Drivetrain reliability improvements from long-term field data processed in the cloud

To cite this article: Cédric Peeters et al 2019 J. Phys.: Conf. Ser. 1222 012044

View the article online for updates and enhancements.
Drivetrain reliability improvements from long-term field data processed in the cloud

Cédric Peeters¹,a, Nicoletta Gioia¹, Pieter-Jan Daems¹, Jonas Verbeke¹, Timothy Verstraeten², Ann Nowé², Jan Helsen¹
¹Vrije Universiteit Brussel, Acoustics and Vibrations Research Group (AVRG)/OWI-lab, Pleinlaan 2, B-1050, Brussel, Belgium
²Vrije Universiteit Brussel, Artificial Intelligence Lab (AI-lab), Pleinlaan 2, B-1050, Brussel, Belgium
E-mail: aCedric.Peeters@vub.be

Abstract. This work describes an autonomous condition monitoring framework to process and analyze data measured on wind turbine gearboxes. Industry 4.0 and the Industrial Internet of Things open the door for much more elaborate measurement and data analysis campaigns thanks to the reduction in cost of sensors and of processing power. This increase in data acquisition and handling potential is especially useful considering that most current state-of-the-art methods in signal processing often lead to large quantities of health indicators due to the multiple processing steps. Such large numbers of indicators become unfeasible to inspect manually when the data volume and the number of monitored turbines increases. Therefore, this paper illustrates a hybrid analysis approach that combines advanced signal processing methods with machine learning and anomaly detection. This approach is validated on an experimental wind turbine gearbox vibration data set.

1. Introduction
The rise of Industry 4.0 and correspondingly the Industrial Internet of Things (IIoT), will facilitate an increase in the amount of instrumentation in wind turbines, either directly mounted by the Original equipment manufacturer (OEM) or installed during its lifetime by owner-operators [1, 2, 3]. The IIoT context in particular promotes sensors that are connected directly to the internet. A cloud data-center can thus continuously receive data coming from such sensors. This vast amount of sensor data can be employed for multiple goals. Continuous data streams have both longer signal length and a finer granularity compared to snapshot measurements which are often still common practice nowadays. These properties enable a more refined analysis and help to better understand the turbine characteristics and behavior changes a healthy turbine experiences.

While Industry 4.0 offers valuable opportunities, there are still some challenges that need to be overcome. Given that measuring continuously results in large data sets, automated processing approaches able to autonomously analyze these large amounts of data are necessary. Additionally, all these analyses result in a multitude of indicators that need to be fused into a joint conclusion about system health by applying sensor fusion approaches.
Gearbox condition monitoring of wind turbines (WT) is becoming increasingly more popular and advanced. Current practice in condition monitoring systems often revolves around tracking time-domain statistical indicators and more component-specific frequency-domain indicators [4, 5, 6]. The advantage of using simple scalar time-domain indicators is that no a-priori knowledge about the characteristic fault frequencies is required and thus the number and complexity of the components is not taken into account. While this does simplify the analysis procedure significantly, it also does not allow pinpointing which component is exhibiting anomalous behavior. Not having any knowledge about the exact location of a fault is a downside, but in an early analysis stage it is often easier and quicker to have a straightforward overview of which measurement sensor is picking up faulty behavior without needing to have an immediate answer regarding the exact component and location of failure. Afterwards, a more in-depth frequency-based analysis can follow up and try to determine the missing failure information.

This paper focuses on the challenges and opportunities related to the Industry 4.0 monitoring trend by demonstrating an automated multi-level approach designed for continuous measurements. The approach specified in Section 2 allows facilitating future steps towards prognosis by linking the specifics of the failure modes to the specifics of each analysis work-flow. The goal of this paper is to clarify the different techniques that are used to monitor the turbine and how to unify them into a single approach. The starting point for this monitoring approach are the failure modes and thus they form the base of what the monitoring scheme tries to track.

Section 3 shows how to use condition indicators optimally for the health monitoring of a WT gearbox. Thanks to the recent advances of machine learning (ML), it has now become more feasible to not only identify relevant condition indicators, but also relate them to the operating regime of the turbine for each measurement. It turns out that most of the commonly used statistical condition indicators are sensitive to changes in the operating parameters and thus can lead to erroneous conclusions if they are not compensated for these changes. This paper illustrates a combined use of statistical indicators with machine learning techniques to allow for normalization of the indicators making them independent of the operating regime and provide a streamlined methodology for efficient early fault analyses. Ideally, the provided indicators measure a certain property in the vibration signal that is linked to a specific failure type in order to provide the diagnostic ML models with physics-based and thus relevant features.

