Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Engineering
Department of Mechanical and Industrial Engineering

Tharanath Thiruthiyappan

Early Fault Detection of Wind Turbine Generator and Rotor Bearings

Master's thesis in Reliability, Availability, Maintainability and Safety (RAMS)
Supervisor: Jørn Vatn
Co-supervisor: Thor Inge Bernhardsen
July 2023

Aneo (https://www.aneo.com/fornybar-kraft/var-kraftproduksjon/bessakerfjellet-vindpark/)

NTNU
Norwegian University of Science and Technology

equinor

Tharanath Thiruthiyappan

Early Fault Detection of Wind Turbine Generator and Rotor Bearings

Master's thesis in Reliability, Availability, Maintainability and Safety (RAMS)
Supervisor: Jørn Vatn
Co-supervisor: Thor Inge Bernhardsen
July 2023

Norwegian University of Science and Technology
Faculty of Engineering
Department of Mechanical and Industrial Engineering

NTNU
Norwegian University of Science and Technology

RAMS
Reliability, Availability,
Maintainability, and Safety

Early Fault Detection of Wind Turbine Generator and Rotor Bearings

Tharanath Thiruthiyappan

June 2023

TPK4950 - Reliability, Availability, Maintainability, and Safety, Master's Thesis

Department of Mechanical and Industrial Engineering
Norwegian University of Science and Technology

Supervisor 1: Jørn Vatn
Supervisor 2: Thor Inge Bernhardsen

“Only a genius can do things his own way. You? You’re no genius.” - *Jeff Haden*

Preface

This master thesis project report is conducted at the Faculty of Engineering Science (IV), as part of the two-year master program in Reliability, Availability, Maintainability and Safety (RAMS) at NTNU, in Trondheim. This thesis was written in the course TPK4950 - Reliability, Availability, Maintainability, and Safety, Master's Thesis, during the spring semester of 2023, to meet the requirements for graduating the Master's program.

This project was carried out as part of the FME NorthWind (Norwegian Research Centre on Wind Energy) program, with NTNU Professor in Reliability and Maintenance Optimization Jørn Vatn from the Department of Mechanical and Industrial Engineering at NTNU as my main supervisor. Industry collaboration with the energy company Equinor headquartered in Stavanger, Norway is established and my co-supervisor from Equinor is Thor Inge Bernhardsen.

RAMS master student's who are currently in their second year of the master degree program should have acquired the required knowledge through the courses TPK4120 - Safety and Reliability Analysis, TPK4140 - Maintenance Management and TPK4450 - Digitalized Solutions to Prognosis, Predictive Maintenance and Safety Analysis, to understand and comprehend this master thesis.

Knowledge of unsupervised fault detection methods, different offshore wind technology, how it operates, common minor and major failure modes and challenges regarding uncertainty with harsh weather conditions related to offshore wind farms can be useful to better understand the scope of this project. Experience in programming languages such as MATLAB is recommended to understand the data analysis and unsupervised fault detection methods used.

Peranamallur, 2023-07-02

Tharanath Thiruthiyappan

Acknowledgment

I would like to sincerely thank the following persons for their great help and support during the completion of this master thesis. My supervisor Jørn Vatn from NTNU and co-supervisor Thor Inge Bernhardsen from Equinor, have been vital in the successful completion of this master thesis. Their guidance through weekly meetings have been crucial and have helped me take the right approach to solving problems and making progress. I would like thank Shen Yin and Yiliu Liu for teaching the RAMS course TPK4450 and for sparking my interest in data driven fault diagnostics and prognostics.

I sincerely thank Viggo Gabriel Borg Pedersen who is a professor at NTNU, for his help with obtaining Bessakerfjellet wind turbine bearing vibration data from NTNU's SKF @ptitude observer. Additionally, I thank Tore Rasmussen from Aneo, for kindly sharing SCADA data of the Bessakerfjellet wind farm. The access to these wind turbine data have been critical for carrying out the master thesis work. I thank Erik Solberg from SKF, for the engaging and informative guidance on bearing vibration analysis.

I would like to thank the following people Eldbjørg Lyssand Eide, Asle Bjørn Emorsten and Olav Tu Husveg from Equinor their help during various stages of my master's thesis. Additionally, I sincerely thank my main contact person Thibaut Forest from the Equinor Renewable's department for the insightful discussions on wind turbine anomaly detection and condition based maintenance, and for his support and guidance for my master thesis. I would like to thank Rohit Agrawal from MathWorks for his in-depth guidance with the MATLAB toolboxes and for the interesting discussions on supervised and unsupervised fault detection models.

Most importantly, I sincerely thank my family for their unconditional love and moral support during the completion of this thesis.

T.T.

Remark:

Given the opportunity here, the RAMS group would recognize Professor Emeritus Marvin Rausand for the work to prepare this template. Some modifications have been proposed by Professor Mary Ann Lundteigen and Professor Jørn Vatn. In the preparation of this revised version important material from Associate Professor Anita Romsdal has also been included.

Executive Summary

The sustainable energy transition depends on the use of renewable energy sources and wind turbines play a central role. Fault diagnostics and prognostics of major failures in offshore and onshore wind turbines, causing long production downtime, have been a key research area. However, far greater attention is required for the early detection of abnormal conditions, minor faults and failures in wind turbines. Since the accumulation of minor failures can gradually lead to unexpected major failures, reducing the wind turbine reliability significantly. Resolving minor failures in offshore wind turbines has higher complexity than wind turbines onshore. The production downtime at offshore locations is amplified when harsh weather conditions, difficulties with maintenance scheduling and safe execution of maintenance is considered. Therefore, research on fault detection in wind turbine components to provide wind farm asset managers with early indication of faults for maintenance planning is the key focus in this master thesis. Fault detection of minor failures can improve all aspects of reliability, availability, maintainability and safety of wind turbines.

Emphasis is placed on offshore wind turbine minor failures such as the main shaft bearings, which have a critical role. Due to generally strict data sharing procedures in the industry for offshore wind turbine components condition monitoring data, onshore wind turbine data had to be used. Unlabeled condition monitoring vibration data of the generator and rotor bearing from 2 onshore wind turbines at the Bessakerfjellet located in Trøndelag, Norway is used. The health condition of the bearings is unknown, due to lack of wind turbine maintenance logs. The raw vibration data is first analyzed using statistical data visualization, then frequency domain vibration analysis and finally a correlation study between the bearing vibration and wind turbine rotor speed. This helped uncover the health condition of the bearings. Major findings respectively are the effect of seasonality aspects showing higher bearing vibration in winter months, developing minor defect in one wind turbine generator bearing inner ring and cage components, and machine resonance occurring at a particular rotor speed in both wind turbines.

Unsupervised fault detection methods are trained using the generator and rotor bearing vibration data of one wind turbine showing the most anomalous data. A contamination factor specifying the percentage of abnormality and outliers contained in the data is set based on the author's analysis and from expert judgement. The models trained are isolation forest, local outlier factor, one class support vector machines and Mahalanobis distance. The performance of the models is evaluated from emphasis on findings from the raw vibration analysis and correlation study. From a visual comparison of the detected outlier and normal data points of the four models using scatter plots of rotor speed and vibration, the Mahalanobis distance-based model was able to detect signs of machine resonance, outliers away from the data cluster and low false detection within the cluster of normal bearing vibration data. The model training has some limitations including low data size used to train the models and the judgement of abnor-

mal, outlier percentage in the unlabeled bearing vibration data using the contamination factor. Despite this, the results from the thesis show a strong early outlier detection performance of the unsupervised fault detection model Mahalanobis distance and has strong implications for the use in the onshore and offshore wind industry. This will support the decision-making process for wind farm asset managers. Evaluation of fault detection model performance using offshore wind turbine generator and rotor bearing condition data is required to further support the results from this thesis.

Contents

Chapter 1

Introduction

The first chapter of this thesis puts forward the background of the problem to be investigated, together with the formulated objectives to be solved. The approach for solving the problem and meeting the objectives are presented. The contributions and limitations of the study, and how the report is organised is presented in the outline.

1.1 Background and motivation

The use of onshore and offshore wind turbines (WTs) for producing clean, renewable energy plays a key role in the transition to sustainable energy. Scaling up the production and use of renewable energy is crucial to a successful energy transition and to meeting the global climate objectives of the 2015 Paris Agreement, which calls for halving the greenhouse gas emissions already by 2030 (Hutchinson and Zhao, 2023).

The onshore wind industry is a proven and a mature technology that has an extensive global supply chain and now the offshore wind industry is expected to grow rapidly in the coming years (Technologies, 2022). Greater consistent wind speeds offshore than when compared with onshore locations, means deploying larger wind farms with higher production capacity in the sea can take advantage of this natural resource (Technologies, 2022). Significantly higher wind power generation is achievable, in contrast to onshore wind farms. Currently, there is an increasing trend towards the development of larger offshore WTs and towards installation at wind farm sites farther out in sea, with limited accessibility especially during the harsh winter months (Fischer and Coronado, 2015). The O&M phase pose great challenges related to offshore wind farms. The harsh weather conditions, largely variable aerodynamic, gravitational, centrifugal and gyroscopic loads induce higher failures rates in offshore WT systems than onshore, resulting in overall higher frequency of faults and failures (Badihi et al., 2022).

Presently, the operational expenditure (OPEX) account for anywhere from 10% to 30% of the total energy generation cost of onshore WTs, whereas in offshore WTs the OPEX can surge up to

25% to 50% (Badihi et al., 2022). Mainly caused by the distance from shore to the offshore wind farm location, the number of WTs installed, the O&M strategy and logistical setup (Christensen, 2022). The remote locations of the offshore WT sites combined with difficult weather conditions directly reduces the time window to perform maintenance tasks safely (Zhang et al., 2022). Safety regulations restricts the maintenance activity to be conducted in daytime, with low air humidity, minimum allowable environment temperature of 10°C and wind speeds must below 12 m/s (Badihi et al., 2022). This together with the enormous size of offshore WTs, high component failure rates and harsh weather conditions, increases the complexity of maintenance scheduling and execution. Resulting in greater logistics and maintenance costs (Badihi et al., 2022). Hence, a cost-optimised maintenance strategy is needed and motivates the extension of condition monitoring (CM) to additional offshore WT components (Fischer and Coronado, 2015).

This issue must be addressed and hence, to upkeep offshore WTs cost-effectively, and to ensure efficient production and financial viability of wind power, it is crucial to maintain the offshore WTs reliability and availability (Badihi et al., 2022). Predictive maintenance (PdM) is an appealing strategy for the offshore wind industry (Zhang et al., 2022). PdM aims to monitor the condition and performance of mechanical components through condition monitoring technologies, and in combination with data collection and analytic's, can provide wind farm asset managers with early fault detection and predict remaining useful life (RUL) estimates WT component failures. Implementing the PdM strategy can reduce the unexpected failures of critical offshore WT components. This will ultimately increase the WT production availability. The number of trips to the sea for maintenance activities can be reduced from better maintenance planning based on the asset condition. The maximum working life of critical WT components can be utilized, catastrophic unexpected damages can be avoided and the safety of personnel can be improved (Fox et al., 2022).

Generator and rotor bearings are critical WT components operating in difficult conditions. The WT main shaft bearings support the weight of the rotor where the rotor blades are assembled, they experience greater loads caused by the chaotic wind, and are required to rotate smoothly and transfer the torque to the generator where electricity is produced (Kihlström, 2019). The research article by Faulstich et al. (Faulstich et al., 2011) show the annual failure rate of minor failures in WT sub-assemblies are much greater than for major failures, while the WT downtime from major failures is much higher than from minor failures. Here the minor failures are classified as causing downtime of ≤ 1 day, and major failures as downtime of > 1 day. For this reason, a majority of the research is focused on the WT sub-assemblies which cause major failures, as found through the literature review on fault diagnostics and prognostics conducted in the specialization project (Thiruthiyappan, 2022) in the autumn semester of 2022. The minor failures in onshore WTs can be resolved quickly, whereas in offshore WTs, the limited accessi-

bility, high waiting, travel and corrective maintenance repair has the potential to amplify the WT downtime and decrease the production availability significantly (Faulstich et al., 2011). In the case of Equinor's Hywind Tampen, the offshore wind farm will be located 140 km from the shore (Tampen, 2021). Faulstich et al. (Faulstich et al., 2011) state the wind farms availability may decline further for wind farms located in the excess of 50 km from shore. Hence, generator and rotor bearing failures can cause long WT downtime and high repair costs. Therefore, focus on reliability improvement, effective condition monitoring with early fault detection, diagnosis and prognosis is required to improve the lifetime of the WT bearings.

From the specialization project (Thiruthiyappan, 2022), an in-depth literature review of different modelling techniques used within fault diagnosis and prognosis were studied. Many authors use various classifications for the methods, however, the review paper by (Gao and Liu, 2021) for wind turbine systems was chosen for its simplicity. For fault diagnosis, model-based, signal-based, knowledge-based and the combination as hybrid method were investigated. For fault prognosis, model-based, data-driven and hybrid approach were investigated. Hybrid models combine the strengths of different fault diagnostic and prognostic techniques, in order to reduce the limitations of individual models. This increases the robustness of fault detection and RUL prediction accuracy with lower uncertainty, to aid the decision making process for maintenance planning (Zhang et al., 2022). The techniques/models within each of the methods are numerous as found for example in knowledge-based fault diagnostics, which use data-driven computational intelligence together with historical WT condition data. It is not straight forward and clear which techniques and models will be the most optimal to use for any given case. Therefore, in relation to the findings from the literature review conducted in (Thiruthiyappan, 2022), further investigation and researching within the fault diagnostics field of testing with different early fault detection methods are required for providing the necessary decision support for wind farm asset managers and is the main focus of this thesis.

Relevant articles which treat problems similar to the focus of this thesis are given in the specialization project (Thiruthiyappan, 2022). Turnbull et al. (Turnbull et al., 2019) apply Fourier analysis and support vector machine (SVM) algorithms for generator bearing failure prediction using high-frequency vibration data. Tutiven et al. (Tutivén et al., 2022) applied one class SVM for early fault diagnosis of the WT generator bearings based on Supervisory control and data acquisition (SCADA) data. Wang et al. (Wang et al., 2022) applied parameter optimized isolation forest for early fault detection of rolling element bearings using normal and abnormal vibration data. The detection results showed more effective fault detection than other methods such as Local outlier factor (LOF), one class SVM and Mahalanobis Distance.

The following book and RAMS course are used for studying and gaining knowledge within the thesis topic.

- System Reliability Theory - Models, Statistical Methods, and Applications. (Rausand et al.,

2020). This book provides a thorough introduction to system reliability theory and introduces many analytical methods for reliability analysis.

- TPK4450 - Digitalized Solutions to Prognosis, Predictive Maintenance and Safety Analysis. (Yin, 2022). This course focuses on the four following topics of competence which are 1) Diagnosis using statistical and data driven methods, 2) Prognosis using model based and data driven methods, 3) Decision optimization and 4) Reliability analysis & standardization of safety-critical systems.

A major challenge found in literature is the comparison challenges between the hybrid models for fault diagnosis and prognosis to study which is more superior in performance. Hence, there is a need for a common evaluation standard/benchmark to analyse and compare the performance of the various hybrid based models together. This will allow researchers and experts to better assess the performance and capability of hybrid models.

For the hybrid-based fault diagnosis and prognosis, the literature showed interesting combinations of the different fault diagnosis methods such as model-based and signal-based methods, and for prognosis, both model-based and data-driven were combined. In addition, it was more common to combine statistical approaches and machine learning methods within data-driven methods for fault prognosis. Further combinations of different methods and techniques is required to increase the possibilities for implementing the most optimal performing hybrid model for different industry use cases. A major limitation of hybrid models is the high required time to design and safely implement the model to the use case (Zhang et al., 2022). This requires additional research to find ways to overcome this limitation. The literature review from the specialization project (Thiruthiyappan, 2022) indicated the hybrid model approach area to be still in the early phase, and hence requires further research and development to reap the benefits.

1.2 Objectives

The main objective of the master thesis is to investigate and compare different early fault detection techniques and models for WT components. This will provide wind farm asset managers with better tools for informed decision making for WT maintenance scheduling and execution.

1. Explore state-of-the-art technology and present the industrial background for wind turbines.

2. Present a theoretical background on fault diagnostics, workflow for fault detection model development and unsupervised fault detection models. Present different maintenance strategies for wind turbines and maintenance practices used in the industry.

3. Present theory of industrial rolling bearing vibration analysis, the bearing fault frequencies and the bearing wear fault stages development.
4. Investigate the acquired WT condition data through statistical data visualization, frequency domain analysis and correlation analysis between the variables to identify trends, outliers, and patterns in the data.
5. Train different fault detection, diagnosis models using wind turbine SCADA, or condition monitoring systems (CMS) data, to provide early warning of anomalies.
6. Compare the different models based on fault detection performance.

1.3 Approach

1. Learn about Anomaly detection, Condition-based maintenance and Predictive analytics workflow used by Renewable's (REN) Integrated Operation Center (IOC) in Equinor.
2. Get access to Equinor operated offshore wind turbines SCADA data, or condition monitoring systems (CMS) data and relevant maintenance logs.
3. Investigate the WT data for preliminary analysis to decide what data set to use for fault detection, prediction model development.
4. Investigate MATLAB programming language, together with relevant toolboxes such as Predictive Maintenance toolbox, Statistics and Machine Learning toolbox and the diagnostic feature designer app for extracting relevant condition indicators.
5. Research and implement different early fault detection models using the WT data.
6. Compare the different models.

1.4 Contributions

- **Improved fault diagnosis techniques**

  Correlation analysis between WT generator and rotor bearing vibration and rotor speed, for identifying early signs of resonances in WT machinery. This allows for timely identification of vulnerable WT components and locate the root cause of the fault. Minor failures which can lead to major failures can be repaired earlier.

Frequency domain vibration analysis for identifying early signs of bearing component defects in wind turbine generator bearing, allowing for early intervention to mitigate the developing faults.

These techniques improves the tools wind farm operators and maintenance crew can use for early fault detection of WT bearings and helps support the decision making process for maintenance planning.

- **Early outlier detection**

Systematically trained unsupervised early fault detection models using unlabeled generator and rotor bearing vibration data, to identify which models perform best in detecting machine resonance and outliers points far from normal bearing vibration data. This contributes to early detection of bearing faults and can improve the reliability of WT bearings used in the industry.

1.5 Limitations

- **Time Constraint**
  The short time frame of the spring semester 2023 is the major limitation to conduct meaningful work and achieve good results.

- **Lack of early access to WT SCADA and CMS data**
  Gaining access to historical wind farm SCADA, CMS data and relevant maintenance logs proved to be a difficult challenge, causing significant delay in the thesis work. The approval process for getting access to live operational and historical SCADA data of Equinor operated offshore wind farms was lengthy process, mainly due to strict data access procedures and policies within the Renewable's department at Equinor. The SCADA data is considered to be confidential information requiring approval from all the partners of each joint venture company, who technically also own the SCADA data.

  As a result of this delay and agreement with my co-supervisor Thor Inge Bernhardsen, focus was shifted during the middle of April to obtaining Bessakerfjellet wind farm generator, rotor bearing vibration data from NTNU's SKF @ptitude observer. Approval to use the data had to be obtained from the wind farm operator Aneo. SCADA data of the onshore wind turbines were acquired at the end of April.

- **Unlabeled WT generator and rotor bearing vibration condition data**
  The already ongoing collaboration with Equinor, resulted in Aneo not having enough resources to assist with providing guidance and maintenance logs of the Bessakerfjellet wind

farm. This increased the fault detection model development difficulty with limited knowledge of when maintenance activities were performed. Separating normal and abnormal vibration behavior became an additional challenge. Resulting in only unsupervised fault detection methods to be used.

1.6 Outline

An overview of how the remaining part of the report is organized:

- Preface: Contains practical information about what has been done, and where the work has been carried out. The assumed background of the reader is specified here.
- Acknowledgments: Gratitude expressed to those who have supported in completing this Master's thesis. Acknowledging both family and professional individuals.
- Executive Summary: Contains a through summary of the work carried out in the thesis and the importance of findings for the wind industry.
- Chapter 1. Introduction: Presents the background of the problem to be investigated, the formulated objectives to be solved. The approach, contributions and limitations of the study and an outline of how the report is organised is presented.
- Chapter 2. Industrial background: Presents the state of the art of offshore wind turbine technology, how they function, wind turbine structures, installation capacities, common faults and failures and condition monitoring techniques.
- Chapter 3. Theoretical background: Presents the theoretical background required for solving the presented problem, which are fault diagnostics, the workflow for fault detection model development, reliability and maintenance strategies for WTs, industrial rolling bearing vibration analysis allowing the reader to follow the work process.
- Chapter 4. Wind Turbine SCADA & CMS data: Presents the data acquisition process, the type and format of wind turbine SCADA, CMS from Equinor and Aneo. The wind farm components of where the data is measured is presented, along with the programs 'SKF @ptitude observer' and 'MATLAB' which required for fault detection model development.
- Chapter 5. Case study: Presents the case study of the wind turbine generator, rotor bearing and shaft vibration data statistical visualization, frequency domain analysis and correlation analysis. The data pre-processing steps and the unsupervised fault detection models training are presented.

- Chapter 6. Results: Presents comparison scatter plots of the detected normal and outlier data points of the unsupervised fault detection models trained.
- Chapter 7: Discussion: Presents an evaluation of supervised fault detection techniques with the assumption of condition labels initially used, the limitations with model training parameters, and the strengths and weakness of the unsupervised fault detection models used.
- Chapter 8. Presents a summary of if the thesis objectives have been met, the key findings and the applications for early fault detection in WT main shaft bearings for the wind industry, and recommendations for future work are given.
- Appendix A: Acronyms
- Appendix B: Supplementary information, figures and MATLAB code
- Bibliography

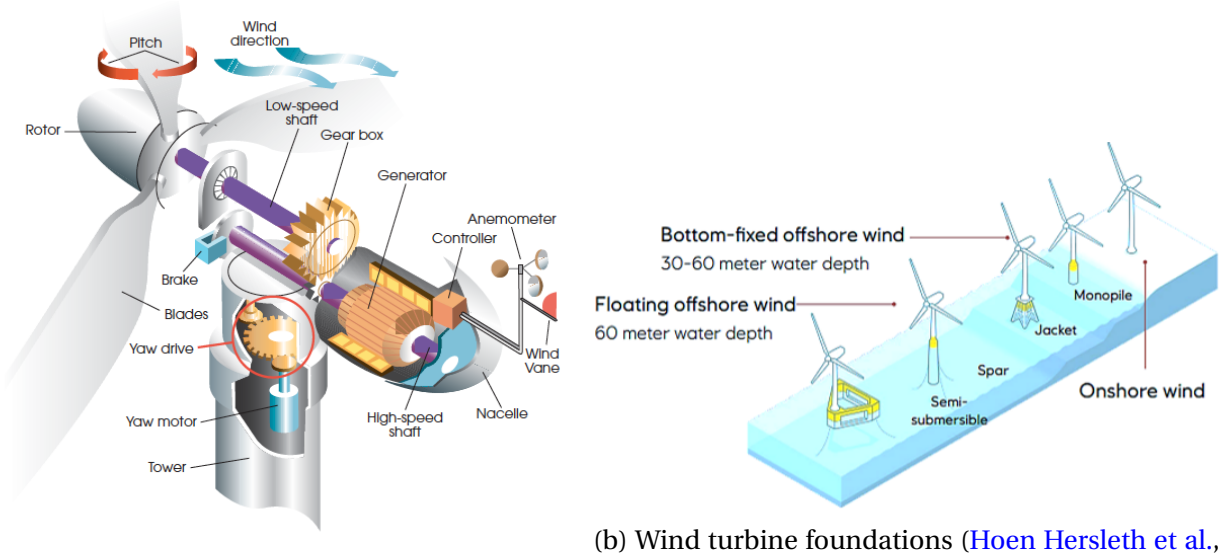
Chapter 2

Industrial background

This chapter puts forward the state of the art of wind turbine technology, how they work, different wind turbine structures, current wind power installation capacities, reliability and maintenance strategies used for wind turbines, common major faults and failures of wind turbines and the commonly used condition monitoring techniques.

2.1 State-of-the-art technology in WTs

2.1.1 How WTs work?



(a) Wind turbine components (Energy, [n. d.])

(b) Wind turbine foundations (Hoen Hersleth et al., 2021)

Figure 2.1: Wind turbines

WTs function by harnessing the power of wind flowing on the blades. Figure 2.1a shows an overview of the main WT components which are required for its function. The kinetic energy

of the wind is converted into mechanical energy with the rotation of the rotor and low-speed shaft. The gearbox receives the input of the rotating low-speed shaft and using a system of gears, the rotation is significantly increased and is directly linked to the high-speed shaft (Martin Gill, 2021). The gearbox is responsible for ensuring that the rotations per minute (rpm) of the high-speed shaft is suitable for operating the turbine generator at its optimum electricity generation speed (Martin Gill, 2021). Hence, the generator converts the mechanical energy into useful electrical energy.

2.1.2 Wind turbine structures

Various WT foundation technologies exist and for offshore wind, there are many types available. The main drivers for the selection of foundation for offshore WTs are from the WT size, the local water depth, the significant wave height conditions, as well as the sea bed soil conditions. Figure 2.1b shows the available foundation types for onshore and offshore wind, and after considering the critical factors mentioned, the most suitable foundation type is selected. From figure 2.1b, monopile and jacket foundation types are permanently fixed to the seabed for water depths below 60 meters. For water depths greater than 60 meters, floating foundation types such as spar and semi-submersible are used together with mooring lines fixed to the seabed. Water depths at Equinor's Hywind Tampen wind farm in the north sea is between 260 and 300 meters, and a spar floating foundation type with concrete structure attached to a shared anchoring system is used (Tampen, 2023).

2.1.3 Wind power installation capacities

The annual global wind report 2023, published by the Global Wind Energy Council (GWEC) shows that in 2022, 77.6 gigawatt (GW) of new wind turbine installations were connected to the electricity grid (Hutchinson and Zhao, 2023). This includes the installation of 8.8 GW of new offshore WTs, increasing the global offshore wind capacity to 64.3 GW (Hutchinson and Zhao, 2023). The total installed wind capacity reached 906 GW at the end of 2022, with a year-on-year (YoY) growth of 9% (Hutchinson and Zhao, 2023). The bar chart in Figure 2.2 shows the combined new onshore and offshore installations outlook from 2023 to 2030 in colour blue, giving the reader an idea of the increasing production and deployment of WTs globally. In addition, an indication of the annual capacity gap to meet net zero goal by 2050 scenarios is shown in colour grey. This indicates the need for further growth in development and deployment of WTs to stay on track to meet the 2015 Paris Agreements.

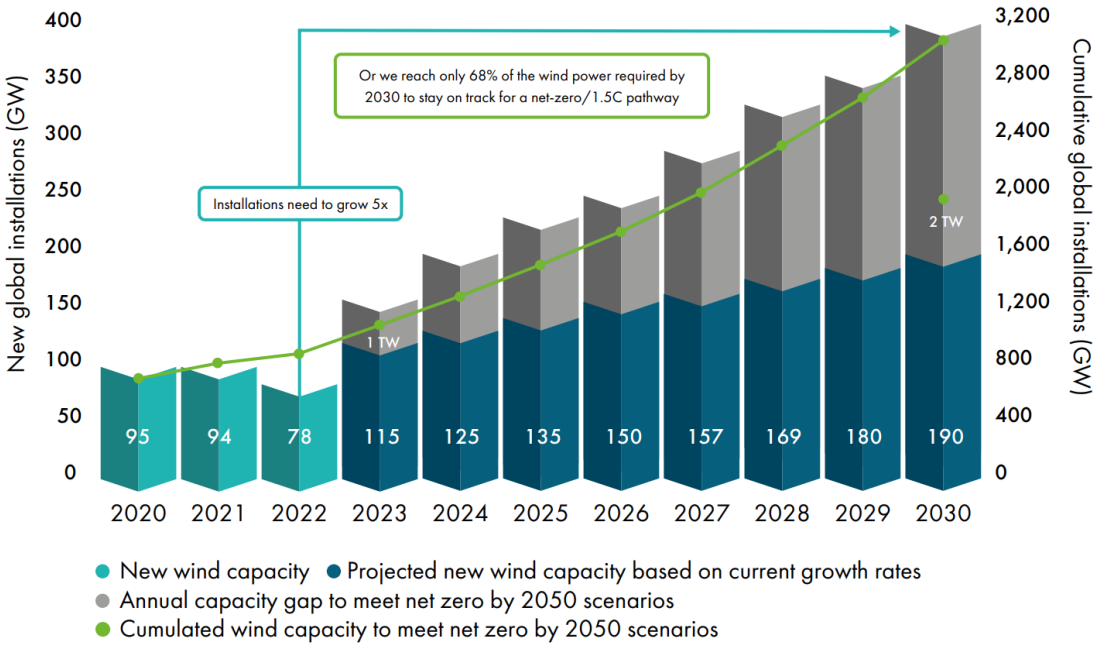


Figure 2.2: Projected and required new wind capacity outlook to 2030 (Hutchinson and Zhao, 2023)

Numerous energy provider companies globally have set concrete goals to reach net-zero emissions. Equinor is determined to become a global offshore wind energy major by investing heavily and have built a robust offshore wind portfolio (wind, 2023). A total of 60 megawatt (MW) of floating offshore wind (FOW) was commissioned by Norway in 2022, resulting in Europe leading the path to FOW (Hutchinson and Zhao, 2023). Equinor has the potential to be a world leader in FOW farm technology. By 2030, Equinor plans to reach an installed net capacity of 12-16 GW for offshore wind (wind, 2023). The latest FOW farm project by Equinor includes Hywind Tampen with a system capacity of 88 MW, which is estimated to provide 35% of the necessary annual electricity demand for the Snorre and Gullfaks oil and gas fields located at the Norwegian North Sea (Tampen, 2023).

2.2 Major and minor faults, failures of WTs

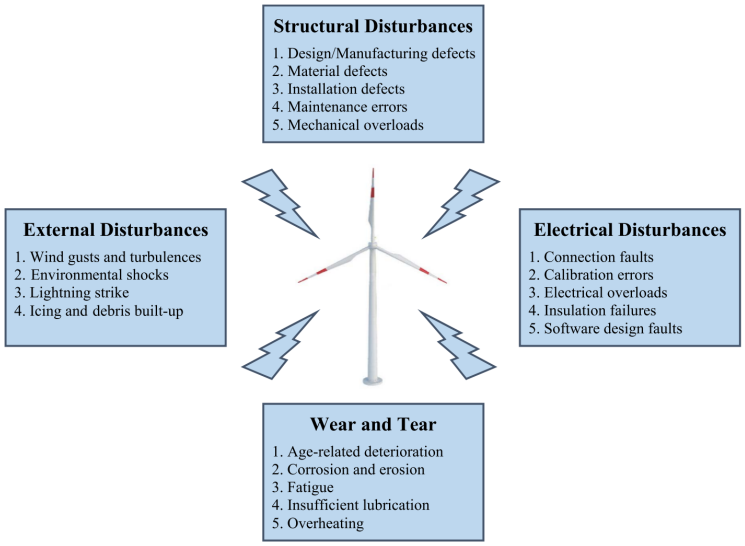


Figure 2.3: Common root causes of faults and failures in wind turbines (Badihi et al., 2022)

Offshore WTs experience a range of external loads from the offshore environment and this directly leads to greater component failure rates than when compared with onshore WTs (Badihi et al., 2022). Figure 2.3 shows a classification of the common failure modes and their root causes into external, structural, electrical disturbances and the ordinary wear and tear. Dao et al. (Dao et al., 2019) state the average failure rate of offshore WTs is greater than onshore WTs, however the onshore average production stop rate is slightly higher than offshore. On the contrary, the total downtime of a stop in production is roughly double that of onshore WTs relatively, mainly due to harsher operating conditions and the additional challenge with accessibility of the WTs for maintenance (Dao et al., 2019). Maintenance during downtime is normally scheduled during winter periods with low wind speed and commonly in the summer, this allows for securing higher wind farm availability during the winter months with greater average wind speeds (Guide_to_an_offshore_wind_farm, 2019).

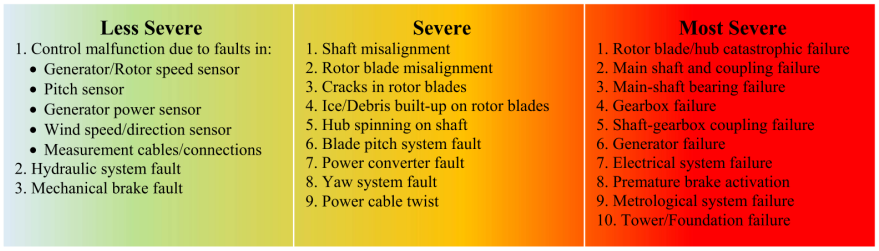


Figure 2.4: Major faults and component failures in wind turbines (Badihi et al., 2022)

Badihi et al. (Badihi et al., 2022) classify the major WT fault and component failures in figure

2.4 in terms of the severity from less severe, severe and most severe. This gives an idea of the seriousness of faults, failures and its effect on the WT. The faults and failures in the most severe list for instance the rotor blade/hub catastrophic failure, are critical leading to production loss of the WT. The faults listed in the less severe and severe contribute only a partial decrease in the WT production output capability, which however may need urgent repair to decrease the damage caused (Badihi et al., 2022).

Faulstich et al. (Faulstich et al., 2011) provide an overview of reliability characteristics of different WT sub-assemblies, based on available onshore WT data. Figure 2.5 shows the major and minor WT annual failure rates and corresponding downtime. A key finding is that mechanical sub-assemblies such as the drive train, support & housing, generator and gearbox fail less frequently than the electrical and electronic sub-assemblies, however the downtime due to mechanical failures are more significant. To conclude, more frequent failures cause less downtime, whereas, less frequent failures result in high downtime.

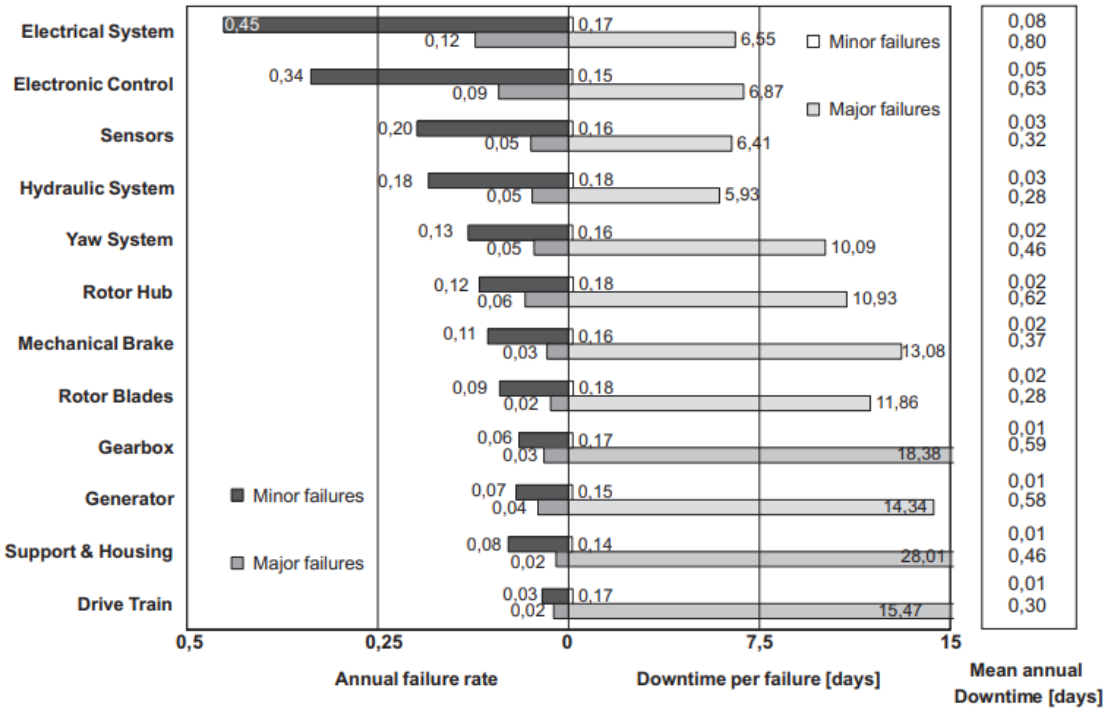


Figure 2.5: Reliability characteristics of different onshore WT sub-assemblies (Faulstich et al., 2011)

2.2.1 Condition monitoring techniques

CBM for WTs require the implementation of efficient and cost-effective condition monitoring systems (CMS), based on available CMS techniques found across in other industries and have recently seen significant development (Badihi et al., 2022). The purpose of condition moni-

Table 2.1: Condition monitoring technologies for WTs (Badihi et al., 2022)

Signal		Monitoring Hardware			Condition Monitoring Capabilities (FDD/LTP)				Monitored Components
Source	Name	Nature	Operation	Rel. Cost	Fault Detection	Fault Location	Fault Identification	Fault Prognosis	
Process parameters monitoring	Vibration	Intrusive	Online	Low to medium	Yes	Yes	Yes	Yes	Rotor blades, generator, drivetrain (gearbox, bearings…), and tower
	Strain	Intrusive	Online	High	Yes	Yes	Yes	Yes	Rotor blades, tower, foundation, and drivetrain
	Torque	Intrusive	Online	High	Yes	Yes	Yes	Yes	Rotor, generator, drivetrain (gearbox, shaft…), and tower
	Shock pulse	Intrusive	Online	Low	Yes	Yes	No	No	Bearings (used in gearbox, yaw system or blades)
	Acoustic emission	Intrusive	Online	High	Yes	Yes	Yes	Yes	Rotor blades, generator, drivetrain (gearbox, bearings…), and tower
	Temperature	Intrusive	Online	Low to medium	Yes	Possible	No	No	Generator, converter, drivetrain (gearbox, bearings…), nacelle, and transformer
	Oil debris/quality parameters	Intrusive	Online /Offline	Medium to high	Yes	Possible	Possible	Possible	Generator and drivetrain (gearbox, bearings…)
	Electrical effects	Non-intrusive	Online	Low	Yes	Yes	Yes	Yes	Rotor, generator, converter, sensor, actuator, drivetrain (gearbox, bearings…), and tower
Non-destructive inspections	Machine vision	Non-intrusive	Online /Offline	Low to medium	Yes	Yes	Possible	No	Rotor blades, nacelle, tower, and foundation
	Ultrasound (ultrasonic scanning/readings)	Intrusive/Non-intrusive	Online /Offline	Medium to high	Yes	Yes	Yes	Possible	Rotor blades, nacelle, drivetrain, and tower
	Thermography	Non-intrusive	Online /Offline	High	Yes	Yes	Yes	Possible	Rotor blades, generator, converter, drivetrain (gearbox, bearings…), nacelle, and transformer
	Radiography (X-ray inspections)	Intrusive/Non-intrusive	Online /Offline	Medium to high	Yes	Yes	Yes	Possible	Rotor blades, nacelle, and tower

Note 1: The provided information includes any technology readiness levels from R&D stage to already commercialized systems.
Note 2: Wind turbine's "bearings" refer to both transmission bearings (generator, gearbox, main bearings) and adjustment bearings (blade and yaw bearings).

toring is detecting the changes in equipment condition, which indicate some developing fault (Fischer and Coronado, 2015).

