END OF DEGREE PROJECT

Degree in Chemical Engineering

STRUCTURAL HEALTH MONITORING FOR OFFSHORE WIND TURBINE FOUNDATIONS THROUGH UNSUPERVISED AND SEMI SUPERVISED MACHINE LEARNING METHODS

Report and Annexes

Author: Clara Rull
Director: Yolanda Vidal
Call: May 2022

Abstract

The current climate crisis requires a shift towards renewable energies. Wind energy generation will play a major role. Offshore wind energy can provide greater output due to more predictable weather conditions compared to onshore wind energy and has one of the lowest lifecycle greenhouse gas emissions for any source of energy. Some of the difficulties in their operation and maintenance lie in the difficulty of accessing the site. Although remote monitoring has become standard in the industry, structural health monitoring and predictive maintenance still present some challenges.

Normally, most or all the available data are of regular operation, thus methods that focus on the data leading to failures end up using only a small subset of the available data. Furthermore, when there is no historical precedent of a type of damage, those methods cannot be used. In addition, offshore wind turbines work under a wide variety of environmental conditions and regions of operation involving unknown input excitation given by the wind and waves. Finally, supervised approaches rely on correctly labelling data, which is not possible in production conditions. Considering the difficulties, the stated strategy in this work is based on unsupervised and semi-supervised approaches and it works under different operating and environmental conditions based only on the output vibration data gathered by accelerometer sensors. The proposed strategy has been tested through experimental laboratory tests on a down-scaled model.

This project applies spectral entropy, a non-standard parameter in vibration analysis, to the studied models. Overall accuracies of 93,88% for Isolation Forest (a semi-supervised method), and 88,67% for One Class Support Vector Machine (a non-supervised method) can be achieved. The accuracies of both models increase to up to 100% when trained against a larger dataset of healthy samples, however achieving these results requires retuning for features and hyperparameters.

For all of this, the use of non-supervised and semi-supervised machine learning models is a realistic approach to structural health monitoring of offshore wind turbines and has obtained promising results when tested against an experimental dataset.

Resum

La crisi climàtica actual requereix un gir cap a les energies renovables. La generació d'energia eòlica hi jugarà un paper important. L'energia eòlica marina pot proporcionar una major producció degut a condicions climàtiques més previsibles en comparació amb l'energia eòlica terrestre i té una de les emissions de gasos d'efecte hivernacle de cicle de vida més baixes en comparació amb qualsevol font d'energia. Algunes de les dificultats en el seu funcionament i manteniment radiquen en la dificultat d'accés al lloc. Si bé la monitorització remot s'ha volgut estàndard a la indústria, la monitorització de la salut estructural i el manteniment predictiu encara presenta algunes dificultats.

Normalment, la majoria o totes les dades disponibles són de l'operació regular, per tant els mètodes enfocats en la utilització de les dades precedents a falles acabant utilitzant només un petit subconjunt de les dades disponibles. A més, quan no hi ha antecedents històrics d'un tipus de dany, no es poden utilitzar aquests mètodes. Encara, les turbines eòliques marines funcionen en una amplia varietat de condicions ambientals i regions d'operació que involucren una excitació d'entrada desconeguda proporcionada pel vent i les onades. Finalment, els enfocaments supervisats es basen en l'etiquetatge correcte de les dades, que no és possible en condicions de producció. Tenint en compte les dificultats, l'estratègia establerta en aquest treball es basa en enfocaments no supervisats i semi-supervisats i funciona sota diferents condicions ambientals i operatives basant-se únicament en les dades de vibració de sortida recopilades pels acceleròmetres. L'estratègia ha estat provada a través d'assajos experimentals de laboratori en un model a escala reduïda.

Aquest projecte aplica l'entropia espectral, un paràmetre no estàndard en l'anàlisi de vibracions, als models estudiats. Es poden aconseguir precisions generals del 93,88 % per a 'Isolation Forest' (un mètode semi supervisat) i del 88,67 % per a 'One Class Support Vector Machine' (un mètode no supervisat). Les precisions dels dos models augmenten fins al 100 % quan s'entrenen amb un conjunt de dades més grans de mostres sanes; tanmateix, per aconseguir aquests resultats és necessari tornar a ajustar les 'features' i els hiperparàmetres.

Per tot això, l'ús de models no supervisats i semi supervisats és un enfoc realista per la monitorització estructural de les turbines de vent marines obtenint resultats prometedors quan s'ha provat contra un conjunt de dades experimental.

Resumen

La actual crisis climática requiere un giro hacia las energías renovables. La generación de energía eólica jugará un papel importante. La energía eólica marina puede proporcionar una mayor producción debido a las condiciones climáticas más predecibles en comparación con la energía eólica terrestre y tiene una de las emisiones de gases de efecto invernadero de ciclo de vida más bajas en comparación cualquier fuente de energía. Algunas de las dificultades en su funcionamiento y mantenimiento radican en la dificultad de acceso al sitio. Si bien el monitoreo remoto se ha vuelto estándar en la industria, el monitoreo de la salud estructural y el mantenimiento predictivo aún presenta algunos desafíos.

Normalmente, la mayoría o todos los datos disponibles son de operación regular, por lo que los métodos que se enfocan en los datos que conducen a fallas terminan usando solo un pequeño subconjunto de los datos disponibles. Además, cuando no existe un antecedente histórico de un tipo de daño, no se pueden utilizar esos métodos. Por añadido, las turbinas eólicas marinas funcionan en una amplia variedad de condiciones ambientales y regiones de operación que involucran una excitación de entrada desconocida proporcionada por el viento y las olas. Finalmente, los enfoques supervisados se basan en el etiquetado correcto de los datos, que no es posible en condiciones de producción. Teniendo en cuenta las dificultades, la estrategia establecida en este trabajo se basa en enfoques no supervisados y semi supervisados y funciona bajo diferentes condiciones operativas y ambientales basadas solo en los datos de vibración de salida recopilados por los sensores del acelerómetro. La estrategia propuesta ha sido probada a través de pruebas experimentales de laboratorio en un modelo a escala reducida.

Este proyecto aplica la entropía espectral, un parámetro no estándar en el análisis de vibraciones, a los modelos estudiados. Se pueden lograr precisiones generales del 93,88 % para 'Isolation Forest' (un método semi supervisado) y del 88,67 % para 'One Class Support Vector Machine' (un método no supervisado). Las precisiones de ambos modelos aumentan hasta un 100 % cuando se entrenan con un conjunto de datos más grande de muestras sanas; sin embargo, para lograr estos resultados es necesario volver a ajustar las 'features' y los hiperparámetros.

Por todo esto, el uso de modelos de aprendizaje automático no supervisados y semi supervisados es un enfoque realista para el monitoreo de la salud estructural de las turbinas eólicas marinas y ha obtenido resultados prometedores cuando se prueba con un conjunto de datos experimental.

Glossary

AI: Artificial intelligence

CF: Crest factor

ML: Machine learning

O&G: oil and gas

O&M: operation and maintenance

PP: Peak-peak

RBF: Radial basis function

RBM: reliability-based maintenance

RMS: Root mean squared

SHM: Structural health monitoring

SVM: Support vector machine

TPM: total productive maintenance

WT: Wind turbine

ZP: Zero-peak

Index

Index of figures

Index of tables

1. Introduction

Within the current climate crisis that the world is facing right now it only makes sense to devout as many resources as possible to make sure that humanity can thrive in a sustainable way. The use of renewable sources of energy is one way to reduce the environmental impact of the current demand for energy.

Offshore wind turbines make use of the high-speed wind currents found at sea. However, their locations also provide a challenge as they can be difficult and costly to service. For this reason, the advances in predictive maintenance and structural health monitoring (SHM) have been of great interest in the industry.

The implementation of the latest maintenance techniques and research to monitor the structural integrity of offshore wind turbines will translate into an optimization of resources. In the last years, the offshore wind turbine industry has started to integrate extensive predictive maintenance plans in order to maximise the performance, minimize costs and increase profits.

1.1. Goals of the project

The main goal of this project is to study the applicability of non-supervised and semi-supervised artificial intelligence models in health monitoring of offshore wind turbines.

This project will introduce wind energy and offshore wind turbines and their relevance. The project will also include an overview of maintenance theory and vibration analysis in wind turbines. There will be a brief introduction to artificial intelligence, machine learning and some non-supervised and semi-supervised methods. Vibration analysis and the previously introduced machine learning methods will be applied to a SHM dataset to create a health monitoring model.

2. Wind energy

One of climate change's main drivers are CO2 emissions. Energy consumption was responsible for 76% of CO2 emissions globally (37.2 GtCO2e) in 2018. Within the energy sector, heat and electricity generation was responsible for 15.6 GtCO2e in 2018, or 31.9% of total greenhouse gas emissions.

Emissions from electricity and heat generation increased 78% from 1990 to 2013, but then dropped by 2.4% between 2013 and 2016. The decrease was driven by various factors, including a shift to natural gas from coal and increased use of renewables. (World Resources Institute 2021) Renewable energy sources have a much lower carbon footprint in comparison to fossil fuels. (Schlömer S. 2014)

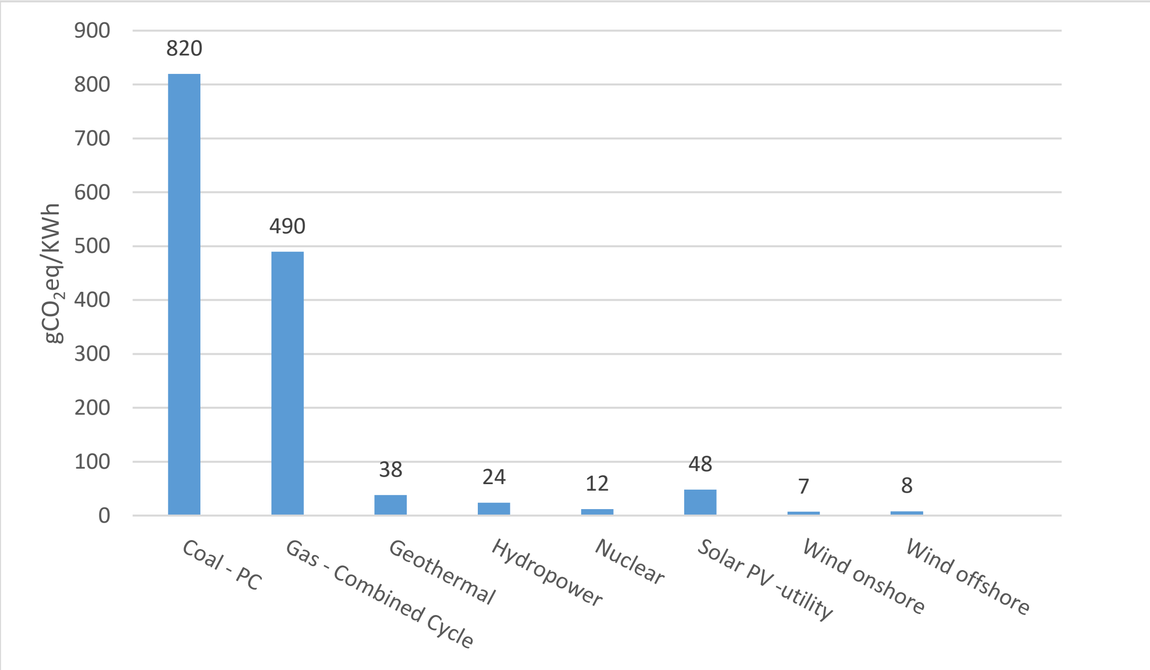

Figure 1 Median emissions of selected electricity supply technologies

In 2019 wind energy generated 1427TWh, a 5% of the total electrical power generated globally. The output has grown year after year for the past three decades and is expected to continue this trend. (IEA 2019)

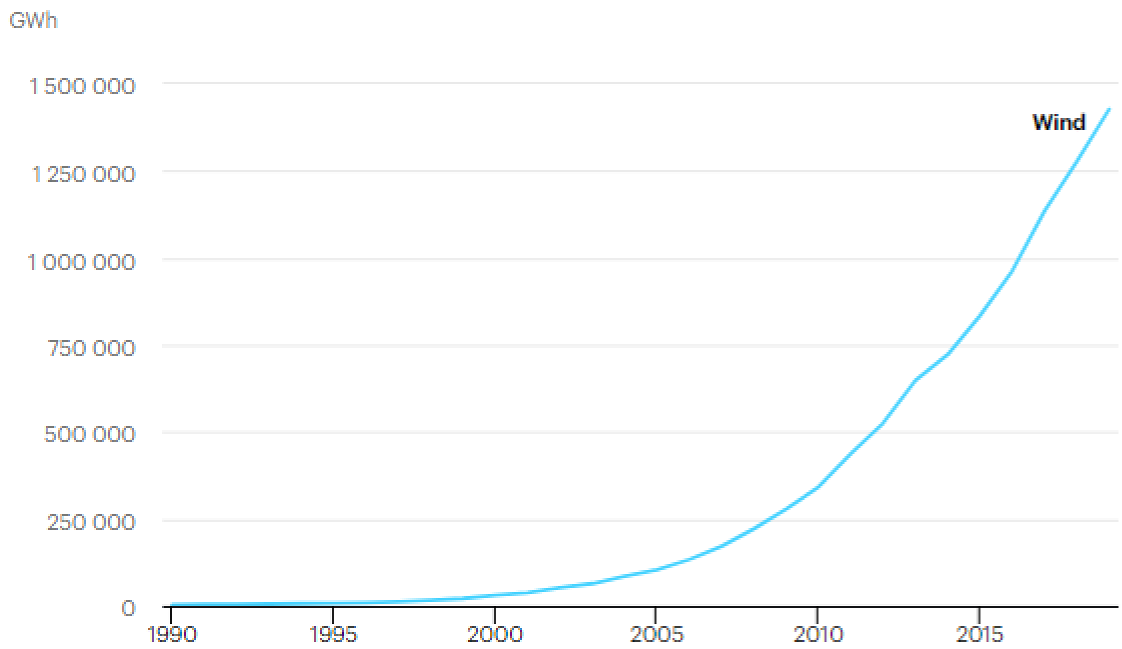

Figure 2 Wind electricity generation, World (1990-2019) (IEA 2018)

2.1. Off-Shore wind turbines

Offshore wind power is a subset of wind power, where wind turbines are placed on bodies of water (usually seas or oceans, but also in lakes). Offshore wind turbines benefit from higher and more predictable wind speeds. However, they also present higher operation and maintenance (O&M) costs compared to onshore wind turbines.

