Wind turbine gearbox fault prognosis using high-frequency SCADA data

Ayush Verma[1], Donatella Zappalá[1], Shawn Sheng[2], Simon J. Watson[1]
[1]TU Delft, Kluyverweg 1, 2629 HS Delft, Netherlands
[2]National Renewable Energy Laboratory, 15013 Denver West Parkway | Golden, CO 80401

E-mail: avvermatudelft@gmail.com

Abstract. Condition-based maintenance using routinely collected Supervisory Control and Data Acquisition (SCADA) data is a promising strategy to reduce downtime and costs associated with wind farm operations and maintenance. New approaches are continuously being developed to improve the condition monitoring for wind turbines. Development of normal behaviour models is a popular approach in studies using SCADA data. This paper first presents a data-driven framework to apply normal behaviour models using an artificial neural network approach for wind turbine gearbox prognostics. A one-class support vector machine classifier, combining different error parameters, is used to analyse the normal behaviour model error to develop a robust threshold to distinguish anomalous wind turbine operation. A detailed sensitivity study is then conducted to evaluate the potential of using high-frequency SCADA data for wind turbine gearbox prognostics. The results based on operational data from one wind turbine show that, compared to the conventionally used 10-min averaged SCADA data, the use of high-frequency data is valuable as it leads to improved prognostic predictions. High-frequency data provides more insights into the dynamics of the condition of the wind turbine components and can aid in earlier detection of faults.
Keywords: Wind turbine, Gearbox failure prognostics, High-frequency SCADA data, Machine learning

1. Introduction

The offshore wind energy is expected to become one of the largest sources of renewable electricity by 2042 [1]. However, the sector faces challenges pertinent to operations and maintenance (O&M), which can make up a significant portion of the total costs associated with any wind farm, with up to 30% for offshore installations [2]. Wind turbine (WT) gearboxes are among the most expensive subsystems and their failures incur the longest downtime [3], therefore, this research focuses on prognostics of gearbox failures.

Condition-based maintenance is an effective strategy to reduce O&M costs and improve availability of WTs through condition monitoring. Commercially available purpose-based CM systems require additional monitoring devices to be installed on the WT. These devices measure and record large volumes of high-frequency data, including vibrations and oil debris measurements. However, their wide-scale deployment is still limited, as their economic benefit to O&M costs depends largely on the fault detection rate [4] and they only records component-specific information. On the other hand, WT condition monitoring using supervisory control and data acquisition (SCADA) systems is seen as a cost-effective and wide-scale approach, as these are pre-installed systems in utility-scale WTs. SCADA systems are rich data sources that usually record 10-min averaged signals which characterize the operational and environmental conditions of the WT [5]. This has motivated the scientific community to investigate their use for CM purposes. When compared with condition monitoring systems, an important drawback of SCADA data is low temporal resolution, which results in a loss of information about the dynamic behaviour of the WTs and can lead to a loss of detection capabilities during the monitoring phase [6]. Nevertheless, SCADA data are a potential solution for WT CM owing to their availability at no additional cost [7]. Furthermore, using high-frequency SCADA data instead of 10-min averaged signals, this research investigates in detail the potential of using high-frequency SCADA data for WT prognostics.

Various recent studies reviewed in [8] have shown that multiple approaches and machine-learning models can be used for condition monitoring of several wind turbine components using SCADA data. One popular approach in studies using SCADA data is through the development of normal behaviour models (NBM). NBMs are based on the idea of modelling the normal (or healthy) behaviour of a WT component and comparing the model predictions with in-field measurements to track error between the signals and detect fault inception. Mathematical modelling methods like artificial neural networks (ANN) can be utilised to develop NBM to analyse WT data [9]. ANNs are advantageous in monitoring any component without an in-depth knowledge of its working principles and are efficient at modelling nonlinear complex systems [10].

Zaher *et al.* [11] used wind turbine field data to train an auto-regressive ANN model to detect failures in gearbox bearings. The anomalies were detected based on analysing the error between the predicted and measured target parameters. Such a method of fault detection is prone to errors, as it does not take into consideration the inherent randomness in the ANN predictions, and a threshold value is desirable to generate automatic alarms. Bangalore *et al.* [12] used a nonlinear auto-regressive with exogenous input ANN configuration to predict the gearbox bearing and lubrication oil temperature. They showed that postprocessing of the error signal using a statistical approach based on the Mahalanobis distance, can be used to improve confidence in the anomaly detection process. Other metrics, such as using simple thresholds that can be set based on the training root-mean-squared-error (RMSE) to determine anomaly rates or calculating health degree based on probability, have also been suggested in the literature [13] [14]. However, these methods tend to focus on only one parameter to describe the error signal, which over any period of time is multifaceted and has a unique error distribution in that interval. An effective way of addressing these limitations and analysing the error signal is to combine different characteristics of its distribution over a selected time interval in order to set a more robust threshold for detecting anomalies. The advantage of such a method has also been demonstrated by Turnbull *et al.* [15].

