Review

General comments

In my opinion, the publication represents a useful contribution to scientific progress in the context of WES. It is of interest to the entire wind power community.
The main objective of the work is anomaly detection using simple Principal Component Analysis (PCA). The topic of artificial intelligence and machine learning is one of the hot topics of the moment. Therefore, it is also important to examine how these methods can lead to improvements in the context of wind turbines.

It is important to the authors to show a simple solution that does not require additional sensors. The approach using PCA and the SCADA data is interesting here. However, I wonder whether, given the value of the turbines and the maintenance costs that may be necessary, additional sensors and higher-quality ML methods would not be more effective. The state of the art is that PCA is not particularly suitable for anomaly detection.

Specific comments

Data preprocessing is not sufficiently described in the paper. If I understand it correctly, a range is specified for the real data and outliers are adjusted accordingly to the damage-free training data. This has several problems: Weak signals are filtered out, the model is only valid for the system under consideration, and the model assumes that the system under consideration is at the bottom of the bathtub curve, i.e. entirely error-free. Overall, due to the low sampling rate, the 10-minute intervals and the averaging over a week, the data appear to me to be very smoothed, which makes it difficult to find anomalies. Since we are dealing here with time series, a simple Pearson correlation is only of limited help (a Spearman's rank correlation should at least be examined here).

The data show a clear seasonal component. The question arises why this was not removed by decomposition, especially since the model is only based on individual data sets. After peaks (Figure 11), the signal drops sharply again for longer periods. What is the reason for this? Since the level seems to be significantly lower in the period from May to November, it is questionable whether incipient damage could be detected here at all.

Overall, despite the previously mentioned criticisms, the work is important because it helps to further advance the topic of machine learning and discuss the benefits and methods.

The anomalies also do not allow any statement on the type of damage present and the severity of the error. At the same time, there is no statement about the historical data and any false alarms. For a scientific consideration, a hit rate AND an error rate must be given.

The metrics in Table 4 need to be explained.

Line 133 and Figure 3: It is not described what kind of damage is typically involved here (lack of lubrication, wear, pitting, ...). Each defect should produce certain characteristics in the measured variables.

The structural, linguistic and graphic quality of the publication is very good. The work is clearly structured and the tables, graphs and pictures are easily recognizable and informative.

Technical corrections

Figure 16 with WT16 should be placed closer to line 326 where it is addressed.