, or main bearing replacement. Thus, preventive maintenance work orders, or minor work orders are not available. As future work, it would be interesting to check whether values of the indicator lower than the threshold but close to it are related to some minor work orders. Finally, other neural networks, such as long short-term memory (LSTM), recurrent neural network (RNN), or one dimension convolutional neural network (1D-CNN), should be

studied, as they have interesting properties that, initially, make them appropriate for the problem under consideration.

**Author Contributions:** Conceptualization, Y.V.; Data curation, Á.E.-D. and B.P.; Formal analysis, C.T.; Funding acquisition, Y.V.; Investigation, Á.E.-D., B.P., B.P. and Y.V.; Methodology, B.P. and Y.V.; Project administration, Y.V.; Software, Á.E.-D. and B.P.; Supervision, B.P. and Y.V.; Validation, Á.E.-D. and B.P.; Writing—original draft, Á.E.-D., B.P., B.P. and Y.V.; Writing—review & editing, Y.V. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work has been partially funded by the Spanish Agencia Estatal de Investigación (AEI)—Ministerio de Economía, Industria y Competitividad (MINECO), and the Fondo Europeo de Desarrollo Regional (FEDER) through the research project DPI2017-82930-C2-1-R; and by the Generalitat de Catalunya through the research project 2017 SGR 388.

**Acknowledgments:** The authors express their gratitude and appreciation to the Smartive company (http://smartive.eu/, accessed on 22 March 2021), as this work would not have been possible without their support and ceded wind farm data.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data required to reproduce these findings cannot be shared at this time as it is proprietary.

**Conflicts of Interest:** The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of the data; in the writing of the manuscript; and in the decision to publish the results.

## References

1. Europe, W. *Wind Energy in Europe in 2019—Trends and Statistics*; Wind Europe: Brussels, Belgium, 2020.
2. Shen, W.; Chen, X.; Qiu, J.; Hayward, J.A.; Sayeef, S.; Osman, P.; Meng, K.; Dong, Z.Y. A comprehensive review of variable renewable energy levelized cost of electricity. *Renew. Sustain. Energy Rev.* **2020**, *133*, 110301. [CrossRef]
3. Tang, M.; Zhao, Q.; Ding, S.X.; Wu, H.; Li, L.; Long, W.; Huang, B. An Improved LightGBM Algorithm for Online Fault Detection of Wind Turbine Gearboxes. *Energies* **2020**, *13*, 807. [CrossRef]
4. Ruiz, M.; Mujica, L.E.; Alferez, S.; Acho, L.; Tutiven, C.; Vidal, Y.; Rodellar, J.; Pozo, F. Wind turbine fault detection and classification by means of image texture analysis. *Mech. Syst. Signal Process.* **2018**, *107*, 149–167. [CrossRef]
5. Pozo, F.; Vidal, Y.; Serrahima, J.M. On real-time fault detection in wind turbines: Sensor selection algorithm and detection time reduction analysis. *Energies* **2016**, *9*, 520. [CrossRef]
6. Hossain, M.L.; Abu-Siada, A.; Muyeen, S. Methods for advanced wind turbine condition monitoring and early diagnosis: A literature review. *Energies* **2018**, *11*, 1309. [CrossRef]
7. Florian, M.; Sørensen, J.D. Wind turbine blade life-time assessment model for preventive planning of operation and maintenance. *J. Mar. Sci. Eng.* **2015**, *3*, 1027–1040. [CrossRef]
8. Tang, J.; Soua, S.; Mares, C.; Gan, T.H. An experimental study of acoustic emission methodology for in service condition monitoring of wind turbine blades. *Renew. Energy* **2016**, *99*, 170–179. [CrossRef]
9. Cheng, F.; Qu, L.; Qiao, W. Fault prognosis and remaining useful life prediction of wind turbine gearboxes using current signal analysis. *IEEE Trans. Sustain. Energy* **2017**, *9*, 157–167. [CrossRef]
10. Dupuis, R. Application of oil debris monitoring for wind turbine gearbox prognostics and health management. In Proceedings of the Annual Conference of the Prognostics and Health Management Society, Portland, OR, USA, 10–16 October 2010; pp. 10–16.
11. Rezamand, M.; Kordestani, M.; Carriveau, R.; Ting, D.S.K.; Saif, M. An integrated feature-based failure prognosis method for wind turbine bearings. *IEEE/ASME Trans. Mechatron.* **2020**, *25*, 1468–1478. [CrossRef]
12. Motahari-Nezhad, M.; Jafari, S.M. ANFIS system for prognosis of dynamometer high-speed ball bearing based on frequency domain acoustic emission signals. *Measurement* **2020**, *166*, 108154. [CrossRef]
13. Elasha, F.; Greaves, M.; Mba, D.; Fang, D. A comparative study of the effectiveness of vibration and acoustic emission in diagnosing a defective bearing in a planetry gearbox. *Appl. Acoust.* **2017**, *115*, 181–195. [CrossRef]
14. Leahy, K.; Gallagher, C.; O'Donovan, P.; O'Sullivan, D.T. Issues with data quality for wind turbine condition monitoring and reliability analyses. *Energies* **2019**, *12*, 201. [CrossRef]
15. Leahy, K.; Hu, R.L.; Konstantakopoulos, I.C.; Spanos, C.J.; Agogino, A.M.; O'Sullivan, D.T. Diagnosing and predicting wind turbine faults from SCADA data using support vector machines. *Int. J. Progn. Health Manag.* **2018**, *9*, 1–11.
16. Chen, B.; Matthews, P.C.; Tavner, P.J. Wind turbine pitch faults prognosis using a-priori knowledge-based ANFIS. *Expert Syst. Appl.* **2013**, *40*, 6863–6876. [CrossRef]

17. Leahy, K.; Gallagher, C.; O'Donovan, P.; Bruton, K.; O'Sullivan, D.T. A robust prescriptive framework and performance metric for diagnosing and predicting wind turbine faults based on SCADA and alarms data with case study. *Energies* **2018**, *11*, 1738. [CrossRef]
18. Theodoropoulos, P.; Spandonidis, C.C.; Themelis, N.; Giordamlis, C.; Fassois, S. Evaluation of Different Deep-Learning Models for the Prediction of a Ship's Propulsion Power. *J. Mar. Sci. Eng.* **2021**, *9*, 116. [CrossRef]
19. Chen, J.; Wang, J.; Zhu, J.; Lee, T.H.; De Silva, C.C. Unsupervised Cross-domain Fault Diagnosis Using Feature Representation Alignment Networks for Rotating Machinery. *IEEE/ASME Trans. Mechatron.* **2020**. [CrossRef]
20. Hart, E.; Clarke, B.; Nicholas, G.; Kazemi Amiri, A.; Stirling, J.; Carroll, J.; Dwyer-Joyce, R.; McDonald, A.; Long, H. A review of wind turbine main bearings: design, operation, modelling, damage mechanisms and fault detection. *Wind. Energy Sci.* **2020**, *5*, 105–124. [CrossRef]
21. Vidal, Y.; Pozo, F.; Tutivén, C. Wind turbine multi-fault detection and classification based on SCADA data. *Energies* **2018**, *11*, 3018. [CrossRef]
22. Jin, X.; Xu, Z.; Qiao, W. Condition Monitoring of Wind Turbine Generators Using SCADA Data Analysis. *IEEE Trans. Sustain. Energy* **2020**, *12*, 202–210. [CrossRef]
23. Leahy, K.; Hu, R.L.; Konstantakopoulos, I.C.; Spanos, C.J.; Agogino, A.M. Diagnosing wind turbine faults using machine learning techniques applied to operational data. In Proceedings of the 2016 IEEE International Conference on Prognostics and Health Management (ICPHM), Ottawa, ON, Canada, 20–22 June 2016; pp. 1–8.
24. Jiang, Z.; Hu, W.; Dong, W.; Gao, Z.; Ren, Z. Structural reliability analysis of wind turbines: A review. *Energies* **2017**, *10*, 2099. [CrossRef]
25. Srinivasan, V. Analysis of dynamic load characteristics on hydrostatic bearing with variable viscosity and temperature using simulation technique. *Indian J. Sci. Technol.* **2013**, *6*, 4797–4803.
26. Hamadache, M.; Lee, D. Wind turbine main bearing fault detection via shaft speed signal analysis under constant load. In Proceedings of the 2016 16th International Conference on Control, Automation and Systems (ICCAS), Gyeongju, Korea, 16–19 October 2016; pp. 1579–1584.
27. Bearing Damage and Failure Analysis. Available online: https://www.skf.com/binaries/pub12/Images/0901d1968064c148-Bearing-failures---14219_2-EN_tcm_12-297619.pdf (accessed on 24 January 2021).
28. Tavner, P.; Edwards, C.; Brinkman, A.; Spinato, F. Influence of wind speed on wind turbine reliability. *Wind Eng.* **2006**, *30*, 55–72. [CrossRef]
29. Marti-Puig, P.; Blanco-M, A.; Cárdenas, J.J.; Cusidó, J.; Solé-Casals, J. Effects of the pre-processing algorithms in fault diagnosis of wind turbines. *Environ. Model. Softw.* **2018**, *110*, 119–128. [CrossRef]
30. Zhang, Z. Missing data imputation: Focusing on single imputation. *Ann. Transl. Med.* **2016**, *4*, 9. [CrossRef]
31. Lu, S.; Wang, Y.; Wu, Y. Novel High-Precision Simulation Technology for High-Dynamics Signal Simulators Based on Piecewise Hermite Cubic Interpolation. *IEEE Trans. Aerosp. Electron. Syst.* **2018**, *54*, 2304–2317. [CrossRef]
32. Sim, K.S.; Yeap, Z.X.; Ting, F.; Tso, C. The performance of adaptive tuning piecewise cubic hermite interpolation model for signal-to-noise ratio estimation. *Int. J. Innov. Comput. Inf. Control.* **2018**, *14*, 1787–1804. [CrossRef]
33. Ilten, P. CIMBA: Fast Monte Carlo generation using cubic interpolation. *Comput. Phys. Commun.* **2021**, *258*, 107622. [CrossRef]
34. Li, H.; Li, L.; zhao, D. An improved EMD method with modified envelope algorithm based on C2 piecewise rational cubic spline interpolation for EMI signal decomposition. *Appl. Math. Comput.* **2018**, *335*, 112–123. [CrossRef]
35. Foresee, F.D.; Hagan, M.T. Gauss-Newton approximation to Bayesian learning. In Proceedings of the International Conference on Neural Networks (ICNN'97), Houston, TX, USA, 12 June 1997; Volume 3, pp. 1930–1935.
36. Leahy, K.; Gallagher, C.V.; Bruton, K.; O'Donovan, P.; O'Sullivan, D.T. Automatically identifying and predicting unplanned wind turbine stoppages using scada and alarms system data: Case study and results. In *Journal of Physics: Conference Series*; IOP Publishing: Bristol, UK, 2017; pp. 1–14.
37. McKinnon, C.; Turnbull, A.; Koukoura, S.; Carroll, J.; McDonald, A. Effect of time history on normal behaviour modelling using SCADA data to predict wind turbine failures. *Energies* **2020**, *13*, 4745. [CrossRef]
38. Levenberg, K. A method for the solution of certain non-linear problems in least squares. *Q. Appl. Math.* **1944**, *2*, 164–168. [CrossRef]
39. Marquardt, D.W. An algorithm for least-squares estimation of nonlinear parameters. *J. Soc. Ind. Appl. Math.* **1963**, *11*, 431–441. [CrossRef]
40. MacKay, D.J. Bayesian interpolation. *Neural Comput.* **1992**, *4*, 415–447. [CrossRef]
41. Shi, G.; Zhang, J.; Li, H.; Wang, C. Enhance the performance of deep neural networks via L2 regularization on the input of activations. *Neural Process. Lett.* **2019**, *50*, 57–75. [CrossRef]
42. Hagen, M.; Menhaj, M. Training multilayer networks with the Marquardt algorithm. *IEEE Trans. Neural Netw.* **1994**, *5*, 989–993. [CrossRef]