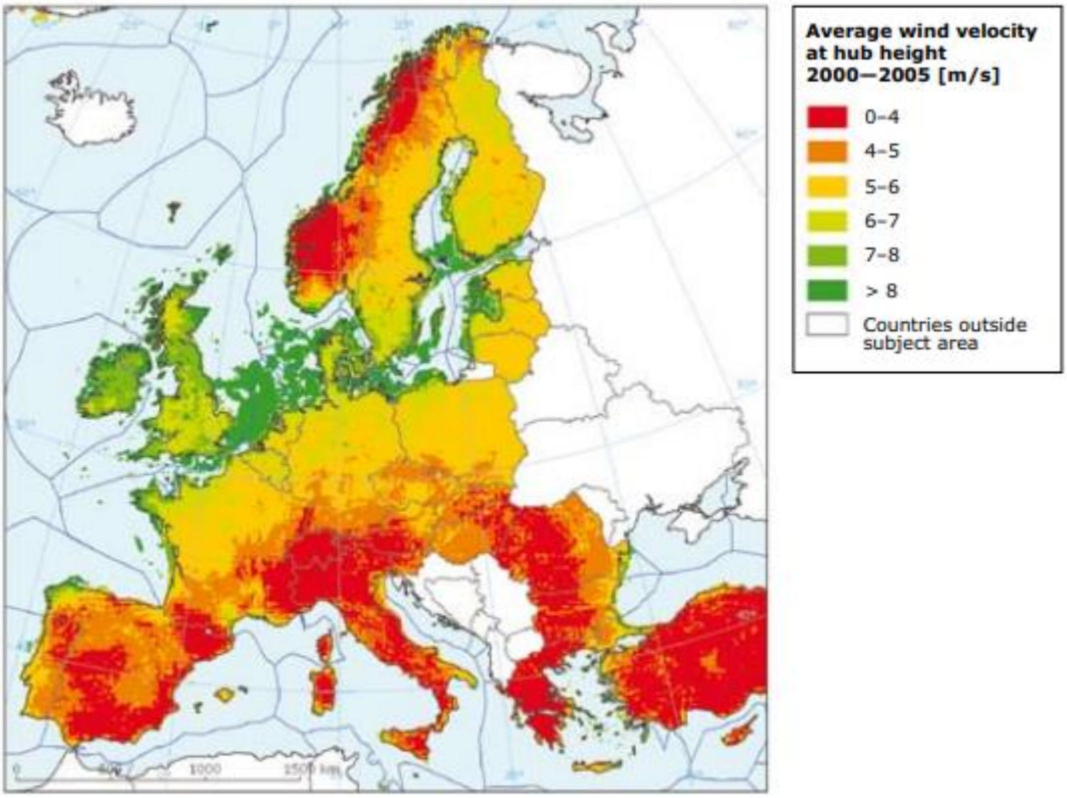

Figure 3 ECMFW wind field data after correction for orography and local roughness (European Environment Agency 2009)

Globally, in 2020 offshore wind capacity passed 35GW and now represents 4.8% of total cumulative wind capacity. GWEC Market Intelligence expects that over 469 GW of new onshore and offshore wind capacity will be added in the next five years - that is nearly 94 GW of new installations annually until 2025, based on present policies and pipelines. (Global Wind Energy Council 2021)

Currently Europe has over 25GW of offshore wind energy capacity, with a total of 5402 grid connected wind turbines delivering power from 116 offshore wind farms in 12 European countries.

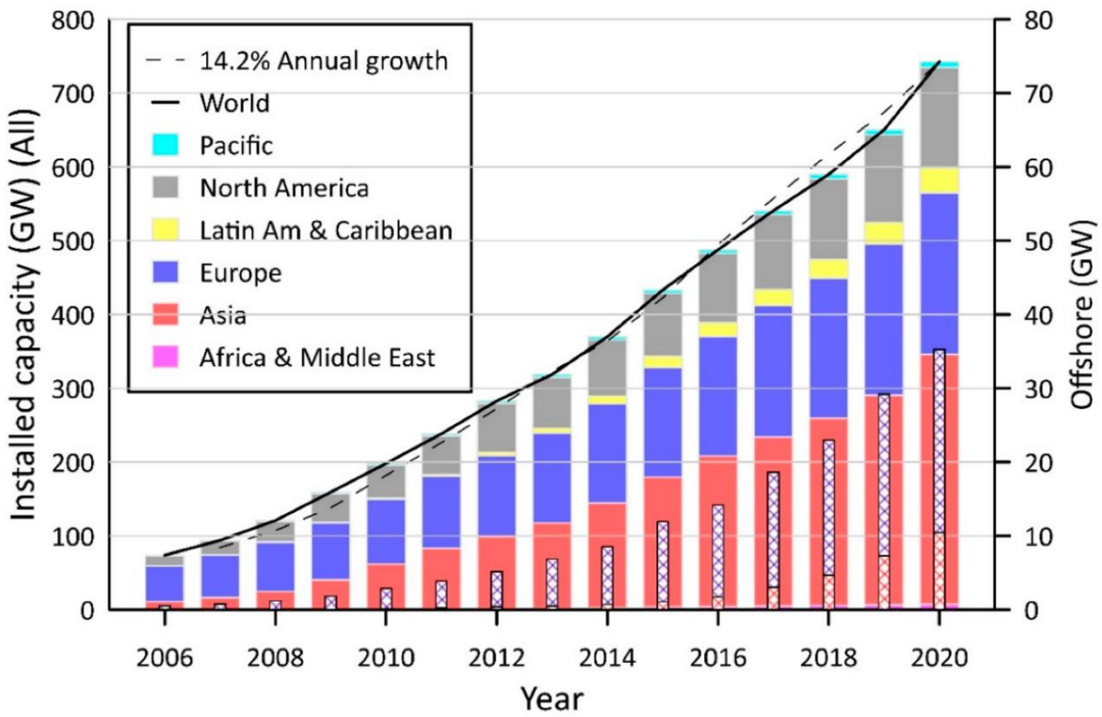

Figure 4 Total installed capacity (IC) of offshore wind energy by region (Barthelmie and Pryor 2021)

2.2. Components of Wind Turbine Installations

This project focuses on horizontal axis upwind turbines installed offshore. In horizontal axis wind turbines, the axis that is connected to the main bearing for electricity generation is parallel to the ground and the main rotor is directed towards the wind.

The design of wind turbines in offshore must consider the harsher conditions compared to onshore wind turbines:

- Strong currents and waves.
- Corrosive environments.
- Harsh climatological conditions, stronger storms, and winds.

Typically, the turbine manufacturer provides the roto-nacelle assembly and the tower. The support structure and base are chosen according to the needs of the project (Bhattacharya 2019).

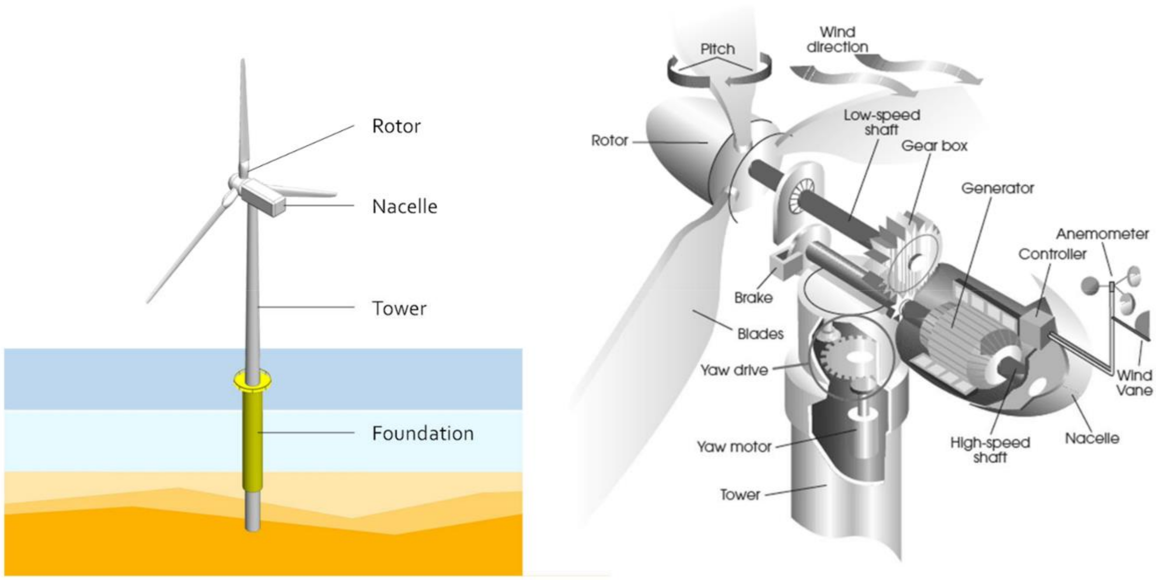

Figure 5 Components of an offshore wind turbine (Li, et al. 2022)

2.2.1. Drivetrain

In wind turbines the power is transmitted from the rotor to the generator through the system composed of the main shaft, friction connection, multiplying gearbox and a flexible coupling. This whole system is known as the drivetrain. (Michal, Gawarkiewicz and Wasilczuk 2015)

The drivetrain may have a gearbox between that main rotor and generator to increase the rotational speed of the rotor to generator speeds. This is the most common design as it allows for use of standard components. Less frequently, drivetrains may be gearless, requiring a multi-pole generator. (Barszcz 2019)

2.2.2. Foundation

The foundations of wind turbines can be classified in two main groups: grounded systems and floating systems. Foundations can be classified as shallow base or deep base. Some examples are the following:

- Monopile structures are deep base structures, where a long steel cylinder of 3 to 7m of diameter is placed up to 40m into the ocean floor. These are the most common kind of foundation.
- Shallow foundation structures, designed to avoid tensions between the foundation structure and the seabed, in order to avoid torsion.

- Suction based foundations are more shallow than monopolar structures but deeper than gravity-based ones. They are formed by a tubular structure topped by a circular side that acts like a suction cup, attaching to the seabed.

2.2.3. Structural Support

Offshore wind turbines require more robust support structures than onshore wind turbines, due to the extreme conditions at sea.

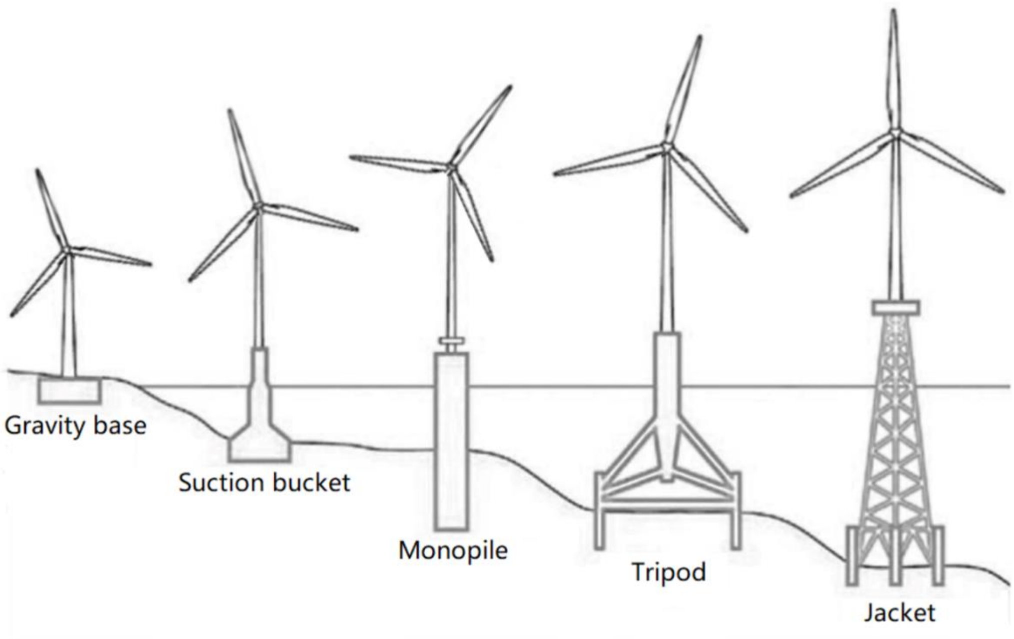

Figure 6 Main types of offshore wind turbine foundations (Xie and Lopez-Querol 2021)

Support structures may be: monopile (essentially an extension of a pile foundation), tripile, tripod, gravity based/shallow foundation or jacketed/latticed. This project focuses on the SHM of jacket structures, more in-depth description of them can be found in the next section.