Table 2.1 summarize the current CMS for WTs with information on monitoring hardware, the condition monitoring capabilities and suitable components for monitoring. The general CMS used include vibration, strain, torque, acoustic emission, temperature oil debris and quality analysis. Fisher and Coronado. (Fischer and Coronado, 2015) state the commercially available CMS in WTs are dominated by cost-effective vibration-based systems and the suitable monitoring components include rotor blades, tower oscillations and bearings in the rotating drive train, generator and gearbox. The vibration signals are mainly analysed using the time and frequency domain techniques for fault diagnosis and prognosis (Crabtree et al., 2014). Frequency domain analysis use spectrum analysis and Fast-Fourier Transform (FFT) (Fischer and Coronado, 2015). It is also common to use time-frequency analysis techniques (Nie and Wang, 2013).

Significant developments in the onshore and offshore wind industry have led to employing condition monitoring technologies with integration to the SCADA system (Pinar Pérez et al., 2013). Condition monitoring under condition-based maintenance comprises of offline periodic inspections and real-time online condition monitoring using sensors with hardware signal-based, mathematical model-based and or hybrid techniques. Offline condition monitoring requires stop in production of WTs for safe inspection. For offshore WTs with long intervals per

inspection, real-time condition monitoring can provide up to date information of WT condition during operation. Hardware signal-based technique for WT condition monitoring analyse the output signals of permanently assembled hardware sensors coming from SCADA systems or other data acquisition systems (Badihi et al., 2022). In addition, the operating and environment conditions including electrical properties, control properties and temperature properties of all WTs are monitored and recorded to the SCADA system roughly every 10-minutes, which can be used for both fault diagnostics and prognostics (Fischer and Coronado, 2015).

Artigao et al. (Artigao et al., 2018) state there currently is no industry agreement with CMS for WT components which is applicable and valid across all WT sizes and technologies, hence they suggest further research and development in CMS.

Chapter 3

Theoretical background

This chapter presents the relevant theory required to understand and solve the problem formulated in chapter 1. The main theory behind fault diagnosis, unsupervised fault detection methods and the workflow for fault detection development are presented. The maintenance strategies used by Equinor for their offshore wind farms and theory about industrial rolling bearings vibration data analysis are presented.

3.1 Fault diagnosis

The aim of fault diagnosis is to monitor whether a system is in a healthy or faulty state, detect the malfunctions early and determine where the fault is and asses the severity of the fault (Gao and Liu, 2021). From the RAMS course TPK4450 compendium (Barros, 2021), the definition of a symptom and fault are

☞ **A symptom**: is the abnormal deviation of an observable quantity.

☞ **A fault**: is an abnormal deviation of at least one characteristic of a component or a system (structural changes, parametric changes).

Fault diagnosis comprises of three different steps which are fault detection, fault isolation and fault estimation. Fault detection determines if a fault is present or not in an equipment and it solely relies on the observation of the symptom(s) in the component or system (Barros, 2021). Fault detection, also referred to as anomaly detection requires only the nominal behaviour of a monitored system and any behaviour outside this is considered as an anomaly (Barros, 2021). Here it is determined the time of the fault occurrence on a system (Blanke and Schröder, 2006). Fault isolation locates the faulty component in a system, and fault estimation assesses the magnitude, the kind and severity of the fault which has occurred (Blanke and Schröder, 2006).

Statistical and data-based approaches are used for fault diagnosis. The three spaces are distinguished as

- The hypothesis/source space $\mathscr{H}$
- The observation space $\mathscr{X}$
- The decision space $\mathscr{D}$

For data-based approaches, it is assumed we do not know the probability density functions $p_i(x)$ which characterize the $n+1$ hypothesis. The hypothesis are instead characterized by a set S of m observations, which constitute samples of the classes ω_i and are respectively associated to the hypothesis H_i (Barros, 2021). Hypothesis testing is used for making decisions and in the case of when $n = 1$ (2 $classes$, 2 $hypothesis$) we have

$$\begin{cases} H_0 & : X \,\epsilon\, \omega_0 \\ H_1 & : X \,\epsilon\, \omega_1 \end{cases} \tag{3.1}$$

Here we want to find a partition $\mathscr{X}_0, \mathscr{X}_1$ of the observation space $\mathscr{X}$ which is defined by the observations made on X (Barros, 2021). The general decision structure becomes

$$\delta_f(x) = \begin{cases} 0 & D_0 : H_0 \; is \; accepted \quad if \quad f(x) < 0 \\ 1 & D_1 : H_1 \; is \; accepted \quad if \quad f(x) \geq 0 \end{cases} \tag{3.2}$$

Here "the function f is determined within a class of functions where Θ is a set of parameters that have to be optimized:" (Barros, 2021).

$$\mathscr{F} \; = \; \{f(x,\theta)/\theta \; \epsilon \; \Theta\} \tag{3.3}$$

To assess the performance, the following confusion matrix in table 3.1 is utilized.

Table 3.1: Confusion Matrix with a simple case where m = 2 (Barros, 2021)

Confusion Matrix		Hypothesis $\mathscr{H}$	
		H_0	H_1
Decision $\mathscr{D}$	D_0		$type \; II \; (\beta)$
	D_1	$type \; I \; (\alpha)$	

The confusion matrix allows for finding out the number of misclassifications classified by two types which are false alarm and non detection. A false alarm is a $type \; I(\alpha)$ and occurs when D_1 is decided, when H_1 is accepted while H_0 is actually true (Barros, 2021). A non detection is a $typeII(\beta)$ and occurs when D_0 is decided, when H_0 is accepted while H_1 is actually true (Barros, 2021). The matrix in table 3.1 shows the type I and II errors.

3.2 PdM algorithm development for fault detection or prediction model

When a good business case is available for implementing CBM and or PdM, developing algorithms for CBM and PdM requires a well planned strategy to precisely assess the machine condition, early detection of machine incipient faults and for forecasting when machine failure will occur (noa, 2023a). Figure 3.1 shows the workflow for PdM in section (a) and a deeper look at the workflow for pre-processing data and to identifying suitable condition indicators on the MATLAB application 'Diagnostic Feature Designer' (DFD) provided in section (b).

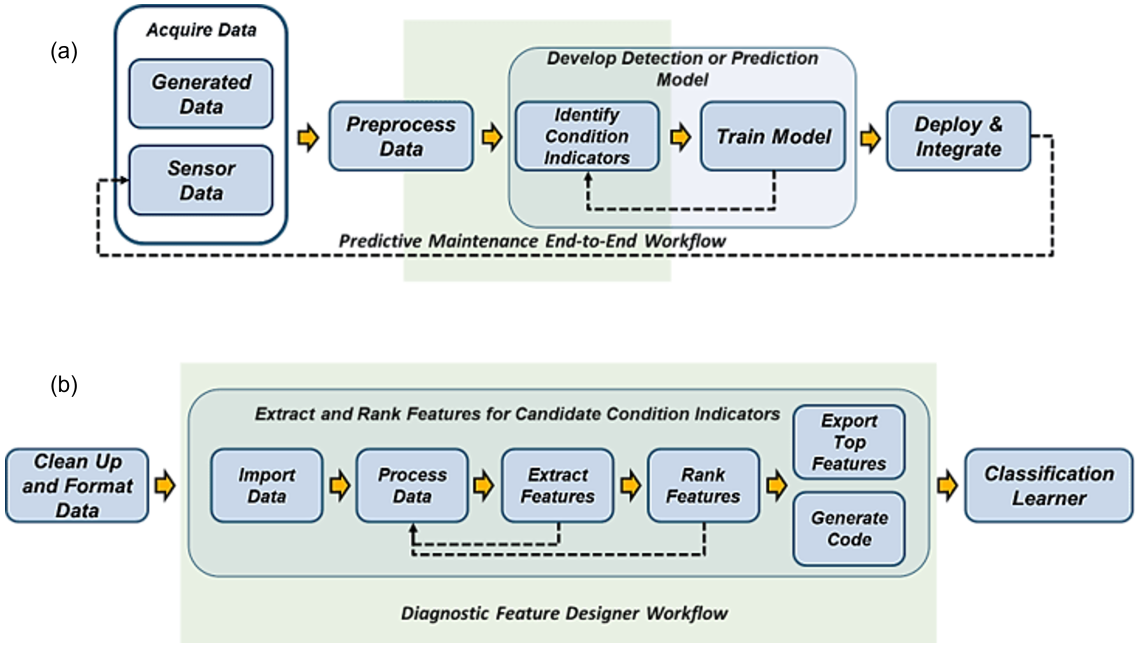


Figure 3.1: PdM workflow with diagnostic feature designer workflow (noa, 2023b)

The PdM workflow is similar for developing either a fault detection model or a RUL prediction model. The workflow starts with acquiring condition monitoring data, which can be from normal system operation, the system operating in a faulty condition and run to failure system data (noa, 2023a). Due to regular preventive maintenance action to keep the system availability high for operation, faulty and especially system failure data is rarely available in industry. This limits what is achievable for PdM, however, by using generated faulty and failure condition data, it becomes possible to overcome these limitations. An in depth knowledge of the operating system is required and it is a challenge to generate realistic fault and failure data.

Pre-processing the data is an essential step for transforming the data to be suitable for effective condition indicator feature extraction. The techniques used for pre-processing can be system specific depending on system application, as well as employing general techniques such as outlier and missing data removal (noa, 2023a).

The extraction of condition indicators is a key step in the workflow, which will determine how the trained model will perform, and this step can be a iterative step to develop the most optimal fault detection or failure prediction model. Condition indicators are essentially features extracted from the system data where it changes in a distinct manner as the system degrades, and can clearly distinguish between normal and faulty operation (noa, 2023a). The mean value which changes over time is a simple feature, while the frequency of peak magnitude in a signal spectrum is a more complex feature (noa, 2023a).

The PdM toolbox in MATLAB has an useful application called 'Diagnostic Feature Designer' which can be used for extracting and ranking features for condition indicators. The workflow for DFD includes importing the condition data to the application, processing the data to a suitable format called file ensemble datastore and extracting time-domain and frequency-domain features. Then the features can be ranked with histograms, using various metrics depending on what category of model is going to be trained. The ranking methods are grouped into supervised, unsupervised and prognostic ranking, to evaluate features best suitable for model development. For unsupervised ranking, Laplacian scores is a useful technique for feature ranking. The method is formed on the real world observation in classification problems that data points from the same variables are usually near other data points, and the ranking of features is assessed by a so called 'locality preserving power' (He et al., 2005). In-depth detail about how Laplacian score functions is found in the research paper by He et al. (He et al., 2005). The top ranking features of interest can be exported for developing the fault detection model.

3.2.1 Principal component analysis (PCA)

PCA is a technique used to reduce the dimensionality of a large multivariate data set, while still containing the most important information (Jaadi, 2023). This method simplifies the machine learning process where a smaller data set can be used, instead of the original large data set. Performing PCA keeps the size of the original data set in terms of rows and columns, however the highest possible variance is explained in the first column, called the first principal component. The second column contains the second principal component which explain the next possible highest variance. Before applying PCA, the first step is to standardize the original data set, since the large and small values should equally contribute to the PCA (Jaadi, 2023). Larger values will dominate over the smaller values, giving biased results when performing PCA and this is not ideal. By standardizing, each value from each variable is subtracted by the mean and divided by the standard deviation (Jaadi, 2023). Then the covariance matrix is computed to identify the correlations between the variables in the data set, allowing for knowing how one variable influences another variable when it is increased or decreased (Jaadi, 2023). Then the eigenvectors and eigenvalues are computed to determine all the principal components from the covariance matrix (Jaadi, 2023). Figure 3.2 shows a scatter plot of a data set, where the first principal com-

ponent is the diagonal line connecting with the purple lines, which maximises the variance. The second principal component is perpendicular in direction and accounts for next highest variance (Jaadi, 2023). Performing PCA in MATLAB returns the principal component coefficients in a matrix, where each column has coefficients for each principal components in descending order (noa, [n. d.]). The scores in a matrix and the variances in a array, which are the eigenvalues from the covariance matrix, are returned for each principal components (noa, [n. d.]).

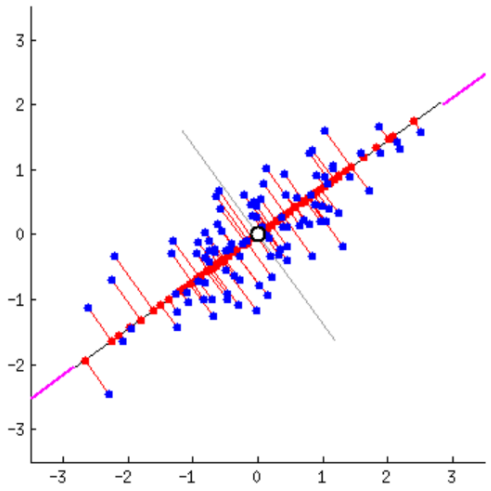


Figure 3.2: First and second principal components (Jaadi, 2023)

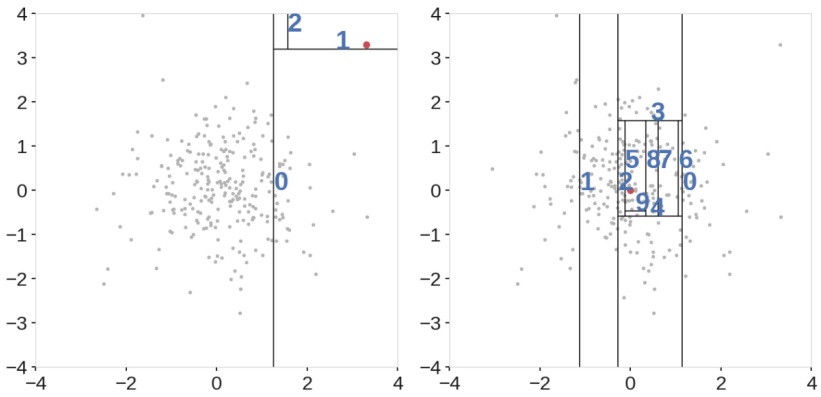
3.2.2 Fault detection

Four unsupervised anomaly detection models for unlabeled multivariate data is presented. These outlier detection methods are available in the Statistics and Machine learning toolbox in MATLAB.

Isolation forest

Liu et al. (Liu et al., 2008) first proposed the Isolation forest algorithm in 2008 for anomaly detection. It is a so called tree-based machine learning algorithm which effectively isolates outliers from a multivariate data set (Alam, 2020b). The algorithm is based on the theory of decision trees and random forests, which splits the data set into two using a random threshold value and continues to do so until all the individual data points are isolated (Alam, 2020b). The data samples are spilt, of which the deepest into the tree are classified to be less likely as anomalies, since more splits are required to finally isolate them (Akshara_416, 2021). Likewise, the data samples which required the least splits to isolate are considered as anomalies and would be the shorter branches of a tree (Akshara_416, 2021). When the isolation forest model has been trained, an anomaly score is given based on the depth of the tree which is required to reach the data point

(Akshara_416, 2021). Figure 3.3 shows two scatter plots where in plot 3.3 (a) shows the anomaly point easily classified in the tree with the branch 1 and 2, and in plot 3.3 (b) shows a normal point which took branch 8 and 9 to classify. This shows the outliers in the data can be isolated quickly.



(a) Anomaly point (b) Nominal point

Figure 3.3: Isolation forest model showing a anomaly and normal point (Akshara_416, 2021)

Here a contamination factor is needed to specify the percentage of outliers in the data and is used for calculating the model score threshold. This is used for detecting the outliers in new data, after the model is trained. The threshold value is set based on the contamination factor used and specifies the percentage of data to be classified as faulty.

Local Outlier Factor (LOF)

The LOF is a unsupervised machine learning algorithm which utilizes the density of the data points in the distribution to classify outliers (Alam, 2020a). The algorithm uses the density of a data points and compares to its neighbouring data points, and since outliers are generally from low density area in the distribution, the ratio is greater for outliers (Alam, 2020a). Hence, the LOF of the outliers are most likely to be high, than when compared with normal points, where they are generally from high density areas, having lower LOF. This technique functions best when the data set is highly spread out, meaning the density is not concentrated in one area (Shukla, 2022). A contamination factor value can be set manually to give the LOF model a proportion of the data set which is considered as outliers (Shukla, 2022). Figure 3.4 shows a scatter plot of LOF with the outlier scores given as a red circle. The data points with a long distance to the other data points, which are the large red circles are considered as most likely to be a outlier.

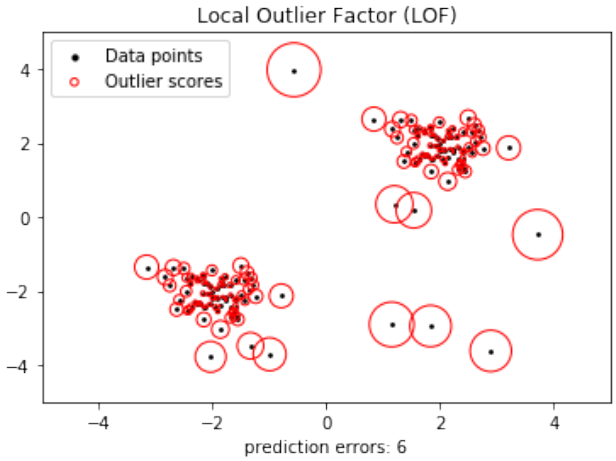


Figure 3.4: Local outlier factor scatter plot with outlier scores (akshayvarma72, 2020)

One-class support vector machine (SVM)

SVM is an algorithm used for classification, and utilizes hyper-planes in a multi-dimensional space to effectively separate a class of observations to another class of observations (Alam, 2020c). One class SVM works by using an hyper-sphere to classify normal and outlier data points. This method can be used to detect outliers when only normal machine condition data is available, where the model then tries to group new data, and if it does not belong to a established group, it is classified as an anomaly (Alam, 2020c). Figure 3.5 shows a plot of red circles, in this case the one class SVM constructs an optimal best fit hyper-sphere from the training data, and data points located outside the hyper-sphere are classified as outliers (Kilaru, 2022).

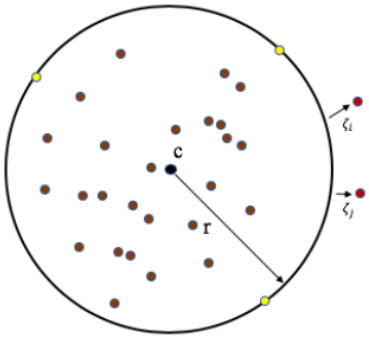


Figure 3.5: One-class support vector machine (SVM) showing best fit hyper-sphere (Kilaru, 2022)

Mahalanobis Distance

The Mahalanobis distance method is used to detect outliers using a multivariate distance metric in a multivariate data set. The multivariate distance metric overcomes the major limitation of using the commonly used Euclidean distance between two data points, which is a straight line

distance (Prabhakaran, 2019). The euclidean distance functions best when the dimensions of the data are equally weighted, as well as being independent of each other (Prabhakaran, 2019). The euclidean distance however gives different values of variables when the unit is changed, even though the physical distance between the data variables remains the same. This can be solved with standardizing the variables, however when the data variables are correlated with each other which is usually the case, incorrect classifications can be made (Prabhakaran, 2019). Figure 3.6 shows two scatter plots where in the left plot, the data variable X and Y are uncorrelated variables and the right plot has correlated variables. In the uncorrelated case, the euclidean distance to the centroid can be used to classify if the points are a member of the distribution (Prabhakaran, 2019). In the correlated case, the euclidean distance is the same from the centroid, however it is only point 1 which is a member of the distribution, while point 2 is clearly outside the main distribution. Using Euclidean distance for correlated variables give incorrect classifications, hence it is not suitable. The multivariate distance metric however measures the distance from a data point to the distribution instead and works by using the covariance matrix of the data set (Prabhakaran, 2019). The mean value is subtracted of each value in the variables and divided with the covariance value. This overcomes the limitation with Euclidean distance and solves the issue with scale and the possible correlation between variables (Prabhakaran, 2019).

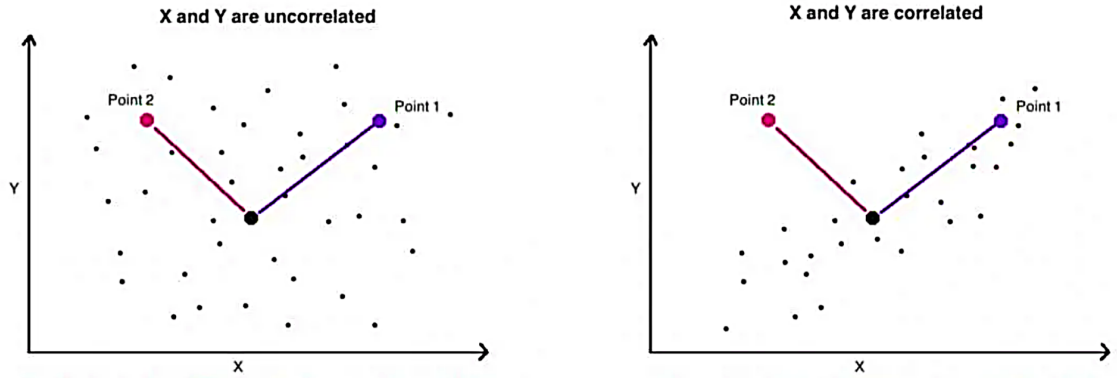


Figure 3.6: Euclidean distance for classifying data points with correlated and uncorrelated data variables (Prabhakaran, 2019)

3.3 Reliability, maintenance objectives and strategies for WTs

The success of wind farms, especially in offshore locations, is highly dependent on the reliability of the WTs. It is critical to the design, O&M, performance assessment and improvement of the WTs (Dao et al., 2019). The definition of reliability given by (Rausand et al., 2020) as

☞ **Reliability**: The ability of an item to perform as required in a stated operating context and

for a stated period of time.

WTs with a low level of reliability may result in high production unavailability due to several component breakdowns requiring extensive maintenance, which gives high O&M costs. On the contrary, WTs with a high level of reliability may have low production unavailability due to reduced breakdown frequency, however it can be prohibitively costly to accomplish (Dao et al., 2019). Pinar Pérez et al. (Pinar Pérez et al., 2013) state that new larger WTs fail more frequently compared to smaller WTs and hence the reliability needs to be greatly improved. The overall offshore WT farm performance and electricity generation is greatly affected by the reliability of the WTs. Figure 3.7 shows a hierarchy of factors affecting the cost of energy produced and the factors highlighted in gray are of interest. Among this, reliability directly affects availability and the energy produced. In addition, the O&M expenditure from component breakdowns and maintenance activity costs contribute to a significant part of the WT lifetime costs, affecting also the cost of energy produced (Dao et al., 2019).

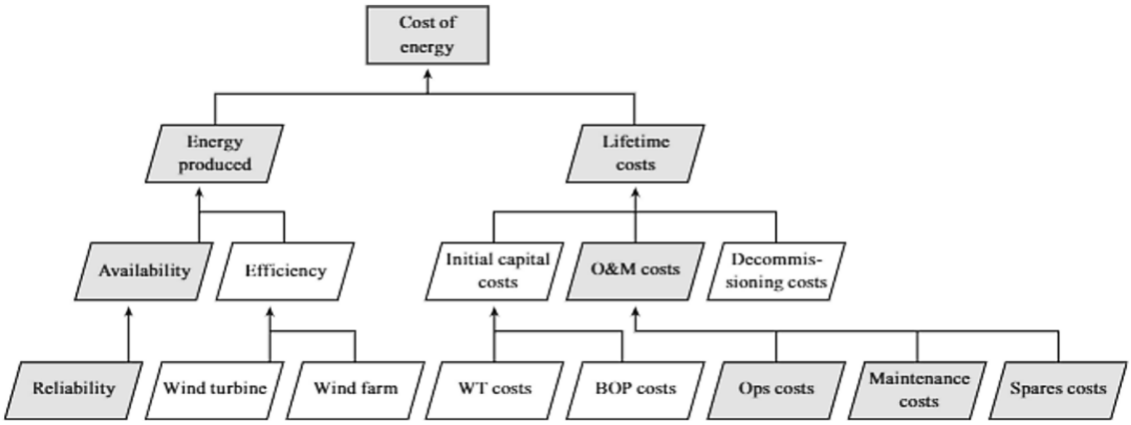


Figure 3.7: Structure of cost of energy of wind turbines (Tavner, 2021)

The definition of maintenance and maintainability given by (Rausand et al., 2020) are

☞ **Maintenance**: the combination of all technical and management tasks intended to retain an item in, or restore it to, a state in which it can perform as required.

☞ **Maintainability**: The ability of an item, under stated conditions of use, to be retained in, or restored to, a state in which it can perform as required, when maintenance is performed under stated conditions and using prescribed procedures and resources.

Figure 3.8 shows the commonly implemented maintenance strategies in the industry such as corrective maintenance (CM), preventive maintenance (PM) which includes time-based main-

tenance (TBM) and condition-based maintenance (CBM). The conventional maintenance strategies for onshore WTs include CM and PM, this approach however for offshore WTs is not sufficient and needs to be improved to ensure lower OPEX (Artigao et al., 2018). Fox et al. (Fox et al., 2022) state that CM and PM dominate the presently used maintenance strategies for offshore WTs.

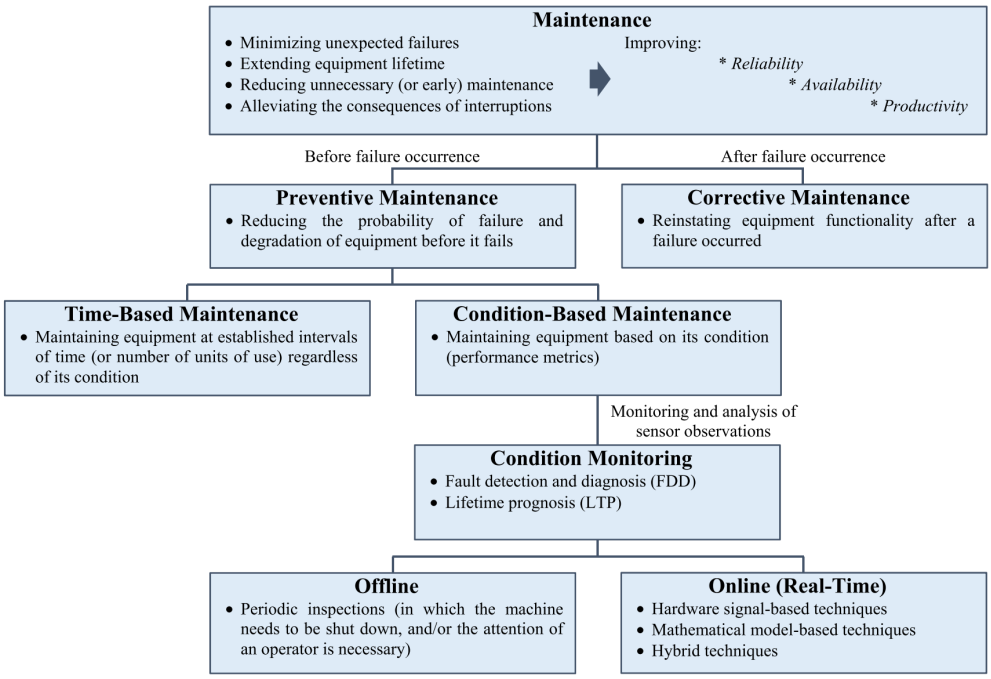


Figure 3.8: Maintenance objectives and strategies (Badihi et al., 2022)

CM is reactive based, and maintenance is performed immediately after a component failure occurs or scheduled at the next suitable time. PM includes TBM and CBM strategies. TBM is a planned maintenance activity for a component or system at pre-selected intervals of time without considering the condition of the equipment. Whereas, CBM takes into consideration the actual condition of the equipment from the use of condition monitoring technologies and when the measured variables reach a pre-selected threshold, the necessary PM activity is scheduled. CBM finds the optimum level between CM and PM, the unnecessary maintenance activities are decreased and is regarded to be essential to reaching high system availability (Artigao et al., 2018).

PdM builds upon CBM with the use of relevant theory and prognosis methods to estimate the remaining useful time of the equipment. This is the time until failure and the necessary PM activity is scheduled before the failure is expected to occur. Hence, for offshore WTs, PdM is critical for achieving high production availability for critical WT components with early fault detection and diagnosis.

3.3.1 Equinor's maintenance strategy for offshore wind farms

Equinor currently operates and as part of a joint venture partnership with external companies, jointly owns a range of offshore wind farms such as the Dogger bank (UK), Sheringham Shoal (UK), Dudgeon (UK), as well as owning its own offshore wind farms such as Hywind Scotland (UK) and Hywind Tampen (Norway) (wind, 2023). In addition, there are ongoing collaboration with more companies worldwide to construct and operate more offshore wind farms.

Figure B.1 provided in the appendix show an overview of Equinor's failure management strategies with a more in-depth breakdown and processes involved than when compared with figure 3.8. At Equinor, a combination of traditional CM, TBM, CBM and PdM are utilized for offshore wind farms. The choice of maintenance strategy is highly dependent on the criticality of the WT component failure, the cost of implementing the maintenance strategy and other factors such as spare part inventory is taken into account. CBM is mainly implemented if there is a good business case and it is cost-beneficial for the equipment, where repair cost and the time required to perform the maintenance activity per WT in the offshore wind farm is considered (Svennevig, 2022). This can more optimally utilize the remaining WT component condition and is expected to increase the profitability of the WT operation. Maintenance optimization is performed by implementing CBM and PdM together with the asset's failure management policy, by processing data from online condition monitoring and using predictive analytic's (PA) (Svennevig, 2022). From Equinor's guidelines for PdM and CBM (Svennevig, 2022), "Predictive Analytics includes a selection of mathematical and statistical techniques within the framework of probability theory. Machine learning apply these techniques to predict the expected future condition of an equipment within a certain probability."

One of Equinor's key priorities from the Maintenance strategy for NES (REN) Operated Assets (Callister, 2021) is "The maintenance process shall have a reliability and availability focused approach, where condition monitoring and proactive maintenance routines are the primary maintenance drivers." This gives the reader an idea of the emphasised approach for condition monitoring. Data driven O&M is a key focus area of Equinor's offshore wind technology strategy (Callister, 2021). The road map for offshore wind technology strategy for 2023 shows that their NES Integrated Operation Centre (IOC) Anomaly detection PowerBI dashboard is up and running, predictive algorithms are running and a shift from TBM to CBM is implemented. The purpose of the REN IOC dashboard is to support the O&M teams with anomaly detection, CBM forecast with trends, predict component failures and optimising the maintenance performance.

3.4 Industrial rolling bearing

Rolling bearings are commonly used in industry for many applications and the following description provides a basic understanding of their function.

> Rolling bearings support and guide, with minimal friction, rotating or oscillating machine elements – such as shafts, axles or wheels – and transfer loads between machine components. Rolling bearings provide high precision and low friction and therefore enable high rotational speeds while reducing noise, heat, energy consumption and wear.
>
> (SKF_Bearing_Basics, 2023)

Figure 3.9 shows the four main components required for the bearing to function which are an outer ring, inner ring, rolling elements and a cage. A seal is used to hold and protect the bearing lubricant from harmful external contaminants. The rolling elements contact surface with the bearing raceway differentiate if the bearing is a ball bearing or a roller bearing. Ball bearing as in figure 3.9 have a point contact surface and a roller bearing have a line contact surface, making it capable of accommodating heavier loads (SKF_Bearing_Basics, 2023).

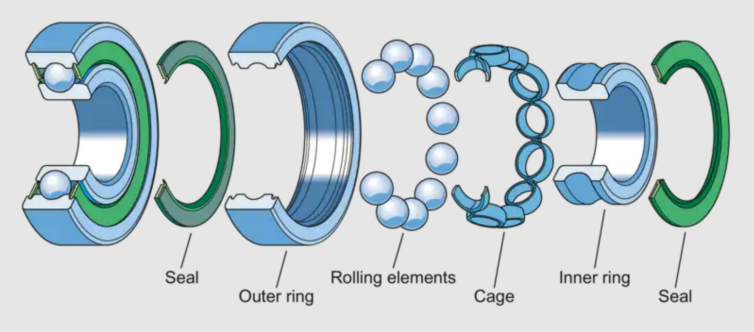


Figure 3.9: Ball bearing components (SKF_Components_and_materials, 2023)

3.4.1 Vibration analysis for fault diagnostics

During operation of industrial machines, all mechanical components vibrate as they interact with both internal and external forces (SKF_Vibration_Diagnostic_Guide, 2011). Vibration is the mechanical oscillation about an equilibrium point (ISO_2041:2018, 2018). Vibration can be used as an indicator of a mechanical components condition, due to the distinct characteristic of excessive vibration which is commonly observed with faults in rotating machinery (SKF_Vibration_Diagnostic_Guide, 2011). In addition, different faults in machines produce usually a distinct vibration pattern, and this can be used for identifying the root cause of the machine problem for performing the necessary maintenance repair (SKF_Vibration_Diagnostic_Guide, 2011). Hence, vibration can provide an insight into the early development of fault conditions and the vibration pattern changes along with the machine condition (Mobius_Institute, 2016).

A vibration signal is analysed by splitting it into two components which are it's amplitude and frequency (SKF_Vibration_Diagnostic_Guide, 2011). These two signal characteristics of a

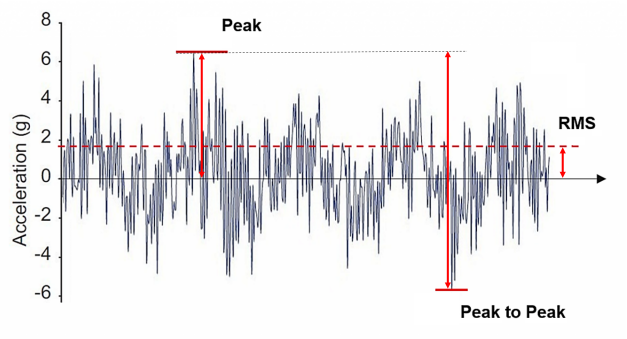


Figure 3.10: Vibration signal scale factors (Adapted from (Litchwark, 2023))

signal are well known to engineers. The formal definitions of amplitude and frequency are:

☞ **Frequency**: is the reciprocal of the period which is the smallest interval of time for which a periodic function repeats itself. It is measured in the unit Hertz (Hz), which corresponds to one cycle per second (ISO_2041:2018, 2018).

☞ **Amplitude**: is the magnitude, size or value of a quantity (ISO_2041:2018, 2018). For vibration, it depends on the severity of the mechanical fault.

To compare and analyse vibration signals overall values, the industry uses scale factors. The three factors commonly utilized are Peak, Peak to Peak and Root Mean Square (RMS). Figure 3.10 shows these methods on a vibration signal measured with an accelerometer vibration sensor. The y-axis unit is given in acceleration (g) and time is used for the x-axis. From the neutral position, or zero reference, the peak amplitude value is the highest point in the top signal waveform, whereas peak to peak amplitude value is the distance from the highest top waveform to the lowest waveform. The RMS value is the square root of the sum of all the squared instantaneous values of the vibration signal, and for a pure sine waveform, the RMS value is simply 0,707 times the peak value (SKF_Vibration_Diagnostic_Guide, 2011). The RMS value of the vibration signal in Figure 3.10 is indicated with a red horizontal dotted line.

The position of the accelerometers on an machine is important to capture the vibrational motion in different directions. Figure 3.11 shows the industry standard for the position of accelerometers on an industrial motor on both the drive and non drive end. Accelerometers are mounted in the axial, vertical and horizontal direction. Due to the assembly of the motor to the ground, the horizontal measurements usually show more radial vibration due to the machine flexibility in the horizontal plane (SKF_Vibration_Diagnostic_Guide, 2011). Whereas, due to the mounting and force of gravity, there is less movement in the vertical direction and hence lower vibration (SKF_Vibration_Diagnostic_Guide, 2011). The function of the motor makes it that the greatest forces are experienced in perpendicular direction to the motor shaft and hence, the

movement in the axial direction is expected to be low under normal operation (SKF_Vibration_Diagnostic_Guide, 2011). Unless if there are other machinery problems present such as misalignment and bent shaft, which lead to greater axial vibration.

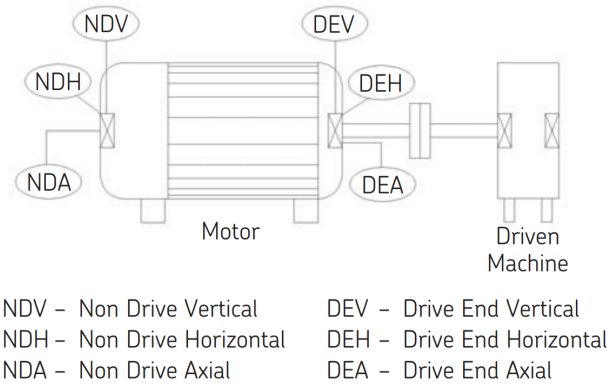


Figure 3.11: Measurement sensor position on industrial motor (SKF_Vibration_Diagnostic_Guide, 2011)

There exists various sensor technologies such as proximity probes, velocimeters and accelerometers which can measure vibration respectively in displacement, velocity and acceleration. Their effectiveness of measuring vibration is different at various frequency ranges. "Displacement is most sensitive to lower frequencies, acceleration is most sensitive to high frequencies and velocity is sensitive to most frequencies but not so great at very high or very low frequencies." (Mobius_Institute, 2016). Velocity vibration sensors lose measurement effectiveness below 10 Hz and above 2 kHz (SKF_Vibration_Diagnostic_Guide, 2011). To overcome this, accelorometers are used, where it's strength is the ability to operate in a wide range in frequency from nearly 0 Hz to above 400 kHz (SKF_Vibration_Diagnostic_Guide, 2011). Specifically a ceramic piezoelectric sensor is used to measure the vibration acceleration in industrial machines (Measuring_Vibration_with_Accelerometers, 2023).

Time Waveform Analysis of the vibration signal can be a starting point to get an idea of the bearing condition. It is a graph of the amplitude of the vibration signal plotted with time. Commonly a short time sample of the raw vibration is available (SKF_Vibration_Diagnostic_Guide, 2011). The time waveform provides a view of the complete vibration picture of the bearing and it is rather difficult to diagnosis any faults. More advanced techniques are required.

Fast Fourier Transform (FFT) spectrum analysis is an important vibration analysis tool, providing the capability of diagnosing various machinery problems (SKF_Vibration_Diagnostic_Guide, 2011). The principal idea of FFT is to break the raw vibration signal down into specific amplitudes at various component frequencies as figure 3.12 is illustrating. FFT transforms the vibration signal from the time domain to the frequency domain, where the x-axis becomes frequency instead of time. FFT allows for the isolating the root cause of the problem such as

for rolling bearings, which have physically corresponding defect fault frequencies. Due to the fact that certain machinery problems occur at specific frequencies, it becomes possible to analyse the FFT spectrum by observing the change in amplitude at the frequency ranges of interest (SKF_Vibration_Diagnostic_Guide, 2011).

