An Investigation on the Usability of High-Frequency Wind Turbine Controller Data for Predictive Maintenance

Evan Dwayne Roberts, M.Sc.[1]*
Björn Roscher[1], Tobias Winnemöller[2], Ralf Schelenz[1]

[1]Center for Wind Power Drives
Campus-Boulevard 61, 52074 Aachen, Germany

[2]E.ON Climate & Renewables GmbH
Brüsseler Platz 1, 45131 Essen, Germany

Contents

1 Introduction

The wind power industry is experiencing significant pressure to reduce the levelized cost of energy. Predictive maintenance is a promising technology for offshore applications. As an operator of several GWs of wind power worldwide, the company E.ON Climate & Renewables (EC&R) is developing predictive maintenance solutions to support its business. In a pursuit to exploit all available data sources, the company has partnered with the Center for Wind Power Drives (CWD) at RWTH Aachen University to conduct research on the usability of high-frequency turbine controller data for predictive maintenance applications.

The wind power industry uses state of the art SCADA technology to remotely monitor its wind turbines, thereby storing 10-minute measurements of signal data as prescribed by the IEC 61400-12 Norm [IEC15]. EC&R has access to the wind turbines' controller data at high sampling rates. More than 9 months' worth of high-frequency SCADA data, including several thousand channels, are available at sample rates of approximately 0.05 to 8 Hz ready to be explored and analyzed. As the analysis and storage of the vast increase in data come at a cost, the industry requires insight as to which signals are relevant, their respective use cases, storage, and costs.

The used wind farm has reoccurring issues with the converter system (e.g. bus bar and cables, and delta modules) as well as the hydraulics and oil system, which account for the majority of faults at Amrumbank West, one of EC&R's offshore wind farm, making up 45% and 9% of the losses respectively [EON18].

EC&R's primary goal in this research is to explore the usability of turbine controller data for predictive maintenance, thus asking whether high-frequency wind turbine controller data can be used to predict failure modes. Specifically:

- Can the faults be detected in the data?
- How far in advance can they be detected?

The hypothesis of this research is that the use of high-frequency signal data, as opposed to the otherwise used 10-minute averages, yields predictive insight and an improved capacity of conditions monitoring, and can be deployed in the company's predictive maintenance tool in the future. This hypothesis is being put to the test by developing normal behavior models to detect pitch tracking faults with the help of data driven modelling based on varying sample rates of the turbine controller data and is furthermore tested on multiple models.

2 State of the Art

A normal behavior model (NBM) is an example of a condition monitoring system that can be used for predictive applications. Typically, non-parametric models such as machine learning (ML) algorithms are used to monitor features that are associated with normal operational behavior [SHO14]. The flexibility offered by these models enables the creation of NBMs at the component level to monitor performance, health, or other conditions that the model intends to reveal [GON17].

For technical systems, time series data extracted from the IoT network is used as input for NBMs. In a first development step data associated with normal or healthy behavior is used to train a ML algorithm. The developed models can be deployed to monitor data and evaluate how much the signals deviate from the expected behavior. If the prediction error deviates past a certain control value, a warning can be issued. The technology comes at a cost, the initial investment costs are usually significant and therefore require long-term planning to amortize over the lifetime of a given system's project [GRO17, GOD15]. Data analytics is often named in the context of big data as more and more data is being produced in industrial infrastructures, i.e. sensor and log data from SCADA systems [DOR15]. The data on wind turbines is measured at critical major and minor components, e.g. the gearbox, generator, and pitch and yaw system [MCM15, GOD15].