This research proposes a data-driven framework that utilises a feed-forward ANN-based NBM to predict gearbox oil temperature and a one-class SVM (OC-SVM) classifier to combine multiple error features to set a threshold for anomaly detection through a complex decision boundary. Furthermore, a sensitivity study based on different SCADA sampling frequencies is carried out to better understand the loss of information due to the data averaging effect. This paper demonstrates the use of high-frequency SCADA data for performing WT gearbox prognostics for potentially performing early fault detection compared to the 10-min averaged data.

2. Methodology

2.1. Data overview

The data set used in this research was collected and shared from the two-bladed Control Advanced Research Turbine (CART2) on the Flatirons Campus of the National Renewable Energy Laboratory (NREL), Colorado, USA. The CART2 turbine has a rotor diameter of 43.3 m and reaches a rated generator power of 600 kW at a wind speed of 11 m/s. The turbine was outfitted with a total of 88 sensors recording measurements including pitch angles, shaft torque and rotational speed, oil temperature and pressure, power output and other control signals [16]. The sensors installed in the CART2 turbine are shown in the Figure 1.

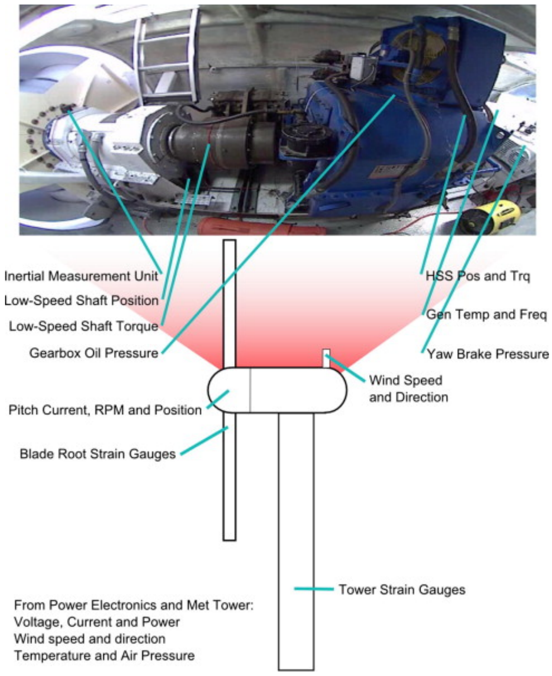


Figure 1: CART2 sensor location [16]

The CART2 turbine experienced a gearbox failure on April 7, 2009, which led to a turbine shutdown. The root cause analysis suggested that the failure was due to the misalignment of gear teeth in the first stage of the gearbox (planetary gear failure). The gearbox was then replaced in the summer of 2009 [17]. Compared to data sets usually used in the literature for the analysis of gearbox failure prognostics, the CART2 data set is unique as the data were recorded at a frequency of 100 Hz rather than 10-min averaged values, which is the conventional industry practice.

2.2. Framework for gearbox failure prognostics

A NBM-based framework was developed for gearbox prognostics and is shown in Figure 2. The entire framework can be segmented into three modules: Module 1 - data preprocessing, Module 2 - normal behaviour model, and Module 3 - anomaly detection and prognosis. It is executed in two phases. First phase (training and validation) encompasses training, testing and validating the employed machine learning models. The trained models are then used to detect anomalies and this characterizes the second phase (application). The training and validation phase is a one-time process, which is done using data representative of the healthy/normal state of the turbine whereas the application phase represents the continuous application for anomaly detection and condition monitoring.

2.2.1. Module 1: Data preprocessing The first task consists in to distinctly categorize the data corresponding to the turbine's healthy and faulty operation."Healthy" instances are data representing the normal behaviour of the turbine (i.e., when no signs of failure were observed and are used to model the WT's normal state). "Faulty" instances comprise data indicating deterioration of the machine component before the fault occurs, hence, containing information of the fault initiation and development. Faulty data is then compared with the NBM output to