Jacketed or latticed structure

The historical precedent for jacket structures in offshore wind foundations are gas and oil extraction platforms. However, their use as a structural support for wind turbines presents some specific particular challenges, the most prominent one being a significant contribution to vibrations due to the impact of wind, while in oil and gas (O&G) extraction platforms waves are the most significant vibration contribution.

These structures typically have 4 supports, which will have pile, gravity bases or suction caissons. There has been increased interest in the use of 3-legged jacket structures as they present lower costs.

As for the dimensions of the jacket structure, more traditional approaches rely on integrated aeroelastic models with a simplified representation of the foundation for calculations. (Agustyn, Nielsen and Pedersen 2017) Some models for a systematic approach for the predesign phase have been developed, but further work from experienced professionals is still required for a complete design. (Häfele, et al. 2018)

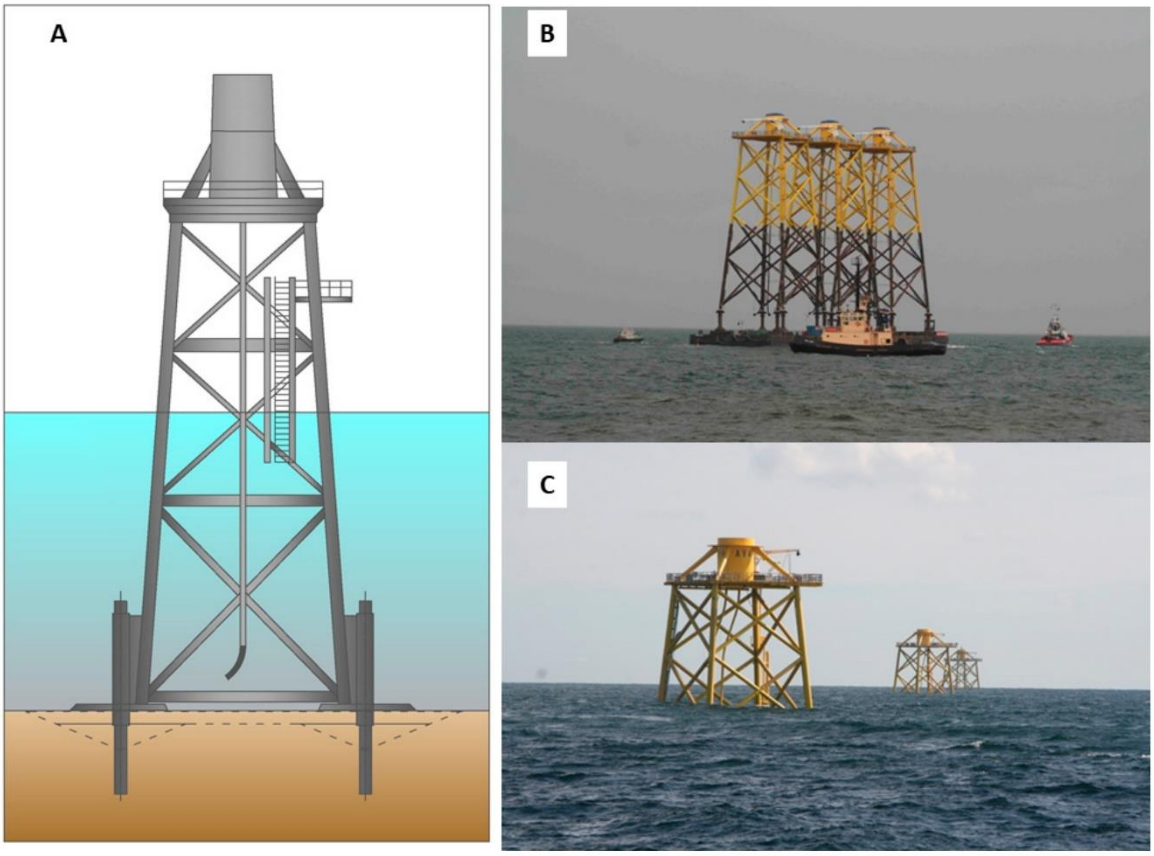

Figure 7 Jacket structure. (A) Scheme (B) Jacket foundation transportation (Alpha Ventus wind farm) (C) Jacket foundations installed (Alpha Ventus wind farm) (Manzano-Agugliaro, et al. 2020)

2.2.4. Floating systems

There has been increasing interest in floating systems to be used when the depth exceeds around 60m.

- Mooring stabilised TLP (tension leg platform) concept
- Ballast stabilised Spar buoy
- Buoyancy stabilised semi-submersible is a combination of the previous two approaches.

Although some offshore wind projects with floating systems have been deployed in Scotland (Scotland Hywind) and Norway (Equinor Tampen), they are still in the minority.

3. Maintenance theory

Maintenance is a highly scoped subject, that includes but is not limited to the maintenance of buildings, the emergency repairs of machines damaged during industrial accidents and the monitorisation of equipment in any industry.

Many companies have started to implement methodologies such as Six-Sigma, or Just in Time in an effort to fulfil the customer demands for high-quality products in a timely manner. This has resulted in a shift of their manufacturing, organizational, and supply chain strategies toward agility, quality, automation, and high performance. This has resulted in very high investments in equipment and people. To achieve the targeted rates of return-on-investment equipment must be reliable and safe to operate without costly work stoppages and repairs. (Duffuaa and Raouf 2015)

In the energy industry, the growing global energy demand, and the inability to store excess energy at a large scale have resulted in the need to minimize downtime in energy production systems, ranging from nuclear reactors to solar panels. These changes have shifted the perception of maintenance from a necessary evil to a key activity in manufacturing and energy production.

The requirements for agility, quality, automation, and high performance have led to the implementation of maintenance methodologies like total productive maintenance (TPM), reliability centred maintenance (RCM), or lean six sigma.

3.1. Maintenance Strategies

In this project maintenance will refer to conservative maintenance, that is, maintenance that is intended to preserve the functionality of a system. However, maintenance may also include improvement maintenance, overhaul maintenance, emergency maintenance and others.

Several different maintenance strategies have been developed since the industrial revolution, with increasing technical complexities, leveraging the latest technical developments in statistical analysis and monitoring capabilities. More simple maintenance strategies must not be disregarded as it is usual for several different strategies to coexist in the maintenance plan of any system.

- **Corrective maintenance**: Maintenance actions are carried out after a breakdown. Upfront costs of this type of maintenance are non-existent, however, long term and for expensive pieces of equipment, it may result in very high costs.

- **Preventive maintenance**: Maintenance actions are carried out at predetermined intervals of time or wear. This approach leads to less equipment downtime and longer asset life, however it is also more labour-intensive and there is potential for over-maintenance.
- **Condition-based maintenance** or **predictive maintenance**: Preventive maintenance that is initiated because of knowledge of the condition equipment through routine (discontinuous) or continuous monitoring. This approach leads to a decrease of maintenance costs of 30% on average (Schallehn, et al. 2018) and reduces the frequency of breakdowns by about 75% (PwC 2018). The complexities in the implementation of predictive maintenance systems in most industries arise from difficulties in developing the models and implementing the infrastructure required for condition monitoring tracking.

3.2. Key performance indicators in maintenance

Different industries will have different definitions for success in maintenance, a common way to define success in relatively standardised way are Key Performance Indicators, or KPIs. Some common KPIs are as follows:

- **Mean time Between Failures (MTBF)** is the average amount of time between breakdowns. The definition of a breakdown can differ. In the case of Offshore WT this is a specially relevant metric as service trips to the turbine farms can be costly and have a high logistical complexity.

$$MTBF = \frac{Total\ Working\ Hours}{Number\ of\ failures}$$

- **Mean time to repair (MTTR)** is the amount of time that it takes, on average, to return a piece of equipment to working conditions after a breakdown.

$$MTTR = \frac{Total\ repair\ time}{Number\ of\ failures}$$

- **Availability** measures the percentage of time that equipment is in working conditions. It gives an idea of the uptime of a piece of equipment. It is especially relevant in renewable energy generation (specifically solar and wind) as a readiness metric for the use of favourable wind conditions, as the throughput relies on external variable factors (e. g. meteorology).

$$Availibility = \frac{MTBF}{MTBF + MTTR}$$

- Overall equipment effectiveness (OEE): it is a measure of quality, performance and availability frequently used in manufacturing. The highest score (100%) is obtained when equipment is operating at the highest performance (number of pieces produces per time unit), with no defective pieces and no unavailability events.

$$OEE = Availibility \cdot Quality \cdot Performance$$

These metrics offer a way to objectively compare different maintenance strategies and the reliability of equipment.

3.3. Maintenance in offshore wind turbines

The increase in offshore wind turbine installations has led to a renewed interest for new and advanced techniques of maintenance for wind turbines. (Costa, et al. 2021)

Operation and maintenance costs represent 25% of energy production costs offshore wind turbine maintenance. This is 15% more than O&M costs for onshore wind turbines.

The reason behind this difference is the technical and logistical complexity of maintenance operations for offshore wind farms, which his higher than for onshore wind turbines. Service visits to offshore wind farms occur approximately once every 6 months (Faulstish, Hahn y Tavner 2011) and up to 5 times per and require 40 to 80 of man-hours to service.

Two decades ago, the improvements were centred on improving the maintainability of the turbines by facilitating access through lifting improvements and onshore farms and condition monitoring happening discontinuously, with measurements taken during service visits exclusively. (van Bussel and Henderson 2001)

Currently, although the evolution of strategies for maintenances is ongoing, it is clearly centred on remote condition monitoring of the turbines through the application of advanced models from vibration and acoustic signals. (van Bussel and Henderson 2001) Specifically vibration analysis represents 58% of the market share of condition monitoring. (Barszcz 2019)

4. Vibration analysis

As stated in the previous section, most current condition monitoring systems use vibration analysis, and there is great interest in its application to smart, remote condition monitoring. This section will give an introduction on how vibration signals are acquired and processed, and how condition monitoring systems use vibration data.

4.1. Data acquisition

The first step for vibration analysis is the acquisition of the vibration data. This is usually done by placing several accelerometers throughout the machine or area to be monitored. The exact placement will depend on the machine, the components that present most wear, and a variety of other factors.

In wind turbines the sensors are usually placed on the drive train, blades, and support structure.

4.2. Vibration Signals

Although the study of vibration signals may start with simple, clean sine waves, vibration signals recorded in real settings are often much noisier, including several overlapping signals of different amplitudes and phases. In any case, vibration signal analysis frequently starts by analysing the signal waveform itself but other methods such as frequency analysis or envelope analysis may be used. Due to the scope of this project, only time domain vibration features will be presented.

4.2.1. Relevant features of vibration signals

The features that will be presented in this section are "broadband" features because they do not use any filtering techniques. Therefore, the information they provide considers all signal components from a large (or "broad") frequency band and provide information of the overall system and not just from the specific mechanical problem that may be malfunctioning. All these features can be easily calculated from the vibration signal. They are:

- Statistical values such as the mean, the standard deviation, and the kurtosis of the signal.
- Root-mean-square (RMS) which describes the area of the signal and therefore its energy $RMS = \sqrt{E(x^2)}$, where E is the mean value operator.
- Peak value, or peak-peak (PP) is a measure of the distance of the maximum peaks of the signal $PP = x_{max} - x_{min}$

- Zero-peak (ZP) may be used instead of peak-peak as it is more easily compared to signal components such as the amplitude and other related signal components. $ZP = \frac{PP}{2}$
- Crest Factor (CF), which is the ratio of the peak amplitude to the RMS value. Usually, a change in the operation mode will change RMS and PP proportionally, therefore it is useful to track the ratio of these two, which should be independent of operation conditions. $CF = \frac{ZP}{RMS}$

4.3. Complex methods

Although graphical comparisons of the waveforms or simple statistical analysis of these features can be enough for some purposes, more complex methods have been developed to determine health status of structures and machinery.

Guided waves are one of such methods. This approach relies on the mathematical modelling of the behaviour of a material when an ultrasonic wave is applied to it via a transducer. Any cracks, delamination or defects will become apparent as a deviation of the mathematical model. Guided wave-based approach for health monitoring of composite structures; Application to wind turbine blades (Shoja 2018) is an example of its application to wind turbines, but the size of wind turbine components is a challenge as only a low number of frequencies can be applied.

Another challenge to SHM of wind turbines (WT) is the fact that the excitation in the structures due to wind and waves cannot be known, and therefore a normal input-output model cannot be used. This can be solved by vibration-response-only SHM where only the response of the structure is studied. (Puruncajas, Vidal and Tutivén 2020)

In some cases, more complex or novel features of the vibration signals are studied. These can include spectral kurtosis, spectral entropy (Sandoval, et al. 2020), and others.

Finally, many artificial intelligence (AI) methods have been applied. Some examples are convolutional neural networks (Puruncajas, Vidal and Tutivén 2020) , supervised machine learning methods such as logistic regression and support vector machine models (Taylor, Beale and Murat 2017) or unsupervised methods like Gaussian mixture models (Statco, et al. 2019).

5. Applied Artificial Intelligence

The British Encyclopaedia defines artificial intelligence as "the ability of a digital computer or computer-controlled robot to perform tasks commonly associated with intelligent beings".