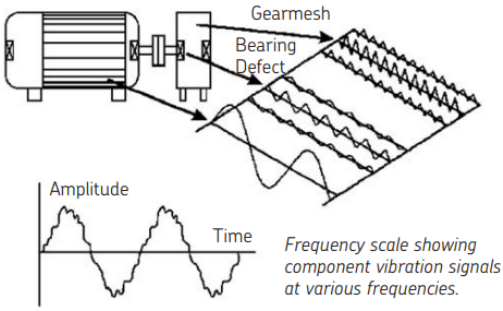


Figure 3.12: FFT spectrum analysis (SKF_Vibration_Diagnostic_Guide, 2011)

An FFT spectrum becomes useful with isolating and diagnosing machinery faults which have developed, however, more advanced techniques are required for the early detection of rolling bearing faults. This would allow for the timely intervention to resolve the fault by taking preventive maintenance action. A common starting point of bearing damage is through improper lubrication, corrosion & contamination to the lubricant and damaging of the bearing surfaces, and excessive load causes surface fatigue (Katz, 2021). "Initial bearing fatigue results in shear stresses cyclically appearing immediately below the load carrying surface. After a time, these stresses cause cracks that gradually extend up to the surface. As rolling elements pass over these cracks, fragments break away. (SKF_Vibration_Diagnostic_Guide, 2011)" This phenomenon is referred to as spalling and figure 3.13 illustrates this.

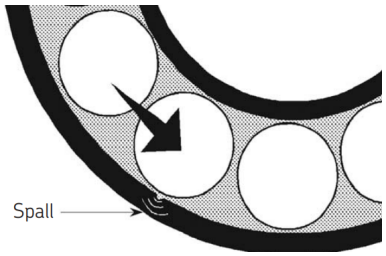


Figure 3.13: Spalling fault in bearing outer ring raceway (SKF_Vibration_Diagnostic_Guide, 2011)

Envelope Detection is an effective technique at detecting the early development of the failure mode spalling. A characteristic behaviour of spalling, is that a low amplitude, high frequency impulse signal is generated each time the rolling element of the bearing rolls over the defected area (Vermeiren and Bark, 2011). This repetitive impulse signal becomes difficult to observe

when it is combined with the normal machine signal noise and the overall structural vibration, resulting in a complex vibration signal (Vermeiren and Bark, 2011). Envelope detection, referred also as 'Acceleration enveloping' relies on a series of process and band-pass filters to remove the low frequency rotational vibration signals and enhance the repetitive impact type signals (SKF_Vibration_Diagnostic_Guide, 2011). Enveloping detection, combined with FFT allows for fault detection, isolation and assessing the severity through fault estimation at the early stages of bearing failure.

3.4.2 Bearing defect frequencies

The geometry of the rolling bearing allows for the calculation of four fault frequencies which correspond physically to the movement the rolling elements make with a shaft rotation. This can be used to find the root cause of the bearing fault through frequency domain vibration analysis. The peak amplitude at these fault frequencies enables to diagnosis the bearing fault. The following variables of the rolling bearing is required to calculate the fault frequencies. Figure 3.14 shows these variables, which are the pitch diameter (P_D), ball diameter (B_D), number of balls (N_B) and the contact angle (β) between the outer and inner race. The rotations per minute (RPM) of the shaft is also necessary.

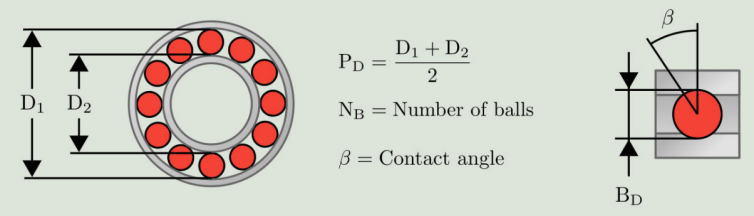


Figure 3.14: Rolling element bearing variables for fault frequency calculation (Fernandez, 2017)

The four distinct fault frequencies together with their formulas from (Fernandez, 2017) are:

The Ball Pass Frequency - Outer race (BPFO) which is the rate of a rolling element passing a point on the outer race of the bearing and in the case of a damage on the outer race, a periodic pulse is created (Mobius_Institute, 2016). The following formula 3.4 can be used

$$BPFO = RPM\frac{N_B}{2}\left(1-\frac{B_D}{P_D}\cos(\beta)\right) \tag{3.4}$$

The Ball Pass Frequency - Inner race (BPFI) which is the rate of a rolling element passing a point on the inner race of the bearing and in the case of a damage on the inner race, a periodic pulse is created (Mobius_Institute, 2016). The following formula 3.5 can be used

$$BPFI = RPM\frac{N_B}{2}\left(1 - \frac{B_D}{P_D}\cos(\beta)\right) \tag{3.5}$$

The Ball Spin Frequency (BSF) is the rate of which a point in the rolling element passes the inner and outer race of the bearing and in the case of a damage on rolling element, a periodic pulse is created each time it strikes the raceways (Mobius_Institute, 2016). The following formula 3.6 can be used

$$BSF = RPM\frac{P_D}{B_D}\left[\left(1 - \frac{B_D}{P_D}\cos(\beta)\right)^2\right] \tag{3.6}$$

The Fundamental Train Frequency (FTF) most commonly referred to as the cage frequency and is the rate of which the bearing cage rotates with each shaft rotation (Fernandez, 2017). The following formula 3.7 can be used

$$FTF = RPM\frac{1}{2}\left(1 - \frac{B_D}{P_D}\cos(\beta)\right) \tag{3.7}$$

When details about the rolling bearing is known, the precise fault frequencies can be calculated for effective fault diagnosis. There exists also simplified formulas which can be used in the case the bearing type details are not known, in this case the number of balls and the RPM are required.

3.4.3 Fault and failure stage development

Bearings which are correctly specified, transported, stored, installed, lubricated and operated, is expected to operate reliably to its design lifetime (Mobius_Institute, 2016). However, the improper handling, poor design and service of a bearing results in it to only operate to around 10% of the designed life time (Mobius_Institute, 2016). Bearing failures can be spilt into four categories being lubrication (36%), fatigue (34%), handling & installation (16%) and contamination (14%) (Mobius_Institute, 2016). The percentages of the bearing failures indicate that lubrication and fatigue of the bearings are accountable for the majority of the failures. Increased load on the bearing which is beyond the design load will reduce the fatigue life and decrease the operational life (Mobius_Institute, 2016). Lubrication of the rolling elements inside the bearing is vital for a smooth operation, however, the type, the quantity of the lubrication being too much or too little can greatly impact the operational life of the bearings (Mobius_Institute, 2016). Contamination in the bearings, by the presence of foreign particles can impact the rolling elements in the bearings, leading to damage for example in the inner and outer race.

Typically the wear of a bearing is spilt into four stages which can be progressed through in different speeds, depending on the use case and environment of the bearings. Figure 3.15 shows these with important bullet points of the stages. Due to the change in vibration between the

stages, different indicators can be used. The vibration can be spilt into four components which are friction between surfaces, stress waves from metal-to-metal contact, periodic vibration from many bearing fault types and resonance from surface impacts (Mobius_Institute, 2016).

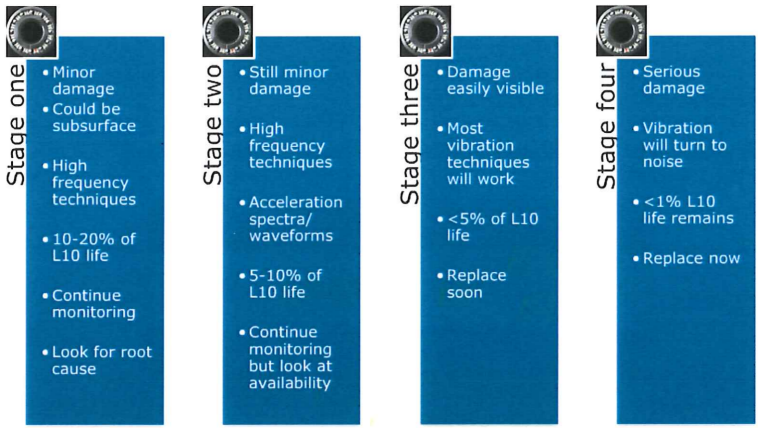


Figure 3.15: Bearing wear stages (Mobius_Institute, 2016)

Stage one bearing faults are only sub-surface damage and are caused by friction and minor impacts of metal to metal contact within the bearing. High frequency vibration greater than 20 kHz with low amplitude is generated from friction and stress waves from very short duration impacts (Mobius_Institute, 2016). Here, high frequency detection with early upward trend and envelope (demodulation) analysis methods with an increase in overall floor noise are most effective at showing the earliest signs of faults. The bearing damage is difficult to see and a possible maintenance action would be to lubricate correctly, while looking for the root cause.

Stage two bearing faults are also sub-surface damage and are caused by friction and minor bearing component impacts. The high frequency vibration continues to increase in amplitude and can cause resonances of the machine at its natural frequency (Mobius_Institute, 2016). Additionally, lower frequency vibration greater than 2 kHz is generated, as the fault develops. Similarly to stage one detection techniques, high frequency detection will show upward trend and envelope analysis will show increase in overall noise. These methods are most effective at showing the signs of faults. Now the acceleration spectrum will indicate fault development and time waveform analysis becomes useful. The bearing damage is still difficult to see and a possible maintenance action would be to check and lubricate correctly, while investigating the root cause.

Stage three bearing faults show signs of more significant damage and the high frequency vibration amplitude increases more in amplitude. The increased wear generates more higher amplitude vibration with lower frequency (Mobius_Institute, 2016). The envelope spectrum will be most effective at fault visibility, and acceleration spectrum may shows signs of increasing peaks at the bearing fault frequencies. The velocity spectrum will shows signs of wear in the

bearing (Mobius_Institute, 2016).

Stage four bearings faults will show significant damage and it is usually that damage in one bearing component will result in the damage of other bearing components and so on, due to the nature of bearing component contacts during operation (Mobius_Institute, 2016). The high frequency vibration will be less and may show a downward trend, due to the fact that constant metal-to-metal impacts have a smoothing effect and the sharp impacts will reduce. This method becomes less effective, whereas, the envelope, acceleration and velocity spectrum's will show the fault condition. The overall floor noise will be high in the envelope spectrum, since, there are many points of damage contacts on the bearing, generating low frequency vibration with high amplitude (Mobius_Institute, 2016). The bearing damage will be easily visible and the recommended maintenance action would be to replace the bearing.

It is important to note that it is usually other machinery problems which results in an bearing defect and the source of the problem is not commonly the bearing defect itself (SKF_Vibration_Diagnostic_Guide, 2011). Therefore, the presence of a developing bearing defect indicates that other machinery problems may be present and should be investigated in parallel, so that both issues can be resolved. Detecting the bearing faults does not automatically improve the reliability and extend the bearing life, the necessary preventive maintenance repair should be taken and the consequences can be reduced (Tranter, 2016).

Chapter 4

Wind turbine SCADA & CMS data

This chapter presents the data acquisition process for offshore and onshore WT SCADA and CMS data. Details about the wind farm, WT component of the data acquired and the applications used to store and analysis the data are presented.

4.1 Data acquisition

Table B.1 shows a timeline summary of the data acquisition process and is provided in the appendix Chapter 4

4.1.1 Equinor - offshore wind farms

This master's thesis is carried out in collaboration with the energy company Equinor, who operates and owns many offshore wind farms. Hence, in relation to the problem formulated in this thesis, it was a natural choice to obtain SCADA and other CMS data of their offshore wind turbines. However, the approval process for getting access to the live operational data and historical SCADA data of Equinor operated offshore wind farms such as Sheringham Shoal (SHS), Hywind Scotland (HYS) and Dudgeon (DOW) was an unusually lengthy process. This was mainly due to strict data sharing procedures and policies within the Renewable's department at Equinor. The offshore wind farm SCADA data is considered to be highly confidential information which would need approval from all the partners of each joint venture company of the wind farms, since they technically own the SCADA data as well. The data approval process plus the mandatory risk assessment was estimated to take months. Due to this and the limited time of the master's thesis, the scope of using Equinor's offshore wind farms SCADA data was not viable. A suggestion was given by the Equinor REN renewable's asset management, of providing access to anonymised and tweaked SCADA data, such that it would not be possible to recognise which wind farm the data was originating from. This as well would have taken many weeks to get

access to and hence it was decided in the end to use the available data from NTNU and Aneo.

Meaningful work on understanding the maintenance strategies implemented by REN Equinor such as anomaly detection, condition-based maintenance and predictive analytic's for offshore wind turbines was carried out in the meantime, while waiting for access to the SCADA data. Preparation to analyse the data and develop the fault diagnosis models and how they can be compared was carried out.

4.1.2 Aneo - onshore wind farms

The renewable energy company Aneo was established in 2022 by TrønderEnergi and HitecVision (noa, 2023c). Aneo is demerged from TrønderEnergi and they now own and operate all the wind power plants from TrønderEnergi's portfolio (noa, 2023c). As part of the RAMS lab at NTNU, several vibration sensors are amounted on the generator and rotor axle bearings of 2 onshore wind turbines at the Bessakerfjellet wind farm in Åfjord municipality. This wind farm is owned and operated by Aneo. These accelerometer sensors were installed by students from the Department of Mechanical and Industrial Engineering (MTP) in April 2022 and started collecting vibration data since then. This data is stored to NTNU's SKF @ptitude observer. With agreement from my co-supervisor Thor Inge Bernhardsen, focus was shifted during the middle of April to obtaining and analysing the generator, rotor bearing vibration data from NTNU's SKF @ptitude observer database. Approval from Aneo was obtained before using the vibration data. SCADA data from Aneo of the onshore wind turbines were obtained at the end of April 2023. This Bessakerfjellet SCADA data was originally shared for a research project in FWE Northwind at NTNU and the same access was provided. The data included the WTs active power, nacelle direction, turbulence, wind direction and wind speed. SCADA data from August 2017 to the end of April 2022 was shared. Due to the already ongoing collaboration with Equinor, Aneo did not have enough resources to assist with providing additional guidance and sharing of maintenance logs of the wind turbines. This increased the fault detection model development difficulty with limited knowledge of when maintenance activities were performed and assumptions for gaps found in the data had to be made. Separating normal and possible faulty vibration became an additional challenge.

4.2 Bessakerfjellet

4.2.1 Description of wind farm

The wind farm officially opened in 2008 and has a total of 25 WTs installed of type Enercon E-70 (NVE, [n. d.]). The average yearly wind speed is 8.5 m/s at the Bessakerfjellet (NVE, [n. d.]). The WTs has a cut-in wind speed of 2.5 m/s, rated wind speed of 15.0 m/s and a cut-out wind speed

of 34 m/s (Bauer and Matysik, 2023). The hub height is 64 meters and the rotor diameter is 71 meters and has a rated generator power is 2.3 MW (Bauer and Matysik, 2023). The wind farm electricity production is on average 175 GWh annually (NVE, [n. d.]).

4.2.2 Rotor and generator axle bearing

The 3 rotor blades of the WT is mounted on a rotor axle which rotates on a double row tapered rolling bearing. This bearing will be referred from here on as rotor bearing. The generator is located further inside the WT which generates electricity with the rotation of the axle. The axle here is supported and rotates on a single row cylindrical roller bearing. This bearing will be referred here on as generator bearing. The inner ring of both the bearings is the rotating component together with the shaft. Table 4.1 shows the rotor and generator bearing model types.

Table 4.1: Bessakerfjellet WT bearing information

Placement	**Rolling bearing type**	**Bearing arrangement**	**Model**
Rotor Axle	Double row tapered roller bearing	Locating support	SKF BT2-8079/HA1 OHNE DICHTUNG
Generator Axle	Single row cylindrical roller bearing	Non-locating support	SKF BC1-8033/HB1VK443

The double row tapered roller bearing used for the rotor bearing provides locating support which accommodates high radial and axial loads in both direction. The single row cylindrical roller bearing used for the generator bearings provide a non-locating support which can accommodate high radial load and locate the shaft axially in one direction. Specifications of the bearings are listed in table 4.2. Details about the rolling element ball diameter was proven difficult to gather and hence, is left blank.

Table 4.2: Bessakerfjellet WT rotor and generator bearing specifications

Specification	**Rotor Bearing**	**Generator Bearing**
Outside diameter (D)	540 mm	1050 mm
Inside bore diameter (d)	400 mm	850 mm
Pitch diameter (P_D)	470 mm	950 mm
Number of balls (N_B)	41	12
Ball diameter B_D		
Contact angle (β)	11.667 °	0°

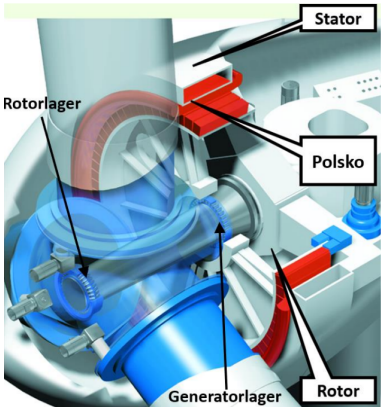


Figure 4.1: Bessakerfjellet wind turbine bearing placement (Meland and Haugan, 2021)

Preventive maintenance action is carried out twice a year, with one in the summer and the other in winter (Karzon et al., 2022). The lubrication pump are filled up in the summer and winter, while grease cartridges are changed only in the summer (Karzon et al., 2022).

4.2.3 Accelerometers

There are in total 3 accelerometers mounted by NTNU on selected WTs. The model type of the accelerometer is SKF CMSS-WIND-100-10 which has a measurement range from 0.5 Hz to 10000 Hz, and is specially designed for the use in WT generators due to its low profile and compact size (Meland and Haugan, 2021). One sensor is mounted on the rotor bearing, measuring radial vibration, which is perpendicular to the shaft. Two sensors are mounted on the generator bearing, one measuring axial and the other measuring radial vibration. A tachometer is used for measuring the number of rotations per minute of the WT axle.

4.3 SKF @ptitude observer

The accelerometer and tachometer sensors are mounted on the rotor and generator bearing of WT 4 and WT 21 in the Bessakerfjellet. These sensors were installed and configured for the first time in April 2022 and hence only started to measure data from this month. The data is collected and stored in 'NTNU - Bessakerfjellet' folder within SKF @ptitude observer database. This application provides the capability of viewing and analysing the measured vibration signals with numerous graphic displays and diagnostic tools. The raw vibration signal can for example be viewed in the time waveform for initial analysis, and more advanced tools such as spectra are available which uses FFT for viewing the signal in the frequency domain.

Figure B.2 provided in the appendix Chapter 4 shows an overview of the application and shows an example the WT 21 rotor axle enveloped radial vibration signal in a trend graph, to-

gether with the rotor speed. Various measuring points are setup per accelerometer sensor, allowing for measuring and capturing the bearing vibration in different frequency ranges. The list below shows the measuring points setup and the information is gathered from the following relevant NTNU project assignment (Karzon et al., 2022).

- Acceleration: 1kHz Dynamic
- Envelope 1: 5 Hz - 100 Hz Dynamic, Envelope
- Envelope 2: 50 Hz - 1 kHz Dynamic, Envelope
- Envelope 3: 500 Hz - 10 kHz Dynamic, Envelope

The first bullet point measures the acceleration of the shaft which can be either in radial or axial direction depending on where the accelerometer is mounted. The name for this measuring point is given in short form as 'ACC (1 kHz)'. The enveloped vibration signal are named in short form as 'ENV 1, 2 or 3' depending on the measurement frequency range. For each sensor whether it is measuring radial or axial vibration, the four measuring points are configured 'ACC (1 kHz)', 'ENV 1', 'ENV 2' and 'ENV 3'. Figure B.2 shows this setup in the hierarchy tree. The letter A or R is used to denote whether it is axial or radial vibration respectively.

For the trend graph in figure B.2 for '3 Rotor aksel R ENV 3', the x-axis is given in Time and for the y-axis is given in 'gE PtP'. The 'g' denotes it is a acceleration signal and the 'E PtP' indicates the signal is transformed with the envelope detection method and shows the peak to peak value of the vibration signal.

4.3.1 Warning and alarm levels

The SKF application allows for setting a warning and alarm level on the trend graph. The colour of warning level is given in orange and the alarm level has colour red. These threshold levels are already preset by previous NTNU bachelor students in MTP department, who have setup the sensors on the Bessakerfjellet WT and connected these to the SKF @ptitude observer. Thresholding is a simple anomaly detection method which can alert an anomaly when the system vibration level exceeds a threshold on a statistical metric (noa, 2023d). The thresholds were set by referring to relevant standards such as ISO 10816 and VDI 3834, and from recommendations from the machine supplier (Meland and Haugan, 2021). ISO 10816 specifies the measurement and evaluation of mechanical vibration of WTs and its components, and splits the WTs into two groups depending on if the WT has a gearbox or not.

The warning level from SKF is referred to as alert in ISO 10816 (ISO_10816-21:2015, 2015) and the purpose is to indicate that a specified vibration limit is reached and remedial measures are necessary. The alert signal is the first zone of an anomaly, requires an heightened awareness and is identified with a yellow indicator (ISO_13372:2012, 2012). Additionally, ISO 10816 points

out that it is generally permissible to allow the WT to operate until the cause of the increase in vibration has been found and any preventive measures have been undertaken (ISO_10816-21:2015, 2015).

The alarm level is similarly referred to as alarm in ISO 10816 and the intent is to protect the WT and its components from function failure which can lead to unsafe operating conditions. It indicates that an selected anomaly has occurred and the necessary corrective actions is required (ISO_13372:2012, 2012). Crossing the threshold value further indicates that operation of the WT will be harmful and damage the WT components (ISO_10816-21:2015, 2015).

4.4 MATLAB

The SKF @ptitude observer lacks functionalities required for multi-dimensional data analysis and data-based approaches using machine learning for advanced fault detection, diagnosis and prognosis model development. Hence, external programming languages such as MATLAB, can utilized to overcome these limitations by importing the necessary condition monitoring data from the SKF @ptitude observer database. MATLAB is developed by MathWorks and has many useful toolboxes such as Statistics and machine learning toolbox and the PdM toolbox. Here it becomes possible to use the vibration data for condition indicator feature extraction, visualization, and ranking the features. Suitable machine learning algorithms can be explored and trained for fault detection models.

Chapter 5

Case study

This chapter presents a extensive case study of generator and rotor bearing raw vibration analysis with statistical data visualization, frequency domain vibration analysis and correlation study with WT rotor speed. The preparation process of the data, and the unsupervised fault detection models trained are presented.

5.1 WT Data

5.1.1 Acquiring data and formatting steps

The enveloped vibration signals and rotor speed data from SKF @ptitude observer is extracted. The sensors started collecting data from April 2022 and data up to the time of extraction April 2023 is collected. All available data for ACC of the shaft, ENV 1, 2 and 3 radial and axial vibration of the rotor and generator bearing for WT 4 and 21 is exported. From previous experience of using the SKF @aptitude observer during the specialization project 2022, it was found that exporting the data in an OpenDocument Spreadsheet (ODS) file gave the correct number format for importing to MATLAB. A few formatting and editing steps are performed in excel to remove unnecessary columns and add column names such as 'Time', 'Vibration', 'Rotor Speed', 'Yellow Warning' and 'Red Warning.' The processed ODS files are imported as tables into the MATLAB workspace.

The SCADA data from Aneo is provided in Excel CSV files from August 2017 to the end of April 2022. It became difficult to make use of this data, since the NTNU mounted sensors on the two WTs at the Bessakerfjellet only started measuring vibration and rotor speed from mid April 2022. Additionally, the measurement sampling rate of the data from Aneo is taken once every minute, while SKF @aptitude observer had an inconsistent measurement sampling rate through the past year, from once every hour to sometimes once every 20 minutes. This is caused by changes in sensor measurement settings. The amount of overlapping SCADA data from Aneo and CMS data

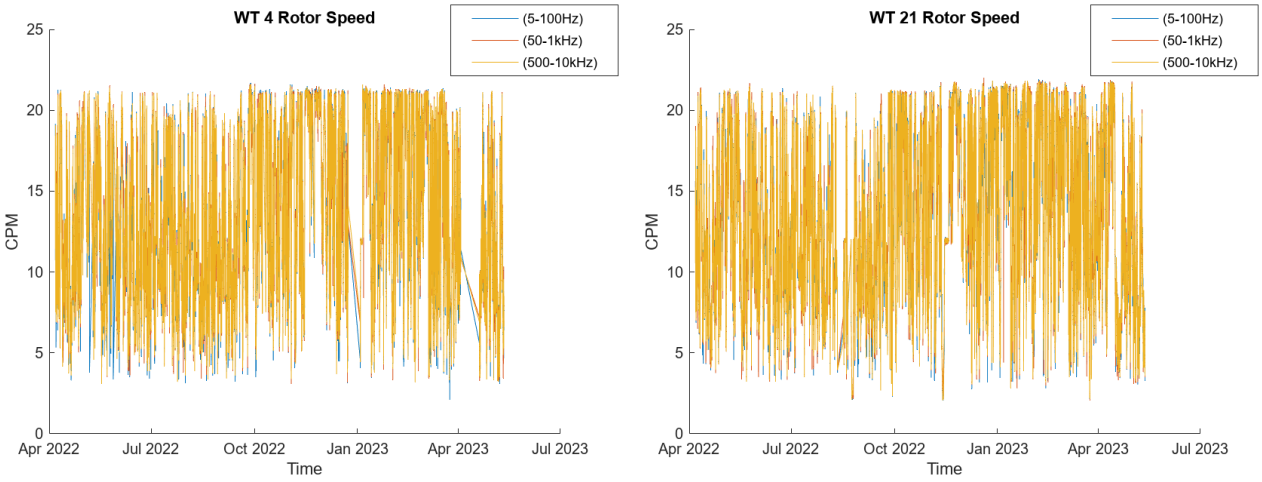
from SKF @ptitude observer of maximum 1.5 weeks in a few cases, resulted in limited value of using the SCADA data for fault detection model development. WT rotor speed data of the shaft is available and stored locally in SKF @ptitude observer with the same sampling frequency as vibration. For this reason, it became more logical to use for example the relationship between vibration and rotor speed for fault detection model development in MATLAB.

5.1.2 Data visualization

The first step is to visualize the rotor speed data, generator and rotor bearing vibration data to determine the condition of the bearings. This helps in making a judgement of whether the vibration levels show normal or degrading conditions in relation to the preset warning and alarm thresholds. Additionally, the wear of the bearings observed through the vibration amplitudes in different frequency envelopes can be used to classify the bearing wear stage.

Rotor Speed

The rotor speed of WT 4 and 21 are extracted individually with ENV 1, 2 and 3 axial and radial vibrations. The time stamps of the vibration and rotor speed data were slightly different than each other for the different envelopes, of a maximum half a minute. The rotor speed of WT 4 and 21 are plotted in figure 5.1 to identify any variations between the envelopes. It is reasonable to state there is little to no variation between the rotor speed data extraction. The y-axis unit for rotor speed in SKF @ptitude observer is given as counts per minute (CPM) and is consistently higher from October 2022 to April 2023, than when compared from April to October 2022. This is most likely from higher wind speeds in the winter months, showing clearly the seasonality aspect in rotor speed. Comparing the rotor speed of both WTs show the WT 21 5.1b has been operational longer than WT 4 in 5.1a, where there are two large gaps in the data in January and in April 2023. A reason for this can be from scheduled maintenance on WT 4. The exact reasons is difficult to know, without the access of maintenance logs.



(a) WT 4 (b) WT 21

Figure 5.1: WT Rotor Speed data

Vibration data

WT 4

The **generator bearing axial vibration** in different measurement frequencies in ENV 1, 2 and 3 together with their corresponding preset warning and alarm thresholds are plotted respectively as 5.2a, 5.2d and 5.2g in figure 5.2. The axial acceleration of the shaft is plotted similarly in section 5.2j. The vibration signals of ENV 1, 2 and 3 show generally a stable vibration and it is clear the amplitudes of the signals are quite different. From the preset thresholds, ENV 1 and 2 indicate normal bearing operating condition, while ENV 3 consistently crosses the alarm threshold. Indicating some possible fault development.

The validity of the warning and alarm thresholds are questionable. In the case of ENV 2, the alarm threshold was increased by 11 fold and the warning threshold by tenfold at the end of April 2022. The reason for this is unclear. In contrast to ENV 3, both thresholds were decreased slightly around May 2022 and the vibration signals are more visibly over passing the alarm threshold consistently.

There are a few spikes as in the case of ENV 1 from December 2022 to February 2023 and in ENV 3, there is a large spike in March 2023. After guidance with analysing vibration signals from an SKF bearing engineer, the enveloped high PtP amplitude vibration spikes can be ignored, which can be caused by a number of reasons, including background noise. However, minor bearing component impacts can be a possible reason.

The axial acceleration of the shaft in plot 5.2j given in the y-axis unit 'g P' where P is the peak amplitude, show the acceleration is well below the warning threshold. It is generally stable from April to October 2022, after which it decreases abruptly afterwards. The exact reasons are unclear and can be from change in sensor measurement settings.

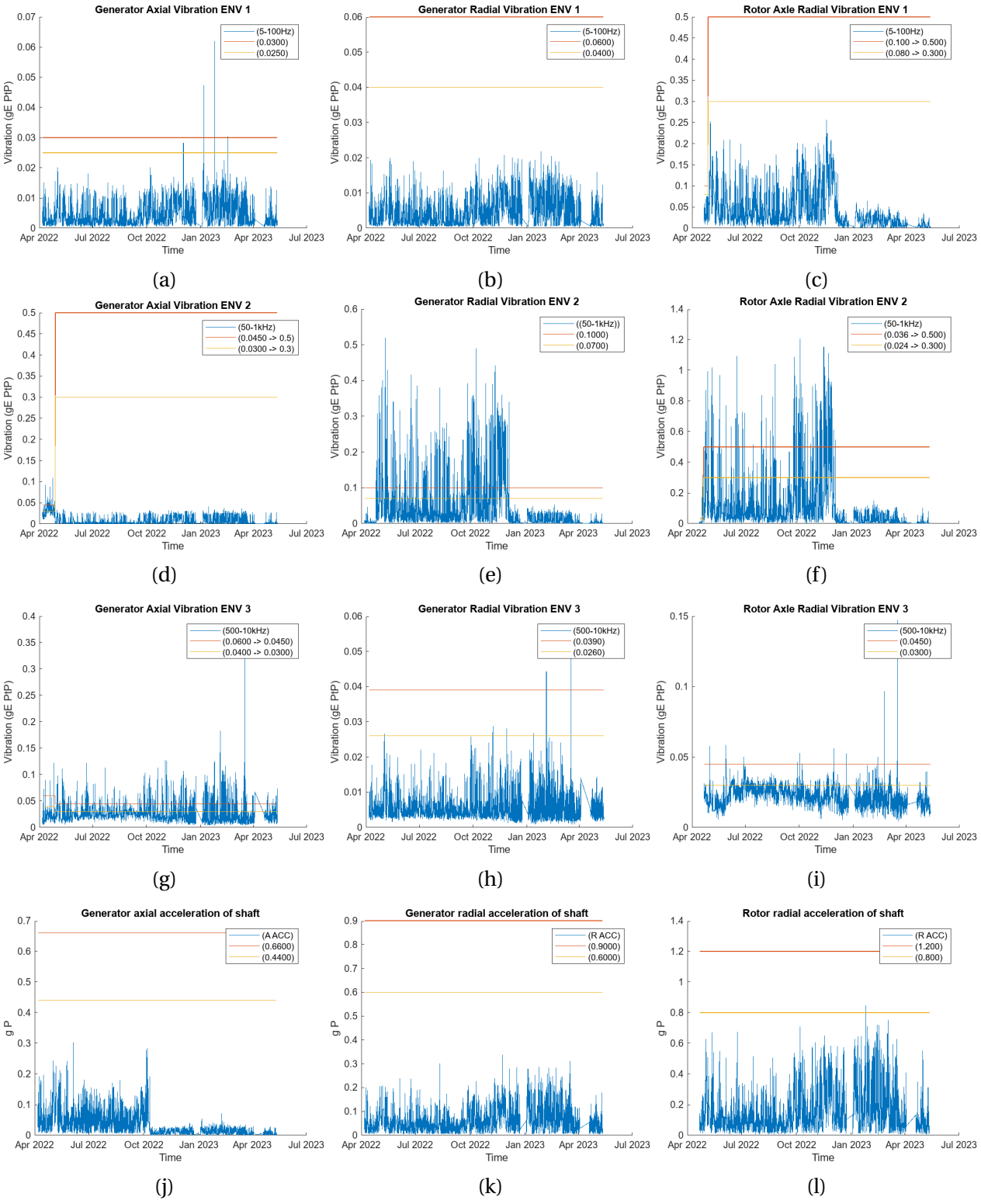


Figure 5.2: WT 4 shaft, generator and rotor bearing vibration data

The **generator bearing radial vibration** in ENV 1, 2 and 3 together with the warning and alarm thresholds are plotted respectively in 5.2b, 5.2e and 5.2h. The radial acceleration of the shaft is plotted in 5.2k. The vibration signals of ENV 1 and 3 are generally lower than the warning thresholds and have similar PtP amplitudes. There are a few anomalous spikes in ENV 3 in February and March 2023. This is similar to the generator axial vibration in ENV 3, which may indicate some minor bearing component impact as in bearing wear stage two.

The ENV 2 vibration signal show a strange behaviour compared to ENV 1 and 3. There are high amplitude spikes consistently from the end of April to December 2022. There seems to be a low radial vibration from the start to the end of April 2022, and after which increases suddenly. The highest amplitudes are 0.5 gE PtP and is clearly above the alarm thresholds. From the start of January 2023, the radial vibration is greatly lower and below the warning thresholds. It is rather unclear the exact reasons for the vibration pattern in ENV 2, without the availability of maintenance logs. One possibility is maintenance service may have been performed at the end of December 2022, due to the high radial vibration. The sudden increase in vibration could have been caused by sensor errors or from changes in sensor measurement settings for example. Assumptions have to be made on the bearing condition vibration data when training an fault detection model. The clear variation in the ENV 2 radial vibration from possible maintenance service, allows for labeling the data into 'before' and 'after' maintenance with condition labels. With this, the fault detection models can be trained using supervised machine learning methods.

The radial acceleration of the shaft in plot 5.2k is generally well below the warning threshold. There are a few clearly visible data gaps. The peak amplitudes after the data gap in April 2023, are roughly half the size compared to before, indicating possible maintenance service. The different rotor speed of WT 4 is visible in figure 5.1a from changing wind speeds from winter to summer months. This can be a reason for lower vibration in the shaft, when overall rotor speeds are decreased.

The **rotor bearing radial vibration** in ENV 1, 2 and 3 are plotted respectively as 5.2c, 5.2f and 5.2i. The radial acceleration ACC of the shaft is plotted in section 5.2l. The vibration signals of ENV 1 is generally lower than the warning thresholds, while ENV 2 and 3 consistently cross it. There are a few high amplitude anomalous spikes in ENV 3 in March 2023. This is similar to what was observed in the generator axial and radial vibration in ENV 3. Similar vibration pattern across the two bearings in the same measurement frequency envelope 3 are most likely be caused other systems within the WT4, such as increased shaft vibration from high rotor speed.

As seen in the generator bearing radial ENV 2 vibration, the rotor bearing ENV 2 show a similar behaviour and it is rather unclear the exact reasons for the vibration pattern without the availability of maintenance logs. It is unexpected since, they are two different bearings showing similar vibration pattern. There are high amplitude spikes consistently from April to December

2022. The highest amplitudes are above 1 gE PtP, clearly crossing the alarm thresholds and is double than what is found in the ENV 2 radial vibration of the generator bearing. From the start of January 2023, the radial vibration is greatly lower and well below the warning threshold. The sudden decrease in PtP amplitude in ENV 2 at the end of December 2022 is also evident in ENV 1, showing a similar vibration signature.

The radial acceleration of the shaft in plot 5.2l is well below the warning threshold, while showing anomalous spikes. Similarly to the generator shaft ACC, the peak amplitudes after the data gap in April 2023 are much lower compared to before. This indicates possible maintenance service or can be due to the lower rotor speeds after the data gap as seen in figure 5.1a.

WT 21

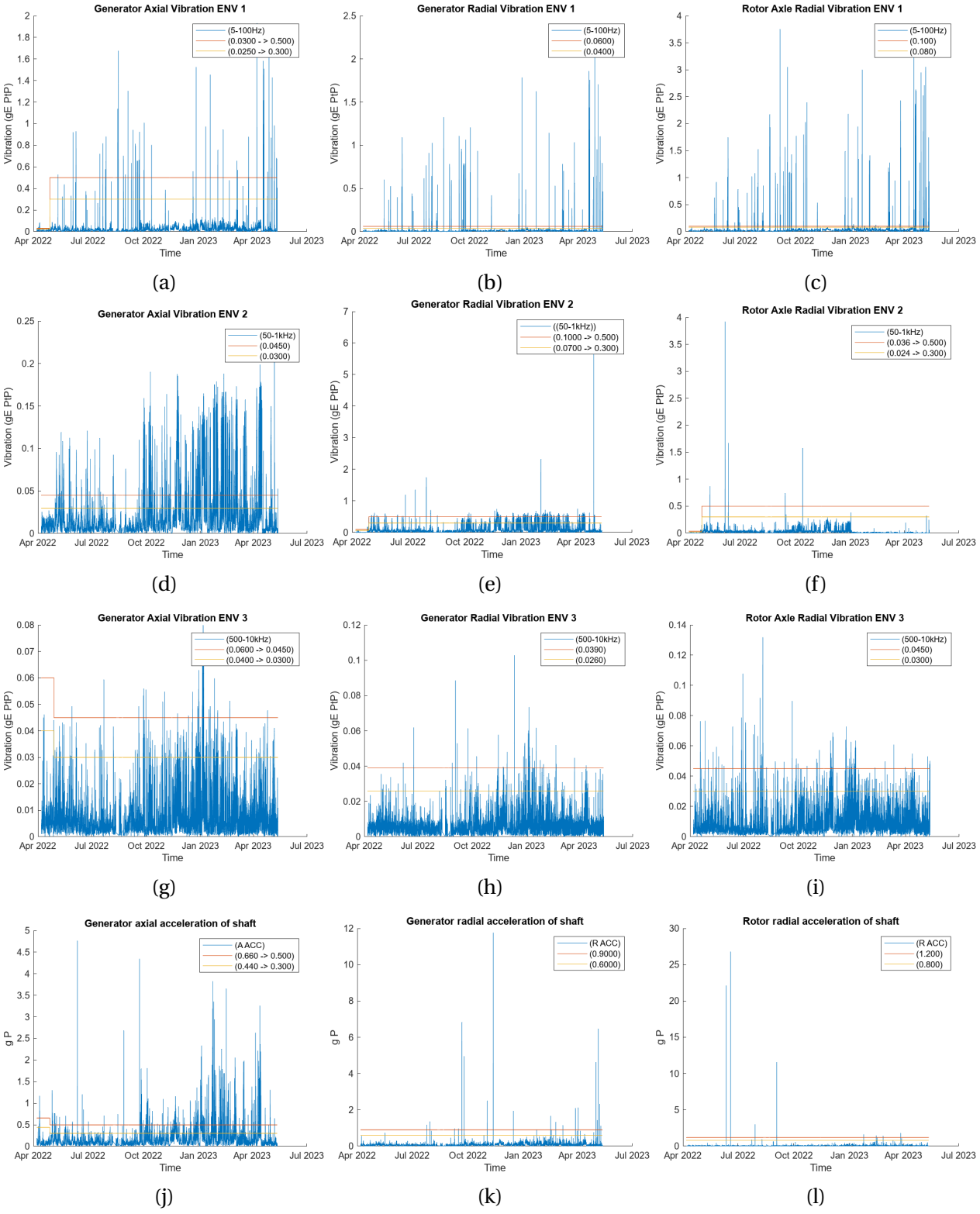


Figure 5.3: WT 21 shaft, generator and rotor bearing vibration data

The **generator bearing axial vibration** in ENV 1, 2 and 3 are plotted respectively as 5.3a, 5.3d and 5.3g. The axial acceleration of the shaft is plotted in 5.3j. The vibration signals of ENV 1, 2 and 3 all show high PtP amplitudes and are in some cases well over the alarm threshold. It is clear the amplitudes of the signals are quite different. ENV 1 show high amplitude spikes of a maximum 1.7 gE PtP, which is the highest compared to ENV 2 and 3. The frequency of the spikes differ greatly between the envelopes.