# Appended Paper IV

# Early Fault Detection in the Main Bearing of Wind Turbines Based on Gated Recurrent Unit (GRU) Neural Networks and SCADA Data

**Authors:**
Ángel Encalada-Dávila, Luis Moyón, Christian tutivén, Bryan Puruncajas, and Yolanda Vidal

**Paper published in:**
*IEEE/ASME Transactions on Mechatronics* 2022
DOI: 10.3390/s21030849
Ranking JCR: Q1 (2022)

**Paper history:**
Submitted: 21-12-2020
Accepted: 23-01-2021
Published: 08-07-2022

# Early Fault Detection in the Main Bearing of Wind Turbines Based on Gated Recurrent Unit (GRU) Neural Networks and SCADA Data

Ángel Encalada-Dávila, Luis Moyón, Christian Tutivén, Bryan Puruncajas, and Yolanda Vidal, *Senior Member, IEEE*

***Abstract*—Failures in the main bearings of wind turbines are critical in terms of downtime and replacement cost. Early diagnosis of their faults would lower the levelized cost of wind energy. Thus, this work discusses a gated recurrent unit (GRU) neural network, which detects faults in the main bearing some months ahead (when the event that initiates/develops the failure releases heat) the actual fatal fault materializes. GRUs feature internal gates that govern information flow and are utilized in this study for their capacity to understand whether data in a time series is crucial enough to preserve or forget. It is noteworthy that the proposed methodology only requires healthy Supervisory Control and Data Acquisition (SCADA) data. Thus, it can be deployed to old wind parks (nearing the end of their lifespan) where specific high frequency condition monitoring sensors are not installed and to new wind parks where faulty historical data do not exist yet. The strategy is trained, validated, and finally tested using SCADA data from an in-production wind park composed of nine wind turbines.**

***Index Terms*—early fault detection, wind turbine, main bearing, gated recurrent unit neural network, anomaly detection, SCADA data.**

## I. Introduction

THE global energy system is undeniably in flux. Renewable energy uptake and utilization are critical to combating climate change and ensuring a long-term future. Renewable electricity will be key to Europe reaching climate neutrality by 2050, according to the European Commission's projections [1]. Wind energy is a critical component in achieving this goal, as it accounts for 50% of the European Union's power mix, with renewables accounting for 81%. The core of the problem in the progress of the wind business, however, is a decrease in the levelized cost of electricity (LCOE). The LCOE of a wind park includes many parameters such as total installed costs, lifetime capacity factor, operation and maintenance (O&M) expenses, project economic lifespan, and cost of capital. While all these criteria play a role in calculating a project's LCOE, some have a greater influence. Wind energy O&M expenses, particularly, account for a major portion of the LCOE (20-25 percent for onshore wind parks and 25-30 percent for offshore wind parks) [2]. Hence, optimizing maintenance procedures forms a critical component of achieving low-cost wind energy.

In any industrial-scale wind park, energy output losses due to unforeseen asset maintenance, as well as component replacement costs, can amount up to millions of euros each year. Hence, it is critical to transition from corrective (replacing failed parts) and preventive maintenance (planned at regular intervals without regard for the asset's actual state) to predictive maintenance, which is based on actual and timely data collected through a network of sensors monitoring the actual asset (performed using high-frequency data of physical quantities) and warns operators in advance before the fatal break-down materializes, enabling them to program, repairs to match with weather or production windows, thus lowering costs. To better use the information in the vast quantity of data (gathered continuously or periodically, online or offline) from diverse sensors obtained from the assets, digitalization and artificial intelligence are crucial technologies. The main concept is to identify changes in the situation that suggest a growing malfunction and reflect departures from typical operational procedures.

This paper provides an early (months ahead of time) defect detection technique for the main bearing of a wind turbine (WT) based on a GRU neural network (NN) that employs just SCADA data to this framework. As SCADA data is primarily used for operation and control rather than condition monitoring (CM), using it for this purpose is a significant issue. SCADA data contains approximately 200 different variables (i.e., it is high dimensional), has a low sampling rate (recorded continuously at 10-minute averaged intervals to minimize data transmission bandwidth and storage), depends on the WT's

The authors are grateful to the Smartive firm since this study would not have been feasible without their provided wind park data. Additionally, this work has been partially funded by the Spanish Agencia Estatal de Investigación (AEI) - Ministerio de Economía, Industria y Competitividad (MINECO), and the Fondo Europeo de Desarrollo Regional (FEDER) through the research project DPI2017-82930-C2-1-R; and by the Generalitat de Catalunya through the research project 2017 SGR 388.

*Corresponding author: Yolanda Vidal (yolanda.vidal@upc.edu)*
Ángel Encalada-Dávila, Christian Tutivén, and Bryan Puruncajas are with the Escuela Superior Politécnica del Litoral, ESPOL, Faculty of Mechanical Engineering and Production Science, FIMCP, Mechatronics Engineering, Campus Gustavo Galindo Km. 30.5 Vía Perimetral, P.O. Box 09-01-5863, Guayaquil, Ecuador (e-mail: angaenca@espol.edu.ec, cjtutive@espol.edu.ec, and bpurunca@espol.edu.ec).
Luís Moyón is with Universidad Ecotec, Km. 13.5 Samborondón, Samborondón, EC092302, Ecuador (e-mail: amoyon@hotmail.com).
Bryan Puruncajas and Yolanda Vidal are with the Research Group of Control, Data, and Artificial Intelligence (CoDAlab), Department of Mathematics, Escola d'Enginyeria de Barcelona Est (EEBE), Universitat Politècnica de Catalunya (UPC), Campus Diagonal-Besós (CDB), Eduard Maristany, 16, 08019 Barcelona, Spain.
Yolanda Vidal is also with the Institute of Mathematics (IMTech), Universitat Politècnica de Catalunya (UPC), Pau Gargallo 14, 08028 Barcelona, Spain (e-mail: yolanda.vidal@upc.edu).

