On the Joint Use of An Ensemble of Linear Residuals to Improve Fault Detection in Wind Turbines

Théodore Raymond, Sylvie Charbonnier, Alexis Lebranchu, Christophe Bérenguer

On the joint use of an ensemble of linear residuals to improve fault detection in wind turbines

Théodore Raymond
Data Engineering, Valemo, France. E-mail: theodore.raymond@valemo.fr

Sylvie Charbonnier
Université Grenoble Alpes, CNRS, Grenoble-INP, GIPSA-Lab, France.
E-mail: sylvie.charbonnier@grenoble-inp.fr

Alexis Lebranchu
Data Engineering, Valemo, France. E-mail: alexis.lebranchu@valemo.fr

Christophe Bérenguer
Université Grenoble Alpes, CNRS, Grenoble-INP, GIPSA-Lab, France.
E-mail: christophe.berenguer@grenoble-inp.fr

One of the biggest levers for reducing the cost of wind power generation is to minimize the replacement frequency of large components. To address this need, researchers have focused on the development of real-time health monitoring of component to perform condition-based maintenance. In a previous work, a fault detection solution based on multi-turbine indicators built from automatically generated linear models has been presented and validated on a converter fault case. However, the application of this method on other faults revealed weaknesses in the detection performance, making the solution unreliable. To address these issues, the solution proposed in this study is to consider an ensemble method to automatically generate a set of tri-variable linear models predicting the evolution of a common variable. The linear models are constructed using a constrained greedy selection algorithm, providing unique sets of model variables. From these models, residual-based multi-turbine health indicators are constructed, and a mean linear residual is considered, computed as the mean of seven different indicators. The comparative analysis of these indicators, carried out based on the area under two receiver operating characteristic curves on two fault cases, shows that the use of a mean linear residual computed from a set of linear residuals significantly improves the global detection performance, and thus the reliability of the condition-based maintenance process under development.

Keywords: Wind energy, Linear multi-variable models, Ensemble methods, Residual-based fault detection, Intelligent variable selection, Automatic model generation

1. Introduction

With the rapid expansion of French wind farms in recent years, the number of component replacement operations is increasing, resulting in a rise in the levelized cost of energy (LCOE). In order to reduce corrective maintenance operations (and thus reduce the LCOE), the solution developed in the literature consists in moving from a periodic maintenance based on a calendar to a condition-based maintenance (or CBM) based on a real-time monitoring of the components' health, allowing the planning of preventive maintenance operations, intervening at a minor stage of a fault, when the wind turbine is still in normal operation.

Most of the CBM methods proposed in the literature for wind turbines are based on supervisory control and data acquisition (or SCADA) data, whose sensors are cheap, in large numbers, and already installed on the various components of a wind turbine (Márquez (2020)). The health status is then modeled by a residual signal, measuring the deviation between a measured value and a value estimated by a model (Maldonado-Correa, Martín-Martínez, Artigao, and Gómez-Lázaro (2020), Tautz-Weinert and Watson (2017)).

In a previous work (Raymond, Lebranchu,

Bérenguer, and Charbonnier (2022)), a data-driven method to automatically generate linear model for the fault detection of wind turbine components was presented and applied to a converter fault case, and showed encouraging results. This method relies on a forward variable selection algorithm to generate simple linear models. This choice of modeling strategy is made because it allows robustness to industrial constraints, regarding data availability and noise. This approach also has the advantage of providing interpretable models, for a low implementation cost, compared to neural network modeling ones (Stetco, Dinmohammadi, Zhao, Robu, Flynn, Barnes, Keane, and Nenadic (2019), Schlechtingen and Ferreira Santos (2011))). The data-driven fault detection method proposed has been designed to be deployable at the scale of a fleet of wind turbines: the generated models are applicable to any component of any wind turbine technology. However, when applying this process to different study cases, it was found that the health indicators remain insensitive to certain faults.

To solve this issue, one solution is to to resort to an ensemble approach to leverage the individual detection performance of a set of health indicators in order to improve the global performances.