Research in this field has been ongoing since the 1950s. The 1960s were a decade of relative success, but at that time most applications focused on following an established set of rules in order to perform a task. However, this meant that AI could not yet be applied to solving problems that could not be defined by simple sets of rules. As computational power became more easily accessible, increasingly complex sets of rules could be applied obtaining results in a timely manner, and neural networks were developed, which allowed to solve problems that could not be encoded in those rules. Today, the amount of data available has meant that AI has been used for problem solving in a variety of fields, ranging from medicine to retail. (Haenlein and Kaplan 2019)

5.1. Machine Learning

Machine Learning is a subset of artificial intelligence based on computers observing data, building a model based on those observations and using this model as a hypothesis for problem-solving. (Norvig and Russell 2021)

There are three main types of Machine Learning:

- Supervised
- Non-supervised
- Semi-supervised
- Reinforcement learning

Supervised models are trained with structured labelled data. This means that the data the model is trained with contains both the input data and the desired output data. The model can then infer the label of a new, unseen data point. It is often expensive and cumbersome to rely on specialists and trained workers to label a dataset. In some cases, labelling may not even be possible

For health monitoring of offshore wind turbines, labelling of data would require collection of data in a wide array of conditions for the specific structure or an identical one. The wide array of possible structural damages as well as the different possible combinations of wind turbine models, foundations and structural support configurations would require an extensive dataset, which would be costly in terms of time, monetary resources, and expertise.

In non-supervised machine learning models, the labels for the data are not provided, so the desired output is not available. The model will look for patterns or clusters to generate a classification algorithm relying solely on the input data. This is a common industry scenario when there are systems in place to record data, such as SCADA systems. This route can be cheaper, as there are no costs in terms of time and expertise to label the data. Due to the lack of labels on the data, it may be more difficult to achieve the same level of accuracy as with supervised machine learning.

Semi-supervised models use a small set of labelled data and a much larger set of non-labelled data for training. This is an option where data can be collected for one class of data, for example collecting data from "normal" behaviour, such as non-fraudulent credit card transactions, or in the case of this project, "healthy" data from a non-damaged structure.

Reinforcement learning models are trained based on some reward or punishment that the algorithm is trained to achieve or avoid. This is used for example to train algorithms that will play a video game. Reinforcement learning is outside the scope of this project.

In this project, non-supervised and semi-supervised models will be used, in an attempt to provide a solution that can be easily applied for SHM of real offshore wind turbine foundations.

5.2. Carrying out Machine Learning Projects

In the previous section several types of machine learning models were discussed. Attaining meaningful results when applying any of these models to a dataset requires extensive work. The following diagram of flow shows the steps in which a data science project is carried out. (Grus 2015)

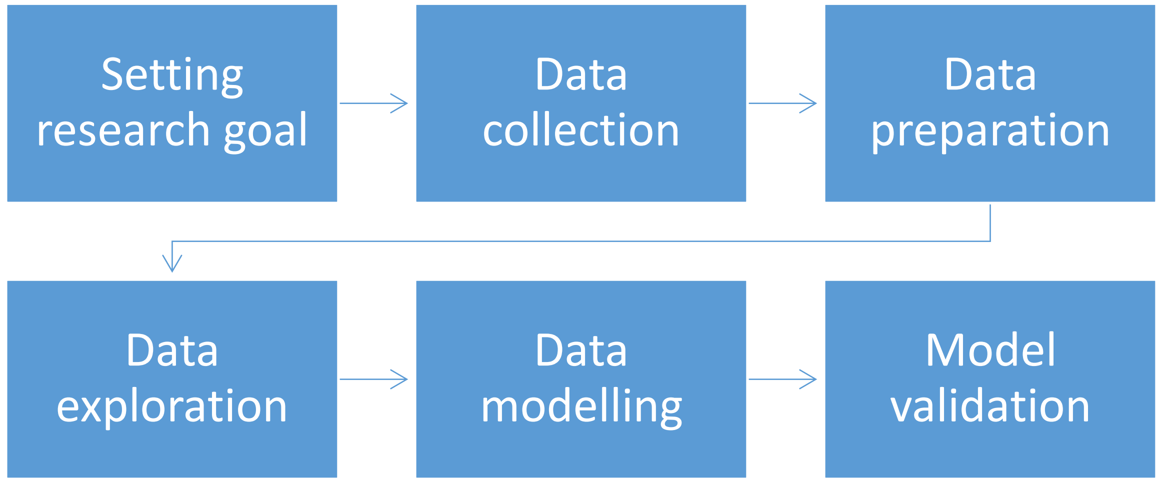

Figure 8 Steps of a Machine Learning project

Setting the research goal

The first step is to decide what will be the goal of the research. This includes not only choosing the topic but also setting metrics by which the success of the outcome will be measured. Some metrics that could be applied to predictive maintenance projects are described in Section 3.2.

Data collection

Data must be acquired before it can be analysed.

In a lot of cases datasets are published by researchers or institutions with hopes that the availability of the data will interest researchers. One example of this is the MNIST data set, a dataset of handwritten characters.

Sometimes, in corporate or government settings data will be collected with the intention of it being analysed. Data collection can be done through surveys in social science studies, or with sensors or other devices in STEM research.

Sometimes real data cannot be collected, in this case, synthetic data may be used. Synthetic data is data manufactured to replicate real world data as closely as possible. Synthetic data is used in cases where real world data is not available due to lack of technical or economical sources.

Data Preparation

Once data has been collected, it must be prepared before further analysis. This may include data cleaning, where malformed or badly recorded data is removed from the data set; or data imputation, where missing data is filled in. It may also include transforming the data by reshaping matrixes or change the format, from audio to image or from the time domain to the frequency domain. This step usually takes up the longest, and often makes up for most of the time spent in any data science project.

Data Exploration

After preparing the data, some preliminary analysis is carried out. This usually entails generating some quick visualisations of the data sets with bar plots or scatter plots, to get a broad understanding of the data.

Data Modelling

In this step, mathematical models, statistical models, and machine learning models are used to model how the system behaves. The goal is to produce a model that will classify, cluster or label new data or accurately predict an outcome.

Model Validation

After the models have been trained, the model is validated by measuring its behaviour against several metrics and measuring its effectiveness in achieving the goals set on the first step of the process.

After these steps, the model goes in “production” where it will be used for the research goal determined in the first step of the process. The model being “in production” does not mean that the data scientist’s work has finished. Changes in the processes, new data becoming available and new developments in the machine learning field men that the model may need to be retrained with a new dataset, retuned to achieve better results, or completely changed if a better alternative model becomes available.

5.3. Models in Machine Learning

Within all these types of machine learning, there are many algorithms or models. A model is essentially a mathematical pattern that performs certain operations to be able to label, classify or cluster data.

Oftentimes a simple machine learning model will be implemented within a software solution that offers a complex functionality. A lot of chatbots are powered by early Natural Language Processing models, like ELIZA.

Sometimes several machine learning models can also be combined or linked to provide more complex outputs. For example, modern object recognition models use first a model that will locate all the objects in the picture. The sub-images of the sections of the objects will then be classified by a second model.

The next sections will introduce the models used in the practical section of this project.

5.3.1. One-Class SVM

One-Class Super Vector Machine (SVM) is an unsupervised model for outlier detection.

It is a variation of the Super Vector Machine. In traditional SVM, samples are separated by a decision boundary that attempts to separate two classes with the maximum margin. The support vectors refer to the data points that lie on the margin to the decision boundary.

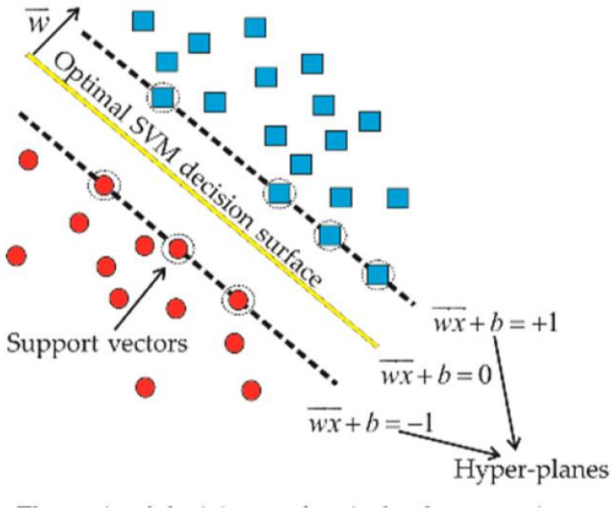

Figure 9 SVM decision boundary and support vectors (Wang, y otros 2019)

In the case of non-linearly separable classes, samples can be projected on a higher dimensional space using a kernel and are then separated by a hyperplane. This way, samples that may not have been separable in the original space, become separable by a decision boundary that is not linear in the original space.

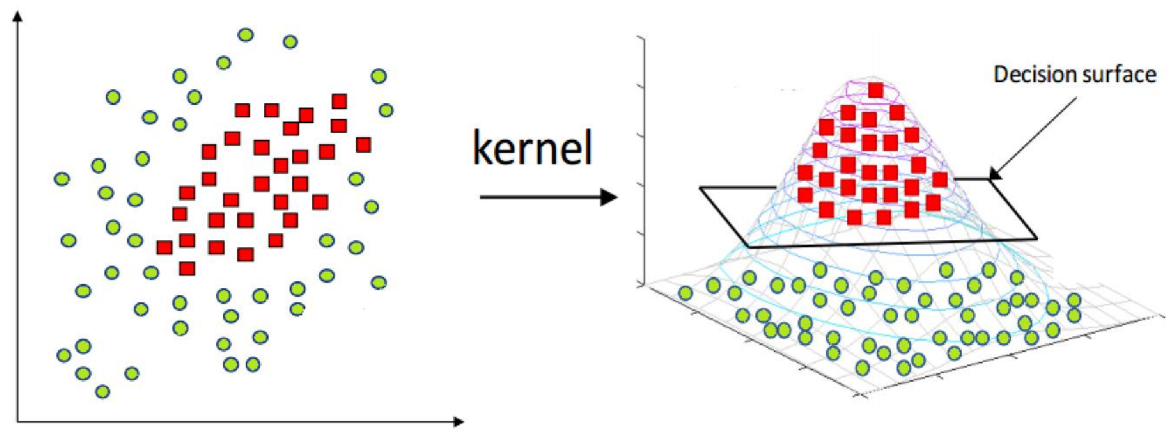

Figure 10 Original and kernelized feature space (Rizwan, et al. 2021)

One-Class SVM is trained with only one class of data. In this case the samples are projected in a higher dimensional space and the hyperplane is set between the origin and the samples, making the region where the samples lie as small as possible. Points that lie on the side of the origin of the hyperplane will be considered outliers. (Scholkopf, et al. 1999)

Kernels

This is the function that performs the projection into a higher dimensional space. Particularly, this project uses the polynomial, sigmoid and radial basis function (RBF) kernels.

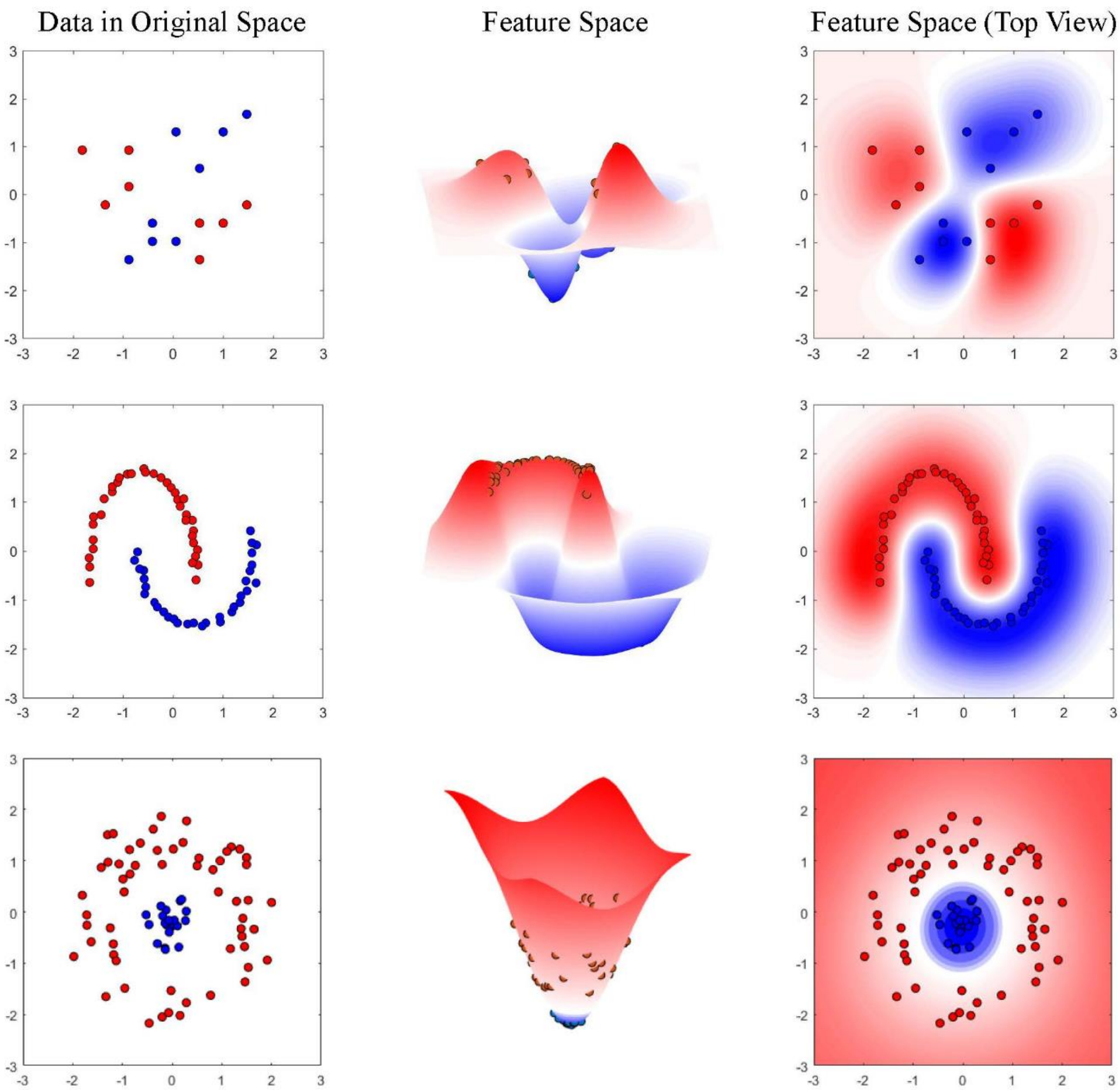

Figure 11 Illustration of non-linear kernel transformations (Ezra Pilario, et al. 2020)

Nu

Nu is the number of samples we allow outside the decision boundary during the initial training. This hyperparameter is useful in the case of a noisy dataset, where although we expect most samples to belong to the "normal" class, we want to allow some samples to lie outside of the class. Otherwise, the decision boundary may include outliers.

