Journal of Physics: Conference Series

PAPER • OPEN ACCESS

On the use of high-frequency SCADA data for improved wind turbine performance monitoring

To cite this article: E Gonzalez *et al* 2017 *J. Phys.: Conf. Ser.* **926** 012009

IOP Conf. Series: Journal of Physics: Conf. Series **926** (2017) 012009 doi:10.1088/1742-6596/926/1/012009

On the use of high-frequency SCADA data for improved wind turbine performance monitoring

E Gonzalez[1], B Stephen[2], D Infield[2], J J Melero[1]

[1]CIRCE - Universidad de Zaragoza, C/ Mariano Esquillor Gómez, 15, 50018 Zaragoza, Spain
[2]Department of Electronic and Electrical Engineering, University of Strathclyde, Glasgow, UK

E-mail: egonzalez@fcirce.es

Abstract. SCADA-based condition monitoring of wind turbines facilitates the move from costly corrective repairs towards more proactive maintenance strategies. In this work, we advocate the use of high-frequency SCADA data and quantile regression to build a cost effective performance monitoring tool. The benefits of the approach are demonstrated through the comparison between state-of-the-art deterministic power curve modelling techniques and the suggested probabilistic model. Detection capabilities are compared for low and high-frequency SCADA data, providing evidence for monitoring at higher resolutions. Operational data from healthy and faulty turbines are used to provide a practical example of usage with the proposed tool, effectively achieving the detection of an incipient gearbox malfunction at a time horizon of more than one month prior to the actual occurrence of the failure.

Keywords: Operation & Maintenance, Wind Turbine, Power Curve, Performance Monitoring, SCADA, High-frequency data, Fault Detection

1. Introduction

High wind farm (WF) operation and maintenance (O&M) costs are a major concern today. The maintenance requirements due to wind turbine (WT) failures, together with poor site accessibility, can lead to rates up to 30% of the levelised cost of energy (LCOE) of an offshore WF [1]. Unscheduled maintenance activities not only increase the O&M costs but also have a strong impact on downtime and thus on annual energy production. There is therefore a need to improve the WT O&M phase.

Early detection of failures is desired to reduce O&M costs while improving reliability. To this end, the use of condition monitoring systems (CMS) is required. Although numerous commercial options and techniques are already available [2, 3], there are still limitations for a wide deployment. Some CMS might outstrip the expense of the required additional equipment, but may also exhibit high rates of false positive alarms while the diagnosis is dedicated to a unique component or assembly rather than system-wide [2, 4]. WT condition monitoring based on the use of data from the supervisory control and data acquisition (SCADA) system [5] is increasingly seen as a cost-effective and promising approach, as SCADA data is available at no additional cost. Within this approach, monitoring WT performance is fundamental, since it is undoubtedly the main feature characterising WT overall operation. Component malfunction may degrade the energy conversion efficiency leading to performance deviations. Performance monitoring can thus play a useful role in global condition monitoring.

Generally, performance monitoring is synonymous with power curve monitoring [6]. The most widely used approach consists on modelling a reference power curve, based on operational data under normal operating conditions (WT normal performance), together with confidence intervals [7] for discriminating between normal and abnormal operation over time. Deviations from this reference, beyond these thresholds, denote a fault in the energy extraction process, e.g. due to incipient failures of components [8]. Although important research has been conducted regarding WT power curve modelling, little emphasis has been placed on the estimation of the related confidence limits.

Both parametric [9, 10, 11, 12] and non-parametric methods [9, 13, 14, 15, 16] have been studied to produce a reference for WT normal performance, based on operational data from the SCADA system. While parametric models fit a functional form to the data by changing one or several parameters, non-parametric methods infer a functional form and do not need to make any assumption about the data. Recent research has highlighted the accuracy of non-parametric and multivariate methods [16] for WT power curve modelling. While most contributions rely on univariate modelling (relating wind speed and WT power output), multivariate models allow to take into account the natural variability of WT performance, greatly influenced by several phenomena.

With regard to the data used, the majority of studies and commercial applications are based on the use of 10-minute aggregated data. As stated by Yang *et al.* [3], this low frequency resolution negatively affects the diagnosis and prognosis capabilities. Indeed, this aggregation may hide short-lived events. High-resolution SCADA data instead, should allow dynamic turbine behaviour to be identified with higher fidelity and thus improve detection capabilities [13, 17, 18].

Finally, as previously mentioned, the assessment of WT performance thresholds has not been extensively studied. The estimation of these confidence intervals is directly linked to power curve modelling related uncertainty. Indeed, WT performance can significantly vary under normal operating conditions [19, 20, 21]. Power curve modelling uncertainty should account for this normal variability to ensure effective abnormal behaviour detection. Even though non-parametric approaches can accurately represent WT performance, the assessment of their related uncertainty is not always straightforward. State-of-the-art methods rely on the assumption of a normal distribution of the modelling error [9, 22]. Recent work published by the authors has shown the heteroscedastic nature of the model residuals [18], that is, characterised by a non-constant variance across the different wind speed regimes. As a result, the assumption of a normal distribution of the modelling error may lead to a loss of detection capabilities and its validity should be questioned. WT normal performance thresholds should rather vary across the different wind speed regimes. This paper introduces the use of Quantile Regression Forests [23, 24] and high-frequency SCADA data for modelling WT normal performance together with its related uncertainty, for the purpose of performance monitoring.

The paper is organised as follows. Section 2 details the suggested approach for WT performance monitoring, with four alternatives to assess WT performance and its related uncertainty. In Section 3, the methodology is first tested in a case study, consisting of a healthy WT; modelling accuracy and detection effectiveness are compared for the different modelling techniques and for SCADA data of high and low frequency data. The best approach is then applied to concurrently monitor healthy and faulty turbines for detection purposes. Advantages and shortcomings are discussed in Section 4, together with future work.

2. Methodology

The suggested methodology for WT performance monitoring includes several phases as one can see in Figure 1. The four phases can be summarised as data filtering, performance modelling and assessment of its related uncertainty, then abnormal performance detection and root-cause diagnosis. Each phase is detailed subsequently.