All authors contributed equally to this work.

region of operation and environmental conditions, and is a time series with strong seasonality. Furthermore, the benefit of keeping consistent maintenance work order records with full fault descriptions was unknown when SCADA systems were established (as it was not envisioned that AI could help in this application). Operators and maintenance contractors record maintenance repair operations in a number of ways, ranging from handwritten forms that detail any work done in an unstructured format to highly automated work orders. Furthermore, most of the data comes from routine operations, resulting in very imbalanced data sets. Despite these challenges, the idea of leveraging SCADA data for predictive maintenance has recently received more attention. Many obstacles, however, remain to be solved in present and future studies. The next paragraphs quickly assesses the most important studies in the field, which demonstrate its potential while also underlining the research hurdles.

For starters, supervised algorithms are used in a large percentage of articles (classification methods). For example, in [3], the main bearing fault is diagnosed using support vector machine (SVM) classifiers; and in [4], simultaneous multiple faults are again diagnosed using SVM. Despite the promising performance of supervised algorithms in research, their deployment in a real application requires domain adaptation techniques that provide the ability to train a model on one dataset (source) for which label is available and secure a good performance on another dataset (target) whose label is not available. Otherwise, it is nearly impossible to apply supervised approaches, as deriving labeled data sets from WT operational data is often difficult (because of the lack of standardization in maintenance records), time-consuming, and error-prone, resulting in a very imbalanced data set. Furthermore, domain adaptation techniques are crucial to deploy supervised approaches to wind parks where the defect has not already occurred. It is noteworthy that there is a consistent portion of the literature that fosters the use of domain adaptation techniques to deal with supervised learning for WT fault diagnosis, as in [5] where a WT gearbox and cross-bearing faults are studied through simulated data, and [6] where a bearing fault is studied through a test bench. Second, as indicated in the review study [7], a substantial number of sources confirm the conclusions using simulated SCADA data (as in [4]) or experimental data (from a test bench) as in [3]. Although this is understandable given that real SCADA data sets are frequently proprietary and not readily available to the scientific community, it is a significant disadvantage because data provided by test rigs or mathematical models may not generalize well to real-world settings [7]. Third, as indicated in the highly referenced study [8], where a detailed analysis regarding utilizing real SCADA data for WT CM is given, the bulk of the references analyzed based their conclusions on a very limited quantity of data, typically just 1 to 4 WTs. Again, this is a significant disadvantage, as it is unclear whether these tactics would scale successfully throughout the whole wind park. Forth, certain references, such as [9], offer tactics that result in a significant number of false alerts, making the contribution inconvenient in the real world, as it would result in alarm fatigue for operators. Fifth, a significant number of studies, such as [10] and [3], detect the fault with less than a week's notice, rendering them useless in a real-world situation where the plant operator needs at least months to program repair to concur with the availability of replacement parts as well as weather or production windows to minimize turbine downtime. Finally, several notable references employ completely unsupervised techniques that have been tested on real wind parks. For instance, in [11], the pitch system CM is stated using isolation forest and validated on ten WTs.

However, with advances in the field of deep learning, progressively, much of the research is focused on capturing relevant features using neural networks with deep hidden layers. Methodologies based on artificial neural networks (ANN) have been proposed, as in [12] and [13], as well as different variations of autoencoders (based on ANNs) have been widely studied. For example, in [14] a gearbox failure detection method is proposed based on a deep joint variational autoencoder. In [15] a multi-level-denoising autoencoder approach is stated and faults at the pitch system and drive train (vibration anomaly) are detected. In [16] a denoising autoencoder with temporal information is proposed to detect anomalies in the generator speed and gearbox filter. Note that in all aforementioned references time-series data is widely adopted, where long-term dependency is essential to form the classifiable features. Because SCADA data is a time series, temporal data is essential for constructing a prognosis model. Furthermore, understanding the changes and trends in variables over time is crucial to constructing the model architecture. That is, the model learns how previous data samples impact future data samples, which is the primary purpose of designing an early defect detection approach. Most basic designs, such as ANNs, do not take a prior data point into account when deciding the next; instead, the model learns from individual samples.

To better address that the traditional ANNs either rely on expert knowledge and handcrafted features or do not fully model long-term dependencies hidden in time-domain signals, recent works focus on the adoption of recurrent neural networks (RNNs). In this case, the recurrent neural network (RNN) model may be useful since it has a recurrent connection and can learn the influence of past and current inputs while forecasting the outcome. On the other hand, RNNs suffer from the problem of vanishing gradients that the LSTM and GRU models solve by employing gates to determine which information should be kept and which should be ignored. GRU networks have a few benefits over LSTM [17] such as a smaller number of parameters and a reduced computational cost, which are relevant in this application, which requires training with data over several years. In this work, the GRU neural network is selected because of its ability to forecast information from time series data (taking into account past information) and its simplicity (associated with a lower computational cost) compared to other types of RNNs. Finally, there are some relevant works related to WT fault diagnosis with RNNs. For example, in [18], fault diagnosis of WT based on long short-term memory (LSTM) networks (a particular type of RNN) is proposed. However, not only SCADA but also high-sampling vibration data is used

and validation is done only with simulated (not real) data. It is noteworthy the work in [19], where the main bearing failure (the fault of interest in the present work) is predicted based on LSTM with more than 90 days on average, but again not only SCADA data is used but also high-sampling vibration data that greatly helps to the obtained performance. Finally, the work in [20] detects the gearbox failure by means of GRU networks, but the used SCADA data is 1-minute averaged instead of the standard 10-minute averaged, thus giving extra predicting capabilities.

All that being said, the novelty of the proposed design is the following. A novel methodology to detect months in advance the main bearing failure (when the event that initiates/develops the failure releases heat) using only standard 10-minute SCADA data and based on a GRU (because of its ability to forecast information from time series data and its low computational cost in comparison to other RNNs) and a novel fault prognosis indicator (FPI) that precludes the problem of alarm fatigue. Furthermore, the stated methodology is validated with real (as opposed to simulated or experimental which greatly simplifies the problem) data from an in production wind farm. The methodology main features are the following: 1) Semi-supervised and based only on healthy data (precluding the problem of imbalanced data sets), thus expanding its range of application to any wind park (although no faults have yet occurred). 2) Reliable predictions with minimum false alarms. 3) Early warning months in advance, providing the plant operator time to program maintenance to match with replacement part availability, as well as weather or production windows to minimize turbine downtime. 4) Robust to seasonality and operating and environmental conditions. 5) Installing extra costly sensors is not needed, as only SCADA data is used. Therefore, the method can be applied to WTs already in operation for life-time extension services (this is relevant, as it is expected that 38 GW of wind parks in Europe will reach their life expectancy in the next five years). 6) Validated on real SCADA data from a wind park composed by nine WTs in production.

The following is how the rest of the article is structured. Section II provides a brief description of the wind park. Following this, the types of main bearing faults are introduced in Section III. The available SCADA data and work order records are then presented in Section IV. Section V discusses the suggested methodology. The findings and discussion in Section VI are used to analyze and convey the performance of the stated strategy. Finally, conclusions are drawn and recommendations for future works presented in Section VII.

## II. Wind park

The wind park is composed of WTs with a diameter of 101 m and 2300 kW of nominal power. Figure 1 displays the main elements of the WT. The energy production starts at an initial wind speed of 3 m/s and reaches its rated power at 12 m/s. Finally, when the wind speed is 20 m/s or more, the WT automatically stops using its braking system. In Table I, the technical specifications of these turbines are summarized.

Notably, the main bearing employed by these WTs is a double-spherical roller type, which are appropriate for applications involving very low speeds (up to 25 rpm) and high loads with varying directions; therefore, they are a reliable choice for the main bearing of a WT. Other characteristics of this type are robustness to withstanding loads of variable direction and low friction, which implies a low energy loss and a longer lifespan.

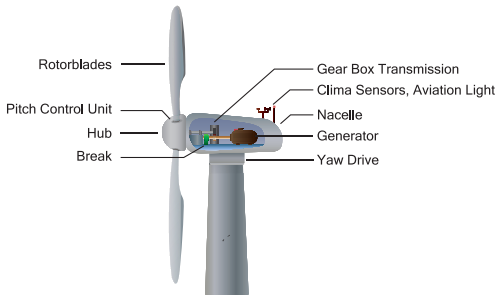



Fig. 1. Wind turbine parts.

TABLE I
Wind turbines' technical specifications.

| Technical specification | Value |
|---|---|
| Number of blades | 3 |
| Nominal power | 2300 kW |
| Rotor diameter | 101 m |
| Cut-in wind speed | 3 m/s |
| Rated wind speed | 12 m/s |
| Cut-out wind speed | 20 m/s |

## III. Main Bearing Faults

Bearings are supports or guides for machine elements such as shafts that need rotation or oscillation and can handle axial loads. Many sectors use them, including transportation, medical (surgical instruments, as well as diagnostic and laboratory equipment), and energy (wind turbines, and solar panels), among others. Typically, a bearing comprises four elements: cage, inner race, outer race, and rolling. During operations, each of these components is subjected to mechanical stress by at least one of the following forces: frictional, impact, centrifugal, and inertial. Therefore, any of them can suffer a breakdown.