In ML, preprocessing time series data can have a big impact on the subsequent forecasting performance. Pre-processing includes filtering noise, dealing with missing data, as well as transforming and scaling data. The process aims at uncovering features that can be "learned" by a model, thereby improving its performance [AHM14]. In the first few steps of data pre-processing**, filtering and cleaning**. The aim of this step is to omit incomplete data points [SHO14] and consider only relevant data points, thus reducing data that can either be regarded as noise [EST96], or for NBM, considering only faultless operation by omitting any abnormal behavioral data [GON17] (e.g. abnormal pitch angle or transient situations, such as ramp up or shut down). Filtering can be done either manually or with the help of unsupervised algorithms, e.g. a density-based clustering algorithm. The algorithm searches for clusters that have a higher density of data points than outside of the cluster. Therefore "outliers" that aren't contained in any of the thresholds are considered noise. The idea behind the algorithm is that each point within a cluster must contain a minimum number of samples in its neighborhood [EST96].

Many ML algorithms only take input data within a certain range. Neural networks, for instance, require input between the values -1 and 1, which requires the data to be **scaled** between these two values [AHM06]. When features for ML contain data that varies on the order of magnitude, scaling the features within a similar range typically yields better results. Various types of scaling methods exist, the most common of which will be presented in the following paragraphs: normalization and standardization. Standardization scales every sample around zero while considers the mean and standard deviation of the data set [GOD15]. Standardization is commonly used and recommended when working with ML algorithms such as a multi-layer perceptron [PED11]. Many ML algorithms struggle with non-linear relationships or heavily skewed data.

Data/ dimensionality reduction techniques are another class of predictor-transformations. These types of transformations reduce the data by generating smaller sets of predictors that seek to capture most of the information contained in the original variables, thereby reducing the computational load. The process is often called signal or feature extraction [KUH16] and a common method is called a principal component analysis (PCA), which drastically reduces dimensionality while preserving the maximum amount of variance (statistical information) [JOL16].

For time series prediction problems, other more manual forms of **feature engineering** exist, including several methods that are applied to the data to improve prediction results. The simplest approach is called lagging values, the process of which includes is shifting time series values by 1 or n-number of timestamps to provide information on expected values when preceded by another value. Time series differencing takes the gradients of preceding values into consideration. This type of feature engineering can be especially well-suited at capturing the dynamics of a technical system, e.g. a wind turbine's transient conditions. Data can be smoothed out (moving averages), thereby reducing noise, allowing the model to focus on global properties of a time series [AHM06].

3 Methodology

The question in the scope of this paper is whether high-frequency turbine controller data can be used for predictive maintenance at wind turbines. To answer this question systematically, the question is broken down into several more manageable questions which can be evaluated independently and are specific demands by the industry partner. The questions can be evaluated with the help of other available data and by applying some of the methods found in today's literature.

This publication deploys various NBMs with varying data input and ML models to generate a matrix of solutions. These can be analyzed and evaluated answering the questions in scope. After that, various NBMs can be developed with a variation of the chosen data input, e.g. the sample rate of the time series. Turbine controller data from one turbine is resampled at various rates – in this case every 10 minutes (IEC standard) and 10 seconds – and used as input for various models that are designed for NBM. Three regression models are used: a random forest regressor (RF), a support vector machine regressor (SVM), and a multi-layer perceptron regressor (MLP).

Active power, wind speed, pitch position, set-point, and hydraulic pressures for all blades are considered in this paper. Figure 1 shows a flow chart that depicts the processing steps and intermediary objects created along the way on the data pre-processing, training and testing steps for a normal behavior model for specific wind turbine components. The data processing flow follows the standard data science steps of developing and training models, evaluating these, and last but not least testing the model.