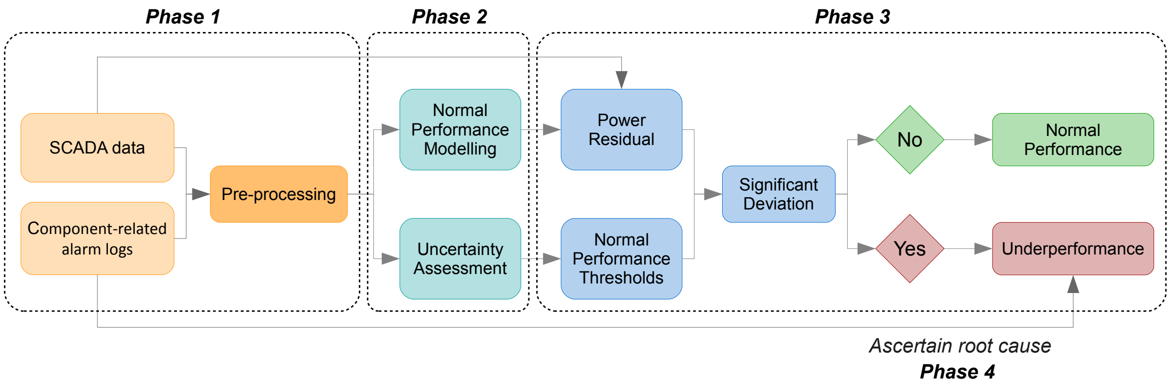

Figure 1. General framework for wind turbine performance monitoring.

During ***Phase 1***, the historical SCADA data used for normal performance modelling should be filtered to ensure faultless representation of WT operation. The pre-processing is addressed in two steps. First, the alarm logs also recorded by the SCADA system are integrated to segregate operational data and flag any abnormal event acknowledged by the system. The SCADA logs are automatically translated into component-related information, based on a modernised WT taxonomy [25]. The interested reader may look in [26] for more information. This way, the SCADA data can be flagged every time the WT was experiencing an issue related to a specific component, to the grid or to extreme weather events; the flagged periods can be then discarded. Since some problems might not be acknowledged by the SCADA system, further actions may be required. In this second step, further abnormal performance filtering is applied based on a multivariate curve approach and machine characteristics, similarly as in [20]. Pitch angle and rotor speed data allow clear identification of the different operating regimes and therefore normal operation, as one can see in Figure 2. All figures presented here are normalised for confidentiality reasons.

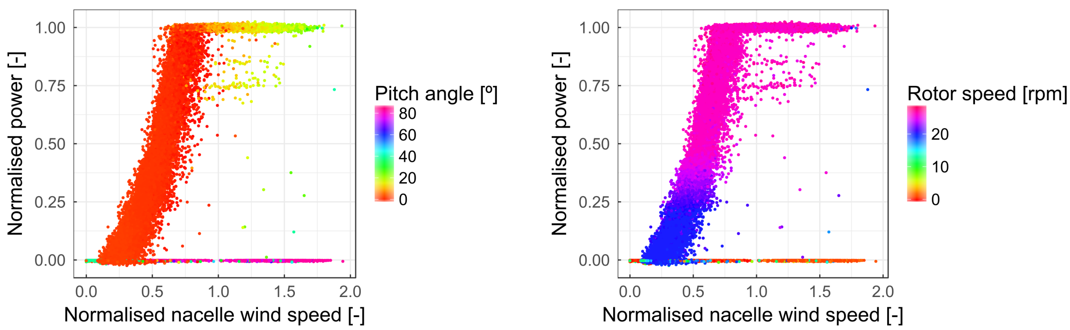

Figure 2. Scatter WT power curve varying depending on blade pitch angle (left) and rotor speed (right).

Phase 2 includes both WT performance modelling and assessment of its related uncertainty. In this paper, four different models are considered for comparison purposes; they are summarised subsequently.

IOP Conf. Series: Journal of Physics: Conf. Series **926** (2017) 012009 doi:10.1088/1742-6596/926/1/012009

- *Method of bins*
  The method of bins (MOB) is the current industry practice for WT performance testing, as outlined in [27, 28]. It relies on the reduction of the dataset into mean values per wind speed intervals, called bins. For each bin, centred on a multiple of 0,5 m/s, the wind speed and power output are reduced to the calculated mean. Based on these bins and corresponding mean values it is then possible to predict the power output given the wind speed. In this paper, a power curve is obtained based on the nacelle wind speed and the power output. Also, as detailed in [28], data normalisation is applied to the nacelle wind speed to account for the air density effect.
  Similarly to the mean, the standard deviation of the power is calculated per bin to assess the uncertainty related to predictions, as illustrated in Figure 3. This model uncertainty is directly related to the statistical variation of the power output, as defined in [28].

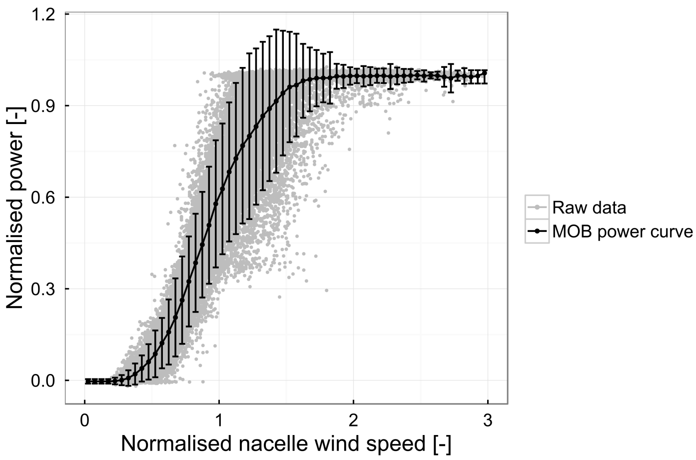

Figure 3. Power curve and related uncertainty derived from the MOB.