The main bearing, which supports the main shaft of a WT, is a large component and this section attempts to offer an insight on the many ways of bearing failure as well as the enormous number of existing patterns, which make forecasting a bearing failure in advance difficult. Following the ISO 15243 standard, the SKF company classified a bearing's different failure modes as follows [21]: fatigue, wear, corrosion, electrical erosion, plastic deformation, and fracture and cracking. Each of these modes has different sources, behavior, and bearing damage. It is beyond the scope of this paper to elaborately explain these failure scenarios. However, a brief overview of some of the aforementioned failure modes is provided in the next paragraph to promote an understanding of the need for considering temperature variations while detecting a range of bearing defects.

First, fatigue is one of the most common failure modes in bearings. Subsurface-initiated fatigue and surface-initiated

fatigue are the two types of fatigue. Both submodes have an accumulation of residual stresses, causing the material to transition from a randomly oriented grain structure to fracture planes. This results in microcracks and beginning flaking, which emit heat. Second, the loss of a bearing surface is referred to as wear, and frictional heat is frequently present. Third, corrosion occurs due to the entry of moisture, water and aggressive contaminating liquids into the bearing. Corrosion leads to premature and extended spalling, as the material undergoes structural change and the load zone surfaces are reduced such that overloading occurs. Thus, the release of heat due to these three types of failure can be detected using temperature variables in the fault prediction model. Fourth, electrical erosion occurs because the electrical current flowing through the rings via rolling elements causes damage. The erosion can be divided into two types: excessive current erosion and current leakage erosion (produced by low-intensity current). Excessive current erosion heats up the material to temperatures between tempering and melting levels, forming craters in the rolling element and deforming the raceways. Again, the release of heat is the first symptom of this type of failure. Lastly, the bearing can fracture and crack, where bearing cracking can be caused by thermal factors. When two surfaces slide past each other, frictional heat is generated. If the sliding is considerable, the heat can crack in the perpendicular direction of the slip. Again, heat release is an initial starting point that can reveal the developing fault.

In summary, all bearing failures have a starting point from which an anomalous behavior begins until the end point, which is when the bearing eventually fails. Symptoms of anomalous behavior include a change in the main bearing temperature, among others. Heat release is a widespread symptom, so this work aims to predict the fault by detecting this first event. Thus, an early warning can be triggered, notifying that within a certain number of months, that element will seriously fail.

## IV. SCADA Data

In this study, the SCADA data is obtained from nine in-production WTs. The available data spans from January 1, 2015, up to December 31, 2018, and shows the continuous operation of each WT every 10 minutes. The SCADA data contain a variety of measurements that may be divided into five categories: environmental (see Table II), electrical, component temperature (see Table III), hydraulic, and control variables. The data is obtained with a 1 Hz sampling frequency and stored on a 10-minute average. The mean, standard deviation, maximum and minimum values are available for each sensor. This work focuses only on the mean values; therefore, the description of the variables in Tables II-III refers to its mean value. Furthermore, only some exogenous variables (environmental) and some temperature variables are employed by the proposed strategy; thus, in this section, only these two sets of variables are explained in depth.

Table II displays various environmental sensors from the SCADA data. In this, the ambient temperature and the wind speed can be found. The first one affects the temperature of all subsystems that change significantly between summer and winter. Finally, given its direct impact on the machine's operation, wind speed is the most crucial exogenous variable associated with a WT.

TABLE II
Environmental variables.

| Variable | Description | Units |
|---|---|---|
| wtc_AmbieTmp_mean | Ambient temperature | °C |
| wtc_PrWindSp_mean | Primary wind speed | m/s |
| wtc_AcWindSp_mean | Actual wind speed | m/s |
| wtc_PriAnemo_mean | Primary wind speed of anemometer | m/s |

TABLE III
Component temperature variables.

| Variable | Description | Units |
|---|---|---|
| wtc_A1IntTmp_mean | Internal temperature A1 | °C |
| wtc_BrkTmpGn_mean | Brake generator temperature | °C |
| wtc_GenBeRTm_mean | Generator bearing temperature | °C |
| wtc_GeOilTmp_mean | Gearbox oil temperature | °C |
| wtc_HSGenTmp_mean | HS generator temperature | °C |
| wtc_HSRotTmp_mean | HS rotator temperature | °C |
| wtc_HubTemp_mean | Hub temperature | °C |
| wtc_NacelTmp_mean | Nacelle temperature | °C |
| wtc_MainBTmp_mean | Main bearing temperature | °C |
| wtc_HydOilTm_mean | Oil temperature for hydraulic system | °C |

Table III shows some variables related to temperature. The main bearing defect is the subject of this research. As a result, it is critical to monitor the temperatures of nearby components, such as the main bearing temperature itself, the gearbox oil temperature, and the generator bearing temperature.

Finally, in addition to the SCADA data, information is available on maintenance and repair actions (work orders) for the different WTs. These data provide the following information: failure type, failure date, work order date, affected subsystems, and actions taken. This information is used in this work to model and test whether the proposed methodology can predict the appearance of main bearing faults months in advance.

## V. Fault Detection Methodology

This section comprehensively describes the stated fault detection strategy. First, the data preprocess, crucial when dealing with real data, is described in Section V-A. Next, the data split is described in Section V-B. This section emphasizes the GRU model strategy and the way to deal with seasonality. Then, in Section V-C, the importance of data normalization is stated, as well as the used normalization technique. Next, in Sections V-D and V-E, the normal behavior model (NBM) -WT's normal/healthy behavior- is constructed based on a GRU neural network, and the GRU architecture and the hyperparameter set-up are comprehensively explained. When healthy data is supplied, the GRU model is trained to predict the main bearing temperature in a certain time step from the input variables, functioning as a virtual sensor. Then, the proposed fault prognosis indicator (FPI) is stated in Section V-F, which employs a moving average strategy to diminish the number of false alerts, avoiding alarm fatigue to the wind park operator.

The proposed methodology can detect any failure associated to the main bearing that initiates or develops with

TABLE IV
SELECTED SCADA VARIABLES.

| Variable | Description | Units |
|---|---|---|
| wtc_MainBTmp_mean | Mean main bearing temperature | °C |
| wtc_GenBeRTm_mean | Mean generator bearing temperature | °C |
| wtc_GeOilTmp_mean | Mean gearbox oil temperature | °C |
| wtc_PrWindSp_mean | Mean primary wind speed | m/s |

an associated heat release. That is because the conceived methodology relies on SCADA data associated to temperature variables to detect the failure. Note that the low sampling rate of SCADA data (10-minute average) hinders information in variables with a fast dynamic (e.g., vibrations), however as temperature variables have a slow dynamic their SCADA data still contains relevant information. Finally, as indicated in Section III, the vast majority of different main bearing failure modes are associated to heat release, thus allowing the proposed methodology to be widely applicable.

### A. Data Preprocess

In Section III, the different types of main bearing failure are described, noting that a wide variety of failure modes induce an increase in temperature when the fault initiates. Consequently, in this work, the temperatures of the components near the main bearing are employed in conjunction with the ambient temperature to preclude data seasonality. The selected variables are shown in Table IV, namely, the main bearing temperature, the generator bearing temperature, the gearbox oil temperature, and the wind speed. It is noteworthy that the ambient temperature is subtracted from all variables related to temperature to avoid the problem of seasonality.

Real data is noisy; thus, a data cleaning strategy is required. Particularly, in this work, a data imputation strategy is used for missing data. Techniques based on statistical variables such as mean, median, and mode are not considered because they could generate a bias in the data's mean and deviation. [22]. Alternatively, the use of the piecewise cubic Hermite polynomial interpolation is proposed [12], as it contributes a continuous first derivative function and maintains the shape and monotonicity. Missing data at the beginning and end of the data set are filled in with their closest value.

### B. Data Split: Train, Validation, and Test

To build a deep learning model, the following steps are taken: i) dividing the data into training, validation, and testing data sets; (ii) training and validating the model using the train and validation data sets, respectively; and (iii) evaluating the obtained model on the test set to investigate the performance of the model predictions when using the new data (that the model has never seen before).

The approach used for data splitting is crucial in developing a machine or deep learning model, as it has a major impact on the overall model. The following assumptions are made to separate the data in this work. To begin with, only healthy data is utilized to train and validate the model because the major goal is to build WT's NBM. Additionally, data from the WT's entire region of operation (range 1: when the wind speed is lower than the cut-in wind speed; range 2, when the wind speed is between the cut-in and rated wind speed and; range 3, when the wind speed is above the rated wind speed) is used to train and validate the model. Thus, the model generated is robust to the WT's varied operating and environmental conditions. Finally, to make the model resilient to seasonality, the training and validation data must span at least one full year. In summary, the data are split in this manner. Data from January 2015 to September 2017 (144576 samples) are used for training, data from October 2017 to December 2017 (13248 samples) are used for validation, and data from January 2018 to December 2018 (52560 samples) are used for testing.