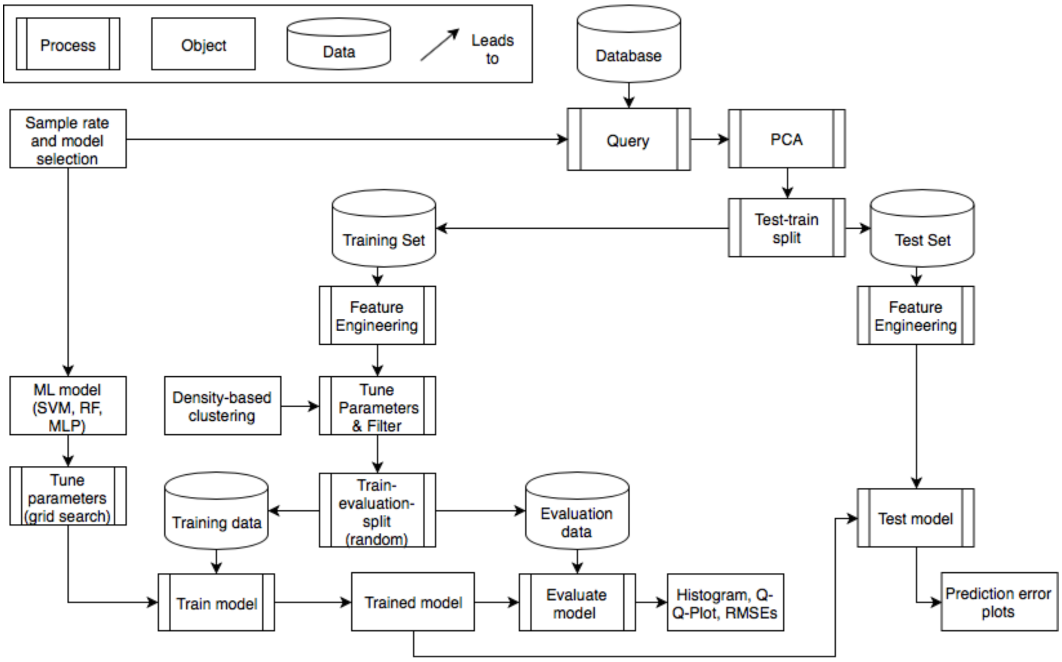

Figure 1: Processing flow chart

The control values (CV) used in this thesis are based on the standard deviation and mean of the prediction error. Three times the RMSE of the prediction error was used to determine an upper and a lower control bound scaled by the mean. Any value that is predicted within these intervals is considered as normal behavior. With respect to a normal distribution these control values will include 99.7 % of the data.

4 Data Analysis

A first smart step in ML is to evaluate over- and underfitting which can be done with the help of the RMSE values of the training and evaluation set. When the difference between the RMSEs on the training and evaluation data is small there is a low probability that the model is overfitted. The SVM showed the best results, although overall its RMSE was among largest. This was true for both high frequency and low frequency data (Figure 2).

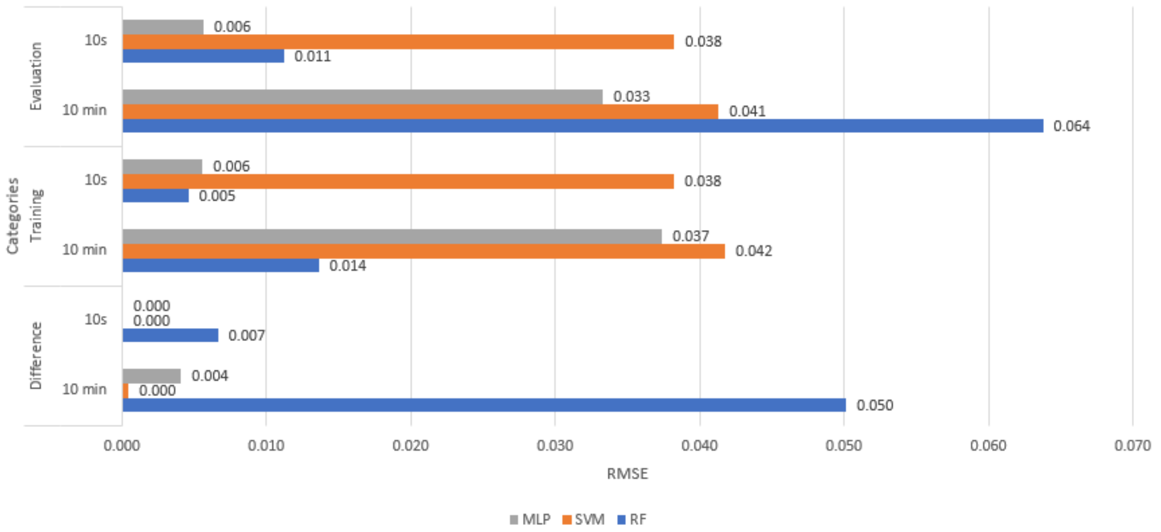

Figure 2: RMSE using training and evaluation data, and the difference between the two