- *k-Nearest Neighbours*
  The k-Nearest Neighbours (kNN) is a non-parametric machine learning tool that predicts a new sample using the k closest samples from the training dataset [29]. A multivariate approach is considered here; the nacelle wind speed, ambient temperature, pitch angle and rotor speed data are used as inputs to predict WT power output. Although considering wind direction data as an additional model input may improve WT power prediction [16], this was discarded in this paper due to a lack of data availability. Nevertheless, WT directional behaviour, including wake affected performance, is considered to be inherent in the WT operational data.
  The value of k, characterising the k-Nearest Neighbours, is selected according to the minimum root mean square error (RMSE) from the testing dataset.
  Uncertainty related to the kNN predictions is assessed based on the state-of-the-art practice [9], that assumes a normal distribution of the training error with mean μ_{train} and standard deviation σ_{train}. As a consequence, the same distribution is considered across the different wind speed regimes. Thresholds for normal performance can be derived from control limits (CL) as defined in Equation 1. The parameter η can be tuned iteratively to attain a level of 95% confidence based on the training dataset.

$$CL = \mu_{train} \pm \eta \frac{\sigma_{train}}{\sqrt{N_{train}}} \tag{1}$$

– *Random Forests*
Random Forests (RF) are also a non-parametric machine learning method built upon multiple regression trees [30]. A conceptual diagram of the algorithm is illustrated in Figure 4, where the left-hand side corresponds to the training phase and the right-hand side to the testing/predicting phase. A single regression tree (e.g. $tree_1$, $tree_2$, or $tree_T$) usually splits the training output values according to conditions of the input values. Different conditions of these input values are represented as branches, while the target values are represented as leaves, corresponding to light blue shaded nodes in Figure 4. RF are considered an ensemble method since they operate by randomly constructing multiple regression trees, avoiding the overfitting problems often related to the use of a single regression tree. As concerns new predictions, RF predict new samples by averaging the individual predicted values from the regression trees, according to conditional inputs; the single predicted values from each tree are highlighted in dark blue in Figure 4.

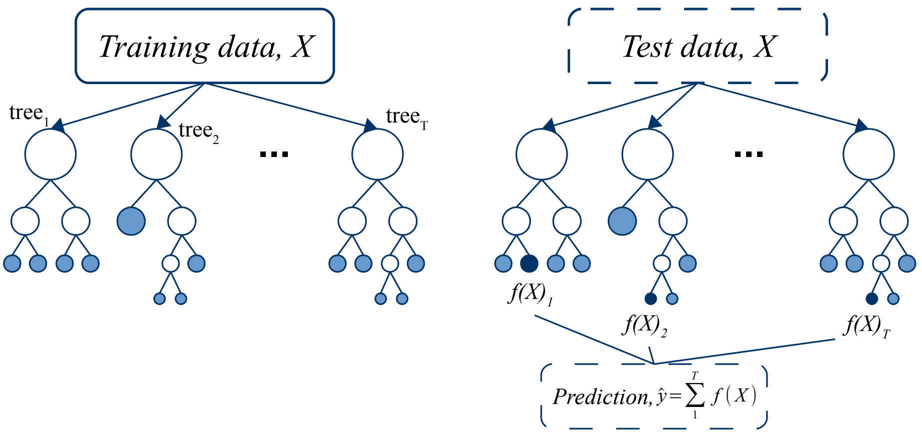

Figure 4. Conceptual diagram of the RF algorithm. Adapted from [31].

Similarly to the previous method, in this paper a power curve is obtained using the nacelle wind speed, ambient temperature, pitch angle and rotor speed data as inputs.
Likewise, uncertainty related to the RF predictions is assessed by the use of control limits derived from the assumed normal distribution of the training error (Equation 1).

– *Quantile Regression Forests*
As can be seen on the right side of Figure 4, WT power output is predicted from RF as a conditional mean of multiple trees responses. One can therefore easily think that considering the full conditional distribution of the response variable rather than only the mean, would allow to properly address WT power variability under similar input conditions [24]. Quantile Regression Forests (QRF) are based on this principle, as a natural generalisation of the RF algorithm. As a result, quantile regression provides information about the spread of the response variable, according to different levels of desired confidence. The slight variation between the classic RF algorithm and the QRF relies on the main predicted value; the first predicts the response based on the conditional mean whereas the latter predicts the response based on the median of the conditional distribution, corresponding to a level of confidence of 50%. To the best of the author's knowledge, no quantile regression has been applied before for WT performance assessment. In this paper, the WT power curve is modelled by the use of QRF with nacelle wind speed, ambient temperature, pitch angle and rotor speed as input data. Four prediction intervals according to different levels of confidence (80%, 85%, 90% and 95%) are used .

As illustrated in Figure 1, ***Phase 3*** consists in the detection of abnormal performance. The difference between the modelled and actual power production is compared to the normal performance thresholds to distinguish between abnormal and expected behaviour. The present work focuses on the detection of WT underperformance. Over and underperformance are here considered to be independent, rather than a common detection as abnormal performance.

Finally, ***Phase 4*** is defined as the diagnosis phase. Once an event of underperformance is detected, the alarms registered by the SCADA system at the same time might be helpful for ascertaining the root cause of the detected deviation. Underperformance events may thereby be attributed to a component-related root cause.

3. Results and discussion

3.1. Data description

SCADA data from an operating WF was used to evaluate the power curve modelling accuracy and the detection capabilities of the described methodology. Two different data time resolutions were considered. First, real measurements recorded at a 4-second resolution, so-called high-frequency data. Second, 10-minute aggregated signals, the current industry practice. One full year of data was taken from three selected turbines. The turbines, of the same type, are located next to each other, at similar altitudes and terrain complexity, hence exposed to similar local conditions. Two of the turbines did not register any failure during the available year, whereas the third registered a major gearbox failure, as indicated in the documented information provided by the WF operator. No information about the specific failure mode was reported. A major failure is understood by the considered operator as entailing either a component repair or replacement.

3.2. Power curve modelling accuracy

High-frequency data from the two healthy turbines (WT1 and WT2) was here used to compare the accuracy of the models described in Section 2. The first month was used to train and test the models, all using a 10-fold cross-validation, by randomly dividing the observations into ten-folds of approximately equal sizes. Modelling accuracy was evaluated by the mean absolute error (MAE) and the root mean squared error (RMSE). Results for the high-frequency data are summarised in Table 1.