Bagging and boosting are among the most popular ensemble methods. However, they are not adapted to industrial data, as shown by Opitz and Maclin (1999). Stacking methods, which consider models from various classifiers (e.g. k-nearest neighbors, support vector machine) to improve the accuracy of the final model are a more relevant solution in this respect (Lima, Blatt, and Fujise (2020),Waqas Khan and Byun (2022),Yuan, Sun, and Ma (2019)). Ensemble methods are very flexible, and to achieve the requirements of modeling accuracy, a very large number of models can be considered, as in Pichika, Meganaa, Geetha Rajasekharan, and Malapati (2022), which performs a fusion of the three approaches mentioned above.

Thus, these study cases are not suitable for industrial use because they can require a high implementation and execution cost, and bring confusion in the interpretation of the models by an expert user. The non-detection of some faults is due to the choice of variables made by the selection algorithm: as the model generation process is based on the estimation performance on healthy data, following the normal behavior modeling (or NBM) approach ((Schlechtingen, Santos, and Achiche (2013), Schlechtingen and Santos (2014))), it can select regressors correlated to the estimated variable during a fault, making the associated health indicator insensitive to it. Thus, using a larger number of variables in the model generation process may be a way to solve the problem. This would allow to take advantage of the diversity of available SCADA database to capture different faulty behaviors. The results of variable selection algorithms depend on the ranking criterion of the models used, and there are nowadays many methods, summarized in the review Li, Cheng, Wang, Morstatter, Trevino, Tang, and Liu (2018). A solution to consider a greater diversity in the variable selection would then be to modify the ranking criterion of the algorithm used for variables selection. However, as the same model can be optimal for several different criteria, nothing guarantees that using different criteria will lead to a greater variability in the selected variables.

In this study, an ad-hoc ensemble method, inspired by the stacking approach, to automatically generate a set of tri-variate linear models, predicting the evolution of a same output variable is proposed. The process is based on a constrained greedy forward selection algorithm, generating five unique linear models. Each model is made different by forcing the variable selection algorithm to use different input variables. An additional model composed of variables linked to the output one, but whose variations are not impacted by wind turbine faults is also considered. Multi-turbine health indicators are then built using these models, together with a signal measuring the deviation of the variation of the variable to be estimated with the median value of the wind farm. A mean residual, measuring the average value of these seven indicators, is then constructed. In order to compare the detection performances of the different linear residuals considered, a receiver operating characteristic (or ROC) (Fawcett (2006)) curve analysis is performed. The

method is evaluated on two fault cases, one where the mono-residual monitoring approach provides good detection performances, and the other where the fault was not detected by the previous approach because of the variable selection result.

The outline of this paper is as follows. In Section 2, the constrained variable selection algorithm is first detailed, together with the linear residual considered. Then the ROC-curve based performance analysis method is presented. The data used and the faults considered are discussed in section 3. Finally, Section 4 and 5 show and discuss the obtained models and detection performances on the two faults.

2. Method

2.1. *Constrained Greedy Forward Selection (CGFS)*

The models are generated using a greedy forward selection (GFS) approach (Pohjankukka, Tuominen, Pitkänen, Pahikkala, and Heikkonen (2018)). To reduce implementation costs, the proposed GFS algorithm uses as an input the ten best variables to predict a given output, selected using a Lasso type linear regression. This set of variables will be named V_{Lasso}. The variable selection process provides tri-variate models applicable to all turbines of a given wind farm by integrating a multi-turbine performance criterion: a given model is ranked according to the median MAE of all turbines of the wind farm, using one year of data where the studied component is known to be healthy for all turbines, following a normal behavior modeling (NBM) approach (Schlechtingen et al. (2013)). More details on the variable selection process can be found in Raymond et al. (2022).

As the GFS algorithm is applied on a fault-free period, the detection performance on faulty period remains unknown, and it can thus retains regressors which allow a very accurate modeling on a healthy period, but correlated with the variable to model in faulty period. The associated residual is then insensitive to the fault. The solution proposed in this paper is to run the GFS algorithm iteratively with a constraint on V_{Lasso}: let $X_0 = (X_{01}, X_{02}, X_{03})$ be the variables retained by the application of the unconstrained GFS algorithm, the process is executed again by removing the variables X_{02}, X_{03} from the input V_{Lasso}, providing a new set of model variables X_1. X_2 is then the variable selection result obtained by deleting X_{01} and X_{03}, and X_3 the one obtained by the removal of X_{01} and X_{02} from V_{Lasso}. The algorithm is executed one last time by deleting from V_{Lasso} all the variables of X_0, providing X_4. In the following, this algorithm will be referred as constrained greedy forward selection or CGFS, and the set of variables removed from V_{Lasso} will be called "constraint". Figure 1 illustrates the full CGFS process.