5.3.2. Isolation forest

Isolation forest is a semi-supervised model for anomaly detection based on the use of decision trees.

The model will create decision trees based on random features of the samples, setting a random threshold for the separation criteria. As anomalies are "few and different" they will be separated early in the tree.

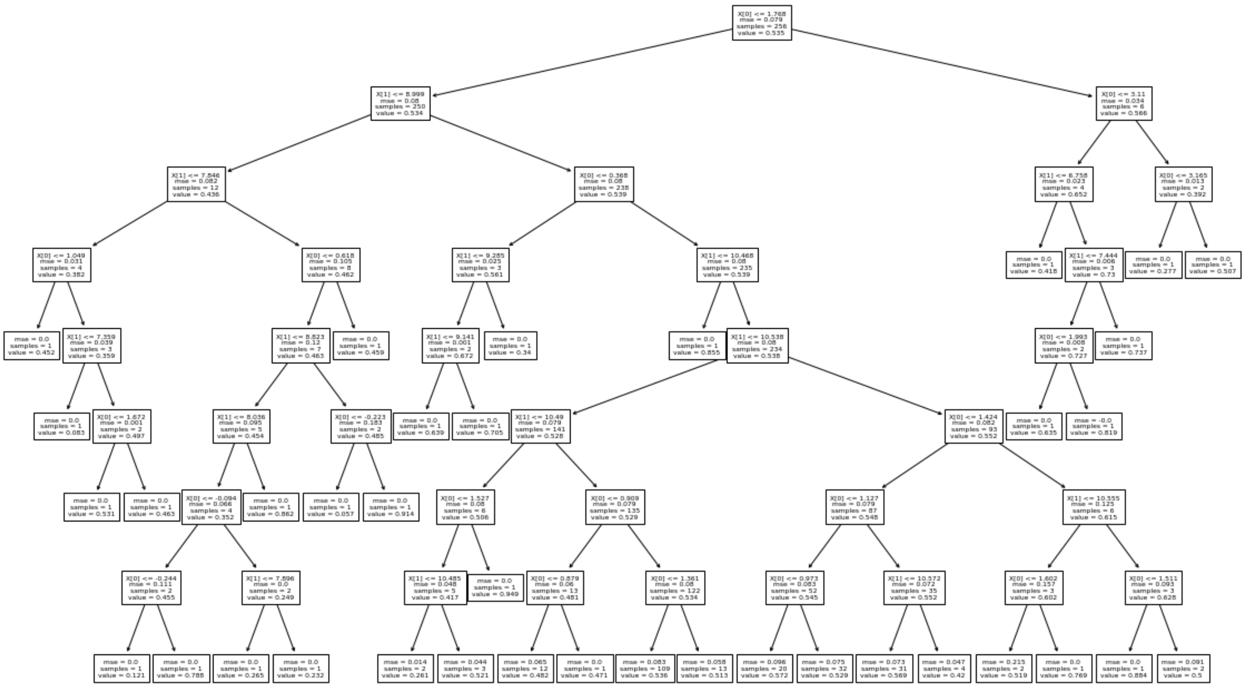

Figure 12 Example of a random tree in an Isolation Forest Model

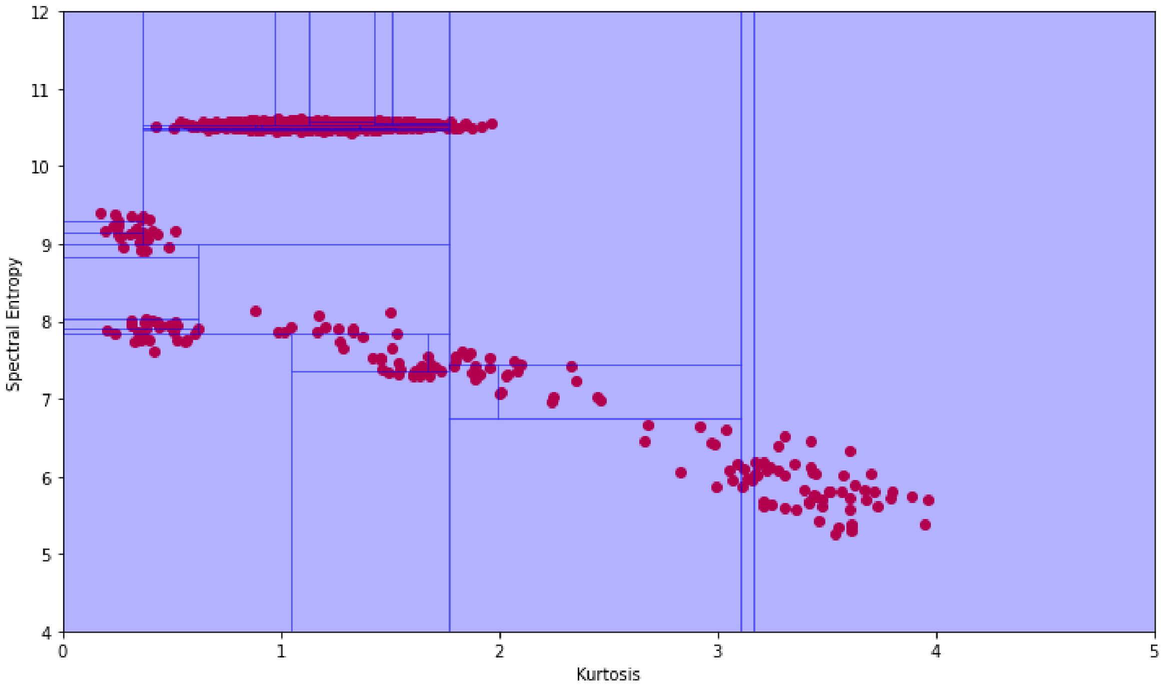

Figure 13 Scatter plot and decision boundaries of a random decision tree in an Isolation Trees model

Instead of relying on one single decision tree, the model generates many isolation trees, and the anomalies will be those that on average, over all of the trees, have short paths. The ensemble of these isolation trees is what the name of the algorithm refers to.

Contamination

The percentage of samples that are expected to be anomalies. It is the one parameter that makes this model into a semi-supervised model, as at least an estimation of the proportion of the two classes must be known beforehand. (Liu, Ting and Zhou 2008)

Number of estimators

The amount of decision trees in the random forest. The number of trees will impact the computation time significantly. However, a larger number of trees will also produce more nuanced anomaly scores when compared to a lower number of trees.

As this is not a deterministic model, with a lower number of trees the results will also be less repeatable. This can be solved by using a pseudo-random generation that can be seeded with a repeatable randomness state.

5.4. Validation metrics in Machine Learning

This section will present some metrics that can be used for the validation of ML models. In regression models some metrics like Error, Mean Square Error or Root Mean Square Error may be used. As this project focuses on the separation of two classes of data, two relevant metrics are:

Accuracy

Accuracy is a relatively simple metric to assess the performance of a ML model. It is the percentage of correct labels predicted for a dataset. Although it is a simple metric it has some shortfalls in the case of unbalanced data. The model may be classifying correctly only one large class and due to class imbalance, a high accuracy could still be obtained.

Confusion matrix

Confusion matrixes are a common way to represent the performance of a ML model. Confusion matrixes have four boxes, and they represent the predicted and true label of a dataset. A good performing model will perfectly map all the samples, so the predicted label matches the true label.

Confusion matrixes also provide insight into false positives and false negatives.

5.5. Applications of Machine Learning in Engineering

Machine Learning has been applied to problems in the engineering domain for many decades now. Currently, the advances in computing, sensor technology and new algorithms are facilitating the implementation of Machine Learning to new industry problems.

5.5.1. Manufacturing Industry

Within the manufacturing industry, Machine Learning has been used for process optimisation (Weichert, et al. 2019) and predictive maintenance and computer vision systems have been implemented for quality control. (Wu and Sun 2013)

5.5.2. Energy industry

Within the energy industry, machine learning is currently being applied to energy demand forecasting (Ahmat and Chen 2018) and predictive maintenance of both electrical distribution (Hoffman, et al. 2020) and energy production assets such as wind turbines.

6. Vibration analysis and ML model application to experimental data

In this section an experimental dataset will be used to train two models (One Class SVM and Isolation Forest) in order to assess if it is feasible to use non-supervised and semi-supervised ML models for SHM of offshore wind turbines.

6.1. Data collection

The dataset used is the same as in Vidal et. al. In the article, eight triaxial accelerometers are placed on a scaled down model of an offshore wind turbine with a jacket structure. The wind conditions are simulated by a modal shaker using several amplitudes (0.5, 1, 2 and 3A) of electrical current as a proxy for wind speeds. Furthermore, data is recorded in 4 scenarios: a healthy bar, a bar with a loose bolt, a bar with a crack, and a replica bar. (Vidal, Rubias and Pozo 2019)

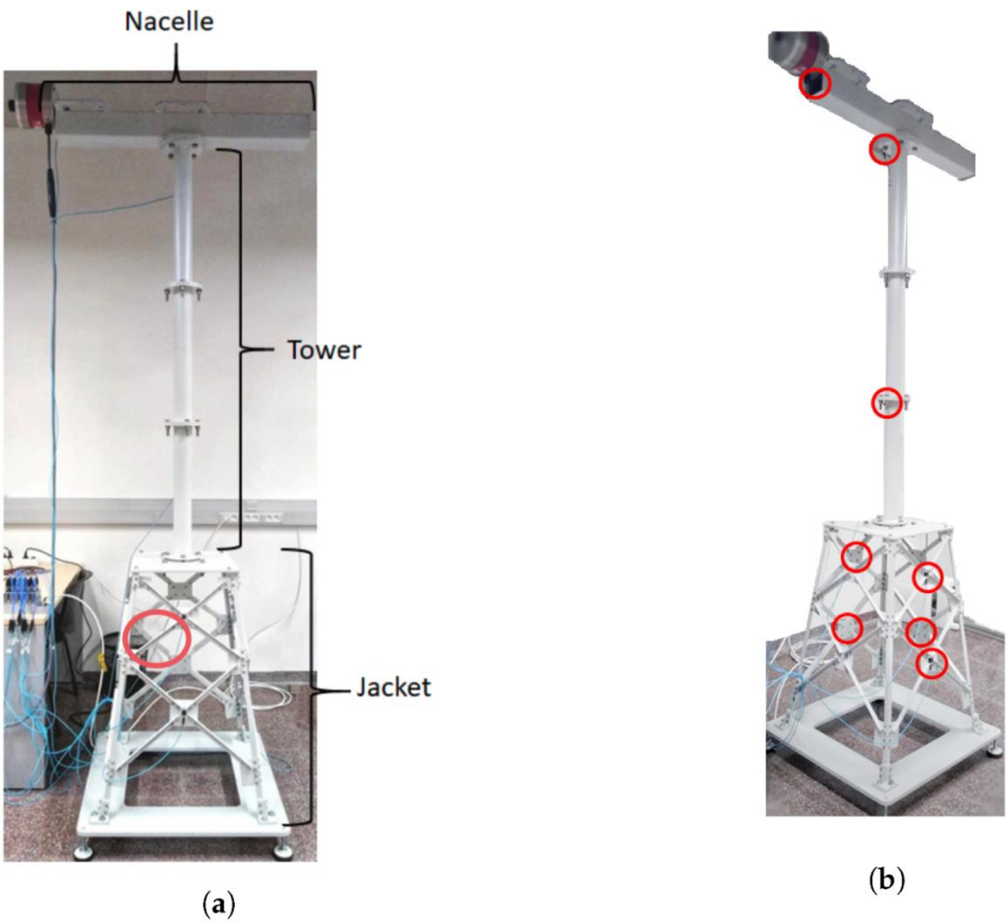

Figure 14 (a) The bench test detailing the location of the bar, and(b) Location of the sensors (Hoxha, Vidal and Pozo 2020)

The data consists of 25 experiments for each amplitude, amounting to a total of 100 experiments:

	Amplitude			
	0.5 A	1 A	2 A	3 A
Healthy bar	10	10	10	10
Replica bar	5	5	5	5
Cracked bar	5	5	5	5
Loose bolt in bar	5	5	5	5

Table 1 Number of experiments by state of bar and amplitude

In each experiment, a time window of 60 seconds is recorded at a frequency of 1651.6129 Hz. Thus, we obtain 99097 data measurements from each of the 24 sensors (8 accelerometers with 3 axis each) for each experiment.