Table 1. Performance metrics for power curve modelling accuracy.

	WT1		WT2	
	MAE (%)	RMSE (%)	MAE (%)	RMSE (%)
MOB	4.40	7.22	4.11	6.76
kNN	2.65	5.00	2.66	4.83
RF	1.39	2.65	1.36	2.52
QRF	2.14	4.44	2.03	4.07

In agreement with the literature, all non-parametric models are highly accurate in representing WT normal performance, and the RF outperforms in comparison. The MOB presents higher errors, especially in terms of the RMSE. The univariate and simplistic approach is not able to take the natural variability of WT performance into consideration. The difference between the QRF and the RF, relying on the consideration of the median over the mean, is translated into a slightly significant variation in the prediction accuracy.

IOP Conf. Series: Journal of Physics: Conf. Series **926** (2017) 012009 doi:10.1088/1742-6596/926/1/012009

3.3. Uncertainty assessment

As detailed in Section 2, the modelling related uncertainty was used to build normal performance thresholds for each method. The model error distribution together with its related normal performance thresholds is depicted in Figures 5 to 8.

Although, the MOB appears to be less accurate in normal performance modelling, the data reduction to a mean and standard deviation per bin accounts for the heteroscedastic nature of the model residual, that is, the variance is not the same for each bin; it therefore produces variable control limits across the different wind speed regimes. The wrong assumption for the uncertainty related to both the kNN and the RF is confirmed in Figure 6 and Figure 7. While the combined distribution may resemble a normal distribution, notwithstanding the high kurtosis, the resulting control limits do not take into account the heteroscedasticity of the residual. This undoubtedly results in a loss of the detection capabilities. Finally, the QRF offers a very flexible way to assess WT performance uncertainty, while having a reasonable modelling accuracy.

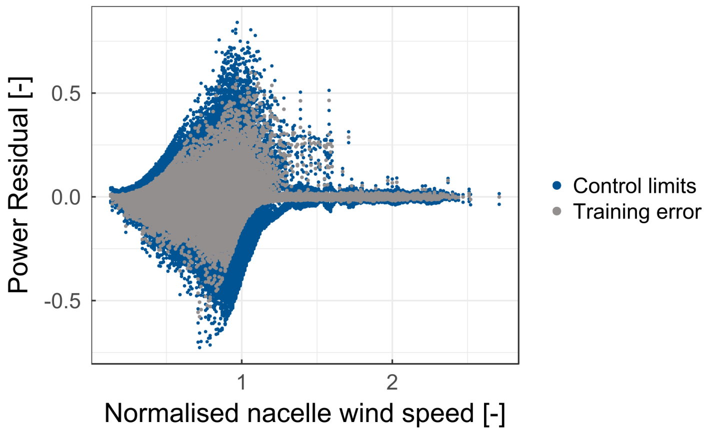

Figure 5. Training error (grey) and related control limits (blue) from the MOB varying depending on the wind speed.

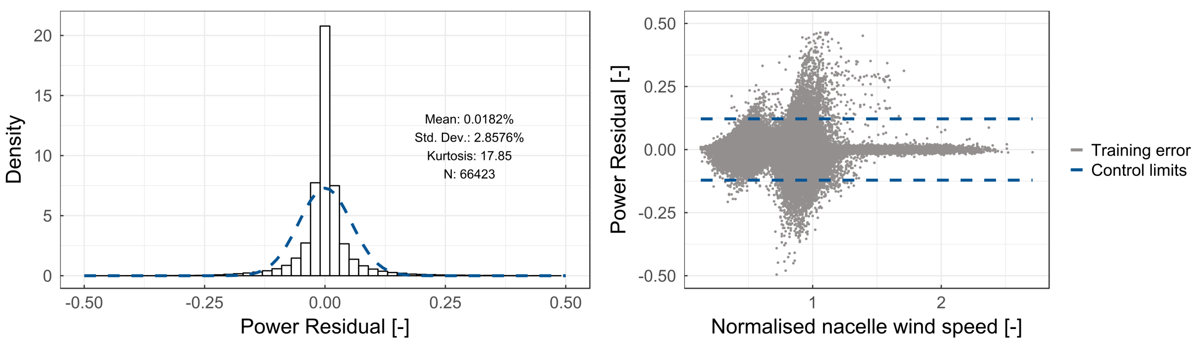

Figure 6. Training error statistics and distribution from kNN model (left side); the related normal distribution is illustrated in blue. Training error (grey) and related control limits (blue) from kNN model varying depending on the wind speed (right side).

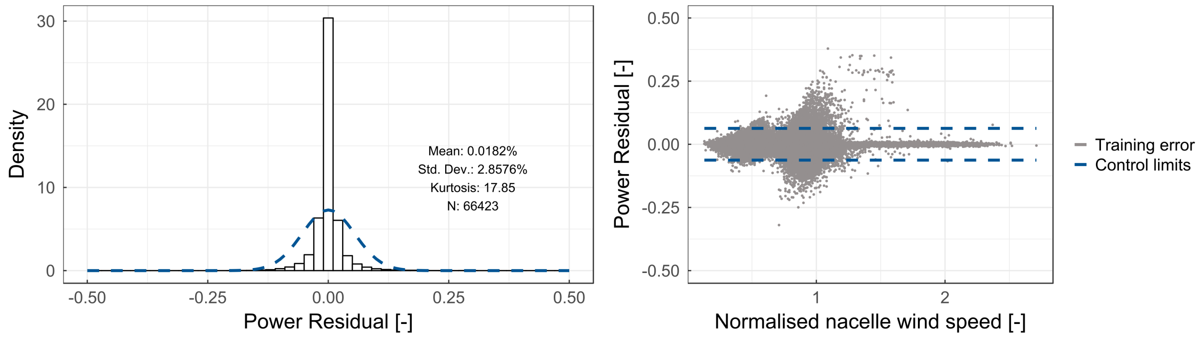

Figure 7. Training error statistics and distribution from RF model (left side); the related normal distribution is illustrated in blue. Training error (grey) and related control limits (blue) from RF model varying depending on the wind speed (right side).