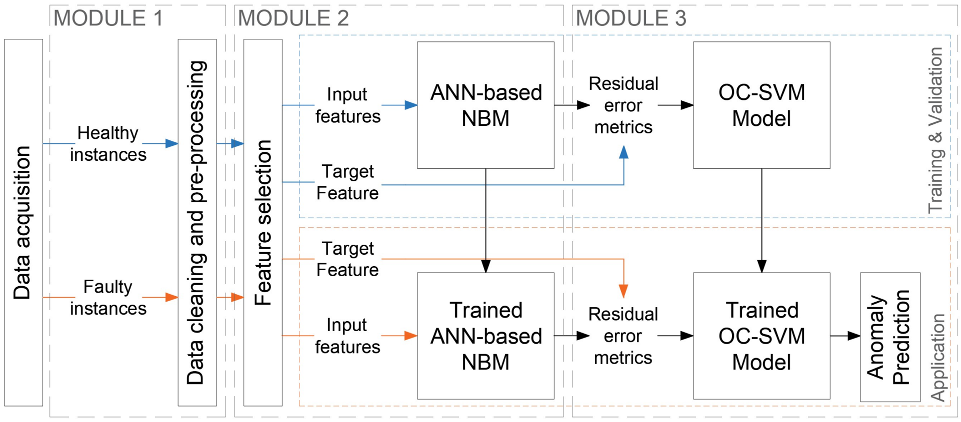


Figure 2: Model architecture for gearbox failure prognostics

identify the probable component degradation by observing the deviation from normal behaviour. SCADA data for 4 months up to the known date of failure were denoted as faulty data and all recorded operational data before this period were identified as healthy data. A 4-month time period was selected, as this provided an appropriate balance between the two data sets and sufficient timeline to understand the progression of the WT gearbox failure.

Because the CART2 turbine was installed to perform advanced wind turbine control research, the data were recorded only when the field tests were conducted [17]. To tackle the issue of discontinuous time-series data, the data set was clustered in weeks of operational data, as shown in Figure 3. The faulty data have been analysed in terms of "weeks to failure" which was done based on the known date of failure (7 April 2009).

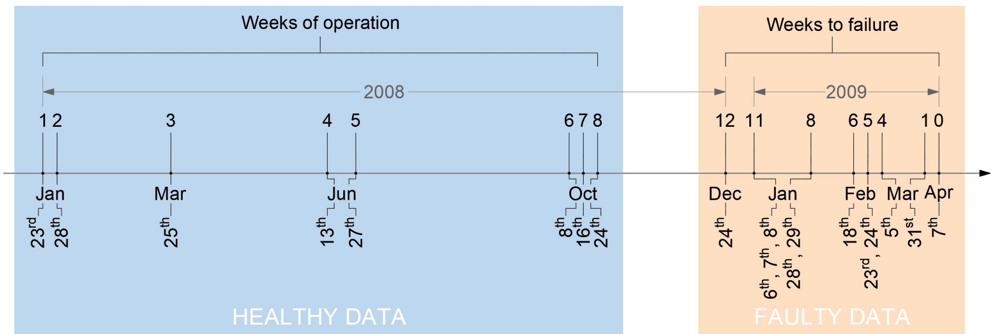


Figure 3: Weeks of operation and corresponding dates in which the data was available for healthy data (blue) and weeks to failure and corresponding dates for faulty data (orange)

Training the machine-learning models with low quality abnormal data might lead to inaccuracies in the model predictions; therefore, it is imperative to filter out the outliers. For the initial phase of model development and testing, the 100-Hz signals were resampled to 1-Hz. This was done to tackle the issue of noise in the high-frequency raw data while keeping their failure features.The down-sampled data set were cleaned of any data corresponding to curtailment or downtime, blade pitch anomalies, and erroneous sensor measurements. Additionally, any power measurement values recorded either below the cut-in or above the cut-out wind speed were filtered out.

A common issue when working with real-world data sets is the imbalance in data distribution across different machine operational regimes which can affect regression tasks [18]. Such imbalance within the data set result in over-fitting for operational regimes with large volumes of data and under-fitting for regimes with relatively less data. This deters the generalization capability of the ML models. An imbalance in the healthy data set analysed in this work can be observed in Figure 4 for the case of healthy data. It can be seen that the original data distribution is left-skewed with a substantially higher number of data points referring to low power output values (i.e., between 0 - 100 kW; low power operational regime). In order to tackle this issue, data were preprocessed using a synthetic minority oversampling technique with introduction of Gaussian Noise (SMOGN) [19]. SMOGN combines a random undersampling strategy with a commonly known oversampling techniques: synthetic minority oversampling technique for regression. To make the oversampling technique more robust, the method adds interpolation using normally distributed Gaussian noise to generate new synthetic samples when the samples in the data set are too far from each other. The SMOGN implementation allowed for the low-power region (0–100 kW) to become undersampled whereas synthetic samples were generated in the high-power region (400–600 kW). There was no reduction in the number of data samples when the healthy data were preprocessed with SMOGN.