### C. Data Normalization

The data for the selected variables originate from several sources; hence, their order of magnitude differs. Scaling the data is widely advised by machine/deep learning approaches to achieve higher performance of the employed optimization algorithm. Based on SCADA data's nature for the specific wind park under study, these data have no significant outliers. Thus, the min-max scaling is directly selected (since there is no problem with outliers' sensibility that this kind of scaling can suffer) that scales data in the [0,1] range using a linear transformation from the initial set's range. Thus, this technique delivers data that are normalized exactly to the same interval for all used variables.

### D. Brief introduction to GRU neural networks

The deep learning model used in this work is a GRU neural network. Although this paper does not intend to provide a detailed understanding of GRU networks, they are briefly explained below to introduce the used notation and render this paper self-contained and understandable.

Cho et al. [23] proposed GRUs to allow relationships in an adaptable manner across temporal scales [24]. The GRU workflow resembles that of a simple recurrent neural network (RNN). It processes the input vector $x_t$ and the hidden state $h_{t-1}$ from the previous timestamp through the gates at each timestamp $t$. It then produces a new hidden state $h_t$, which contains both short- and long-term states and is transmitted to the next timestamp. Regarding the manner in which the flow of information is regulated by gating units, the GRU is comparable to a long short-term memory (LSTM) unit and is often thought of as a simpler version of LSTM. The amount of gating units and the state vector are the main distinctions between them. The GRU only has two gating signals (update and reset), whereas LSTM has three (input, forget, and output). Therefore, the GRU has fewer parameters and lower computational cost. Note that the GRU also combines both short-term and long-term state vectors into a single vector; see [24], [25].

In Figure 2, the gated GRU architecture is shown. The input and output vectors are $x_t$ and $y_t$, the reset and update gates are $r_t$ and $z_t$, and the activation and candidate activation are $h_t$ and $\tilde{h}_t$, respectively. The following are the definitions of the two gates:

$$z_t = \sigma(W_z x_t + U_z h_{t-1} + b_z), \tag{1}$$

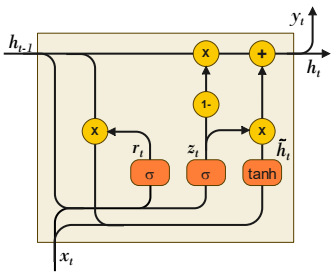



Fig. 2. Architecture of GRU cell.

$$r_t = \sigma(W_r x_t + U_r h_{t-1} + b_r). \tag{2}$$

The update gate, $z_t$, controls which bits of the long-term state should be inserted and which parts should be removed. The reset gate, $r_t$, determines which part of the prior state is displayed to the main layer or candidate activation, ($\tilde{h}_t$). Then, the GRU model can be formulated as follows:

$$\tilde{h}_t = \tanh\left(W_h x_t + U_h(r_t \cdot h_{t-1}) + b_h\right), \tag{3}$$

$$h_t = (1 - z_t) \cdot h_{t-1} + z_t \cdot \tilde{h}_t, \tag{4}$$

where the candidate activation, $\tilde{h}_t$, analyzes the input vector $x_t$ and the previous short-term $h_{t-1}$, and only the most relevant parts are stored in the vector $h_t$. On the other hand, the activation, $h_t$, represents the output state of the model and, as stated before, it contains the short-term and long-term state. From the above equations,

- $W_z, W_r, W_h$ are the weight matrices for each of the three layers in relation to the input vector $x$.
- $U_z, U_r, U_h$ are the weight matrices for each of the three layers in relation to the prior short-term state, $h_{t-1}$.
- $b_z, b_r, b_h$ are the bias terms for each of the three layers.

The GRU network architecture utilizes the four variables displayed in Table IV. The mean main bearing temperature is the output of the GRU model at time $t$, and the inputs are the mean generator bearing temperature, the mean gearbox oil temperature, and the mean primary wind speed. Thus, there are three inputs to the network and one output, i.e., the network is considered as a many-to-one structure. The inputs are given as a sequence of 144 consecutive time steps, that is, data that comprise 24 hours (recall samples are given each 10 minutes).

The GRU architecture and selected hyperparameters are comprehensively explained in the next section.

### E. GRU proposed architecture

The GRU architecture requires a set of different hyperparameters for its optimal performance according to the problem to be solved. This section describes the selected hyperparameters in detail.

First, the number of hidden layers, that is, the layers between the input and output of the network, which contain the GRU cells, is stated. In this work, it is defined as one single hidden layer, as testing with two layers did not improve performance in relation to early fault detection and thus did not justify the additional computational cost.

Second, the hidden size determines the number of features in the hidden state. In other words, it defines the number of neurons or GRU cells in the hidden layer. This hyperparameter determines the learning power of the model. In this paper, 128 neurons are selected in the hidden state due to the amount of samples that the model has to process.

Third, the batch size is defined in 128 samples, where each sample contains 144-time steps. An initial learning rate of 0.001 is defined, and 50 epochs is used to train the network.

Finally, a loss function must be defined. In this work, the mean squared error (MSE) is employed, as large errors are important (may represent a fault) with respect to small errors (maybe due to model error). To sum up, Table V details the hyperparameter values for the proposed GRU architecture.

In a nutshell, the diagram shown in Figure 3 summarizes the stated approach from the initial data preprocess, through data imputation, data normalization, and, finally, the training stage of the GRU model.

TABLE V
SETUP OF GRU HYPERPARAMETERS.

| Hyperparameter | Value |
|---|---|
| Number of hidden layers | 1 |
| Number of neurons in the hidden state | 128 |
| Batch size | 128 |
| Epoch size | 50 |
| Initial learning rate | 0.001 |
| Loss function | MSE |

### F. Fault Prognosis Indicator (FPI)

The FPIs are typically defined using a residual and establishing a decision threshold. When the residual of a sample exceeds the threshold, an alert is activated. In this work, the natural residual to be employed would be the square of the difference between the real SCADA main bearing temperature, $T$, and that predicted by the GRU network, $\hat{T}$, i.e., $(T - \hat{T})^2$. However, when using this residual straightforward to define a threshold, an overwhelming number of false alarms (false positives) would be triggered, leading to alarm fatigue in wind turbine operators and rendering the method useless.

To prevent this issue, the long persistence in time of the residual over the threshold has to be monitored, to discriminate false alarms from real fault detection alarms. Thus, in this section, to accomplish this aim, an FPI is stated, which filters the initial obtained residual with a moving average (MA) to smooth the original spiky residual.

Remember that a MA of $k$ observations smooths a time series by computing the average of the $k$ most recent observations; as a result, each new entering observation drives the oldest in the group out of the computation [26]. According to [27], the MA may generate cyclic and trend-like displays even when the source data are separate random occurrences with a fixed mean. As a result, this property limits its use as a control mechanism, and the exponentially weighted moving average (EWMA) emerges. This has the property of allocating less weight to data as it ages. A point on an EWMA chart

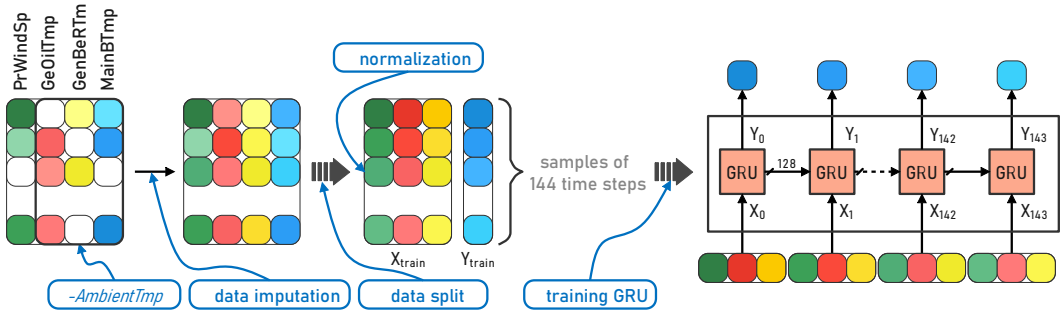



Fig. 3. Flowchart for the first phase of the proposed methodology until the training of the GRU model.

can be assigned a long or short memory. The EWMA is the best depicted one-time position ahead of the most recent observation; consequently, this statistic may be used to forecast the future observation. In this work, the EWMA is selected to be applied to the prediction errors as follows,

$$\text{EWMA} = \hat{T}_{t+1} = \hat{T}_t + \lambda e_t = \hat{T}_t + \lambda(T_t - \hat{T}_t), \quad (5)$$

where $\hat{T}_{t+1}$ is the predicted value at time $t+1$. Similarly, $\hat{T}_t$ is the predicted value at time $t$, $T_t$ is the measured real SCADA value at time $t$, and $e_t = T_t - \hat{T}_t$ is the prediction error at time $t$. Finally, $\lambda$ is a parameter ($0 < \lambda < 1$) that determines the memory depth of the EWMA. This parameter is empirically selected using its relation with the span ($s \geq 1$) [28]. Equation 6 shows this relation:

$$\lambda = \frac{2}{s+1}. \quad (6)$$

Finally, after averaging the initial residuals with the EWMA, a threshold is defined using the training and validation data. In this study, the threshold is the limit where a residual would be considered within normal behavior, so the mean $\mu_h$ and standard deviation $\sigma_h$ (this must not be confused with the sigmoid function) of the EWMA residuals of training and validation data must be taken into account. It is important to emphasize that, since the aim is to model the WT's normal behavior, the $\mu_h$ and $\sigma_h$ are obtained only from healthy data (from training and validation). Finally, the threshold is designated:

$$\text{threshold} = \mu_h + \kappa\sigma_h, \quad (7)$$

where $\kappa$ is a constant which establishes the threshold value. In this work, two values of $\kappa$ are used, as one defines a warning threshold and the other a fault alarm threshold. To sum up, Figure 4 shows the second phase of the proposed methodology, which details the computation of the warning and alarm thresholds after the EWMA filtering of the prediction error.