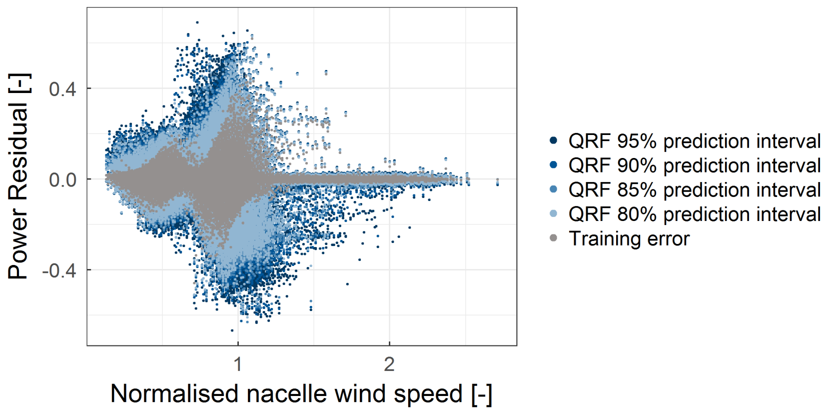

Figure 8. Training error (grey) and related control limits (4 shades of blue) from the QRF varying depending on the wind speed.

3.4. Time resolution detection capabilities

The detection capabilities of each model were tested using two different data time resolutions. To this end, the first month of the year was used to train the models and the subsequent months were used to generate operating residuals. For every model, an underperformance event was deemed to be every time the actual power was below the related low performance limit.

The control limits for normal performance modelled by the kNN and the RF were obtained based on Equation 1 and the following. First, η was obtained from a theoretical definition of the desired level of confidence (95%) given the assumed normal distribution, so-called *Normal error*. Second, η was adjusted empirically so that the resulting control limits encompassed 95% of the training observations, so-called *Adjusted error*.

The cumulative daily number of underperformance events detected was considered for WT performance monitoring purposes. The results for one of the healthy turbines using the high-frequency SCADA data and the different modelling approaches are shown in Figure 9. Similarly, the results using 10-minute data are shown in Figure 10.

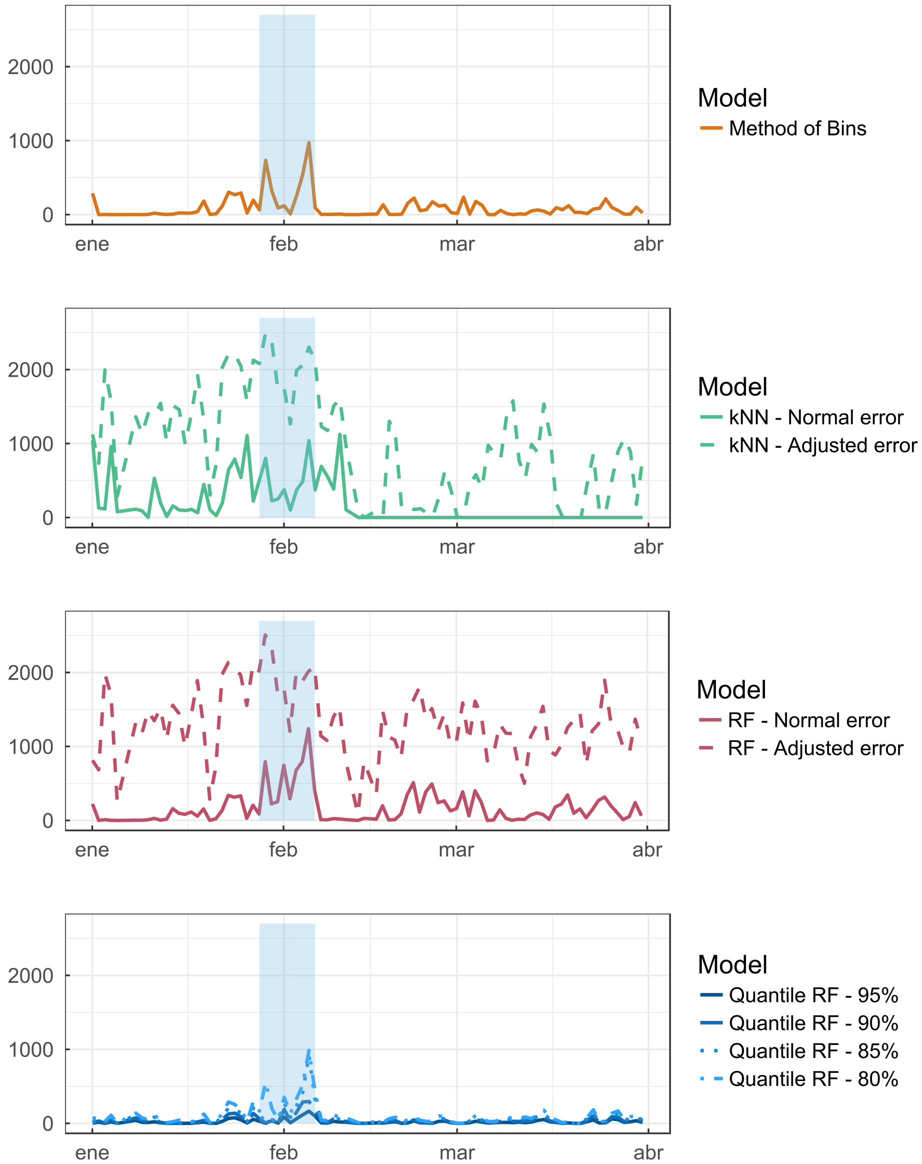

Figure 9. Daily monitoring of a healthy WT using high-frequency SCADA data: cumulative daily number of underperformance events detected with four different modelling approaches. The blue shaded area represents a storm event.

As reported in the alarm logs, a storm occurred in the beginning of February, highlighted in blue in both Figures 9 and 10. When using SCADA data of high-frequency, this event was effectively detected by the MOB and the QRF with confidence levels below 85%, while the latter showed the lowest level of false positives. The use of fixed control limits for the kNN and the RF, with the different uncertainty assessments, gives however a high level of false positives. Therefore, WT performance is not properly monitored when using these two approaches.