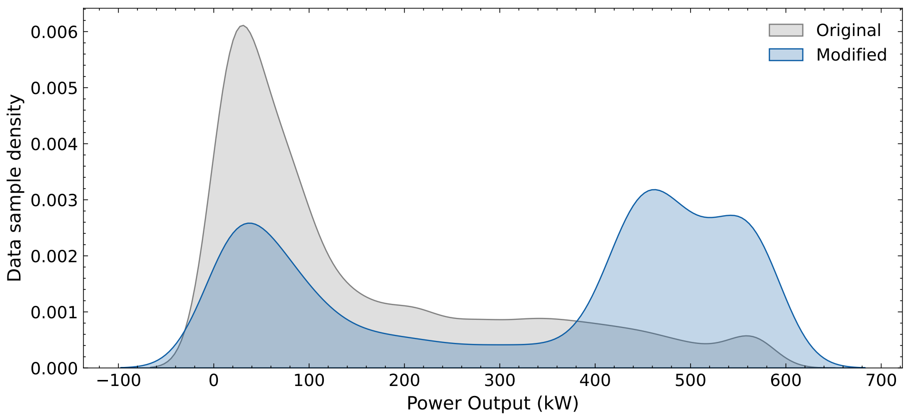


Figure 4: Data density of the original and modified (after SMOGN application) healthy data set

2.2.2. Module 2: Normal behaviour model The ANN-based NBM takes a variety of input features to predict a target feature. Those features have been selected based on the understanding of the gearbox physics. The gearbox lubrication oil temperature was selected as the target feature, among the available signals, because it is the most sensitive to the condition of the gearbox; hence, any deviations from healthy conditions would manifest themselves into anomalies in temperature values [20]. The input features include signals describing the environmental conditions (wind speed, and ambient temperature) and the WT operating conditions (power output, high-speed shaft rotation speed, and nacelle temperature). In summary, the NBM utilises five input features to predict one target feature as summarized in Table 1. In order to model the NBM, a three-layer feed-forward neural network was built to predict the target feature from the five input features. The key characteristics of the ANN model are described in Table 2.

Out of the available samples in the healthy data, 70% were selected for training the ANN model and 15% for validation, with the remaining 15% used to test the model independently. Such a split provides a good balance to compare the model performance in all phases [21]. Once the neural network was trained and optimized, the same input features from the faulty data set were then fed into the ANN model to predict the gearbox lubrication oil temperature.

Table 1: Model features for the ANN-based NBM

Features	Description	Layer
U_{wind}	Wind speed	Input layer
P_{out}	Power output	Input layer
U_{Hss}	High-speed shaft rotation speed	Input layer
T_{nac}	Nacelle temperature	Input layer
T_{amb}	Ambient temperature	Input layer
T_{gb}	Gearbox lubrication oil temperature	Output layer

Table 2: Key characteristics of the ANN model

Attributes	Value
Number of input neurons	5
Number of hidden neurons	343
Number of layers	3 (1 input, 1 hidden, and 1 output layer)
Loss function	Adaptive moment estimation (Adam)
Learning rate	0.001 (constant)

2.2.3. Module 3: Anomaly detection and prognosis After training and optimizing the ANN, the NBM predictions are compared with the field measurements to track anomalies and determine a robust threshold that can sufficiently distinguish between normal and anomalous behaviour. A OC-SVM classifier was used to combine multiple error features and set complex boundaries to describe the fault threshold. The error features which best described the unique error distribution over a given period of time were selected. The approach of using multiple error features to set a threshold to detect anomalies mitigates the risk of capturing the fluctuations in the ANN model prediction.In the study presented in this paper, 4 parameters–namely, the root-mean-squared-error (RMSE), the minimum error, the maximum error, and the standard deviation of error distribution–are computed over a period of 1 minute and used as inputs to the OC-SVM model. Using the healthy data set, the OC-SVM model is trained to recognise up to 1% of data as anomalies in the training period and, therefore, a similar percentile would be expected moving forward in absence of any fault in the system.

3. Results and Discussion

3.1. ANN model performance

The ANN performance was evaluated for each phase of the model development (training, validation and testing) using the RMSE and the R-squared values (R^2). The low values of RMSE ranging between 0.03 - 0.06 for different model development phases demonstrate that the predicted values of the target feature are indeed close to the actual values. The R^2 values close to 1 (0.95 - 0.97) and their 1.7% deviation between data sets demonstrate that the model fits well and generalises relatively well. Figure 5 compares the observed and ANN-predicted values of the gearbox oil temperature for the healthy data set. As a consequence of the SMOGN preprocessing the model performs equally well for periods of low and high operational power. The maximum deviation is observed when there is a sudden change in output power of the turbine, due to the variability in wind speed.