6.2. Data transformation

As explained in the previous section, for each of the 25 experiments performed we obtain a matrix of shape [999097x24]. However, since the sampling frequency is very high compared to an industry setting, the data is subsampled in a 1:6 ratio. Therefore, our new subsampled matrixes are of shape [166517x24] which is equivalent to a sampling frequency of 256 Hz and a time window of 60 seconds.

$$\begin{bmatrix} x_{(1,1)} & \cdots & x_{(1,24)} \\ \vdots & \ddots & \vdots \\ x_{(999097,1)} & \cdots & x_{(999097,24)} \end{bmatrix} \xrightarrow{Subsampling\ to\ lower\ freq.} \begin{bmatrix} x_{(1,1)} & \cdots & x_{(1,24)} \\ \vdots & \ddots & \vdots \\ x_{(166517,1)} & \cdots & x_{(166517,24)} \end{bmatrix}$$

However, we can expect to obtain results with a shorter time window, so the data is reshaped in order to obtain 664 samples from each experiment, which equates using a time window of 0.090361 seconds. Therefore, each row (sample) will contain 199 timestamps for 24 sensors, for a total length of 4776 datapoints). We can stack the samples in a matrix of shape [664x4776] for each experiment, and furthermore stacking samples of several experiments, although each sample will be processed separately.

$$\begin{bmatrix} x_{(1,1)} & \cdots & x_{(1,24)} \\ \vdots & \ddots & \vdots \\ x_{(166517,1)} & \cdots & x_{(166517,24)} \end{bmatrix} \xrightarrow{Splitting\ the\ matrix} \begin{bmatrix} [x_{(1,1)} & \cdots & x_{(199,24)}] \\ & \vdots & \\ [x_{(166318,1)} & \cdots & x_{(166517,24)}] \end{bmatrix}$$

The matrixes are then scaled with a standard scaler fitted column wise to the “healthy” dataset.

6.3. Features

For each of the samples obtained in the previous section, we calculate the average, standard deviation, kurtosis, RMS, PP, ZP, CF (defined in Section 4.2.1) and spectral entropy.

6.4. Model training

The healthy samples are split into a training set (80%) and a validation set (20%). In the case of One Class SVM all other states are used only as validation data and not used for training. In the case of Isolation Forest, a random set of 8,73% the size of the healthy sample training set is drawn and included in the training set, and the complete set of other states is used for training.

Then we train the model with the training set for a range of values on the hyperparameters for both models.

Hyperparameter	*Values*
Kernel	RBF, Polynomic, Sigmoid
Nu	0.0001, 0.01, 0.1, 0.25
Tolerance	0.01, 0.001, 0.0001
Gamma	Scale, Automatic
Degree (only for polynomic kernel)	2

Table 2 Hyperparameter values tested for One Class SVM

Hyperparameter	*Values*
Number of estimators	5, 10, 50, 100
Contamination	$N_{random\ outliers}/N_{total\ training\ data}$
Random state	32

Table 3 Hyperparameter values tested for Isolation Forest

6.5. Model validation

We perform validation against the samples of classes 1, 3 and 4. Both models obtain 100% accuracy in several cases. We enclose two particular sets of hyperparameters and features that achieved this accuracy. The full results can be found in the GitHub repository in https://github.com/clara-9/TFG_public.

Parameters	Number of Estimators	Overall accuracy
Standard Deviation, Kurtosis, Spectral Entropy	50	100%
RMS, Zero Peak, Spectral Entropy	100	100%
Standard Deviation, Zero Peak, Spectral Entropy	100	100%
Standard Deviation, Peak-Peak, Spectral Entropy	100	100%
Mean, RMS, Spectral Entropy	50	100%
Mean, Kurtosis, Spectral Entropy	100	100%
Kurtosis, Spectral Entropy	50	100%
Kurtosis, Spectral Entropy	100	100%
Standard Deviation, Kurtosis, Spectral Entropy	100	100%
Mean, Kurtosis, Spectral Entropy	50	100%
Mean, RMS, Spectral Entropy	100	100%

Table 4 Selection of hyperparameters and features for Isolation Forest models with 100% accuracy

Parameters	Tolerance	Nu	Gamma	Overall accuracy
Zero Peak, Spectral Entropy	0.0010	0.0001	auto	100%
Zero Peak, Spectral Entropy	0.0100	0.0001	auto	100%
Zero Peak, Spectral Entropy	0.0001	0.0001	scale	100%
Zero Peak, Spectral Entropy, Crest Factor	0.0010	0.0001	auto	100%
Zero Peak, Spectral Entropy	0.0010	0.0001	scale	100%
Zero Peak, Spectral Entropy, Crest Factor	0.0100	0.0001	auto	100%
Zero Peak, Spectral Entropy	0.0001	0.0001	auto	100%
Zero Peak, Spectral Entropy, Crest Factor	0.0001	0.0001	auto	100%

Table 5 Selection of hyperparameters and features for One Class SVM models kernel RBF with 100% accuracy

Parameters	Tolerance	Nu	Gamma	Overall accuracy
Mean, Spectral Entropy	0.0001	0.0001	auto	100%
RMS, Spectral Entropy	0.0001	0.0001	auto	100%
Kurtosis, Spectral Entropy, Crest Factor	0.0001	0.0001	auto	100%
Kurtosis, Spectral Entropy, Crest Factor	0.0001	0.0001	scale	100%
Standard Deviation, Spectral Entropy	0.0001	0.0001	scale	100%
Mean, Spectral Entropy	0.0001	0.0001	scale	100%
Standard Deviation, Spectral Entropy,	0.0001	0.0001	auto	100%

Crest Factor				
RMS, Spectral Entropy	0.0001	0.0001	scale	100%

Table 6 Selection of hyperparameters and features for One Class SVM models kernel sigmoid with 100% accuracy

Parameters	Tolerance	Nu	Gamma	Overall accuracy
Mean, Spectral Entropy	0.0001	0.0001	scale	100%
RMS, Spectral Entropy	0.0001	0.0001	scale	100%
RMS, Spectral Entropy	0.0001	0.0001	auto	100%
Mean, Spectral Entropy	0.0001	0.0001	auto	100%
RMS, Spectral Entropy	0.0010	0.0001	scale	100%
Mean, Spectral Entropy	0.0010	0.0001	auto	100%
Mean, Spectral Entropy	0.0100	0.0001	auto	100%
RMS, Spectral Entropy	0.0100	0.0001	scale	100%

Table 7 Selection of hyperparameters and features for One Class SVM models kernel polynomic with 100% accuracy

We will further analyse one of the sets of hyperparameters and features that achieved a 100% accuracy for each model. Specifically, for isolation forest, we will analyse the pair of features kurtosis and spectral entropy with 100 estimators, and for one class SVM kernel RBF, gamma "scale", tolerance 0.0001, nu 0.0001 with zero-peak and kurtosis

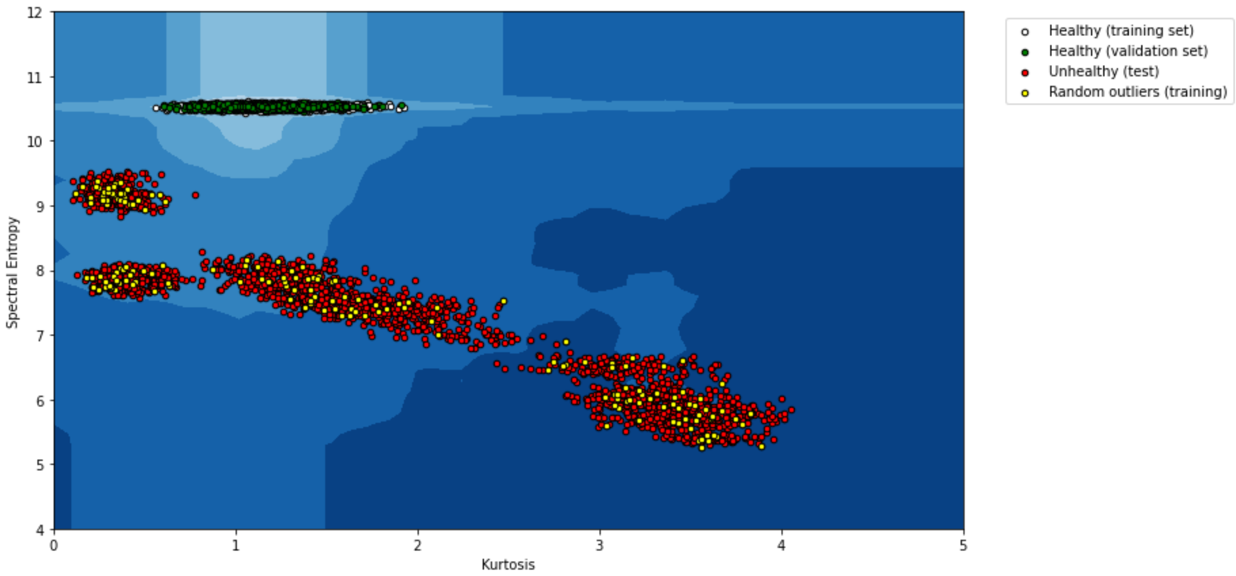

Figure 15 Visualisation of Isolation Model selected for further analysis

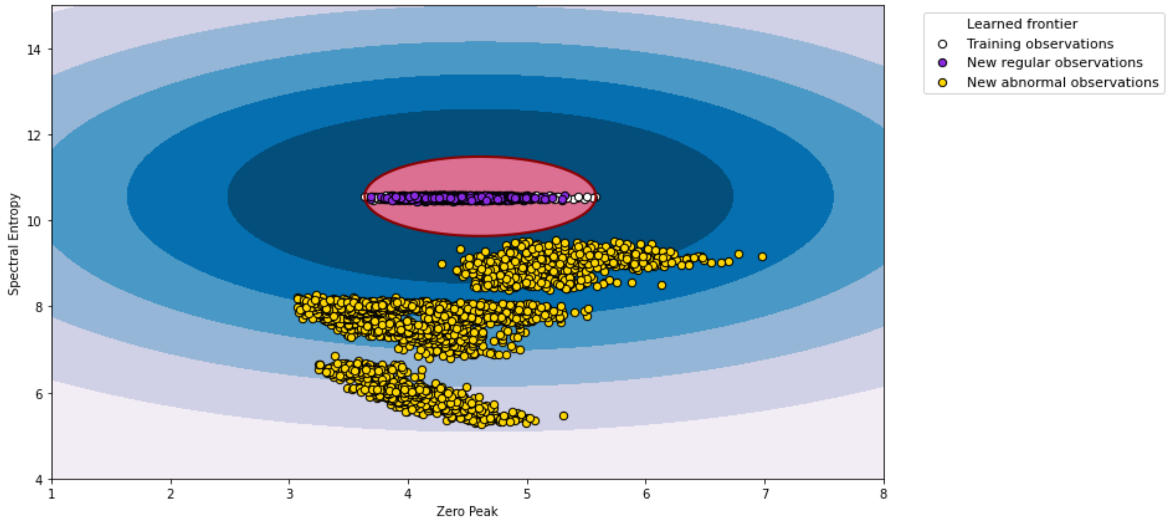

Figure 16 Visualisation of One Class SVM selected for further analysis

6.6. Model validation with replica bar

In the previous section the model was validated only against classes 1, 3, and 4. That is, against the healthy class and two different unhealthy classes.

In this section the selected models are tested against a second healthy class, the replica bar.

This would be comparable to a scenario where there is damage and the bar is replaced for a new one, so the system goes from an unhealthy to a healthy state again, but this second healthy state isn't necessarily identical to the first case.

	Isolation Forest	*One Class SVM*
Healthy class accuracy	100%	100%
Replica class accuracy	24,57%	0%
Unhealthy class accuracy	100%	100%

Table 8 Selected models accuracy by class

In both models the accuracies for this new class were low, obtaining a maximum of accuracy of 24,57% for the replica class for One Class SVM and 0% accuracy for the replica class for Isolation Forest.

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Escola d'Enginyeria de Barcelona Est

By visualizing the data, we can see a much higher overlap of the replica samples and the unhealthy samples in the parameters chosen. Furthermore, in the case of the Isolation Forest, the samples for the replica class lay completely outside of the decision boundary.