It is clear the trend for the axial vibration of the generator bearing from all 3 envelopes and the axial acceleration ACC of the shaft is increasing steadily and notably from October 2022 to February 2023. A few infrequent high amplitude spikes in the graph can be from external factors and background noise, which can be treated as outliers. However, it is clear the general trend is increasing. The peak amplitude of the axial acceleration of the shaft in February 2023 is over 3 g P, which is significantly higher than when compared to WT 4 generator axial acceleration of the shaft in 5.2j. After February 2023, a decreasing trend is visible from all four plots of the generator bearing and shaft axial vibration, however a large spike in vibration amplitude is seen in 5.3d and 5.3j. These increasing and decreasing trends in vibration are clearly similar to the trend seen in the WT 21 rotor speed plot in 5.1b, from October 2022 to April 2023. This strongly suggests the axial vibration generated is highly dependent on the rotor speed, which itself is dependent on the seasonality aspects.

In comparison to WT4, the vibration patterns clearly indicate more degradation and movement in the axial direction of the bearing and shaft. From chapter 3, the greatest forces are experienced perpendicular to the motor shaft, which is in radial direction. The axial vibration is expected to be low. In this case however, the axial vibration levels are generally high as in ENV 1, 2 and ACC of the shaft. Hence, this most likely is the result of misalignment or even a bent shaft, causing greater axial movement in this rotatory machinery.

The **generator bearing radial** vibration in ENV 1, 2 and 3 are plotted respectively in 5.3b, 5.3e and 5.3h. The radial acceleration of the shaft is plotted in 5.3k. The vibration signals of ENV 1 show signs of infrequent high amplitude spikes, while the majority of the signals are within the warning threshold. Minor bearing component impacts from lubrication issues can cause the high amplitude spikes, however, they should be treated as outliers when this pattern is not clearly visible in the other envelopes.

In ENV 2 and 3, the signals show an increasing trend, notably from October 2022 to April 2023. This is similar to what is observed with the WT 21 generator axial vibration in ENV 2 5.3d and ENV 3 5.3g. Indicating the bearing may most likely be experiencing greater degradation. The signals are below the warning threshold in the beginning of October 2022 and gradually increase to cross the alarm threshold. There are also some infrequent high amplitude spikes in ENV 2 and 3, which can be ignored from the general trend.

The radial acceleration of the shaft in plot 5.3k show the majority of the signals are below

the warning thresholds, in spite of the infrequent high amplitude spikes above 6 g P. The increasing radial acceleration of the shaft from October 2022 to February 2023, is similar to what is observed with the bearing radial vibration in the same time frame of ENV 2 5.3e and ENV 3 5.3h. After February 2023, there is an decreasing trend. The seasonality aspects are visible here, as observed in the generator axial vibration of WT 21.

Compared to WT 4 generator bearing, where the vibration amplitudes in ENV 1, 2 and 3 and ACC of the shaft are much lower, indicates the generator bearing of WT 21 may have experienced greater degradation causing higher radial vibration. The axial and radial vibration signals show signs of stage 2 to 3 bearing wear where the increased wear generates higher amplitude vibration with lower frequency, as observed in ENV 1 5.3b.

The **rotor bearing radial vibration** in ENV 1, 2 and 3 are plotted respectively as 5.3c, 5.3f and 5.3i. The radial acceleration ACC of the shaft is plotted in 5.3l. The vibration signals of ENV 1 show signs of infrequent high amplitude spikes of over 2 gE PtP, while the majority of the signals are low amplitude and within the warning threshold. In ENV 2, the vibration signals are stable and low from April 2022 to January 2023, after which decrease suddenly and have similar signal amplitude as in April 2022. The reason for this is unclear and may likely be due to changes in sensor measurement settings. A sudden drop is vibration signals are more likely to be sensor errors, rather than after maintenance service, where in this case, there are not any significant data gaps in January 2023. This signal pattern is similar to the rotor bearing radial vibration of WT 4 in ENV 2 5.2f. For ENV 3, the vibration signals measured in the higher frequency range, show a few vibration spikes from April to October 2022 crossing the alarm threshold. Minor bearing component impacts can be a reason.

The radial acceleration of the shaft in plot 5.3l show the majority of the signals are well below the warning threshold, while showing a few clear outliers of over 20 g P amplitude in June 2022. The gradual increase of the signal amplitude is notable from October 2022 to March 2023, showing the effect on seasonality.

The observations and findings from axial and radial vibration of the bearings and shaft of WT 4 and WT 21 can be used to decide what data can be used for training an fault detection model. It is clear from lack of data gaps in WT 21, that it has been operational longer than WT 4. The notable higher axial and radial vibration signal amplitudes found in the generator bearing in WT21 compared with WT 4, indicates it has experienced more degradation. The WT 4 rotor bearing and shaft radial vibration have similar signal amplitudes as in WT 21, when the signal outliers are disregarded, indicating the rotor bearing and shaft have more or less similar health conditions.

5.2 Frequency domain analysis

SKF @ptitude observer allows for analysing the vibration signals in the frequency domain using the 'Spectra' tool. The purpose is to identify the current condition of the bearings by isolating and diagnosing any bearing faults, using industry practice methods such as looking for the vibration amplitude at the four bearing fault frequencies. Spalling faults in the bearing can be easily identified when the magnitude of the vibration amplitude is high at the fault frequencies. The lack of maintenance logs, means FFT spectrum analysis is the only method available to find any possible developed bearing fault. The spectra graphs for the rotor and generator bearing of WT 4 and 21 are given in appendix chapter 5. Numerous bearing fault frequencies are already calculated and stored in a catalogue within SKF @ptitude observer. This eliminates the need to calculate the fault frequencies by hand, when not all bearing specification data are readily available online. Table 5.1 shows the four bearing fault frequencies of the Bessakerfjellet WT rotor and generator bearings.

Table 5.1: Bessakerfjellet WT rotor and generator bearing fault frequencies

Fault type	**Rotor Bearing frequencies**	**Generator Bearing frequencies**
Outer race (BPFO)	22.5966	20.7102
Inner race (BPFI)	25.4034	23.2902
Roller (BSF)	16.68	16.995
Cage (FTF)	0.471	0.4704

SKF @ptitude observer allows for plotting together the spectra graph of multiple different measurement frequency envelopes, which effectively simplifies the fault isolation and estimation process. For WT 4, figures B.3 and B.4 provided in the appendix shows the generator bearing and rotor bearing spectra graphs respectively. All the available axial and radial vibration envelopes are plotted together. The four fault frequencies marked with a vertical line at the respective frequencies. The vibration amplitude at the fault frequencies for both bearings are generally low, below 0.002 gE PtP. This indicates there are not any developed bearing faults present at the fault frequencies. There are however, higher vibration amplitudes at other frequencies, showing no clear pattern. The ENV 3 axial shows higher amplitudes across the frequency domain, due to the high raw vibration amplitude as seen in 5.2g. It is difficult to interpret such patterns in the spectra graph without more frequency domain vibration analysis knowledge. From the bearing fault frequencies, it is reasonable to conclude both bearing show normal healthy conditions.

For WT 21, figures B.5 and B.6 provided in the appendix show the generator and rotor bearing spectra graphs respectively. The radial envelopes vibration amplitude across the rotor bearing spectra graph are stable and there are not increased amplitudes at the fault frequencies. For the

generator bearing however, there are clearly increased vibration amplitudes at the inner race (BPFI) and cage (FTF) fault frequencies, suggesting some spalling faults are present. Only the generator bearing axial envelope 1 and 2, show the BPFI and FTF fault development. The amplitudes at the fault frequencies are not significantly high as they are only around 0.0045 gE PtP, suggesting the severity of the BPFI and FTF faults are low. However, there is most likely some minor defect in the inner race and possibly the cage causing the increase in vibration amplitudes. Figure 5.4 show signs of harmonics caused by the BPFI fault in the generator bearing spectra graph, which slowly decrease for every four order. These harmonics further damage the bearing components.

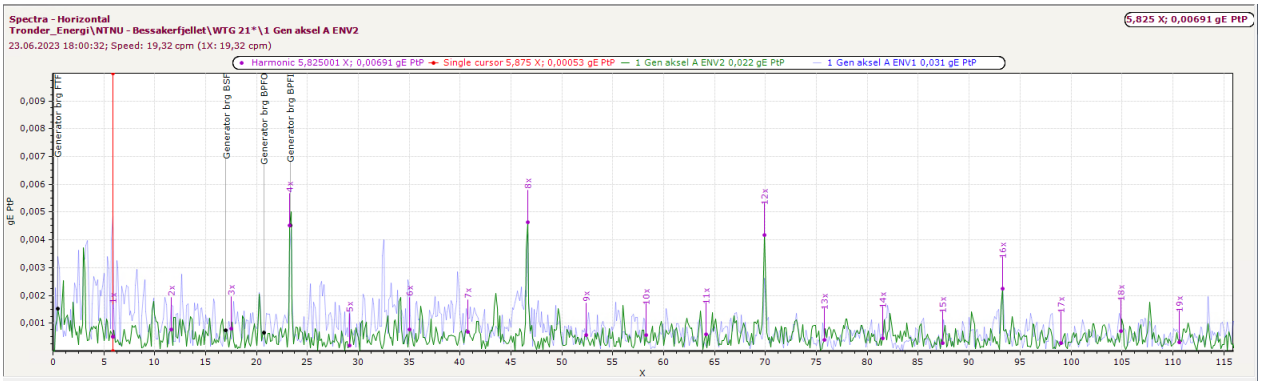


Figure 5.4: WT 21 Generator Bearing Axial ENV 1 and 2 BPFI Fault Frequency Harmonics

5.3 Correlation study

A correlation study similar to what I conducted in the specialization project (Thiruthiyappan, 2022) is carried out here, since more vibration and rotor speed data is available. The generator bearing and shaft radial and axial vibration, and rotor bearing radial vibration in different measurement frequency envelopes, together with the WT rotor speed is used in the correlation analysis. The purpose is to identify the relationship between the variables and better understand the condition of the bearings and shaft. Plotting together with the rotor speed gives the vibration data another dimensionality, with the aim of giving more insight.

WT 4

All the vibration data available from WT 4 from April 2022 to April 2023 is used. Figure 5.5 show scatter plots from the analysis, where the vibration is plotted in the y-axis and the rotor speed on the x-axis. The scatter plots are arranged so the first column has the generator bearing and shaft axial vibration plotted with the rotor speed. The second column shows the generator bearing and shaft radial vibration. Similarly, the rotor bearing and shaft radial vibration in the third

column. The correlation coefficients are provided in the legend in the scatter plots.

Both linear and non-linear patterns can be identified. There seems to be a similar linear pattern across the generator bearing ENV 1 axial and radial vibration in 5.5d and 5.5e. They both have a very strong correlation coefficient above 0.8. Both scatter plots however show a distinct spike in vibration amplitude at a particular rotor speed of roughly 11.8 cpm. This is rather strange behaviour and was not previously identifiable from their individual generator bearing ENV 1 axial and radial vibration plots in 5.2a and 5.2b. A characteristic of stage two bearing wear from chapter 3 is the development of resonances of the machine, caused by high frequency vibration increasing in amplitude. The increase in axial and radial vibration amplitude of the generator bearing in ENV 1 at a particular rotor speed, indicates the presence of resonance in the machine. It is however, difficult to deduce if the resonance is solely caused by the WT shaft rotating at a particular rotor speed of 11.8 cpm. The distinct spike in vibration amplitude is similarly visible in the rotor bearing ENV 1 radial vibration 5.5f at a rotor speed of 11.8 cpm. The shaft at the generator and rotor bearing radial acceleration in 5.5b and 5.5c similarly show signs of a spike in vibration amplitude at a rotor speed of 11.8 cpm. It is however, not clearly visible in the shaft axial acceleration at the generator in 5.5a.

The generator and rotor bearing ENV 2 radial vibration in 5.5h and 5.5i show two distinct clusters in the scatter plot. This supports the hypothesis of possible maintenance service, previously deduced from the raw generator and rotor bearing ENV 2 radial vibration analysis. The top cluster shows a exponential growth in vibration, when rotor speed increases, while the bottom cluster shows a minimal linear increase in vibration as rotor speed increases. This may possibly be caused by change in sensor measurement settings as well.

The lowest correlation coefficients are found in ENV 3, generator bearing axial and radial vibration in 5.5j and 5.5k, as well as in the rotor bearing radial vibration in 5.5l. This indicates the high frequency vibration generated is similar in amplitude when the rotor speed increases, which is a good sign. There is however a visible spike in the vibration amplitude at the highest rotor speeds of above 20 cpm.

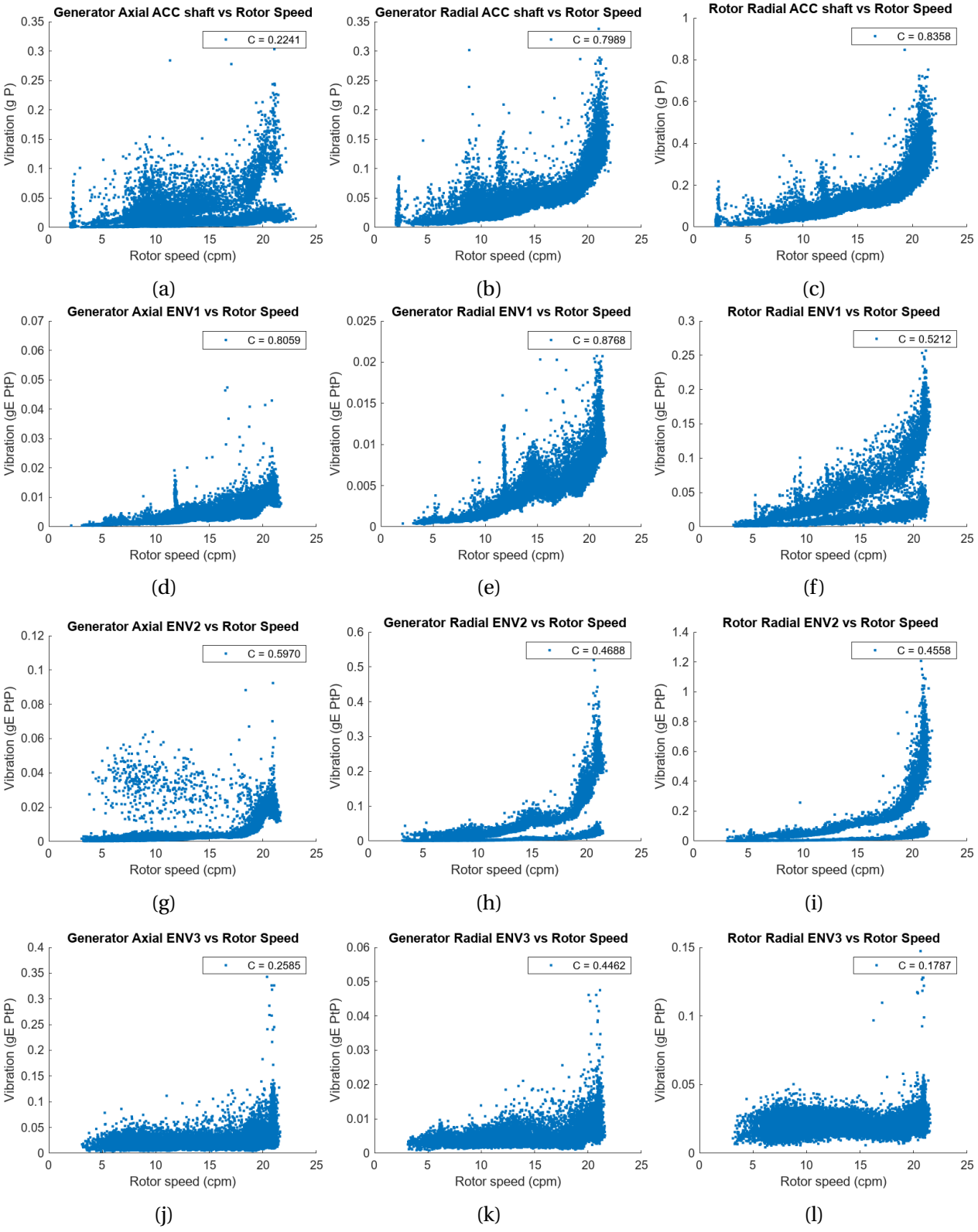


Figure 5.5: WT4 correlation analysis

WT 21

Similarly to WT 4, all the vibration data available from April 2022 to April 2023 is used. Figure 5.6 show scatter plots from the analysis, where the vibration is plotted in the y-axis and the rotor speed on the x-axis. The scatter plots have the same arrangement as in figure 5.5. The correlation coefficients are provided in the legend of the scatter plots.

Non-linear patterns can be identified. Similarly to WT 4, there seems to be a similar pattern across the generator bearing ENV 1 axial and radial vibration in 5.6d and 5.6e. The rotor bearing ENV 1 radial vibration in 5.6f shows a similar pattern. The correlation coefficients are very low, with the presence of outliers which are similar in the three scatter plots 5.6d, 5.6e and 5.6f. These outliers have high amplitude and occur from rotor speeds from 5 to 10 cpm. A upper limit was applied to the y-axis to better analysis the vibration patterns in ENV 1, hence, the outliers are not visible. Similarly to WT 4, the ENV 1 vibration scatter plots of the generator and rotor bearing show a distinct spike in vibration amplitude at the particular rotor speed of roughly 11.8 cpm. This is rather strange behaviour and was also not previously identifiable from their individual generator bearing ENV 1 axial and radial raw vibration plots in 5.3a and 5.3b, and rotor bearing ENV 1 radial vibration in 5.3c. The stage two bearing wear characteristic of the development of resonances in the machine, caused by high frequency vibration increasing in amplitude is clearly visible here. The increase in axial and radial vibration amplitude of bearings in ENV 1 at a particular rotor speed, supports the hypothesis of the development of resonance in the machine. As mentioned, it is difficult to deduce if the resonance is solely caused by the WT shaft rotating at a particular rotor speed of 11.8 cpm. For WT 21 however, the distinct spike in vibration amplitude is not clearly visible in the shaft axial and radial direction at the generator and rotor. The ENV 3 plots in this case do show a small visible spike at a rotor speed of 11.8 cpm, in 5.6k and 5.6l.

Similarly to WT 4, the generator and rotor bearing ENV 2 radial vibration in 5.6h and 5.6i show two distinct clusters in the scatter plot. The top cluster shows a exponential growth in vibration, when rotor speed increases, while the bottom cluster shows a minimal linear increase in vibration as rotor speed increases. This suggests possible maintenance service, however, due to no data gaps in WT 21, while it is present in WT 4, it is unlikely to be caused by maintenance action. The change in sensor measurement settings for ENV 2, is more likely to have caused the two distinct clusters.

The correlation coefficients found in ENV 3 generator bearing axial and radial vibration in 5.6j and 5.6k, as well as in the rotor bearing radial vibration in 5.6l, are low as similar to WT 4. This indicates the high frequency vibration generated is similar in amplitude when the rotor speed increases, which is a good sign. There is however a visible spike in the vibration amplitude at the highest rotor speeds of above 20 cpm as in WT 4.

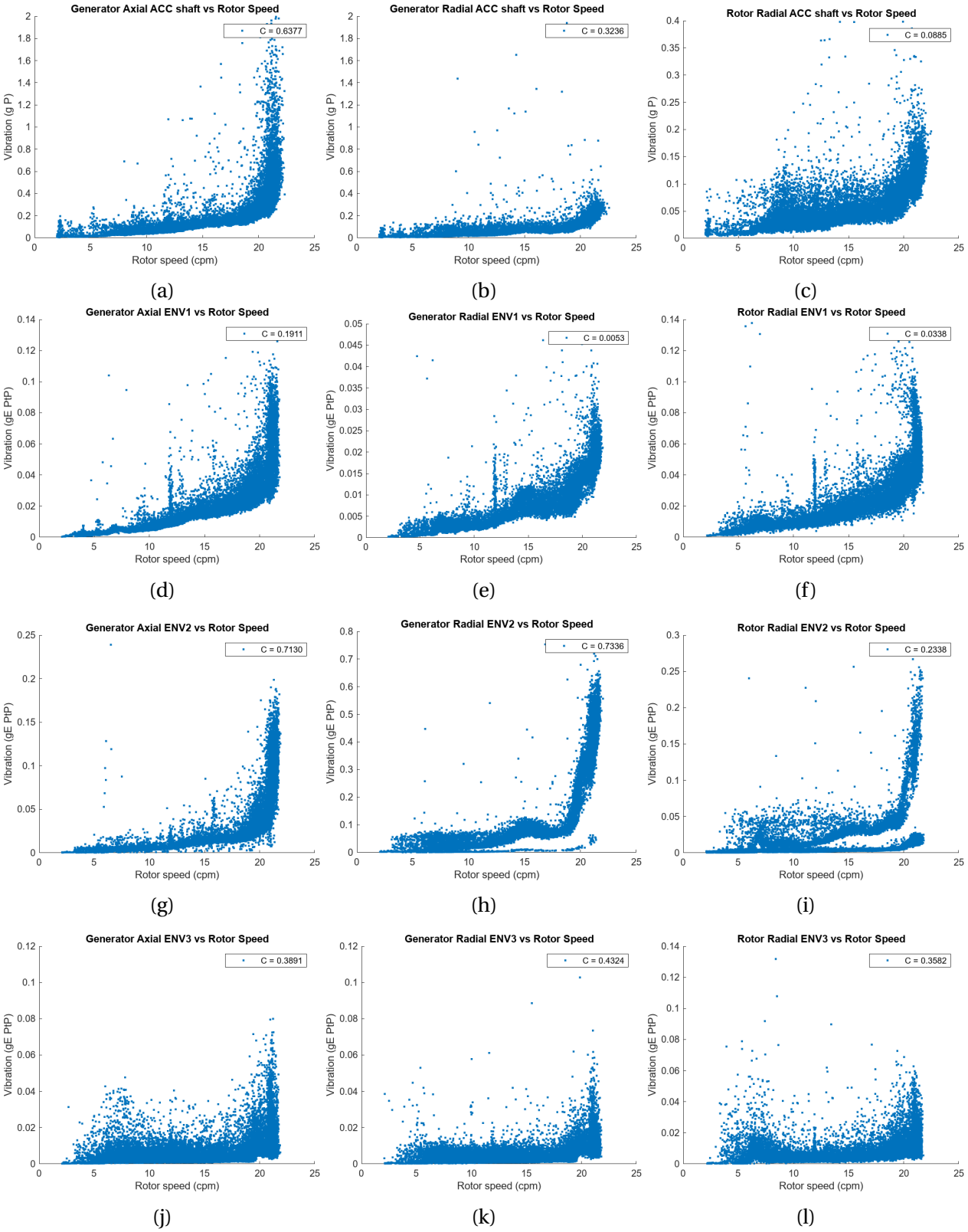


Figure 5.6: WT21 correlation analysis

5.4 Data preprocessing

The generator and rotor bearing ENV 2 radial vibration data are disregarded, due to the changes in sensor measurement settings giving incorrect formatted data. This is confirmed through a guidance session on vibration analysis by a SKF bearing engineer.

5.4.1 Data cleaning

The acceleration enveloped bearing axial and radial vibration in different envelopes go through a series of band pass filters which remove the low frequency machinery noise from the raw vibration. This isolates the periodic and repetitive vibrations in the bearing. The signal is rectified (demodulated) where the negative amplitudes become positive, increasing the signal density. The acceleration ACC of the shaft goes through the same enveloping method. This process is performed automatically in SKF @ptitude observer and additional signal cleaning steps are required to remove outlier and background noise data as seen in many of the enveloped vibration signal graphs of WT 4 and especially of WT 21. Missing data fields are removed and signal smoothing is performed. MATLAB's 'Clean data' tool is used to perform the data cleaning process. To remove the sensor outlier data as verified by an SKF Bearing engineer, the default settings in MATLAB 'clean data' tool involve a moving median using a threshold factor of 3 and moving window size of 24 days. This technique effectively removes the majority of high amplitude signal spikes. Signal smoothing with a moving mean method of smoothing factor 0.05 is used to further reduce the high amplitude spikes and filter out the background noise. The cleaned data set of the WTs are used for fault detection model development. The rotor speed data is used as extracted from SKF @ptitude observer. A MATLAB function for the data cleaning process is made, to perform the same process for future extracted data.

5.4.2 Data format

A file ensemble datastore structure of the WT vibration data is required to use the MATLAB Diagnostic feature designer (DFD) tool for feature extraction. The vibration data is stored in a table of three columns, which are date, vibration and rotor speed. The table is converted to a timetable, so the time stamp in the date column becomes associated with the rotor speed and vibration data rows. The shaft, generator and rotor bearing vibration data are stored in separate time tables as extracted from SKF @ptitude observer. A for loop is hence written to combine the generator bearing and generator shaft day to day data in one ensemble datastore. This is done similarly for the rotor bearing and the rotor shaft vibration data.

Figure 5.7 shows the generator bearing, shaft axial and radial vibration data, where the first column has the data for one day stored in a timetable for each variable. The second row has the

next day worth of data and so on. Additionally, the sample rate of the data is different for the variables in a few cases and are evident in figure 5.7.

Generator21ARENV123ACC_ensemble_table = 394×8 table

	RotorSpeed	AENV1	AENV2	AENV3	AACC	RENV1	RENV3	RACC
1	10×1 timetable	10×1 timetable	11×1 timetable	11×1 timetable	11×1 timetable	9×1 timetable	10×1 timetable	11×1 timetable
2	23×1 timetable	23×1 timetable	23×1 timetable	24×1 timetable	22×1 timetable	20×1 timetable	24×1 timetable	22×1 timetable
3	24×1 timetable	24×1 timetable	24×1 timetable	23×1 timetable	23×1 timetable	23×1 timetable	20×1 timetable	23×1 timetable
4	24×1 timetable	24×1 timetable	24×1 timetable	24×1 timetable	24×1 timetable	24×1 timetable	24×1 timetable	24×1 timetable
5	24×1 timetable	24×1 timetable	24×1 timetable	20×1 timetable	24×1 timetable	24×1 timetable	23×1 timetable	24×1 timetable
6	24×1 timetable	24×1 timetable	24×1 timetable	22×1 timetable	23×1 timetable	24×1 timetable	24×1 timetable	24×1 timetable
7	23×1 timetable	23×1 timetable	22×1 timetable	21×1 timetable	23×1 timetable	23×1 timetable	22×1 timetable	23×1 timetable
8	20×1 timetable	20×1 timetable	20×1 timetable	20×1 timetable	20×1 timetable	20×1 timetable	19×1 timetable	18×1 timetable
9	24×1 timetable	24×1 timetable	24×1 timetable	24×1 timetable	24×1 timetable	24×1 timetable	23×1 timetable	24×1 timetable

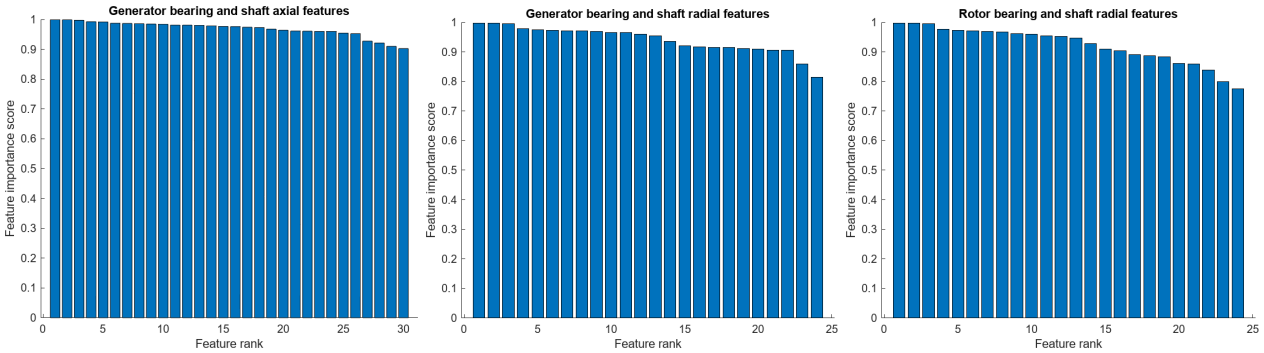
Figure 5.7: Ensemble datastore of WT 21 Generator bearing, shaft axial and radial vibration data

5.4.3 Diagnostic feature designer

The two ensemble datastore structures of the generator and rotor bearing along with the shaft vibration data is imported to DFD. Due to the low sample rate of the enveloped vibration data, only the time signal statistical metrics such as mean, root mean square (RMS), standard deviation, shape factor, kurtosis and skewness were extracted. More time signal metrics are available in DFD, however, these metrics give no values (NaN) or $-\infty$ for a many instances. This gives issues for the fault detection model development, and therefore only the six time signal statistical metrics are used. The features are exported as tables. Only radial vibration is measured for the rotor bearing and since both radial and axial measurement are taken for the generator bearing, it is decided to spilt it into two separate feature tables. The fault detection model functions in MATLAB only work with a numeric array, so the three tables are converted from table to a numeric array and any missing data fields are removed.

Feature ranking

The unsupervised Laplacian score feature ranking method is used to select the top important features. For the generator bearing and shaft, there are originally 30 feature columns, while the rotor bearing and shaft has 24 columns. The 'fsulaplacian' function in MATLAB is used and the top 15 ranked features are selected. Figure 5.8 shows three bar charts from the Laplacian score feature ranking. Overall, the first 15 features have a feature importance score of above 0.9 and are selected for training the fault detection models. Features from all the variables are found in the top 15 features.



(a) (b) (c)

Figure 5.8: WT21 Laplacian score feature ranking

5.5 Fault detection model training

The available vibration data of the generator and rotor bearing and shaft, extracted from SKF @ptitude observer is unlabeled data. This means the exact condition of the bearings is unknown, as it is not clearly stated anywhere in the program. There is no direct integration in SKF @ptitude observer with the Bessakerfjellet WT 4 and 21 maintenance records. Therefore, it is not possible to know if any maintenance has been performed when looking at the data gaps in the historical data. The lack of accessibility to the Bessakerfjellet wind farm maintenance logs information and other supplementary information of WT 4 and 21 from the wind farm operator limits what can be achievable in the fault detection model development here.

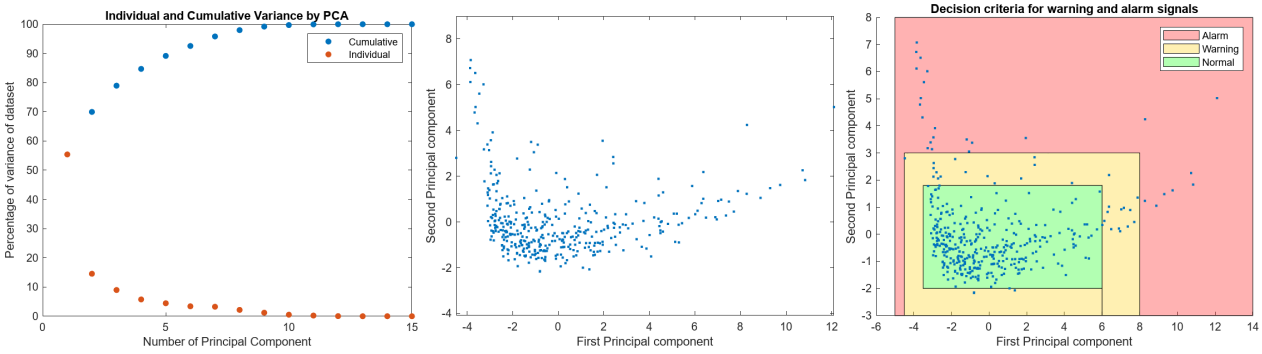
The WT 4 and 21 raw vibration data visualization for the past year show mainly signs of the seasonality aspects. Generally higher vibration of the shaft, generator and rotor bearing for WT 4 and especially WT 21, from higher wind speeds in the winter months is visible. The frequency domain vibration analysis showed signs of minor BPFI and FTF faults in the generator bearing of WT 21. These are important findings and increases the knowledge of the generator bearing condition at the time of analysis. It is however highly difficult to concretely label when these faults may have begun to develop from the raw vibration data. For these reasons, only unsupervised fault detection models will be applied. Additionally, the vibration amplitudes at the fault frequencies are still substantially low, indicating only a very minor defect of the inner race and cage. The fault severity is hence low. Compared to WT 4 where the vibration data show generally stable levels with little to no clear trend than in WT 21. The sign of developing fault gives a good case for detecting anomalies. Hence, it is decided to use the data from WT 21 for training the fault detection models. MATLAB code from an unsupervised anomaly detection tutorial in (Unsupervised_Anomaly_Detection, 2023), on using different models such as isolation forest, local outlier factor, one class SVM and Mahalanobis distance is adapted and used. This is presented in the appendix B.

The correlation study of both WTs shows signs of higher vibration with increasing rotor speed as expected, in both linear and non-liner correlations. An important finding in both WTs is the sign of machine resonance from higher vibration amplitudes at rotor speeds from 11.5 to 12 CPM. Resonance in a rotatory machine can greatly lead to high damage in the shaft and bearing components, and the surrounding sub-assemblies in the WT. Emphasis of identifying the early signs of machine resonance is placed on the fault detection models.

5.5.1 Principal component analysis

PCA is performed initially to assess if it is suitable for using as a unsupervised fault detection technique using the first and second principal component, together with a decision criteria for warning and alarm levels. This technique is found in the MATLAB tutorial by (Filion, 2019). The MATLAB code adapted and used from this tutorial for this case is presented in the appendix B. Before performing PCA, the 3 top features table from WT 21 are standardized using the 'normalize' MATLAB function, to have all the values of the variables to contribute equally in PCA. After performing PCA, the principal components coefficients, scores and variances are returned.

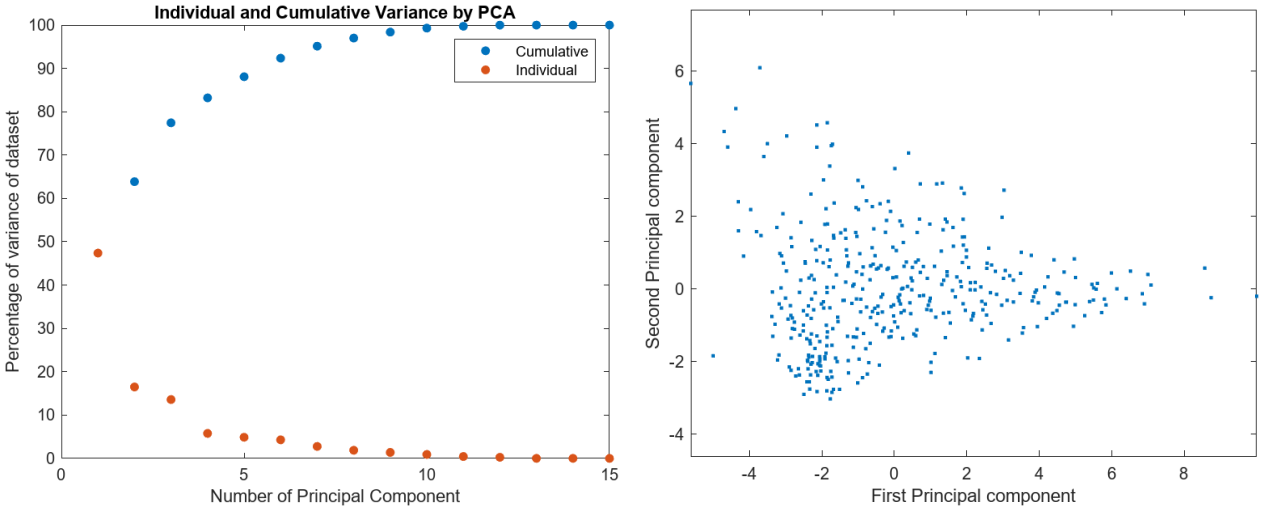
Figure 5.9 shows the PCA plots for WT 21 generator bearing and shaft axial vibration. A scatter plot of total variance which is explained by the individual principal components and the cumulative of the principal components is plotted. Here the x-axis represents the number of principal components and the y-axis is the percentage of variance of the data set. In plot 5.9a, the individual variances of the principal components are marked with orange circles and the cumulative variances are marked with blue circles. Here, the first two principal components can only explain roughly 70 % of the variance in the data, and it takes minimum 5 principal components to explain at least 90 %. The scatter plot 5.9b shows the first and second principal component scores plotted together, which shows a group of data points clustering in the third quadrant of the graph. From here, there are data points which drift upwards to the second quadrant and also to the right in the first quadrant. This gives an early indication, the data points may drift away from the group of cluster in the third quadrant, to the first and second quadrant, as a fault begins to develop. Using this scatter plot, another plot in 5.9c with an initial decision criteria for the warning and alarm signals are created from how the first and second principal component scores data are clustered. The initial criteria is set from how the data points in 5.9b is spread out and hence when more data of the bearing and shaft in a degraded state to allow for setting the warning and alarm thresholds more reasonably. After guidance with Rohit Agrawal from MathWorks, the total variance of the first and second principal components should be expected to explain a minimum of 80 %. In this case, the cumulative variance is only about 70 % and is not the most optimal to use with the decision criteria.



(a) (b) (c)

Figure 5.9: WT21 Generator Bearing and Shaft Axial PCA

Figure 5.10 shows the PCA plots for the generator bearing and shaft radial vibration. In plot 5.10a, the first two principal components can only explain about 64 % of the variance in the data. This is lower than for the generator bearing and shaft axial first and second principal component variance. The scatter plot 5.10b shows the first and second principal component scores, and the data points are clustered in all four quadrants. It is not immediately clear how the cluster will move when fault is developing. Since the cumulative variance of the first two principal components is only 64 %, it would not be optimal to use a decision criteria for warning and alarm thresholds, when the remaining variance in the data can not be captured.