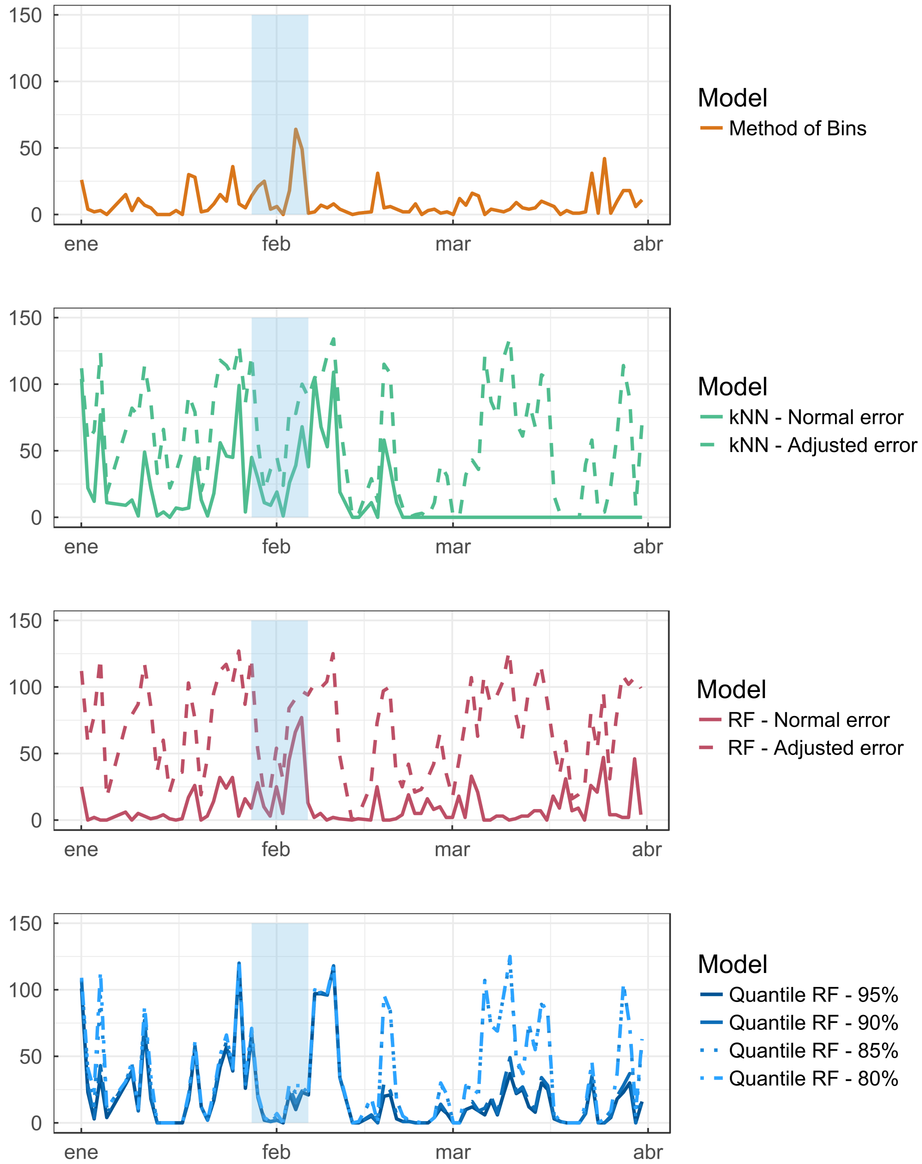

Figure 10. Daily monitoring of a healthy WT using 10-minute SCADA data: cumulative daily number of underperformance events detected with four different modelling approaches. The blue shaded area represents a storm event.

Similar results are observed in Figure 10 for kNN and the RF approaches using 10-minute data. Nevertheless, no significant deviation is detected by the QRF in Figure 10. The abnormal behaviour previously recognised is obscured with many irrelevant events, hindering the appearance of clear patterns of underperformance. As concerns the MOB, the deviation previously detected when using SCADA data of high-frequency is not as significant as when 10-minute data are used. Similarly, no clear patterns are identified. Summarising, WT performance was more accurately monitored in this case study by using high-frequency SCADA data together with QRF for performance assessment.

This storm event reported in the SCADA alarm logs was further investigated in terms of wind speed behaviour as illustrated in Figure 11. As can be seen, the wind speed recorded was not always higher than the cut-out during the whole event. As a result, the underperformance events successfully detected by the QRF using high-frequency SCADA data in Figure 9 were not all due to a normal stoppage occurring for wind speeds beyond the cut-out. Rather, these could also be related, for instance, to high turbulence intensity, and hence extreme loads, to high variations in wind direction or to high wind hysteresis effects.

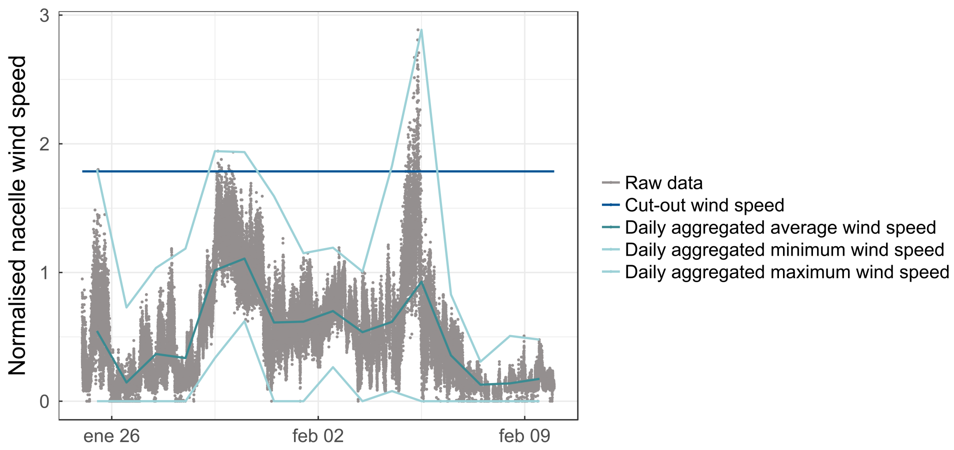

Figure 11. Nacelle wind speed data recorded during the storm event as reported in the alarm logs.