The prediction error distribution is analyzed in the next step (**Figure 3** & **Figure 4**). In addition to the models' characteristic of producing predictable error values (i.e. a normal distribution), the suitability of the method for determining the control values can be evaluated. **Figure 3** and **Figure 4** show the distributions for both sample rates using the SVM. From left to right they contain a prediction error histogram, a Q-Q-plot and a scatter plot of the prediction errors vs. the actual values. All contain horizontal green lines for the upper and lower CV. The red line in the Q-Q-plot represents the theoretical quantiles vs. the expected values for a theoretical normal distribution. The distributions are quite normally distributed, and few data points are classified as outliers. With distributions as these, the chosen method for the CV certainly makes sense and can be expected to produce meaningful classifications of abnormal behavior. Both sample rates produce similar and reliable error distributions.

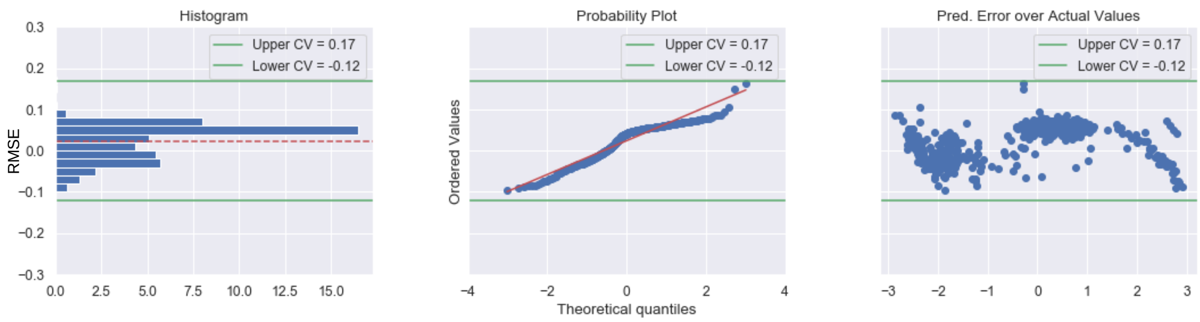

Figure 3: Evaluation plots (PC1: turbine: A18, model: SVM, sampling rate: 10min)

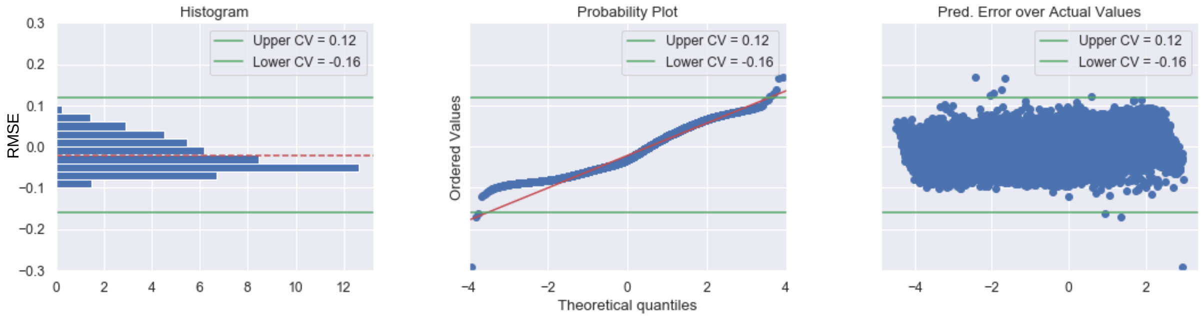

Figure 4: Evaluation plots (PC1: turbine: A18, model: SVM, sampling rate: 10s)

Table 1 shows the average coefficient of determination for all three ML models at both sample rates for a theoretical normal distribution (red diagonal line in Q-Q-plot) and is treated as a measure of normality on a scale of 0 to 1. It shows that the SVM produces the most normal prediction errors and therefore is most suitable for the CV method because it produces the most reliable prediction errors.