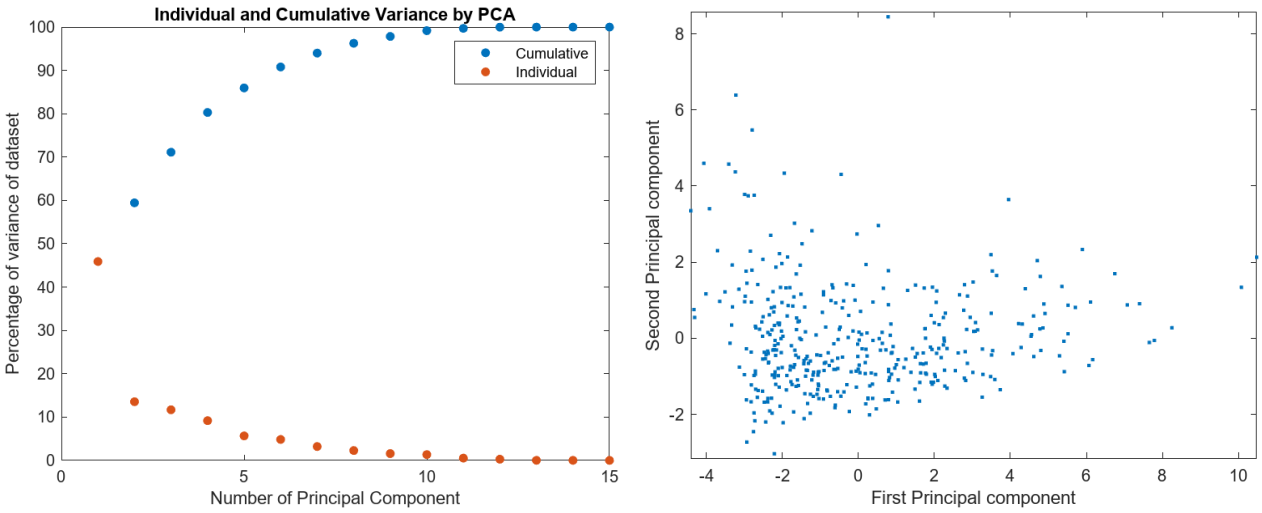


(a) (b)

Figure 5.10: WT21 Generator Bearing and Shaft Radial PCA

Figure 5.11 shows the PCA plots for the rotor bearing and shaft radial vibration. In plot 5.11a, the first two principal components can only explain about 60 % of the variance in the data. This is lower than for the generator bearing and shaft axial and also the radial first and second prin-

cipal component variance. The scatter plot 5.11b shows the first and second principal component scores, and the data points here are also clustered in all four quadrants as in 5.10b. It is not immediately clear how the cluster will move when fault is developing. Since the cumulative variance of the first two principal components is only 60 %, it is not optimal to use a decision criteria for warning and alarm thresholds, when the remaining variance in the data can not be captured.



(a) (b)

Figure 5.11: WT21 Rotor Bearing and Shaft Radial PCA

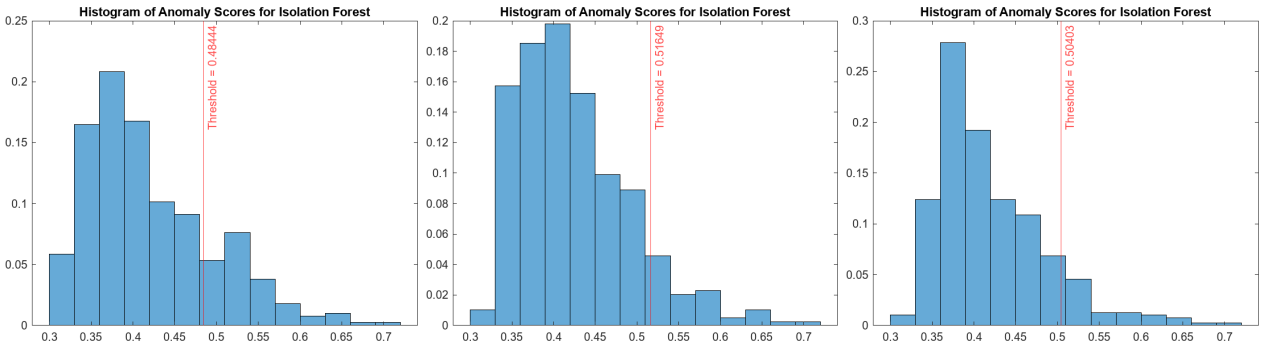
5.5.2 Unsupervised model training parameters

Four different fault detection models are trained with the data from WT 21. A contamination factor is needed to specify the percentage of outliers in the data and is used for calculating the different models score threshold. This is used for detecting the outliers in new data, after the model has been trained. A model threshold factor is set based on the contamination factor used and will detect outliers in the data from this specified fraction. A best judgement from the writer and with guidance from Rohit Agrawal, an individual contamination factor is set for the three cases, which are the generator bearing and shaft axial vibration, generator bearing and shaft radial vibration and rotor bearing and shaft radial vibration. The raw vibration analysis of WT 21 in figure 5.3 is used to specify the contamination factors. For the generator bearing and shaft axial vibration, most of the vibration signals over the alarm thresholds set. For this reason, 20% is used as the contamination factor. For the generator bearing and rotor bearing shaft radial, only a limited amount of vibration signals are over the warning and alarm thresholds in both cases as seen in figure 5.3. Hence, 10% is used as the contamination factor in both cases for training the models. The choice of contamination factors will influence how the models perform

in detecting outliers in new data. All models used are trained with the contamination factors for the three cases, allowing to compare how the different models perform in detecting outliers.

5.5.3 Isolation forest

The isolation forest models trained for the three cases using the contamination factor set and the models return a threshold factor, anomaly scores and anomaly indicators of the data. Histogram of the anomaly scores for all three cases together with the threshold value as a red vertical line is plotted in figure 5.12. Plot 5.12a, where a 20% contamination factor is set, shows a lower threshold value, than in 5.12b and 5.12c, where 10% contamination factor is used. Hence, a larger number of observations are over the threshold value in plot 5.12a. The fraction of the observations detected as outliers is checked for all three cases, by dividing the sum of the anomaly indicators by the number of rows in the features table used. For the isolation forest model in 5.12a, the outlier fraction is 0.2005, for 5.12b is 0.0990 and in 5.12c is 0.0987. These are close to the contamination factors used.



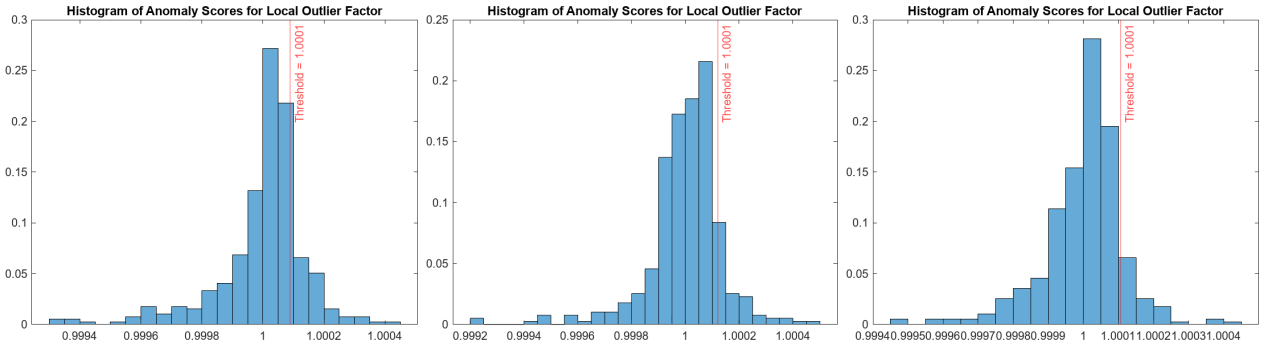
(a) Generator Bearing and Shaft Axial IF (b) Generator Bearing and Shaft Radial IF (c) Rotor Bearing and Shaft Radial IF

Figure 5.12: WT21 Histogram of Anomaly scores for Isolation forest

5.5.4 Local outlier factor (LOF)

Similar to the isolation forest, LOF models are trained for the three cases using the contamination factors set and the models return a threshold factor, anomaly scores and anomaly indicators of the data. Histogram of the anomaly scores for all three cases together with the threshold value is plotted in figure 5.13. Here the threshold value is similar for all three cases, however the distribution of the data is different as expected. Hence, for plot 5.13a,, a larger number of observations are over the threshold value and classified as outliers, while in plots 5.13b and 5.13c, the majority of the observations are below the model threshold value. The fraction of the observations detected as outliers is checked for all three cases. For the LOF model in 5.13a, the outlier

fraction is 0.2005, for 5.13b is 0.0990 and in 5.13c is 0.0987. These outlier fractions are similar to that of the isolation forest models and are close to the contamination factors used.

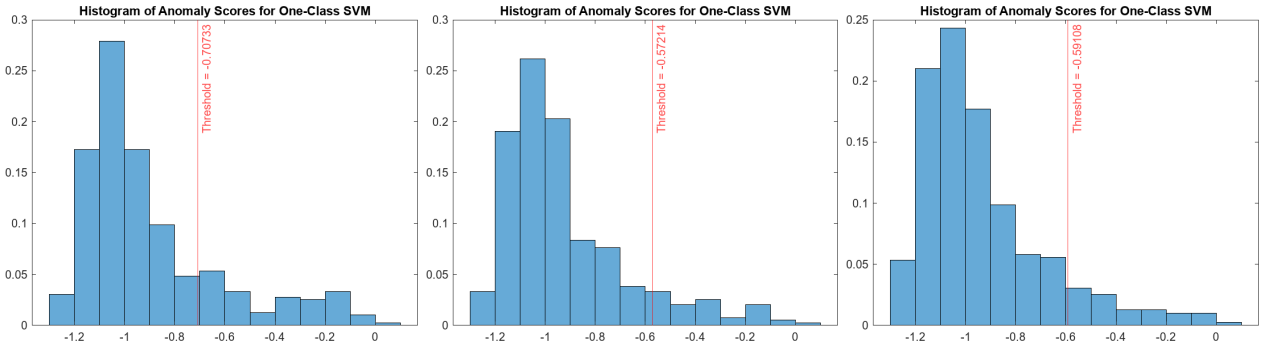


(a) Generator Bearing and Shaft Axial LOF (b) Generator Bearing and Shaft Radial LOF (c) Rotor Bearing and Shaft Radial LOF

Figure 5.13: WT21 Histogram of Anomaly scores for Local Outlier Factor

5.5.5 One-class support vector machine (SVM)

One-class SVM models are trained for the three cases using the contamination factors set and the models return a threshold factor, anomaly scores and anomaly indicators of the data. Histogram of the anomaly scores for all three cases together with the threshold value is plotted in figure 5.14. Plot 5.14a, where a 20% contamination factor is set, shows a higher threshold value, than in 5.14b and 5.14c, where 10% contamination factor is used. For plot 5.14a, a larger number of observations are over the threshold value and classified as outliers, while in plots 5.14b and 5.14c, most of the observations are below the threshold value. The fraction of the observations detected as outliers is checked for all three cases. For the SVM model in 5.14a, the outlier fraction is 0.2005, for 5.14b is 0.0990 and in 5.14c is 0.0987. These outlier fractions are similar to the isolation forest and LOF models and are close to the contamination factors used.

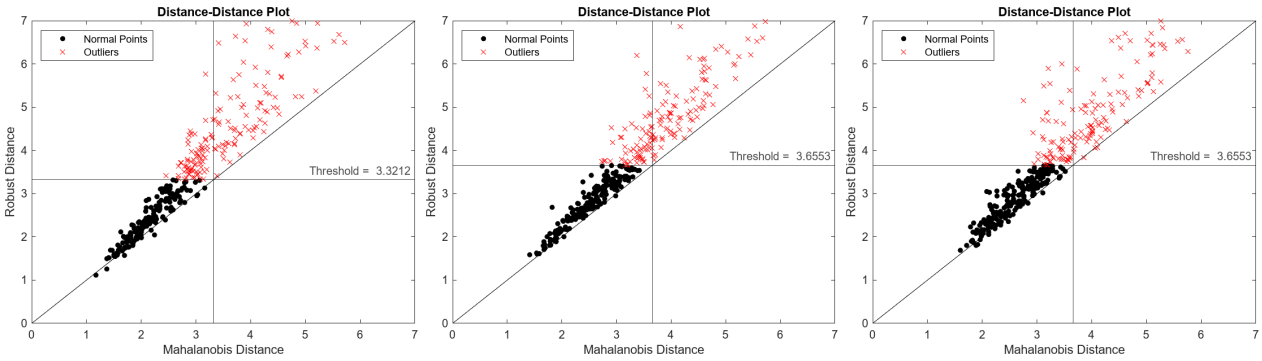


(a) Generator Bearing and Shaft Axial OCSVM (b) Generator Bearing and Shaft Radial OCSVM (c) Rotor Bearing and Shaft Radial OCSVM

Figure 5.14: WT21 Histogram of Anomaly scores for One-class support vector machine

5.5.6 Mahalanobis distance

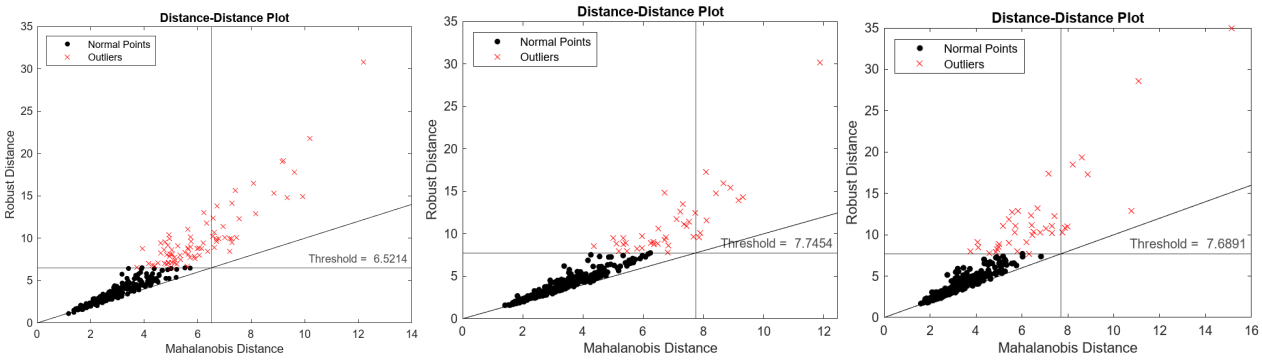
For each case, the robust Mahalanobis distance and the robust estimates for the mean and covariance of the data is calculated using the 'robustcov' MATLAB function. The Mahalanobis distance from features table data to the distribution of the features table data is calculated. The contamination factor decided for each case is used and the 'robustcov' function minimizes the covariance determinant to over 1 - contamination factor of the observations. If the contamination factor is 10%, the function minimizes the covariance determinant over 90% of the observations. The function computes the outlier indicators by default with assuming the data set follows a multivariate normal distribution and identifies 2.5% of the observations as outliers based on the critical values from the chi-square distribution (Unsupervised_Anomaly_Detection, 2023).



(a) Generator Bearing and Shaft Axial MD (b) Generator Bearing and Shaft Radial MD (c) Rotor Bearing and Shaft Radial MD

Figure 5.15: WT21 Histogram of Anomaly scores for Mahalanobis Distance

The multivariate normality of the data is checked with a distance-distance (DD) plot and are plotted for the three cases in 5.15. A 45 degree reference line is given to check if the data set used follows a multivariate normal distribution. Since the data set does not cluster around this reference line, the data set does not follow a multivariate normal distribution. To solve this, the quantile of the distance values for the cumulative probability which is 1 minus the contamination factor used to calculate the new threshold (Unsupervised_Anomaly_Detection, 2023). This is done for all three cases, as the data points are not clustered along the reference line. The anomaly indicators are again obtained with the new threshold. DD plots with the new threshold values is plotted in 5.16. The new thresholds are much higher than the original threshold values, and the number of detected outliers is lower. The fraction of detected outliers is checked, for 5.16a, the outlier fraction is 0.2005, for 5.16b is 0.0990 and in 5.16c is 0.0987. This is the same as all the other models.



(a) Generator Bearing and Shaft Axial MD new (b) Generator Bearing and Shaft Radial MD new (c) Rotor Bearing and Shaft Radial MD new

Figure 5.16: WT21 Histogram of Anomaly scores for Mahalanobis Distance with new threshold

Chapter 6

Results

This chapter presents comparison scatter plots of the detected normal and outlier data points of the unsupervised fault detection models trained.

6.1 Unsupervised fault detection models

The four unsupervised fault detection models are compared using the top feature from two variables found from the Laplacian score feature ranking. The top feature of the rotor speed is used for the y-axis and the top feature for the vibration is used for the x-axis. The observation values of the normal and outlier points classified by each fault detection model is plotted together. This helps to visualize where the fault detection model classifies data points as outliers.

6.1.1 Generator bearing and shaft axial vibration

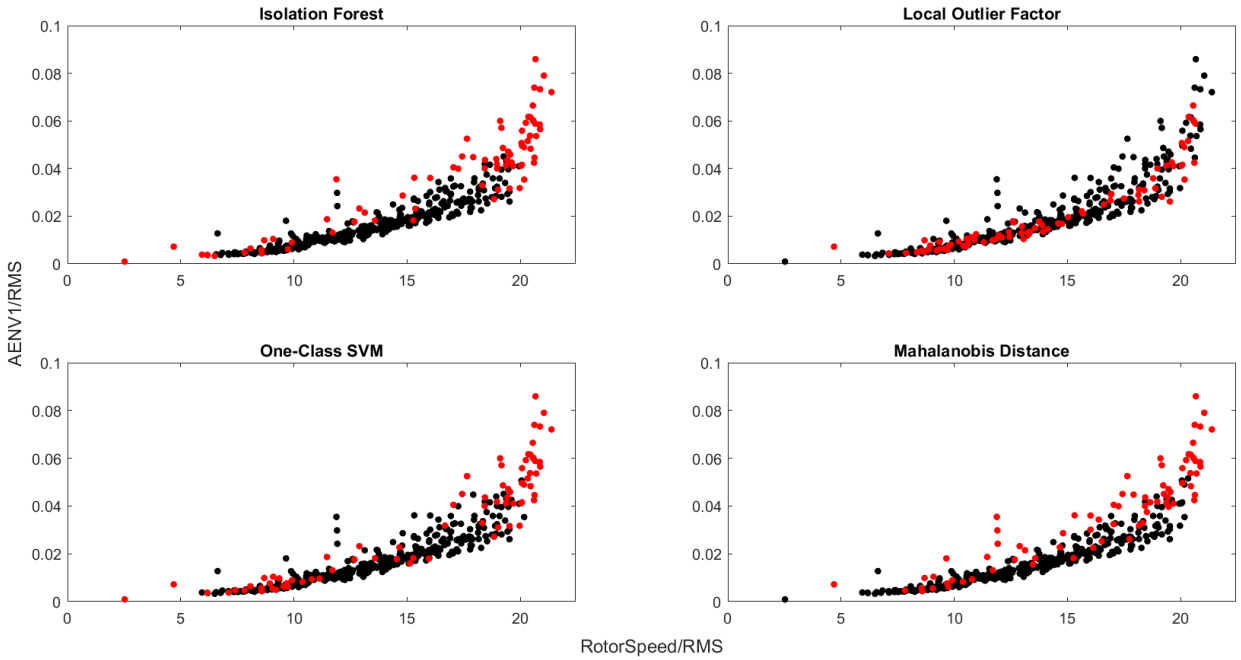


Figure 6.1: Generator Bearing and Shaft Axial fault detection models top variable features visual comparison

Here the y axis is the root mean square (RMS) of the generator bearing ENV 1 axial vibration. The measurement frequency is from 5 Hz to 100 Hz, which are the low frequency, high amplitude vibrations. The y-axis is the RMS of the rotor speed.

6.1.2 Generator bearing and shaft radial vibration

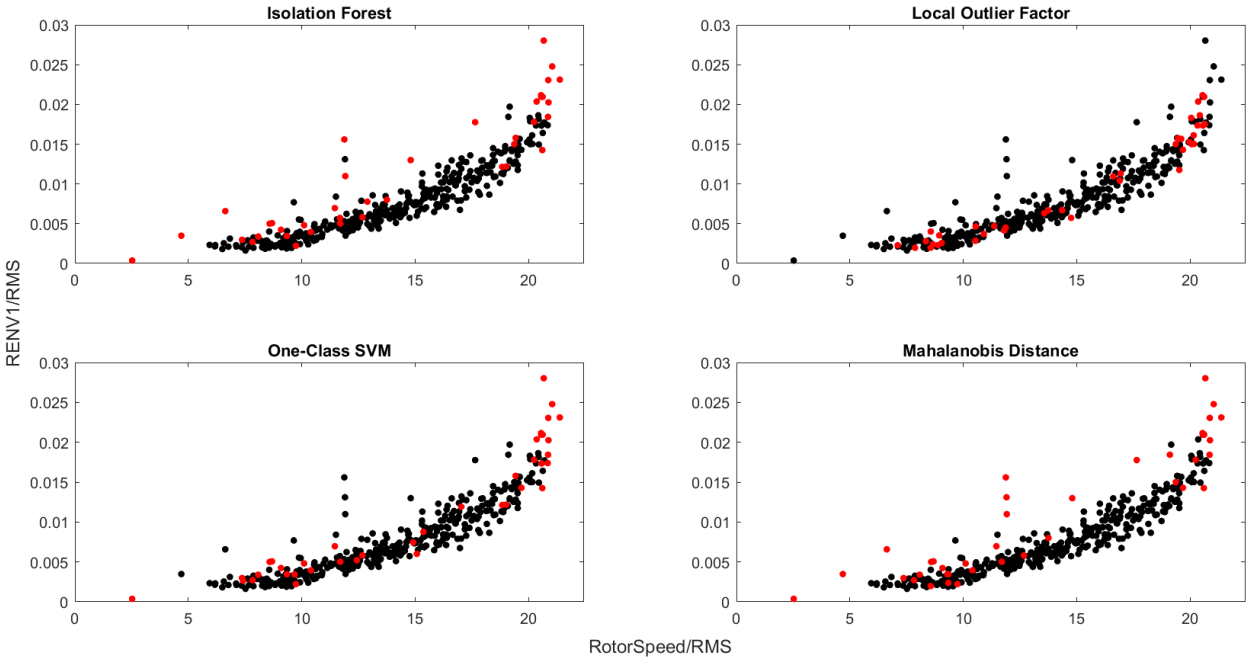


Figure 6.2: Generator Bearing and Shaft Radial fault detection models top variable features visual comparison

Here the y axis is the RMS of the generator bearing ENV 1 radial vibration. This is similar to the figure 6.1 where the top feature is RMS. The y-axis is also the RMS of the rotor speed.

6.1.3 Rotor bearing and shaft radial vibration

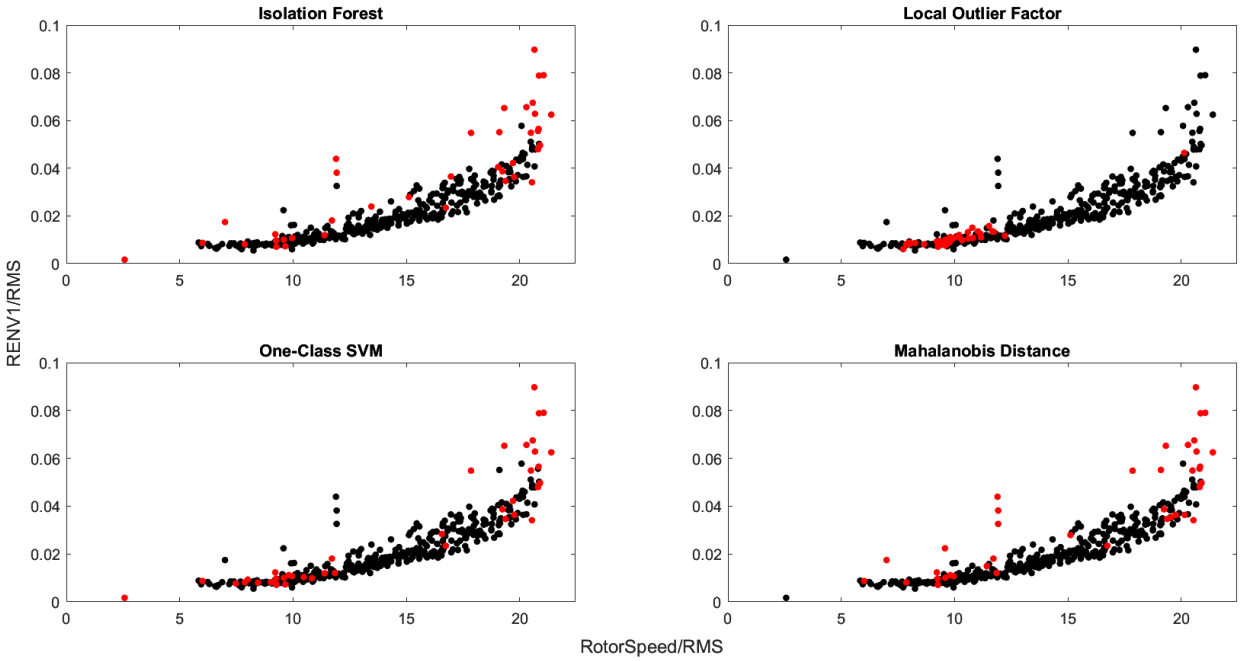


Figure 6.3: Rotor Bearing and Shaft Radial fault detection models top variable features visual comparison

Similar to figure 6.2, the y-axis is the RMS value and here it is for the rotor bearing ENV 1 radial vibration. The y-axis is also the RMS of the rotor speed.

Chapter 7

Discussion

This chapter presents an in-depth evaluation of the initial use of supervised fault detection techniques with the assumption of condition labels, the limitations with model training with low data size and accounting for seasonality aspects and the strengths and weakness of the unsupervised fault detection models used.

7.1 Early supervised fault detection model development

Since only vibration data with unlabeled bearing condition is available, it was first decided to visualize all the available data and look for trends, patterns and outliers. From this visual analysis, assumptions of the raw vibration data condition can be made for developing a fault detection model.

In the case of WT 4 generator and rotor bearing ENV 2 radial vibration respectively in plot 5.2e and 5.2f, there is a clear visible change in vibration amplitudes as previously discussed. From May to the end of December 2022, high vibration amplitudes crossing the alarm thresholds exist and from January 2023, the vibration amplitudes decrease significantly and stay generally low within the warning threshold. This distinct change in bearing vibration, supported the idea of using condition labels as 'before' and 'after' maintenance service. Here 'before' classifies the bearing has a fault present, while 'after' means the bearing is in normal operating condition after maintenance has been performed. Additionally, there are visible data gaps found in the vibration data at the end of December 2022 in WT 4, from possible maintenance service, which further supported the idea of using the condition labels.

The PdM fault detection model workflow in figure 3.1 (a) and the MATLAB diagnostic feature designer application workflow in 3.1 (b) was followed thoroughly. Relevant time signal features were extracted and ranked using supervised machine learning ranking methods such as T-Test. The top ranking features were exported to train a fault detection model. Vibration as the response variable and rotor speed as the predictor variable was used in the classification learner.

To avoid overfitting, cross validation with 5 fold and 20% of data was kept our for testing. Various supervised machine learning techniques such as SVM, neural networks, decision trees, Nearest neighbors (kNN) and Ensemble Classification methods were trained. An optimal decision tree fault detection model with a test accuracy of roughly 85%, and only minor type I and type II misclassification errors was trained.

After guidance with an SKF bearing engineer about vibration data in the SKF @ptitude observer application, it was only found lately that the distinct change from the WT 4 generator and rotor bearing ENV 2 vibration amplitudes was solely caused by changes in the sensor measurement settings. The unit of measurement was changed from gE PtP, to m/s^2E PtP, and then back to gE PtP. The change in measurement unit caused the vibration levels to be 9.81 times higher. The correct sensor measurement settings as recommended from SKF @ptitude observer is gE PtP, as used for the other measurement envelopes. This meant the raw vibration data of the WT 4, generator and rotor bearing ENV 2 radial vibration had to be excluded from the fault detection model development. The vibration measured does not reflect the condition of the bearings and is solely related to change in sensor measurement settings. As a result of this finding, the work on using supervised machine learning methods with bearing condition label assumptions had to be disregarded, since the work is no longer valid.

After guidance with Thibaut Forest from REN Equinor and Rohit Agrawal from MathWorks on the WT data I have access to, it was recommended to use unsupervised machine learning techniques such as cluster analysis and anomaly detection models for detecting outliers. This allowed for implementing suitable fault detection model and the following methods were therefore used; isolation forest, local outlier factor, one-class support vector machine (SVM) and Mahalanobis distance.

7.2 Model training

7.2.1 Seasonality aspects

Ideally, the fault detection model should be trained on the full year worth of data, to account and capture the seasonality aspects. Winter months often have much higher wind speeds than when compared with the summer months. Due to this, the rotor speed of the WTs are greater. Hence, the seasonality aspect needs to be captured between rotor speed and vibration. For example, a fault detection model should be trained on one whole year worth of data and then be tested on the next year, to identify any faults.

In our case, there is lack of vibration and rotor speed sensor data for the first three months of 2022, meaning the model can be only trained from April to December for 2022. Due to this, we will not be able to capture the seasonality aspects to the full extent, and this will be unfair

for our model, mainly due to different environmental conditions. Since, we have data for 2023 from January to April 2023, we can compensate for this, by using the data from 2023. However, it is a caveat to the model and the results, the validity of the model is affected.

If more data is available of the WTs at the Bessakerfjellet, it will be possible to identify the seasonality aspects and its influence in the data measured. This is crucial for modelling an optimal and robust fault detection model, by taking into account the effect of seasonality on the data. This will allow the fault detection model to more precisely detect faults. With the vibration data available, it is possible to observe signs of increasing vibration from higher rotor speeds in the winter months. More vibration data, will help verify if this is the case and will help with training our model.

7.2.2 Resolution of data

A major limitation of the data available is the sampling rate is varied for the different envelopes of the axial and radial vibration. The sample rate of the generator and rotor bearing, and shaft axial and radial vibration is different in many instances throughout the time period from April 2022 to April 2023. The highest sample rate commonly observed per day is 72 measurements while the lowest is 24 measurements per day. This is already evident in the first 9 days of the ensemble datastore for WT 21 in figure 5.7. This results in the table row size is not being the same when comparing many of the different measurement frequency envelopes. Due to this, it is not possible to contain all the different generator bearing raw axial vibration in different envelopes in one table for example, to use for performing PCA and implementing unsupervised fault detection modes.

A solution to overcome this, is to use some statistical metric calculation of the data per day of the different variables for the whole year. Using this method allows for getting the same table row size of all the variables. First, it was decided to calculate the mean of the data per day of the different variables, however, it is difficult to access if it can capture the trends and patterns in the data thoroughly. Hence, six statistical metrics using the DFD was extracted for all the variables such as rotor speed and the different envelopes axial and radial vibration. These metrics are mean, root mean square (RMS), standard deviation, shape factor, kurtosis and skewness. It is not immediately clear which features best explain the available data and the unsupervised machine learning ranking method Laplacian score is used to select the top features.

The resolution of the data is greatly reduced to only one data value per day. This results in the loss of information and all the variations, trends and patterns in the data can not be captured thoroughly. Most of the outliers will in addition be lost. Hence, the fault detection models trained will not be robust enough take into account the variations in data and detect outliers more precisely. This low resolution of data used is a major limitation and the sensor measurement settings in the SKF @ptitude observer needs to be standardized to be similar to all the

different envelopes.

7.3 Fault detection models

7.3.1 Principal component analysis

The PCA fault detection technique as found in the MATLAB tutorial (Filion, 2019) did not perform as optimal in this case. The top 15 ranking features from all the variables are used to perform PCA, however, in all three cases the cumulative variances of the first and second principal components are below 80 % and they are not able to capture a large amount of variance in the data. It is therefore not suitable to use the first two principal components scores in a scatter plot, as a robust fault detection technique using warning and alarm level decision criteria in this case. It is however expected the cumulative variance of the first and second principal components to be much higher, if only the raw vibration data of all the different envelopes and acceleration of the shaft is used for performing PCA. The only major limitation here is the sample rate difference found in many cases of the different variables, resulting in different table row sizes.

Performing PCA is a great technique for reducing the data dimensionality size of the bearing radial and axial vibration in combination with the rotor speed. The most important and essential data is retained and is analyzed. However, a limitation of the PCA performed, is that it is a linear technique, meanwhile real-world processes are most generally nonlinear as seen evidently from the correlation analysis of rotor speed and vibration data. This reduces the effectiveness of PCA for using it as a fault detection technique. Hence, more advanced techniques than PCA are required such as the unsupervised fault detection models.

7.3.2 Unsupervised fault detection models

The four unsupervised fault detection models trained for the three different cases are presented with a visual comparison of the normal and outliers data points in figure 6.1, 6.2 and 6.3. Commonly for all three figures, the RMS of the ENV 1 axial and radial vibration is the top vibration feature from the Laplacian score feature ranking and is used for the y-axis. With the RMS of the rotor speed on the y-axis, it allows for comparing the outlier classification of the different models more easily.

Emphasis is placed on if the fault detection models are able to detect the machine resonance in the generator and rotor bearing identified from the WT 21 correlation analysis in chapter 5. At rotor speeds from 11.5 to 12 CPM, there is increased vibration amplitudes showing a clear spike in the ENV 1 vibration levels. The machine resonance is easily identifiable in plots 5.6d, 5.6e and 5.6f. Since, the model comparison figures in 6.1, 6.2 and 6.3 use the same variables, we can evaluate the performance of the models through the detection of the machine resonance

and the overall outliers from the general cluster in the data. An important observation when comparing the ENV 1 vibration plots from the correlation study and in the fault detection model comparisons is the majority of noise and general outliers in the data are removed through the data cleaning process. The decreased resolution in the data used for training the models is also visible from the fault detection model comparisons.

Detection of machine resonance

In figure 6.1, there are only three data points showing the machine resonance. This is not ideal and is mainly caused by the low resolution of data used. This issue can be resolved if the sample rate of all the different measurement frequency envelopes is the same in SKF @ptitude observer. This is a particularly weak point and greatly influences the performance of the model in detecting outliers in the data. From figures 6.1, 6.2 and 6.3, it is only the Mahalanobis distance and isolation forest models which are able to detect the machine resonance in the data. The local outlier factor and one class SVM models however are not able to detect the resonance in all three cases at all. The Mahalanobis distance model detects all the data points in the vibration spike as outliers, then the isolation forest model, where either 1 or 2 data point are not detected. Hence, from this initial analysis, the Mahalanobis distance model performs best in detecting the machine resonance found from the generator and rotor bearing and shaft axial and radial vibration. Higher resolution in the data can however help confirm the performance of the models more accurately.

General outliers in the data

The general outliers here are classified as data points deviating and located far from the cluster of the data points. From figures 6.1, 6.2 and 6.3, the local outlier factor model is not able to detect any of the data points deviating from the general clusters as outliers. The outliers detected are mainly within the clusters of the data. This is due to fundamentally how the model works. This observation is similar to what is found in the research paper by Wang et al. (Wang et al., 2022). The model compares the density of the local neighboring data points with the density around the data points. This greatly limits the performance of the model with outliers located far from the general cluster, since only local anomalies are detected. As Wang et al. (Wang et al., 2022) suggest, distance-based methods such as the Mahalanobis distance are more effective in these cases.

The isolation forest, one class SVM and Mahalanobis distance models perform significantly better than the local outlier factor model in detecting the outliers far from the cluster of data, as in the cases where rotor speed is highest at above 20 CPM. Here the outliers have high vibration amplitude and are spread out at the highest rotor speeds. Greater axial and radial vibration

are expected when rotor speeds are high, however, the isolation forest, one class SVM and Mahalanobis distance models comparison plots confirm that they are able to detect these high vibration amplitude data points at the higher rotor speeds as outliers. Hence, when more data is available of the generator and rotor bearing of when they become degraded and produce greater vibration signal amplitudes at higher rotor speeds, it is expected these three models to be able to detect these outliers.

The three models which are the isolation forest, one class SVM and Mahalanobis distance, however, do detect a few data points located within the cluster of data as outliers. This is visible in all three figures 6.1, 6.2 and 6.3. These false alarm detections of outliers range from a rotor speed from 6 and up to 20 CPM. These outlier detections within the data is challenging to explain and is not ideal for a fault detection model. The best performance of these three models needs to be judged from the detection of general outliers away from the main data clusters and the local false outliers detected within the cluster. With these criteria, it is the isolation forest and Mahalanobis distance which are able to achieve this in comparison to one class SVM, where a greater portion of outliers within the data cluster exist. This is clearly evident in figures 6.1 of the generator bearing and in figure 6.3 of the rotor bearing. Tutiven et al. (Tutivén et al., 2022) applied one class SVM for early fault diagnosis of the WT generator bearings based on only normal condition SCADA data such as wind speed, ambient temperature and the main bearing temperature. Their results however showed very few false alarms, and this is not the case here, where a high amount of misclassification of outliers within the cluster of data exist.

When comparing the isolation forest and Mahalanobis distance models, both function differently in classifying outliers from normal data, however in this case the detected outliers are nearly identical, and it is difficult to pick the best performing model. When emphasis is placed on detecting early signs of resonance, which can damage WT components significantly, then the clear choice is the Mahalanobis distance model. Both resonance and outliers far from the data cluster is detected, while a few falsely detected outliers exist. This model provides greater insight from the detected outliers than the warning and alarm thresholds used in SKF @ptitude observer.

Contamination factor

The use of contamination factor on training the fault detection models can significantly influence the number of detected outliers in the data. Additionally, the number of false alarms and non-detection of outliers due to faults is influenced. A high contamination factor will give higher false alarms, when the condition of the bearing is normal, however the fault detection model detects a fault is present. A lower contamination factor will mean, the fault detection model will be trained with faulty data treated as normal. This leads to higher non-detection of possible faults in the bearings which may be present. The number of false alarms and non-detection of bearing

faults should be ideally low.

Access to the Bessakerfjellet wind farm maintenance logs and other supplementary information such as results from bearing oil analysis, can help in making a more informed decision of the choice of contamination factor for unsupervised fault detection models. The percentage of abnormality and outliers in the data due to possible faults can be more accurately decided. Additionally, confirmation of the bearing condition allows for the possibility of using supervised fault detection methods and the use of a contamination factor can be avoided. If there have been bearing faults which have occurred and have been repaired with preventive maintenance for example, supervised fault detection methods such as SVM, neural networks, decision trees and nearest neighbors (kNN) can be utilized. A challenge here is if the vibration data can be accurately labeled with condition labels such as normal and bearing fault types, as the classification models will be trained using this information. Consequently, the performance of the model is dependent on the quality of the data used.

A whole year worth of unlabeled generator and rotor bearing vibration data is available from the Bessakerfjellet, and it is clear the vibration amplitudes are higher in the winter months. However, it is difficult to determine if the higher vibration of the generator and rotor bearing are solely caused by the increased rotor speeds of the WT. It is highly likely some other bearing defect may have developed, as found in the case of WT 21 generator bearing, where the frequency domain analysis showed signs of beginning bearing damage in the inner race and cage of the bearing. From this finding, it is however not possible to determine how long the signs of these bearing faults have existed and if any minor damage in the bearing has occurred. This is a particularly weak point, without the access to the wind farm maintenance logs.