In the next section, an analysis is conducted on how the selection of the span number, $s$, and the $\kappa$ values to determine the thresholds (warning and alarm) affects the proposed methodology.

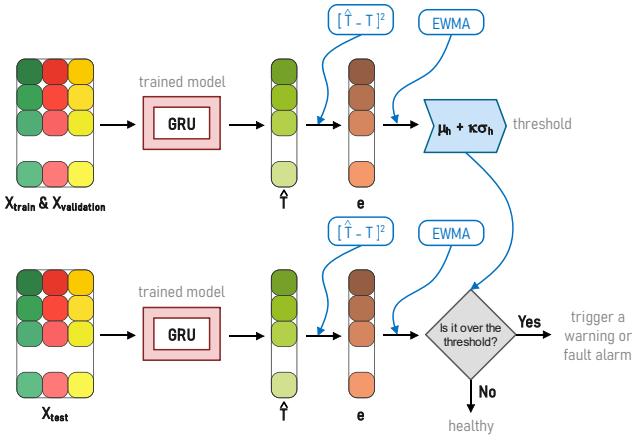



Fig. 4. Flowchart for the second phase of the proposed methodology, considering the computation of the threshold.

## VI. Results and Discussion

This section provides and examines the outcomes of the suggested fault prediction approach for a real-world wind park in operation.

Figure 5 shows the GRU estimations for two different WTs over the train and test data sets. Remember that the test data set contains the data of the year 2018. Figures 5 (a) and (b) display the estimated value $\hat{T}$ and the SCADA target of WT1, a healthy WT. The estimation is close to the real SCADA value in the two figures, and very few samples have disparate values. The performance of the GRU for this WT is shown in Figures 5 (c) and (d), where the absolute error between SCADA and prediction is graphed for the training and test sets. In contrast, Figures 5 (e) and (f) present the GRU estimated value and SCADA value for WT2, a WT that suffered a main bearing fault on June 11, 2018. The prediction over the training set is close to the target, contrary to the prediction over the test set. Particularly, in the test set, the prediction does not fit well for months before the fault occurs, and after it (when the fault is corrected), this error decreases. Figures 5 (g) and (h) show the residual on the training and test data sets, confirming the good performance over the training set, while in the test set, the residual increases a month before the fault and decreases

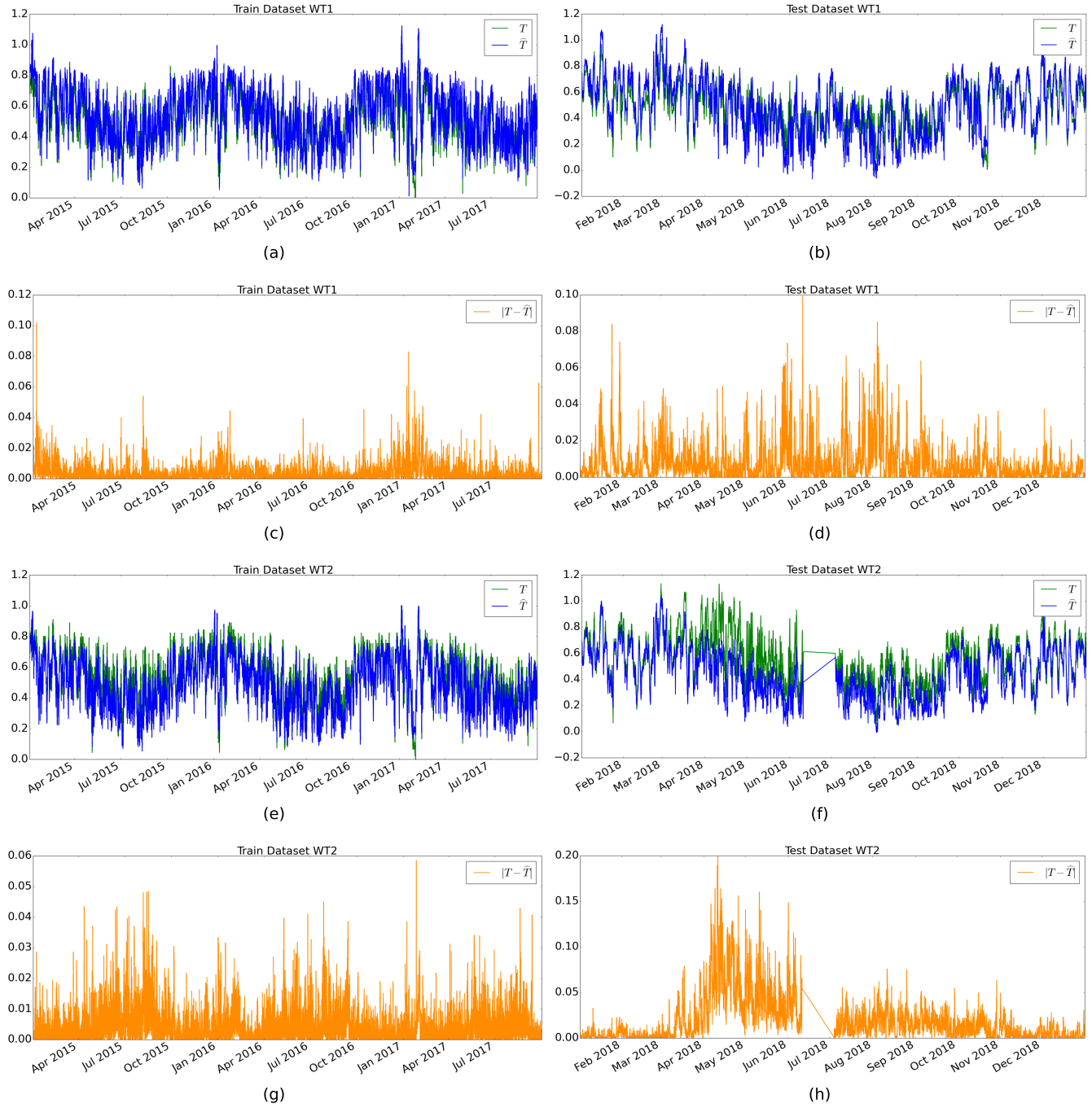



Fig. 5. (a) GRU estimated $\widehat{T}$, and SCADA real $T$, for WT1 over the train set. (b) GRU estimated $\widehat{T}$, and SCADA real $T$, for WT1 over the test set. (c) Error absolute value, $|T-\widehat{T}|$, for WT1 over the train set. (d) Error absolute value, $|T-\widehat{T}|$, for WT1 over the test set. (e) GRU estimated $\widehat{T}$, and SCADA target value $T$, for WT2 over the train set. (f) GRU estimated $\widehat{T}$, and SCADA target $T$, for WT2 over the test set. (g) Error absolute value, $|T-\widehat{T}|$, for WT2 over the train set. (h) Error absolute value, $|T-\widehat{T}|$, for WT2 over the test set.

after it. As comprehensively explained in Section V-F, the EWMA is applied over the prediction errors to measure the behavior and trend of the data as a function of persistence over time and, thus, to diminish the number of false alerts. The historical data in this study corresponds to samples taken every 10 minutes during a year, and even though temperature is a slow-changing-rate variable by nature, the prediction error must be filtered with an EWMA to obtain data with better defined patterns and behaviors. As shown in Equation 5, the depth of memory must be selected and, as stated in Equation 6, that parameter is established as a function of the span, $s$. For this study, two spans are considered and analyzed: $s$ =144 (one day), and $s$ =1008 (one week). These values are chosen to measure the persistence of the data trend (successive peaks, monotony) through time. For instance, if the EWMA is weekly based, perhaps the data trend is more visible and smoother