3.2. Anomaly detection using OC-SVM

Once both models–ANN and OC-SVM–were trained with the healthy data, all the steps in the framework were implemented with the faulty data set and the residual error over the 4 months

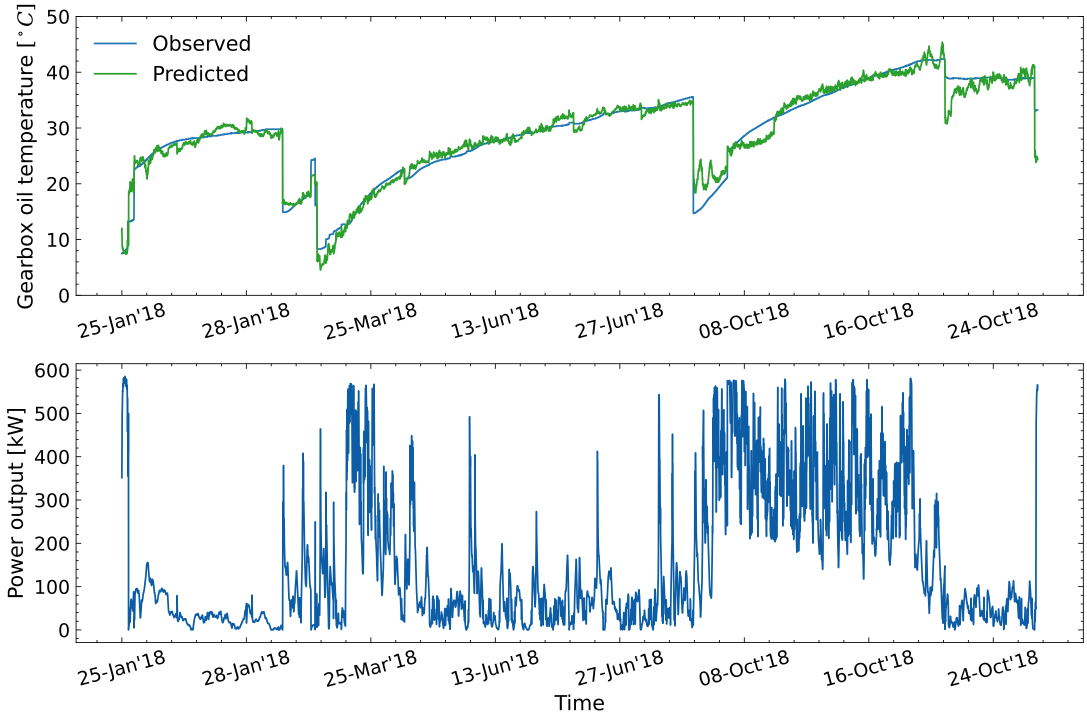


Figure 5: (Top) Observed and predicted gearbox oil temperature values (Bottom) Power output values in the case of healthy data set

leading up to the failure was evaluated. For each 1-min time step, the error features, introduced in section 2.2.3 were calculated and fed into the OC-SVM model. Figure 6 shows the maximum error plotted against the RMSE for both healthy and faulty data. The OV-SVM classifier is able to learn the boundary of healthy operation and gives a prediction value of +1 if a data point falls within it. Any data point far away from this boundary is assigned a value of -1 and is interpreted as representative of anomalous turbine operation. Figure 7 shows the data classification ("normal operation" vs. "anomaly") done by the OC-SVM for the faulty data set.

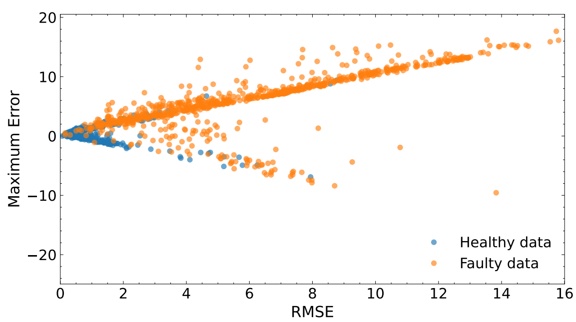


Figure 6: Maximum error vs RMSE (Healthy and Faulty data)

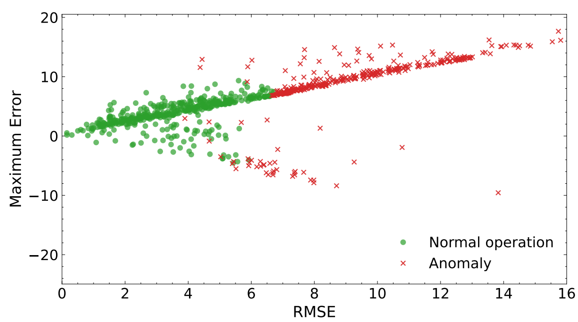


Figure 7: Anomalous and normal operation points classified by OC-SVM (Faulty data)

The percentage of anomalies in each week of turbine operation leading up to failure was then calculated by considering the ratio between the number of data points classified as anomaly by OC-SVM and the total number of data points recorded in that respective week. Figure 8 shows the progression of the anomaly percentage as the turbine gearbox approaches failure. The results show a gradual increase in the number of anomalies with a sharper change one month before failure (4 weeks to failure) and, thereafter, an increasing trend showing 100% anomalies detected on the day of failure. Between 12 and 5 weeks to failure, the percentage of anomalies detected increase on average by 8% with a substantial increase of 32% between 5 and 4 weeks before the failure. In terms of detection time, the first consistent increase in anomalies is seen almost 8 weeks before failure, providing a lead time of 2 months to plan and execute maintenance activities. The sharp increase in percentage of anomalies 4 weeks before failure shows that the fault has become more severe and urgent action is required to prevent turbine shutdown.