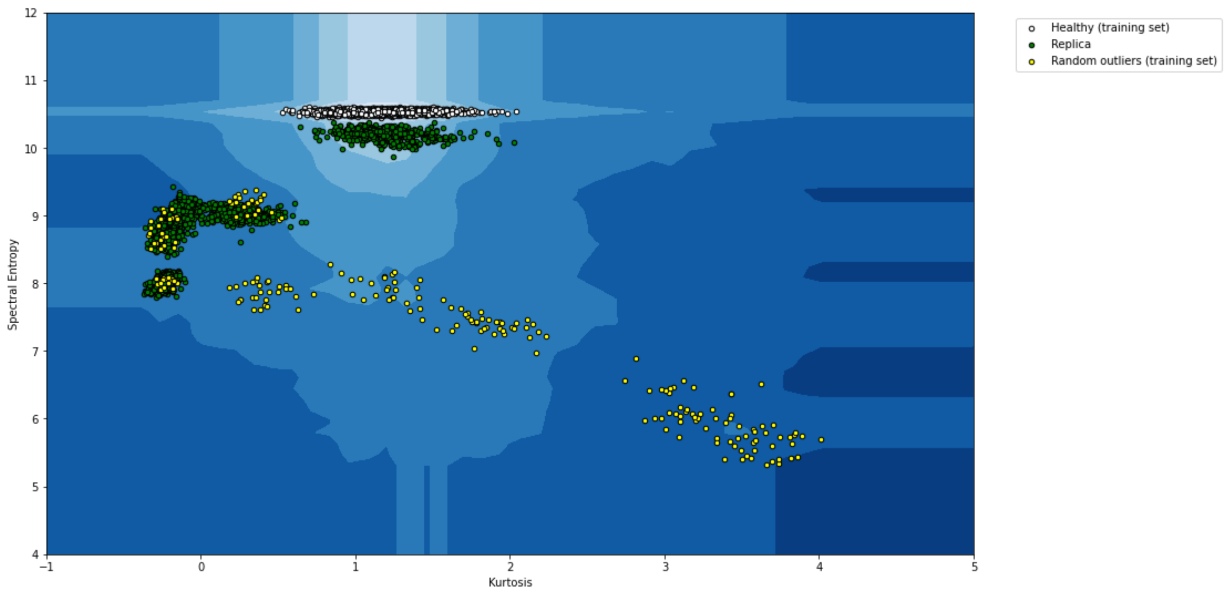

Figure 17 Selected Isolation Forest model validated against replica class

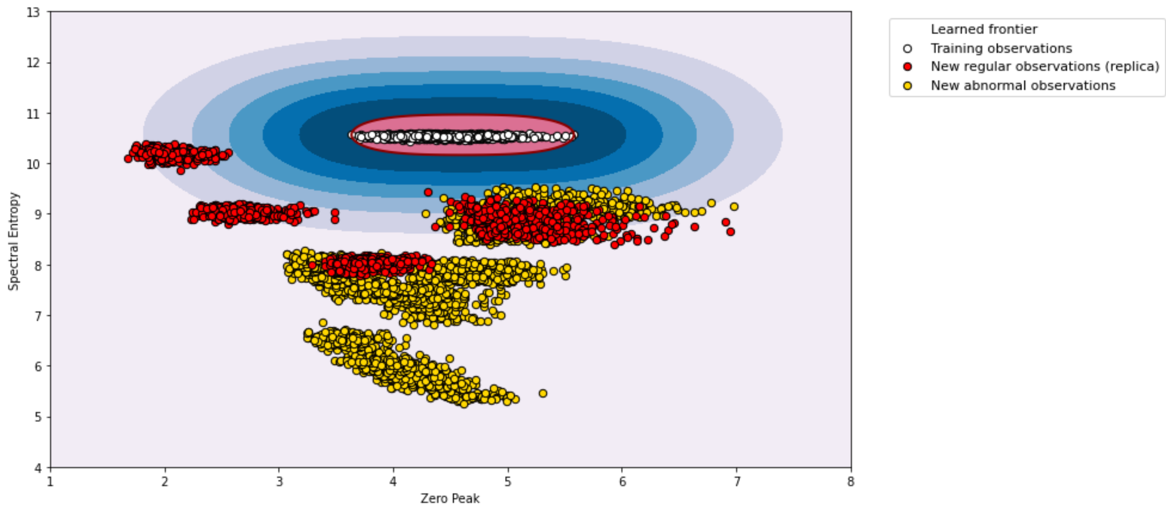

Figure 18 Selected One Class SVM model validated against replica class

6.7. Retrained model validation with replica bar

In a production environment, repairs will be known. The model could then be retrained using the "replica bar" dataset and not including the first "healthy bar" dataset.

A first approach would be to use the selected models but use the replica class for training. This yields the following results:

	Isolation Forest	*One Class SVM*
Training class accuracy (replica)	95,52%	99,62%
Validation class accuracy (replica)	96,39%	98,50%
Unhealthy class accuracy	64,25%	59,91%

Table 9 Accuracy of selected models by class when trained with replica class

As can be seen, although the decision borders of the models shift due to the training with the replica class, the overlap of the samples in the two parameters makes a clear separation impossible and limits the accuracy of the models.

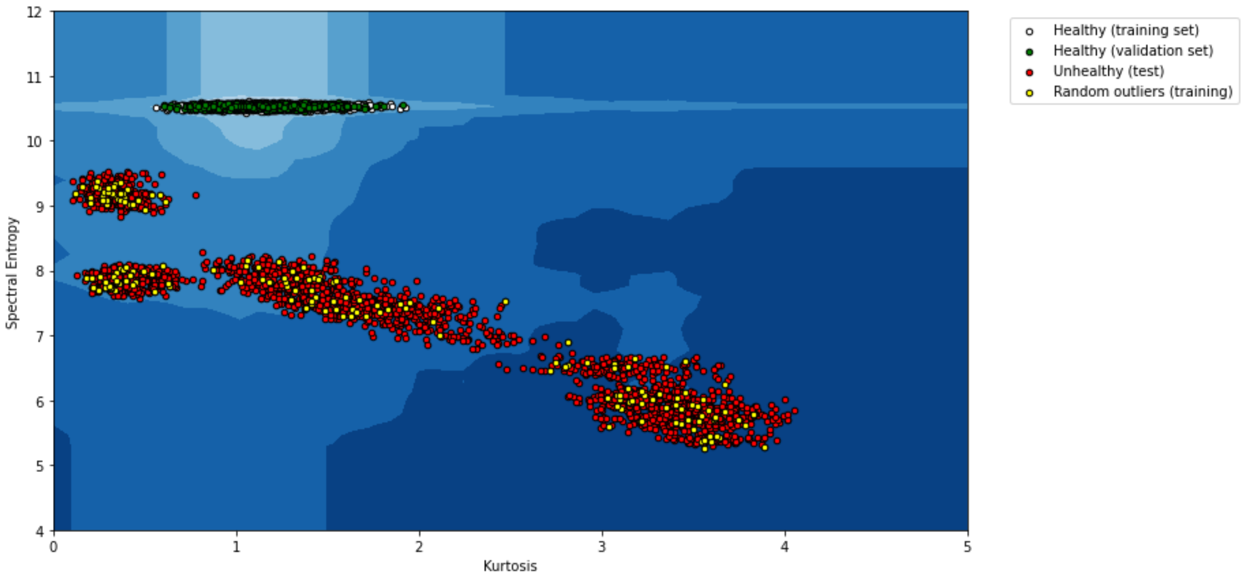

Figure 19 Selected Isolation Forest model trained with replica class

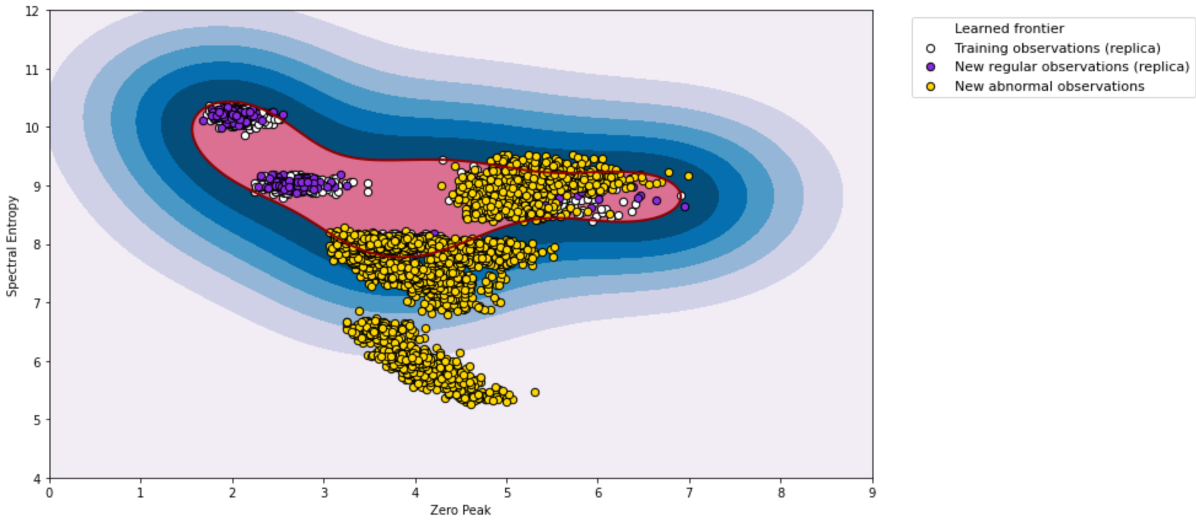

Figure 20 Selected One Class SVM model trained with replica class

6.8. Retuning the model for replica bar

Another option would be to retune the hyperparameters. As can be seen, the best results are achieved by using different parameters.

Parameters	Number of Estimators	Overall accuracy
Mean, Standard Deviation, Kurtosis	50	93,88%
Mean, Standard Deviation, Kurtosis	10	93,76%
Mean, Standard Deviation	50	93,57%
Mean, Standard Deviation	100	93,53%

Table 10 Selection of hyperparameters and features for Isolation Forest models with best accuracy (replica class)

Parameters	Kernel	Tolerance	Nu	Gamma	Overall Accuracy
Mean, Kurtosis	RBF	0.0001	0.0001	auto	88,67%
Standard Deviation, Crest Factor, Spectral Entropy	Sigmoid	0.01	0.25	auto	75,44%
Mean, Kurtosis	Polynomial	0.1	0.10	auto	87,13%

Table 11 Selection of hyperparameters and features for One Class SVM models, with best accuracy (replica class)

The visualisation for the best performing two dimensional models, although not the best performing one in the case of Isolation Forest, are enclosed below for ease of comparison with the previous sections.

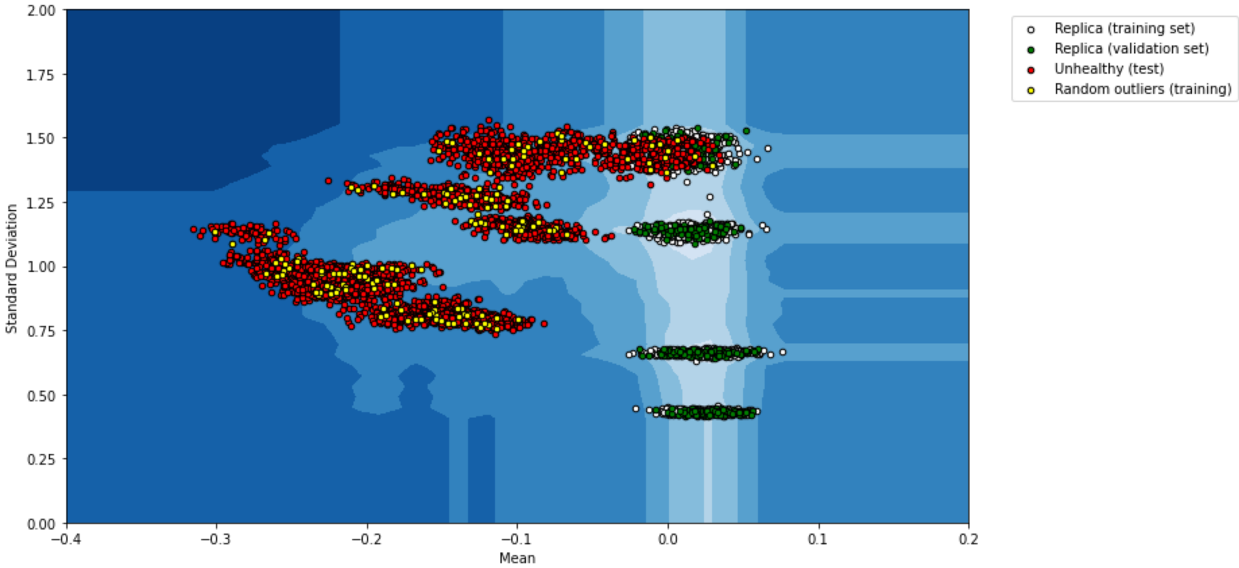

Figure 21 Isolation Forest trained with replica class, Standard Deviation, Kurtosis and 50 estimators

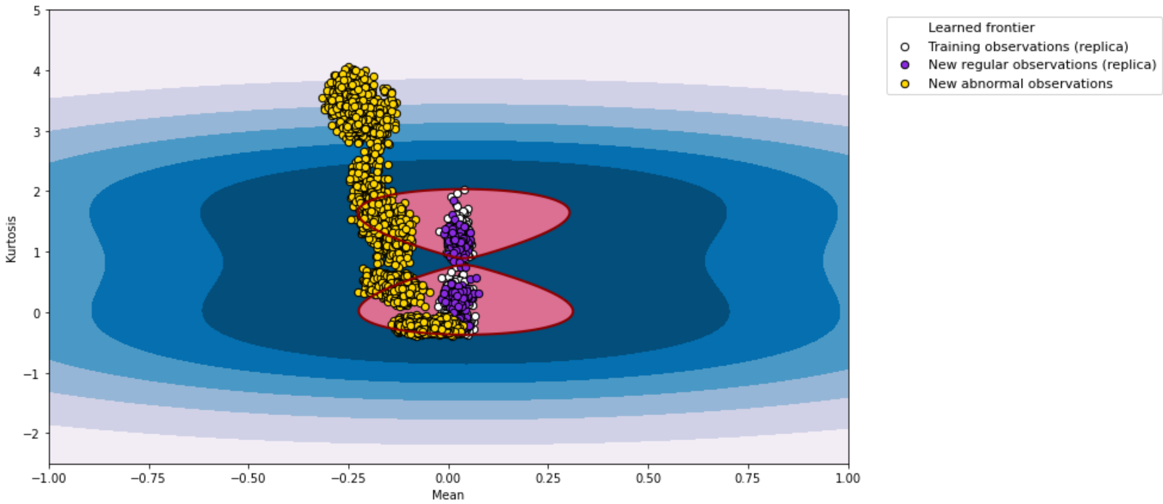

Figure 22 Best Performing One Class SVM with replica class

	Isolation Forest	*One Class SVM*
Training class accuracy (replica)	98,95%	99,62%
Validation class accuracy (replica)	98,49%	97,29%
Unhealthy class accuracy	90,93%	83,43%

Table 12 Accuracies by class of the best performing models when trained with the replica class (hyperparameters can be found in Table 10 and Table 11)

These new results are slightly inferior to the results of training with the “healthy bar” class only. One possible explanation would be that the new healthy dataset is 1/4th the size of the original dataset and this decreases the accuracy of the model. This could be solved with some data augmentation techniques such as changing the values taken during the subsampling for a lower frequency.