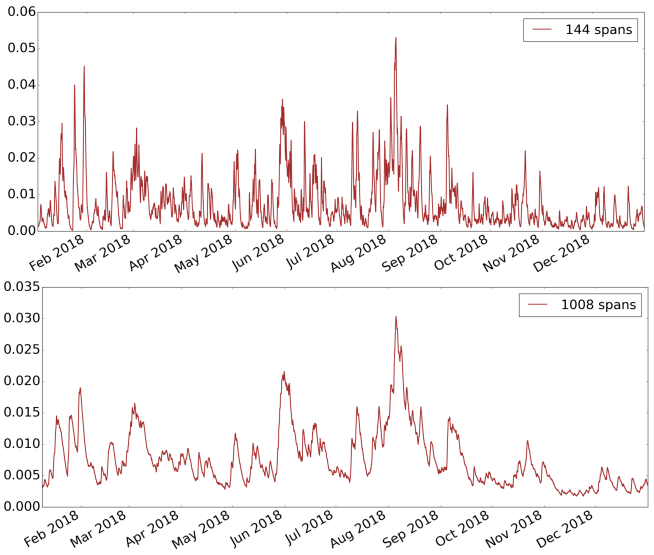



Fig. 6. EWMA filtered prediction errors for a WT, considering the span values $s = 144$, and $s = 1008$.

than if it is daily based. It helps to avoid the noise caused by isolated hourly peaks, which is very common in real SCADA data. Actually, this selection is influenced by the findings of McKinnon et al. [29]. Their research, on the influence of time history on WT failures using SCADA data, tests three distinct moving windows (MWs): daily, weekly, and monthly. In comparison to the others, the weekly MW has the best performance in identifying failures. On the one hand, a daily MW contains too much noise, leading to a large percentage of false alarms. On the other hand, a monthly MW removes much information and does not allow any specification of when an anomaly occurred.

Figure 6 shows the EWMA computation considering the established spans. After the EWMA computation, the definition of the thresholds is required, as stated in Section V. The aim is to propose two thresholds: one used as a warning and another to trigger a fault alarm. Table VI shows the detailed false-positive alarms over the training and validation dataset for each WT, considering the two proposed spans' values and two values for $\kappa$: 12 and 15. Thus, note that the selected values for $\kappa$ are just based on the observation of the training and validation dataset where the WTs are healthy. The value of $\kappa$ is set to minimize the number of false alarms over these datasets. Therefore, there is no information from the test set (or from the knowledge of the occurred fault on the test set) used to decide the $\kappa$ values. The results demonstrate that using 1008 spans no false-positive alarms are reached over the training and validation datasets.

Finally, Figure 7 shows the results on the test dataset for the entire wind park, obtained with the proposed FPI. A warning is triggered when the prediction error crosses $\mu_h + 12\sigma_h$ (green dotted line), while a fault alarm is triggered when the error exceeds $\mu_h + 15\sigma_h$ (red dotted line). Recall that the aforementioned mean and standard deviation are computed over the filtered prediction errors from train and validation. For WT2, both the warning and fault alarm were triggered

TABLE VI
FALSE-POSITIVE ALARMS (X-MARK) OVER THE TRAINING AND VALIDATION DATASETS.

| WT ID | 144 spans | | 1008 spans | |
|---|---|---|---|---|
| | $\mu_h + 12\sigma_h$ | $\mu_h + 15\sigma_h$ | $\mu_h + 12\sigma_h$ | $\mu_h + 15\sigma_h$ |
| **WT1** | | | | |
| **WT2** | | | | |
| **WT3** | X | | | |
| **WT4** | X | | | |
| **WT5** | X | X | | |
| **WT6** | | | | |
| **WT7** | X | | | |
| **WT8** | X | | | |
| **WT9** | X | X | | |

in the first week of June 2018. Considering that for WT2, the main bearing fault occurred on June 11, 2018, the early detection through this methodology is accomplished roughly two months in advance. Additionally, note that there is a clear trend of residual's increasing and then a decreasing. When bearing failure initiates (or develops) there is usually a brief heat release rendered as temperature increasing. As it is stated in [21], almost all bearing failure modes (excessive current, fatigue fracture, thermal cracking, etc.) are driven by unforeseen heat release. After that, the temperature goes back to normal i.e. crack is not growing. Thus, the methodology's approach is to predict this typical heat release in advance, before the bearing is entirely damaged. The case for WT3 and WT5 is interesting, as both a warning and a fault alarm were triggered in a week, but this WTs are healthy during 2018. Thus, they represent a false alarm. However, note that the alarm has a very short period length (less than a week), in both cases, in contrast to WT2 where the alarm is activated for two months. Therefore, in short, the model successfully detects the expected fault onset on the WT that contains the main bearing fault. As it has been before-explained, the onset defines the start point of the component's degradation through time. Thus, it is crucial that despite the residuals go back under the threshold, the triggered alarm must be kept active. It allows maintenance operators to plan ahead on-site revisions and give all needed attention to these component and perform required actions to extend its lifespan and not compromise WTs uptime.

## VII. CONCLUSIONS

In this work, an early fault detection approach for main bearing failures in WTs is devised and verified using a GRU neural network and just SCADA data. The model is built entirely from healthy data and is robust to all operational and environmental variations. The approach has been tested at a wind park with nine WTs. The findings show that the system produces minimal false alarms and that the defect of concern is predicted months in advance. Unfortunately, there is just one major bearing failure in the investigated wind park data, which is insufficient for statistical analysis. To investigate more extensively and draw conclusions such as a predicted time and confidence level, it is necessary to apply the model to other cases with this issue. Finally, it is convenient to study and test other sequential learning models that have been developed

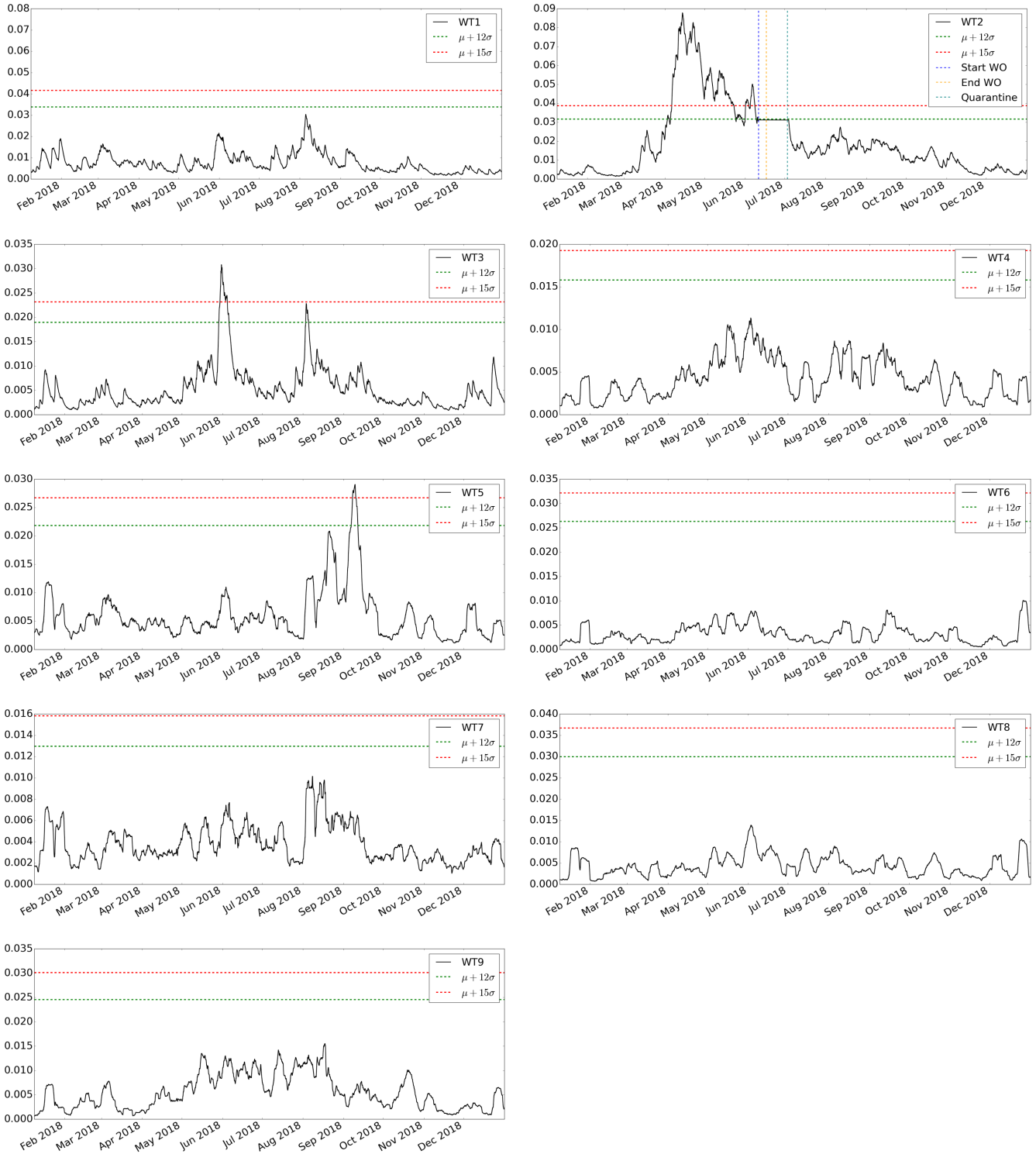



Fig. 7. EWMA on residual errors for the WT's test dataset (using 1008 spans), where the green line represents a fault warning and the red line indicates a definite fault.

in recent years. Specifically, time convolution neural networks and transformer networks, both of which have shown to be very powerful for some specific problems.