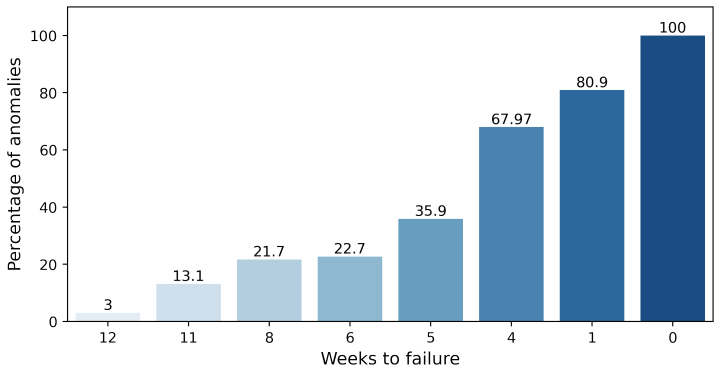


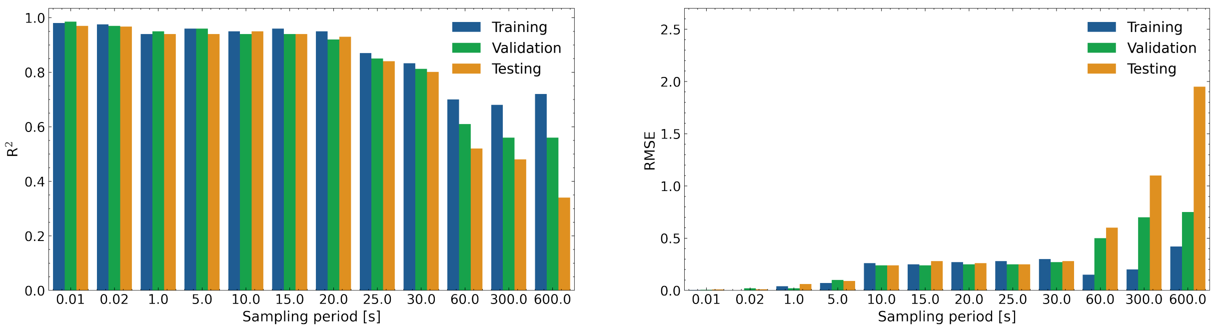
Figure 8: Percentage of anomalies detected in each week leading up to the failure

3.3. Sensitivity study

The key contribution of this paper lies in the use of high-frequency SCADA data rather than the routinely used 10-min averaged data. The use of high-frequency SCADA data introduces additional noise in the signals which is usually smoothed by the averaging process over conventionally used 10-min periods. However, the data averaging also leads to a loss of information that might lead to false/missed alarms. To investigate the potential of using high-frequency data, a sensitivity study was performed. The framework as described in Section 2.2 was implemented using data sampled at 12 different rates: 0.01 s (100 Hz), 0.02 s (50 Hz), 1 s (1 Hz), 5 s, 15 s, 20 s, 25 s, 30 s, 60 s, 300 s, 600 s.

The first objective of the sensitivity study was to investigate how the ANN performance varies with the data sampling period. This has been done by calculating two metrics, R^2 and RMSE. Results are shown in Figure 9a and Figure 9b, respectively. With R^2 values ranging between 0.9 and 1, the model performs well for sampling periods below 30 s but deteriorates for sampling periods higher than a minute. A similar trend can be observed in the RMSE values in Figure 9b, where there is a considerable increase for sampling periods greater than 30 s. Additionally, it can be seen that the model performance does not vary considerably for high sampling frequencies. This trend in model performance is a consequence of two main factors: (1) the gearbox oil temperature is a slow changing parameter and (2) the performance improvement of a configured ANN model plateaus after a certain dataset size [22]. This poses a disadvantage in using SCADA data sampled at such high frequencies (0.01 s and 0.02 s), as storing and operating large data sets would require extensive resources and computational power with no major improvements in the ANN model performance.

Another interesting point to observe in Figure 9 is how the values of the two metrics vary with the model development phase in the case of sampling periods higher than 30 s. The metrics for the model training phase being considerably higher, in the case of R^2 and lower, in the case of RMSE than those for the other two phases is an indication of model overfitting the training data



(a) ANN model performance: R^2 (b) ANN model performance: RMSE

Figure 9: ANN performance metrics for different sampling frequencies

and its inability to generalize its learning to the testing data set. This is due to the lower number of data samples available for model training for the case of these longer sampling periods. This limitation stems from the CART2 data set itself, as the turbine was only operated in periods when testing and research were conducted, resulting in a low number of operational hours and a small amount of data samples when aggregated for higher sampling periods.