2.2. *Linear residuals*

Once the five sets of model variables are selected by the CGFS algorithm described in Section 2.1, the models are learned over a one-year fault-free period, and the health indicator is constructed according to the multi-turbine approach presented in Lebranchu, Charbonnier, Bérenguer, and Prevost (2019), and detailed in Equation 1, where N_T is the number of wind turbines of the considered wind farm, WT_i a given wind turbine, and ε^{WT_i} the mono-turbine residual, defined as the deviation between the measured value of an output variable and its value estimated by a GFS model for the data from WT_i:

$$\varepsilon(t) = \varepsilon^{WT_i}(t) - \text{median}_{j \in [\![1;N_T]\!]}(\varepsilon^{WT_j}(t)) \quad (1)$$

In addition to these five linear residuals, another residual, called ε_{X_r} and built according to Equation 1 is considered. It uses as model variables the power produced p_{apparent}, the measured wind

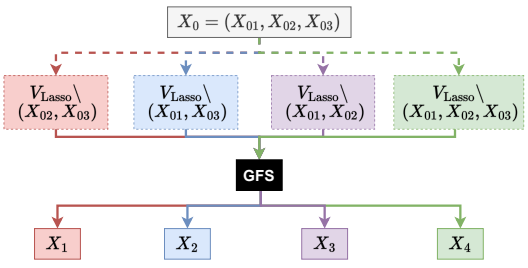

Fig. 1. Constrained GFS process

speed s_{wind} and the outdoor temperature t_{ext}. These variables are relevant because they can be related to a component heating, and since they are insensitive to a fault in the operation of a turbine, the risk of correlation with the modeled variable during a fault is avoided.

The last indicator considered for the computation of the mean linear residual measures the deviation between the value of a variable to be modeled of a given wind turbine and the median value at the scale of the wind farm. The advantage of this indicator, which will be called ε_{md}, is that it is only composed of temperature variables that have a linear relationship. ε_{md} is therefore expected to present a lower false alarm rate than the other indicators in normal operating conditions.

2.3. *Performance evaluation using ROC curves*

To evaluate the detection performance of the indicators, three metrics are used: the false positive rate (FPR), the true positive rate (TPR), and the advance detection time (ADT).

FPR measures the probability that the health indicator is above a detection threshold τ during a fault-free period of one-year H_0, while TPR is the probability that the health indicator is above τ during the faulty period H_1. The end of H_1 is set at a time when a fault has been detected by the manufacturer or when the turbine has been shut down for maintenance, while the beginning of H_1 is defined approximately given a WT expert operator knowledge. The FPR and the TPR are calculated from the area under the curve (AUC) of the probability density function of the indicator on these two periods, estimated by the kernel density estimation method (Silverman (2017)).

For a given τ, the first detection date is set as the first day from which the residual has crossed the detection threshold at least five days of the previous week during the faulty period. The ADT is then defined as the number of days between the first detection date and the end of H_1.

Two ROC curves $\mathrm{TPR}(\mathrm{FPR})$ and $\mathrm{ADT}(\mathrm{FPR})$ are then built using sets of FPR, TPR and ADT computed for values of τ chosen from the values of the residual indicator on H_0 and H_1 with a step of 0.1.

The global detection performances of the different linear residuals are evaluated using the AUC of the two ROC curves (Park, Goo, and Jo (2004)). The AUC is computed using the trapezoidal rule (Atkinson (1991)).

In particular, the AUC of the two ROC curves of ε_{X_0}, the residual built using the GFS algorithm without constraints on V_{Lasso}, will be compared to those of the mean residual $\overline{\varepsilon}$, calculated from the values of the seven residuals considered, to evaluate the contribution of the ensemble method proposed in this study.