6.9. Conclusion

Satisfactory results were achieved with both models, but Isolation Forest was a better performer.

The use of spectral entropy, a non-traditional feature in vibration analysis, provided reliable results when the training set was larger.

The selected models needed to be trained and tuned again (with new features and hyperparameters) in order to have satisfactory results for a new “healthy” state. Even in that case, the performance achieved was lower, possibly due to the effects of a smaller data set for training.

7. Environmental impact analysis

Currently offshore wind energy generation has a higher environmental impact than onshore wind energy generation. Some estimates indicate that the emissions of greenhouse gases amounted to less than 7 g CO2-eq/kWh for onshore and 11 g CO2-eq/kWh for offshore. (Bounou, Laurent and Olsen 2016) Offshore wind energy can also lead to marine habitat loss and ecosystem degradation in a variety of ways. (Hernandez, Shadman and Maali 2021)

However, appropriate maintenance can lead to a smaller environmental footprint of systems. In fact, badly maintained systems can lead to a higher energy consumption. (Jasiulewicz-Kaczmarek and Drożyner 2013)

In the case of offshore wind turbines, the environmental impact of an improved maintenance strategy is twofold:

- A higher availability of the wind turbines will lead to a higher generation of renewable energy, enabling displacement fossil fuel-based energy generation. (Snyder and Kaiser 2009)
- Improved maintenance leads to more reliable systems. A higher reliability would allow for a lower frequency of servicing, which is usually done by boat or sometimes helicopter. The reduction in these servicing trips would reduce carbon emissions.

For these reasons the application of structural health monitoring to offshore wind turbines could overall lead to a decrease in greenhouse gas emissions and have a positive environmental impact.

Conclusions

The current climate crisis requires a shift towards renewable energies. Wind energy generation Will play a major role. Offshore wind energy can provide greater output due to more predictable weather conditions compared to onshore wind energy and has one of the lowest lifecycle greenhouse gas emissions for any source of energy. For these reasons, there is increased interest and investment in offshore wind turbines.

Some of the difficulties in their operation and maintenance lie in the difficulty of accessing the site. Although remote monitoring has become standard in the industry, structural health monitoring and predictive maintenance still presents some challenges.

Most predictive maintenance strategies in the industry rely on vibration analysis, this work introduces some of the most standard, broadband features to study vibrations in the industry, but it also introduces some novel features that have shown promising results in the field of SHM of WT. Specifically, this project uses spectral entropy as an additional feature to the classical features.

Regarding machine learning, this project features the use of Isolation Forest (a semi supervised method) and One Class SVM (a non-supervised method) as a more realistic approach to SHM of WT compared to supervised methods due to the difficulty of labelling data. The strategy tested in this project (available in https://github.com/clara-9/TFG_public) works under different operating and environmental conditions and provides results based only on the output vibration data gathered by accelerometer sensors.

When the strategy is tested against an initial dataset of experimental data accuracies of 100% are achieved with both models. However, when the accuracies for a different, previously unseen healthy dataset obtains lower accuracies and the features and hyperparameters of the model must be retuned. Excluding the initial healthy state achieves overall accuracies of 93,88% for Isolation Forest and 88,67% for One Class Support Vector.

For all of this, the use of non-supervised and semi-supervised machine learning models is realistic approach to structural health monitoring of offshore wind turbines and has obtained good results when tested against an experimental dataset based on a scaled model.

Economic analysis

The resources required to develop this project are as follows:

	Time	**Salary**	**Cost**
Researcher hours	24 ECTS at 60h/ECTS	760€/80 h monthly	12680€
Supervisor hours	10% of researcher hours	3040€/160 h monthly	2736€
Computer resources	-	-	700€
		Total cost	**16116€**

Table 13 Economic analysis of the project

This calculation considers the net salary for a20h/week researcher position at UPC, multiplied by 1.3 to take into account taxes. In the case of the supervisor, the salary has been calculated by doubling the hourly rate.

The experimental data used for the development of the model was generated in a previous study and will be released in an open-source journal. All cited articles were accessed through open-source journals or access was provided by the university. These costs have not been considered.

The economic impact of the research goes beyond the direct cost of the project. Operation and maintenance costs represent 25% of energy production costs offshore wind turbine maintenance. (van Bussel and Henderson 2001) This is 15% than Operation and Maintenance costs for onshore wind turbines. Moreover, Turnbull reports that up to 8% of these costs can be saved through early maintenance intervention. (Turnbull and Carrol 2021)

Bibliography

Agustyn, Dawid, Martin B. Nielsen, and Ronnie R. Pedersen. “Desing of offshore wind turbine jacket foundations.” *Eurosteel*, 2017.

Ahmat, Tanveer, and Huanxin Chen. “Utility companies strategy for short-term energy demand forecasting using machine learning based models.” *Sustainable Cities and Society*, 2018: 401-417.

Barszcz, Thomas. *Vibration-Based Condition Monitoring ofWind Turbines.* Krakow: Springer, 2019.

Barthelmie, Rebecca J., and Sarah C. Pryor. “Climate Change Mitigation Potential of Wind Energy.” *Climate*, 2021.

Bhattacharya, Buhamoy. *Design of Foundations for Offshore Wind Turbines.* Wiley, 2019.

Bounou, Alexandra, Alexis Laurent, and Stig I. Olsen. “Life cycle assessment of onshore and offshore wind energy-from theory to application.” *Applied Energy*, 2016.

Costa, Ángel M., José A. Orosa, Diego Vergara, and Pablo Fernández Arias. “New Tendencies in Wind Energy Operation and Maintenance.” *Applied Sciences* 11, no. 4 (2021).

Duffuaa, Salih O., and A. Raouf. *Planning and control of maintenance systems.* 2015.

European Environment Agency. “Europe's onshore and offshore wind energy potential:An assessment of environmental and economic constraints.” 2009.

Ezra Pilario, Karl, Mahmood Sahfiee, Yi Cao, Yao Liyun, and Yang Shuang-Hua. “A Review of Kernel Methods for Feature Extraction in Nonlinear Process Monitoring.” *Processes*, 2020.

Faulstish, S., Berthold Hahn, and P.J. Tavner. “Wind Turbine downtime and its importance for offshore deployment.” *Wind Energy*, 2011.

Global Wind Energy Council. “Global Wind Report 2021.” 2021.

Grus, Joel. *Data Science from Scratch.* O'Rielly, 2015.

Haenlein, Michael, and Andreas Kaplan. “A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence.” *California Management Review*, 2019.

Häfele, Jan, Rick R. Dammiani, Ryan N. King, Cristian G. Gebhardt, and Raimund Rolfes. “A systematic approach to offshore wind turbine jacket predesign and optimization: geometry, cost, and surrogate structural code check models.” *Wind Energy Science* 3 (2018): 553–572.

Hernandez, Mauricio, Milad Shadman, and Mojtaba Maali. “Environmental impacts of offshore wind installation, operation and maintenance, and decommissioning activities: A case study of Brazil.” *Renewable and Sustainable Energy Reviews*, 2021.

Higgins, Lindley R., R. Keith Mobley, and Smithm Ricky. *MAintenance Engineering Handbook.* McGraw-Hill, 2002.

Hoffman, Martin W., Stephan Wildermuth, Ralph Gitzel, and Aydin Boyaci. “Integration of Novel Sensors and Machine Learning for Predictive Maintenance in Medium Voltage Switchgear to Enable the Energy and Mobility Revolutions.” *Sensors*, 2020.

Hoxha, Ervin, Yolanda Vidal, and Francesc Pozo. “Damage Diagnosis for Offshore Wind Turbine Foundations Based on the Fractal Dimension.” *Applied Sciences*, 2020.

IEA. *Wind electricity generation, World 1990-2019.* 2019. https://www.iea.org/data-and-statistics/data-browser?country=WORLD&fuel=Renewables%20and%20waste&indicator=WindGen (accessed 10 17, 2021).

—. “World Energy Outlook.” 2018.

Jasiulewicz-Kaczmarek, Małgorzata, and Przemysław Drożyner. “The Role of Maintenance in Reducing the Negative Impact of a Business on the Environment.” In *Sustainability Appraisal: Quantitative Methods and Mathematical Techniques for Environmental Performance Evaluation*, by Marina G Erechtchoukova, Peter A Khaiter and Paulina Golinska, 141-166. Springer, 2013.

Li, Jiawen, Zhenni Li, Yichen Jiang, and Tang Yougang. “Typhoon Resistance Analysis of Offshore Wind Turbines: A Review.” *Athmosphere*, 2022.

Liu, Fei Tony, Kai Ming Ting, and Zhi-Hua Zhou. “Isolation Forest.” *2008 Eighth IEEE International Conference on Data Mining.* 2008.

Manzano-Agugliaro, Francisco, Miguel Sánchez-Calero, Alfredo Alcayde, Carlos San-Antonio-Gómez, and Alberto Jesús, Samerón-Manzano, Ester Perea-Moreno. “Wind Turbines Offshore Foundations and Connections to Grid.” *Inventions*, 2020.

Michal, Wasilczuk¸, Rafal Gawarkiewicz, and Filip Wasilczuk. “Drivetrain of a Wind Turbine.” *Machine Dynamics Research*, 2015.

Norvig, Peter, and Stuart Russell. *Artificial Intelligence: A Modern Approach.* Pearson, 2021.

Puruncajas, Bryan, Yolanda Vidal, and Chritian Tutivén. “Vibration-Response-Only Structural Health Monitoring for Offshore Wind Turbine Jacket Foundations via Convolutional Neural Networks.” *Sensors*, 2020.

PwC. “Predictive Maintenance 4.0.” 2018.

Rizwan, Atif, Naeem Iqbal, Rashid Ahmad, and Kim Do-Hyeum. “WR-SVM Model Based on the Margin Radius Approach for Solving the Minimum Enclosing Ball Problem in Support Vector Machine Classification.” *Applied Sciences*, 2021.

Sandoval, Diego, Urko Letiurondo, Pozo Frnacesc, and Yolanda Vidal. “Low-Speed Bearing Fault Diagnosis Based on Permutation and Spectral Entropy Measures.” *Applied Sciences*, 2020.

Schallehn, Michael, Christopher Schorling, Peter Bowen, and Oliver Straehle. *Beyond Proofs of Concept: Scaling the Industrial IoT.* Bain & Company, 2018.

Schlömer S., T. Bruckner, L. Fulton, E. Hertwich, A. McKinnon, D. Perczyk, J. Roy, R. Schaeffer, R. Sims, P. Smith, and R. Wise. “Technology-specific cost and performance parameter.” Chap. Annex III in *Climate Change*, by T. Bruckner, L. Fulton, E. Hertwich, A. McKinnon, D. Perczyk, J. Roy, R. Schaeffer, R. Sims, P. Smith, and R. Wise Schlömer S. Cambridge University Press, 2014.

Scholkopf, Bernhard, Robert Williamson, Alex Smola, Alex Shawe-Taylor, and John Platt. “Support Vector Method for Novelty Detection.” *Neurips*, 1999: 582-588.

Shoja, Slavash. *Guided wave-based approach for health monitoring of composite structures, Application to wind turbine blades.* Department of Mechanics and Maritime Sciences, Chalmers University of Technology, 2018.

Snyder, Bryan, and Mark J. Kaiser. “Ecological and economic cost-benefit analysis of offshore wind energy.” *Renewable Energy*, 2009.

Statco, Adrian, Fateme Dinmohammadi, Xingyu Zhao, and Valentin Robu. “Machine learning methods for wind turbine condition monitoring: A review.” *Renewable Energy*, 2019: 620-635.

Taylor, Regan, Christopher Beale, and Inalpolat Murat. “Wind Turbine Blade Damage Detection Using Supervised Machine Learning Algorithms.” *Journal of Vibration and Acoustincs* (ASME), 2017.

Turnbull, Alan, and James Carrol. “Cost Benefit of Implementing Advanced Monitoring and Predictive Maintenance Strategies for Offshore Wind Farms.” *Energies*, 2021.

van Bussel, G.J.W., and A. R. Henderson. “State of the art and Technology Trends for Offshore Wind Energy: Operation and Maintenance Issues.” *Offshore Wind Energy*, 2001.

Vidal, Yolanda, José L. Rubias, and Francesc Pozo. “Wind Turbine Health Monitoring Based on Accelerometer Data.” *IX ECCOMAS Thematic Conference on Smart Structures and Materials.* 2019.

Wang, Dahai, Jung Peng, Quian Yu, Yuanyuan Chen, and Hanghang Yu. “Support Vector Machine Algorithm for Automatically Identifying Depositional Microfacies Using Well Logs.” *Sustainability*, 2019.

Weichert, Dorina, Patrick Link, Anke Stoll, Stefan Rüping, Steffen Ihlendfeldt, and Stefan Wrobel. “A review of machine learning for the optimization of production processes.” *The International Journal of Advanced Manufacturing Technology* 104 (2019): 1889–1902.

World Resources Institute. “Climate Watch Historical GHG Emissions.” Washingto D.C., 2021.

Wu, Di, and Da-Wen Sun. “Colour measurements by computer vision for food quality control – A review.” *Trends in Food Science & Technology* 21 (2013): 5-20.

Xie, Mian, and Susana Lopez-Querol. “Numerical Simulations of the Monotonic and Cyclic Behaviour of Offshore Wind Turbine Monopile Foundations in Clayey Soils.” *Journal of Marine Science and Engineering*, 2021.