The second objective was to assess how the trend in percentage of anomalies detected in the WT operational period before failure varies with the data sampling frequency. Figure 10 shows the heat map of the percentages of anomalies that were detected by OC-SVM for each week before the failure for all different sampling rates. As we approach the failure, we can observe an increasing trend in the percentage of anomalies detected, progressing to 100% for the week of failure for all sampling periods except for the longer sampling intervals of 60 to 600 s. In fact, the percentage of anomalies detected when using data with these high sampling periods is extremely low for any week before failure, with almost no signs of failure until 5 weeks before failure. This could be consequence of several factors such as: (1) loss of information about the condition of the gearbox because of the data averaging effects; (2) poor ANN model performance due to a low number of data samples, which leads to model overfitting to training data; (3) high range of prediction error values, which could increase the OC-SVM threshold boundary, resulting in anomalies being misclassified as normal operational points.

In contrast, the percentage of anomalies observed when using data sampled at high frequency of 0.01 and 0.02 s is quite high, even 12 weeks before the failure. This can be accredited to the fact that there is more information available due to the sampling frequency of the data. However, considering that this data often entails higher noise, such high percentages could also be attributed to the outliers/noise in the signals being misclassified as anomalies. In the latter case, this might even lead to false alarms. Additionally, it can observed that the percentage of anomalies decreases with increasing sampling periods with slight oscillations of around 5-6% in the case of 8 to 6 weeks before failure. Such oscillations could be due to various reasons, such as the data resampling, and the ANN and/or the SVM model performance. Lastly, the data aggregated for 30 s periods shows the sudden increase in the percentage of anomalies progressing from 5 to 4 weeks before failure, as was noted with the 1-Hz data. This indicates that despite the averaging effect, the aggregated data still retain information about the gearbox fault that can be extracted efficiently through the framework, and a maintenance alarm could be triggered a month in advance.

4. Conclusion

In this research, the potential for using high-frequency SCADA data for the purpose of WT condition monitoring has been thoroughly investigated. A framework combining ANN-based

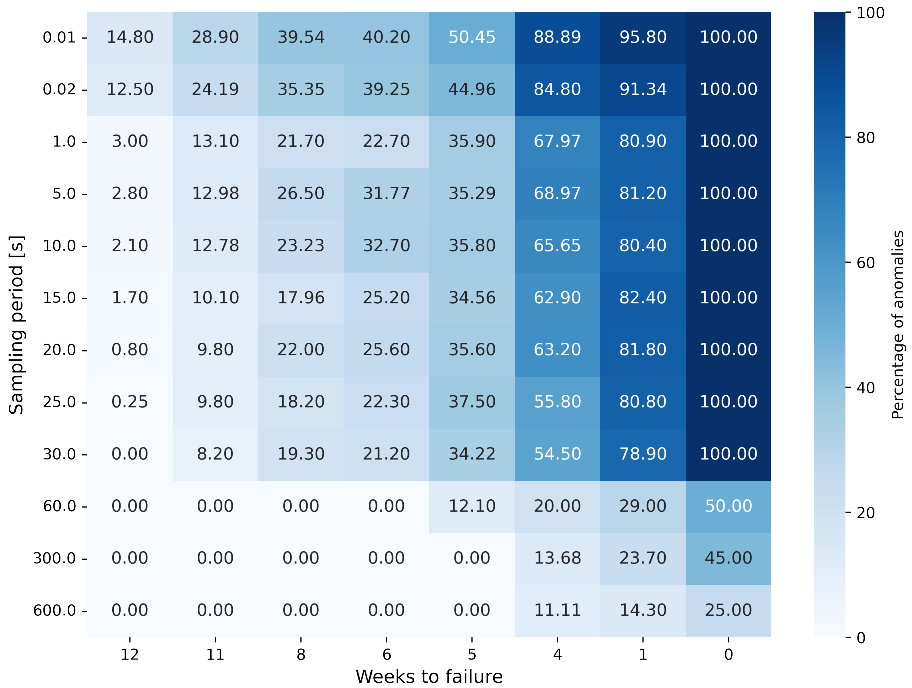


Figure 10: Heat map of the percentage of anomalies detected each week before failure for different SCADA sampling rates