It is necessary to recall that the GFS algorithm generates models from fault-free data, and therefore no information about their detection performances on H_1 is available at the model generation step. The interest of using an ensemble method is therefore to find models that will have potentially better detection performances on H_1 than the model built using the mono-residual approach from Raymond et al. (2022). Since the existence and occurrence of such models resulting from the proposed process is not known, the computation of the mean of the linear residuals is a relevant choice.

3. Resources

3.1. *Data*

A subset of low frequency (10 minutes) SCADA variables from the manufacturer's database is used in this study. Most of these operating variables are temperatures from components, but the database also includes mechanical variables such as rotation speed and torque, electrical variables or measured wind speed.

3.2. *Faults*

Two fault cases will be considered in this study. The first one, $F^A_{\mathrm{generator}}$, is an overheating of the generator rear bearing of a turbine from the wind farm A. This anomaly led to a replacement operation that lasted one month. This fault is monitored using the modeling of the generator rear bearing temperature $t_{\mathrm{generatorRearBearing}}$. The health indicator constructed from the unconstrained GFS

algorithm provides good detection performance on H_0 and H_1.

The second one, F^B_{gearbox}, is an abnormal heating of components of the gearbox system of a turbine from the wind farm B. This anomaly, monitored using the modeling of the gearbox oil temperature $t_{\text{gearboxoil}}$, did not result in the replacement of the component. However, the energy losses are significant due to the recurrent clamping of the wind turbine during the faulty period. The model from the unconstrained GFS provides a residual insensitive to the fault, resulting in very poor detection performance on H_1.

4. Results

4.1. *CGFS variable selection*

Tables 1 and 2 present the variable selection results of the CGFS algorithm for respectively the wind farm A and B databases. In Table 1, it can be seen that the variables $s_{\text{fastshaft}}$ and t_{transfo} are regular choices of regressors to replace those in X_0.

Concerning the wind farm B database for the modeling of the oil gearbox temperature, Table 2 presents a more diverse choice of variables among the generated models, including electrical variables, and several mechanical ones ($\rho_{\text{gearboxoil}}$,$s_{\text{fastshaft}}$). This result can be explained by the fact that the database contains more variables belonging to the gearbox than to the generator system.

Table 1. constrained GFS variable selection results on wind farm A

Model label	Model variables
X_0	$t_{\text{generatorFrontBearing}}, t_{\text{stator}}, t_{\text{int}}$
X_1	$t_{\text{generatorFrontBearing}}, s_{\text{fastshaft}}, t_{\text{transfo}}$
X_2	$t_{\text{stator}}, s_{\text{fastshaft}}, t_{\text{transfo}}$
X_3	$t_{\text{transfo}}, s_{\text{fastshaft}}, t_{\text{int}}$
X_4	$t_{\text{transfo}}, s_{\text{fastshaft}}, t_{\text{gearboxambient}}$
X_r	$p_{\text{apparent}}, s_{\text{wind}}, t_{\text{ext}}$

4.2. *Linear residuals comparative analysis*

A line chart of the eight different residuals on periods H_0 (green) and H_1 (red) is presented in Figures 2 and 4. Figures 3 and 5 show the two ROC curves TPR(FPR) and ADT(FPR), drawn using the sets of metric performances computed following the method detailed in Section 2.3, of the eight residuals, with a highlighting on the measurement points associated with the performance constraint $\text{FPR} = 5\%$, which represents a classic detection performance objective when evaluating a health indicator. The ROC curves from the seven linear residuals are shown with a dashed line, while a solid line is used for the mean residual. Finally, Table 3 and 4 shows the AUC of the two ROC curves TPR(FPR) and ADT(FPR) from the eight considered residuals on the two fault cases.