The performance of the methodology is illustrated on an experimental vibration data set measured on a WT gearbox. Short measurements were made for five years on the planetary stage of the gearbox providing enough data to extract a healthy behavior baseline. At the end of these measurements the diagnostic models trigger an alarm for anomalous behavior related to the planetary gear stages. This observation is then corroborated by a manual inspection of the gearbox which reveals severe indenting and micro pitting on the gears. This case study confirms that using a more advanced approach combining statistical condition indicators and machine learning can deliver on the promise of early detection of planetary gear damage in WT gearboxes. In turn, this will lead to a reduction in the turbine downtime and cost-of-energy. This paper serves as an overview of how a complex, hybrid monitoring approach can look like in the modern IIoT context, but it does not claim to cement the definite framework of a continuous monitoring scheme. Improvements and extensions to the framework will be implemented in the future to further increase the reliability and efficacy of the proposed approach.

2. Approach
The proposed monitoring approach consists out of multiple steps:

- Important failure modes of the turbine are extracted from a historical failure database.
The advantage of ranking the failure modes is that the monitoring system can target only
the most relevant failure modes of the wind turbine. Furthermore, it allows adapting the
monitoring methods for each failure mode specifically.

• Automated event detection is implemented in order to get a summary about all the
impactful and dynamic events a wind turbine experiences. An example of such a dynamic
event is the emergency braking of a turbine during grid loss. Automatically detecting these
events and classifying their severity provides continuous insight into the loads that the
drivetrain undergoes during its lifetime. Afterwards these events can be linked to specific
failure modes.

• Features are calculated using vibration-based signal processing algorithms. Complex
pipelines of signal processing algorithms are combined in order to extract characteristic
health features.

• Information about the loads is used to normalize the extracted indicators from the vibration-
based signal processing methods and the abnormal behavior health indicators. If there is
no load information directly available, the load is extracted from operational parameters
coming from SCADA data.

• Lastly, feature fusion approaches can be used now to integrate the health indicators and
information from the previous steps into a joint assessment of turbine health. A single
health indicator is defined using a fusion of the various calculated indicators by model
ensemble. An overall health indicator for the turbine is defined by cascading the individual
failure mode alarms.

3. Methodology
In order to be able to integrate the different monitoring approaches the data needs to be made
available in a central data-storage location such that dedicated data processing pipelines can
transform the raw data into health indicators. To facilitate this, we make use of an integrated
big data platform. However, to keep the length of this paper concise, the interested reader is
referred to [7] for more information about the practical implementation of this platform. This
section provides an overview of the techniques used in the data processing and analysis scheme.

3.1. Vibration-based signal processing
Vibration signal processing allows for extracting complex health features from raw data [8].
Signals are processed by a sequence of data-cleaning and filtering steps. A typical flow is
shown in Fig. 1. In case of wind turbine vibration signals, specific attention should be given
to the compensation of speed variations, the variability of the wind, the operational speed will
continuously change in the vicinity of the target value set by the controller [9]. Instantaneous
speed estimation techniques are employed to estimate the speed signal without any need for an
encoder signal.

The Multi-Order Probabilistic Approach (MOPA), originally developed by Leclère et
al. [10][11], is employed for this purpose. Figure 2a shows the spectrogram of a turbine gearbox
signal [10]. Based on the spectrogram, a probability density function (pdf) map is constructed
that combines contributions of all harmonics. This map for the considered signal is shown
in Fig. 2b. Extracting the most probable estimate results in the speed signal illustrated in
Fig. 2c. The MOPA method results in a more accurate speed estimation compared to a classical
approach such as band-pass filtering. Figure 2d illustrates the effect of using the extracted speed
to resample the signal in the angular domain, which leads to a substantial decrease in smearing.

Other processing methods are used to further clean the signal from unwanted components.
Cepstrum-based techniques are employed for removing deterministic signal content [12][13][14].
Filtering the data is also performed to track the statistics in different frequency bands.
Figure 1. Schematic representation of the processing procedure for bearing failure detection.