	10 min	10s
RF	0,69	0,32
SVM	0,94	0,97
MLP	0,70	0,78

Table 1: Coefficient of determination of the Q-Q plot prediction

The NBM can be deployed on test data. The test set includes two pitch tracking events that are associated with abnormal operational behavior. The idea is that predicted values that are associated with normal behavior will show no deviation greater than the determined CV (purple horizontal lines in each of the following figures). If the NBM ingests data associated with abnormal behavior, the prediction errors will be larger.If the deviation exceeds the CV a deviation can be registered by the system. (Exceeding the CV are marked with a red vertical line).

The next two figures (Figure 5 and Figure 6) are zoomed-in plots of the fault detection on March 24. At both resolutions, the plots reveal that the detection of pitch tracking faults is possible in real-time (green line). The prediction error (blue line) seems to cross the lower control value before the fault occurs. It must be denoted that the blue line is, in fact, not a true representation of the values, but merely a linearly interpolated line connecting individual measurements which are plotted. While this detection is significant and a valuable insight, the lack of predictive insight leading up to the fault over the previous hours makes this detection more a case of a reverse engineered alarm.

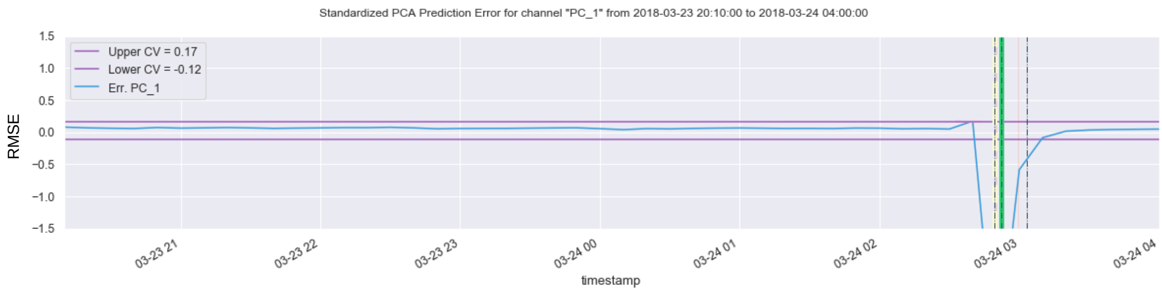

Figure 5: Prediction error plot (PC1, turbine: A18, model: SVM, sampling rate: 10min)

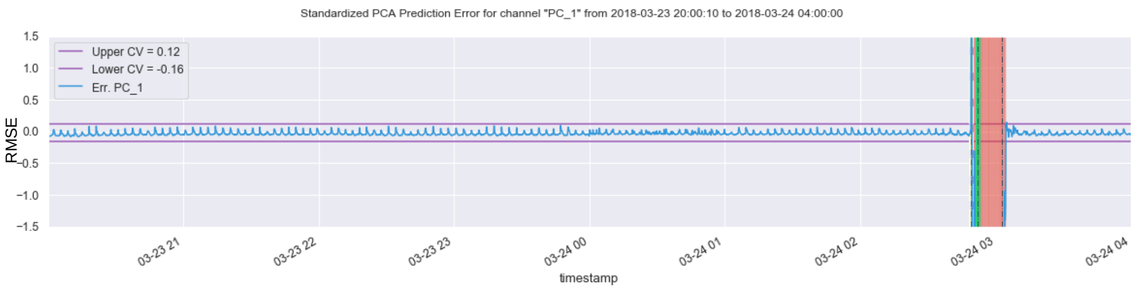

Figure 6: Prediction error plot (PC1, turbine: A18, model: SVM, sampling rate: 10s)

The figures show that the detections are independent of the sample rate. Both sample rates can make the detection equally well and do not show trend behavior leading up the fault.