Table 2. constrained GFS variable selection results on wind farm B

Model label	Model variables
X_0	$t_{\text{gearboxbearing2}}, t_{\text{gearboxambient}}, t_{\text{stator1}}$
X_1	$t_{\text{gearboxbearing2}}, \rho_{\text{gearboxoil}}, t_{\text{gearboxbearing1}}$
X_2	$i_{\text{produced}}, t_{\text{gearboxambient}}, s_{\text{fastshaft}}$
X_3	$t_{\text{stator1}}, s_{\text{fastshaft}}, \rho_{\text{gearboxoil}}$
X_4	$i_{\text{produced}}, t_{\text{slowshaftbearing}}, \rho_{\text{gearboxoil}}$
X_r	$p_{\text{apparent}}, s_{\text{wind}}, t_{\text{ext}}$

For $F^A_{\text{generator}}$, as can be seen in Figure 3 and Table 3, the residual from the unconstrained GFS algorithm ε_{X_0} provides an excellent global detection performance for both the TPR(FPR) and ADT(FPR) ROC curves. The other residuals generated using the variables selected by the Constrained GFS algorithm lead to unequal global performance: some of which are slightly worse than ε_{X_0} (ε_{X_1}, ε_{X_4} and ε_{md}), while the others show better overall performance for a single ROC curve. Table 3 shows that although ε_{X_0} already provides excellent detection performance, the mean residual $\overline{\varepsilon}$ provides a slight improvement in global detection performances for both ROC curves, with a respective gain of 1.64% for TPR(FPR) and 0.65% for ADT(FPR).

Concerning F^B_{gearbox}, by analyzing the model variables in Table 2 and the line chart of Figure 4, it can be noticed that the fault-sensitive linear residuals $\varepsilon_{X_1}, \varepsilon_{X_3}, \varepsilon_{X_4}, \varepsilon_{X_r}$ and ε_{md} are the ones that do not contain $t_{\text{gearboxambient}}$, while the others residuals do not detect the fault. In Figure 5, it can be seen that the ROC curves of the residual ε_{X_0} and ε_{X_2} illustrate the non-detection of the anomaly F^B_{gearbox} with an almost linear

TPR(FPR) curve (corresponding to an AUC close to 50%, as can be verified in Table 4). In Table 4, it can be observed that all the other residuals greatly improve the global detection performance with respect to the three considered metrics. Consequently, the mean residual $\overline{\varepsilon}$ also provides a significant gain in global performance, of respectively more than 35% for the TPR(FPR) curve and more than 25% for the ADT(FPR) curve.

Table 3. AUC results of ROC curves from $F^A_{\text{generator}}$

ROC curve	ε_{X_0}	ε_{X_1}	ε_{X_2}	ε_{X_3}	ε_{X_4}	ε_{X_r}	ε_{md}	$\overline{\varepsilon}$
TPR(FPR)	92.0	85.2	92.5	93.8	85.9	96.7	89.4	93.6
ADT(FPR)	94.0	86.5	93.2	93.7	87.0	27.4	89.7	94.7

Table 4. AUC results of ROC curves from F^B_{gearbox}

ROC curve	ε_{X_0}	ε_{X_1}	ε_{X_2}	ε_{X_3}	ε_{X_4}	ε_{X_r}	ε_{md}	$\overline{\varepsilon}$
TPR(FPR)	51.8	78.9	48.6	82.8	89.6	91.0	85.9	87.1
ADT(FPR)	72.3	92.8	86.9	97.5	99.7	99.5	98.6	98.9

5. Discussion

In addition to the significant improvement in global detection performance over mono-residual analysis, the multi-residual monitoring approach presented in this paper has other benefits in an industrial application context. Indeed, the availability of the mean residual is increased in situations of missing data as only one of the seven linear residuals is needed for $\overline{\varepsilon}$ to be evaluated.

By considering more complex machine learning methods such as support vector machine or neural network, it is very likely that the AUC