The contamination factor used for training the fault detection models were solely based on the raw vibration analysis of WT 21 in 5.3. Vibration signals which had crossed the alarm thresholds are considered to be abnormal and regarded as outliers, while signals below the alarm threshold is treated as normal. The major limitation of this method is in the uncertainty regarding the choice of value for the warning and alarm thresholds. These threshold values were extracted from SKF @ptitude observer along with the vibration and rotor speed data. From the raw vibration analysis, there are many instances where the warning and alarm thresholds were increased or decreased, and the reasons for this is unclear. Hence, there is underlying uncertainties with these thresholds values and the choice should be standardized for all WTs. Since, the threshold values differ for WT 4 and 21 in a few instances such as the rotor bearing ENV 1 radial vibration. Guidance from Rohit Agrawal from MathWorks was required to set a reasonable contamination factor for the WT 21 generator and rotor bearing and shaft vibration data. The fault detection models trained such as the isolation forest and Mahalonobis distance using this contamination factor, were able to detect the resonance and the high vibration amplitudes at high rotor speeds as outliers. The isolation forest model has a few non-detection errors of

the resonance data points. However, both models had false alarm detection of outliers which where generally within the cluster of data as seen in figure 6.1, 6.2 and 6.3. The contamination factors used are reasonable from this initial analysis, and when more vibration data is available, the values can be adjusted to reduce the false alarm detections and non-detection of possible bearing faults.

Chapter 8

Conclusions, Discussion, and Recommendations for Further Work

This final chapter presents a through summary of if the thesis objectives have been met, the key findings from the thesis work and the applications it has for early fault detection in WT main shaft bearings in the wind industry, and recommendations for extension to work carried out in this theses are given.

8.1 Summary and Conclusions

The main focus of this mater thesis was to investigate and compare different early fault detection techniques for the use in WT generator and rotor bearings using a relevant case study. Six objectives were presented. The first objective's aim was educating the reader and giving a broader context of WTs. This is met in Chapter 2, which presents the state-of-the-art and industrial background of WTs, how they function, WT structures, offshore and onshore installation capacities and available condition monitoring technologies. The major, minor faults and failures of WTs are presented.

The second objective's aim was to provide a through theoretical background for the reader to follow the work carried out in the thesis, especially the case study. This objective is met in Chapter 3, where theory about fault diagnostics, fault detection model development workflow, four unsupervised fault detection models and maintenance strategies for WTs and of which are used in the industry. Here the maintenance practices and key priorities for offshore wind farms is presented.

The third objective's aim was to educate and provide the reader with knowledge of industrial roller bearing, which is a critical WT component and is relevant to the thesis topic. They are required for the smooth rotation of the rotor blades and shaft. The third objective is met in Chapter 3, where knowledge of how bearings function, components of a vibration signal, sen-

sor measurement positions in industrial machinery, vibration analysis methods used in fault diagnostics are presented. The four distinct bearing fault frequencies, and how the faults and failures in bearings develop in four stages are provided. This extensive knowledge of bearings is required to support the reader in understanding the thesis work and analysis carried out in Chapter 5.

The fourth objective's aim was to thoroughly examine the WT SCADA, and or CMS data acquired, before implementing the fault detection models. This objective is met in Chapter 5 case study, where the raw axial and radial vibration data in different envelope frequencies of the Bessakerfjellet WT 4 and 21 generator and rotor bearing and shaft is presented. The data is investigated first with statistical visualization analysis by plotting the available vibration data with the warning and alarm thresholds found in SKF @ptitude observer. The next analysis method was the frequency domain vibration analysis in SKF @ptitude observer using the 'spectra' tool of the generator and rotor bearings. Lastly the correlation analysis between the WT rotor speed and vibration data is performed. From all three analysis methods, a through investigation was conducted to identify and uncover the trends, outliers and patterns in the acquired data. Three outcomes from the analysis respectively are identifying the seasonality aspects in vibration levels, observation of developing defect in the inner race and cage of the WT 21 generator bearing and discovery of machine resonance occurring at a particular rotor speed in both WTs.

The fifth objective's aim was to use suitable fault detection methods to detect outliers and abnormal data using the acquired WT SCADA, and or CMS data. Greater emphasis on offshore wind turbine early fault detection of bearing minor failures was placed in Chapter 1, however, it proved to be challenging to get access to relevant offshore wind turbines SCADA and CMS data, along with maintenance logs to determine the condition of the WT components. Strict offshore wind farm historical and operational data sharing procedures resulted in extensive data access delay. To compensate for this, vibration data from generator and rotor bearing from onshore wind turbines at the Bessakerfjellet was used. The unlabeled bearing condition data, and unavailability of the wind farm maintenance log, allowed for only implementing unsupervised fault detection models. Fault diagnosis could not be implemented with the model, since the condition of the bearing could not be accurately labeled in the acquired data. Therefore, unsupervised fault detection is used, where the models are trained with parameters such as the contamination factor to detect outliers at the specified fraction in the data. The root cause of the outliers is highly difficult to be determined just from the fault detection. Hence, this objective is partially met in Chapter 5, where only fault detection could be achieved, and diagnosis of the detected fault could not be achieved. This is primarily due to the lack of labeled WT bearing condition data such as normal and faulty types.

The sixth objective's aim was to compare the performance of the fault detection models used. The results from objective 4 analysis of the acquired data performed in Chapter 5 is uti-

lized for judging the performance of the fault detection models trained. Since the fault detection models were trained using the contaminated data, where outliers are present as found from the vibration analysis, only a visual analysis of the model performance was possible. This was hence carried out using scatter plots of the normal and detected outliers of the models, where the top-ranking time signal statistical metric for the rotor speed and vibration is used. The fault detection model normal and outlier comparison scatter plots for the three cases which are the generator bearing and shaft axial vibration, generator bearing and shaft radial vibration, and the rotor bearing and shaft radial vibration are presented in Chapter 6. The performance of the models outlier detection are discussed thoroughly in Chapter 7.

8.2 Discussion

The minor failures of offshore WT components such as main shaft bearings causing longer production downtime than for onshore WTs, require further rigorous research in academia for early fault detection, diagnostics and prognostics. This will improve all aspects of reliability, availability, maintainability and safety for offshore WT components, ensuring a sustainable energy transition, greater production of low-cost renewable energy and reduction in environmental impacts from optimized 'just in time' maintenance strategy. Onshore WT minor failure of sub-assemblies are less complex to perform corrective maintenance than for offshore WTs and effective maintenance planning is the key to achieving a lower OPEX. Early fault detection supports the decision-making process for wind farm asset managers, by providing indications of developing faults and performing timely maintenance action. The maintenance scheduling can be optimized for safe execution at offshore wind farms.

As found from the specialization project (Thiruthiyappan, 2022), fault diagnostics and prognostics of major failures in WT sub-assemblies have been a key research area, while minor failures are studied less frequently. To address these issues, the thesis focused on early fault detection of minor failures in offshore WTs. However, CMS data from onshore WTs had to be used due to extensive lack of access to offshore WT SCADA, and CMS data. The acquired vibration data of the generator and rotor bearings on onshore WTs in the Bessakerfjellet is used for training fault detection models to provide the wind farm asset managers with early indication of bearing faults. Unlabeled vibration data and lack of wind farm maintenance logs allowed for implementing only unsupervised fault detection models. These were isolation forest, local outlier factor, one class SVM and Mahalanobis distance.

The detected outlier result of the fault detection models from the case study, showed strong outlier detection performance from the isolation forest and Mahalanobis distance, with minor false detections within the normal vibration data cluster. The one class SVM and local outlier factor had limitations in detecting the outliers found in the vibration analysis. Emphasis was

placed on detecting the machine resonance, outlier data points of high vibration amplitude at higher rotor speeds and general deviation from the data clusters. From these criteria, it was evaluated the Mahalanobis distance-based method performed robustly in detecting outliers in all three cases. The Mahalanobis distance measures the distance from a data point to the distribution of the data points and uses the covariance matrix of the data set. This initial analysis with unlabeled WT CMS data shows the unsupervised fault detection method Mahalanobis distance, is capable of detecting early abnormal conditions in onshore WT generator and rotor bearings. Hence, provides strong implications for the use in onshore wind industry. The original intent of the thesis of using offshore WT SCADA, CMS data was not achieved due to data access issues. Nevertheless, the author believes the findings from the unsupervised fault detection models performance comparison from using onshore WT CMS data, useful applications in the offshore wind industry, for detecting early abnormal conditions in offshore WT bearings.

An unexpected and important finding from the case study is the signs of resonance from the correlation analysis of vibration and rotor speed. The result is unforeseen and is present in both WTs. This finding has strong implications for the wind industry. Resonance is capable of doing extensive damage and can severely reduce the useful life of critical WT components. Early detection of harmful machine resonance is vital in the wind industry, for taking the necessary mitigation steps to prolong the assets lifetime.

A major limitation of the study is the low resolution of data used for training the fault detection models. One data value per day of the time period of data acquired is used. The sample rate variations from SKF @ptitude observer, showed some vibration envelopes had much higher sample rate than other envelopes, giving large differences in the table row size of the variables in the vibration data. Due to this, statistical metric calculations of the data per day of the different variables had to be performed to have the data in one table with the same row size. This greatly influences the fault detection model training, and it is highly possible, the other models such as isolation forest, one class SVM and local outlier factor may show promising results than the Mahalanobis distance, if only raw data was used. The findings from this initial analysis can be utilized and implemented by the Bessakerfjellet wind farm asset managers for early fault detection, using the similar data resolution used to train the models. However, using greater resolution of the vibration data, may likely result in the model giving higher false alarm and non-detections of possible bearing faults, hence, the model should be retrained.

A minor limitation in the unsupervised fault detection model training is the assumption of contamination factor. There are no set guidelines choosing such contamination factor for any given case and it was chosen based on best judgement and along with guidance from an external supervisor. The choice of the value has influence in the performance of the fault detection models and is therefore a limitation in using the results.

The findings from this thesis can be employed in other domains, where unlabeled condition

data of industrial roller bearings are available. The unsupervised fault detection model Mahalanobis distance can be trained for early detection of abnormal conditions. The performance of the fault detection should be evaluated, and the necessary parameters of the model should be adjusted to optimize the model performance. Relevant domains include the automotive, power generation, marine and offshore, as well as the aerospace industries.

8.3 Recommendations for Further Work

- The sample rate of the vibration measurement data collection for the generator and rotor bearings in SKF @ptitude observer needs to be standardized for all WTs. This will give the same data table dimensions of the 3 vibration measurement envelopes. The collected data needs to be quality checked for measurement accuracy.
- The PCA is expected to have higher cumulative variance in the first and second principal components, which can be used for fault detection with decision criteria. This needs further investigation and is expected to perform better than the individual warning and alarm thresholds which are set in the SKF @ptitude observer.
- The unsupervised fault detection models are expected to yield more robust outlier detection performance from training with higher resolution raw vibration data. The models can be evaluated more thoroughly based on the false and non-detections through visual analysis. This is allow for choosing the most optimal fault detection model.
- Request access to the wind farm maintenance logs and other relevant information to accurately determine the condition of the bearings. This will allow for using supervised machine learning techniques with condition labels for fault detection. These methods are expected to more accurately detect faults of the bearings and requires further investigation.
- Use more CMS data from a higher number of WTs at the Bessakerfjellet to train and evaluate the fault and abnormal condition detection of the unsupervised fault detection methods. This will allow for finding the most optimal model.
- Investigate and mitigate the effects of resonance found in the Bessakerfjellet WT 4 and 21 rotor and generator bearings. Further investigation is required to identify if the resonance is caused solely by the shaft rotation at a particular rotor speed from 11.5 to 12 CPM.
- Get access to offshore WT SCADA, and or CMS data of the generator and rotor bearing, to investigate if the findings from this thesis of using unsupervised early fault detection

method with onshore WT CMS data is applicable to the offshore WTs. Compare the fault detection performance of the different models.

Appendix A

Acronyms

CBM Condition based maintenance

CM Corrective maintenance

CMS Condition monitoring system

DFD Diagnostic feature designer

FFT Fast Fourier Transform

FOW Floating offshore wind

LOF Local outlier factor

OPEX Operational expenditure

PCA Principal component analysis

PdM Predictive maintenance

PM Preventive mainteance

RAMS Reliability, availability, maintainability, and safety

RUL Remaining useful life

SCADA Supervisory control and data acquisition

SVM Support vector machine

TBM Time-based maintenance

WT Wind turbines

Appendix B

Supplementary information & MATLAB code

B.1 Chapter 3

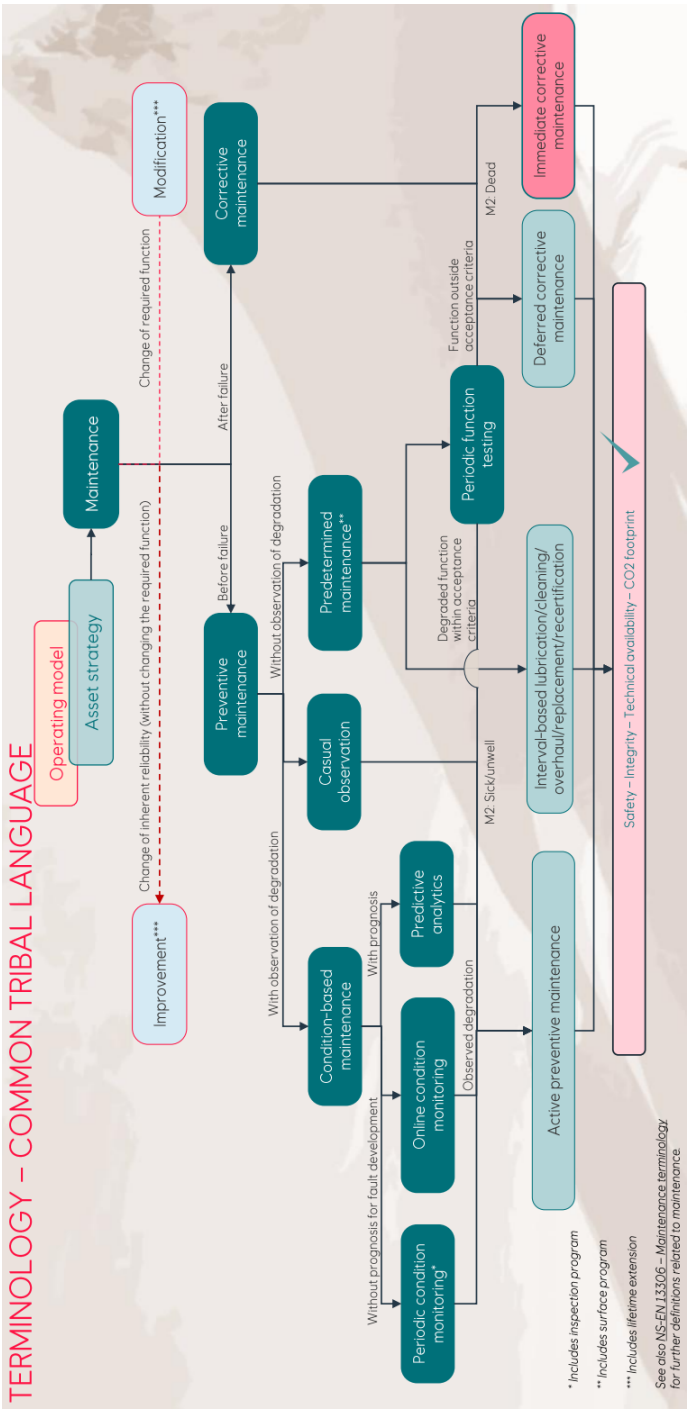


Figure B.1: Overview of Equinor's failure management strategies (Svennevig, 2022)

B.2 Chapter 4

Table B.1: Data acquisition process timeline for wind farm SCADA & CMS data

Date	Company	Progress
27th March 2023	Equinor	Requested access for Equinor operated offshore wind farms such as Sheringham Shoal (SHS), Hywind Scotland (HYS) and Dudgeon (DOW)
13th April 2023	Equinor	Due to delay with wind farm SCADA data acquisition from Equinor, it was suggested by my co-supervisor Thor Inge Bernhardsen to use vibration data from NTNU's SKF @ptitude observer, in the case access to Equinor SCADA data would take more time.
19th April 2023	Aneo (TrønderEnergi)	Approval received to use generator and rotor bearing vibration and rotor speed data from SKF @ptitude observer.
20th April 2023	Equinor	Meeting with SCADA data owner and risk assessment performer together with co-supervisor to discuss about access to Equinor wind farm SCADA data and eventually anonymised and tweaked data. It was informed that it may take a few more weeks to access anonymised data and hence we reached the decision to use vibration data from onshore wind turbines at the Bessakerfjellet, stored in SKF @ptitude observer.
25th April 2023	Aneo (TrønderEnergi)	Signed thesis collaboration and confidentiality agreement and access given to Bessakerfjellet onshore wind turbines SCADA data such as active power, nacelle direction, turbulence, wind direction and wind speed.
3rd May 2023	Aneo (TrønderEnergi)	Access to wind turbine maintenance log data not approved due to lack of resources and the already ongoing collaboration with Equinor.

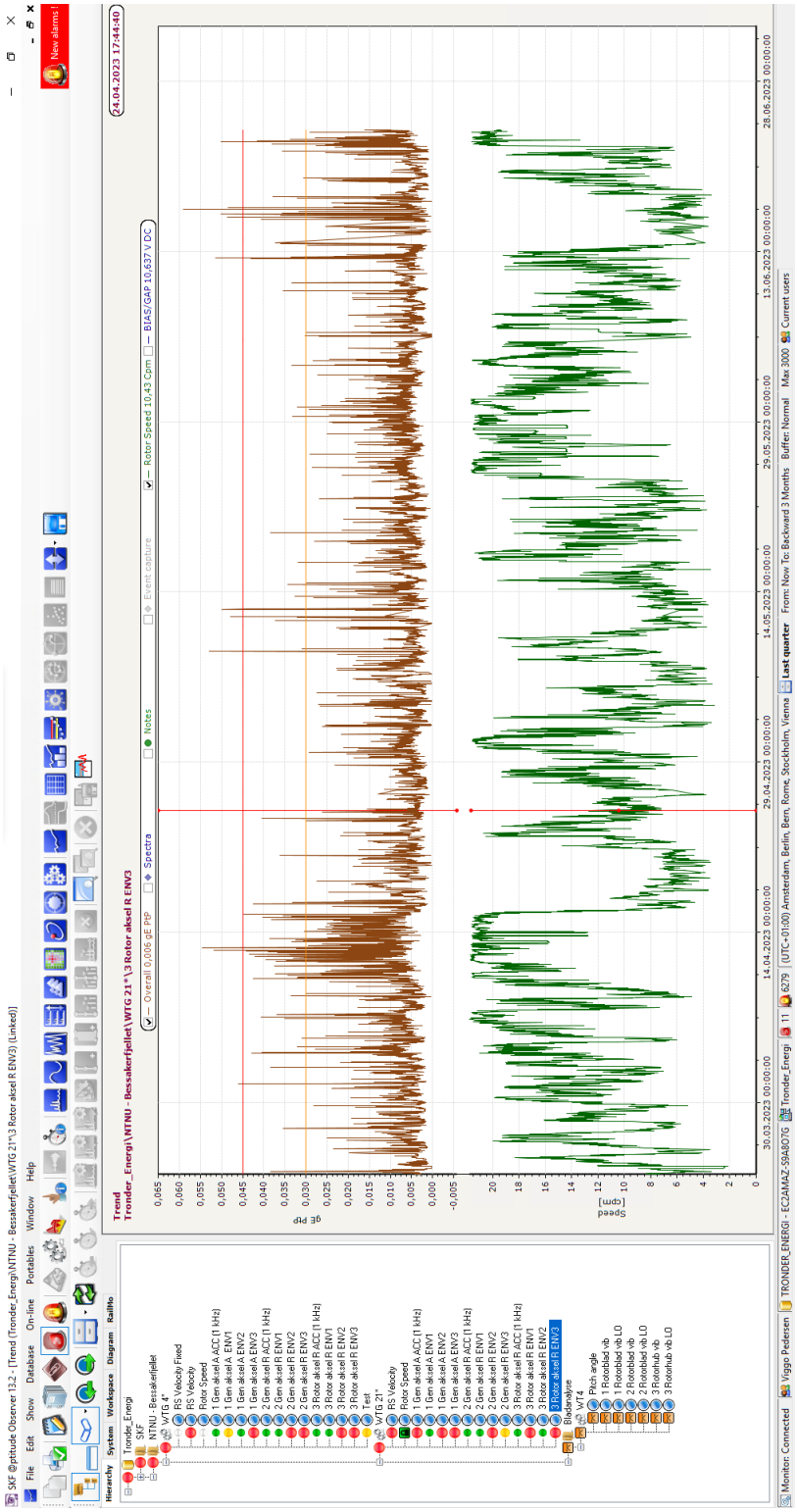


Figure B.2: SKF @ptitude observer program

B.3 Chapter 5

WT 4

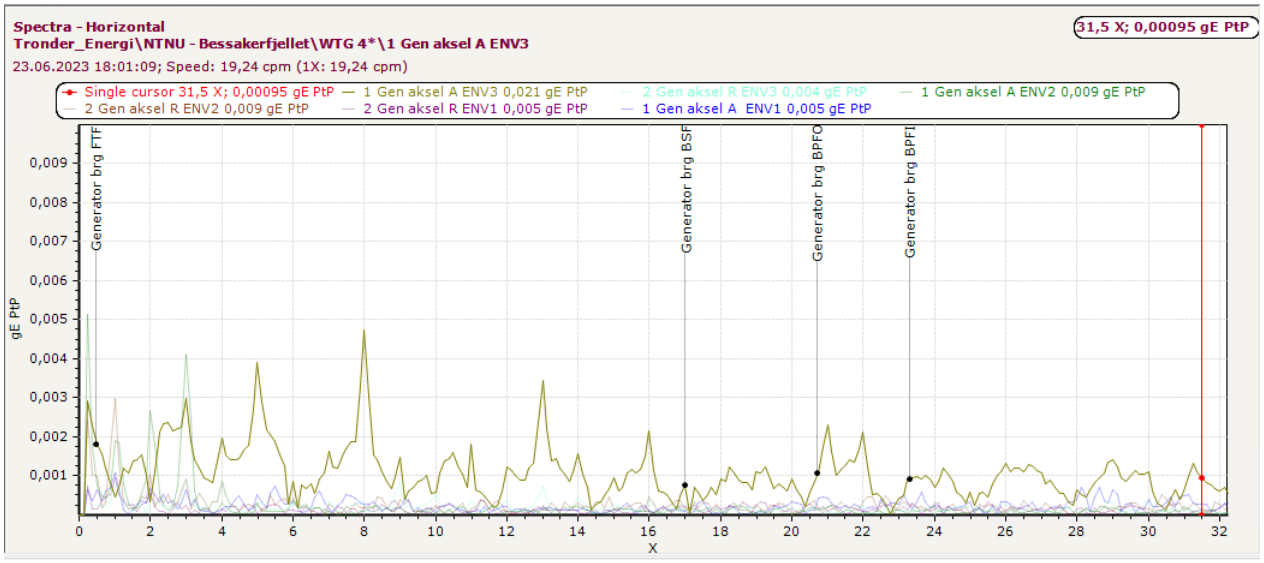


Figure B.3: WT 4 Generator Bearing Fault Frequencies

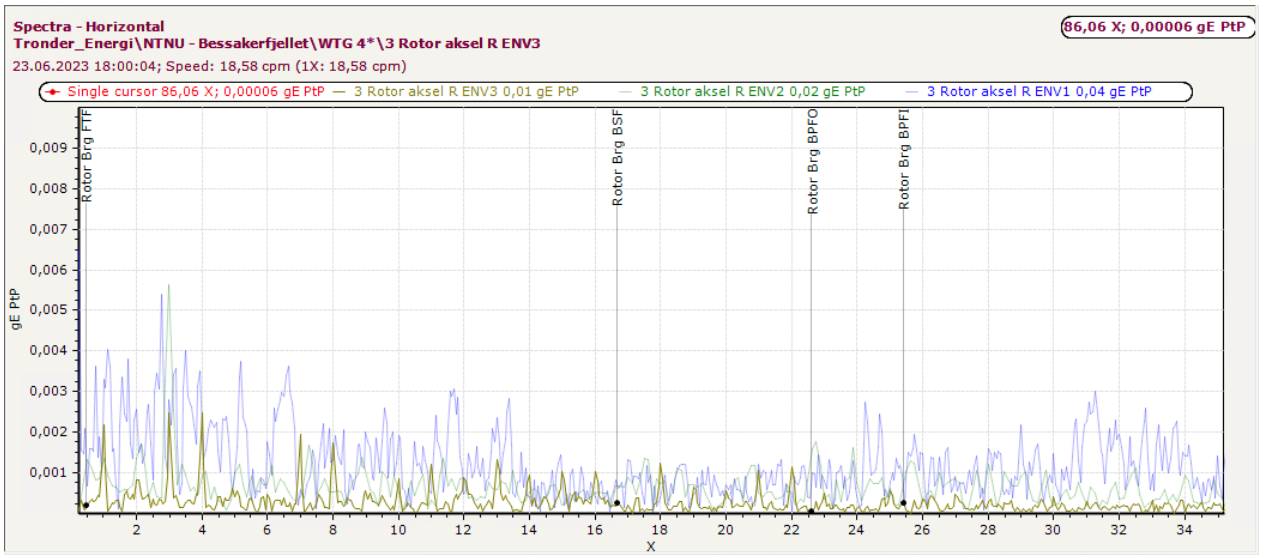


Figure B.4: WT 4 Rotor Bearing Fault Frequencies

WT 21

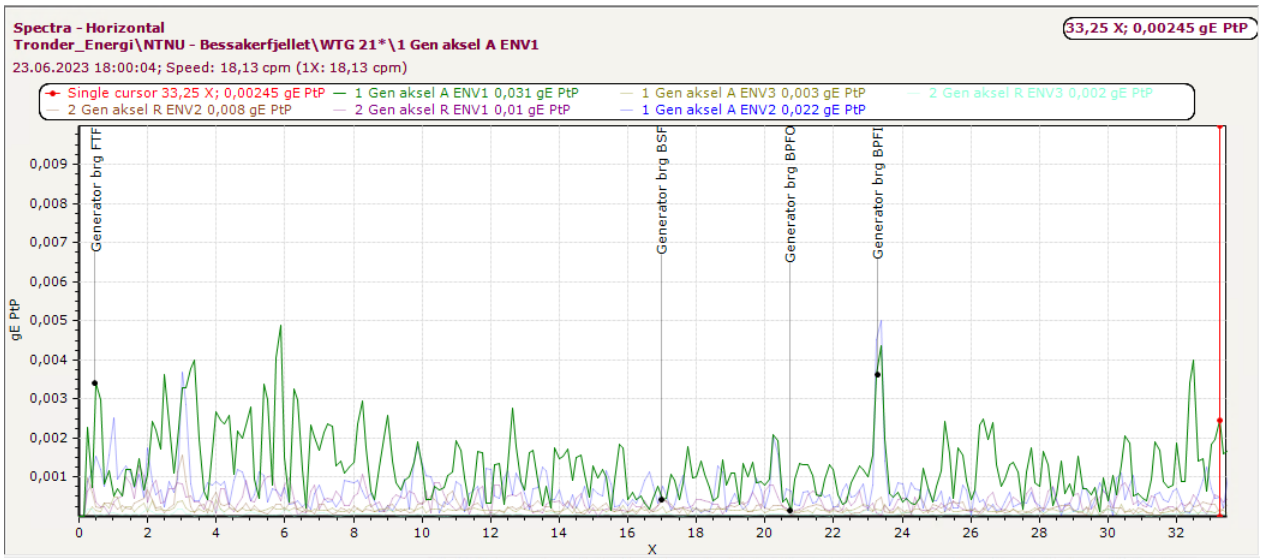


Figure B.5: WT 21 Generator Bearing Fault Frequencies

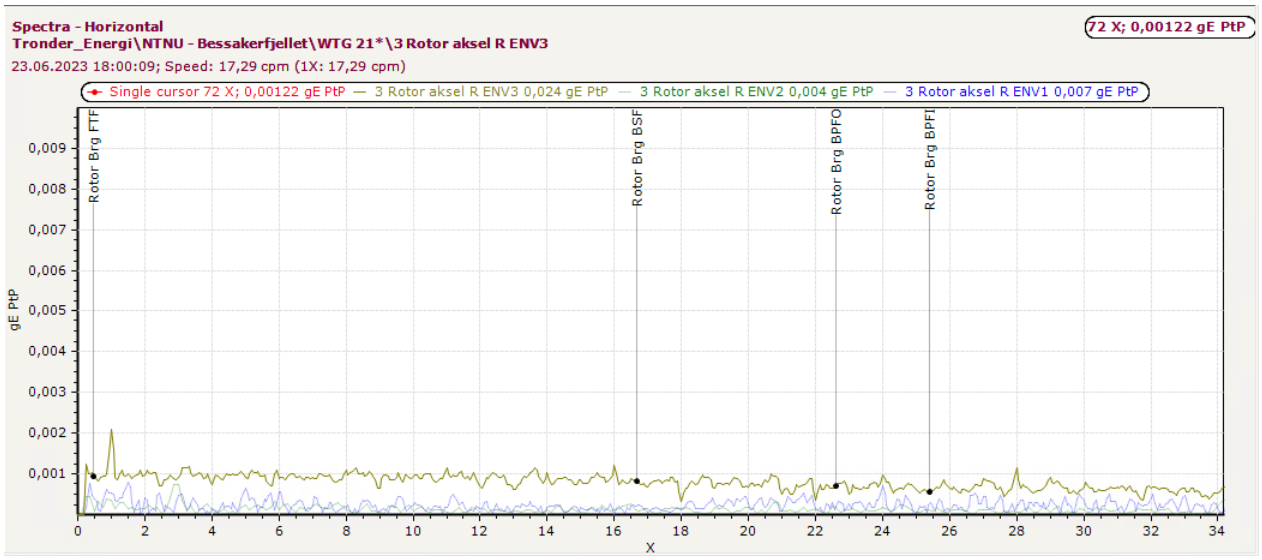


Figure B.6: WT 21 Rotor Bearing Fault Frequencies

WTG 21 - Generator Aksel Axial & Radial ENV 1, 2, 3 and ACC of shaft

Table of Contents

Data Cleaning

Vibration and ACC of shaft data cleaned using the MATLAB Data Cleaner. Missing data fields cleaned by filling missing values with linear interpolation. Cleaned outlier data. Smoothed data with moving mean method with a smoothing factor of 0.05.

To do:

- Samples within envelopes where at least one input or output is missing (Checked when exporting data from SKF Cloud Solutions, where the vibration and Rotor speed are measured with same frequency)

Functions for data cleaning

```
% % Vibration
% function WT41GenAkselAENV1LastYear = VibrationDataClean(WT41GenAkselAENV1LastYear)
%       % Fill missing data
%       WT41GenAkselAENV1LastYear = fillmissing(WT41GenAkselAENV1LastYear,"linear",...
%           "DataVariables",["Vibration"]);
%       % Remove outliers
%       WT41GenAkselAENV1LastYear = rmoutliers(WT41GenAkselAENV1LastYear,"movmedian",24,...
%           "DataVariables",["Vibration"]);
%       % Smooth input data
%       WT41GenAkselAENV1LastYear = smoothdata(WT41GenAkselAENV1LastYear,...
%           "movmean","SmoothingFactor",0.05,"DataVariables",["Vibration"]);
% end
%
% % ACC
% function WT41GenAkselAACC1kHzLastYear = ACCDataClean(WT41GenAkselAACC1kHzLastYear)
%       % Fill missing data
%       WT41GenAkselAACC1kHzLastYear = fillmissing(WT41GenAkselAACC1kHzLastYear,"linear",...
%           "DataVariables",["Velocity"]);
%       % Remove outliers
%       WT41GenAkselAACC1kHzLastYear = rmoutliers(WT41GenAkselAACC1kHzLastYear,"movmedian",24,...
%           "DataVariables",["Velocity"]);
%       % Smooth input data
%       WT41GenAkselAACC1kHzLastYear = smoothdata(WT41GenAkselAACC1kHzLastYear,...
%           "movmean","SmoothingFactor",0.05,"DataVariables",["Velocity"]);
% end
```

Data Cleaning of Generator and Rotor Bearing Vibration, and ACC of Shaft:

WT 21

```
% GenAkselA
WT211GenAkselAENV1CD = VibrationDataClean(WT211GenAkselAENV1LastYear);
WT211GenAkselAENV2CD = VibrationDataClean(WT211GenAkselAENV2LastYear);
WT211GenAkselAENV3CD = VibrationDataClean(WT211GenAkselAENV3LastYear);
WT211GenAkselAACC1kHzCD = ACCDataClean(WT211GenAkselAACC1kHzLastYear);

% GenAkselR
WT212GenAkselRENV1CD = VibrationDataClean(WT212GenAkselRENV1LastYear);
WT212GenAkselRENV2CD = VibrationDataClean(WT212GenAkselRENV2LastYear);
WT212GenAkselRENV3CD = VibrationDataClean(WT212GenAkselRENV3LastYear);
WT212GenAkselRACC1kHzCD = ACCDataClean(WT212GenAkselRACC1kHzLastYear);

% RotorAkselR
WT213RotorAkselRENV1CD = VibrationDataClean(WT213RotorAkselRENV1LastYear);
WT213RotorAkselRENV2CD = VibrationDataClean(WT213RotorAkselRENV2LastYear);
WT213RotorAkselRENV3CD = VibrationDataClean(WT213RotorAkselRENV3LastYear);
WT213RotorAkselRACC1kHzCD = ACCDataClean(WT213RotorAkselRACC1kHzLastYear);
```

Convert From Table to TimeTable

WT 21

```
% GenAkselA
WT211GenAkselAENV1CDTT = table2timetable(WT211GenAkselAENV1CD);
WT211GenAkselAENV2CDTT = table2timetable(WT211GenAkselAENV2CD);
WT211GenAkselAENV3CDTT = table2timetable(WT211GenAkselAENV3CD);
WT211GenAkselAACC1kHzCDTT = table2timetable(WT211GenAkselAACC1kHzCD);

% GenAkselR
WT212GenAkselRENV1CDTT = table2timetable(WT212GenAkselRENV1CD);
WT212GenAkselRENV2CDTT = table2timetable(WT212GenAkselRENV2CD);
WT212GenAkselRENV3CDTT = table2timetable(WT212GenAkselRENV3CD);
WT212GenAkselRACC1kHzCDTT = table2timetable(WT212GenAkselRACC1kHzCD);

% RotorAkselR
WT213RotorAkselRENV1CDTT = table2timetable(WT213RotorAkselRENV1CD);
WT213RotorAkselRENV2CDTT = table2timetable(WT213RotorAkselRENV2CD);
WT213RotorAkselRENV3CDTT = table2timetable(WT213RotorAkselRENV3CD);
WT213RotorAkselRACC1kHzCDTT = table2timetable(WT213RotorAkselRACC1kHzCD);
```

Ensemble Datastore structure for Diagnostic Feature Designer (DFD)

WT 21

Generator Bearing Axial, Radial, ACC of Shaft data

**RENV2 -> Discarded due to sensor data acquisition settings change, resulting in incorrect data*

```
Generator21ARENV123ACC_ensemble_table = table();
timestart = datetime(2022,04,05);
timeend = datetime(2022,04,06);
day = 0;
ca = "After";
cb = "Before";
for x= 1:401
    day = day + 1;
%   RotorSpeed
    r1perday = table(WT211GenAkselAENV1CDTT.Date(timerange(timestart + day,timeend + day)), ...
        WT211GenAkselAENV1CDTT.RotorSpeed(WT211GenAkselAENV1CDTT.Date(timerange(timestart + day,timeend + day))), ...
        'VariableNames',{'Time','RotorSpeed'});
    r1perdayTT = table2timetable(r1perday);
%   AENV1
    av1perday = table(WT211GenAkselAENV1CDTT.Date(timerange(timestart + day,timeend + day)), ...
        WT211GenAkselAENV1CDTT.Vibration(WT211GenAkselAENV1CDTT.Date(timerange(timestart + day,timeend + day))), ...
        'VariableNames',{'Time','Vibration'});
    av1perdayTT = table2timetable(av1perday);
%   AENV2
    av2perday = table(WT211GenAkselAENV2CDTT.Date(timerange(timestart + day,timeend + day)), ...
        WT211GenAkselAENV2CDTT.Vibration(WT211GenAkselAENV2CDTT.Date(timerange(timestart + day,timeend + day))), ...
        'VariableNames',{'Time','Vibration'});
    av2perdayTT = table2timetable(av2perday);
%   AENV3
    av3perday = table(WT211GenAkselAENV3CDTT.Date(timerange(timestart + day,timeend + day)), ...
        WT211GenAkselAENV3CDTT.Vibration(WT211GenAkselAENV3CDTT.Date(timerange(timestart + day,timeend + day))), ...
        'VariableNames',{'Time','Vibration'});
    av3perdayTT = table2timetable(av3perday);
%   AACC
    aaccperday = table(WT211GenAkselAACC1kHzCDTT.Date(timerange(timestart + day,timeend + day)), ...
        WT211GenAkselAACC1kHzCDTT.Velocity(WT211GenAkselAACC1kHzCDTT.Date(timerange(timestart + day,timeend + day))), ...
        'VariableNames',{'Time','Velocity'});
```

```
    aaccperdayTT = table2timetable(aaccperday);
%   RENV1
    rv1perday = table(WT212GenAkselRENV1CDTT.Date(timerange(timestart + day,timeend + day)), ...
        WT212GenAkselRENV1CDTT.Vibration(WT212GenAkselRENV1CDTT.Date(timerange(timestart + day,timeend + day))), ...
        'VariableNames',{'Time','Vibration'});
    rv1perdayTT = table2timetable(rv1perday);
%   RENV3
    rv3perday = table(WT212GenAkselRENV3CDTT.Date(timerange(timestart + day,timeend + day)), ...
        WT212GenAkselRENV3CDTT.Vibration(WT212GenAkselRENV3CDTT.Date(timerange(timestart + day,timeend + day))), ...
        'VariableNames',{'Time','Vibration'});
    rv3perdayTT = table2timetable(rv3perday);
%   RACC
    raccperday = table(WT212GenAkselRACC1kHzCDTT.Date(timerange(timestart + day,timeend + day)), ...
        WT212GenAkselRACC1kHzCDTT.Velocity(WT212GenAkselRACC1kHzCDTT.Date(timerange(timestart + day,timeend + day))), ...
        'VariableNames',{'Time','Velocity'});
    raccperdayTT = table2timetable(raccperday);

    Generator21ARENV123ACC_ensemble_table(day,:) = table({r1perdayTT},{av1perdayTT},{av2perdayTT},{av3perdayTT},{aaccperdayTT},{rv1perdayT
end
% Remove empty rows
Generator21ARENV123ACC_ensemble_table([129:132,138,139,222],:) = [];
```