Figure 2. Speed estimation using MOPA, a) Spectrogram of WT gearbox vibration data, b) PDF map, c) Estimated speed profile, d) Spectra before and after angular resampling.

Demodulation of the signal is done after finding the optimal demodulation frequency band using cyclic spectral coherence [15]. Figure 3 shows an example coherence map detecting an outer race fault frequency in a vibration signal measured on a test turbine gearbox installed at the National Renewable Energy Laboratory (NREL). Lastly, condition indicators are calculated on the pre-processed signals to track the state of the WT gearbox. Both time- and frequency-domain indicators are extracted from the processed data. Examples of calculated indicators include the RMS, peak-to-peak value, crest factor, (Moors) kurtosis, and the peak energy index.

3.2. Event detection
The main goal of this step is to automatically detect events such as emergency braking, coast-downs and run-ups in time series data originating from SCADA 1-second from continuous encoder measurements. The developed approach is illustrated by an example where multiple months of SCADA of different turbines of an offshore wind farm have been automatically processed. Sequential pattern mining algorithms are used to extract relevant patterns from learned events. The main idea is to afterwards perform pattern recognition on speed profiles of the SCADA 1-second to detect all similar events automatically.

An example of an automatically annotated speed profile is shown in Fig. 4. The variability in the extracted run-up signals for the different turbines across the wind farm is shown in Fig. 5.
Figure 3. 2D-map of the cyclic spectral coherence. An outer race fault of the planet carrier bearing is detected at its characteristic frequency, of 8.84Hz around a carrier frequency of 4.6 kHz (red line) [12].

It can be seen that the variability of this event is quite large: there is almost 40% variation in the duration between the shortest and longest identified run-up. During the different run ups, the turbine decreases its acceleration around a normalized speed value of around 0.675. This corresponds to the first continuous operating point of the turbine. Afterwards, the acceleration increases once again until the turbine reaches its second continuous operating point. This information can be condensed into relevant summary statistics regarding typical event occurrence rates, maximum loads and event durations. This analysis step in the monitoring scheme provides us with the opportunity to detect events, classify them, and link them to failure modes of the drivetrain.

Figure 4. Example of an automatically annotated speed signal in which different events have been identified.

3.3. Health indicator fusion and anomaly detection

It is clear that automatic monitoring frameworks require all the indicators to be converted autonomously into a single health conclusion. Different fusion approaches are available to achieve this, e.g. principal component analysis and neural networks. We opt for an approach that is more physics-based and understandable: the model ensemble. Indicator processing pipelines are fed with different subsets of data. Each of the pipelines performs feature calculation and
anomaly assessment. Finally, a majority vote among the features decides on the anomalous state of the turbine. A health indicator is defined for each turbine subcomponent. These are propagated to the turbine level to indicate the overall health.

4. Experimental application

The proposed vibration-based anomaly detection method is validated on a case of damage initiation in a planet gear of a multi-megawatt WT gearbox. Due to confidentiality reasons, we cannot disclose any concrete details of the WT type or gearbox kinematics and the figures are displayed using anonymized axes. The measurements themselves span multiple years and consist of ten second vibration signals, captured every two to three days. Different measurement channels were mounted spread over the gearbox housing to increase the sensitivity for signatures coming from different gearbox components. Considering the fact that every single-channel, ten-second signal leads to a large number of indicators due to the pre-processing, it is impractical to manually assess all of the indicator trends for all channels.

The case study shown in this work describes incipient deterioration on the planetary gear stage that has a high potential to develop into much more severe damage. This damage was detected using the proposed methodology and validated afterwards. All measurement channels were employed in this analysis. For the analysis performed here, every signal leads to sixty (2x5x6) calculated indicators due to the pre-processing as follows:

- Raw data and cepstrum liftered data (x2).
- Unfiltered data and four frequency-filtered signals (x5).
- Six indicators per pre-processed signal (x6).

While for one turbine it is possible to manually go through all indicator trends of all measurement channels, it becomes much more cumbersome when this needs to be done for a whole wind farm. Therefore the sixty different indicator trends are used as input for the approach as described in Section 3.3.