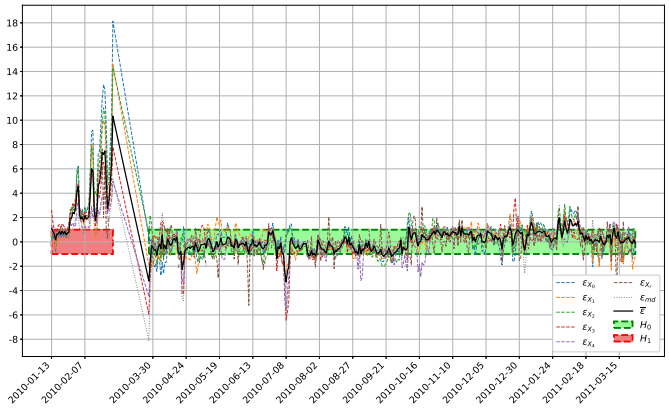

Fig. 2. Line charts of the eight linear residuals $F^A_{\text{generator}}$

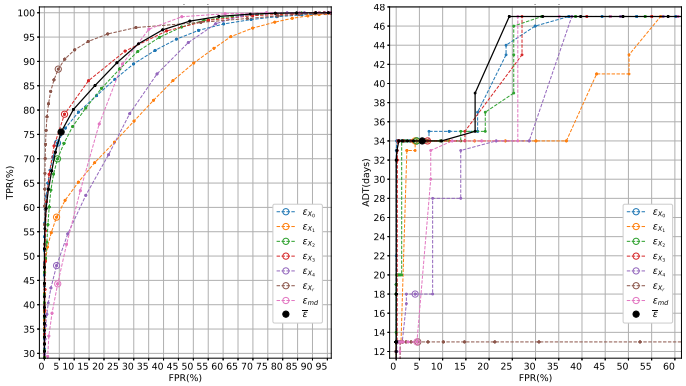

Fig. 3. ROC curves TFP(FPR) and ADT(FPR) on $F^A_{\text{generator}}$

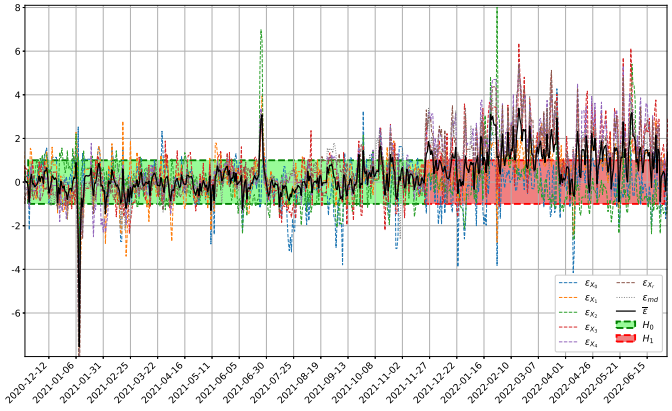

Fig. 4. Line charts of the eight linear residuals F^B_{gearbox}

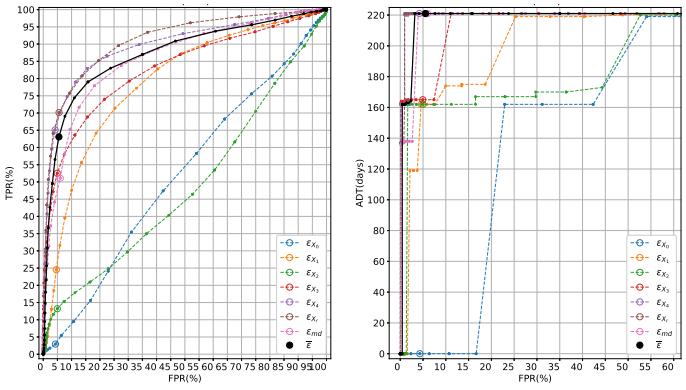

Fig. 5. ROC curves TFP(FPR) and ADT(FPR) on F^B_{gearbox}

will be improved. However, these methods are less robust to industrial constraints, and provide less interpretable models than the proposed linear approach (Schlechtingen and Ferreira Santos (2011)) (one can indeed notice that although more variables are considered for the monitoring of a component, the variable selection result remains physically consistent).

6. Conclusion

This paper presents a solution to improve the detection performance of the CBM process proposed in Raymond et al. (2022), considering an ensemble of linear residuals.

Seven residuals are considered, five of them coming from the GFS variable selection process. The application of constraints on the set of input variables of the GFS algorithm, V_{Lasso}, allows to obtain a greater diversity in the model variables. These seven signals are then used to construct a mean residual, whose detection performance is compared with that of the mono-residual approach from Raymond et al. (2022).

The global detection performance is evaluated and compared using the AUC of two ROC curves, constructed from three performance metrics.