Rotor Bearing Radial and ACC of Shaft data

**RENV2 -> Discarded due to sensor data acquisition settings change, resulting in incorrect data*

```
Rotor21RENV13ACC_ensemble_table = table();
timestart = datetime(2022,04,05);
timeend = datetime(2022,04,06);
day = 0;
for x= 1:402
    day = day + 1;
%   RotorSpeed
    r1perday = table(WT213RotorAkselRENV1CDTT.Date(timerange(timestart + day,timeend + day)), ...
        WT213RotorAkselRENV1CDTT.RotorSpeed(WT213RotorAkselRENV1CDTT.Date(timerange(timestart + day,timeend + day))), ...
        'VariableNames',{'Time','RotorSpeed'});
    r1perdayTT = table2timetable(r1perday);

%   RENV1
    rv1perday = table(WT213RotorAkselRENV1CDTT.Date(timerange(timestart + day,timeend + day)), ...
        WT213RotorAkselRENV1CDTT.Vibration(WT213RotorAkselRENV1CDTT.Date(timerange(timestart + day,timeend + day))), ...
        'VariableNames',{'Time','Vibration'});
    rv1perdayTT = table2timetable(rv1perday);
%   RENV3
    rv3perday = table(WT213RotorAkselRENV3CDTT.Date(timerange(timestart + day,timeend + day)), ...
        WT213RotorAkselRENV3CDTT.Vibration(WT213RotorAkselRENV3CDTT.Date(timerange(timestart + day,timeend + day))), ...
        'VariableNames',{'Time','Vibration'});
    rv3perdayTT = table2timetable(rv3perday);
%   RACC
    raccperday = table(WT213RotorAkselRACC1kHzCDTT.Date(timerange(timestart + day,timeend + day)), ...
        WT213RotorAkselRACC1kHzCDTT.Velocity(WT213RotorAkselRACC1kHzCDTT.Date(timerange(timestart + day,timeend + day))), ...
        'VariableNames',{'Time','Velocity'});
    raccperdayTT = table2timetable(raccperday);

    Rotor21RENV13ACC_ensemble_table(day,:) = table({r1perdayTT},{rv1perdayTT},{rv3perdayTT},{raccperdayTT},'VariableNames',{'RotorSpeed','
end
% Remove empty rows
Rotor21RENV13ACC_ensemble_table([129:132,138,139,222],:) = [];
```

DFD

```
diagnosticFeatureDesigner
```

Time-Domain Features extraction

WT 21

Generator

Statistical metrics -> Mean, RMS, Standard Deviation, Shape Factor, Kurtosis and Skewness

```
[Generator21FeatureTable,Generator21OutputTable] = diagnosticFeaturesGeneratorTimeSingalFeatures(Generator21ARENV123ACC_ensemble_table);
```

Rotor

Statistical metrics -> Mean, RMS, Standard Deviation, Shape Factor, Kurtosis and Skewness

```
[Rotor21FeatureTable,Rotor21OutputTable] = diagnosticFeaturesRotorTimeSingalFeatures(Rotor21RENV13ACC_ensemble_table);
```

```
% Convert to numeric array and remove missing data fields
Generator21TimeSignalFeatures_NumericArray = table2array(Generator21FeatureTable);
Generator21TimeSignalFeatures_NumericArray = rmmissing(Generator21TimeSignalFeatures_NumericArray);
```

```
% Spilt into generator singal features into the axial and radial tables
Generator21AxialFeaturesTable = Generator21FeatureTable(:,[1:30]);
Generator21RadialFeaturesTable = Generator21FeatureTable(:,[1:6,31:48]);
```

```
% Convert to numeric array and remove missing data fields
Generator21AxialFeaturesTable_NumericArray = table2array(Generator21AxialFeaturesTable);
Generator21AxialFeaturesTable_NumericArray = rmmissing(Generator21AxialFeaturesTable_NumericArray);
% Convert to numeric array and remove missing data fields
Generator21RadialFeaturesTable_NumericArray = table2array(Generator21RadialFeaturesTable);
Generator21RadialFeaturesTable_NumericArray = rmmissing(Generator21RadialFeaturesTable_NumericArray);
```

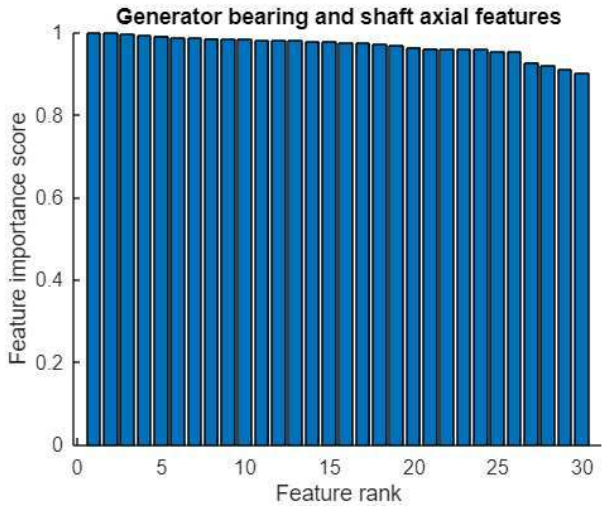
```
% Convert to numeric array and remove missing data fields
Rotor21TimeSignalFeatures_NumericArray = table2array(Rotor21FeatureTable);
Rotor21TimeSignalFeatures_NumericArray = rmmissing(Rotor21TimeSignalFeatures_NumericArray);
```

Feature ranking

Most important features selected by the fsulaplacian function which rank features for unsupervised learning using Laplacian scores

Generator Bearing Axial and ACC of Shaft data

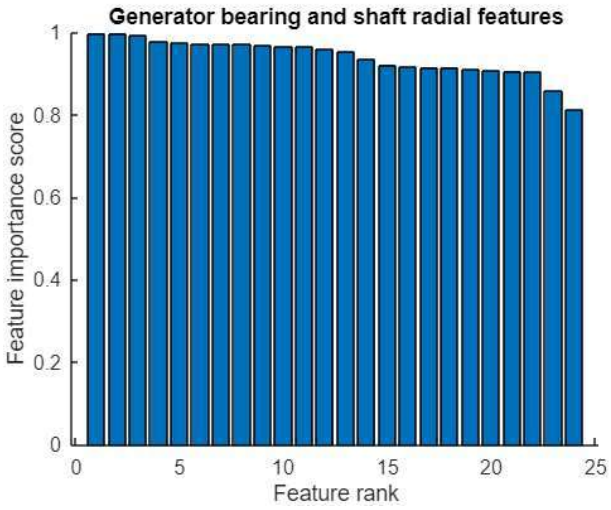
```
[gaidx,scores] = fsulaplacian(Generator21AxialFeaturesTable_NumericArray);
figure
hold on
bar(scores(gaidx))
xlabel('Feature rank')
ylabel('Feature importance score')
title('Generator bearing and shaft axial features')
hold off
```




```
Gen21AxialTopFeatures = gaidx(1:15);
Generator21AxialTopFeaturesTable_NumericArray = Generator21AxialFeaturesTable_NumericArray(:,[Gen21AxialTopFeatures]);
```

Generator Bearing Radial and ACC of Shaft data

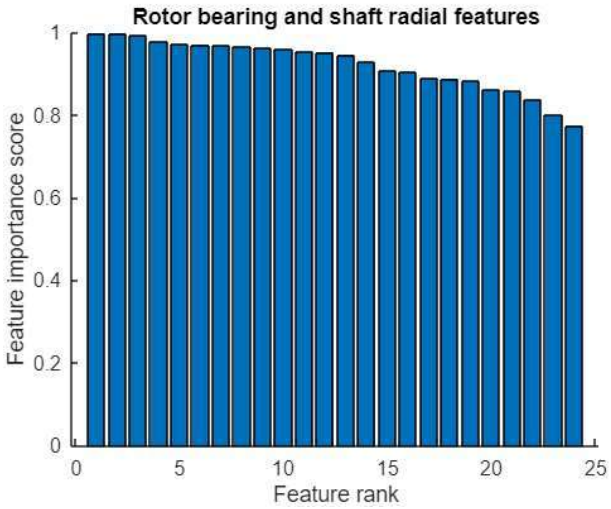
```
[gridx,scores] = fsulaplacian(Generator21RadialFeaturesTable_NumericArray);
figure
hold on
bar(scores(gridx))
xlabel('Feature rank')
ylabel('Feature importance score')
title('Generator bearing and shaft radial features')
hold off
```




```
Gen21RadialTopFeatures = gridx(1:15);
Generator21RadialTopFeaturesTable_NumericArray = Generator21RadialFeaturesTable_NumericArray(:,[Gen21RadialTopFeatures]);
```

Rotor Bearing Radial and ACC of Shaft data

```
[ridx,scores] = fsulaplacian(Rotor21TimeSignalFeatures_NumericArray);
figure
hold on
bar(scores(ridx))
xlabel('Feature rank')
ylabel('Feature importance score')
title('Rotor bearing and shaft radial features')
hold off
```




```
Rotor21RadialTopFeatures = ridx(1:15);
Rotor21TimeSignalTopFeatures_NumericArray = Rotor21TimeSignalFeatures_NumericArray(:,[Rotor21RadialTopFeatures]);
```

Unsupervised Machine Learning for Anomaly detection

Code adapted and used from MATLAB tutorial for PCA fault detection technique: [Predictive Maintenance: Unsupervised and Supervised Machine Learning - YouTube]

```
% https://www.youtube.com/watch?v=AS0H43hMoWM
```

Model 1: Principal Component Analysis (PCA)

WT 21

Generator Bearing Axial and ACC of Shaft data

Normalize the Data, so large values do not dominate items with small values.

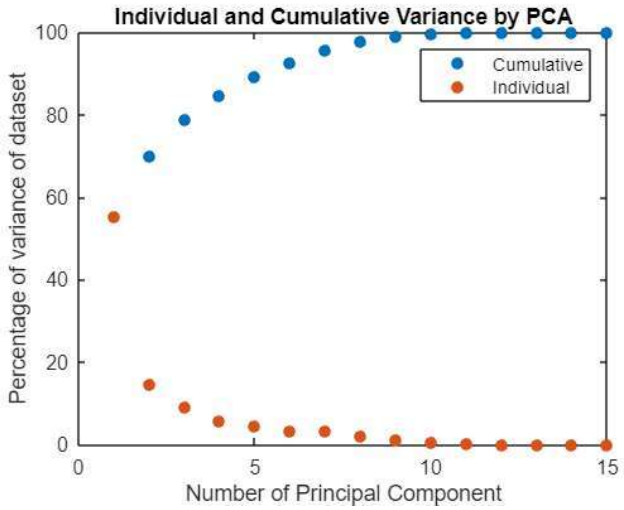
```
Gen21BearingAxialPCA_NumericArrayN = zscore(Generator21AxialTopFeaturesTable_NumericArray);
%GeneratorARENV123ACC_mean_array
```

Perform PCA calculation:

```
[ga21coeff,ga21score,ga21latent] = pca(Gen21BearingAxialPCA_NumericArrayN);
```

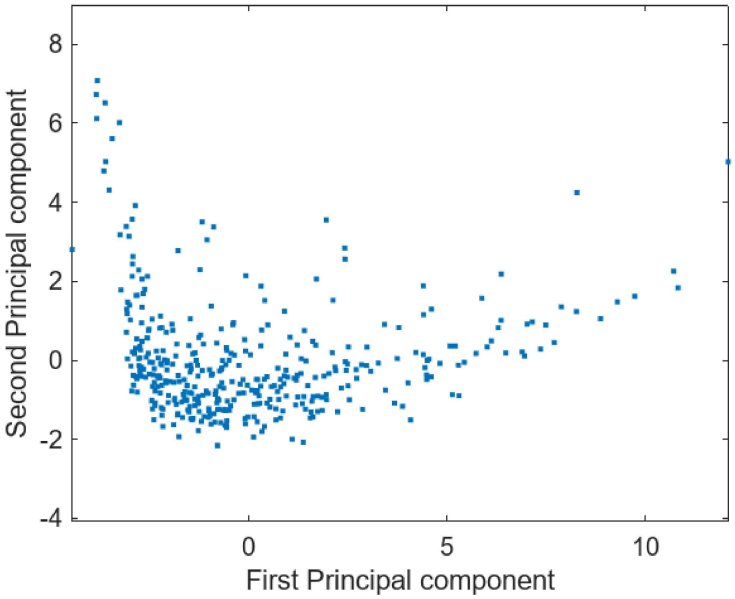
Next plot the total variance explained by each individual principal component together with the cumulative total which has been explained.

```
figure
plot([cumsum(ga21latent(1:15))/sum(ga21latent) ga21latent(1:15)/sum(ga21latent)]*100,'.','MarkerSize',18)
xlabel('Number of Principal Component');
ylabel('Percentage of variance of dataset');
legend('Cumulative','Individual')
title('Individual and Cumulative Variance by PCA')
```



Plot the first and second principal components. Only 70% of the total information in the data can be explained in the plot with the first and second principal components.

```
figure
plot(ga21score(:,1),ga21score(:,2),'.')
axis equal
xlabel('First Principal component')
ylabel('Second Principal component')
```

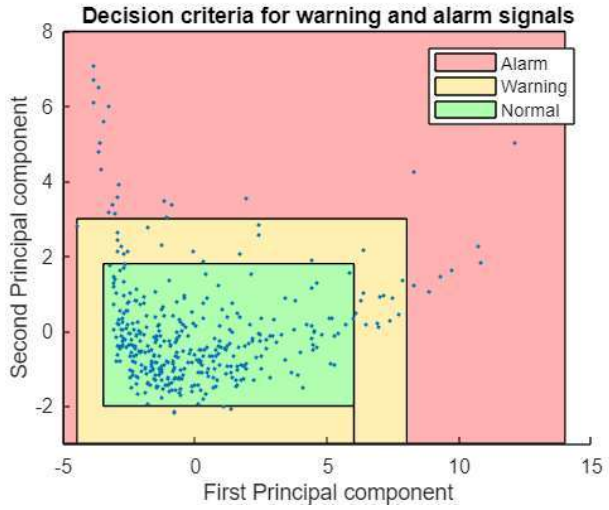


Develop a initial criteria for warning and alarm signals on the plot.

```
Alarm1 = ga21score(:,1) > 8 | ga21score(:,1) < -3.5 | ga21score(:,2) > 3 | ga21score(:,2) < -3;
Warning1 = ga21score(:,1) > 3 | ga21score(:,1) < -3 | ga21score(:,2) > 1.8 | ga21score(:,2) < -2 & ~Alarm1;

figure
hold on
% Red
```

```
patch([-5;-5;14;14;8;8;-4.5;-4.5;-4.5],[-3,8,8,-3,-3,3,3,-3,-3],'r','FaceAlpha',0.3)
% Yellow
patch([-4.5,-4.5,8,8,6,6,-3.5,-3.5,6,6],[-3,3,3,-3,-3,1.8,1.8,-2,-2,-3],'y','FaceColor',[1 .8 0],'FaceAlpha',0.3)
% Green
patch([-3.5,-3.5,6,6,-3],[-2,1.8,1.8,-2,-2],'g','FaceAlpha',0.3)
plot(ga21score(:,1),ga21score(:,2),'.')
ylim([-3 8])
hold off
xlabel('First Principal component')
ylabel('Second Principal component')
legend('Alarm','Warning','Normal')
title('Decision criteria for warning and alarm signals')
```



Generator Bearing Radial and ACC of Shaft data

Normalize the Data, so large values do not dominate items with small values.

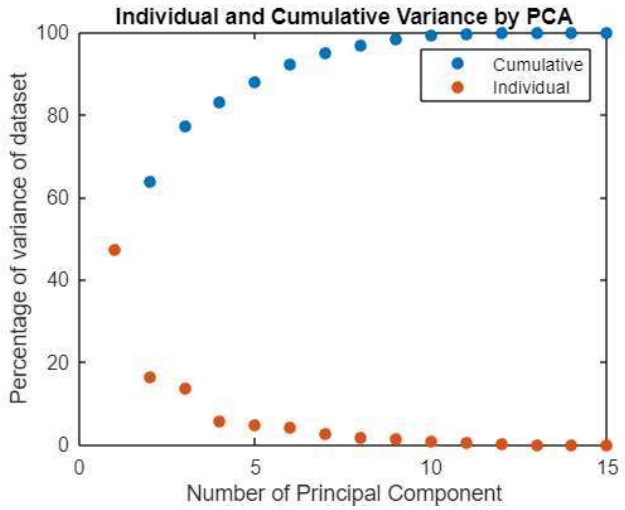
```
Gen21BearingRadialPCA_NumericArrayN = normalize(Generator21RadialTopFeaturesTable_NumericArray);
%GeneratorARENV123ACC_mean_array
```

Perform PCA calculation:

```
[gr21coeff,gr21score,gr21latent] = pca(Gen21BearingRadialPCA_NumericArrayN);
```

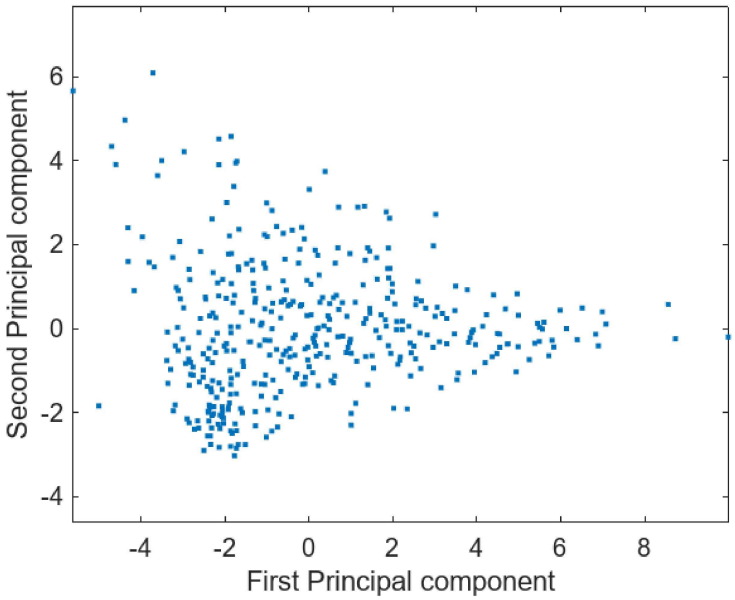
Next plot the total variance explained by each individual principal component together with the cumulative total which has been explained.

```
figure
plot([cumsum(gr21latent(1:15))/sum(gr21latent) gr21latent(1:15)/sum(gr21latent)]*100,'.','MarkerSize',18)
xlabel('Number of Principal Component');
ylabel('Percentage of variance of dataset');
legend('Cumulative','Individual')
title('Individual and Cumulative Variance by PCA')
```



Plot the first and second principal components. Only 64% of the total information in the data can be explained in the plot with the first and second principal components.

```
figure
plot(gr21score(:,1),gr21score(:,2),'.')
axis equal
xlabel('First Principal component')
ylabel('Second Principal component')
```



Develop a initial criteria for warning and alarm signals on the plot.

```
% Alarm2 = gr21score(:,1) > 8 | gr21score(:,1) < -3.5 | gr21score(:,2) > 3 | gr21score(:,2) < -3;
% Warning2 = gr21score(:,1) > 3 | gr21score(:,1) < -3 | gr21score(:,2) > 1.8 | gr21score(:,2) < -2 & ~Alarm2;
%
% figure
% hold on
% % Red
% patch([-7;-7;12;12;8;8;-4.5;-4.5;-4.5],[-4,7,7,-4,-4,3,3,-4,-4],'r','FaceAlpha',0.3)
% % Yellow
% patch([-4.5,-4.5,8,8,6,6,-3.5,-3.5,6,6],[-4,3,3,-4,-4,1.8,1.8,-3,-3,-4],'y','FaceColor',[1 .8 0],'FaceAlpha',0.3)
% % Green
% patch([-3.5,-3.5,6,6,-3],[-3,1.8,1.8,-3,-3],'g','FaceAlpha',0.3)
% plot(gr21score(:,1),gr21score(:,2),'.')
% xlim([-7 12])
% ylim([-4 7])
% hold off
% xlabel('First Principal component')
% ylabel('Second Principal component')
% legend('Alarm','Warning','Normal')
% title('Decision criteria for warning and alarm signals')
```

Rotor Bearing Radial and ACC of Shaft data

Normalize the Data, so large values do not dominate items with small values.

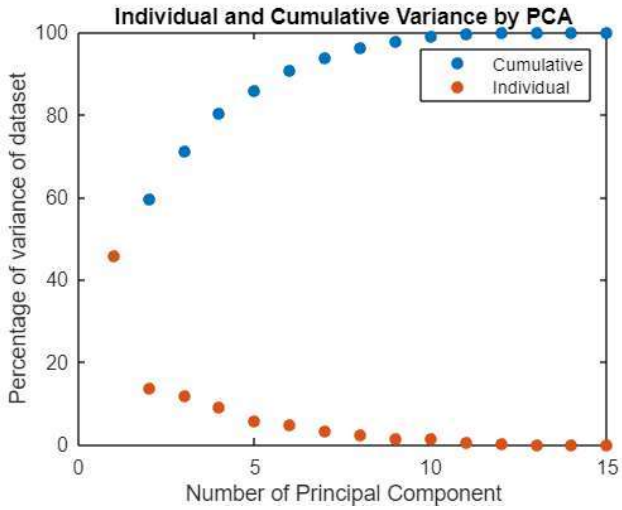
```
Rotor21BearingPCA_NumericArrayN = normalize(Rotor21TimeSignalTopFeatures_NumericArray);
```

Perform PCA calculation:

```
[r21coeff,r21score,r21latent] = pca(Rotor21BearingPCA_NumericArrayN);
```

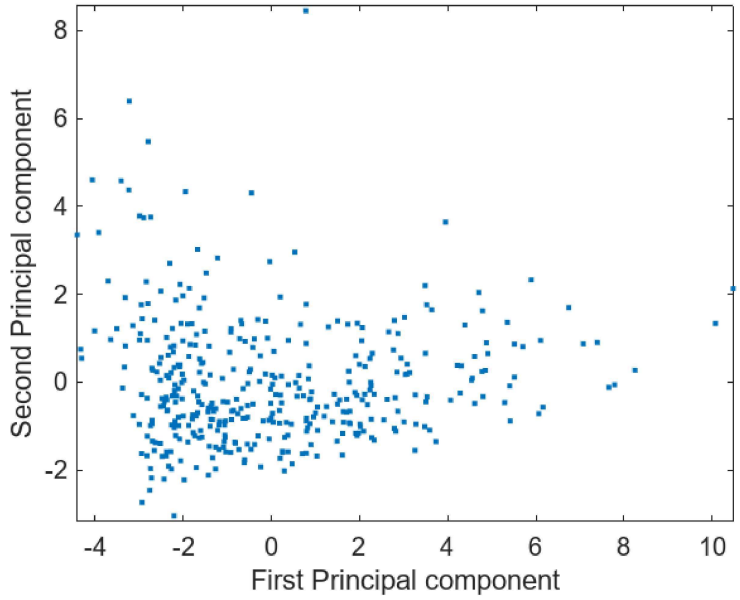
Next plot the total variance explained by each individual principal component together with the cumulative total which has been explained.

```
figure
plot([cumsum(r21latent(1:15))/sum(r21latent) r21latent(1:15)/sum(r21latent)]*100,'.','MarkerSize',18)
xlabel('Number of Principal Component');
ylabel('Percentage of variance of dataset');
legend('Cumulative','Individual')
title('Individual and Cumulative Variance by PCA')
```



Plot the first and second principal components. Nearly 60% of the total information in the data can be explained in the plot.

```
figure
plot(r21score(:,1),r21score(:,2),'.')
axis equal
xlabel('First Principal component')
ylabel('Second Principal component')
```



Develop a initial criteria for warning and alarm signals on the plot.

```
% Alarm3 = r21score(:,1) > 6 | r21score(:,1) < -3.5 | r21score(:,2) > 3 | r21score(:,2) < -3;
% Warning3 = r21score(:,1) > 3 | r21score(:,1) < -3 | r21score(:,2) > 1.8 | r21score(:,2) < -2 & ~Alarm3;
%
% figure
% hold on
%
% % Red
% patch([-7;-7;12;12;8;8;-4.5;-4.5;-4.5],[-4,7,7,-4,-4,3,3,-4,-4],'r','FaceAlpha',0.3)
% % Yellow
% patch([-4.5,-4.5,8,8,6,6,-3.5,-3.5,6,6],[-4,3,3,-4,-4,1.8,1.8,-3,-3,-4],'y','FaceColor',[1 .8 0],'FaceAlpha',0.3)
% % Green
% patch([-3.5,-3.5,6,6,-3],[-3,1.8,1.8,-3,-3],'g','FaceAlpha',0.3)
% plot(r21score(:,1),r21score(:,2),'.')
% xlim([-7 12])
% ylim([-4 7])
% hold off
% xlabel('First Principal component')
% ylabel('Second Principal component')
```

```
% legend('Alarm','Warning','Normal')
% title('Decision criteria for warning and alarm signals')
```

Model 2: Isolation Forest

MATLAB code here on is adapted and used from MATLAB example: Unsupervised Anomaly Detection - MATLAB & Simulink - MathWorks Nordic

```
% https://se.mathworks.com/help/stats/unsupervised-anomaly-detection.html
```

Used for Isolation forest, Local outlier factor, one-class support vector machine (SVM) and Mahalanobis Distance

Generator Bearing Axial and ACC of Shaft data

Contamination fraction of 0.20 is used

```
ga21contaminationFraction = 0.20;
```

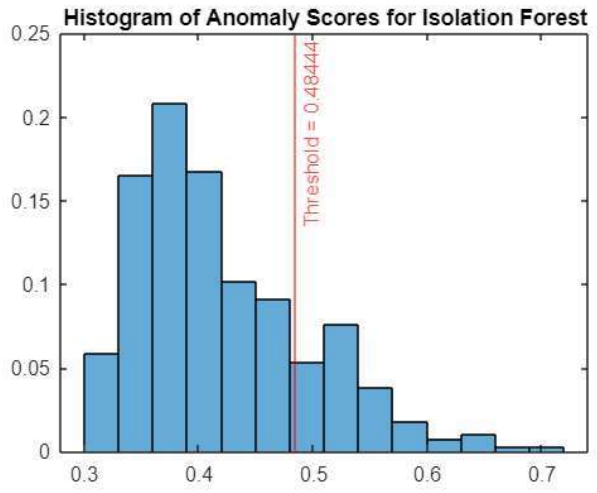
Isolation forest model training

```
[ga21N,ga21D] = size(Generator21AxialTopFeaturesTable_NumericArray);
rng("default") % For reproducibility
[ga21forest,ga21tf_forest,ga21s_forest] = iforest(Generator21AxialTopFeaturesTable_NumericArray,ContaminationFraction=ga21contaminationFra
```

Here, the anomaly indicator is gatf_forest and anomaly score is gas_forest

Histogram of the anomaly scores. A vertical line is placed at the score threshold corresponding to the specified fraction.

```
figure
histogram(ga21s_forest,Normalization="probability")
xline(ga21forest.ScoreThreshold,"r-", ...
    join(["Threshold =" ga21forest.ScoreThreshold]))
title("Histogram of Anomaly Scores for Isolation Forest")
```



Fraction of detection anomalies in the Generator bearing and shaft data.

```
ga21OF_forest = sum(ga21tf_forest)/ga21N
```

```
ga21OF_forest = 0.2005
```

Generator Bearing Radial and ACC of Shaft data

Contamination fraction of 0.10 is used

```
gr21contaminationFraction = 0.10;
```

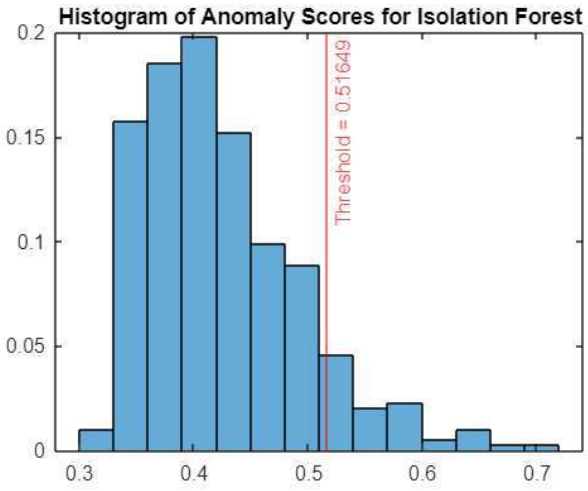
Isolation forest model training

```
[gr21N,gr21D] = size(Generator21RadialTopFeaturesTable_NumericArray);
rng("default") % For reproducibility
[gr21forest,gr21tf_forest,gr21s_forest] = iforest(Generator21RadialTopFeaturesTable_NumericArray,ContaminationFraction=gr21contaminationFr
```

Here, the anomaly indicator is grtf_forest and anomaly score is grs_forest

Histogram of the anomaly scores. A vertical line is placed at the score threshold corresponding to the specified fraction.

```
figure
histogram(gr21s_forest,Normalization="probability")
xline(gr21forest.ScoreThreshold,"r-", ...
    join(["Threshold =" gr21forest.ScoreThreshold]))
title("Histogram of Anomaly Scores for Isolation Forest")
```



Fraction of detection anomalies in the Generator bearing and shaft data.

```
gr21OF_forest = sum(gr21tf_forest)/gr21N
```

```
gr21OF_forest = 0.0990
```

Rotor Bearing Radial and ACC of Shaft data

Contamination fraction of 0.10 is used

```
r21contaminationFraction = 0.10;
```

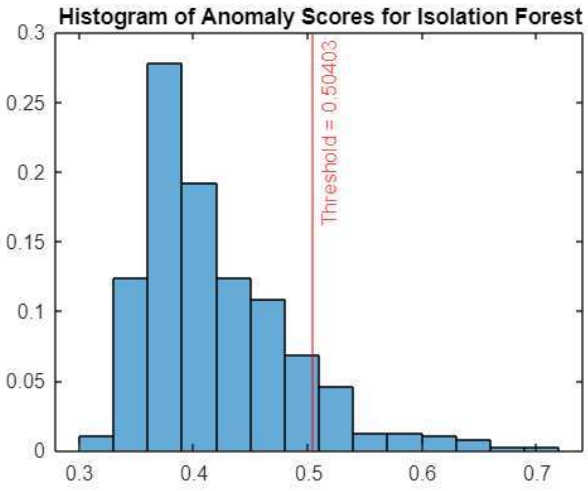
Isolation forest model training

```
[r21N,r21D] = size(Rotor21TimeSignalTopFeatures_NumericArray);
rng("default") % For reproducibility
[r21forest,r21tf_forest,r21s_forest] = iforest(Rotor21TimeSignalTopFeatures_NumericArray,ContaminationFraction=r21contaminationFraction);
```

Here, the anomaly indicator is rtf_forest and anomaly score is rs_forest

Histogram of the anomaly scores. A vertical line is placed at the score threshold corresponding to the specified fraction.

```
figure
histogram(r21s_forest,Normalization="probability")
xline(r21forest.ScoreThreshold,"r-", ...
    join(["Threshold =" r21forest.ScoreThreshold]))
title("Histogram of Anomaly Scores for Isolation Forest")
```



Fraction of detected anomalies in the Rotor bearing and shaft data.

```
r21OF_forest = sum(r21tf_forest)/r21N
```

```
r21OF_forest = 0.0987
```

Model 3: Local Outlier Factor

Generator Bearing Axial and ACC of Shaft data

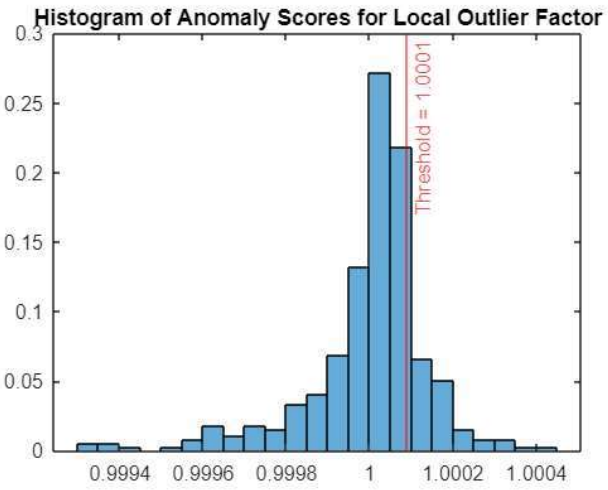
Train the local outlier factor model

```
[ga21LOFObj,ga21tf_lof,ga21s_lof] = lof(Generator21AxialTopFeaturesTable_NumericArray, ...
    ContaminationFraction=ga21contaminationFraction, ...
    NumNeighbors=500,Distance="mahalanobis");
```

Here, the anomaly indicator is gatf_lof and anomaly score is gas_lof

Histogram of the anomaly scores. A vertical line is placed at the score threshold corresponding to the specified fraction.

```
figure
histogram(ga21s_lof,Normalization="probability")
xline(ga21LOFObj.ScoreThreshold,"r-", ...
    join(["Threshold =" ga21LOFObj.ScoreThreshold]))
title("Histogram of Anomaly Scores for Local Outlier Factor")
```



Fraction of detection anomalies in the Generator bearing and shaft data.

```
OF_lof = sum(ga21tf_lof)/ga21N
```

```
OF_lof = 0.2005
```

Generator Bearing Radial and ACC of Shaft data

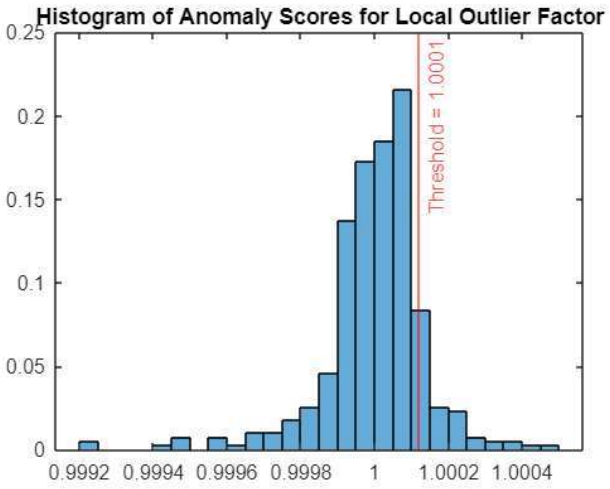
Train the local outlier factor model

```
[gr21LOFObj,gr21tf_lof,gr21s_lof] = lof(Generator21RadialTopFeaturesTable_NumericArray, ...
    ContaminationFraction=gr21contaminationFraction, ...
    NumNeighbors=500,Distance="mahalanobis");
```

Here, the anomaly indicator is grtf_lof and anomaly score is grs_lof

Histogram of the anomaly scores. A vertical line is placed at the score threshold corresponding to the specified fraction.

```
figure
histogram(gr21s_lof,Normalization="probability")
xline(gr21LOFObj.ScoreThreshold,"r-", ...
    join(["Threshold =" gr21LOFObj.ScoreThreshold]))
title("Histogram of Anomaly Scores for Local Outlier Factor")
```



Fraction of detection anomalies in the Generator bearing and shaft data.

```
OF_lof = sum(gr21tf_lof)/gr21N
```

```
OF_lof = 0.0990
```

Rotor Bearing Radial and ACC of Shaft data

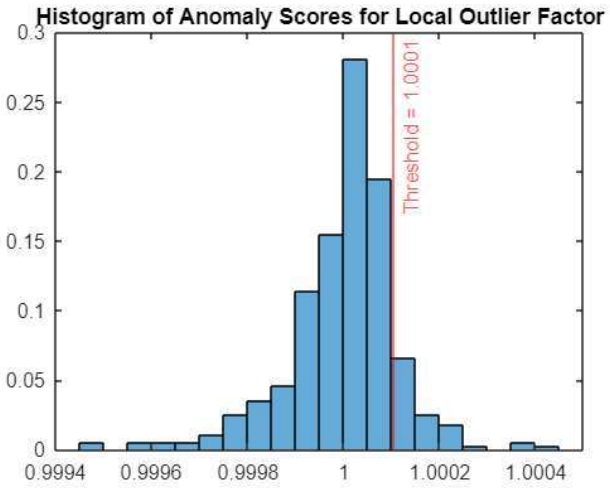
Train the local outlier factor model

```
[r21LOFObj,r21tf_lof,r21s_lof] = lof(Rotor21TimeSignalTopFeatures_NumericArray, ...
    ContaminationFraction=r21contaminationFraction, ...
    NumNeighbors=500,Distance="mahalanobis");
```

Here, the anomaly indicator is rtf_lof and anomaly score is rs_lof

Histogram of the anomaly scores. A vertical line is placed at the score threshold corresponding to the specified fraction.

```
figure
histogram(r21s_lof,Normalization="probability")
xline(r21LOFObj.ScoreThreshold,"r-", ...
    join(["Threshold =" r21LOFObj.ScoreThreshold]))
title("Histogram of Anomaly Scores for Local Outlier Factor")
```



Fraction of detection anomalies in the Rotor bearing and shaft data.

```
OF_lof = sum(r21tf_lof)/r21N
```

```
OF_lof = 0.0987
```

Model 4: One-class support vector machine (SVM)

Generator Bearing Axial and ACC of Shaft data

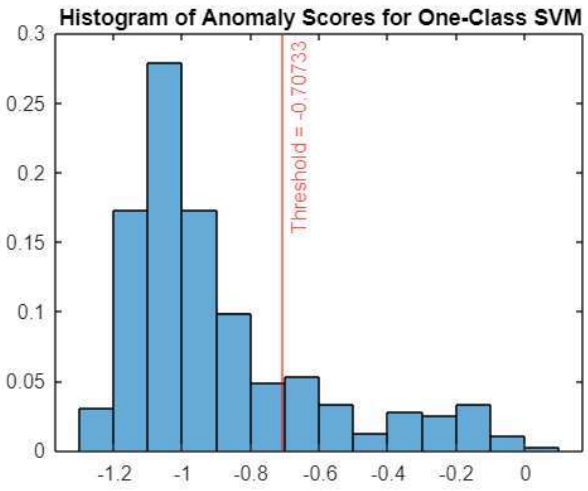
Train the one-class SVM model

```
[ga21Mdl,ga21tf_OCSVM,ga21s_OCSVM] = ocsvm(Generator21AxialTopFeaturesTable_NumericArray, ...
    ContaminationFraction=ga21contaminationFraction, ...
    KernelScale="auto",StandardizeData=true);
```

Here, the anomaly indicator is gatf_OCSVM and anomaly score is gas_OCSVM

Histogram of the anomaly scores. A vertical line is placed at the score threshold corresponding to the specified fraction.

```
figure
histogram(ga21s_OCSVM,Normalization="probability")
xline(ga21Mdl.ScoreThreshold,"r-", ...
    join(["Threshold =" ga21Mdl.ScoreThreshold]))
title("Histogram of Anomaly Scores for One-Class SVM")
```



Fraction of detection anomalies in the Generator bearing and shaft data.

```
OF_OCSVM = sum(ga21tf_OCSVM)/ga21N
```

```
OF_OCSVM = 0.2005
```

Generator Bearing Radial and ACC of Shaft data

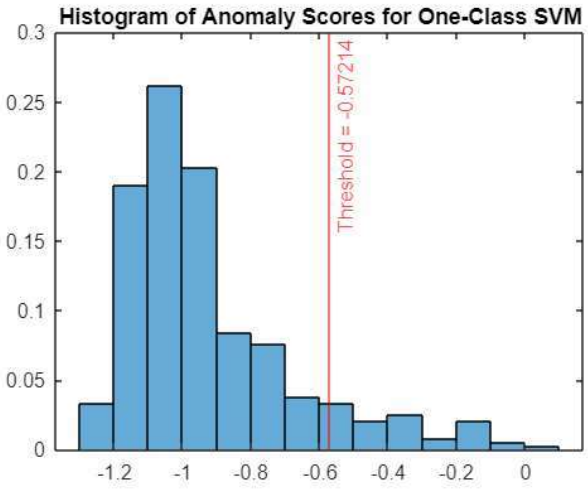
Train the one-class SVM model

```
[gr21Mdl,gr21tf_OCSVM,gr21s_OCSVM] = ocsvm(Generator21RadialTopFeaturesTable_NumericArray, ...
    ContaminationFraction=gr21contaminationFraction, ...
    KernelScale="auto",StandardizeData=true);
```

Here, the anomaly indicator is grtf_OCSVM and anomaly score is grs_OCSVM

Histogram of the anomaly scores. A vertical line is placed at the score threshold corresponding to the specified fraction.

```
figure
histogram(gr21s_OCSVM,Normalization="probability")
xline(gr21Mdl.ScoreThreshold,"r-", ...
    join(["Threshold =" gr21Mdl.ScoreThreshold]))
title("Histogram of Anomaly Scores for One-Class SVM")
```



Fraction of detection anomalies in the Generator bearing and shaft data.

```
OF_OCSVM = sum(gr21tf_OCSVM)/gr21N
```

```
OF_OCSVM = 0.0990
```

Rotor Bearing Radial and ACC of Shaft data

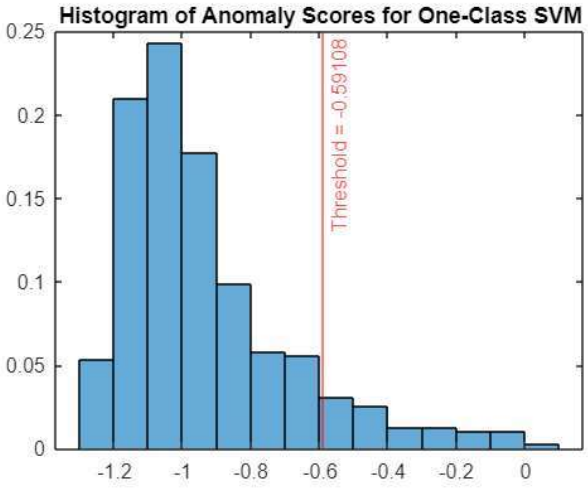
Train the one-class SVM model

```
[r21Mdl,r21tf_OCSVM,r21s_OCSVM] = ocsvm(Rotor21TimeSignalTopFeatures_NumericArray, ...
    ContaminationFraction=r21contaminationFraction, ...
    KernelScale="auto",StandardizeData=true);
```

Here, the anomaly indicator is rtf_OCSVM and anomaly score is rs_OCSVM

Histogram of the anomaly scores. A vertical line is placed at the score threshold corresponding to the specified fraction.

```
figure
histogram(r21s_OCSVM,Normalization="probability")
xline(r21Mdl.ScoreThreshold,"r-", ...
    join(["Threshold =" r21Mdl.ScoreThreshold]))
title("Histogram of Anomaly Scores for One-Class SVM")
```



Fraction of detection anomalies in the Rotor bearing and shaft data.

```
OF_OCSVM = sum(r21tf_OCSVM)/r21N
```

```
OF_OCSVM = 0.0987
```

Model 5: Mahalanobis Distance

Generator Bearing Axial and ACC of Shaft data

Compute the robust Mahalanobis distances and robust estimates for mean and covariance of the data.