5 Conclusion

The results of this thesis do not support the assumptions that higher frequency data would yield better results in predictive modeling. The assumption was that sample rates greater than 10-minute averages of the operational data that includes a larger selection of signals would greatly improve predictive insight and more accurately detect failure modes of wind turbines.

The paper showed that while detections can be made, there are no significant differences between the 10-minute averages and the higher resolution 10-second averages. For the higher resolution NBM, abnormal deviations were detected in real-time while the alarms occurred. There was a slightly delayed detection with 10-minute average given the nature of averaging data. Given the right signal data, pre-processing, model selection, and tuning higher frequency data can be used to detect known failures. The prediction error is viewed as a level of deviation from the norm. The CV show when that abnormality threshold is met. The detections made in this paper exceeded CV were sudden and showed no previous trend that could be interpreted or used. Therefore, no predictive insight could be shown in this research.

It could not be shown that predictive insight can be gained by deploying NBMs with higher- frequency turbine controller data, that is, by using the methods used in this thesis. Alternative methods could be tested, but the more plausible solution is that the failure modes in this paper either occur rapidly with no symptoms leading up to them, or the available signal data is simply a strong limitation to accurately detecting the root-cause of the faults. In future cases, additional 10-minute averages can be used to create predictive applications. If a solid case can be made for higher-frequency data, e.g. vibrational data, a new study would be required.

6 Bibliography

[IEC15] *IEC 61400-12-1: Wind energy generation systems Part 12-1: Power performance measurements of electricity producing wind turbines*, IEC 61400.

[EON18] E.ON, *Comprehensive Failure Analysis*, 2018.

[GON17] E. Gonzalez, B. Stephen, D. Infield, and J. J. Melero, “On the use of high-frequency SCADA data for improved wind turbine performance monitoring,” *J. Phys.: Conf. Ser.*, vol. 926, p. 12009, 2017.

[SHO14] S. Shokrzadeh, M. Jafari Jozani, and E. Bibeau, “Wind Turbine Power Curve Modeling Using Advanced Parametric and Nonparametric Methods,” *IEEE Trans. Sustain. Energy*, vol. 5, no. 4, pp. 1262–1269, 2014.

[GRO17] S. N. Grösser, A. Reyes-Lecuona, and G. Granholm, Eds., *Dynamics of long-life assets: From technology adaptation to upgrading the business model*. Cham, Switzerland: Springer Open, 2017.

[GOD15] Godwin, Jamie, Leigh, “Exploiting Robust Multivariate Statistics and Data Driven Techniques for Prognosis and Health Management,” *Durham theses, Durham University.,* http://etheses.dur.ac.uk/11157/, 2015.

[DOR15] J. Dorschel, *Praxishandbuch Big Data*. Wiesbaden: Springer Fachmedien Wiesbaden, 2015.

[MCM15] D. McMillan, S. Thöns, and A. May, “Economic analysis of condition monitoring systems for offshore wind turbine sub-systems,” *IET Renewable Power Generation*, vol. 9, no. 8, pp. 900–907, 2015.

[AHM06] N. Ahmed, A. Atiya, and N. Gayar, “An Empirical Comparison of Machine Learning Models for Time Series Feorecasting,” 2006.

[EST96] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, “A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise,” (Portland, Oregon 02-0.8. 1996

[PED11] F. Pedregosa *et al.,* “Scikit-learn: Machine Learning in Python,” *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[KUH16] M. Kuhn and K. Johnson, Eds., *Applied predictive modeling,* 5th ed. New York: Springer, 2016.

[JOL16] I. T. Jolliffe and J. Cadima, “Principal component analysis: a review and recent developments,” (eng), *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, vol. 374, no. 2065, p. 20150202, 2016.