The multi-residual analysis is applied on two fault cases: for the first fault, the mean residual slightly improves the global detection performances of the mono-residual approach, which initially provided very good performance. For the second fault, not detected by the application of the mono-residual approach, the mean residual allows a detection of the fault, and a great improvement

of the global performance, both in the fault-free and faulty periods.

References

Atkinson, K. (1991, January). *An Introduction to Numerical Analysis*. John Wiley & Sons.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters 27*(8), 861–874.

Lebranchu, A., S. Charbonnier, C. Bérenguer, and F. Prevost (2019, April). A combined mono- and multi-turbine approach for fault indicator synthesis and wind turbine monitoring using SCADA data. *ISA Transactions 87*, 272–281.

Li, J., K. Cheng, S. Wang, F. Morstatter, R. P. Trevino, J. Tang, and H. Liu (2018, November). Feature Selection: A Data Perspective. *ACM Comput. Surv. 50*(6), 1–45.

Lima, L. A. M., A. Blatt, and J. Fujise (2020, September). Wind Turbine Failure Prediction Using SCADA Data. *J. Phys.: Conf. Ser. 1618*, 022017.

Maldonado-Correa, J., S. Martín-Martínez, E. Artigao, and E. Gómez-Lázaro (2020, January). Using SCADA Data for Wind Turbine Condition Monitoring: A Systematic Literature Review. *Energies 13*(12), 3132.

Márquez, F. P. G. (2020). *MAINTENANCE MANAGEMENT OF WIND TURBINES*. S.l.: MDPI AG. OCLC: 1191152052.

Opitz, D. and R. Maclin (1999, August). Popular Ensemble Methods: An Empirical Study. *jair 11*, 169–198.

Park, S. H., J. M. Goo, and C.-H. Jo (2004). Receiver Operating Characteristic (ROC) Curve: Practical Review for Radiologists. *Korean J Radiol 5*(1), 11.

Pichika, S., G. Meganaa, S. Geetha Rajasekharan, and A. Malapati (2022). Multi-component fault classification of a wind turbine gearbox using integrated condition monitoring and hybrid ensemble method approach. *Applied Acoustics 195*(11).

Pohjankukka, J., S. Tuominen, J. Pitkänen, T. Pahikkala, and J. Heikkonen (2018, October). Comparison of estimators and feature selection procedures based on ALS and digital aerial imagery. *Scandinavian Journal of Forest Research 33*(7), 681–694.

Raymond, T., A. Lebranchu, C. Bérenguer, and S. Charbonnier (2022). Data-driven Model Generation Process for Thermal Monitoring of Wind Farm Main Components through Residual Indicators Analysis. *ESREL 32*(8), 1393–1400.

Schlechtingen, M. and I. Ferreira Santos (2011, July). Comparative analysis of neural network and regression based condition monitoring approaches for wind turbine fault detection. *Mechanical Systems and Signal Processing 25*(5), 1849–1875.

Schlechtingen, M. and I. F. Santos (2014, January). Wind turbine CBM using NBM. Part 2: Application. *Applied Soft Computing 14*, 447–460.

Schlechtingen, M., I. F. Santos, and S. Achiche (2013, January). Wind turbine CBM using NBM. Part 1: System description. *Applied Soft Computing 13*(1), 259–270.

Silverman, B. W. (2017, October). *Density Estimation for Statistics and Data Analysis*. New York: Routledge.

Stetco, A., F. Dinmohammadi, X. Zhao, V. Robu, D. Flynn, M. Barnes, J. Keane, and G. Nenadic (2019, April). Machine learning methods for wind turbine condition monitoring: A review. *Renewable Energy 133*, 620–635.

Tautz-Weinert, J. and S. J. Watson (2017, March). Using SCADA data for wind turbine condition monitoring – a review. *IET Renewable Power Generation 11*(4), 382–394.

Waqas Khan, P. and Y.-C. Byun (2022, September). Multi-Fault Detection and Classification of Wind Turbines Using Stacking Classifier. *Sensors 22*(18), 6955.

Yuan, T., Z. Sun, and S. Ma (2019, January). Gearbox Fault Prediction of Wind Turbines Based on a Stacking Model and Change-Point Detection. *Energies 12*(22), 4224.