3.5. WT performance monitoring

High-frequency data and the QRF suggested approach were used to monitor the performance of the three turbines described in Section 3.1. WT performance was monitored during the three months prior to the occurrence of a major gearbox assembly failure. As detailed in Section 2, the alarms registered by the SCADA system were translated into component-related information and monitored concurrently to WT performance, to address the root-cause of the detected deviations.

The cumulative daily detected events of underperformance are illustrated in Figure 12 for the the three turbines, together with the alarms recorded at the same time by the SCADA system. The storm event detected in Figure 9 for WT2 is highlighted in blue. The methodology achieves its detection for the three turbines. Moreover, most of the alarms registered during this event were categorised as weather related events. It can be said that this extreme weather event was successfully captured as WT underperformance by the methodology.

Furthermore, another significant deviation was found by the methodology by the end of February in WT3. This event is also highlighted in red in Figure 12. A significant number of alarms were recorded by the SCADA system at the same time, all related to gearbox malfunction. As a result, it is very likely that this significant event of underperformance was due to an abnormal condition of the gearbox assembly. This significant event of underperformance, apparently due to an abnormal condition of the gearbox assembly, was detected 34 days before the actual appearance of the gearbox failure. A potential explanation for this performance deviation detected could be the occurrence of a torque reversal due to the transient loads caused by the developing failure of the gearbox assembly. This would also explain the transitional nature of the detected event. Nevertheless, this could not be confirmed with certainty due to the lack of data.

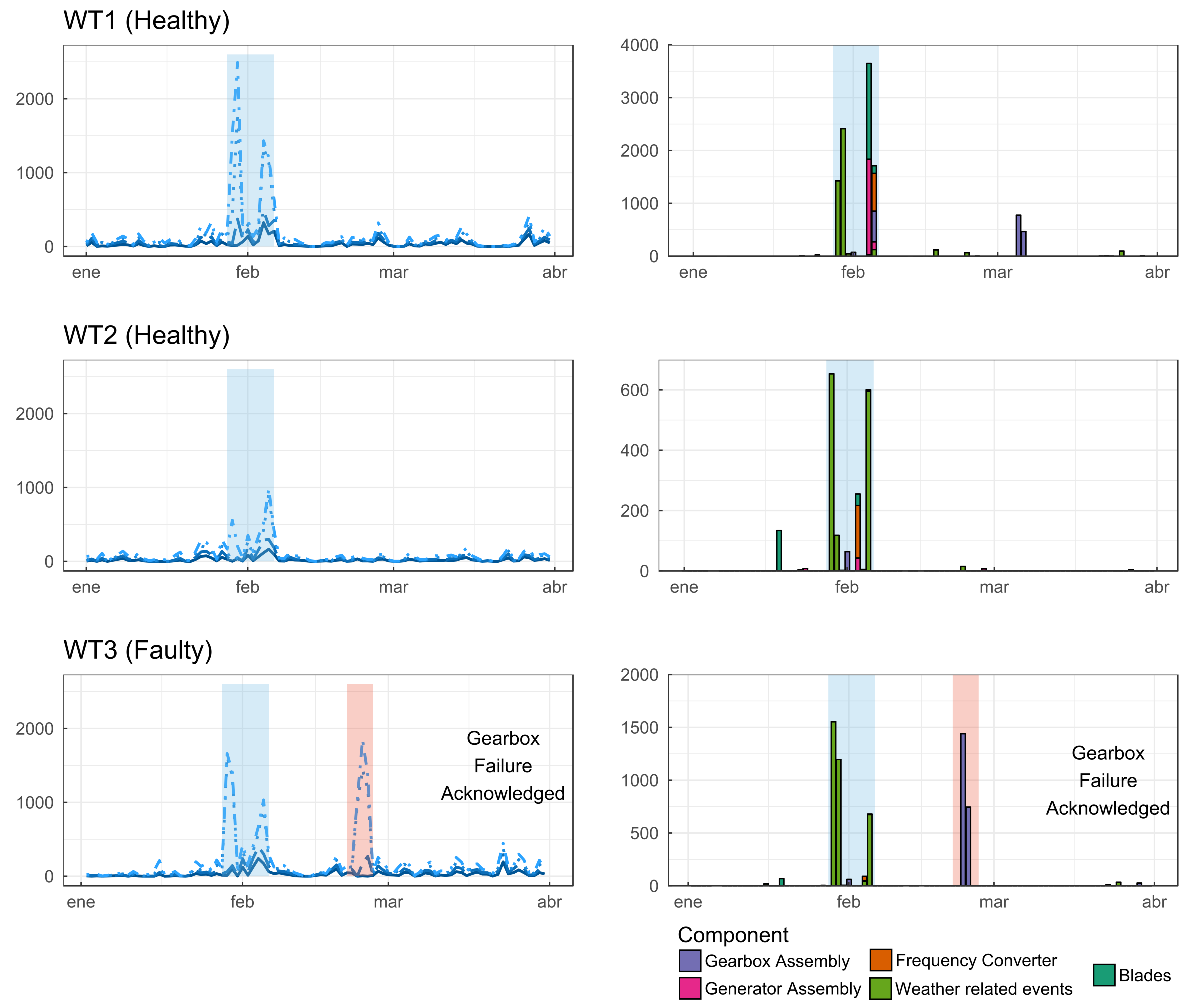

Figure 12. Detection of daily underperformance events (left) and corroborated component-related root cause (right). Alignment of the two are noted by their co-occurence in the shaded regions; the blue shaded area corresponds to the storm event and the red shaded area corresponds to the underperformance detected in the faulty WT.

In any case, the methodology appears to be successful in the monitoring of WT performance, in the detection of performance deviations and in the assessment of their related root cause in the presented case study. The transitional nature of the detected events confirms the importance of using high-frequency data.

4. Conclusions and outlook

This paper presents a methodology where high-frequency SCADA data and QRF are used to monitor WT performance and ascertain the root-cause in the event of underperformance. The results confirm the effectiveness of the methodology and its usefulness as a condition monitoring tool through several case studies.