The anomaly detection algorithm is validated through an in-depth analysis over all statistical features. In Figure 6, we present an example taken from this analysis, demonstrating the validity of the proposed method. It illustrates the anomaly detection and alarm fusion mechanism. The first level shows two peak-to-peak features A and B for the planetary stage channel and the same feature B for the high-speed stage channel with normalized values. Each data point in the time series is colored by the alarm level predicted by the ML model. At the bottom, the fused indicator over all peak-to-peak features and all channels is given. The y-value and color
of the fused indicator shows the severity of the alarm. All time series span over multiple years. The results show a significant change in feature values for the planetary stage channel. This is successfully detected by the ML method. The high-speed stage channel is unrelated to the observed failure and therefore only shows stable behavior. The fused alarm is computed by taking the average alarm level over all features per statistical indicator type (e.g., peak-to-peak) and takes the maximum over all channels. This fusion is carried out over all 8 channels and 60 features, which results into 6 fused alarms (one per indicator type) and thus significantly reduces the workload of the analysis.

The effects of the failure on the behavior of the statistical features can be observed in the fusion, reaching an alarm severity of about 70%. A threshold can be set by the WT operator to automatically signify potential failures. The number of figures that need to be inspected also gets drastically reduced thanks to the feature fusion.

5. Conclusions
This paper illustrates an approach to extract event and turbine health information from large data measurement streams in a automated way. Advanced vibration-based signal processing methods are used to extract complex health indicators related to the gears and bearings in the turbine gearbox. Pattern recognition methods show the potential for automatically detecting transient events and thus gaining better insights into what a turbine experiences during its lifetime. Finally, these indicators are fused using an ensemble of models combined with majority voting to obtain an overall turbine health indication. The main goal of this paper is to provide a guideline for a scalable and effective automated monitoring approach of wind turbine drivetrains. Key monitoring aspects are illustrated on experimental wind turbine data.

Figure 6. Feature fusion results for different channels and per feature type.
Acknowledgments
The authors of VUB would like to thank FWO for their support with an SB Ph.D. fellow grant and the agency for Innovation by Science and Technology in Belgium for supporting the SBO HYMOP project. They would also like to thank FWO for providing funding for a long stay abroad of Cédric Peeters at the university of INSA Lyon.
References

[1] Lee J, Kao H A and Yang S 2014 Procedia Cirp 16 3–8
[2] Jeschke S, Brecher C, Meisen T, Özdemir D and Eschert T 2017 Industrial Internet of Things (Springer) pp 3–19
[3] Sadiku M N, Wang Y, Cui S and Musa S M 2017 IJASRE 3
[4] de Azevedo H D M, Araújo A M and Bouchonneau N 2016 Renewable and Sustainable Energy Reviews 56 368–379
[5] Lu B, Li Y, Wu X and Yang Z 2009 Power Electronics and Machines in Wind Applications, 2009. PEMWA 2009. IEEE (IEEE) pp 1–7
[6] Tchakoua P, Wamkeue R, Ouhrouche M, Slaoui-Hasnaoui F, Tameghe T A and Ekemb G 2014 Energies 7 2595–2630
[7] Helsen J, Peeters C, Doro P, Ververs E and Jordaens P J 2017 Big Data Computing Service and Applications (BigDataService), 2017 IEEE Third International Conference on (IEEE) pp 179–184
[8] Carden E P and Fanning P 2004 Structural health monitoring 3 355–377
[9] Yang W, Tavner P J, Crabtree C J, Feng Y and Qiu Y 2014 Wind Energy 17 673–693
[10] Leclere Q, Andrè H and Antoni J 2016 Mechanical Systems and Signal Processing 81 375386
[11] Peeters C, Leclere Q, Antoni J, Guillaume P and Helsen J 2017 Journal of Physics: Conference Series vol 842 (IOP Publishing) p 012053
[12] Peeters C, Guillaume P and Helsen J 2018 Renewable Energy 116 74–87
[13] Peeters C, Guillaume P and Helsen J 2017 Mechanical Systems and Signal Processing 91 354–381
[14] Borghesani P, Pennacchi P, Randall R, Sawalhi N and Ricci R 2013 Mechanical Systems and Signal Processing 36 370–384
[15] Antoni J 2007 Journal of Sound and vibration 304 497–529