## REFERENCES

[1] WindEurope, "Getting fit for 55 and set for 2050, electrifying europe with wind energy," 2021.
[2] IRENA, "Renewable power generation costs in 2020," 2021.
[3] H. Zhao, Y. Gao, H. Liu, and L. Li, "Fault diagnosis of wind turbine bearing based on stochastic subspace identification and multi-kernel

support vector machine," *Journal of Modern Power Systems and Clean Energy*, vol. 7, no. 2, pp. 350–356, 2019.
[4] Y. Fu, Z. Gao, A. Zhang, and X. Liu, "Fault classification for wind turbine benchmark model based on hilbert-huang transformation and support vector machine strategies," in *2021 IEEE 19th International Conference on Industrial Informatics (INDIN)*. IEEE, 2021, pp. 1–8.
[5] Y. Qin, Q. Qian, J. Luo, and H. Pu, "Deep joint distribution alignment: A novel enhanced-domain adaptation mechanism for fault transfer diagnosis," *IEEE Transactions on Cybernetics*, 2022.
[6] Z.-H. Liu, B.-L. Lu, H.-L. Wei, X.-H. Li, and L. Chen, "Fault diagnosis for electromechanical drivetrains using a joint distribution optimal deep domain adaptation approach," *IEEE Sensors Journal*, vol. 19, no. 24, pp. 12 261–12 270, 2019.
[7] A. Stetco, F. Dinmohammadi, X. Zhao, V. Robu, D. Flynn, M. Barnes, J. Keane, and G. Nenadic, "Machine learning methods for wind turbine condition monitoring: A review," *Renewable energy*, vol. 133, pp. 620–635, 2019.
[8] 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, 2017.
[9] J. L. Godwin and P. Matthews, "Classification and detection of wind turbine pitch faults through scada data analysis," *IJPHM Special Issue on Wind Turbine PHM*, p. 90, 2013.
[10] K. Leahy, R. L. Hu, I. C. Konstantakopoulos, C. J. Spanos, A. M. Agogino, and D. T. O'Sullivan, "Diagnosing and predictingwind turbine faults from scada data using support vector machines," *International Journal of Prognostics and Health Management*, vol. 9, no. 1, 2018.
[11] C. McKinnon, J. Carroll, A. McDonald, S. Koukoura, and C. Plumley, "Investigation of isolation forest for wind turbine pitch system condition monitoring using scada data," *Energies*, vol. 14, no. 20, p. 6601, 2021.
[12] Á. Encalada-Dávila, B. Puruncajas, C. Tutivén, and Y. Vidal, "Wind turbine main bearing fault prognosis based solely on scada data," *Sensors*, vol. 21, no. 6, p. 2228, 2021.
[13] Z. Zhang *et al.*, "Automatic fault prediction of wind turbine main bearing based on scada data and artificial neural network," *Open Journal of Applied Sciences*, vol. 8, no. 06, p. 211, 2018.
[14] L. Yang and Z. Zhang, "Wind turbine gearbox failure detection based on scada data: A deep learning-based approach," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–11, 2020.
[15] X. Wu, G. Jiang, X. Wang, P. Xie, and X. Li, "A multi-level-denoising autoencoder approach for wind turbine fault detection," *Ieee Access*, vol. 7, pp. 59 376–59 387, 2019.
[16] G. Jiang, P. Xie, H. He, and J. Yan, "Wind turbine fault detection using a denoising autoencoder with temporal information," *IEEE/Asme transactions on mechatronics*, vol. 23, no. 1, pp. 89–100, 2017.
[17] A. Flores, H. Tito-Chura, and V. Yana-Mamani, "Wind speed time series prediction with deep learning and data augmentation," in *Intelligent Systems and Applications*, K. Arai, Ed. Cham: Springer International Publishing, 2022, pp. 330–343.
[18] J. Lei, C. Liu, and D. Jiang, "Fault diagnosis of wind turbine based on long short-term memory networks," *Renewable energy*, vol. 133, pp. 422–432, 2019.
[19] J. Herp, N. L. Pedersen, and E. S. Nadimi, "A novel probabilistic long-term fault prediction framework beyond scada data-with applications in main bearing failure," in *Journal of Physics: Conference Series*, vol. 1222, no. 1. IOP Publishing, 2019, p. 012043.
[20] L. Xiang, X. Yang, A. Hu, H. Su, and P. Wang, "Condition monitoring and anomaly detection of wind turbine based on cascaded and bidirectional deep learning networks," *Applied Energy*, vol. 305, p. 117925, 2022.
[21] "Bearing damage and failure analysis," https://www.skf.com/binaries/pub12/Images/0901d1968064c148-Bearing-failures—14219_2-EN_tcm_12-297619.pdf, accessed: 2021-08-12.
[22] Z. Zhang, "Missing data imputation: Focusing on single imputation," *Annals of translational medicine*, vol. 4, 2016.
[23] K. Cho, B. van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using RNN encoder–decoder for statistical machine translation," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Doha, Qatar: Association for Computational Linguistics, Oct. 2014, pp. 1724–1734. [Online]. Available: https://aclanthology.org/D14-1179
[24] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," *CoRR*, vol. abs/1412.3555, 2014. [Online]. Available: http://arxiv.org/abs/1412.3555
[25] R. Dey and F. M. Salem, "Gate-variants of gated recurrent unit (gru) neural networks," in *2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS)*, 2017, pp. 1597–1600.
[26] J. S. Hunter, "The exponentially weighted moving average," *Journal of Quality Technology*, vol. 18, no. 4, pp. 203–210, 1986. [Online]. Available: https://doi.org/10.1080/00224065.1986.11979014
[27] L. S. Nelson, "The deceptiveness of moving averages," *Journal of Quality Technology*, vol. 15, no. 2, pp. 99–100, 1983. [Online]. Available: https://doi.org/10.1080/00224065.1983.11978852
[28] P. Cisar and S. M. Cisar, "Ewma statistic in adaptive threshold algorithm," in *2007 11th International Conference on Intelligent Engineering Systems*, 2007, pp. 51–54.
[29] C. McKinnon, A. Turnbull, S. Koukoura, J. Carroll, and A. McDonald, "Effect of time history on normal behaviour modelling using scada data to predict wind turbine failures," *Energies*, vol. 13, no. 18, p. 4745, 2020.

## VIII. Biography Section

**Ángel Encalada-Dávila** received his bachelor's degree in Mechatronics Engineering in 2021 from Escuela Superior Politécnica del Litoral. He is currently working as a Data Scientist in an international data analytics company. Moreover, he is editor of the CIDiER 2021 Conference Proceedings (Book of Series "Green Energy and Technology", Springer). He is the author of 5 journal articles, and 7 conference papers.

**Luis Moyón** received his bachelor's degree in Mechatronics Engineering in 2022 from Escuela Superior Politécnica del Litoral. Currently, he works in an international software company and collaborates with Universidad Ecotec. His areas of research interest are robotics, automation and control systems, artificial intelligence, and software development.

**Christian Tutivén** received the B.E. degree in Telecommunications in 2007 from Universidad Católica Santiago de Guayaquil, Ecuador, and the Ph.D. in 2018 from the Universitat Politècnica de Catalunya, Barcelona, Spain. Since 2019, he has been with the Faculty of Mechanical Engineering and Production Science and Mechatronics Engineering at Escuela Superior Politécnica del Litoral, where he is an associate professor. He is the author of 10 journal articles, and 28 conference papers.

**Bryan Puruncajas** is a Mechatronics engineer, graduated from the Universidad de las Fuerzas Armadas ESPE, Ecuador, in 2015. He was granted a scholarship by the National Taipei University of Technology, Taiwan, to pursue a master's degree in 2016. He is currently a PhD student at the Universitat Politècnica de Catalunya, Spain, and lecturer at the Escuela Superior Politécnica del Litoral, Ecuador. He is the author of 4 journal articles, and 5 conference papers.

**Yolanda Vidal** received the B.E. degree in mathematics in 1999 and the Ph.D. degree in 2005 from the Universitat Politècnica de Catalunya (UPC), Barcelona, Spain. Since 2001, she has been with the Department of Mathematics and the Barcelona East School of Engineering (EEBE), at UPC, where she is an associate professor. She is an IEEE Senior member and is the author of 50 journal articles, and more than 100 conference papers.