Monitoring performance is fundamental during the O&M phase, to ensure every WT in a WF is operating safely and efficiently. However, modelling a reference performance together with its related uncertainty is still challenging. Multivariate non-parametric models are confirmed in this paper as providing accurate representation of WT performance. Moreover, the use of QRF is first introduced for a probabilistic assessment of WT performance. This empirical approach allows the model to take into account the heteroscedastic nature of the model residuals and introduces the concept of probabilistic assessment of WT performance.

The use of high-frequency data provides a deeper understanding of WT performance and its natural variability. As illustrated in several case studies it improves the detection capabilities of the suggested approach.

Moreover, the integration of the alarm logs to ascertain the root-cause for the underperformance events detected may turn the suggested approach into a high-level condition monitoring tool with great potential of applicability. Indeed, the presented methodology may be particularly useful for very large WFs, as found offshore. This tool could rapidly identify underperforming WTs that may be in need of a more in-depth analysis. Furthermore, the suggested approach does not require any additional devices and investments, such as LIDAR or meteorological masts.

Further testing and evaluation are needed to confirm the detection capabilities of the suggested methodology; that is, in the case of failures of other components, or different failure modes of the same component. Indeed, questions need to be asked about the relation between WT abnormal performance and specific failure modes. Unfortunately, this investigation is subject to real data availability. Current research is also performed to compare the effectiveness of the suggested approach with other methods for WT performance probabilistic assessment.

Finally, few operators are aware of the potential for SCADA data to reinforce CMS because of its low sampling rate. WF operators are therefore highly encouraged to store high-frequency data to build effective predictive maintenance strategies in order to optimise the O&M phase.

Acknowledgments

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 642108.

The authors would like to thank Enel Green Power for their cooperation and supply of the datasets. The authors gratefully acknowledge Tim Rubert (University of Strathclyde) for his helpful suggestions and comments. The anonymous reviewers are also acknowledged for their comments.

References

[1] Carroll J, McDonald A, Dinwoodie I, McMillan D, Revie M and Lazakis I 2017 *Wind Energy* **20** 361–378
[2] García Márquez F P, Tobias A M, Pinar Pérez J M and Papaelias M 2012 *Renewable Energy* **46** 169–178
[3] Yang W, Tavner P J, Crabtree C J, Feng Y and Qiu Y 2014 *Wind Energy* **17** 673–693
[4] McMillan D, Thöns S and May A 2015 *IET Renewable Power Generation* **9** 900–907
[5] Tautz-Weinert J and Watson S J 2017 *IET Renewable Power Generation* **11** 382–394 ISSN 1752-1416
[6] Kusiak A, Zhang Z and Verma A 2013 *Energy* **60** 1–12
[7] Cambron P, Lepvrier R, Masson C, Tahan a and Pelletier F 2016 *Renewable Energy* **94** 126–135
[8] Gill S, Stephen B and Galloway S 2012 *IEEE Transactions on Sustainable Energy* **3** 94–101
[9] Kusiak A, Zheng H and Song Z 2009 *Renewable Energy* **34** 1487–1493
[10] Uluyol O, Parthasarathy G, Foslien W and Kim K 2011 *Annual Conference of the Prognostics and Health Management Society* August pp 1–8
[11] Lydia M, Selvakumar A I, Kumar S S and Kumar G E P 2013 *IEEE Transactions on Sustainable Energy* **4** 827–835
[12] Taslimi-Renani E, Modiri-Delshad M, Elias M F M and Rahim N A 2016 *Applied Energy* **177** 544–552
[13] Schlechtingen M, Santos I F and Achiche S 2013 *IEEE Transactions on Sustainable Energy* **4** 671–679
[14] Karlsson D 2014 *Wind Turbine Performance Monitoring using Artificial Neural Networks With a Multi-Dimensional Data Filtering Approach* Master's thesis Chalmers University of Technology
[15] Jin Zhou, Peng Guo and Xue-Ru Wang 2014 *2014 International Conference on Machine Learning and Cybernetics* (IEEE) pp 71–76
[16] Janssens O, Noppe N, Devriendt C, de Walle R V and Hoecke S V 2016 *Engineering Applications of Artificial Intelligence* **55** 331–338

[17] Wilkinson M 2016 *Technical Workshop on Analysis of Operating Wind Farms* ed WindEurope (Bilbao: WindEurope)
[18] Gonzalez E and Melero J J 2017 *30th International Congress & Exhibition on Condition Monitoring and Diagnostic Engineering, COMADEM*
[19] Zhang J, Jain R and Hodge B M 2016 *Energy* **112** 1139–1152
[20] St Martin C M, Lundquist J K, Clifton A, Poulos G S and Schreck S J *Wind Energy Science* 221–236
[21] Wagner R, Antoniou I, Pedersen S M, Courtney M S and Jørgensen H E 2009 *Wind Energy* **12** 348–362
[22] Kusiak A and Verma A 2013 *IEEE Transactions on Sustainable Energy* **4** 192–199
[23] Koenker R and Hallock K F 2001 *Journal of Economic Perspectives* **15** 143–156
[24] Meinshausen N 2006 *The Journal of Machine Learning Research* **7** 983–999
[25] Reder M, Gonzalez E and Melero J J 2016 *Journal of Physics: Conference Series* **753** 72027
[26] Gonzalez E, Reder M and Melero J J 2016 *Journal of Physics: Conference Series* **753** 72019
[27] International Eletrotechnical Comission 2005 IEC 61400-12-1 Wind turbines Part 12-1: Power performance of electricity-producing wind turbines. Standard
[28] International Eletrotechnical Comission 2013 IEC 61400-12-2 Wind turbines Part 12-2: Power performance of electricity-producing wind turbines based on nacelle anemometry. Standard
[29] Witten I H and Frank E 2005 *Data Mining: Practical Machine Learning Tools and Techniques* 2nd ed (San Francisco, CA: Morgan Kaufmann)
[30] Breiman L 2001 *Random Forests* vol 45 (Springer)
[31] Mennitt D, Sherrill K and Fristrup K 2014 *The Journal of the Acoustical Society of America* **135** 2746–2764