NBM and OC-SVM was developed for WT gearbox failure prognostics. It was shown that SMOGN is an effective data preprocessing approach that can be employed to tackle the issue of imbalanced operational regimes when working with real-world data sets and improves the generalization capabilities of ANN. Through the use of an OC-SVM classifier, multiple metrics were combined to describe the error distribution over a selected time period, and a complex threshold boundary was formed to distinguish between normal and anomalous WT operation. The results confirm the effectiveness of this framework and its usefulness for monitoring the condition of WT gearboxes. The sensitivity study demonstrates that the use of high-frequency data provides a deeper understanding of WT condition and improves the detection capabilities of the suggested approach. Nevertheless, it was also seen that the issue of noise in data becomes more prominent with very high sampling rates, which can lead to misclassified anomalies, resulting in false/missed alarms. Such a study can be used to determine the optimal frequency for recording SCADA data to build effective predictive maintenance strategies that can forestall high O&M costs. Unfortunately, this research could not derive any definitive conclusions due to the limitations posed by the CART2 data set wherein there were too few data samples available for analysis for sampling periods over 30 s. Determining the optimal SCADA data frequency would require further research and performing trade-off analysis of key factors such as the implementation costs, the data storage capacities, the data analysis complexity and its applicability for early fault detection.

Although this research demonstrates the advantage of using SCADA data to predict gearbox failure, it should be noted that the root cause of the failure, which in this case was gear teeth misalignment, could not be derived using this approach. This is because SCADA systems only record general information about the WT rather than component-specific data, which are obtained through condition monitoring systems. However, this framework can still potentially identify any gearbox fault that directly or indirectly results in a rise in the oil temperature values such as gear misalignment, shaft misalignment, wearing of component parts, etc. Furthermore, even though this framework can be tailored to give early fault detection results for other gearbox related faults, to prove its efficacy further testing with different failure modes and turbines should be carried out. Future work should focus on developing a combination of physical and data-driven models to understand the failure causes in WT components and derive the root cause of occurred faults. Moreover, a more comprehensive high-frequency data set could be used to advance the sensitivity study presented in this paper to determine an optimal SCADA data sampling frequency that could be used for wind turbine condition monitoring.

Acknowledgements

This work was authored [in part] by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Wind Energy Technologies Office. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes. We also acknowledge the wind turbine control research team at NREL for their data collection and data sharing support.

References

[1] Wind Europe 2019 Our energy, our future: How offshore wind will help europe go carbon-neutral tech. rep. WindEurope Association
[2] Crabtree C J, Zappalá D and Hogg S I 2015 *Proceedings of the Institution of Mechanical Engineers, Part A* **229** 727–746
[3] Reder M D, Gonzalez E and Melero J J 2016 *J. Phys.: Conf. Ser.* **753** 072027
[4] May A, McMillan D and Thöns S 2015 *IET Renewable Power Generation* **9** 900–907
[5] Yang W, Court R and Jiang J 2013 *Renewable Energy* **53** 365–376
[6] Yang W, Tavner P J, Crabtree C J, Feng Y and Qiu Y 2014 *Wind Energy* **17** 673–693
[7] Wang K S, Sharma V S and Zhang Z Y 2014 *Advances in Manufacturing* **2** 61–69
[8] Maldonado-Correa J, Martín-Martínez S, Artigao E and Gómez-Lázaro E 2020 *Energies* **13** 3132
[9] Tautz-Weinert J and Watson S J 2017 *IET Renewable Power Generation* **11** 382–394
[10] Wilkinson M, Darnell B, Van Delft T and Harman K 2014 *IET Renewable Power Generation* **8** 390–397
[11] Zaher A, McArthur S, Infield D and Patel Y 2009 *Wind Energy: An International Journal for Progress and Applications in Wind Power Conversion Technology* **12** 574–593
[12] Bangalore P, Letzgus S, Karlsson D and Patriksson M 2017 *Wind Energy* **20** 1421–1438
[13] Zhang Z Y and Wang K S 2014 *Advances in Manufacturing* **2** 70–78
[14] Li J, Lei X, Li H and Ran L 2014 *Electric Power Components and Systems* **42** 1201–1212
[15] Turnbull A, Carroll J and McDonald A 2021 *Wind Energy* **24** 197–211
[16] Johnson K E and Fleming P A 2011 *Mechatronics* **21** 728–736
[17] Bossanyi E, Wright A and Fleming P 2010 *Controller field tests on the NREL CART2 turbine* ((No. NREL/TP-5000-49085) National Renewable Energy Lab.(NREL), Golden, CO (United States))
[18] Krawczyk B 2016 *Progress in Artificial Intelligence* **5** 221–232
[19] Branco P, Torgo L and Ribeiro R P 2017 *Proceedings of Machine Learning Research* **74** 36–50
[20] Feng Y, Qiu Y, Crabtree C J, Long H and Tavner P J 2013 *Wind Energy* **16** 728–740
[21] Crowther P S and Cox R J 2005 *International conference on knowledge-based and intelligent information and engineering systems* **3684** 1–7
[22] Guo Q, Jin S, Li M, Yang Q, Xu K, Ju Y, Zhang J, Xuan J, Liu J, Su Y *et al.* 2020 *Science China Earth Sciences* **63** 1457–1474