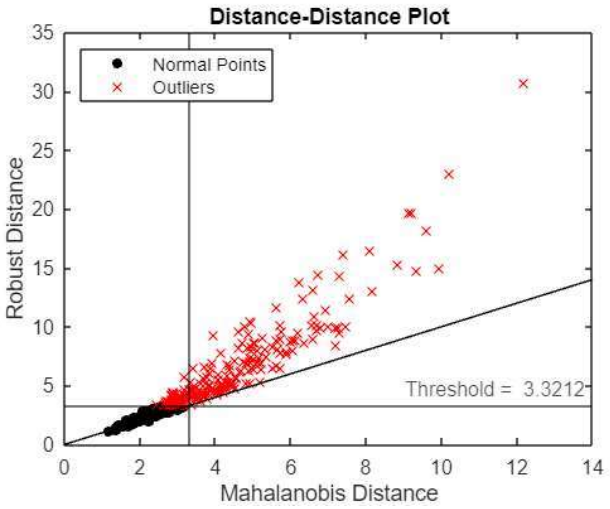
```
[ga21sigma,a21mu,ga21s_robustcov,ga21tf_robustcov_default] = robustcov(Generator21AxialTopFeaturesTable_NumericArray, ...
    OutlierFraction=gacontaminationFraction);
ga21s_robustcov_threshold = sqrt(chi2inv(1-ga21contaminationFraction,D))
```

```
ga21s_robustcov_threshold = 3.3212
```

```
ga21tf_robustcov = ga21s_robustcov > ga21s_robustcov_threshold;
```

Distance-distance plot for checking multivariate normality of the data.

```
figure
ga21d_classical = pdist2(Generator21AxialTopFeaturesTable_NumericArray,mean(Generator21AxialTopFeaturesTable_NumericArray),"mahalanobis");
gscatter(ga21d_classical,ga21s_robustcov,ga21tf_robustcov,"kr",".x")
xline(ga21s_robustcov_threshold,"k-")
yline(ga21s_robustcov_threshold,"k-", ...
    join(["Threshold = " ga21s_robustcov_threshold]));
l = refline([1 0]);
l.Color = "k";
xlabel("Mahalanobis Distance")
ylabel("Robust Distance")
legend("Normal Points","Outliers",Location="northwest")
title("Distance-Distance Plot")
```



The data set does not follow multivariate normal distribution, as points are not cluster along the diagonal reference line. Hence, use new threshold by using the quantile of the distance values for the cumulative probability (`1 - contaminationfactor`) to find a threshold.

```
ga21s_robustcov_threshold = quantile(ga21s_robustcov,1-ga21contaminationFraction);
```

Obtain new anomaly indicators

```
ga21tf_robustcov = ga21s_robustcov > ga21s_robustcov_threshold;
```

Fraction of detected anomalies in the data.

```
OF_robustcov = sum(ga21tf_robustcov)/ga21N
```

```
OF_robustcov = 0.2005
```

Generator Bearing Radial and ACC of Shaft data

Compute the robust Mahalanobis distances and robust estimates for mean and covariance of the data.

```
[gr21sigma,r21mu,gr21s_robustcov,gr21tf_robustcov_default] = robustcov(Generator21RadialTopFeaturesTable_NumericArray, ...
    OutlierFraction=gr21contaminationFraction);
gr21s_robustcov_threshold = sqrt(chi2inv(1-gr21contaminationFraction,D))
```

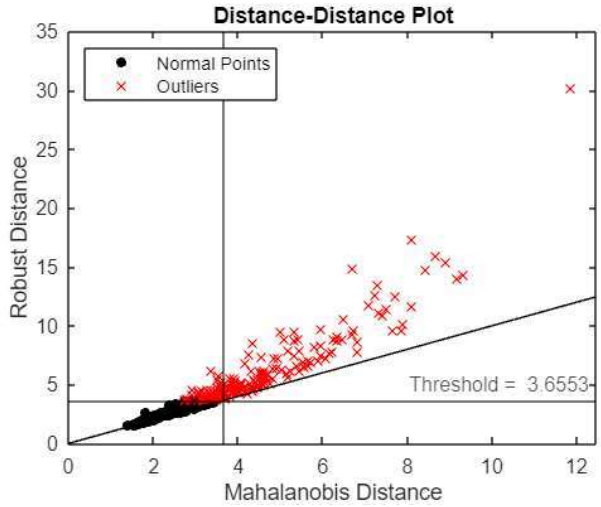
```
gr21s_robustcov_threshold = 3.6553
```

```
gr21tf_robustcov = gr21s_robustcov > gr21s_robustcov_threshold;
```

Distance-distance plot for checking multivariate normality of the data.

```
figure
gr21d_classical = pdist2(Generator21RadialTopFeaturesTable_NumericArray,mean(Generator21RadialTopFeaturesTable_NumericArray),"mahalanobis'
gscatter(gr21d_classical,gr21s_robustcov,gr21tf_robustcov,"kr",".x")
xline(gr21s_robustcov_threshold,"k-")
yline(gr21s_robustcov_threshold,"k-", ...
    join(["Threshold = " gr21s_robustcov_threshold]));
l = refline([1 0]);
```

```
l.Color = "k";
xlabel("Mahalanobis Distance")
ylabel("Robust Distance")
legend("Normal Points","Outliers",Location="northwest")
title("Distance-Distance Plot")
```



The data set does not follow multivariate normal distribution, as points are not cluster along the diagonal reference line. Hence, use new threshold by using the quantile of the distance values for the cumulative probability (`1 - contaminationfactor`) to find a threshold.

```
gr21s_robustcov_threshold = quantile(gr21s_robustcov,1-gr21contaminationFraction);
```

Obtain new anomaly indicators

```
gr21tf_robustcov = gr21s_robustcov > gr21s_robustcov_threshold;
```

Fraction of detected anomalies in the data.

```
OF_robustcov = sum(gr21tf_robustcov)/gr21N
```

```
OF_robustcov = 0.0990
```

Rotor Bearing Radial and ACC of Shaft data

Compute the robust Mahalanobis distances and robust estimates for mean and covariance of the data.

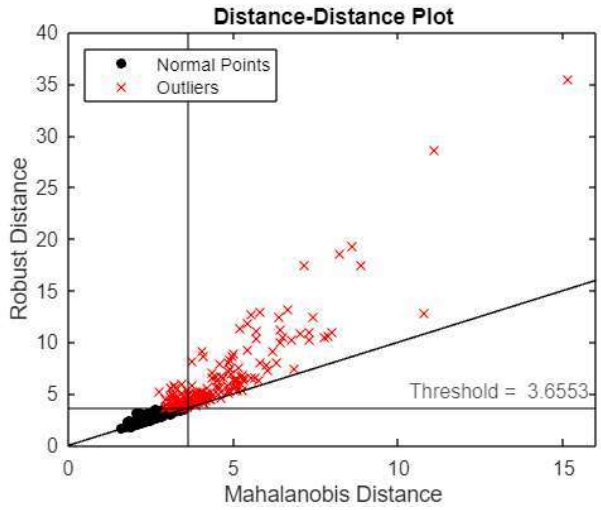
```
[r21sigma,rr21mu,r21s_robustcov,r21tf_robustcov_default] = robustcov(Rotor21TimeSignalTopFeatures_NumericArray, ...
    OutlierFraction=r21contaminationFraction);
r21s_robustcov_threshold = sqrt(chi2inv(1-r21contaminationFraction,D))
```

```
r21s_robustcov_threshold = 3.6553
```

```
r21tf_robustcov = r21s_robustcov > r21s_robustcov_threshold;
```

Distance-distance plot for checking multivariate normality of the data.

```
figure
r21d_classical = pdist2(Rotor21TimeSignalTopFeatures_NumericArray,mean(Rotor21TimeSignalTopFeatures_NumericArray),"mahalanobis");
gscatter(r21d_classical,r21s_robustcov,r21tf_robustcov,"kr",".x")
xline(r21s_robustcov_threshold,"k-")
yline(r21s_robustcov_threshold,"k-", ...
    join(["Threshold = " r21s_robustcov_threshold]));
l = refline([1 0]);
l.Color = "k";
xlabel("Mahalanobis Distance")
ylabel("Robust Distance")
legend("Normal Points","Outliers",Location="northwest")
title("Distance-Distance Plot")
```



The data set does not follow multivariate normal distribution, as points are not cluster along the diagonal reference line. Hence, use new threshold by using the quantile of the distance values for the cumulative probability (`1 - contaminationfactor`) to find a threshold.

```
r21s_robustcov_threshold = quantile(r21s_robustcov,1-r21contaminationFraction);
```

Obtain new anomaly indicators

```
r21tf_robustcov = r21s_robustcov > r21s_robustcov_threshold;
```

Fraction of detected anomalies in the data.

```
OF_robustcov = sum(r21tf_robustcov)/r21N
```

```
OF_robustcov = 0.0987
```

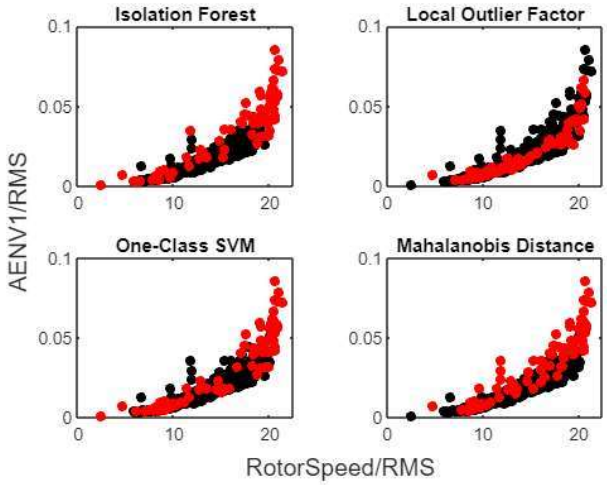
Comparison of the outlier detection models

Generator Bearing Axial and ACC of Shaft data

Compare the detected outliers

Visualize the normal and outliers detected.

```
figure
t = tiledlayout(2,2);
nexttile
gscatter(Generator21AxialTopFeaturesTable_NumericArray(:,1),Generator21AxialTopFeaturesTable_NumericArray(:,10),ga21tf_forest,"kr",[],[],'
title("Isolation Forest")
% nexttile
% gscatter(GeneratorARENV123ACC_mean_array(:,idx(1)),GeneratorARENV123ACC_mean_array(:,idx(2)),gtf_rforest,"kr",[],[],"off")
% title("Robust Random Cut Forest")
nexttile
gscatter(Generator21AxialTopFeaturesTable_NumericArray(:,1),Generator21AxialTopFeaturesTable_NumericArray(:,10),ga21tf_lof,"kr",[],[],"off
title("Local Outlier Factor")
nexttile
gscatter(Generator21AxialTopFeaturesTable_NumericArray(:,1),Generator21AxialTopFeaturesTable_NumericArray(:,10),ga21tf_OCSVM,"kr",[],[],"o
title("One-Class SVM")
nexttile
gscatter(Generator21AxialTopFeaturesTable_NumericArray(:,1),Generator21AxialTopFeaturesTable_NumericArray(:,10),ga21tf_robustcov,"kr",[],[
title("Mahalanobis Distance")
% l = legend("Normal Points","Outliers");
% l.Layout.Tile = 3;
xlabel(t,join(["RotorSpeed/RMS"]))
ylabel(t,join(["AENV1/RMS"]))
```

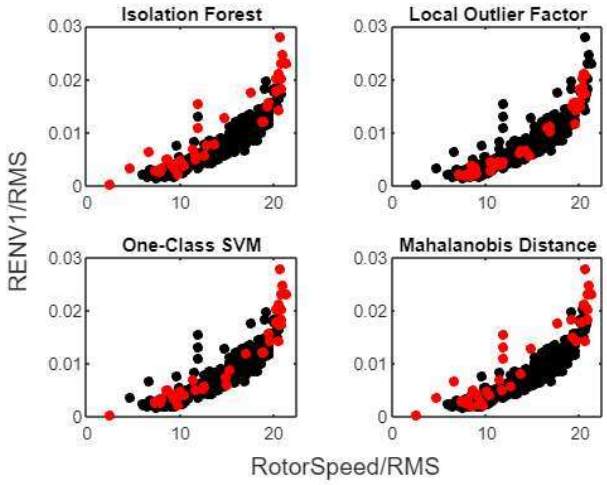



```
% AENV1Kurtosis
```

Generator Bearing Radial and ACC of Shaft data

Compare the detected outliers

```
figure
t = tiledlayout(2,2);
nexttile
gscatter(Generator21RadialTopFeaturesTable_NumericArray(:,1),Generator21RadialTopFeaturesTable_NumericArray(:,7),gr21tf_forest,"kr",[],[],
title("Isolation Forest")
% nexttile
% gscatter(GeneratorARENV123ACC_mean_array(:,idx(1)),GeneratorARENV123ACC_mean_array(:,idx(2)),gtf_rforest,"kr",[],[],"off")
% title("Robust Random Cut Forest")
nexttile
gscatter(Generator21RadialTopFeaturesTable_NumericArray(:,1),Generator21RadialTopFeaturesTable_NumericArray(:,7),gr21tf_lof,"kr",[],[],"of
title("Local Outlier Factor")
nexttile
gscatter(Generator21RadialTopFeaturesTable_NumericArray(:,1),Generator21RadialTopFeaturesTable_NumericArray(:,7),gr21tf_OCSVM,"kr",[],[],'
title("One-Class SVM")
nexttile
gscatter(Generator21RadialTopFeaturesTable_NumericArray(:,1),Generator21RadialTopFeaturesTable_NumericArray(:,7),gr21tf_robustcov,"kr",[],
title("Mahalanobis Distance")
% l = legend("Normal Points","Outliers");
% l.Layout.Tile = 3;
xlabel(t,join(["RotorSpeed/RMS"]))
ylabel(t,join(["RENV1/RMS"]))
```



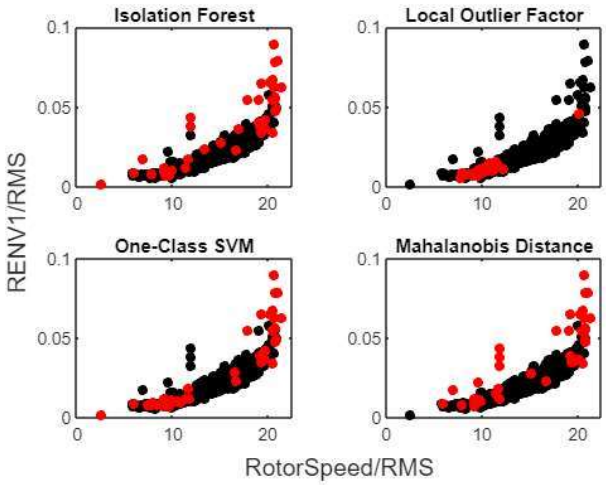

```
% RENV1_Mean
% gridx
```

Rotor Bearing Radial and ACC of Shaft data

Compare the detected outliers

Visualize the normal and outliers detected.

```
figure
t = tiledlayout(2,2);
nexttile
gscatter(Rotor21TimeSignalTopFeatures_NumericArray(:,1),Rotor21TimeSignalTopFeatures_NumericArray(:,11),r21tf_forest,"kr",[],[],"off")
title("Isolation Forest")
% nexttile
% gscatter(GeneratorARENV123ACC_mean_array(:,idx(1)),GeneratorARENV123ACC_mean_array(:,idx(2)),gtf_rforest,"kr",[],[],"off")
% title("Robust Random Cut Forest")
nexttile
gscatter(Rotor21TimeSignalTopFeatures_NumericArray(:,1),Rotor21TimeSignalTopFeatures_NumericArray(:,11),r21tf_lof,"kr",[],[],"off")
title("Local Outlier Factor")
nexttile
gscatter(Rotor21TimeSignalTopFeatures_NumericArray(:,1),Rotor21TimeSignalTopFeatures_NumericArray(:,11),r21tf_OCSVM,"kr",[],[],"off")
title("One-Class SVM")
nexttile
gscatter(Rotor21TimeSignalTopFeatures_NumericArray(:,1),Rotor21TimeSignalTopFeatures_NumericArray(:,11),r21tf_robustcov,"kr",[],[],"off")
title("Mahalanobis Distance")
% l = legend("Normal Points","Outliers");
% l.Layout.Tile = 3;
xlabel(t,join(["RotorSpeed/RMS"]))
ylabel(t,join(["RENV1/RMS"]))
```




```
% RENV1_kurtosis
```

```
function [featureTable,outputTable] = diagnosticFeaturesGeneratorTimeSingalFeatures(inputData)
%DIAGNOSTICFEATURES recreates results in Diagnostic Feature Designer.
%
% Input:
%  inputData: A table or a cell array of tables/matrices containing the
%  data as those imported into the app.
%
% Output:
%  featureTable: A table containing all features and condition variables.
%  outputTable: A table containing the computation results.
%
% This function computes features:
%  RotorSpeed_sigstats/Kurtosis
%  RotorSpeed_sigstats/Mean
%  RotorSpeed_sigstats/RMS
%  RotorSpeed_sigstats/ShapeFactor
%  RotorSpeed_sigstats/Skewness
%  RotorSpeed_sigstats/Std
%  AENV1_sigstats/Kurtosis
%  AENV1_sigstats/Mean
%  AENV1_sigstats/RMS
%  AENV1_sigstats/ShapeFactor
%  AENV1_sigstats/Skewness
%  AENV1_sigstats/Std
%  AENV2_sigstats/Kurtosis
%  AENV2_sigstats/Mean
%  AENV2_sigstats/RMS
%  AENV2_sigstats/ShapeFactor
%  AENV2_sigstats/Skewness
%  AENV2_sigstats/Std
%  AENV3_sigstats/Kurtosis
%  AENV3_sigstats/Mean
%  AENV3_sigstats/RMS
```

Bibliography

[n. d.]. Principal component analysis of raw data. https://se.mathworks.com/help/stats/pca.html#d124e743782

2023a. Designing Algorithms for Condition Monitoring and Predictive Maintenance. https://se.mathworks.com/help/predmaint/gs/designing-algorithms-for-condition-monitoring-and-predictive-maintenance.html

2023b. Explore Ensemble Data and Compare Features Using Diagnostic Feature Designer. https://se.mathworks.com/help/predmaint/ug/explore-ensemble-data-and-compare-features-using-diagnostic-feature-designer.html

2023c. Om Aneo. https://www.aneo.com/om-oss/om-aneo/

2023d. What is Anomaly Detection? https://se.mathworks.com/discovery/anomaly-detection.html

Akshara_416. 2021. Anomaly detection using Isolation Forest - A Complete Guide. https://www.analyticsvidhya.com/blog/2021/07/anomaly-detection-using-isolation-forest-a-complete-guide/

akshayvarma72. 2020. Local outlier factor. https://www.geeksforgeeks.org/local-outlier-factor/

Mahbubul Alam. 2020a. Anomaly detection with Local Outlier Factor (LOF). https://towardsdatascience.com/anomaly-detection-with-local-outlier-factor-lof-d91e41df10f2

Mahbubul Alam. 2020b. Isolation Forest: A Tree-based Algorithm for Anomaly Detection. https://towardsdatascience.com/isolation-forest-a-tree-based-algorithm-for-anomaly-detection-4a1669f9b782

Mahbubul Alam. 2020c. Support Vector Machine (SVM) for Anomaly Detection. https://towardsdatascience.com/support-vector-machine-svm-for-anomaly-detection-73a8d676c331

Estefania Artigao, Sergio Martín-Martínez, Andrés Honrubia-Escribano, and Emilio Gómez-Lázaro. 2018. Wind turbine reliability: A comprehensive review towards effective condition monitoring development. *Applied Energy* 228 (Oct. 2018), 1569–1583. https://doi.org/10.1016/j.apenergy.2018.07.037

H Badihi, YM Zhang, B Jiang, P Pillay, and S Rakheja. 2022. A Comprehensive Review on Signal-Based and Model-Based Condition Monitoring of Wind Turbines: Fault Diagnosis and Lifetime Prognosis. *PROCEEDINGS OF THE IEEE* 110, 6 (June 2022), 754–806. https://doi.org/10.1109/JPROC.2022.3171691

Anne Barros. 2021. *TPK4450 - Data Driven Prognostic and Predictive Maintenance* (1st ed.). Norwegian University of Science and Technology, Department of Mechanical and Industrial Engineering, Trondhem, Norway.

Lucas Bauer and Silvio Matysik. 2023. Enercon E-70 E4 2.300. https://en.wind-turbine-models.com/turbines/69-enercon-e-70-e4-2.300

Mogens Blanke and Jochen Schröder (Eds.). 2006. *Diagnosis and fault-tolerant control* (2nd ed ed.). Springer, Berlin ; New York. OCLC: ocm71336524.

Rebecca Callister. 2021. Maintenance strategy for NES (REN) Operated Assets 2021-2024.

Leoni Christensen. 2022. [Update 2022] OPEX Benchmark – An insight into the operational expenditures of European offshore wind farms. https://peak-wind.com/insights/opex-benchmark-an-insight-into-operational-expenditures-of-european-offshore-wind-farms/

Christopher J Crabtree, Donatella Zappalá, and Peter J Tavner. 2014. Survey of Commercially Available Condition Monitoring Systems for Wind Turbines. (2014).

Cuong Dao, Behzad Kazemtabrizi, and Christopher Crabtree. 2019. Wind turbine reliability data review and impacts on levelised cost of energy. *Wind Energy* 22, 12 (Oct. 2019), 1848–1871. https://doi.org/10.1002/we.2404 _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/we.2404.

Offshore Wind Energy. [n. d.]. Offshore Wind Energy. https://theliquidgrid.com/marine-clean-technology/offshore-wind/

S. Faulstich, B. Hahn, and P. J. Tavner. 2011. Wind turbine downtime and its importance for offshore deployment. *Wind Energy* 14, 3 (April 2011), 327–337. https://doi.org/10.1002/we.421

Alfonso Fernandez. 2017. Rolling element bearing components and failing frequencies. https://power-mi.com/content/rolling-element-bearing-components-and-failing-frequencies

Adam Filion. 2019. Predictive Maintenance: Unsupervised and Supervised Machine Learning. https://www.youtube.com/watch?v=AS0H43hMoWM

Katharina Fischer and Diego Coronado. 2015. Condition monitoring of wind turbines: State of the art, user experience and recommendations. *VGB PowerTech* 2015 (July 2015), 51–56.

Harriet Fox, Ajit C. Pillai, Daniel Friedrich, Maurizio Collu, Tariq Dawood, and Lars Johanning. 2022. A Review of Predictive and Prescriptive Offshore Wind Farm Operation and Maintenance. *Energies* 15, 2 (Jan. 2022), 504. https://doi.org/10.3390/en15020504

Zhiwei Gao and Xiaoxu Liu. 2021. An Overview on Fault Diagnosis, Prognosis and Resilient Control for Wind Turbine Systems. *Processes* 9, 2 (Feb. 2021), 300. https://doi.org/10.3390/pr9020300

Guide_to_an_offshore_wind_farm. 2019. *Guide to an offshore wind farm*. Guide. The Crown Estate and the Offshore Renewable Energy Catapult. 128 pages. https://guidetoanoffshorewindfarm.com/

Xiaofei He, Deng Cai, and Partha Niyogi. 2005. Laplacian score for feature selection. In *Proceedings of the 18th International Conference on Neural Information Processing Systems (NIPS'05)*. MIT Press, Cambridge, MA, USA, 507–514.

Halvor Hoen Hersleth, Vivek Mahajan, Laura Victoria Mejia, and Anders Ystad. 2021. *NES Offshore Wind Sessions*. Pre-read. Equinor. 26 pages.

Mark Hutchinson and Feng Zhao. 2023. *GWEC | Global Wind Report 2023*. Technical Report. GWEC. 120 pages.

ISO_10816-21:2015. 2015. Mechanical vibration — Evaluation of machine vibration by measurements on non-rotating parts — Part 21: Horizontal axis wind turbines with gearbox. https://www.iso.org/standard/60328.html

ISO_13372:2012. 2012. Condition monitoring and diagnostics of machines — Vocabulary. https://www.iso.org/standard/52256.html

ISO_2041:2018. 2018. Mechanical vibration, shock and condition monitoring — Vocabulary. https://www.iso.org/standard/68734.html#:~:text=ISO%20-%20ISO%202041%3A2018%20-%20Mechanical%20vibration%2C%20shock,areas%20of%20mechanical%20vibration%2C%20shock%20and%20condition%20monitoring.

Zakaria Jaadi. 2023. Principal Component Analysis (PCA) Explained |. https://builtin.com/data-science/step-step-explanation-principal-component-analysis

Abbas Karzon, Kjersti Drugli, Marthe Solvåg, Roy Kevin Wang, and Ingrid Hynne. 2022. Bruk av vibrasjonssensorer i prediktivt vedlikehold.

Steve Katz. 2021. Common Causes of Bearing Failure and How to Avoid Them. https://blog.emersonbearing.com/blog/common-causes-of-bearing-failure-and-how-to-avoid-them

Nia Kihlström. 2019. Wind turbine main shaft bearing design considerations. https://www.skf.com/group/news-and-events/news/2019/2019-10-08-wind-turbine-main-shaft-bearing-design-considerations

Vasudeva Kilaru. 2022. One Class Classification Using Support Vector Machines. https://www.analyticsvidhya.com/blog/2022/06/one-class-classification-using-support-vector-machines/

James Litchwark. 2023. Good Vibrations: Controlling creping blade vibration for better day-to-day and long-term performance. https://www.tissuestory.com/2019/09/03/controlling-creping-blade-vibration-for-better-day-to-day-and-long-term-performance/

Fei Tony Liu, Kai Ming Ting, and Zhi-Hua Zhou. 2008. Isolation Forest. In *2008 Eighth IEEE International Conference on Data Mining*. 413–422. https://doi.org/10.1109/ICDM.2008.17 ISSN: 2374-8486.

John Martin Gill. 2021. The Parts of a Wind Turbine: Major Components Explained. https://energyfollower.com/parts-of-a-wind-turbine/ Section: Wind.

Measuring_Vibration_with_Accelerometers. 2023. Measuring Vibration with Accelerometers. https://www.ni.com/en-no/shop/data-acquisition/sensor-fundamentals/measuring-vibration-with-accelerometers.html

Emil Meland and Mats Erik Haugan. 2021. *Bruk av vibrasjonsmonitoreringssystem for å avdekke begynnende degradering av generatorlager i vindturbin.* Bachelor Thesis. NTNU, Trondheim. https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/2781784

Mobius_Institute. 2016. *Vibration Analysis Training Manual ISO Category ll* (4.0 ed.). Mobius Institute. https://www.mobiusinstitute.com/product/vibration-analysis-iso-category-ii-online/

Mengyan Nie and Ling Wang. 2013. Review of Condition Monitoring and Fault Diagnosis Technologies for Wind Turbine Gearbox. *Procedia CIRP* 11 (Jan. 2013), 287–290. https://doi.org/10.1016/j.procir.2013.07.018

NVE. [n. d.]. Bessakerfjellet Vindkraftverk. https://www.nve.no/energi/energisystem/vindkraft/utbygde-vindkraftverk/vindkraftverk/?id=10413

Jesús María Pinar Pérez, Fausto Pedro García Márquez, Andrew Tobias, and Mayorkinos Papaelias. 2013. Wind turbine reliability analysis. *Renewable and Sustainable Energy Reviews* 23 (July 2013), 463–472. https://doi.org/10.1016/j.rser.2013.03.018

Selva Prabhakaran. 2019. Mahalanobis Distance - Understanding the math with examples (python). https://www.machinelearningplus.com/statistics/mahalanobis-distance/

Marvin Rausand, Anne Barros, and Arnljot Hoyland. 2020. *System Reliability Theory* (3rd ed. ed.). Wiley.

Shailesh Shukla. 2022. Outliers Detection Using IQR, Z-score, LOF and DBSCAN. https://www.analyticsvidhya.com/blog/2022/10/outliers-detection-using-iqr-z-score-lof-and-dbscan/

SKF_Bearing_Basics. 2023. Bearing basics. https://www.skf.com/us/products/rolling-bearings/principles-of-rolling-bearing-selection/general-bearing-knowledge/bearing-basics

SKF_Components_and_materials. 2023. Components and materials. https://www.skf.com/us/products/rolling-bearings/principles-of-rolling-bearing-selection/general-bearing-knowledge/bearing-basics/components-and-materials

SKF_Vibration_Diagnostic_Guide. 2011. *Vibration Diagnostic Guide*. SKF Reliability Systems, San Diego, California 92123 USA. https://skftechnicalsupport.zendesk.com/hc/en-us/article_attachments/360042513054/CM5003_EN_Vibration_Diagnostic_Guide_050511.pdf

Tom Nåstad Svennevig. 2022. Guidelines for Predictive and Condition Based Maintenance.

Tampen. 2021. Hywind Tampen approved by Norwegian authorities - equinor.com. https://www.equinor.com/news/archive/2020-04-08-hywind-tampen-approved

Hywind Tampen. 2023. Hywind Tampen. https://www.equinor.com/energy/hywind-tampen

Peter Tavner. 2021. *Offshore Wind Power: Reliability, availability and maintenance* (2nd ed. ed.). The Institution of Engineering and Technology.

Wind Fuels & Technologies. 2022. Wind - Fuels & Technologies. https://www.iea.org/fuels-and-technologies/wind

Tharanath Thiruthiyappan. 2022. *Predictive Maintenance for Offshore Wind Farms*. Specialization project. Norwegian University of Science and Technology, Trondheim. 111 pages.

Jason Tranter. 2016. Vibration Analysis - Bearing Failure Analysis by Mobius Institute. https://www.youtube.com/watch?v=dEn2Qvh_qjc

Alan Turnbull, James Carroll, Sofia Koukoura, and Alasdair McDonald. 2019. Prediction of wind turbine generator bearing failure through analysis of high-frequency vibration data and the application of support vector machine algorithms. *The Journal of Engineering* 2019, 18 (2019), 4965–4969. https://doi.org/10.1049/joe.2018.9281 _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1049/joe.2018.9281.

Christian Tutivén, Yolanda Vidal, Andres Insuasty, Lorena Campoverde-Vilela, and Wilson Achicanoy. 2022. Early Fault Diagnosis Strategy for WT Main Bearings Based on SCADA Data and One-Class SVM. *Energies* 15, 12 (Jan. 2022), 4381. https://doi.org/10.3390/en15124381

Unsupervised_Anomaly_Detection. 2023. Unsupervised Anomaly Detection. https://se.mathworks.com/help/stats/unsupervised-anomaly-detection.html

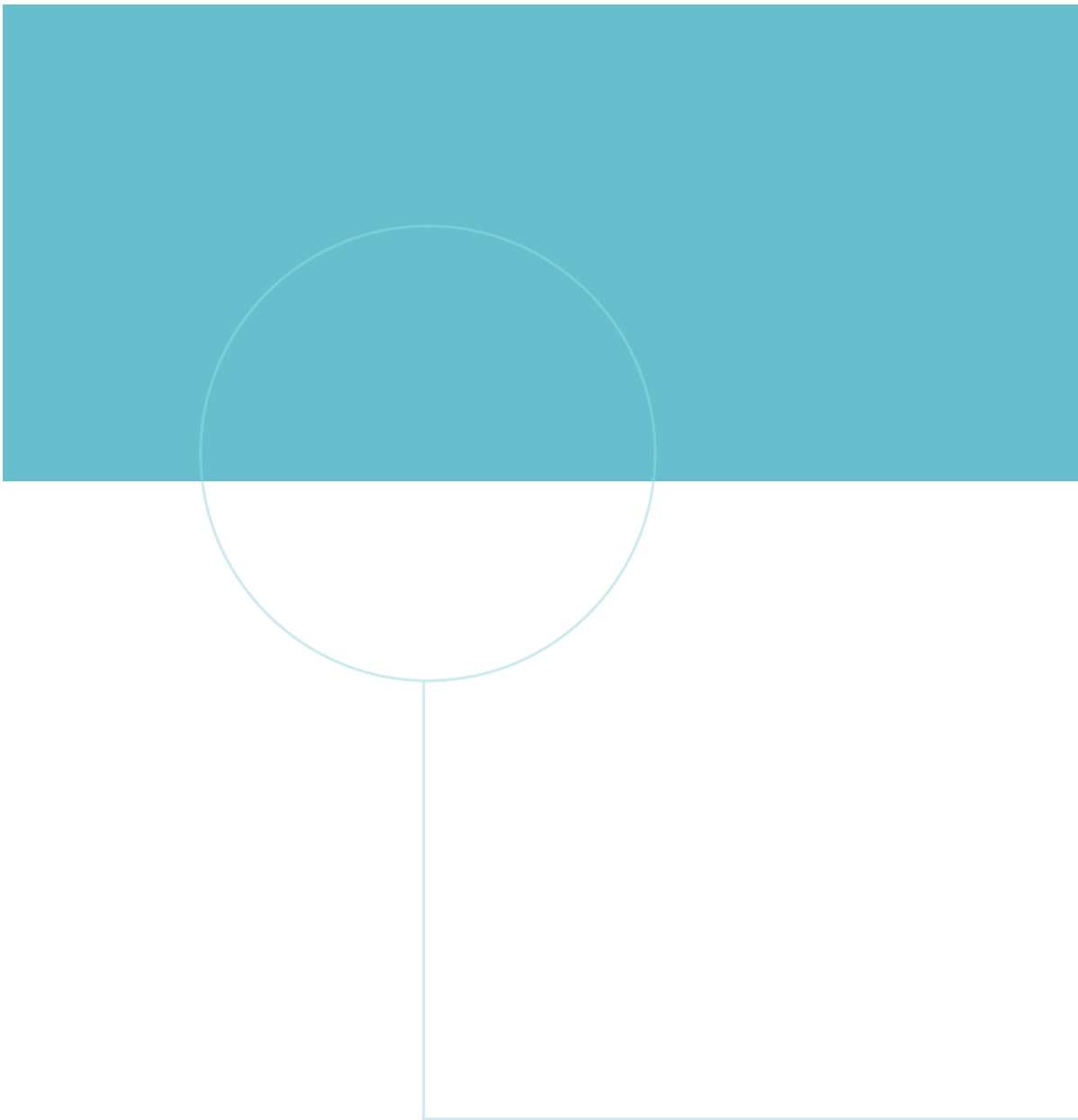
Karel Vermeiren and Torsten Bark. 2011. Enveloped Acceleration with the SKF Machine Condition Transmitter (MCT). https://skftechnicalsupport.zendesk.com/hc/en-us/articles/360033075433-Enveloped-Acceleration-with-the-SKF-Machine-Condition-Transmitter-MCT-

Haiming Wang, Qiang Li, Yongqiang Liu, and Shaopu Yang. 2022. Anomaly Data Detection of Rolling Element Bearings Vibration Signal Based on Parameter Optimization Isolation Forest. *Machines* 10, 6 (June 2022), 459. https://doi.org/10.3390/machines10060459

Offshore wind. 2023. Offshore wind. https://www.equinor.com/energy/offshore-wind

Shen Yin. 2022. Digitalized Solutions to Prognosis, Predictive Maintenance and Safety Analysis. https://www.ntnu.edu/studies/courses/TPK4450#tab=omEmnet

Wanwan Zhang, Jørn Vatn, and Adil Rasheed. 2022. A review of failure prognostics for predictive maintenance of offshore wind turbines. *Journal of Physics: Conference Series* 2362, 1 (Nov. 2022), 012043. https://doi.org/10.1088/1742-6596/2362/1/012043 Publisher: IOP Publishing.



NTNU
Norwegian University of
Science and Technology

equinor